ORCA Manual

Release 6.0

Max-Planck-Institut für Kohlenforschung

Sep 08, 2024

PREFACE

- An *ab initio*, DFT and semiempirical SCF-MO package -

Version 6.0

Design and Scientific Directorship: **Frank Neese**
Technical Directorship: **Frank Wennmohs**

Max-Planck-Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1, 45470 Mülheim a. d. Ruhr, Germany
department-mts@kofo.mpg.de

With contributions from:
Daniel Aravena, Michael Atanasov, Alexander A. Auer, Ute Becker, Giovanni Bistoni, Dmytro Bykov, Marcos Casanova, Vijay G. Chilkuri, Pauline Colinet, Dipayan Datta, Achintya Kumar Dutta, Sebastian Ehlert, Nicolas Foglia, Marvin Friede, Dmitry Ganyushin, Miquel Garcia, Tiago L. C. Gouveia, Yang Guo, Andreas Hansen, Ingolf Harden, Benjamin Helmich-Paris, Lee Huntington, Róbert Izsák, Riya Kayal, Emily Kempfer, Christian Kollmar, Axel Koslowski, Simone Kossmann, Lucas Lang, Marvin Lechner, Spencer Leger, Dagmar Lenk, Dimitrios G. Liakos, Dimitrios Manganas, Dimitrios A. Pantazis, Anastasios Papadopoulos, Taras Petrenko, Peter Pinski, Philipp Pracht, Christoph Reimann, Marius Retegan, Christoph Riplinger, Michael Roemelt, Masaaki Saitow, Barbara Sandhöfer, Kantharuban Sivalingam, Bernardo de Souza, Georgi Stoychev, Van Anh Tran, Willem Van den Heuvel, Zikuan Wang, Hang Xu

And contributions from our collaborators:
Mihály Kállay, Stefan Grimme, Edward Valeev, Garnet Chan, Jiri Pittner, Martin Brehm, Lars Goerigk, Vilhjálmur Åsgeirsson, Liviu Ungur

Additional contributions to the manual from:
Wolfgang Schneider

ORCA 6.0 FOREWORD

Welcome to ORCA 6 – we sincerely hope that you will enjoy it!

ORCA 6.0 is a major turning point for the ORCA project and consequently, it seems appropriate to dwell a little bit on how we got to this point in this foreword.

The ORCA program suite started its life around 1995 as a semi-empirical program written in Turbo Pascal and designed to calculate some magnetic and optical spectra of open-shell transition metal complexes in enzyme active sites. It was unimaginable at the time that it could possibly grow into a major, large-scale software that is used by tens or thousands of people world-wide.

In its evolution the ORCA package probably had a similar trajectory to many other programs: it started with good intentions, courage paired with a healthy dose of ignorance and a vision of a few concrete tasks that it should be able to perform. And it did that. But right there at the beginning, when the foundations were laid for years to come, there was no master plan. There also was no experience – neither with larger-scale software development, nor with quantum chemistry in general.

The logical consequence of the absence of a master plan was, that the program grew in a way that was rather need-driven and short-time goal oriented. The original infrastructure was not horrifically bad, but it was not designed in a way that was strongly suggestive of healthy growth for decades to come. Not surprisingly, after 5-6 generations of Ph.D. students and postdocs working on it, some individuals more disciplined than others (including the original author), the inevitable happened: the code started to look intimidating to new students entering the project since the code started to be clumsy and convoluted.

The consequence of convoluted code is that new programmers start to copy and paste large sections of code which adds significantly to the overhead and made the code even less readable. One immediate consequence of such organically grown code is that it is exceedingly difficult to properly adapt it to the challenges of a rapidly changing hardware landscape and we eventually had to realize that this is also true for the ORCA code.

In the year 2020, when the pandemic hit globally and travelling ceased, there came a time of relative calmness that allowed for contemplation and also concentrated, continued work. At this time, the SHARK package was created based on an idea that I had back in 2016. It turned out to be very successful and led to a highly performant integral code that also was very compact thanks to the loop-kernel-consumer (LKC) concept proposed by Frank Wennmohs. Together with other innovations, for example large improvements in the chain of spheres exchange (COSX) approximation by Robert Izsak, Benjamin Helmich Paris and Bernardo de Souza as well the integration grids in ORCA (by Bernardo de Souza). SHARK, COSX and improved numerical integration formed the core of what was released as ORCA 5.0 on July 1st 2021. The improved performance by the program was received very favorably by the user community and led to an explosive growth of the user base. At the point of writing (2024), the number of ORCA users is increasingly roughly quadratically with time. At the day of writing, ORCA has ~70000 academically registered users and an unknown (but large) number of users in industry.

While we were proud of what we had achieved with ORCA 5.0, it was clear that we had just seen the tip of the iceberg. What was lurking underneath was a complete redesign of the infrastructure, not just patching SHARK into the strategic places. This task amounts to basically a complete rewrite of the entire code and a redesign of the flow of information. This is obviously an intimidatingly complex and large project and something that – to the best of my knowledge – has never been done in the history of quantum chemistry: take a major, large-scale program package and redesign it from scratch – but this time with the hindsight and insights from close to 3 decades of doing it.

Talking to many colleagues a common statement is "if I could start over again, I would do XYZ, but". In the case of ORCA, it actually happened – we did just that! It was a long and occasionally painful road and it would be a lie to not admit that there were moments were I felt like giving up on the idea. But after three long years, it was finally seen through to the end (well, almost) thanks to the tireless efforts and the development team and the enthusiasm and patience that the members have contributed.

It was clear to me that the project "rewrite ORCA" is a bad project for Ph.D. students and postdocs and consequently, I have taken a large part of the tedious work on myself in the hope that the other developers could focus on continuing doing great science - and they did! And they did by embracing and using the emerging new infrastructure which was no small feat since the new infrastructure was a moving target for years and the developers had to work around

bugs, mistakes and incompleteness of the new infrastructure. But they did do that and showed great dedication, appreciation and skill in doing so. And of course, a number of individuals also helped with the tedious part of the whole project and to all those, I am particularly grateful.

The result of our efforts, you now hold in your hands: the ORCA program, version 6.0. But as I explained above, ORCA 6.0 is not just an update to the program, it is essentially an entirely new quantum chemistry package – but one that was designed with a master plan, a vision of how such a package could or should be organized. This led to a highly streamlined and highly efficient new infrastructure that will greatly facilitate future developments.

Please allow me a few personal words: This release of ORCA is a turning point and also a very emotional moment for me. Recreating the ORCA infrastructure and deleting much of the legacy code amounted to reliving a large part of my scientific life in fast forward. Many memories were tied to specific code parts and so many images returned along the way of how life was when this or that was written and what that world looked like back then. Hitting the "delete" button did not come easy and there might have been a tear or two lurking here and there, especially on the last weekend before the initial code freeze where I deleted more than 250000 lines of legacy code and edited over 500 files of source code.

ORCA 6 was the result of the work of a large number of outstanding and dedicated individuals. Unfortunately, It is impossible to individually mention all of them here (for this, please check the credit section at the beginning of the output), but I do want to ensure *all* ORCA developers of my deepest sympathy, my admiration and my gratitude for staying on path, for their hard work, for their creativity, for their intellectual brilliance, for the dedication and for sharing in the vision. Especially the latter was not a given, in particular in those moments where things were broken that once upon a time were working perfectly. Specifically, I am indebted to Frank Wennmohs for his long-term friendship, for enduring my stubbornness in pursuing this project and for his decisive contributions in important moments. I also want to praise Dagmar Lenk for running our testsuite with almost 2000 jobs every night, analyze the results with superhuman patience and patiently going after the people that were supposed to fix the errors. And of course, my very special thanks and deepest gratitude also goes to Ute Becker. Ute has been a member of the team since the early 2000s. She has single-handedly parallelized ORCA and in all these years, she always had everybody's back – implementing, helping, testing, cleaning up behind people without ever complaining and with laser precision and the highest efficiency. Ute will formally retire by the end of 2024 but we consider ourselves lucky that she has agreed to keep working with us on the next generations of ORCA, at least for a while.

I am also deeply indebted to the members of FAccTs. Ever since the foundation of FAccTs, it has been continuously growing and is now very successful in the market. This is largely due to Christoph Riplinger's ingenuity, vision and insightful leadership. It is a major pleasure to see FAccTs bloom and grow, drive technology and also assemble a significant number of the most talented individuals that passed through the group in Mülheim. Importantly, several FAccTs members have made major contributions to the release of ORCA 6, in particular Bernardo de Souza, Georgi Stoychev and Miquel Garcia Rates have written extremely effective and important code and have also been instrumental in streamlining and optimizing the infrastructure.

Now that ORCA 6 has become reality we are highly excited to give it you and we sincerely hope that you embrace it and make good use of it. Thank you for staying with us through the long wait that led to ORCA 6. We believe that we have made the program fit for the next decades to come and to be a great platform for keeping up with the rapidly changing hard- and software landscape. The efficient new infrastructure that Orca 6 is based on will allow for much improved development speed and consequently, we are looking highly forward to giving you the next ORCA versions with exciting new functionality in due course.

Thank you for your support!

Frank Neese, on behalf of the development team on July 17, 2024

0.1 Highlights

0.1.1 SCF and Single Reference

- LeanSCF: reduced memory, more robust convergence
- Electric field optimizations
- General ROHF implementation (SCF/Gradient) with all approximations
- General CSF ROHF
- New density functionals
- Delta-SCF
- UHF STEOM-CCSD
- UHF-IP-EOM-CCSD
- UHF-EA-EOM-CCSD
- Regularized OOMP2
- Solvation in OOMP2
- Improved stability analysis featuring all approximations, solvation etc
- MixGuess to converge to biradicaloid open shell singlet type broken symmetry solutions
- Approximate Spin Projection Method for broken symmetry calculations (SCF/Gradient)

0.1.2 Multi Reference

- TRAH, AVAS, MCRPA
- Linear response CASSCF
- Vastly improved Recursive CI coupling coefficient generation

0.1.3 Automatic Code generation

- MPn
- CCSD(T) gradients
- CCSDT

0.1.4 Relativity

- X2C
- New and consistent DKH infrastructure

0.1.5 Solvation & Embedding

- DLPNO-CCSD(T) PTES Approach
- SMD analytical Hessian
- Dynamically adjusted radii: DRACO
- Improved surface grids
- Interface to openCOSMO-RS
- Explicit Solvator
- Molecule Docker
- FMM implementation for embedding
- CIM implementation works with DLPNO-CC, DLPNO-MP2

0.1.6 Optimization

- More robust optimizer (fewer cycles, fewer cases with negative frequencies)
- GOAT global optimizer and conformer generator
- Basis set limit extrapolated optimizations through compound scripts
- Extrapolation with counterpoise correction through compound scripts

0.1.7 Hessian

- Group parallelization
- Performance improvements

0.1.8 Excited States

- Analytical gradient for meta-GGA functionals

0.1.9 New Spectroscopic Properties

- VCD implementation at the SCF level
- MCD with vibronic structure
- General spin ROCIS
- Higher order moments and exact field matter coupling
- Spin rotation constants

0.1.10 Misc Properties

- Local dipole moments and polarizabilities
- Frequency dependent electric properties
- VPT2 enhancements
- Restructured NMR simulation program `orca_nmrspectrum`
- MBIS charges

0.1.11 Workflow & Interfacing

- Property file: Machine readable, Human readable summary of ORCA run
- Compound: vastly improved Syntax, features, optimizations, ...
- `orca_2json`: generate integrals, property file, run backwards to get MOs into ORCA
- Citation tool for helping find the right references

0.2 Other Changes

0.2.1 SCF and Infrastructure

- Significant improvements to the SOSCF solver to make it more robust, preventing huge steps that break the SCF. Overall improvements on the DIIS solvers.
- Due to the SCF updates, the AutoTRAH is now not so often needed and will start now only from above 50 cycles (`AutoTRAHIter`).
- Improvements to the memory handling of TD-DFT, CP-SCF and the Hessian

0.2.2 Basis sets

- def-TZVP and ma-def-TZVP pseudo-potential basis sets for the actinides (Z = 89, Ac - 103, Lr)
- Lehtola's hydrogenic gaussian basis set family (HGBS) including polarized (HGBSP) and augmented (AHGBS, AHGBSP) variants for all elements up to Oganesson (Z = 118)
- def2-SVPD, def2-TZVPD, def2-TZVPPD, def2-QZVPD, def2-QZVPPD basis sets for lanthanoids
- `!MINIX` now correctly activates the corresponding ECP
- Added user-specified L-limit to AutoAux `AutoAuxLLimit`
- Fixed segfault in dhf-ECP
- Fix for `DelECP` in `%coords`
- Added `ReadFragBasis` keywords read fragment-specific basis sets from a file

0.2.3 Solvation

- New charge correction / compensation algorithm (corrected charges printed in an additional file)
- C-PCM/B scheme for QM/MM calculations
- DDCOSMO and CPCM/X available for XTB calculations and QM/MM calculations
- Generalization of names within all solvation models (C-PCM/SMD/ALPB/DDCOSMO/CPCM-X)
- New discretization scheme for the cavity (C-PCM) based on a constant number of charges per unit of area

0.2.4 DFT

- Allow LibXC functional customization via external parameters
- Simple input keywords added for some LibXC functionals
- Added wB97M(2) functional parameters: must be used with wB97M-V orbitals in a two-step job (compound script available)
- Bugfixes for LibXC combined `*_xc_*` functionals
- Fixed crash for D4 + ghost atoms

0.2.5 Excited states

- Analytical gradient for meta-GGA functionals
- Small bugfix to spin-adapted triplets and NACMEs.
- The FolllowIRoot for excited state optimization uses now a much more robust algorithm.

0.2.6 Relativity

- Enabled `NumGrad` with relativistic methods
- Second order DKH picture-change correction of contact density
- Minor fixes in DKH picture-change corrections of magnetic properties
- Picture change corrections are activated automatically

0.2.7 Multiscale

- Reading PDB files for 10k+ atoms with HETATMs now possible
- Enabled correct FlipSpin behavior with QMMM
- More efficient MM Module
- Implemented wall potential

0.2.8 Coupled cluster / DLPNO

- Implemented energy ordering for PNO generation
- Added semicore treatment for DLPNO
- Enable DLPNO-CCSD(T) calculations to run DLPNO-CCSD unrelaxed densities

0.2.9 MP2

- Corrected memory estimates and batching in response and gradient
- Removed the slow and limited analytic (RI-)MP2 Hessian code
- Removed non-default Gamma-in-core option for RI-MP2 response
- Disabled single-precision calculations
- Disabled SemiDirect option in AO-MP2
- Enabled range-separated DHDFT gradients with RIJDX

0.2.10 NEB

- Improved IDPP initial path
- More efficient GFN-xTB runs for NEB

0.2.11 COSX

- Improvements to numerical integration grids, both for DFT and COSX
- Faster grid step
- Improved performance and accuracy in COSX, also for the gradient and Hessian

0.2.12 Properties

- NMR spin-spin coupling:
  - Added `SpinSpinElemPairs` and `SpinSpinAtomPairs` keywords to limit which couplings are computed
  - Reduced the number of CP-SCF perturbations necessary via a stochastic selection
  - DSO term was transposed.
  - Off-diagonal PSO elements had the wrong sign
  - Efficiency improvement: solve SD/FC CP-SCF equations in restricted mode for RHF, instead of always using UHF
- Optimized numeric integration for HFC gauge correction
- Removed `RITRAFO` option for CP-SCF
- Switched to `tau=Dobson` as default handling of the kinetic energy density in meta-GGA magnetic properties with GIAOs

0.2.13 Hessian

- Improvements to the Hessian to avoid accumulation on numerical noise and reduce the number of spurious negative frequencies.

0.2.14 Geometry Optimization

- Several improvements to the geometry optimization, making is much more stable. Complete redesign of the Cartesian optimizer (!COPT), making it quick enough to be used together with faster methods.
- Fallbacks in the geometry optimization in case something fails, e.g. if the internal coordinates are unacceptable.
- Arbitrary spherical, ellipsoidal or box-like wall potentials can be added, which will reflect on the energy and gradients and can be used during geometry optimization.

0.2.15 Miscellaneous

- CHELPG charges that reproduce the ESP together with the molecular dipole moment
- Fixed issues with constraints in multi-step jobs
- Molden output: store ECP info in `[Pseudo]` block, set point charge atomic number to 0, handling of ghost atoms
- Made the `ExtOpt` interface easier to use
- Store energy from NEB and IRC in the XYZ file

0.3 Why do some of my calculations give slightly different results with ORCA-5?

Besides the new grids there has been some change to the default settings in ORCA-6.

0.4 Why is ORCA called ORCA?

Frank Neese made the decision to write a quantum chemistry program in the summer of 1999 while finishing a postdoc at Stanford University. While thinking about a name for the program he wanted to write he decided against having yet another "whatever-Mol-something". The name needed to be short and signify something strong yet elegant.

During this time in the US Frank went on a whale watching cruise at the California coast—the name "ORCA" stuck. It is often asked whether ORCA is an acronym and over the years, various people made suggestions what acronym this could possibly be. At the end of the day it just isn't an acronym which stands for anything. It stands for itself and the association which comes with it.

0.5 How do I install ORCA on Linux / MacOS / Windows?

On Linux and MacOS the most convenient way to install ORCA is by using the command-line installer. You have to download a .run file, e.g. `orca_6_0_0_linux_x86-64_shared_openmpi416.run`. Then make the file executable, that is, open a terminal, enter the directory in which you stored the file and enter:

```
chmod a+x orca_6_0_0_linux_x86-64_shared_openmpi416.run
```

Afterwards you can simple execute the file, like

```
./orca_6_0_0_linux_x86-64_shared_openmpi416.run
```

The installer will install orca in a user directory, as well as set the path to include the orca directory. After *opening a new window* ORCA can be used as indicated.

The installer has a few more options, like extract-only, setting the path interactively, and setting the path by option:

```
-i        : Set a different install path interactively
-p <path> : Set a different install path
-x        : Just extract
```

The options have to be given *after* a double dash, e. g.

```
./orca_6_0_0_linux_x86-64_shared_openmpi416.run -- -p /my/home/orca/dir
```

0.6 How do I install the xtb module?

Please download xtb version 6.7.1 from the Grimme lab's repository Grimme lab's repository and copy the xtb binary into the orca directory.

0.7 I've installed ORCA, how do I start it?

First and most importantly, ORCA is invoked from the command line on all platforms. A simple click on a binary or an input file won't start a calculation. Under Linux and MacOS you need to open a terminal instance and navigate to the folder containing an example.inp file. You can run an ORCA calculation with the command:

```
<full orca binary folder path>/orca example.inp > example.out
```

Similarly, under Windows you need to open a command prompt (Win7, Win8) or a power shell (Win10), navigate to said directory and execute the following command:

```
<full orca binary folder path>/orca example.inp > example.out
```

0.8 How do I cite ORCA?

Please do **NOT** just cite the generic ORCA reference given below but also cite in addition our original papers! We give this program away for free to the community and it is our pleasure and honour to do so. Our payment are your citations! This will create the visibility and impact that we need to attract funding which in turn allows us to continue the development. So, **PLEASE**, go the extra mile to look up and properly cite the papers that report the development and ORCA implementation of the methods that you have used in your studies!

The generic reference for ORCA is:
Neese,F.; Wennmohs,F.; Becker,U.; Riplinger,C. "The ORCA quantum chemistry program package" *J. Chem. Phys.*, **2020** 152 Art. No. L224108 doi.org/10.1063/5.0004608

There has been an update for ORCA 5.0:
Neese, F. "Software update: The ORCA program system—Version 5.0" *Wiley Interdisciplinary Reviews: Computational Molecular Science*, **2022**, Vol. 12, Issue 5, p. e1606.

0.9 Are there recommended programmes to use alongside ORCA?

As a matter of fact there are: We make extensive use of Chemcraft. It is interesting to note that it works well in MacOS or Linux (using Wine or a virtual machine). Another popular visualization programme is Chimera together with the SEQCROW plugin.

OpenBabel is very useful for file conversion to various chemical formats.

Finally, Avogadro is an excellent tool to edit molecular geometries. It is also able to generate ORCA input files. The Avogadro version with the latest ORCA modifications is available on the ORCA download site.

For other valuable questions/suggestions, please check out the ORCA forum.

0.10 My old inputs don't work with the new ORCA version! Why?

Please be aware that between ORCA revisions keywords and defaults might have changed or keywords may have been deprecated (for detailed information please see the Release Notes). It is therefore not unexpected that the same inputs will now either give slightly different results, or will totally crash. If you are unsure about an input, please consult the manual. It is provided by the ORCA developers and should contain all information implemented in the published version of ORCA.

0.11 My SCF calculations suddenly die with 'Please increase MaxCore'! Why?

The SCF cannot restrict its memory to a given MaxCore, which, in the past, has led to crashes due to lack of memory after many hours of calculation. To prevent this, the newer ORCA versions will try to estimate the memory needed at an early stage of the calculation. If this estimation is smaller than MaxCore, you are fine. If it is larger than MaxCore, but smaller than 2*MaxCore, ORCA will issue a warning and proceed. If the estimation yields a value that's larger than 2*MaxCore, ORCA will abort. You will then have to increase MaxCore. Please note, that MaxCore is the amount of memory dedicated to each process!

0.12 When dealing with array structures, when does ORCA count starting from zero and in which cases does counting start from one?

Since ORCA is a C++ based program its internal counting starts from zero. Therefore all electrons, atoms, frequencies, orbitals, excitation energies etc. are counted from zero. User-based counting such as the numeration of fragments is counted from one.

0.13 How can I check that my SCF calculation converges to a correct electronic structure?

The expectation value $\langle S^2 \rangle$ is an estimation of the spin contamination in the system. It is highly recommended for open-shell systems, especially with transition metal complexes, to check the UCO (unrestricted corresponding orbitals) overlaps and visualise the corresponding orbitals. Additionally, spin-population on atoms that contribute to the singly occupied orbitals is also an identifier of the electronic structure.

0.14 I can't locate the transition state (TS) for a reaction expected to feature a low/very low barrier, what should I do?

For such critical case of locating the TS, running a very fine (e.g. 0.01 Å increment of the bond length) relaxed scan of the key reaction coordinate is recommended. In this way the highest energy point on a very shallow surface can be identified and used for the final TS optimisation.

0.15 During the geometry optimisation some atoms merge into each other and the optimisation fails. How can this problem be solved?

This usually occurs due to the wrong or poor construction of initial molecular orbital involving some atoms. Check the basis set definition on problematic atoms and then the corresponding MOs!

0.16 While using MOREAD feature in ORCA, why am I getting an error saying, "no orbitals found in the .gbw file"?

ORCA produces the `.gbw` file immediately after it reads the coordinates and basis set information. If you put a `.gbw` file from a previous calculation with same base name as your current input into the working directory, it will be overwritten and the previous orbital data will be lost. Therefore, it is recommended to change the file name or `.gbw` extension to something else (`.gbw.old`, for example).

0.17 With all the GRID and RI and associated basis set settings I'm getting slightly confused. Can you provide a brief overview?

Hartree–Fock (HF) and DFT require the calculation of Coulomb and exchange integrals. While the Coulomb integrals are usually done analytically, the exchange integrals can be evaluated semi-numerically on a grid. Here, the pure DFT exchange is calculated on one type of grid (controlled through the `GRID` keyword) while the HF exchange can be evaluated on a different, often smaller grid (`GRIDX`). For both parts, further approximations can be made (RI-J and RI-K[1] or COSX, respectively). When RI is used, auxiliary basis sets are required (<basis >/ J for RI-J and <basis >/ JK for RI-JK). The following possible combinations arise:

- HF calculation
  - Exact J + exact K: no auxiliary functions and no grids needed.
  - RIJ + exact K (`RIJONX`, `RIJDX`): <basis >/ J auxiliaries, no grids.
  - RIJ + RIK = RIJK: <basis >/JK auxiliaries, no grids.
  - RIJ + COSX: <basis >/ J auxiliaries, COSX grid controlled by the `GRIDX` keyword.
- GGA DFT functional
  - Exact J + GGA-XC: no auxiliary functions needed, DFT grid controlled by the `GRID` keyword.
  - RIJ + GGA-XC: <basis >/ J auxiliaries, DFT grid controlled by the `GRID` keyword.
- Hybrid DFT functional
  - Exact J + exact K + GGA-XC: no auxiliary functions needed, DFT grid controlled by the `GRID` keyword.
  - RIJ + exact K (`RIJONX`, `RIJDX`) + GGA-XC: <basis >/ J auxiliaries, DFT grid controlled by the `GRID` keyword.
  - RIJ + RIK (RIJK) + GGA-XC: <basis >/ JK auxiliaries, DFT grid controlled by the `GRID` keyword.
  - RIJ + COSX + GGA-XC: <basis >/ J auxiliaries, COSX grid controlled by the `GRIDX` keyword, DFT grid controlled by the `GRID` keyword.

[1] Note that ORCA can only use RI-K in conjunction with RI-J; hence the combination RI-JK.

0.18 There are a lot of basis sets! Which basis should I use when?

ORCA offers a variety of methods and a large choice of basis sets to go with them. Here is an incomplete overview:

Method	Approximation	basis set (and auxiliaries)
CASSCF/NEVPT2		<basis >
CASSCF/NEVPT2	RI-JK	<basis >+ <basis >/JK
CASSCF/NEVPT2	RIJCOSX	<basis >+ <basis >/J + <basis >/C
CASSCF/NEVPT2	TrafoStep RI	<basis >+ <basis >/JK or <basis >/C
NEVPT2-F12	TrafoStep RI	<basis >-F12 + <basis >-F12/CABS + <basis >/JK or <basis >/C
TDDFT		<basis >
TDDFT	Mode RIInts	<basis >+ <basis >/C
MP2		<basis >
F12-MP2		<basis >-F12 + <basis >-F12/CABS
RI-MP2		<basis >+ <basis >/C
HF+RI-MP2	RIJCOSX	<basis >+ <basis >/C + <basis >/J
F12-RI-MP2		<basis >-F12 + <basis >-F12/CABS + <basis >/C
DLPNO-MP2		<basis >+ <basis >/C
HF+DLPNO-MP2	RI-JK	<basis >+ <basis >/C + <basis >/JK
F12-DLPNO-MP2		<basis >-F12 + <basis >-F12/CABS + <basis >/C
CCSD		<basis >
RI-CCSD		<basis >+ <basis >/C
(D)LPNO-CCSD		<basis >+ <basis >/C
HF+(D)LPNO-CCSD	RIJCOSX	<basis >+ <basis >/C + <basis >/J
F12-CCSD		<basis >-F12 + <basis >-F12/CABS
F12-RI-CCSD		<basis >-F12 + <basis >-F12/CABS + <basis >/C
HF+F12-RI-CCSD	RI-JK	<basis >-F12 + <basis >-F12/CABS + <basis >/C + <basis >/JK

CHAPTER

ONE

GENERAL INFORMATION

1.1 Program Components

The program system ORCA consists of several separate programs that call each other during a run. The following basic modules are included in this release (listed in alphabetical order):

orca	Main input+driver program
orca_ailft	Ab Initio Ligand Field Theory
orca_autoci	CI type program using the automated generation environment (ORCA-AGE)
orca_casscf	Main program for CASSCF driver
orca_casscfresp	CASSCF static linear response
orca_cclib	Precalculation of one particle coupling coefficients for ACCCI
orca_ciprep	Preparation of data for MRCI calculations (frozen core matrices and the like)
orca_cipsi	Iterative Configuration Expansion Configuration Interaction (ICE-CI)
orca_cis	Excited states via CIS and TD-DFT
orca_crystalprep	Set up an embedding calculation from a crystal structure file
orca_eca	Auxiliary program for solving magnetic model Hamiltonians
orca_esd	Excited state dynamics program for calculating vibronic spectra, resonance Raman, ...
orca_guess	Generation of an intial guess for SCF and CASSCF
orca_leanscf	Self consistent field program for HF and DFT
orca_lft	Ligand Field Theory
orca_loc	Calculation of localized molecular orbitals
orca_mcrpa	CASSCF linear response for excited states
orca_md	Molecular dynamics program
orca_mdci	Matrix driven correlation program: CI, CEPA, CPF, QCISD, CCSD(T)
orca_mp2	MP2 program (conventional, direct and RI)
orca_mrci	MRCI and MRPT calculations (individually selecting)
orca_ndoint	Calculates semiempirical integrals and gradients
orca_nmrspectrum	Simulates NMR spectra from calculated NMR parameters
orca_numfreq	Numerical Hessian computation
orca_pc	Addition of point charge terms to the one-electron matrix
orca_plot	Generation of orbital and density plots
orca_pop	External program for population analysis on a given density
orca_prop	Calculation of molecular properties
orca_propint	Calculation of property integrals
orca_rel	(Quasi) Relativistic corrections
orca_rocis	Excited states via the ROCIS method
orca_scfgrad	Analytic derivatives of SCF energies (HF and DFT)
orca_scfresp	self consistent field response
orca_startup	Calculation of startup data before each single point
orca_vpot	Calculation of the electrostatic potential on a given molecular surface
orca_vpt2	VPT2 analysis

Utility programs:

orca_2aim	Produces WFN and WFX files suitable for AIM analysis
orca_2json	Converts information from the gbw-file into JSON files
orca_2mkl	Produces an ASCII file to be read by molekel, molden or other visualization programs
orca_asa	Calculation of absorption, fluorescence and resonance Raman spectra
orca_chelpg	Electrostatic potential derived charges
orca_euler	Calculate Euler angles from `.property.txt` file
orca_exportbasis	Prints out any basis set in ORCA or GAMESS-US format
orca_fitpes	Simple program to fit potential energy curves of diatomics
orca_mapspc	Produces files for transfer into plotting programs
orca_mergefrag	Merges MO coefficients from two independent .gbw files
orca_pltvib	Produces files for the animation of vibrations
orca_pnmr	Calculation of paramagnetic NMR shielding tensors
orca_vib	Calculation of vibrational frequencies from a completed frequency run (also used for isotope shift calculations)
otool_gcp	Geometrical Counterpoise Correction
otool_xtb	For this release, xtb 6.7.1 needs to be downloaded from the Grimme lab

Friends of ORCA:

gennbo	The NBO analysis package of Weinhold (must be purchased separately from the university of Wisconsin; older versions available for free on the internet may also work)
Molekel	Molecular visualization program (see *Interface to Molekel*)
gOpenMol	Molecular visualization program (see *Interface to gOpenMol*)
Avogadro	Molecular builder and visualization program with ORCA support (see download page and original repository)
Chemcraft	Molecular builder and visualization program with ORCA support (see download page)
SEQCROW	Molecular builder and visualization program with ORCA support (see repository)

In principle every individual module can also be called "standalone". However, it is most convenient to do everything via the main module.

There is no real installation procedure. Just copy the executables wherever you want them to be and make sure that your path variable contains a reference to the directory where you copied the files. This is important to make sure that the programs can call each other (but you can also tell the main program the explicit position of the other programs in the input file as described below). The xtb tool (recommended version 6.7.1 or higher) needs to be downloaded separately from the Grimme lab's repository. The xtb binary needs to be copied to the directory to which the orca binaries were copied to.

1.2 Units and Conversion Factors

Internally the program uses atomic units. This means that the unit of energy is the Hartree (Eh) and the unit of length is the Bohr radius (a_0). The following conversion factors to other units are used:

1 Eh	=	27.2113834 eV	
1 eV	=	8065.54477 cm^{-1}	= 23.0605 $\frac{kcal}{mol}$
1 cm^{-1}	=	29979.2458 MHz	
1 a_0	=	0.5291772083 Å	
1 a.t.u.	=	2.4188843 10^{-17} s	

CHAPTER

TWO

THE ARCHITECTURE OF ORCA

In this chapter we describe the broad features of ORCA's internal structure. This knowledge is not necessary to carry out calculations with ORCA. It may, however, help to understand which modules are being called in which order and why this is happening in the sequence it does.

2.1 The structure of the ORCA source code

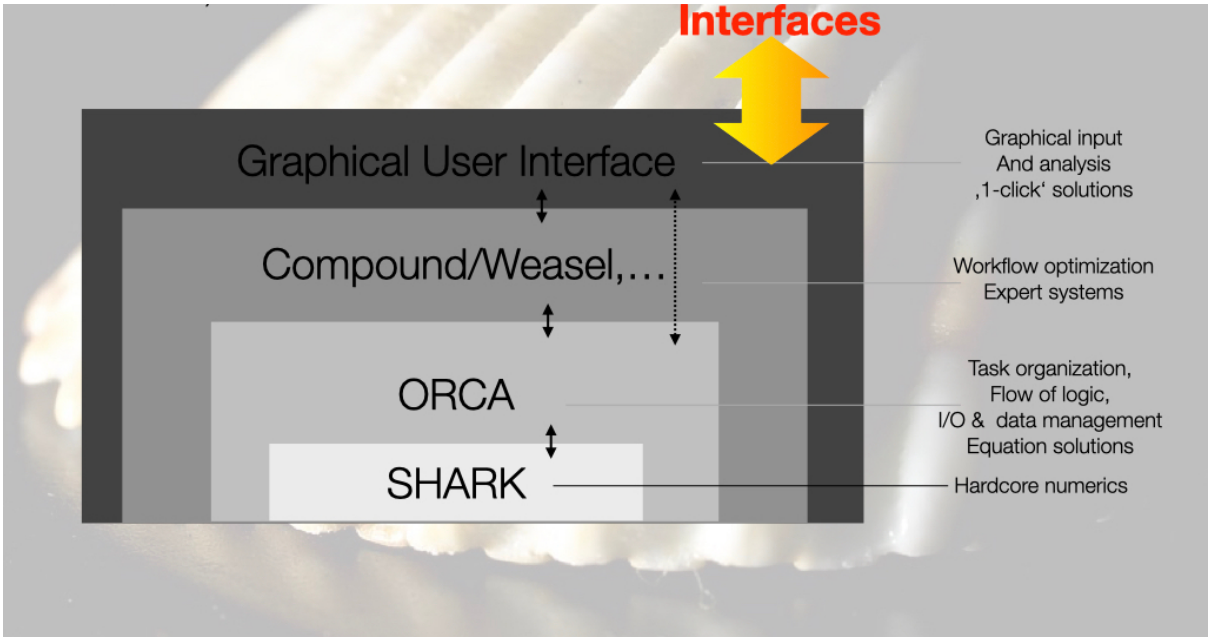
The source code of ORCA is broadly structured in three parts:

1) The main program.
2) The code for the "ORCA tool library".
3) The code for individual modules.

The code in the ORCA tool library is being compiled into one library file that subsequently is linked with all ORCA modules. The code for the individual modules can make use of everything that is in the library, but the modules are not supposed to link to or use any code of the main program or any other module. This way of structuring the code ensures that the modules remain maintainable and that there are no complex and unwanted interdependencies that would make it eventually impossible to exchange modules for new code.

2.2 The shell structure of ORCA

The whole organization of the code and the information flow can be thought of as consisting of a shell structure where a shell is a layer of software with well-defined functions and a well-defined interface to the shells above and below it.



The lowest layer of the shell consists of the SHARK integral package. This is the "motor" of the program that takes care of most compute intensive tasks and essentially all parts that involve integrals over basis functions. This amounts to calculating all one- and two-electron integrals, forming Fock and Fock-like matrices and performing integral transformations using all kinds of basis sets and kernels. Most of the SHARK code is independent of the remaining ORCA source code infrastructure.

The layer above SHARK is the actual ORCA source code. ORCA interacts with the user through the interface and it orchestrates the entire workflow of the calculation. At this point in time, it also performs compute intensive tasks like forming diagonalizing and handling Fock-operators , solving large linear equation systems, forming sigma vectors, densities etc in correlated calculations and similar tasks. The ORCA source code interacts with SHARK through a small set of DRIVERS. The drivers feature genuine ORCA data structures as well as SHARK infrastructure. The DRIVERS call the genuine SHARK functions (ORCA calls SHARK directly only in a few places) and, most importantly, the DRIVERS take care of finding their way through the approximation jungle. By this we mean they handle the necessary basis sets, auxiliary basis sets, grids, integral approximations, solvation, relativity, point charges etc.

Concentrating these important and highly error prone tasks in a number of well documented routines provides a highly transparent way of identifying and properly maintaining the functionality that is at the heart of the program.

Above the ORCA software layer there are tools that orchestrate workflows. Workflows typically consisting of a number of interdepend computational steps that are later combined into a single meaningful chemical result. For example, one may optimize the geometry with a DFT functional and calculate zero-point and thermodynamic corrections at the same level and follow it up with a single point calculation at the coupled cluster level that may or may not feature another correction for core correlation at the MP2 level or complete basis set extrapolation. Such tasks and many others, like running series of calculations on a test-sets of molecules or permuting calculation options like functionals or basis sets on a given test system can be addressed with workflow tools. Inside ORCA, there is the very powerful compound scripting language to achieve and organize such tasks. On the commercial branch, FAccTs develops the Weasel tool that is a powerful and highly reliable workflow organizer.

At the final layer, one might envision a graphical user interface (GUI) that helps building molecules and facilitates running calculations and analyzing results. At this point in time, ORCA does not have a dedicated GUI. There are many free and commercial solutions that interface to ORCA. This form of interfacing is facilitated by the `orca_2json` interface and the property file that ORCA produces. We hope that the transparency of this interface motivates GUI developers to provide ever improved GUIs that interface ORCA. We do not exclude the possibility that ORCA will feature it's own GUI sometime down the road.

2.3 The master/slave concept and the calling sequence

Within ORCA, the code follows a "master/slave" concept in which the main program is allowed to know everything about every module while the modules are not allowed to know anything about the main program. Thus, the main program is the piece of software that orchestrates the entire ORCA run and the interaction with the user. The main program is, however, not supposed to carry out compute intensive tasks or even parallelized tasks by itself.

An ORCA calculation commences with reading and analyzing the ORCA input file. The plausibility of this input is checked by an elaborate "maincheck" routine. Quite obviously, the number of possible combinations of ORCA features is far larger than what could possibly be checked with realistic effort. However, the maincheck routine has evolved over the year in a way that allows to detect the most common combinations of invalid or inconsistent input keywords and take appropriate action (that can be changing some parameters or aborting the job).

After through maincheck the basis sets used in the calculations and the coordinates are completely known. This is then the time to initialize the SHARK integral package and will then carry out the bulk of the computation heavy tasks.

The first module that is being called is `orca_startup`. This module takes care of calculating the one electron integrals such as the HCORE and overlap matrices, the metric matrices in RI calculations, the pre-screening matrices for direct SCF and possibly also the two-electron integrals for integral conventional runs.

The next module downstream is the `orca_guess` program. This module has the task to produce an initial set of molecular orbitals and also an initial density matrix from these orbitals. To this end, there are a number of options out of which the most common ones are to a model density guess (PMODEL) or to read orbitals from a previous calculation (MOREAD). In the latter case, the calculation may feature a different geometry and/or a different basis set but the number, nature and ordering of atoms need to be consistent with the target system.

In the third step, the program branches out into either the SCF or the CASSCF module (`orca_leanscf` or `orca_casscf`). These modules produces converged orbitals and a self-consistent field energy as well as one-particle density matrix. The latter is stored in in the `DensityContainer`, which is a centralized storage facility for densities that will also be left over after the calculation and can be accessed by the users.

The next step of the calculation consists of calling various post-SCF programs like `orca_mp2`, `orca_mdci` or `orca_autoci`. NEVPT2 and CASPT2 are calculated by code that is integrated with `orca_casscf`. These calculate correlation energies and excited states among many other things.

There are many different pathways that the program can take next, for example geometry optimization or embedding calculations or molecular dynamics to only name a few. Rather than going into each and every one of the possible pathways, we will describe the calculation of molecular properties inside ORCA.

2.4 The calculation of molecular properties

For the calculation of molecular properties, ORCA has a fairly unique and strongly streamlined infrastructure that is focusing on the commonalities in the calculated properties and that are independent of the underlying electronic structure method used.

There are three qualitatively different sets of properties that are covered in ORCA property calculations:

1) Response properties. These are properties that can be formulated as derivatives of the total energy.

2) QDPT properties. These are properties that are calculated by quasi-degenerate perturbation theory (QDPT). This amount to diagonalizing the Hamiltonian containing external fields and relativistic corrections over a number of roots delivered by the underlying electronic structure method

3) Excited state properties: these are, at least at this point in time, transition moments between different electronic states of the system.

2.4.1 Response properties

A large number of properties can be written as derivatives of the total energy. These are first order properties:

$$X_M = \frac{\partial E}{\partial M} = \sum_{\mu\nu} P^{\pm}_{\mu\nu} \left\langle \mu \middle| \widehat{h}_M \middle| \nu \right\rangle$$

Here, $P^{\pm}_{\mu\nu}$ is the first-order electron ('+') or spin ('-') density $\widehat{h}_M$ is the operator that is describing the property of interest (e.g. the dipole moment operator) and the basis set is $\{\mu\}$.

Second order properties are:

$$X_{MN} = \frac{\partial^2 E}{\partial M \partial N} = \sum_{\mu\nu} P^{\pm}_{\mu\nu} \left\langle \mu \middle| \widehat{h}_{MN} \middle| \nu \right\rangle + \sum_{\mu\nu} \frac{\partial P^{\pm}_{\mu\nu}}{\partial N} \left\langle \mu \middle| \widehat{h}_M \middle| \nu \right\rangle$$

Here the important quantity is $\frac{\partial P^{\pm}_{\mu\nu}}{\partial N}$, the first derivative (or "response") of the electron or spin density with respect to perturbation N. It can be shown that the order of perturbations M and N is immaterial and hence one can choose the more convenient perturbation to calculate the response density for.

It is important to point out that the equations for response properties can *always* be brought into this form, irrespective of whether the calculation is a Hartree-Fock, DFT, MP2, coupled cluster, CASSCF or full-CI wavefunction.

Given the considerable generality of the property equations, it seems logical to create an infrastructure in which these similarities are exploited to the largest possible extent. In ORCA the central place where all densities and response densities are stored is the `DensityContainer`. This is used intensely throughout the actual calculation and left on disk as `BaseName.densities` where it can be used for visualization.

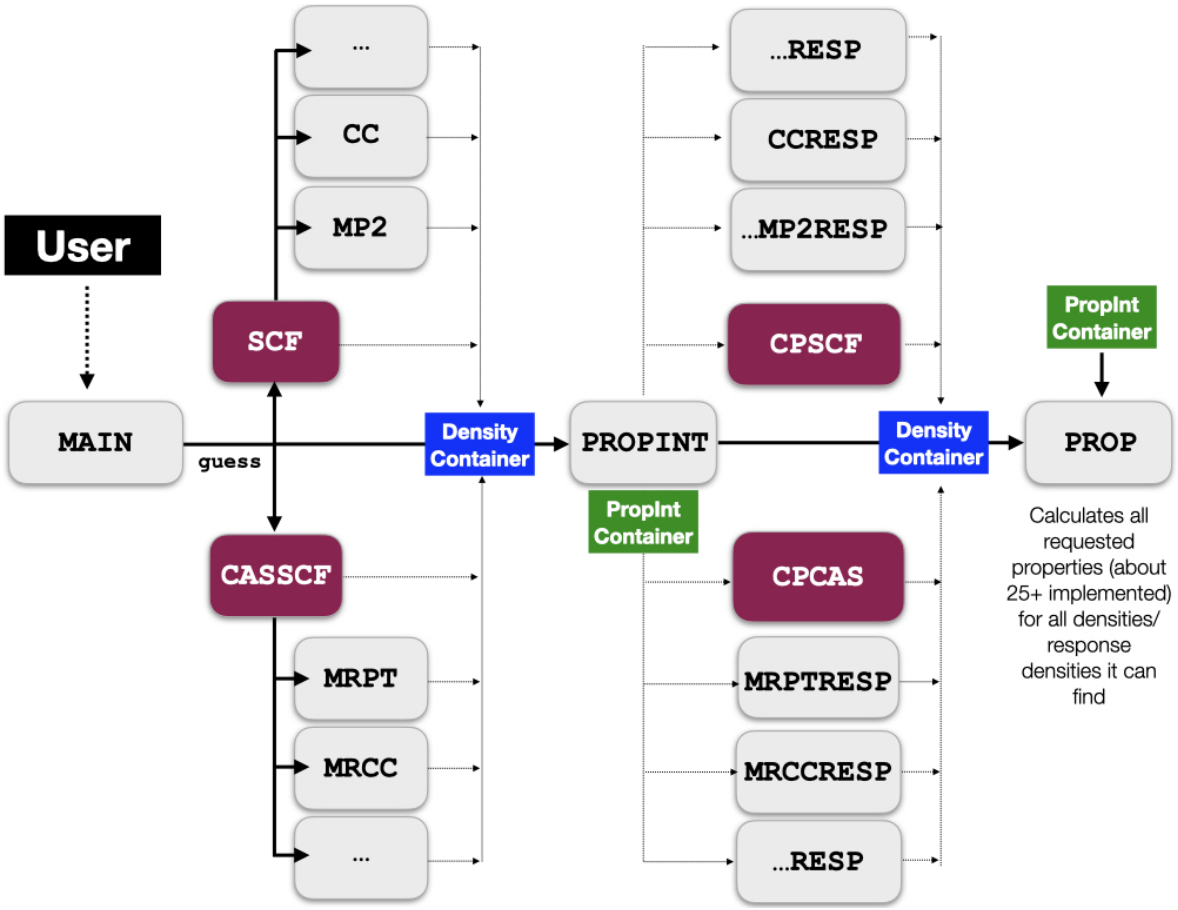
In terms of the calculation flow the main program contains all the logic to drive this calculation. It first drives the SCF and possibly the post-SCF calculations. After this is done, the SCF program collects the information about which property integrals will be needed down the road and call the property integral program `orca_propint` that will calculate the necessary integrals and stores them in another central storage container, the "property integral container".

The next step of the calculation is that the main program determines for which properties response densities are needed. It collects these perturbations and calls the response modules. In the case of a SCF wavefunction (HF or DFT), this is the `orca_scfresp` program. This program will divide the requested perturbations into different types like, real-symmetric, imaginary-antisymmetric, triplet etc. Then it solves the response equations for all types of perturbations simultaneously. This leads to large efficiency gains since the same integrals are needed on the left-hand side of the coupled-perturbed SCF (CP-SCF) equations irrespective of the actual perturbations. Hence, pooling all perturbations of a given type together is much more efficient than treating these perturbations one at a time. This leads to high efficiency gains, in particular for properties like NMR parameters where nucleus dependent perturbations are required.

For a number of other electronic structure methods like MP2, CASSCF or AutoCI the respective modules can be run in "response mode" where instead of solving the equations for the energy, they pick up the amplitudes that were determined in the energy run and use them together with the property integrals in order to produce response densities.

The response densities are then stored in the density container too. At this point, one has everything that is required in order to calculate the actual properties – a task that is performed by `orca_prop`.

The `orca_prop` program works by browsing through the density container and looks for densities that are appropriate for calculating a requested property. As soon as it finds an appropriate density (or combination of unperturbed and response density), it will calculate the property. For example, if you have calculated SCF, MP2 and CCSD in the same calculation and have asked for the calculation of all three densities, `orca_prop` will calculate three dipole moments and print them right next to each other.



One beneficial side effect of this organization is that it is very well suited for future extensions. If a new method becomes available in ORCA that is able to produce densities and possibly response densities, the properties for this method will be fully available without any programmer needing to write a single line of additional code. It also ensures that all properties are calculated in a consistent fashion and with a consistent printout.

The results of all property calculations are stored in the central property file that will be left after the calculation. Users interested in reading properties, it should be read from the property file or its JSON translation.

2.4.2 QDPT properties

Some properties are not calculated as energy derivatives but from quasi degenerate perturbation theory. In this method one diagonalizes the QDPT matrix:

$$Q_{IJ} = \left\langle \Psi_I^{SM} \middle| E_I \delta_{IJ} + \widehat{H}_{SOC} + \widehat{H}_{SS} + \widehat{H}_{LB} + \ldots \middle| \Psi_J^{S'M'} \right\rangle$$

Here $\left|\Psi_I^{SM}\right\rangle$ are the roots of a given method that can generate excited states with energy E_I. For example, these can be TD-DFT roots, CASSCF roots, CASSCF roots with energy corrections from NEVPT2 or CASPT2, MRCI-roots, EOM-CCSD roots etc. and $\widehat{H}_{SOC}$ is the spin-orbit coupling (SOC) operator, $\widehat{H}_{SS}$ the electron-electron spin-spin coupling operator, $\widehat{H}_{LB}$ the molecular Zeeman operator etc. In practice, only the principle component $M = S$ (i. e. Ψ_I^{SS}) is calculated and the necessary matrix elements are generated using the Wigner-Eckart theorem.

The diagonalization produces the complex valued relativistic eigenstates as linear combinations of the non-relativistic or scalar relativistic eigenstates. Using the eigenstates relativistic densities or transition densities can be calculated that subsequently can be used to calculate magnetic properties or "relativistic" optical or magneto-optical spectra.

The procedure obviously suffers from a truncation error because only a finite number of roots can be calculated in practice. However, results indicate that the QDPT generated properties compare often very favorable with experimental data provided that the underlying electronic structure method delivers reasonable results.

In ORCA, all QDPT properties are calculated in a consistent infrastructure and then also leads to consistent printing and reporting of the results.

2.4.3 Excited state properties

Closely related to the QDPT infrastructure is the "one-photon absorption" (OPA) infrastructure. This is an infrastructure that coordinates the calculation of transition moments using a set of transition densities as input. These might have been stored on disk or might have been generated on the fly. In a similar way to the QDPT infrastructure, these transition moments are generated in a consistent manner throughout all modules of ORCA that can generate excited states.

The exact field/matter coupling Hamiltonian is:

$$A(\mathbf{r},t) = -A_0 \varepsilon \exp(i\mathbf{kr} - \omega t)$$

Here $\mathbf{k}$ is the wavevector of the radiation with frequency ω and ε is its polarization. A_0 is the intensity of the radiation, $\mathbf{r}$ the electronic coordinate and t is the time.

Evaluation of the matrix element $\left\langle \Psi_I^{SM} \middle| \sum_i A(\mathbf{r}_i,t) \middle| \Psi_J^{S'M'} \right\rangle$ amount to generating the "exact" field-matter transition moments. This can be requested by setting the flag `DoFullSemiClassical = true` in the appropriate ORCA input block.

A series expansion of the cosine term yields the familiar electric dipole, electric quadrupole, magnetic dipole etc. contributions. Interestingly, the direct evaluation of the electric dipole term would yield it in the gauge independent "velocity" form. They are related to the more familiar "length" form dipole matrix elements by:

$$\left\langle \Psi_I^{SM} \middle| \sum_i \mathbf{p}_i \middle| \Psi_J^{S'M'} \right\rangle = \frac{-i}{\hbar}(E_J - E_I) \left\langle \Psi_I^{SM} \middle| \sum_i \mathbf{r}_i \middle| \Psi_J^{S'M'} \right\rangle$$

The results of the length and velocity transition moment evaluation are expected to match in the basis set limit if the electronic structure method at hand satisfies certain conditions. In practice, they usually do not agree very well, even if large basis sets are used.

In order to generate electric length and velocity as well as higher moment evaluations, you can use the keywords

```
DoDipoleLength   = true
DoDipoleVelocity = true
DoHigherMoments  = true
```

in the appropriate ORCA input blocks that trigger the respective excited state calculation.

CHAPTER

THREE

CALLING THE PROGRAM (SERIAL AND PARALLEL)

3.1 Calling the Program

Under Windows the program is called from the command prompt! (Make sure that the PATH variable is set such that the orca executables are visible)

```
orca MyMol.inp > MyMol.out
```

Under UNIX based operating systems the following call is convenient[^1] (here also: make sure that the PATH variable is set to the directory where the orca executables reside):

```
orca MyMol.inp >& MyMol.out &
```

The program writes to *stdout* and *stderr*. Therefore the output must be redirected to the file *MyMol.out* in this example. *MyMol.inp* is a free format ASCII file that contains the input description. The program will produce a number of files *MyMol.x.tmp* and the file *MyMol.gbw*. The "*.gbw" file contains a binary summary of the calculation. GBW stands for "Geometry-Basis-Wavefunction". Basically this together with the calculation flags is what is stored in this file. You need this file for restarting SCF calculations or starting other calculations with the orbitals from this calculation as input. The "*.tmp" files are temporary files that contain integrals, Fock matrices etc. that are used as intermediates in the calculation. If the program exits normally all of these files are deleted. If it happens to crash you have to remove the files manually (`rm MyMol*.tmp` under Unix or `del MyMol*.tmp` under Windows). In case you want to monitor the output file while it is written, you can use the command (under Unix):

```
tail -f MyMol.out
```

to follow (option -f) the progress of the calculation. Under Windows you have to either open another command shell and use:

```
type MyMol.out
type MyMol.out |more
```

or you have to copy the output file to another file and then use any text editor to look at it.

```
copy MyMol.out temp.out
edit temp.out
```

you cannot use `edit MyMol.out` because this would result in a sharing violation.

3.2 Calling the Program with Multiple Processes

There are parallel versions for Linux, MAC and Windows computers (thanks to the work of Ms Ute Becker) which make use of OpenMPI (open-source MPI implementation) and Microsoft MPI (Windows only). Most of the important modules in ORCA can run in parallel or in multi-process mode:

3.2.1 List of Parallelized Modules

The following modules and utility programs are presently parallelized/usable in multi-process mode:

AUTOCI
CASSCF / NEVPT2 / CASSCFRESP
CIPSI
CIS/TDDFT
GRAD (general Gradient program)
GUESS
LEANSCF (memory conserving SCF solver)
MCRPA
MDCI (Canonical- and DLPNO-Methods)
MM
MP2 and RI-MP2 (including Gradients)
MRCI
PC
PLOT
PNMR
POP
PROP
PROPINT
REL
ROCIS
SCFGRAD
SCFRESP (with SCFHessian)
STARTUP
VPOT
Numerical Gradients, Frequencies, Overtones-and-Combination-Bands
VPT2
NEB (Nudged Elastic Band)

Thus, all major modules are parallelized in the present version. The efficiency is such that for RI-DFT perhaps up to 16 processors are a good idea while for hybrid DFT and Hartree-Fock a few more processors are appropriate. Above this, the overhead becomes significant and the parallelization loses efficiency. Coupled-cluster calculations usually scale well up to at least 8 processors but probably it is also worthwhile to try 16. For Numerical Frequencies or Gradient runs it makes sense to choose nprocs = 4 or 8 times 6*Number of Atoms. For VPT2 on larger systems you may well even try 16 times 6*Number of Atoms - if you use multiple processes per displacement. (Please check out the section *Hints on the Use of Parallel ORCA* what you have to take care of for such kind of calculations.)

If you run a queuing system you have to make sure that it works together with ORCA in a reasonable way.

> **Note**
>
> Parallelization is a difficult undertaking and there are many different protocols that work differently for different machines. Please understand that we can not provide support for each and every platform. We are trying our best to make the parallelization transparent and provide executables for various platforms but we can not possibly guarantee that they always work on every system. Please see the download information for details of the version.

3.2.2 Hints on the Use of Parallel ORCA

Many questions that are asked in the discussion forum deal with the parallel version of ORCA. Please understand that we cannot possibly provide one-on-one support for every parallel computer in the world. So please, make every effort to solve the technical problems locally together with your system administrator. Here are some explanations about what is special to the parallel version, which problems might arise from this and how to deal with them:

1. Parallel ORCA can be used with OpenMPI (on Linux and MAC) or MS-MPI (on windows) only. Please see the download information for details of the relevant OpenMPI-version for your platform.

   The OpenMPI version is configurable in a large variety of ways, which cannot be covered here. For a more detailed explanation of all available options, cf. http://www.open-mpi.org

2. Please note that the OpenMPI version is dynamically linked, that is, it needs at runtime the OpenMPI libraries (and several other standard libraries)! If you compile MPI on your own computer, you also need to have a fortran compiler, as `mpirun` will contain fortran bindinngs.

   (Remember to set `PATH` and `LD_LIBRARY_PATH` to mpirun and the mpi libraries)

3. Many problems arise, because parallel ORCA does not find its executables. To avoid this, it is crucial to call ORCA with its complete pathname. The easiest and safest way to do so is to include the directory with the orca-executables in your `$PATH`. Then start the calculation:

   ```
   - interactively:
     start orca with full path: "/mypath_orca_executables/orca MyMol.inp"
   - batch :
     set your path: `export PATH=/mypath_orca_executables:$PATH` (for bash) then
     start orca with full path: "$PATH/orca $jobname.inp"
   ```

   This seems redundant, but it really is important if you want to run a parallel calculation to call ORCA with the full path! Otherwise it will not be able to find the parallel executables.

4. Assuming that the MPI libraries are installed properly on your computer, it is fairly easy to run the parallel version of ORCA. You simply have to specify the number of parallel processes, like:

   ```
   ! PAL4  # everything from PAL2 to PAL8 and Pal16, Pal32, Pal64 is recognized
   ```

   or

   ```
   %pal nprocs 4 # any number (positive integer)
        end
   ```

   The parallelized modules of ORCA are started by the (serial) ORCA-Driver. If the driver finds `PAL4` or `%pal nprocs 4 end` (e.g.) in the input, it will start up the parallel modules instead of the serial ones. So - please do not start the driver with mpirun! (Please see below for what else has to be taken care of for a successfull parallel run.)

5. It is recommended to run orca in local (not nfs-mounted) scratch-directories, (`/tmp1` or `/usr/local` e.g.) and to renew these directories for each run to avoid confusion with left-overs of a previous run.

6. It has proven convenient to use “wrapper” scripts. These scripts should

   ```
   - set the path
   - create local scratch directories
   - copy input files to the scratch directory
   - start orca
   - save your results
   - remove the scratch directory
   ```

   A basic example of such a submission script for the parallel ORCA version is shown below (this is for the Torque/PBS queuing system, running on Apple Mac OS X):

```
#!/bin/zsh

setopt EXTENDED_GLOB
setopt NULL_GLOB
#export MKL_NUM_THREADS=1

b=${1:r}

#get number of procs.... close your eyes... (it really works!)
if [[ ${$(grep -e '^!' $1):u} == !*(#b)PAL(<0-9>##)* ]]; then
  nprocs=$match
let "nodes=nprocs"
elif [[ ${(j: :) $(grep -v '^#' $1):u} == *%(#b)PAL*NPROCS' '#(<0-9>##)* ]]; then
  nprocs=$match
  let "nodes=nprocs"
fi

cat > ${b}.job <<EOF
#!/bin/zsh
#PBS -l nodes=1:ppn=${nodes:=1}
#PBS -S /bin/zsh
#PBS -l walltime=8760:00:00

setopt EXTENDED_GLOB
setopt NULL_GLOB
export PATH=$PBS_O_PATH

logfile=$PBS_O_WORKDIR/${b}.log
tdir=$(mktemp -d /Volumes/scratch/$USER/${b}__XXXXXX)

trap '
echo "Job terminated from outer space!" >> $logfile
rm -rf $tdir
exit
' TERM

cp $PBS_O_WORKDIR/$1 $tdir
foreach f ($PBS_O_WORKDIR/*.gbw $PBS_O_WORKDIR/*.pot) { cp $f $tdir }
cd $tdir

echo "Job started from ${PBS_O_HOST}, running on $(hostname) in $tdir using
$(which orca)" > $log
file
=orca $1 1>>$logfile 2>&1

cp ^(*.(inp|tmp*))  $PBS_O_WORKDIR/
rm -rf $tdir

EOF

qsub -j oe -o ${b}.job.out ${b}.job
```

7. Parallel ORCA distinguishes the following cases of disk availability:

   - each process works on its own (private) scratch directory (the data on this directory cannot be seen by any other process). This is flagged by **"working on local scratch directories"**
   - all processes work in a common scratch directory (all processes can see all file-data) ORCA will distinguish two situations:
     - all processes are on the same node - flagged by **"working on a common directory"**
     - the processes are distributed over multiple nodes but accessing a shared filesysten - flagged by **"working on a shared directory"**

- there are at least 2 groups of processes on different scratch directories, one of the groups consisting of more than 1 process - flagged by **"working on distributed directories"**

Parallel ORCA will find out, which of these cases exists and will handle the I/O respectively. If ORCA states disk availability differently from what you would expect, check the number of available nodes and/or the distribution pattern (fill_up/round_robin)

8. If Parallel ORCA finds a file named "`MyMol.nodes`" in the directory where it's running, it will use the nodes listed in this file to start the processes on, provided your input file was "`MyMol.inp`". You can use this file as your machinefile specifying your nodes, using the usual OpenMPI machinefile notation.

```
node1 cpu=2
node2 cpu=2
```

or

```
node1
node1
node2
node2
```

> **Note**
>
> If you run the parallel ORCA version on only one computer, you do not need to provide a nodefile, and neither have to enable an rsh/ssh access, as in this case the processes will simply be forked! If you start ORCA within a queueing system you also don't need to provide a nodefile. The queueing system will care for it.

9. It is possible to pass additional MPI-parameters to mpirun by adding these arguments to the ORCA call - all arguments enclosed in a single pair of quotes:

```
/mypath_orca_executables/orca MyMol.inp "--bind-to core"
```

– or – for multiple arguments

```
/mypath_orca_executables/orca MyMol.inp "--bind-to core --verbose"
```

10. If the MPI-environment variables are not equally defined on all participating compute nodes it might be advisable to export these variables. This can be achieved by passing the following additional parameters to mpirun via the ORCA call:

```
/mypath_orca_executables/orca MyMol.inp "-x LD_LIBRARY_PATH -x PATH"
```

11. An additional remark on multi-process numerical calculations (frequencies, gradient, hybrid Hessian): The processes that execute these calculations do not work in parallel, but independently, often in a totally asynchronous manner. The numerical calculations will start as many processes, as you dedicated for the parallel parts before and they will run on the same nodes. If your calculation runs on multiple nodes, you have to set the environment variable `RSH_COMMAND` to either "`rsh`" or "`ssh`". If `RSH_COMMAND` is not defined, ORCA will abort. This prevents that all processes of a multi-node run are started on the 'master'-node.

12. On multiple user request the 'parallelization' of NumCalc has been made more flexible. If before ORCA would start nprocs displacements with a single process each, the user can now decide on how many processes should work on a single displacement.

    For this the nprocs keyword got a sibling:

```
%pal nprocs       32 # or nprocs_world - total number of parallel processes
     nprocs_group  4 #                 - number of parallel processes per sub-task
     end
```

This setting will ORCA make use 32 processes, with 4 processes working on the same displacement, thus running 8 displacements simultaneously. The methods that can profit from this new feature are

- all NumCalc-methods: as NumGrad, NumFreq, VPT2, Overtones, NEB, and GOAT.
- the analytical Hessian, leading to a nice increase of parallel performance for really large calculations.

It is highly recommended to choose nprocs_group to be an integer divisor of nprocs_world!

For convenient use a couple of standard 'groupings' are made available via simple input keyword:

```
PAL4(2x2)  - 2 groups a 2 workers
PAL8(4x2)  - 4 groups a 2 workers
PAL8(2x4)  - ...
PAL16(4x4)
PAL32(8x4)
PAL32(4x8)
PAL64(8x8)
```

Note

If your system-administration does not allow to connect via rsh/ssh to other compute nodes, you unfortunately cannot make use of parallel sub-calculations within NumCalc runs. This affects NEB as well as GOAT, VPT2, Overtone-and-Combination-Bands, as well as Numerical Frequencies and Gradients.

CHAPTER

FOUR

GENERAL STRUCTURE OF THE INPUT FILE

In general, the input file is a free format ASCII file and can contain one or more keyword lines that start with a "!" sign, one or more input blocks enclosed between an "%" sign and "end" that provide finer control over specific aspects of the calculation, and finally the specification of the coordinates for the system along with the charge and multiplicity provided either with a `%coords` block, or more usually enclosed within two "*" symbols. Here is an example of a simple input file that contains all three input elements:

```
! HF def2-TZVP

%scf
   convergence tight
end

* xyz 0 1
C  0.0  0.0  0.0
O  0.0  0.0  1.13
*
```

Comments in the file start by a "#". For example:

```
# This is a comment. Continues until the end of the line
```

Comments can also be closed by a second "#", as the example below where `TolE` and `TolMaxP` are two variables that can be user specified:

```
TolE=1e-5;    #Energy conv.#  TolMaxP=1e-6; #Density conv.#
```

The input may contain several blocks, which consist of logically related data that can be user controlled. The program tries to choose sensible default values for all of these variables. However, it is impossible to give defaults that are equally sensible for all systems. In general the defaults are slightly on the conservative side and more aggressive cutoffs etc. can be chosen by the user and may help to speed things up for actual systems or give higher accuracy if desired.

4.1 Input Blocks

The following blocks exist:

`autoci`	Controls autogenerated correlation calculations
`basis`	Basis sets are specified
`casresp`	Control of CASSCF static linear response calculations
`casscf`	Control of CASSCF/NEVPT2 and DMRG calculations
`cipsi`	Control of Iterative-Configuration Expansion Configuration Interaction calculation
`cim`	Control of Cluster In Molecules calculations
`cis`	Control of CIS and TD-DFT calculations (synonym is `tddft`)
`coords`	Input of atomic coordinates

continues on next page

Table 4.1 – continued from previous page

compound	Control of compound
cosmors	Control of ORCA/COSMO-RS calculations
cpcm	Control of the Conductor-like Polarizable Continuum Model
elprop	Control of electric property calculations
eprnmr	Control of EPR and NMR calculations
esd	Control of ESD calculations
freq	Control of frequency calculations
geom	Control of geometry optimization
irc	Control of intrinsic reaction coordinate calculations
loc	Localization of orbitals
mcrpa	Control CASSCF linear response calculations
md	Control of molecular dynamics simulation
mdci	Controls single reference correlation methods
method	Here a computation method is specified
mp2	Controls the details of the MP2 calculation
mrcc	Control of multi-reference CC calculations
mrci	Control of MRCI calculations
neb	Control of NEB calculations
numgrad	Control of numerical gradients
nbo	Controls NBO analysis with GENNBO
output	Control of output
pal	Control of parallel jobs
paras	Input of semi-empirical parameters
plots	Control of plot generation
rel	Control of relativistic options
rocis	Control of restricted-open-shell CIS
rr	Control of resonance Raman and absorption/fluorescence band-shape calculations
scf	Control of the SCF procedure
symmetry	Control of spatial symmetry recognition

Blocks start with "%" and end with "end". Note that input is **not** case sensitive. For example:

```
%method  method HF
         end
```

No blocks *need* to be present in an input file but they *can* be present if detailed control over the behavior of the program is desired. Otherwise all normal jobs can be defined via the keywords described in the next section. Variable assignments have the following general structure:

```
VariableName Value
```

Some variables are actually arrays. In this case several possible assignments are useful:

```
Array[1]  Value1
Array[1]  Value1,Value2,Value3
Array  Value1,Value2
```

Note

Arrays always start with index 0 in ORCA (this is because ORCA is a C++ program). The first line in the example gives the value "Value1" to Array[1], which is the *second* member of this array. The second line assigns Value1 to Array[1], Value2 to Array[2] and Value3 to Array[3]. The third line assigns Value1 to Array[0] and Value2 to Array[1].

Strings (for examples filenames) must be enclosed in quotes. For example:

```
MOInp  "Myfile.gbw"
```

In general the input is not case sensitive. However, inside strings the input is case sensitive. This is because on unix systems `MYFILE.GBW` and `MyFile.gbw` are different files. Under Windows the file names are not case sensitive.

4.2 Keyword Lines

It is possible to give a line of keywords that assign certain variables that normally belong to different input blocks. The syntax for this "simple input" is line-oriented. A keyword line starts with the "!" sign.

```
! Keywords
```

4.2.1 Main Methods and Options

Table 4.2 provides a list of keywords that can be used within the "simple input" keyword line to request specific methods and/or algorithmic options. Most of them are self-explanatory. The others are explained in detail in the section of the manual that deals with the indicated input block.

Table 4.2: Main keywords that can be used in the simple input of ORCA.

Keyword	Input block	Variable	Comment
HF	METHOD	METHOD	Selects the Hartree-Fock method
DFT			Selects the DFT method (see section *Density Functional Methods* for a list of functionals)
FOD			FOD analysis (see *Fractional Occupation Number Weighted Electron Density (FOD)*) employing default settings (TPSS/def2-TZVP, TightSCF, SmearTemp = 5000 K)

Runtypes

Keyword	Input block	Variable	Comment
ENERGY or SP	METHOD	RUNTYP	Selects a single point calculation
OPT			Selects a geometry optimization calculation (using 2022 internal redundant coordinates)
COPT			Optimization in Cartesian coordinates (if you are desperate)
ENGRAD			Selects an energy and gradient calculation
NUMGRAD			Numerical gradient (has explicitly to be asked for, if analytic gradient is not available)
NUMFREQ			Numerical frequencies
NUM-NACME			Numerical non-adiabatic coplings (only for CIS/TD-DFT)
MD			Molecular dynamic simulation
CIM			Cluster-In-Molecule calculation

Atomic mass/weight handling

Keyword	Input block	Variable	Comment
Mass2016	METHOD	AMASS	Use the latest (2016) atomic masses of the most abundant or most stable isotopes instead of atomic weights.
PAF			Shift and rotate the molecule into its principle axis frame using the 2016 atomic masses by default.

Symmetry handling

Keyword	Input block	Variable	Comment
UseSym			Turns on the use of molecular symmetry (see section *ORCA and Symmetry*). *THIS IS VERY RUDIMENTARY!*
NoUseSym			Turns symmetry off

Second order many body perturbation theory

Keyword	Input block	Variable	Comment
MP2			Selects Method=HF and DoMP2=true
MP2RI or RI-MP2			Select the MP2-RI method
SCS-MP2			Spin-component scaled MP2
RI-SCS-MP2			Spin-component scaled RI-MP2
			(synonym is SCS-RI-MP2)
OO-RI-MP2			Orbital optimized RI-MP2
OO-RI-SCS-MP2			Orbital optimized and spin-component scaled RI-MP2
MP2-F12			MP2 with F12 correction
			(synonym is F12-MP2)
MP2-F12-RI			MP2-RI with RI-F12 correction
MP2-F12D-RI			MP2-RI with RI-F12 correction employing the D approximation (less expensive), (synonyms are F12-RI-MP2, RI-MP2-F12)

High-level single reference methods

These are implemented in the MDCI module. They can be run in a number of technical variants.

Keyword	Input block	Variable	Comment
CCSD	MDCI	CITYPE	Coupled-cluster singles and doubles
CCSD(T)			Same with perturbative triples correction
CCSD-F12			CCSD with F12 correction
CCSD(T)-F12			CCSD(T) with F12 correction
CCSD-F12/RI			CCSD with RI-F12 correction
CCSD-F12D/RI			CCSD with RI-F12 correction employing the D approximation (less expensive)
CCSD(T)-F12/RI			CCSD(T) with RI-F12 correction
CCSD(T)-F12D/RI			CCSD(T) with RI-F12 correction employing the D approximation (less expensive)
QCISD			Quadratic Configuration interaction
QCISD(T)			Same with perturbative triples correction
QCISD-F12			QCISD with F12 correction
QCISD(T)-F12			QCISD(T) with F12 correction
QCISD-F12/RI			QCISD with RI-F12 correction
QCISD(T)-F12/RI			QCISD(T) with RI-F12 correction
NCPF/1			A “new” modified coupled-pair functional
CEPA/1			Coupled-electron-pair approximation
NCEPA/1			The CEPA analogue of NCPF/1
RI-CEPA/1-F12			RI-CEPA with F12 correction
MP3			MP3 energies
SCS-MP3			Grimme’s refined version of MP3

Other coupled-pair methods are available and are documented later in the manual in detail (section 7.8) In general you can augment the method with RI-METHOD in order to make the density fitting approximation operative; RI34-METHOD does the same but only for the 3- and 4-external integrals). MO-METHOD performs a full four index transformation and AO-METHOD computes the 3- and 4-external contributions on the fly. With AOX-METHOD this is is done from stored AO integrals.

AUTOCI single reference methods

These single reference correlation methods are available in the AUTOCI.

Keyword	Input block	Variable	Comment
AUTOCI-CID	AUTOCI	CITYPE	Configuration Interaction with doubles
AUTOCI-CISD			CI with singles and doubles
AUTOCI-CISDT			CI with singles, doubles and triples
AUTOCI-CEPA(0)			Zeroth-order Coupled-Electron pair approximation
AUTOCI-CCD			Coupled-Cluster with doubles
AUTOCI-CCSD			Coupled-Cluster with singles, doubles
AUTOCI-CCSDT			Coupled-Cluster with singles, doubles, triples
AUTOCI-CCSDT-1			Approximate CCSDT-1 model
AUTOCI-CCSDT-2			Approximate CCSDT-2 model
AUTOCI-CCSDT-3			Approximate CCSDT-3 model
AUTOCI-CCSDT-4			Approximate CCSDT-4 model
AUTOCI-QCISD			Quadratic CISD model
AUTOCI-CC2			Approximate CC with singles, doubles
AUTOCI-CC3			Approximate CC with singles, doubles, triples
AUTOCI-CCSD(T)			Coupled-Cluster with singles, doubles and pertubative triples
AUTOCI-CCSD[T]			Coupled-Cluster with singles, doubles and pertubative triples
AUTOCI-MP2			Second-order Moller-Plesset PT
AUTOCI-MP3			Third-order Moller-Plesset PT
AUTOCI-MP4(SDQ)			Fourth-order Moller-Plesset PT without triples
AUTOCI-MP4			Forth-order Moller-Plesset PT
AUTOCI-MP5			Fifth-order Moller-Plesset PT

Local correlation methods

These are local, pair natural orbital based correlation methods. They must be used together with auxiliary correlation fitting basis sets. Open-shell variants are available for some of the methods, for full list please see section *Coupled-Cluster and Coupled-Pair Methods*.

Keyword	Input block	Variable	Comment
DLPNO-CCSD			Domain based local pair natural orbital coupled-cluster method with single and double excitations (closed-shell only)
DLPNO-CCSD(T)			DLPNO-CCSD with perturbative triple excitations
DLPNO-CCSD(T1)			DLPNO-CCSD with iterative perturbative triple excitations
DLPNO-MP2	MP2	Various	Local (DLPNO) MP2
DLPNO-SCS-MP2			Spin-component scaled DLPNO-MP2
			(a synonym is SCS-DLPNO-MP2)
DLPNO-MP2-F12			DLPNO-MP2 with F12 correction employing an efficient form of the C approximation
DLPNO-MP2-F12/D			DLPNO-MP2-F12 with approach D (less expensive than the C approximation)
DLPNO-CCSD-F12			DLPNO-CCSD with F12 correction employing an efficient form of the C approximation
DLPNO-CCSD-F12/D			DLPNO-CCSD-F12 with approach D (less expensive than the C approximation)
DLPNO-CCSD(T)-F12			DLPNO-CCSD(T) with F12 correction employing an efficient form of the C approximation
DLPNO-CCSD(T)-F12/D			DLPNO-CCSD(T)-F12 with approach D (less expensive than the C approximation)
DLPNO-CCSD(T1)-F12			DLPNO-CCSD(T1) with F12 correction employing an efficient form of the C approximation
DLPNO-CCSD(T1)-F12/D			DLPNO-CCSD(T1)-F12 with approach D (less expensive than the C approximation)
DLPNO-NEVPT2			DLPNO-NEVPT2 requires a CASSCF block

Accuracy control for local correlation methods

These keywords select predefined sensible sets of thresholds to control the accuracy of DLPNO calculations. See the corresponding sections on local correlation methods for more details.

Keyword	Input block	Variable	Comment
LoosePNO	MDCI, MP2	Various	Selects loose DLPNO thresholds
NormalPNO			Selects default DLPNO thresholds
TightPNO			Selects tight DLPNO thresholds
DLPNO-HFC1			Tightened truncation setting for DLPNO-CCSD hyperfine coupling constants calculation
DLPNO-HFC2			Tighter truncation setting than for DLPNO-HFC1

Automatic basis set extrapolation

Keyword	Input block	Variable	Comment
Extrapolate (n/m, bas)			Extrapolation of the basis set family "bas" (bas=cc,aug-cc, cc-core, ano, saug-ano, aug-ano, def2; if omitted "cc-pVnZ" is used) for cardinal numbers n,m (n<m=2,3,4,5), e.g. Extrapolate(2/3,cc) extrapolates the SCF, MP2 and MDCI energies to the basis set limit. "core" refers to basis sets with core correlation function. In this case the frozen core approximation is - by default - turned off. This setting can be overridden in the "methods" block if one just wants to use the basis set with core correlation functions (steep primitives) but without unfreezing the core electrons.
Extrapolate (n, bas)			Calculate the first n-energies for member of the basis set family basis, e.g. Extrapolate(3) is doing calculations with cc-pVDZ, cc-pVTZ and cc-pVQZ.
ExtrapolateEP2 (n/m, bas,[method,method-details])			Similar: performs SCF, MP2 and MDCI calculations. The higher basis set can only be done with DLPNO-CCSD(T) or MP2 methods and then used to extrapolate the MDCI calculation to the basis set limit.
ExtrapolateEP3 (bas,[method,method details])			Similar to EP2: for the high basis set method we go one cardinal number higher.

High-level methods for excited states as implemented in the MDCI module

An additional block input to define the number of roots is required. The EOM family of methods feature IP and EA extensions. The list below is incomplete as some methods need more refined settings such as the Hilbert space MRCC approaches (MkCCSD/BWCCSD). Note that excited states can also be computed with CIS, RPA, ROCIS and TD-DFT. Please check the excited states section of the manual for details.

Keyword	Input block	Variable	Comment
EOM-CCSD	MDCI	NRoots	Equation of Motion CCSD
bt-PNO-EOM-CCSD			back-transformed PNO approximation
STEOM-CCSD			Similarity Transformed Equation of Motion CCSD
bt-PNO-STEOM-CCSD			back-transformed PNO approximation
STEOM-DLPNO-CCSD			DLPNO approximation
IH-FSMR-CCSD			Fock-Space CCSD using an intermediate Hamiltonian
bt-PNO-IH-FSMR-CCSD			back-transformed PNO approximation

CASSCF related options

All of them require the CASSCF block as minimal input.

Keyword	Input block	Variable	Comment
DMRG	CASSCF	CIStep	Sets DMRG as "CIStep" in CASSCF
NEVPT2		PTMethod	SC NEVPT2
SC-NEVPT2			SC-NEVPT2 same as NEVPT2
RI-NEVPT2			SC-NEVPT2 with the RI approximation
FIC-NEVPT2			FIC-NEVPT2 aka PC-NEVPT2
DLPNO-NEVPT2			FIC-NEVPT2 in the framework of DLPNO
CASPT2			FIC-CASPT2
RI-CASPT2			FIC-CASPT2 with the RI approximation
DCD-CAS(2)		DCD-CAS	2nd order Dynamic Correlation Dressed CAS
RI-DCD-CAS(2)			2nd order Dynamic Correlation Dressed CAS with RI approximation

(internally contracted) Multireference methods beyond NEVPT2/CASPT2

If specified in a single keyword all information about reference spaces, number of roots etc. is taken from the CASSCF module that is assumed to be run in advance. These methods reside in the autoci module. More refined settings require the autoci block in the input.

Keyword	Input block	Variable	Comment
FIC-MRCI	AUTOCI	CIType	Invokes the fully internally contracted MRCI
FIC-DDCI3			Fully internally contracted DDCI3
FIC-DDCI3-C0			Fully internally contracted DDCI3-C0
FIC-CEPA(0)			Fully internally contracted CEPA(0)
FIC-ACPF			Fully internally contracted ACPF
FIC-AQCC			Fully internally contracted AQCC
FIC-MRCC			Fully internally contracted MRCCSD

(uncontracted) Multireference methods

If specified in a single keyword all information about reference spaces, number of roots etc. is taken from the CASSCF module that is assumed to be run in advance. In general, these calculations are of the individually selecting type and are very time consuming. Very many flags can be set and modified for these methods and in general using these methods requires expert users! In general see the variables `Tsel`, `Tpre` and `Tnat` that define the individual selection process. All of these methods can be used with RI integrals by using RI-MRCI etc. However, then the calculations become even more time consuming since integrals are made one- by one on the fly. Non-RI calculations will be pretty much limited to about 200-300 orbitals that are included in the CI.

Keyword	Input block	Variable	Comment
MRCI	MRCI	CIType	Initiates a multireference configuration interaction calculation with single and double excitations
MRCI+Q			Same with multireference Davidson correction for unlinked quadruples
MRACPF			Average coupled-pair functional
MRAQCC			Average quadratic coupled-cluster
MRDDCI1			Difference dedicated CI with one degree of freedom
MRDDCI2			Same with two degrees of freedom
MRDDCI3			Same with three degrees of freedom
MRDDCIn+Q			MRDDCI with Davidson correction
SORCI			Spectroscopy oriented CI

Frozen core features

Note

This deviates from previous versions of ORCA! We are now counting core electrons rather than using an energy window. If you do want to use an orbital energy window use `%method FrozenCore FC_EWIN end`. Otherwise the EWin commands will be ignored! (alternatives are `FC_ELECTRONS`(default) and `FC_NONE`).

Keyword	Input block	Variable	Comment
FROZENCORE	METHOD	FrozenCore	Use a frozen core. By default this is done by counting the number of chemical core electrons
NOFROZENCORE			Do not use a frozen core

Semiempirical methods

Keyword	Input block	Variable	Comment
ZINDO/S			Selects the ZINDO/S method
ZINDO/1			Selects the ZINDO/1 method
ZINDO/2			Selects the ZINDO/2 method
NDDO/1			Selects the NDDO/1 method
NDDO/2			Selects the NDDO/2 method
MNDO			Selects the MNDO method
AM1			Selects the AM1 method
PM3			Selects the PM3 method

Algorithmic variations, options, add-ons, modifiers, ...

Keyword	Input block	Variable	Comment
RHF or RKS	SCF	HFTYP	Selects closed-shell SCF
UHF or UKS			Selects spin unrestricted SCF
ROHF or ROKS			Selects open-shell spin restricted SCF
AllowRHF	METHOD	ALLOWRHF	Allow a RHF calculation even if the system is open-shell (Mult>1). Default is to switch to UHF then
RI	METHOD	RI	Sets RI=true to use the RI approximation in DFT calculations. Default to Split-RI-J
NORI			Sets RI=false
RIJCOSX	METHOD/ SCF	RI, KMatrix	Sets the flag for the efficient RIJCOSX algorithm (treat the Coulomb term via RI and the Exchange term via seminumerical integration)
RI-JK	METHOD/ SCF	RI, KMatrix	Sets the flag for the efficient RI algorithm for Coulomb and Exchange. Works for SCF (HF/DFT) energies and gradients. Works direct or conventional.
SPLITJ	SCF	JMATRIX	Select the efficient Split-J procedure for the calculation of the Coulomb matrix in non-hybrid DFT (rarely used)
SPLIT-RI-J	SCF	JMATRIX, RI	Select the efficient Split-RI-J procedure for the improved evaluation of the RI-approximation to the Coulomb-matrix
NoSplit-RI-J	SCF	JMATRIX, RI	Turns the Split-RI-J feature off (but does not set the RI flag to false!)
RI-J-XC	SCF	JMATRIX, KMATRIX, RI	Turn on RI for the Coulomb term *and* the XC terms. This saves time when the XC integration is significant but introduces another basis set incompleteness error. (rarely used)
DIRECT	SCF	SCFMODE	Selects an integral direct calculation
CONV			Selects an integral conventional calculation
NOITER	SCF	MAXITER	Sets the number of SCF iterations to 0. This works together with MOREAD and means that the program will work with the provided starting orbitals.

Initial guess options

In most cases the default PMODEL guess will be adequate. In some special situations you may want to switch to a different choice.

Keyword	Input block	Variable	Comment
PATOM	SCF	GUESS	Selects the polarized atoms guess
PMODEL			Selects the model potential guess
HUECKEL			Selects the extended Hückel guess
HCORE			Selects the one-electron matrix guess
MOREAD			Read MOs from a previous calulation (use `%moinp "myorbitals.gbw"` in a separate line to specify the GBW file that contains these MOs to be read)
AUTOSTART		AUTOSTART	Try to start from the existing GBW file of the same name as the present one (only for single-point calculations)
NOAUTOSTART			Don't try to do that

Basis-set related keywords

Keyword	Input block	Variable	Comment
DecontractBas	BASIS	DecontractBas	Decontract the basis set. If the basis set arises from general contraction, duplicate primitives will be removed.
NoDecontractBas		NoDecontractBas	Do not decontract the basis set
DecontractAuxJ		DecontractAuxJ	Decontract the AuxJ basis set
NoDecontractAuxJ		NoDecontractAuxJ	Do not decontract the AuxJ basis
DecontractAuxJK		DecontractAuxJK	Decontract the AuxJK basis set
NoDecontractAuxJK		NoDecontractAuxJK	Do not decontract the AuxJK basis
DecontractAuxC		DecontractAuxC	Decontract the AuxC basis set
NoDecontractAuxC		NoDecontractAuxC	Do not decontract the AuxC basis
Decontract		Decontract	Decontract all (orbital and auxiliary) basis sets

Relativistic options

There are several variants of scalar relativistic Hamiltonians to use in all electron calculations. See *Relativistic Options* for details.

Keyword	Input block	Variable	Comment
DKH or DKH2	REL	METHOD/ORDER	Selects the scalar relativistic Douglas–Kroll–Hess Hamiltonian of 2nd order
ZORA	REL	METHOD	Selects the scalar relativistic ZORA Hamiltonian
X2C	REL	METHOD	Selects the scalar relativistic X2C Hamiltonian
DLU-X2C	REL	METHOD/DLU	Selects the scalar relativistic X2C Hamiltonian with the diagonal local approximation to the unitary transformation matrix

Grid options

Keyword	Input block	Variable	Comment
DEFGRIDn ($n = 1$–3)	METHOD	GRID	Selects the integration grids
NOFINALGRIDX			Turn off the final grid in COSX (not recommended)

Convergence thresholds

These keywords control how tightly the SCF and geometry optimizations will be converged. The program makes an effort to set the convergence thresholds for correlation modules consistently with that of the SCF.

Keyword	Input block	Variable	Comment
NORMALSCF	SCF	CONVERGENCE	Selects normal SCF convergence
LOOSESCF			Selects loose SCF convergence
SLOPPYSCF			Selects sloppy SCF convergence
STRONGSCF			Selects strong SCF convergence
TIGHTSCF			Selects tight SCF convergence
VERYTIGHTSCF			Selects very tight SCF convergence
EXTREMESCF			Selects "extreme" convergence. All thresholds are practically reduced to numerical precision of the computer. Only for benchmarking (very expensive).
SCFCONVn			Selects energy convergence check and sets $ETol$ to 10^{-n} (n = 6–10). Also selects appropriate *thresh*, *tcut*, and *bfcut* values.
VERYTIGHTOPT	GEOM	TolE, TolRMSG	Selects very tight optimization convergence
TIGHTOPT		TolMaxG	Selects tight optimization convergence
NORMALOPT		TolRMSD, TolMaxD	Selects default optimization convergence
LOOSEOPT			Selects loose optimization convergence

Convergence acceleration

The default is DIIS which is robust. For most closed-shell organic molecules SOSCF converges somewhat better and might be a good idea to use. For "trailing convergence", KDIIS or the trust-region augmented Hessian procedures TRAH-SCF might be good choices.

Keyword	Input block	Variable	Comment
DIIS	SCF	DIIS	Turns DIIS on
NODIIS			Turns DIIS off
KDIIS	SCF	KDIIS	Turns Kollmar's DIIS on
TRAH	SCF	TRAH	Turns trust-region augmented Hessian SCF on
NOTRAH			Turns trust-region augmented Hessian SCF off
SOSCF	SCF	SOSCF	Turns SOSCF on
NOSOSCF			Turns SOSCF off
DAMP	SCF	CNVDAMP	Turns damping on
NODAMP			Turns damping off
LSHIFT	SCF	CNVSHIFT	Turns level shifting on
NOLSHIFT			Turns level shifting off

Convergence strategies

(does not modify the convergence criteria)

Keyword	Input block	Variable	Comment
EasyConv			Assumes no convergence problems.
NormalConv			Normal convergence criteria.
SlowConv			Selects appropriate SCF converger criteria for difficult cases. Most transition metal complexes fall into this category.
VerySlowConv			Selects appropriate SCF converger criteria for very difficult cases.
CPCM(solvent)	CPCM		Invoke the conductor-like polarizable continuum model with a standard solvent (see section *[sec:solvationmodels.detailed]* for a list of solvents). If no solvent is given, infinity (a conductor) is assumed.

Spin-orbit coupling

Keyword	Input block	Variable	Comment
SOMF(1X)	REL	SOCType, SOCFlags	Invokes the SOMF(1X) treatment of the spin-orbit coupling operator.
RI-SOMF(1X)			Invokes the SOMF(1X) treatment of the spin-orbit coupling operator, with RI for the Coulomb part.
SOMF(4X)	REL	SOCType, SOCFlags	Invokes the SOMF(4X) treatment of the spin-orbit coupling operator.
RI-SOMF(4X)			Invokes the SOMF(4X) treatment of the spin-orbit coupling operator, with RI for the Coulomb part.
SOMF(4XS)	REL	SOCType, SOCFlags	Invokes the SOMF(4XS) treatment of the spin-orbit coupling operator.
RI-SOMF(4XS)			Invokes the SOMF(4XS) treatment of the spin-orbit coupling operator, with RI for the Coulomb part.
VEFF-SOC			Invokes the VEFF-SOC treatment of the spin-orbit coupling operator.
VEFF(-2X)-SOC			Invokes the VEFF(-2X)-SOC treatment of the spin-orbit coupling operator.
AMFI			Invokes the AMFI treatment of the spin-orbit coupling operator.
ZEFF-SOC			Uses effective nuclear charges for the spin-orbit coupling operator.

Miscellaneous options

Keyword	Input block	Variable	Comment
ANGS	COORDS	UNITS	Select angstrom units
BOHRS			Select input coordinates in atomic units
FRACOCC	SCF	FRACOCC	Turns the fractional occupation option on (FOD is always calculated in this case)
NoPropFile	Method	Method	Turns writing to property file off. By default is on for everything, except MD and L-Opt calculations
SMEAR	SCF	SMEARTEMP	Temperature for occupation number smearing on (default is 5000 K; FOD (see *Fractional Occupation Number Weighted Electron Density (FOD)*) is always calculated in this case)
NOSMEAR			Turn occupation number smearing off
KEEPINTS	SCF	KEEPINTS	Keep two electron integrals on disk
NOKEEPINTS			Do not keep two electron integrals
KEEPDENS	SCF	KEEPDENS	Keep the density matrix on disk
NOKEEPDENS			Do not keep the density matrix
KEEPTRANSDENSITY		KEEPTRANSDENSITY	Keep the transition density matrices on disk
READINTS	SCF	READINTS	Reading of two electron integrals on
NOREADINTS			Reading of two electron integrals off
CHEAPINTS	SCF	USECHEAPINTS	Use the cheap integral feature in direct SCF calculations
NOCHEAPINTS			Turn that feature off
FLOAT	SCF	VALFORMAT	Set storage format for numbers to single precision (SCF, RIMP2, CIS, CIS(D), MDCI)
DOUBLE	SCF	VALFORMAT	Set storage format for numbers to double precision (default)
UCFLOAT	SCF	VALFORMAT COMPRESSION	Use float storage in the matrix containers without data compression
CFLOAT	SCF	VALFORMAT COMPRESSION	Use float storage in the matrix containers with data compression
UCDOUBLE	SCF	VALFORMAT COMPRESSION	Use double storage in the matrix containers without data compression
CDOUBLE	SCF	VALFORMAT COMPRESSION	Use double storage in the matrix containers with data compression

Output control

Keyword	Input block	Variable	Comment
NORMALPRINT	OUTPUT	PRINTLEVEL	Selects the normal output
MINIPRINT			Selects the minimal output
SMALLPRINT			Selects the small output
LARGEPRINT			Selects the large output
PRINTMOS	OUTPUT	Print[p_MOS]	Prints MO coefficients
NOPRINTMOS	OUTPUT		Suppress printing of MO coefficients
PRINTBASIS	OUTPUT	Print[p_basis]	Print the basis set in input format
PRINTGAP	OUTPUT	Print[p _homolu-mogap]	Prints the HOMO/LUMO gap in each SCF iteration. This may help to detect convergence problems
ALLPOP	OUTPUT	Print[. . .]	Turns on all population analysis
NOPOP			Turns off all populaton analysis
MULLIKEN			Turns on the Mulliken analysis
NOMULLIKEN			Turns off the Mulliken analysis
LOEWDIN			Turns on the Loewdin analysis
NOLOEWDIN			Turns off the Loewdin analysis
MAYER			Turns on the Mayer analysis
NOMAYER			Turns off the Mayer analysis
NPA			Turns on interface for the NPA analysis using the GENNBO program
NBO			Turns on the interface for the NPA plus NBO analysis with the GENNBO program
NONPA			Turns off NPA analysis
NONBO			Turns of NBO analysis
REDUCEDPOP			Prints Loewdin reduced orb.pop per MO
NOREDUCED-POP			Turns this feature off
UNO	SCF	UNO	Produce UHF natural orbitals
AIM			Produce a WFN file
XYZFILE	OUTPUT	XYZFILE	Produce an XYZ coordinate file
PDBFILE		PDBFILE	Produce a PDB file

Nudged Elastic Band methods

Keyword	Input block	Variable	Comment
NEB			Selects standard NEB method
ZOOM-NEB			ZOOM-NEB method
NEB-IDPP			Initial path NEB calculation
NEB-CI			Climbing image NEB calculation
ZOOM-NEB-CI			Zoom version of NEB-CI
NEB-MMFTS			NEB + subsequent MMF-TS optimization
NEB-TS			NEB + subsequent transition state optimization
ZOOM-NEB-TS			ZOOM-NEB + subsequent transition state optimization
FLAT-NEB-TS			
FAST-NEB-TS			NEB with one iteration + subsequent transtion state optimization
LOOSE-NEB-TS			
TIGHT-NEB-TS			Select tight convergence criteria

Compression and storage

The data compression and storage options deserve some comment: in a number of modules including RI-MP2, MDCI, CIS, (D) correction to CIS, etc. the program uses so called "Matrix Containers". This means that the data to be processed is stored in terms of matrices in files and is accessed by a double label. A typical example is the exchange operator $\mathbf{K^{ij}}$ with matrix elements $K^{ij}(a,b) = (ia|jb)$. Here the indices i and j refer to occupied orbitals of the reference state and a and b are empty orbitals of the reference state. Data of this kind may become quite large (formally N^4 scaling). To store the numbers in single precision cuts down the memory requirements by a factor of two with (usually very) slight loss in precision. For larger systems one may also gain advantages by also compressing the data (e.g. use a "packed" storage format on disk). This option leads to additional packing/unpacking work and adds some overhead. For small molecules UCDOUBLE is probably the best option, while for larger molecules UCFLOAT or particularly CFLOAT may be the best choice. Compression does not necessarily slow the calculation down for larger systems since the total I/O load may be substantially reduced and thus (since CPU is much faster than disk) the work of packing and unpacking takes less time than to read much larger files (the packing may reduce disk requirements for larger systems by approximately a factor of 4 but it has not been extensively tested so far). There are many factors contributing to the overall wall clock time in such cases including the total system load. It may thus require some experimentation to find out with which set of options the program runs fastest with.

Caution

- It is possible that FLOAT may lead to unacceptable errors. Thus it is not the recommended option when MP2 or RI-MP2 gradients or relaxed densities are computed. For this reason the default is DOUBLE.
- If you have convinced yourself that FLOAT is OK, it may save you a factor of two in both storage and CPU.

Global memory use

Some ORCA modules (in particular those that perform some kind of wavefunction based correlation calculations) require large scratch arrays. Each module has an independent variable to control the size of these dominant scratch arrays. However, since these modules are never running simultaneously, we provide a global variable `MaxCore` that assigns a certain amount of scratch memory to all of these modules. Thus:

```
%MaxCore 4000
```

sets 4000 MB (= 4 GB) as the limit for these scratch arrays. **This limit applies per processing core**. Do not be surprised if the program takes more than that – this size only refers to the dominant work areas. Thus, you are well advised to provide a number that is significantly less than your physical memory. Note also that the memory use of the SCF program cannot be controlled: it dynamically allocates all memory that it needs and if it runs out of physical memory you are out of luck. This, however, rarely happens unless you run on a really small memory computer or you are running a gigantic job.

4.2.2 Density Functional Methods

For density functional calculations a number of standard functionals can be selected via the "simple input" feature. Since any of these keywords will select a DFT method, the keyword "DFT" is not needed in the input. Further functionals are available via the `%method` block. References are given in Section *[sec:model.dft.functionals.detailed]*

Local and gradient corrected functionals

HFS	Hartree–Fock–Slater Exchange only functional
LDA or LSD	Local density approximation (defaults to VWN5)
VWN or VWN5	Vosko-Wilk-Nusair local density approx. parameter set "V"
VWN3	Vosko-Wilk-Nusair local density approx. parameter set "III"
PWLDA	Perdew-Wang parameterization of LDA
BP86 or BP	Becke '88 exchange and Perdew '86 correlation
BLYP	Becke '88 exchange and Lee-Yang-Parr correlation
OLYP	Handy's "optimal" exchange and Lee-Yang-Parr correlation
GLYP	Gill's '96 exchange and Lee-Yang-Parr correlation
XLYP	The Xu and Goddard exchange and Lee-Yang-Parr correlation
PW91	Perdew-Wang '91 GGA functional
mPWPW	Modified PW exchange and PW correlation
mPWLYP	Modified PW exchange and LYP correlation
PBE	Perdew-Burke-Erzerhoff GGA functional
RPBE	"Modified" PBE
REVPBE	"Revised" PBE
RPW86PBE	PBE correlation with refitted Perdew '86 exchange
PWP	Perdew-Wang '91 exchange and Perdew '86 correlation

Hybrid functionals

B1LYP	The one-parameter hybrid functional with Becke '88 exchange and Lee-Yang-Parr correlation (25% HF exchange)
B3LYP and B3LYP/G	The popular B3LYP functional (20% HF exchange) as defined in the TurboMole program system and the Gaussian program system, respectively
O3LYP	The Handy hybrid functional
X3LYP	The Xu and Goddard hybrid functional
B1P	The one-parameter hybrid version of BP86
B3P	The three-parameter hybrid version of BP86
B3PW	The three-parameter hybrid version of PW91
PW1PW	One-parameter hybrid version of PW91
mPW1PW	One-parameter hybrid version of mPWPW
mPW1LYP	One-parameter hybrid version of mPWLYP
PBE0	One-parameter hybrid version of PBE
REVPBE0	"Revised" PBE0
REVPBE38	"Revised" PBE0 with 37.5% HF exchange
BHANDHLYP	Half-and-half hybrid functional by Becke

Meta-GGA and hybrid meta-GGA functionals

TPSS	The TPSS meta-GGA functional
TPSSh	The hybrid version of TPSS (10% HF exchange)
TPSS0	A 25% exchange version of TPSSh that yields improved energetics
M06L	The Minnesota M06-L meta-GGA functional
M06	The M06 hybrid meta-GGA (27% HF exchange)
M062X	The M06-2X version with 54% HF exchange
PW6B95	Hybrid functional by Truhlar
B97M-V	Head-Gordon's DF B97M-V with VV10 nonlocal correlation
B97M-D3BJ	Modified version of B97M-V with D3BJ correction by Najibi and Goerigk
B97M-D4	Modified version of B97M-V with DFT-D4 correction by Najibi and Goerigk
SCANfunc	Perdew's SCAN functional
r2SCAN	Regularized and restored SCAN functional by Furness, Sun et. al.
r2SCANh	Global hybrid variant of r^2 SCAN with 10% HF exchange
r2SCAN0	Global hybrid variant of r^2 SCAN with 25% HF exchange
r2SCAN50	Global hybrid variant of r^2 SCAN with 50% HF exchange

Range-separated hybrid functionals

wB97	Head-Gordon's fully variable DF ωB97
wB97X	Head-Gordon's DF ωB97X with minimal Fock exchange
wB97X-D3	Chai's refit incl. D3 in its zero-damping version
wB97X-D4	Modified version of ωB97X-V with DFT-D4 correction by Najibi and Goerigk
wB97X-D4rev	Modified version of ωB97X-V with DFT-D4 correction by Grimme et al.
wB97X-V	Head-Gordon's DF ωB97X-V with VV10 nonlocal correlation
wB97X-D3BJ	Modified version of ωB97X-V with D3BJ correction by Najibi and Goerigk
wB97M-V	Head-Gordon's DF ωB97M-V with VV10 nonlocal correlation
wB97M-D3BJ	Modified version of ωB97M-V with D3BJ correction by Najibi and Goerigk
wB97M-D4	Modified version of ωB97M-V with DFT-D4 correction by Najibi and Goerigk
wB97M-D4rev	Modified version of ωB97M-V with DFT-D4 correction by Grimme et al.
CAM-B3LYP	Handy's fit
LC-BLYP	Hirao's original application
LC-PBE	range-separated PBE-based hybrid functional with 100% Fock exchange in the long-range regime
wr2SCAN	Range-separated hybrid variant of r^2 SCAN with 0-100% HF exchange

Perturbatively corrected double-hybrid functionals

Add the prefix RI- or DLPNO- to use the respective approximation for the MP2 part.

B2PLYP	Grimme's mixture of B88, LYP, and MP2
mPW2PLYP	mPW exchange instead of B88, which is supposed to improve on weak interactions.
B2GP-PLYP	Gershom Martin's "general purpose" reparameterization
B2K-PLYP	Gershom Martin's "kinetic" reparameterization
B2T-PLYP	Gershom Martin's "thermochemistry" reparameterization
PWPB95	Goerigk and Grimme's mixture of modified PW91, modified B95, and SOS-MP2
PBE-QIDH	Adamo and co-workers' "quadratic integrand" double hybrid with PBE exchange and correlation
PBE0-DH	Adamo and co-workers' PBE-based double hybrid
DSD-BLYP	Gershom Martin's "general purpose" double-hybrid with B88 exchange, LYP correlation and SCS-MP2 mixing, i.e. not incl. D3BJ correction
DSD-PBEP86	Gershom Martin's "general purpose" double-hybrid with PBE exchange, P86 correlation and SCS-MP2 mixing, i.e. not incl. D3BJ correction
DSD-PBEB95	Gershom Martin's "general purpose" double-hybrid with PBE exchange, B95 correlation and SCS-MP2 mixing, i.e. not incl. D3BJ correction
revDSD-PBEP86/2021, revDSD-PBEP86-D4/2021	Double-Hybrid Functional with with PBE exchange, B95 correlation and SCS-MP2 Mixing
revDOD-PBEP86/2021, revDOD-PBEP86-D4/2021	Double-Hybrid Functional with with PBE exchange, B95 correlation and SOS-MP2 Mixing
Pr2SCAN50	Global SOS-double-hybrid variant of r^2 SCAN with 50% HF exchange
Pr2SCAN69	Global SOS-double-hybrid variant of r^2 SCAN with 69% HF exchange
kPr2SCAN50	Global SOS-double-hybrid variant of r^2 SCAN with 50% HF exchange and kappa-regularized MP2

Range-separated double-hybrid functionals

Add the prefix RI- or DLPNO- to use the respective approximation for the MP2 part.

wB2PLYP	Goerigk and Casanova-Páez's range-separated DHDF, with the correlation contributions based on B2PLYP, optimized for excitation energies
wB2GP-PLYP	Goerigk and Casanova-Páez's range-separated DHDF, with the correlation contributions based on B2GP-PLYP, optimized for excitation energies
RSX-QIDH	range-separated version of the PBE-QIDH double-hybrid by Adamo and co-workers
RSX-0DH	range-separated version of the PBE-0DH double-hybrid by Adamo and co-workers
wB88PP86	Casanova-Páez and Goerigk's range-separated DHDF based on Becke88 exchange and P86 correlation, optimized for excitation energies
wPBEPP86	Casanova-Páez and Goerigk's range-separated DHDF based on PBE exchange and P86 correlation, optimized for excitation energies
wB97M(2)	Mardirossian and Head-Gordon's ωB97M(2) range-separated meta-GGA DHDF including VV10 non-local correlation: **must be used with ωB97M-V orbitals!** See *DFT Calculations with Second Order Perturbative Correction (Double-Hybrid Functionals)*.
wPr2SCAN50	Range-separated SOS-double-hybrid variant of r^2 SCAN with 50-100% HF exchange

Global and range-separated double-hybrid functionals with spin-component and spin-opposite scaling

Add the prefix RI- or DLPNO- to use the respective approximation for the MP2 part.

wB97X-2	Chai and Head-Gordon's ωB97X-2(TQZ) range-separated GGA-based DHDF with spin-component scaling
SCS/SOS-B2PLYP21	spin-opposite scaled version of B2PLYP optimized for excited states by Casanova-Páez and Goerigk (SCS fit gave SOS version; SOS only applies to the CIS(D) component)
SCS-PBE-QIDH	spin-component scaled version of PBE-QIDH optimized for excited states by Casanova-Páez and Goerigk (SCS only applies to the CIS(D) component)
SOS-PBE-QIDH	spin-opposite scaled version of PBE-QIDH optimized for excited states by Casanova-Páez and Goerigk (SOS only applies to the CIS(D) component)
SCS-B2GP-PLYP21	spin-component scaled version of B2GP-PLYP optimized for excited states by Casanova-Páez and Goerigk (SCS only applies to the CIS(D) component)
SOS-B2GP-PLYP21	spin-opposite scaled version of B2GP-PLYP optimized for excited states by Casanova-Páez and Goerigk (SOS only applies to the CIS(D) component)
SCS/SOS-wB2PLYP	spin-opposite scaled version of ωB2PLYP optimized for excited states by Casanova-Páez and Goerigk (SCS fit gave SOS version; SOS only applies to the CIS(D) component)
SCS-wB2GP-PLYP	spin-component scaled version of ωB2GP-PLYP optimized for excited states by Casanova-Páez and Goerigk (SCS only applies to the CIS(D) component)
SOS-wB2GP-PLYP	spin-opposite scaled version of ωB2GP-PLYP optimized for excited states by Casanova-Páez and Goerigk (SOS only applies to the CIS(D) component)
SCS-RSX-QIDH	spin-component scaled version of RSX-QIDH optimized for excited states by Casanova-Páez and Goerigk (SCS only applies to the CIS(D) component)
SOS-RSX-QIDH	spin-opposite scaled version of RSX-QIDH optimized for excited states by Casanova-Páez and Goerigk (SOS only applies to the CIS(D) component)
SCS-wB88PP86	spin-component scaled version of ωB88PPBE86 optimized for excited states by Casanova-Páez and Goerigk (SCS only applies to the CIS(D) component)
SOS-wB88PP86	spin-opposite scaled version of ωB88PPBE86 optimized for excited states by Casanova-Páez and Goerigk (SOS only applies to the CIS(D) component)
SCS-wPBEPP86	spin-component scaled version of ωPBEPPBE86 optimized for excited states by Casanova-Páez and Goerigk (SCS only applies to the CIS(D) component)
SOS-wPBEPP86	spin-opposite scaled version of ωPBEPPBE86 optimized for excited states by Casanova-Páez and Goerigk (SOS only applies to the CIS(D) component)

Composite Methods

HF-3c	HF-based composite method by Grimme et al. emplyoing the MINIX basis set
B97-3c	GGA composite method by Grimme et al. employing a modified def2-mTZVP basis set
R2SCAN-3c	meta-GGA composite method by Grimme et al. employing a modified def2-mTZVPP basis set
PBEh-3c	Hybrid (42% HF exchange) composite method by Grimme et al. employing a modified def2-mSVP basis set
wB97X-3c	Range-separated hybrid composite DFT method by Grimme et al. employing a polarized valence double-ζ basis set

Dispersion corrections

See *DFT Calculations with Atom-pairwise Dispersion Correction* and *Treatment of Dispersion Interactions with DFT-D3* for details.

D4	density dependent atom-pairwise dispersion correction with Becke-Johnson damping and ATM
D3BJ	Atom-pairwise dispersion correction to the DFT energy with Becke-Johnson damping
D3ZERO	Atom-pairwise dispersion correction with zero damping
D2	Empirical dispersion correction from 2006 (not recommended)

Non-local correlation

See *DFT Calculations with the Non-Local, Density Dependent Dispersion Correction (VV10): DFT-NL* for details.

NL	Does a post-SCF correction on the energy only
SCNL	Fully self-consistent approach, adding the VV10 correlation to the KS Hamiltonian

4.3 Basis Sets

4.3.1 Standard basis set library

There are standard basis sets that can be specified via the "simple input" feature in the keyword line. However, any basis set that is not already included in the ORCA library can be provided either directly in the input or through an external file. See the BASIS input block for a full list of internal basis sets and various advanced aspects (section *Choice of Basis Set*). Effective core potentials and their use are described in section *Effective Core Potentials*.

Pople-style basis sets

STO-3G	Minimal basis set(H–I)
3-21G	Pople 3-21G (H–Cs)
3-21GSP	Buenker 3-21GSP (H–Ar)
4-22GSP	Buenker 4-22GSP (H–Ar)
6-31G	Pople 6-31G and its modifications (H–Zn)
m6-31G	Modified 6-31G for 3d transition metals (Sc–Cu)
6-311G	Pople 6-311G and its modifications (H–Br)

Polarization functions for the 6-31G basis set:

* or (d)	One set of first polarization functions on all atoms except H
** or (d,p)	One set of first polarization functions on all atoms
Further combinations:	(2d), (2df), (2d,p), (2d,2p), (2df,2p), (2df,2pd)

Polarization functions for the 6-311G basis set: All of the above plus (3df) and (3df,3pd)

Diffuse functions for the 6-31G and 6-311G basis sets:

+ "G"	before	Include diffuse functions on all atoms except H (e.g. 6-31+G)
++ "G"	before	Include diffuse functions on all atoms. Works only when H polarization is already included, e.g. 6-31++G(d,p)

The def2 basis sets of the Karlsruhe group

These basis sets are all-electron for elements **H-Kr**, and automatically load Stuttgart-Dresden effective core potentials for elements **Rb-Rn**.

def2-SVP	Valence double-zeta basis set with "new" polarization functions.
def2-SV(P)	The above with slightly reduced polarization.
def2-TZVP	Valence triple-zeta basis set with "new" polarization functions. Note that this is quite similar to the older ("def") TZVPP for the main group elements and TZVP for hydrogen.
def2-TZVP(-f)	TZVP with f polarization removed from main group elements.
def2-TZVPP	TZVPP basis set with "new" polarization functions.
def2-QZVP	Polarized quadruple-zeta basis.
def2-QZVPP	Accurate doubly polarized quadruple-zeta basis.

Older ("def") Ahlrichs basis sets

ECP basis sets for elements **Fr-Lr**. This basis set automatically employs the original def-ECP.

def-TZVP	Valence triple-zeta basis set with polarization functions.
ma-def-TZVP	Minimally augmented def-TZVP variant with diffuse s and p functions according to Truhlar[917].

*All-electron basis sets for elements **H-Kr***:

SV	Valence double-zeta basis set.
SV(P)	Valence double-zeta with polarization only on heavy elements.
SVP	Polarized valence double-zeta basis set.
TZV	Valence triple-zeta basis set.
TZV(P)	Valence triple-zeta with polarization on heavy elements.
TZVP	Polarized valence triple-zeta basis set.
TZVPP	Doubly polarized triple-zeta basis set.
QZVP	Polarized valence quadruple-zeta basis set.
QZVPP	Doubly polarized quadruple-zeta basis set.

Note

Past versions of ORCA used to load all-electron basis sets also for elements Rb-I with the above keywords for double- and triple-zeta basis sets. The Rb-I basis sets originated from non-relativistic all-electron basis sets of the Turbomole library (such as "TZVPAlls"). **This automatic substitution is now deprecated**. However, we offer temporarily the ability to reproduce that behavior by adding the prefix "old-" to the above keywords, e.g. `old-TZVP`.

Diffuse def2 basis sets

Minimally augmented def2 basis sets:	Augmented def2 basis sets by diffuse s and p functions according to Truhlar[917]. Recommended for general use.
ma-def2-SVP	Minimally augmented def2-SVP basis set.
ma-def2-SV(P)	Minimally augmented def2-SV(P) basis set.
ma-def2-TZVP	Minimally augmented def2-TZVP basis set.
ma-def2-TZVP(-f)	Minimally augmented def2-TZVP(-f) basis set.
ma-def2-TZVPP	Minimally augmented def2-TZVPP basis set.
ma-def2-QZVPP	Minimally augmented def2-QZVPP basis set.
Rappoport property-optimized diffuse def2 basis sets:	Augmented def2 basis sets by diffuse functions according to Rappoport et al.[709, 710]
def2-SVPD	Diffuse def2-SVP basis set for property calculations
def2-TZVPD	Diffuse def2-TZVP basis set for property calculations
def2-TZVPPD	Diffuse def2-TZVPP basis set for property calculations
def2-QZVPD	Diffuse def2-QZVP basis set for property calculations
def2-QZVPPD	Diffuse def2-QZVPP basis set for property calculations

Karlsruhe basis sets with Dirac–Fock ECPs

These basis sets are derived from the def2-XVP ones with small modifications for 5s, 6s, 4d, and 5d elements and iodine.[883] They are optimized for the revised Dirac-Fock ECPs (dhf-ECP) as opposed to the Wood–Boring ones (def2-ECP). Versions for two-component methods are also available, e.g. `dhf-TZVP-2c`, **however, such methods are currently not implemented in ORCA**.

dhf-SV(P)	based on def2-SV(P)
dhf-SVP	based on def2-SVP
dhf-TZVP	based on def2-TZVP
dhf-TZVPP	based on def2-TZVPP
dhf-QZVP	based on def2-QZVP
dhf-QZVPP	based on def2-QZVPP

Relativistically recontracted Karlsruhe basis sets

For use in DKH or ZORA calculations we provide adapted versions of the def2 basis sets for the elements **H-Kr** (i.e., for the all-electron def2 basis sets). These basis sets retain the original def2 exponents but have only one contracted function per angular momentum (and hence are somewhat larger), with contraction coefficients suitable for the respective scalar relativistic Hamiltonian. These basis sets can be called with the prefix DKH- or ZORA-, and can be combined with the SARC basis sets for the heavier elements.

DKH-def2-SVP and ZORA-def2-SVP
DKH-def2-SV(P) and ZORA-def2-SV(P)
DKH-def2-TZVP and ZORA-def2-TZVP
DKH-def2-TZVP(-f) and ZORA-def2-TZVP(-f)
DKH-def2-TZVPP and ZORA-def2-TZVPP
DKH-def2-QZVPP and ZORA-def2-QZVPP

Minimally augmented versions:

ma-DKH-def2-SVP and ma-ZORA-def2-SVP
ma-DKH-def2-SV(P) and ma-ZORA-def2-SV(P)
ma-DKH-def2-TZVP and ma-ZORA-def2-TZVP
ma-DKH-def2-TZVP(-f) and ma-ZORA-def2-TZVP(-f)
ma-DKH-def2-TZVPP and ma-ZORA-def2-TZVPP
ma-DKH-def2-QZVPP and ma-ZORA-def2-QZVPP

The same functionality is offered for the "def" basis sets, e.g. "ZORA-TZVP". In this case too, the relativistically recontracted versions refer to the elements **H-Kr**. To replicate the behavior of past ORCA versions for elements Rb-I, the prefix "old-" can be used with these keywords as in the non-relativistic case.

Warning

Previous verions of ORCA made extensive use of automatic basis set substitution and aliasing when the use of the DKH or ZORA Hamiltonians was detected. This is no longer the case! Relativistic versions of Karlsruhe basis sets now have to be requested explicitly with the appropriate prefix. SARC basis sets also have to be requested explicitly

All-electron Karlsruhe basis sets up to Rn for the exact two-component (X2C) Hamiltonian.[689] The "-s" variants, e.g. "def2-TZVPall-s", are augmented with additional tight functions for NMR shielding calculations.[275] The "-2c" variants, e.g. "def2-TZVPall-2c", are intended for two-component calculations including spin-orbit coupling (**Note that two-component calculations are not implemented in *ORCA***).

x2c-SV(P)all	2c version: x2c-SV(P)all-2c, NMR version: x2c-SV(P)all-s
x2c-SVPall	2c version: x2c-SVPall-2c, NMR version: x2c-SVPall-s
x2c-TZVPall	2c version: x2c-TZVPall-2c, NMR version: x2c-TZVPall-s
x2c-TZVPPall	2c version: x2c-TZVPPall-2c, NMR version: x2c-TZVPPall-s
x2c-QZVPall	2c version: x2c-QZVPall-2c, NMR version: x2c-QZVPall-s
x2c-QZVPPall	2c version: x2c-QZVPPall-2c, NMR version: x2c-QZVPPall-s

SARC basis sets

[62, 640, 641, 642, 643, 728]

Segmented all-electron relativistically contracted basis sets for use with the DKH2 and ZORA Hamiltonians. Available for elements beyond Krypton.

SARC-DKH-TZVP
SARC-DKH-TZVPP
SARC-ZORA-TZVP
SARC-ZORA-TZVPP

Note

SARC/J is the general-purpose Coulomb-fitting auxiliary for all SARC orbital basis sets.

SARC2 basis sets for the lanthanides

[52]

SARC basis sets of valence quadruple-zeta quality for lanthanides, with NEVPT2-optimized (3g2h) polarization functions. Suitable for accurate calculations using correlated wavefunction methods.

SARC2-DKH-QZVP
SARC2-ZORA-QZVP

Note

Can be called without the polarization functions using ...-QZV. Each basis set has a large dedicated /JK auxiliary basis set for simultaneous Coulomb and exchange fitting.

Jensen basis sets

pc-n	($n = 0, 1, 2, 3, 4$) "Polarization-consistent" generally contracted basis sets (H–Kr) of up to quintuple-zeta quality, optimized for SCF calculations
aug-pc-n	As above, augmented by diffuse functions
pcseg-n	Segmented PC basis sets (H–Kr), DFT-optimized
aug-pcseg-n	As above, augmented by diffuse functions
pcSseg-n	Segmented contracted basis sets (H–Kr) optimized for nuclear magnetic shielding
aug-pcSseg-n	As above, augmented by diffuse functions
pcJ-n	Segmented contracted basis sets (H–Ar) optimized for spin-spin coupling constants
aug-pcJ-n	As above, augmented by diffuse functions

Lehtolas hydrogenic Gaussian basis sets

[505]

HGBS-m	($m = 5, 7, 9$) Lehtolas hydrogenic Gaussian basis sets optimized to the energy threshold m (H-Og)
HGBSPn-m	($n = 1, 2, 3$; $m = 5, 7, 9$) Variants with n polarization shells

Augmented versions:

AHGBS-m	($m = 5, 7, 9$) Variants augmented by diffuse functions
AHGBSPn-m	($n = 1, 2, 3$; $m = 5, 7, 9$)

Sapporo basis sets

Sapporo-nZP-2012	(n = D, T, Q) All-electron generally contracted non-relativistic basis sets (H–Xe)
Sapporo-DKH3-nZP-2012	(n = D, T, Q) All-electron basis sets optimized for the DKH3 Hamiltonian and finite nucleus (K–Rn)

Correlation-consistent basis sets

cc-pVDZ	Dunning correlation-consistent polarized double-zeta
cc-pVTZ	Dunning correlation-consistent polarized triple-zeta
cc-pVQZ	Dunning correlation-consistent polarized quadruple-zeta
cc-pV5Z	Dunning correlation-consistent polarized quintuple-zeta
cc-pV6Z	Dunning correlation-consistent polarized sextuple-zeta
aug-cc-pVnZ	(n = D, T, Q, 5, 6) Augmented with diffuse functions
cc-pCVnZ	(n = D, T, Q, 5, 6) Core-polarized basis sets
aug-cc-pCVnZ	(n = D, T, Q, 5, 6) as above, augmented with diffuse functions
cc-pwCVnZ	(n = D, T, Q, 5) Core-polarized with weighted core functions
aug-cc-pwCVnZ	(n = D, T, Q, 5) as above, augmented with diffuse functions
cc-pVn(+d)Z	(n = D, T, Q, 5) with tight d functions

Partially augmented correlation-consistent basis sets

[645]

apr-cc-pV(Q+d)Z	Augmented with sp diffuse functions on Li–Ca
may-cc-pV(n+d)Z	(n = T, Q): sp (T), spd (Q) on Li–Ca
jun-cc-pV(n+d)Z	(n = D, T, Q): sp (D), spd (T), spdf (Q) on Li–Ca
jul-cc-pV(n+d)Z	(n = D, T, Q): spd (D), spdf (T), spdfg (Q) on Li–Ca
maug-cc-pV(n+d)Z	same as jun-, may-, and apr- for n = D, T, and Q, respectively

DKH versions of correlation-consistent basis sets

cc-pVnZ-DK	(n = D, T, Q, 5) Correlation-consistent all-electron basis sets for use with the 2nd-order Douglas-Kroll-Hess Hamiltonian
aug-cc-pVnZ-DK	(n = D, T, Q, 5) as above, augmented with diffuse functions
cc-pwCVnZ-DK	(n = D, T, Q, 5) DK versions of weighted core correlation-consistent basis sets
aug-cc-pwCVnZ-DK	(n = D, T, Q, 5) weighted-core DK basis sets with diffuse functions
cc-pVnZ-DK3	(n = D, T, Q) Correlation-consistent all-electron basis sets for lanthanides and actinides with the 3rd-order Douglas-Kroll-Hess Hamiltonian
cc-pwCVnZ-DK3	(n = D, T, Q) DK versions of weighted core correlation-consistent basis sets for lanthanides and actinides

ECP-based versions of correlation-consistent basis sets

cc-pVnZ-PP	(n = D, T, Q, 5) Correlation-consistent basis sets combined with SK-MCDHF-RSC effective core potentials
aug-cc-pVnZ-PP	(n = D, T, Q, 5) as above, augmented with diffuse functions
cc-pwCVnZ-PP	(n = D, T, Q, 5) with weighted core functions
aug-cc-pwCVnZ-PP	(n = D, T, Q, 5) as above, augmented with diffuse functions

F12 and F12-CABS basis sets

cc-pVnZ-F12	(n = D, T, Q) Special orbital basis sets for F12 calculations (larger than the regular D, T, Q-zeta basis sets!)
cc-pCVnZ-F12	(n = D, T, Q) with core polarization functions
cc-pVnZ-PP-F12	(n = D, T, Q) ECP-based versions
cc-pVnZ-F12-CABS	(n = D, T, Q) Near-complete auxiliary basis sets for F12 calculations
cc-pVnZ-F12-OptRI	(n = D, T, Q) identical to the cc-pVnZ-F12-CABS basis above
cc-pCVnZ-F12-OptRI	(n = D, T, Q)
cc-pVnZ-PP-F12-OptRI	(n = D, T, Q)
aug-cc-pVnZ-PP-OptRI	(n = D, T, Q, 5)
aug-cc-pwCVnZ-PP-OptRI	(n = D, T, Q, 5)

Atomic Natural Orbital basis sets

ANO-pVnZ	(n = D, T, Q, 5, 6). Our newly contracted ANO basis sets on the basis of the cc-pV6Z (or pc-4 where missing) primitives. These are very accurate basis sets that are significantly better than the cc-pVnZ counterparts for the same number of basis functions (but much larger number of primitives of course).
saug-ANO-pVnZ	(n = D, T, Q, 5) augmentation with a single set of sp functions. Greatly enhances the accuracy of the SCF energies but not for correlation energies.
aug-ANO-pVnZ	(n = D, T, Q, 5) full augmentation with spd, spdf, spdfg set of polarization functions. Almost as expensive as the next higher basis set. In fact, aug-ANO-pVnZ = ANO-pV($n+1$)Z with the highest angular momentum polarization function deleted.

Relativistic contracted ANO-RCC basis sets:

ANO-RCC-FULL	The complete ANO-RCC basis sets (H-Cm). Some default contractions are provided for convenience with the keywords:
ANO-RCC-DZP	
ANO-RCC-TZP	
ANO-RCC-QZP	

Miscellaneous and specialized basis sets

D95	Dunning's double-zeta basis set (H–Cl).
D95p	Polarized version of D95.
MINI	Huzinaga's minimal basis set.
MINIS	Scaled version of the MINI.
MIDI	Huzinaga's valence double-zeta basis set.
MINIX	Combination of small basis sets by Grimme (see Table Table 7.8).
vDZP	Molecule-optimized polarized valence double-ζ basis set by Grimme et al. designed for ωB97X-3c[540].
Wachters+f	First-row transition metal basis set (Sc–Cu).
Partridge-n	(n = 1, 2, 3, 4) Uncontracted basis sets by Partridge.
LANL2DZ	Los Alamos valence double-zeta with Hay–Wadt ECPs.
LANL2TZ	Triple-zeta version.
LANL2TZ(f)	Triple-zeta plus polarization.
LANL08	Uncontracted basis set.
LANL08(f)	Uncontracted basis set + polarization.
EPR-II	Barone's basis set (H, B–F) for EPR calculations (double-zeta).
EPR-III	Barone's basis set for EPR calculations (triple-zeta).
IGLO-II	Kutzelnigg's basis set (H, B-F, Al–Cl) for NMR and EPR calculations.
IGLO-III	Larger version of the above.
aug-cc-pVTZ-J	Sauer's basis set for accurate hyperfine coupling constants.

4.3.2 Auxiliary basis sets

Auxiliary basis sets for the RI-J and RI-MP2 approximations can also be specified directly in the simple input:

Auxiliary basis sets for Coulomb fitting

Def/J	Weigend's "universal" Coulomb fitting basis that is suitable for all def type basis sets. Assumes the use of ECPs beyond Kr (do not use with DKH/ZORA).
Def2/J	Weigend's "universal" Coulomb fitting basis that is suitable for all def2 type basis sets. Assumes the use of ECPs beyond Kr (do not use with DKH/ZORA).
SARC/J	General-purpose Coulomb fitting basis set for all-electron calculations. Consists of the decontracted def2/J up to Kr and of our own auxiliary basis sets for the rest of the periodic table. Appropriate for use in DKH or ZORA calculations with the recontracted versions of the all-electron def2 basis sets (up to Kr) and the SARC basis sets for the heavier elements.
x2c/J	Weigend's Coulomb fitting basis for the all-electron x2c-XVPall basis sets.

Auxiliary basis sets for simultaneously fitting Coulomb and exchange

Fitting basis sets developed by Weigend for fitting simultaneously Coulomb and exchange energies. They are quite large and accurate. They fit SCF energies very well but even if they are large they do not fit correlation as well as the dedicated "/C" auxiliary basis sets.

Def2/JK	Coulomb+Exchange fitting for all def2 basis sets
Def2/JKsmall	reduced version of the above
cc-pVnZ/JK	(n = T, Q, 5) for the respective cc-pVnZ orbital basis
aug-cc-pVnZ/JK	(n = T, Q, 5) for the respective aug-cc-pVnZ orbital basis

Auxiliary basis sets for correlation calculations

Def2-SVP/C	Correlation fitting for the def2-SVP orbital basis
Def2-TZVP/C	for the def2-TZVP orbital basis
Def2-TZVPP/C	for the def2-TZVPP orbital basis
Def2-QZVPP/C	for the def2-QZVPP orbital basis
Def2-SVPD/C	for the def2-SVPD orbital basis
Def2-TZVPD/C	for the def2-TZVPD orbital basis
Def2-TZVPPD/C	for the def2-TZVPPD orbital basis
Def2-QZVPPD/C	for the def2-QZVPPD orbital basis
cc-pVnZ/C	(n = D, T, Q, 5, 6) for the respective cc-pVnZ orbital basis
aug-cc-pVnZ/C	(n = D, T, Q, 5, 6) for the respective aug-cc-pVnZ orbital basis
cc-pwCVnZ/C	(n = D, T, Q, 5) for the respective cc-pwCVnZ orbital basis
aug-cc-pwCVnZ/C	(n = D, T, Q, 5) for the respective aug-cc-pwCVnZ orbital basis
cc-pVnZ-PP/C	(n = D, T, Q) for the respective cc-pVnZ-PP orbital basis
aug-cc-pVnZ-PP/C	(n = D, T, Q) for the respective aug-cc-pVnZ-PP orbital basis
cc-pwCVnZ-PP/C	(n = D, T, Q) for the respective cc-pwCVnZ-PP orbital basis
aug-cc-pwCVnZ-PP/C	(n = D, T, Q) for the respective aug-cc-pwCVnZ-PP orbital basis
cc-pVnZ-F12-MP2fit	(n = D, T, Q) for the respective cc-pVnZ-F12 orbital basis
cc-pCVnZ-F12-MP2fit	(n = D, T, Q) for the respective cc-pCVnZ-F12 orbital basis
cc-pVnZ-PP-F12-MP2fit	(n = D, T, Q) for the respective cc-pVnZ-PP-F12 orbital basis
AutoAux	Automatic construction of a general purpose auxiliary basis for simultaneously fitting Coulomb, exchange and correlation calculations. See section *Automatic generation of auxiliary basis sets* for details.

Note

ORCA versions before 4.0 allowed the use of multiple keywords to invoke the same def2 Coulomb or Coulomb+exchange fitting basis set of Weigend. To avoid confusion all these keywords are now deprecated and the auxiliary basis sets are simply called using "`def2/J`" and "`def2/JK`".

Note

Starting from version 4.1 ORCA internally stores up to five basis sets for each calculation: the obligatory orbital basis set; an `AuxJ` Coulomb-fitting basis for the RI-J, RIJDX/RIJONX, and RIJCOSX approximations; an `AuxJK` Coulomb- and exchange-fitting basis used for RIJK; an `AuxC` auxiliary basis for the RI approximation in dynamical electron correlation treatments (such as RI-MP2, RI-CCSD, and DLPNO methods); and a complementary auxiliary basis set (`CABS`) for F12 methods. "/J" basis sets given in the simple input are assigned to `AuxJ` and likewise for the other types. Non-standard assignments like `AuxJ="def2/JK"` are possible only through the `%basis` block input (see section *Built-in Basis Sets*).

4.3.3 Use of scalar relativistic basis sets

For DKH and ZORA calculations ORCA provides relativistically recontracted versions of the Karlsruhe basis sets for elements up to Kr. These can be requested by adding the prefix DKH- or ZORA- to the normal basis set name. Note that for other non-relativistic basis sets (for example Pople-style bases) no recontraction has been performed and consequently such calculations are inconsistent! The basis set and the scalar relativistic Hamiltonian are specified in the keyword line, for example:

```
! B3LYP ZORA ZORA-TZVP  ...
```

If an auxiliary basis set is required for these recontracted Karlsruhe basis sets, we recommend the use of the decontracted def2/J. This can be obtained simply by using the keyword "`! SARC/J`" (instead of the equivalent "`!`

def2/J DecontractAuxJ") and is the recommended option as it simultaneously covers the use of SARC basis sets for elements beyond Krypton.

```
! TPSS ZORA ZORA-def2-TZVP SARC/J  ...
```

For all-electron calculations with elements heavier than Krypton we offer the SARC (segmented all-electron relativistically contracted) basis sets [62, 640, 641, 642, 643, 728]. These were specifically developed for scalar relativistic calculations and are individually adapted to the DKH2 and ZORA Hamiltonians. In this case the Coulomb-fitting auxiliary basis set *must* be specified as SARC/J, or alternatively the `AutoAux` keyword (*Automatic generation of auxiliary basis sets*) can be employed to create auxiliary basis sets.

```
! PBE DKH SARC-DKH-TZVP SARC/J  ...
```

Specifically for wavefunction-based calculations of lanthanide systems we recommend the more heavily polarized SARC2 basis sets [52].

Other basis sets suitable for scalar relativistic calculations are various versions of the all-electron correlation-consistent basis sets that are optimized for the DKH2 Hamiltonian and can be called with the suffix "-DK". The relativistically contracted atomic natural orbital (ANO-RCC) basis sets of Roos and coworkers were also developed for the DKH2 Hamiltonian and have almost complete coverage of the periodic table (up to Cm).

For calculations with the X2C Hamiltonian, all-electron basis sets with the prefix "x2c-" (e.g. x2c-TZVPall) developed by Weigend and coworkers are available.[275, 689] The matching AuxJ basis set is "x2c/J" and AutoAux can be used as well.

4.3.4 Effective Core Potentials

Starting from version 2.8.0, ORCA features effective core potentials (ECPs). They are a good alternative to scalar relativistic all-electron calculations if heavy elements are involved. This pertains to geometry optimizations and energy calculations but may not be true for property calculations.

In order to reduce the computational effort, the usually highly contracted and chemically inert core basis functions can be eliminated by employing ECPs. ECP calculations comprise a "valence-only" basis and thus are subject to the frozen core approximation. Contributions due to the core orbitals are accounted for by an effective one-electron operator U^{core} which replaces the interactions between core and valence electrons and accounts for the indistinguishability of the electrons. Its radial parts $U_l(r)$ are generally expressed as a linear combination of Gaussian functions, while the angular dependence is included through angular momentum projectors $|S^l_m\rangle$.

$$U^{\text{core}} = U_L(r) + \sum_{l=0}^{L-1} \sum_{m=-l}^{l} \left|S^l_m\right\rangle \left[U_l(r) - U_L(r)\right] \left\langle S^l_m\right|$$

$$U_l = \sum_k d_{kl} r^{n_{kl}} \exp(-\alpha_{kl} r^2)$$

The maximum angular momentum L is generally defined as $l^{\text{atom}}_{\text{max}} + 1$. The parameters n_{kl}, α_{kl} and d_{kl} that are necessary to evaluate the ECP integrals have been published by various authors, among them the well-known Los Alamos (LANL) [367] and Stuttgart–Dresden (SD) [41, 94, 137, 138, 217, 218, 219, 220, 221, 222, 223, 261, 262, 279, 280, 281, 347, 348, 409, 432, 433, 437, 507, 508, 509, 523, 524, 560, 579, 580, 593, 594, 632, 670, 671, 672, 673, 674, 758, 773, 823, 824, 872, 879, 880, 900] parameter sets. Depending on the specific parametrization of the ECP, relativistic effects can be included in a semiempirical fashion in an otherwise nonrelativistic calculation. Introducing U^{core} into the electronic Hamiltonian yields two types of ECP integrals, the local (or type-1) integrals that arise because of the maximum angular momentum potential U_L and the semi-local (or type-2) integrals that result from the projected potential terms. The evaluation of these integrals in ORCA proceeds according to the scheme published by Flores-Moreno et al.[266].

A selection of ECP parameters and associated basis sets is directly accessible in ORCA through the internal ECP library (see Table 4.3 for a listing of keywords).

Table 4.3: Overview of library keywords for ECPs and associated basis sets available in ORCA.

ECP keyword	Core size[Page 47, 1]	Elements	Valence basis sets
		Recommended	
def-ECP	78	Fr–Ra	Karlsruhe basis sets: def-TZVP, ma-def-TZVP
	60	Ac–Lr	
def2-ECP	28	Rb–Xe	Karlsruhe basis sets: def2-SVP, def2-TZVP, etc. def2-SVPD, def2-TZVPD, etc. ma-def2-SVP, ma-def2-TZVP, etc.
	46	Cs–La	
	28	Ce–Lu	
	60	Hf–Rn	
SK-MCDHF-RSC	10	Ca, Cu–Kr	Correlation-consistent basis sets: cc-pVnZ-PP, aug-cc-pVnZ-PP, cc-pCVnZ-PP, aug-cc-pCVnZ-PP, cc-pwCVnZ-PP, aug-cc-pwCVnZ-PP (n = D, T, Q, 5) cc-pVnZ-PP (n = D, T, Q)
	28	Sr–Xe	
	46	Ba	
	60	Hf–Rn	
	78	Ra	
	60	U	
HayWadt[2]	10	Na–Cu	LANL-type basis sets: LANL2DZ, LANL2TZ, LANL2TZ(f), LANL08, LANL08(f)
	18	Zn	
	28	Ga–Ag	
	36	Cd	
	46	In–La	
	60	Hf–Au	
	68	Hg–Tl	
	78	Pb–Bi, U–Pu	
dhf-ECP	28	Rb–Xe	dhf-type Karlsruhe basis sets: dhf-SVP, dhf-TZVP, etc.
	46	Cs–Ba	
	60	Hf–Rn, U	
vDZP-ECP	2	B–Mg	vDZP basis set. *uniquely compiled for the use with vDZP*
	10	Al–Zn	
	28	Ga–Cd	
	46	In–Lu	
	60	Hf–Hg	
	78	Tl–Rn	
		Legacy definitions	
def2-SD	28,MWB	Rb–Cd	
	28,MDF[3]	In–Xe	
	46,MWB	Cs–La	
	60,MWB	Hf–Pt	
	60,MDF[4]	Au–Rn	
def-SD	28,MWB	Rb–Cd	
	46,MWB	In–La	
	28,MWB	Ce–Lu	
	60,MWB	Hf–Pt	
	60,MDF[Page 47, 4]	Au, Hg, Rn	
	78,MWB	Tl–At	
	78,MDF	Fr, Ra	
	60,MWB	Ac–Lr	
SDD	2,SDF	Li, Be	
	2,MWB	B–Ne	
	10,SDF	Na, Mg	
	10,MWB	Al–Ca	

continues on next page

Table 4.3 – continued from previous page

ECP keyword	Core size[1]	Elements	Valence basis sets
	10,MDF	Sc–Ni	
	10,MWB	Cu-Zn	
	28,MWB	Ga–Sr	
	28,MHF	Y–Cd	
	28,MDF	Ge–Br, Rb–Xe	
	46,MWB	In–Ba	
	28,MWB	La–Lu	
	60,MWB	Hf–Hg	
	78,MWB	Tl–Rn	
	60,MWB	Ac–Lr	
LANL1	10	Na–Ar	
	18	K–Zn	
	28	Ga–Kr	
	36	Rb–Cd	
	46	In–Xe	
	54	Cs–La	
	68	Hf–Tl	
	78	Pb, Bi	
LANL2	10	K–Cu	
	28	Rb–Ag	
	46	Cs–La	
	60	Hf–Au	

Note

Some basis sets assign an ECP by default when requested through the simple input (but not through the `%basis` block): for example, "def2" basis sets use the def2-ECP. For others, see the footnotes under Table 7.10.

The simplest way to assign ECPs is by using the ECP keyword within the keyword line, although input through the `%basis` block is also possible (*Advanced Specification of Effective Core Potentials*). The ECP keyword itself assigns *only* the effective core potential, not a valence basis set! As an example for an explicitly named ECP you could use

```
! def2-TZVP def2-SD
```

This would assign the def2-SD ECP according to the definition given in the table above. Without the def2-SD keyword ORCA would default to def2-ECP.

[1] Where applicable, reference method and data are given (S: single-valence-electron ion; M: neutral atom; HF: Hartree-Fock; WB: quasi-relativistic; DF: relativistic).
[2] Corresponds to LANL2 and to LANL1 where LANL2 is unavailable.
[3] I: OLD-SD(28,MDF) for compatibility with TURBOMOLE.
[4] Au, Hg: OLD-SD(60,MDF) for compatibility with TURBOMOLE.

4.4 Numerical Integration in ORCA

Starting from its version 5.0, ORCA has a new scheme for the quadratures used in numerical integration. It is based on the same general ideas which were used for the old grids, except that we used machine learning methods, together with some final hands-on optimization, to find the optimal parameters for all atoms up to the 6th row of the periodic table, with the 7th row being extrapolated from that. For further details look at Ref. [383]. We also realized that the COSX and DFT grids have overall different requirements, and these were optimized separately.

The big advantage of this new scheme is that it is significantly more accurate and robust than the old one, even if having the same number of grid points. We tested energies, geometries, frequencies, excitation energies and properties to develop three new grid schemes named: DEFGRID1, DEFGRID2 and DEFGRID3, that will automatically fix all grids that are used in the calculations. DEFGRID1 behaves essentially like the old defaults, but it is more robust. The second is the new default, and is expected to yield sufficiently small errors for all kinds of applications (see Section *Details on the numerical integration grids* for details). The last is a heavier, higher-quality grid, that is close to the limit if one considers an enormous grid as a reference.

In order to change from the default DEFGRID2, one just needs to add !DEFGRID1 or !DEFGRID3 to the main input.

It is also important to note that the COSX approximation is now the default for DFT, whenever HFexchange is neede. This can always be turned off by using !NOCOSX.

4.5 Input priority and processing order

In more complicated calculations, the input can get quite involved. Therefore it is worth knowing how it is internally processed by the program:

- First, all the simple input lines (starting with "!") are collected into a single string.
- The program looks for all known keywords in a predefined order, regardless of the order in the input file.
- An exception are basis sets: if two different orbital basis sets (e.g. `! def2-SVP def2-TZVP`) are given, the latter takes priority. The same applies to auxiliary basis sets of the same type (e.g. `! def2/J SARC/J`).
- Some simple input keywords set multiple internal variables. Therefore, it is possible for one keyword to overwrite an option, set by another keyword. We have tried to resolve most such cases in a reasonable way (e.g. the more "specific" keyword should take precedence over a more "general" one) but it is difficult to forsee every combination of options.
- Next, the block input is parsed in the order it is given in the input file.
- Most block input keywords control a single variable (although there are exceptions). If a keyword is duplicated, the latter value is used.

Consider the following (bad) example:

```
! def2-TZVP UKS
%method
  functional BP86
  correlation C_LYP
  SpecialGridAtoms[1]  26, 27
  SpecialGridIntacc 8,  8,  8
  SpecialGridAtoms 28, 29
end
! PBE def2-SVP RKS
```

Using the rules above, one can figure out why it is equivalent to this one:

```
! UKS BLYP def2-SVP
%method
  SpecialGridAtoms 28, 29, 27
```

(continues on next page)

(continued from previous page)

```
  SpecialGridIntacc 8,  8,  8
end
```

4.6 ORCA and Symmetry

For most of its life, ORCA did not take advantage of molecular symmetry. Starting from version 2.8 (released in September 2010), there has been at least limited use. On request (using the simple keyword `UseSym` for instance, see below), the program detects the point group, orients the molecule, cleans up the coordinates and produces symmetry-adapted molecular orbitals.

Only for geometry cleanup the full point group is taken into account. For all other purposes such as the construction of symmetry-adapted molecular orbitals and or to describe electronic states, only D_{2h} and subgroups are currently supported. Here the use of symmetry helps to control the calculation and the interpretation of the results.

4.6.1 Getting started

Utilization of symmetry is turned on by the simple keyword `UseSymmetry` (which may be abbreviated by `UseSym`), or if a `%Symmetry` (or `%Sym`) input block is present in the input. ORCA will then automatically determine the point group, reorient and center the molecule to align its symmetry elements with the coordinate system, and replace the input structure by a geometry that corresponds exactly to this point group and which minimizes the sum of square distances between the atoms of both structures.

Any program that attempts to find the point group of an arbitrary atom cluster must be prepared to cope with some amount of numerical noise in the atom coordinates. ORCA by default allows each atom to deviate at most 10^{-4} atomic units from the ideal position that is consistent with the point group being examined. The rationale behind this value is the rounding error that occurs when the user feeds Cartesian coordinates with five significant digits after the decimal point into the program which otherwise represent an exact (symmetry-adapted) geometry. A threshold that is about one order of magnitude higher than the numerical noise in the coordinates is usually very safe.

If the maximum error in the Cartesian coordinates exceeds these 10^{-4} atomic units, the symmetry module in ORCA will fail to recognize the expected point group. The user is strongly advised to always make sure that the detected point group meets their expectations. If the point group reported by the symmetry module appears to be too low, the user may try to increase the detection threshold to 10^{-3} or 10^{-2} Bohr radii using option `SymThresh` in the `%Symmetry` input block:

```
%Sym SymThresh 0.01 End
```

A great method to obtain a structure with perfect symmetry avoiding any expensive calculation is to use the simple keywords `! NoIter XYZFile` with an appropriate threshold. The structure in the resulting file with the extension `.xyz` may then be used as input for the actual calculation.

To give an illustrative example, coordinates for staggered ethane have been obtained by geometry optimization *without* using symmetry. If symmetry is turned on, point group C_i is recognized instead of the expected point group D_{3d} due to the remaining numerical noise. To counter this, the detection threshold is increased to 10^{-2} a. u. and a coordinate file with perfect symmetry is produced by the following input:

```
! RHF SVP NoIter XYZfile
%sym SymThresh 1.0e-2 end
*xyz 0 1
  C    -0.002822   -0.005082   -0.001782
  C    -0.723141   -1.252323   -0.511551
  H     0.017157    0.029421    1.100049
  H     1.042121    0.030085   -0.350586
  H    -0.495109    0.917401   -0.350838
  H    -0.743120   -1.286826   -1.613382
  H    -0.230855   -2.174806   -0.162495
```

(continues on next page)

(continued from previous page)

```
  H    -1.768085   -1.287489   -0.162747
*
```

If ORCA fails to find the expected point group even though a value of 10^{-2} atomic units has been selected for SymThresh, the user is strongly advised to take a careful look at the structure by means of their favorite visualization tool before increasing this value any further. Look for any obvious distortions or even missing atoms. An especially tricky point may be the orientation of methyl groups or the conformation of floppy side chains. A small rotation about a single bond may be enough to push some atom positions above the limit. If the conformational deviations cannot be fixed using a molecular editor or modelling program, a possible alternative may be to pre-optimize the structure without symmetry using a less expensive method like PB86 and a small basis set like def2-SVP. Even several passes of pre-optimization and structure editing may be considered until all symmetry-equivalent side chains are locked in the same conformation so that ORCA finally detects the correct point group.

It is not recommended to run calculations using a value of SymThresh which is much too high or much too small since this may result in some really strange behavior of the symmetry module. Consider for instance the following input file which contains a perfectly octahedral geometry of a sulfur hexafluoride molecule. Its coordinates may be easily created by hand by placing the sulfur atom into the origin and two fluorine atoms on each coordinate axis at equal distances r from the origin ($r = 1.56$ Å or approximately 2.95 atomic units). Using a value for SymThresh as large as 0.1 Bohr radii works fine in this case, resulting in the correct point group O_h.

```
# Sulfur hexafluoride (SF6), point group Oh.
! BP86 def2-SVP
%Sym SymThresh 0.1 End
* xyz 0 1
S   0.00   0.00   0.00
F   1.56   0.00   0.00
F  -1.56   0.00   0.00
F   0.00   1.56   0.00
F   0.00  -1.56   0.00
F   0.00   0.00   1.56
F   0.00   0.00  -1.56
*
```

However, if SymThresh is increased further to $t = 0.5$ atomic units, the point group detection algorithm breaks down (strange warnings are printed as a consequence) and the reported point group decreases to C_i (in which the center of inversion is the only non-trivial symmetry element). This is because the center of inversion is easy to detect and this is done by one of the early checks. The breakdown of the point group recognition may be explained as follows. During the process of point group detection the symmetry module is of course unaware that the given input geometry is exact. Hence it will be treated as any other input structure. A value of $t = 0.5$ Bohr radii for SymThresh means that the unknown exact atom position is located within a sphere of radius $t = 0.5$ atomic units around the input atom position. The input distance $a = \sqrt{2}\,r$ between two adjacent fluorine atoms is approximately $a \approx 2.21$ Å ≈ 4.17 a. u., so their unknown exact distance d may vary in the following interval (see the diagram in Fig. 4.1):

$$d_{\min} = a - 2t = 3.17 \text{ a.u.} \leq d \leq d_{\max} = a + 2t = 5.17 \text{ a.u.}$$

Analogously, the unknown exact distance d' between two opposite fluorine atoms with the input distance $a' = 2r = 5.90$ a. u. is:

$$d'_{\min} = a' - 2t = 4.90 \text{ a.u.} \leq d' \leq d'_{\max} = a' + 2t = 6.90 \text{ a.u.}$$

Since the possible intervals of d and d' overlap (due to $d_{\max} > d'_{\min}$), *all* fifteen F–F distances are considered equal. Since there is no solid with six vertices and fifteen equal inter-vertex distances in three dimensions, the point group detection algorithm fails.

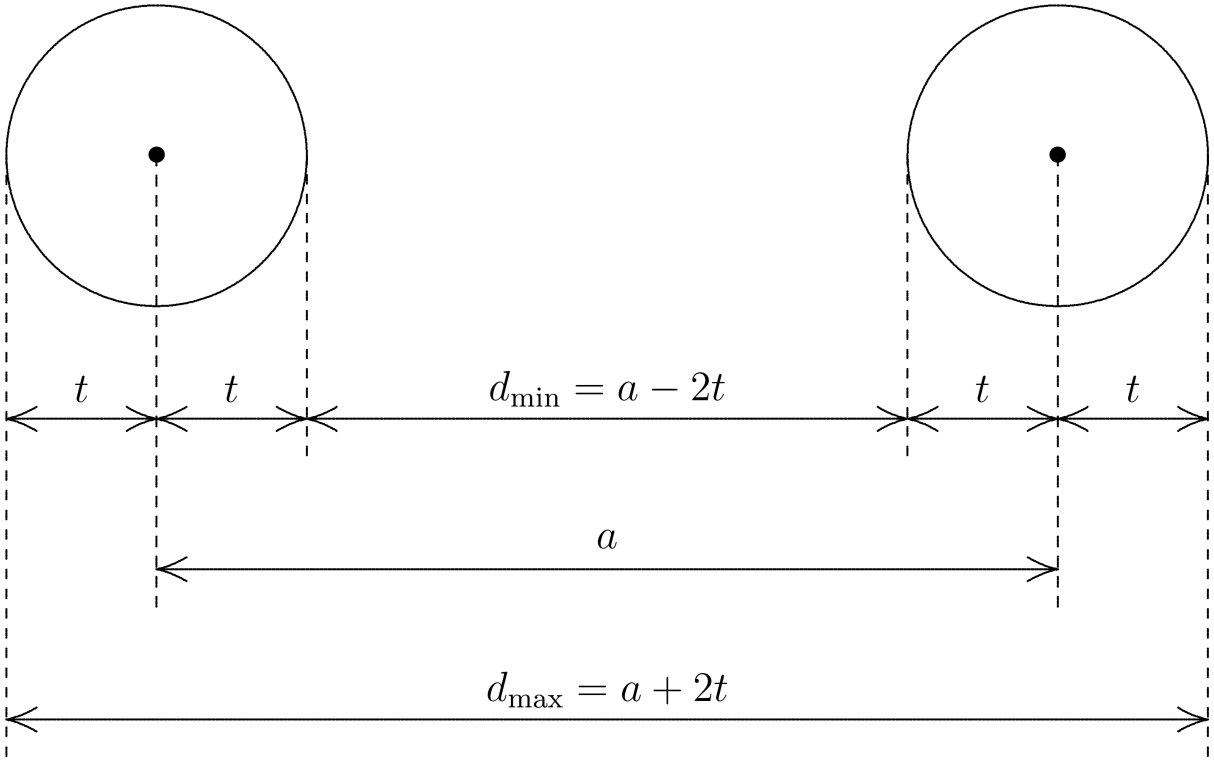


Fig. 4.1: The relation between the value t of `SymThresh`, the distance a of some input atom pair, and the allowed interval $[d_{\min}, d_{\max}]$ for the distance d between the exact atom positions. This interval has the width $d_{\max} - d_{\min} = 4t$.

4.6.2 Geometry optimizations using symmetry

If a geometry optimization is performed with symmetry turned on, ORCA will first determine the point group of the starting structure and replace the geometry that is presumed to contain numerical noise with one that has perfect symmetry. Starting with ORCA 6, the optimizer will clean up the gradient at every step of the optimization if requested by setting option `CleanUpGradient true` in the `%Symmetry` input block. The gradient cleanup is done by projecting out all components that are not totally symmetric. This way the symmetry of the molecule cannot decrease during the optimization.

By default, the point group is determined from scratch again after the geometry has been updated at every step of the optimization. This behaviour may be switched off by setting option `SymRelaxOpt false` in the `%Symmetry` input block. In this case the point group of the molecule is actually frozen during the entire optimization.

The following table summarizes the behaviour of the optimizer depending on the options `SymRelaxOpt` and `CleanUpGradient`:

SymRelaxOpt	CleanUpGradient	Behaviour
true	true	Symmetry may increase but not decrease.
true	false	Symmetry may change freely.
false	true	Symmetry will be frozen.
false	false	Setting not recommended.

Setting both switches `false` would allow the point group to change during the optimization but at the same time, a change would be impossible to detect. Therefore this setting is strongly discouraged.

4.6.3 Default alignment of the symmetry elements with the coordinate system

If ORCA determines the point group of a molecule and the user has not selected any special options, the following principles apply to the manner in which the symmetry elements of the full point group are aligned with the coordinate system:

1. The center of mass of the molecule will be shifted into the origin by default.[5] If the point group leaves one *unique* vertex invariant to all symmetry operations, the center of mass agrees with this vertex. This is the case for all point groups except C_s, C_n $(n \geq 1)$, C_{nv} $(n \geq 2)$, and $C_{\infty v}$.
2. If the molecule exhibits a unique axis of symmetry with the highest number of positions, this axis will become the z axis. This applies to all point groups except C_1, C_i, C_s, D_2, D_{2h}, the cubic point groups, and K_h.
3. For point group C_s, the mirror plane will become the xy plane.
4. For point groups C_{nv} $(n \geq 2)$, one of the vertical mirror planes will become the xz plane.
5. For point groups D_n $(n \geq 3)$, D_{nh} $(n \geq 3)$, and D_{nd} $(n \geq 2)$, one of the two-fold rotation axes perpendicular to the axis with the highest number of positions will become the x axis.
6. For point groups D_2, D_{2h}, T, and T_h, the three mutually orthogonal C_2 axes will become the coordinate axes.
7. For point groups T_d, O, and O_h, the three mutually orthogonal four-fold rotation or rotation-reflection axes will become the coordinate axes.
8. Finally, for point groups I and I_h, one of the five sets of three mutually orthogonal C_2 axes will become the coordinate axes. The pair of C_5 or S_{10} axes closest to the z axis will be located in the yz plane.
9. In general the orientation of the molecule will be changed as little as possible to meet the criteria above. If the input geometry meets these criteria already, the molecule will not be moved or rotated at all.

If the point group of the system is D_{nd} with $n \geq 2$ or T_d and the user has selected subgroup C_{2v} using option `PreferC2v`, the following rules apply instead:

- For point group D_{nd} with $n \geq 2$, one of the diagonal mirror planes will become the xz plane.
- For point group T_d, one of the diagonal mirror planes containing the z axis will become the xz plane, i. e. the molecule will be rotated by 45 degrees about the z axis compared to the default orientation.

Table 4.4 gives an overview over all point groups and the way in which the symmetry elements of the reduced point group (the largest common subgroup of D_{2h}) are aligned with the coordinate system.

Table 4.4: Point groups and corresponding subgroups suitable for electronic-structure calculations.

Full point group	Index n	Unique center[Page 53, 6]	Consistent with planar molecule[Page 53, 7]	Chosen subgroup	Alignment of the subgroup[Page 53, 8]
C_1		no	no	C_1	
C_i		i	no	C_i	
C_s		no	yes	C_s	
C_n	odd	no	no	C_1	
	even	no	no	C_2	z axis
C_{nv}	odd	no	no	C_s	xz plane
	even	no	for $n = 2$	C_{2v}	z, xz, yz
C_{nh}	odd	yes	yes	C_s	xy plane
	even	i	yes	C_{2h}	z, xy
D_n	odd	yes	no	C_2	x axis
	even	yes	no	D_2	
D_{nh}	odd	yes	yes	C_{2v}	x, xy, xz
	even	i	yes	D_{2h}	
D_{nd}	odd	i	no	C_{2h}	x, yz

continues on next page

[5] In the very special case that the Z matrix contains no atoms with mass, the geometric center will be used instead.

Table 4.4 – continued from previous page

Full point group	Index n	Unique center[6]	Consistent with planar molecule[7]	Chosen subgroup	Alignment of the subgroup[8]
	even	yes	no	D_2	
				C_{2v}	z, xz, yz
S_{2n}	odd	i	no	C_i	
	even	yes	no	C_2	z axis
T		yes	no	D_2	
T_h		i	no	D_{2h}	
T_d		yes	no	D_2	
				C_{2v}	z, xz, yz
O		yes	no	D_2	
O_h		i	no	D_{2h}	
I		yes	no	D_2	
I_h		i	no	D_{2h}	
$C_{\infty v}$		no	no	C_{2v}	z, xz, yz
$D_{\infty h}$		i	no	D_{2h}	
K_h		i	no	D_{2h}	

4.6.4 Irreducible representations of D_{2h} and subgroups

Table 4.5, Table 4.6, and Table 4.7 contain lists of the irreducible representations (also called species) and the corresponding characters of the point groups supported for electronic structure calculations in ORCA, and the product tables of these irreducible representations. Where the data depends on the alignment of the symmetry elements with the coordinate system, Mulliken's recommendations [599] are followed. This approach is in line with the recommendations by the IUPAC [770].

Table 4.5: Species and species product table of point group C_{2v}. The species table for C_{2v} corresponds to Table III in [599]. The directions of the two-fold axis and the mirror planes in each column are related to each other by cyclic permutations.

C_{2v}	E	$C_2(z)$ $C_2(x)$ $C_2(y)$	$\sigma_v(xz)$ $\sigma_v(xy)$ $\sigma_v(yz)$	$\sigma_v(yz)$ $\sigma_v(xz)$ $\sigma_v(xy)$
A_1	$+1$	$+1$	$+1$	$+1$
A_2	$+1$	$+1$	-1	-1
B_1	$+1$	-1	$+1$	-1
B_2	$+1$	-1	-1	$+1$

$\times$	A_1	A_2	B_1	B_2
A_1	A_1	A_2	B_1	B_2
A_2	A_2	A_1	B_2	B_1
B_1	B_1	B_2	A_1	A_2
B_2	B_2	B_1	A_2	A_1

Table 4.6: Species and species product table of point group D_2. The species table for D_2 has been obtained by dropping the center of inversion and the mirror planes from the species table for D_{2h} (see Table 4.7).

D_2	E	$C_2(z)$	$C_2(y)$	$C_2(x)$
A	$+1$	$+1$	$+1$	$+1$
B_1	$+1$	$+1$	-1	-1
B_2	$+1$	-1	$+1$	-1
B_3	$+1$	-1	-1	$+1$

$\times$	A	B_1	B_2	B_3
A	A	B_1	B_2	B_3
B_1	B_1	A	B_3	B_2
B_2	B_2	B_3	A	B_1
B_3	B_3	B_2	B_1	A

[6] A center of inversion is denoted i. "yes" indicates the existence of a *unique* vertex that remains invariant to all symmetry operations of the point group.

[7] This column indicates whether the given point group may be the *full* point group of a planar molecule.

[8] This column contains the elements (axes or planes) of the coordinate system that coincide with the symmetry elements of the reduced point group (the largest common subgroup of the full point group and D_{2h}) by *default*. If the full point group contains a unique principle axis of symmetry (with the highest number of positions), this axis is presumed to coincide with the z axis.

Table 4.7: Species and species product table of point group D_{2h}. The species table for D_{2h} corresponds to Table IV in [599].

D_{2h}	E	$C_2(z)$	$C_2(y)$	$C_2(x)$	i	$\sigma(xy)$	$\sigma(xz)$	$\sigma(yz)$
A_g	$+1$	$+1$	$+1$	$+1$	$+1$	$+1$	$+1$	$+1$
B_{1g}	$+1$	$+1$	-1	-1	$+1$	$+1$	-1	-1
B_{2g}	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1
B_{3g}	$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$
A_u	$+1$	$+1$	$+1$	$+1$	-1	-1	-1	-1
B_{1u}	$+1$	$+1$	-1	-1	-1	-1	$+1$	$+1$
B_{2u}	$+1$	-1	$+1$	-1	-1	$+1$	-1	$+1$
B_{3u}	$+1$	-1	-1	$+1$	-1	$+1$	$+1$	-1

$\times$	A_g	B_{1g}	B_{2g}	B_{3g}	A_u	B_{1u}	B_{2u}	B_{3u}
A_g	A_g	B_{1g}	B_{2g}	B_{3g}	A_u	B_{1u}	B_{2u}	B_{3u}
B_{1g}	B_{1g}	A_g	B_{3g}	B_{2g}	B_{1u}	A_u	B_{3u}	B_{2u}
B_{2g}	B_{2g}	B_{3g}	A_g	B_{1g}	B_{2u}	B_{3u}	A_u	B_{1u}
B_{3g}	B_{3g}	B_{2g}	B_{1g}	A_g	B_{3u}	B_{2u}	B_{1u}	A_u
A_u	A_u	B_{1u}	B_{2u}	B_{3u}	A_g	B_{1g}	B_{2g}	B_{3g}
B_{1u}	B_{1u}	A_u	B_{3u}	B_{2u}	B_{1g}	A_g	B_{3g}	B_{2g}
B_{2u}	B_{2u}	B_{3u}	A_u	B_{1u}	B_{2g}	B_{3g}	A_g	B_{1g}
B_{3u}	B_{3u}	B_{2u}	B_{1u}	A_u	B_{3g}	B_{2g}	B_{1g}	A_g

4.6.5 Options available in the `%Symmetry` input block

Table 4.8 contains a list of the options available in the `%Symmetry` (or `%Sym`) input block. Options `SymThresh` and `SymRelax` (same as `SymRelaxSCF` below) can also be accessed in the `%Method` input block for backward compatibility. This use is deprecated and not recommended in new input files, however.

Table 4.8: List of options in the `%Symmetry` (`%Sym`) input block

Option	Type	Default	Description
`UseSymmetry`	Boolean	`True`	By setting this option to `False`, symmetry may be switched off even though the `%Symmetry` block is present in the input file.
`UseSym`	Boolean	`True`	Same as `UseSymmetry`.
`SymThresh`	Real	10^{-4}	Two vertices with a distance shorter than this threshold (in atomic units) are considered identical during point group recognition.
`PreferC2v`	Boolean	`False`	Indicates whether to prefer subgroup C_{2v} over D_2 for electronic-structure calculations where both choices are appropriate (point groups D_{nd} with odd n and T_d).
`PointGroup`	String	Empty string	If the user specifies a point group using this option, point group recognition will be skipped and the user must make sure that the molecule is oriented in the coordinate system in agreement with the conventions in Section 4.6.3. Note that the point group label must be enclosed in double quotes. Otherwise ORCA will complain about an invalid assignment.
`SymRelaxSCF`	Boolean	`False`	Indicates whether orbital occupation numbers of each irreducible representation are allowed to change during SCF.

continues on next page

Table 4.8 – continued from previous page

Option	Type	Default	Description
SymRelaxOpt	Boolean	True	Indicates whether the point group will be determined from scratch in every step of a geometry optimization. A value of True will allow the point group to change in an arbitrary manner. Otherwise the initial point group will be imposed at every step no matter how far the distances between the current and the ideal structure exceed SymThresh.
CleanUpCoords	Boolean	True	Determines whether the molecular geometry will be cleaned up using the automatically detected or user-specified point group. Even if CleanUpCoords is False, symmetrized coordinates will still be computed temporarily and a warning will be printed if the largest deviation from the original geometry exceeds SymThresh.
CleanUpGeom	Boolean	True	Same as CleanUpCoords.
CleanUpGrad	Boolean	True	Indicates whether the full point group of the molecule shall be used to remove all non-totally symmetric components from the gradient. This ensures that the point group will not decrease throughout the optimization.
CleanUpGradient	Boolean	True	Same as CleanUpGrad.
Print	Integer	1	Determines the output size for symmetry handling in general and point group detection in particular; 0 – No output during point group detection; 1 – Normal output; 2 – Detailed information; 3 – Debug print.
PrtSALC	Integer	0	Specifies the output size for the construction of symmetry-adapted linear combinations (SALCs) of atomic orbitals; 0 – No output for symmetry-adapted orbitals; 1 – Normal output; 2 – Detailed information (e. g. the SALCs themselves); 3 – Debug print.

4.7 Jobs with Multiple Steps

ORCA supports input files with multiple jobs. This feature is designed to simplify series of closely related calculations on the same molecule or calculations on different molecules. The objectives for implementing this feature include:

- Calculate of a molecular property using different theoretical methods and/or basis sets for one molecule.
- Calculations on a series of molecules with identical settings.
- Geometry optimization followed by more accurate single points and perhaps property calculations.
- Crude calculations to provide good starting orbitals that may then be used for subsequent calculations with larger basis sets.

For example consider the following job that in the first step computes the g-tensor of BO at the LDA level, and in the second step using the BP86 functional.

```
# ------------------------------------------------------
! LSD DEF2-SVP TightSCF KeepInts
# ------------------------------------------------------
%eprnmr gtensor 1 end

* int 0 2
   B  0  0  0   0      0  0
   O  1  0  0   1.2049 0  0
*
```

(continues on next page)

(continued from previous page)

```
# ************************************************
# ****** This starts the input for the next job   *
# ************************************************
$new_job
# ------------------------------------------------
! BP86 DEF2-SVP SmallPrint ReadInts NoKeepInts
# ------------------------------------------------
%eprnmr gtensor 1 end

* int 0 2
   B  0  0  0   0      0  0
   O  1  0  0   1.2049 0  0
*
```

What happens if you use the `$new_job` feature is that all calculation flags for the actual job are transferred from the previous job and that only the changes in the settings must be input by the user. Thus if you turn on some flags for one calculation that you do not want for the next, you have to turn them off again yourself (for example the use of the RI approximation)! In addition, the default is that the new job takes the orbitals from the old job as input. If you do not want this you have to overwrite this default by specifying your desired guess explicitly.

4.7.1 Changing the default BaseName

Normally the output files for `MyJob.inp` are returned in `MyJob.xxx` (any xxx, for example xxx=out). Sometimes, and in particular in multistep jobs, you will want to change this behavior. To this end there is the variable `%base` that can be user controlled. All filenames (also scratch files) will then be based on this default name. For example, using the following setting, the output files for the current job would be `job1.xxx` (e.g. `job1.gbw`, `job1.densities`, etc.).

```
%base "job1"
```

CHAPTER

FIVE

INPUT OF COORDINATES

Coordinates can be either specified directly in the input file or read from an external file, and they can be in either Cartesian ("xyz") or internal coordinate format ("Z-matrix").

5.1 Reading coordinates from the input file

The easiest way to specify coordinates in the input file is by including a block like the following, enclosed by star symbols:

```
* CType Charge Multiplicity
...
coordinate specifications
...
*
```

Here `CType` can be one of `xyz`, `int` (or `internal`), or `gzmt`, which correspond to Cartesian coordinates, internal coordinates, and internal coordinates in Gaussian Z-matrix format.

The input of Cartesian coordinates in the "`xyz`" option is straightforward. Each line consists of the label for a given atom type and three numbers that specify the coordinates of the atom. The units can be either Ångström or Bohr. The default is to specify the coordinates in Ångströms (this can be changed through the keyword line or via the variable `Units` in the `%coords` main block described below).

```
* xyz Charge Multiplicity
Atom1   x1  y1  z1
Atom2   x2  y2  z2
 ...
*
```

For example for CO^+ in a $S = 1/2$ state (multiplicity = $2 \times 1/2 + 1 = 2$)

```
* xyz 1 2
C  0.0  0.0  0.0
O  0.0  0.0  1.1105
*
```

Internal coordinates are specified in the form of the familiar "Z-matrix". A Z-matrix basically contains information about molecular connectivity, bond lengths, bond angles and dihedral angles. The program then constructs Cartesian coordinates from this information. Both sets of coordinates are printed in the output such that conversion between formats is facilitated. The input in that case looks like:

```
* int Charge Multiplicity
Atom1  0  0  0    0.0   0.0   0.0
Atom2  1  0  0    R1    0.0   0.0
Atom3  1  2  0    R2    A1    0.0
Atom4  1  2  3    R3    A2    D1
. . .
```

(continues on next page)

(continued from previous page)

```
AtomN  NA NB NC   RN    AN    DN
*
```

The rules for connectivity in the "`internal`" mode are as follows:

- `NA`: The atom that the actual atom has a distance (`RN`) with.
- `NB`: The actual atom has an angle (`AN`) with atoms `NA` and `NB`.
- `NC`: The actual atom has a dihedral angle (`DN`) with atoms `NA`, `NB` and `NC`. This is the angle between the actual atom and atom `NC` when looking down the `NA-NB` axis.
- Note that - contrary to other parts in ORCA - atoms are counted starting from 1.

Angles are always given in degrees! The rules are compatible with those used in the well known MOPAC and ADF programs.

Finally, `gzmt` specifies internal coordinates in the format used by the Gaussian program. This resembles the following:

```
* gzmt 0 1
    C
    O   1   4.454280
    Si  2   1.612138    1   56.446186
    O   3   1.652560    2   114.631525  1   -73.696925
    C   4   1.367361    3   123.895399  2   -110.635060
    ...
*
```

An alternative way to specify coordinates in the input file is through the use of the `%coords` block, which is organized as follows:

```
%coords
 CTyp   xyz     # the type of coordinates = xyz or internal
 Charge 0       # the total charge of the molecule
 Mult   2       # the multiplicity = 2S+1
 Units  Angs    # the unit of length = angs or bohrs

 # the subblock coords is for the actual coordinates
 # for CTyp=xyz
  coords
     Atom1   x1  y1  z1
     Atom2   x2  y2  z2
  end
 # for CTyp=internal
  coords
     Atom1  0  0  0    0.0   0.0   0.0
     Atom2  1  0  0    R1    0.0   0.0
     Atom3  1  2  0    R2    A1    0.0
     Atom4  1  2  3    R3    A2    D1
      . . .
     AtomN  NA NB NC   RN    AN    DN
  end
end
```

5.2 Reading coordinates from external files

It is also possible to read the coordinates from external files. The most common format is a `.xyz` file, which can in principle contain more than one structure (see section *Multiple XYZ File Scans* for this multiple XYZ feature):

```
* xyzfile Charge Multiplicity Filename
```

For example:

```
* xyzfile 1 2 mycoords.xyz
```

A lot of graphical tools like Gabedit, molden or Jmol can write Gaussian Z-Matrices (`.gzmt`). ORCA can also read them from an external file with the following

```
* gzmtfile 1 2 mycoords.gzmt
```

Note that if multiple jobs are specified in the same input file then new jobs can read the coordinates from previous jobs. If no filename is given as fourth argument then the name of the actual job is automatically used.

```
... specification for the first job

$new_job
! keywords
* xyzfile 1 2
```

In this way, optimization and single point jobs can be very conveniently combined in a single, simple input file. Examples are provided in the following sections.

5.3 Special definitions

- **Dummy atoms** are defined in exactly the same way as any other atom, by using "DA", "X", or "Xx" as the atomic symbol.
- **Ghost atoms** are specified by adding ":" right after the symbol of the element (see *Counterpoise Correction*).
- **Point charges** are specified with the symbol "Q", followed by the charge (see *Inclusion of Point Charges*).
- **Embedding potentials** are specified by adding a ">" right after the symbol of the element (see *Embedding Potentials*).
- **Non-standard** isotopes or nuclear charges are specified with the statements "`M = ...`" and "`Z = ...`", respectively, after the atomic coordinate definition.

> **Note**
>
> 1. The nuclear charge can adopt non-integer values
> 2. When the nuclear charge is modified throughca "`Z = ...`" statement, the total charge of the system should still be calculated based on the unmodified charge. For example, for a calculation of a single hydrogen atom whose Z is set to 1.5, a charge of 0 and a spin multiplicity of 2 should be entered into the charge and multiplicity sections of the input file, despite that the actual total charge is 0.5.

- **Fragments** can be conveniently defined by declaring the fragment number a given atom belongs to in parentheses "`(n)`" following the element symbol (see *Fragment Specification*).
- **Frozen coordinates**, which are not changed during optimizations in Cartesian coordinates, are defined with a "$" symbol after the X, Y, and/or Z coordinate value (cf. constraints on all 3 Cartesian components *Constrained Optimizations*).

CHAPTER

SIX

RUNNING TYPICAL CALCULATIONS

Before entering the detailed documentation of the various features of ORCA it is instructive to provide a chapter that shows how "typical" tasks may be performed. This should make it easier for the user to get started on the program and not get lost in the details of how-to-do-this or how-to-do-that. We hope that the examples are reasonably intuitive.

6.1 Single Point Energies and Gradients

6.1.1 Hartree-Fock

Standard Single Points

In general single point calculations are fairly easy to run. What is required is the input of a method, a basis set and a geometry. For example, in order run a single point Hartree-Fock calculation on the CO molecule with the DEF2-SVP basis set type:

```
#
# My first ORCA calculation :-)
#
! HF DEF2-SVP
* xyz 0 1
  C  0  0  0
  O  0  0  1.13
*
```

As an example consider this simple calculation on the cyclohexane molecule that may serve as a prototype for this type of calculation.

```
# Test a simple direct HF calculation
! HF DEF2-SV(P)
* xyz 0 1
C   -0.79263     0.55338    -1.58694
C    0.68078     0.13314    -1.72622
C    1.50034     0.61020    -0.52199
C    1.01517    -0.06749     0.77103
C   -0.49095    -0.38008     0.74228
C   -1.24341     0.64080    -0.11866
H    1.10490     0.53546    -2.67754
H    0.76075    -0.97866    -1.78666
H   -0.95741     1.54560    -2.07170
H   -1.42795    -0.17916    -2.14055
H   -2.34640     0.48232    -0.04725
H   -1.04144     1.66089     0.28731
H   -0.66608    -1.39636     0.31480
H   -0.89815    -0.39708     1.78184
H    1.25353     0.59796     1.63523
```

(continues on next page)

(continued from previous page)

```
H    1.57519    -1.01856     0.93954
H    2.58691     0.40499    -0.67666
H    1.39420     1.71843    -0.44053
*
```

Basis Set Options

There is extensive flexibility in the specification of basis sets in ORCA. First of all, you are not only restricted to the basis sets that are built in ORCA, but can also read basis set definitions from files. In addition there is a convenient way to change basis sets on certain types of atoms or on individual atoms. Consider the following example:

```
# CuCl$_4$
! HF
%basis  basis  "SV"
        newGTO Cl "DUNNING-DZP" end
        end
* xyz -2 2
  Cu  0     0     0  newGTO "TZVPP" end
  Cl  2.25  0     0
  Cl -2.25  0     0
  Cl  0     2.25  0
  Cl  0    -2.25  0
*
```

In this example the basis set is initialized as the Ahlrichs split valence basis. Then the basis set on all atoms of type Cl is changed to SVP and finally the basis set for only the copper atom is changed to the more accurate TZVPP set. In this way you could treat different atom types or even individual groups in a molecule according to the desired accuracy. Similar functionality regarding per-element or per-atom assignments exists for effective core potentials. More details are provided in section *Choice of Basis Set*.

Sometimes you will like to change the ordering of the starting orbitals to obtain a different electronic state in the SCF calculation. For example, if we take the last input and want to converge to a ligand field excited state this can be achieved by:

```
! HF SV
%basis  newGTO Cl "Dunning-DZP" end
        end
%scf    rotate {48, 49, 90, 1, 1} end
        end
* xyz -2 2
  Cu  0     0     0   newGTO "TZVPP" end
  Cl  2.25  0     0
  Cl -2.25  0     0
  Cl  0     2.25  0
  Cl  0    -2.25  0
*
```

In the present case, MO 48 is the spin-down HOMO and MO49 the spin-down LUMO. Since we do a calculation on a Cu(II) complex (d^9 electron configuration) the beta LUMO corresponds with the "SOMO". Thus, by changing the SOMO we proceed to a different electronic state (in this case the one with the "hole" in the "d_{xy}" orbital instead of the "$d_{x^2-y^2}$" orbital). The interchange of the initial guess MOs is achieved by the command `rotate {48, 49, 90, 1, 1}` end. What this does is the following: take the initial guess MOs 48 and 49 and rotate them by an angle of 90 degree (this just interchanges them). The two last numbers mean that both orbitals are from the spin-down set. For RHF or ROHF calculations the operator would be 0. In general you would probably first take a look at the initial guess orbitals before changing them.

SCF and Symmetry

Upon request, the SCF program produces symmetry adapted orbitals. This can help to converge the SCF on specific excited states of a given symmetry. Take for example the cation H_2O^+: We first run the simple job:

```
! SVP UseSym

* xyz 1 2
  O     0.000000    0.000000    0.068897
  H     0.000000    0.788011   -0.546765
  H     0.000000   -0.788011   -0.546765
*
```

The program will recognize the C_{2v} symmetry and adapt the orbitals to this:

```
------------------
SYMMETRY DETECTION
------------------
The point group will now be determined using a tolerance of 1.0000e-04.
Splitting atom subsets according to nuclear charge, mass and basis set.
Splitting atom subsets according to distance from the molecule's center.
Identifying relative distance patterns of the atoms.
Splitting atom subsets according to atoms' relative distance patterns.
Bring atoms of each subset into input order.
The molecule is planar.
There is at least one atom subset not centered around the molecule's center.
The molecule does not have a center of inversion.
Analyzing the first atom subset for its symmetry.
Testing point group C2v.
Success!
This point group has been found:    C2v
Largest non-degenerate subgroup:    C2v


Mass-centered symmetry-perfected Cartesians (point group C2v):

Atom          Symmetry-perfected Cartesians (x, y, z; au)
  0       0.000000000000    0.000000000000    0.130195951333
  1       0.000000000000    1.489124980517   -1.033236619729
  2       0.000000000000   -1.489124980517   -1.033236619729

------------------
SYMMETRY REDUCTION
------------------
ORCA supports only abelian point groups.
It is now checked, if the determined point group is supported:
Point Group ( C2v   ) is          ... supported

(Re)building abelian point group:
Creating Character Table          ... done
Making direct product table       ... done
Constructing symmetry operations  ... done
Creating atom transfer table      ... done
Creating asymmetric unit          ... done

----------------------
ASYMMETRIC UNIT IN C2v
----------------------
  \#  AT     MASS              COORDS (A.U.)             BAS
   0 O   15.9990   0.00000000   0.00000000   0.13019595   0
   1 H    1.0080   0.00000000   1.48912498  -1.03323662   0
```

(continues on next page)

(continued from previous page)

```
----------------------
SYMMETRY ADAPTED BASIS
----------------------
The coefficients for the symmetry adapted linear combinations (SALCS)
of basis functions will now be computed:
Number of basis functions         ...    24
Preparing memory                  ... done
Constructing Gamma(red)           ... done
Reducing Gamma(red)               ... done
Constructing SALCs                ... done
Checking SALC integrity           ... nothing suspicious
Normalizing SALCs                 ... done

Storing the symmetry object:
Symmetry file                     ... C05S01_030.sym.tmp
Writing symmetry information      ... done
```

The initial guess in the SCF program will then recognize and freeze the occupation numbers in each irreducible representation of the C_{2v} point group.

```
The symmetry of the initial guess is 2-B1
Irrep occupations for operator 0
    A1 -    3
    A2 -    0
    B1 -    1
    B2 -    1
Irrep occupations for operator 1
    A1 -    3
    A2 -    0
    B1 -    0
    B2 -    1
```

The calculation converges smoothly to

```
Total Energy       :          -75.56349710 Eh           -2056.18729 eV
```

With the final orbitals being:

```
SPIN UP ORBITALS
  NO   OCC          E(Eh)            E(eV)    Irrep
   0   1.0000     -21.127827      -574.9174    1-A1
   1   1.0000      -1.867576       -50.8193    2-A1
   2   1.0000      -1.192139       -32.4397    1-B2
   3   1.0000      -1.124657       -30.6035    1-B1
   4   1.0000      -1.085062       -29.5260    3-A1
   5   0.0000      -0.153303        -4.1716    4-A1
   6   0.0000      -0.071324        -1.9408    2-B2
...
                 SPIN DOWN ORBITALS
  NO   OCC          E(Eh)            E(eV)    Irrep
   0   1.0000     -21.081198      -573.6486    1-A1
   1   1.0000      -1.710193       -46.5367    2-A1
   2   1.0000      -1.152855       -31.3708    1-B2
   3   1.0000      -1.032556       -28.0973    1-B1
   4   0.0000      -0.306683        -8.3453    3-A1
   5   0.0000      -0.139418        -3.7937    4-A1
   6   0.0000      -0.062261        -1.6942    2-B2
   7   0.0000       0.374727        10.1968    3-B2
...
```

Suppose now that we want to converge on an excited state formed by flipping the spin-beta HOMO and LUMO that have different symmetries.

```
! SVP UseSym
! moread
%moinp "Test-SYM-H2O+.gbw"
%scf rotate {3,4,90,1,1}
            end
     end
* xyz 1 2
  O     0.000000    0.000000    0.068897
  H     0.000000    0.788011   -0.546765
  H     0.000000   -0.788011   -0.546765
*
```

The program now finds:

```
Irrep occupations for operator 0
    A1 -    3
    A2 -    0
    B1 -    1
    B2 -    1
Irrep occupations for operator 1
    A1 -    2
    A2 -    0
    B1 -    1
    B2 -    1
```

And converges smoothly to

```
Total Energy       :          -75.48231924 Eh           -2053.97833 eV
```

Which is obviously an excited state of the H_2O^+ molecule. In this situation (and in many others) it is an advantage to have symmetry adapted orbitals.

SymRelax. Sometimes, one may want to obtain the ground state of a system but due to a particularly bad initial guess, the calculation converges to an excited state. In such cases, the following option can be used:

```
%method SymRelax True
        end
```

This will allow the occupation numbers in each irreducible representation to change if and only if a virtual orbital has a lower energy than an occupied one. Hence, nothing will change for the excited state of H_2O^+ discussed above. However, the following calculation

```
! SVP UseSym
! moread
%moinp "Test-SYM-H2O+.gbw"
%scf rotate {3,13,90,1,1}
            end
     end
* xyz 1 2
  O     0.000000    0.000000    0.068897
  H     0.000000    0.788011   -0.546765
  H     0.000000   -0.788011   -0.546765
*
```

which converges to a high-lying excited state:

```
Total Energy       :          -73.87704009 Eh           -2010.29646 eV
...
                 SPIN UP ORBITALS
  NO   OCC          E(Eh)            E(eV)    Irrep
   0   1.0000     -21.314859      -580.0068    1-A1
   1   1.0000      -1.976707       -53.7889    2-A1
```

(continues on next page)

(continued from previous page)

```
   2   1.0000      -1.305096       -35.5135    3-A1
   3   1.0000      -1.253997       -34.1230    1-B2
   4   1.0000      -1.237415       -33.6718    1-B1
   5   0.0000      -0.122295        -3.3278    4-A1
   6   0.0000      -0.048384        -1.3166    2-B2
...
                 SPIN DOWN ORBITALS
  NO   OCC          E(Eh)            E(eV)    Irrep
   0   1.0000     -21.212928      -577.2331    1-A1
   1   1.0000      -1.673101       -45.5274    2-A1
   2   1.0000      -1.199599       -32.6427    1-B2
   3   1.0000       0.727889        19.8069    1-A2
   4   0.0000      -0.449647       -12.2355    3-A1
   5   0.0000      -0.371861       -10.1189    1-B1
   6   0.0000      -0.106365        -2.8943    4-A1
...
```

would revert to the ground state with the `SymRelax` option.

SCF and Memory

As the SCF module cannot restrict its use of memory to MaxCore we introduced an estimation of the expected memory consumption. If the memory needed is larger than MaxCore ORCA will abort.

To check, if a certain job can be run with a given amount of MaxCore, you can ask for the estimation of memory requirements by

```
%scf DryRun true
     end
```

ORCA will finish execution after having printed the estimated amount of memory needed.

If you want to run the calculation (if doable), and only are interested in the estimated memory consumption, you can ask for the printing via

```
%scf Print[P_SCFMemInfo] 1
     end
```

Note

The estimation is given per process. If you want to run a parallel job, you will need the estimated memory × number of parallel processes.

6.1.2 MP2

MP2 and RI-MP2 Energies

You can do conventional or integral direct MP2 calculations for RHF, UHF or high-spin ROHF reference wavefunctions. MP3 functionality is not implemented as part of the MP2 module, but can be accessed through the MDCI module. Analytic gradients are available for RHF and UHF. The analytic MP2-Hessians have been deprecated with ORCA-6.0. The frozen core approximation is used by default. For RI-MP2 the $\langle \hat{S}^2 \rangle$ expectation value is computed in the unrestricted case according to [531]. An extensive coverage of MP2 exists in the literature.[96, 187, 258, 359, 370, 450, 473, 495, 575, 693, 736, 839, 881, 882]

```
! MP2 def2-TZVP TightSCF
%amp2 MaxCore
```

(continues on next page)

(continued from previous page)

```
      end
%paras  rCO = 1.20
        ACOH = 120
        rCH  = 1.08
        end
* int 0 1
C 0 0 0 0.00 0.0 0.00
O 1 0 0 {rCO} 0.0 0.00
H 1 2 0 {rCH} {ACOH} 0.00
H 1 2 3 {rCH} {ACOH} 180.00
*
```

Note

There are two algorithms for MP2 calculations without the RI approximation. The first one uses main memory as much as possible. The second one uses more disk space and is usually faster (in particular, if you run the calculations in single precision using `! FLOAT, UCFLOAT or CFLOAT`). The memory algorithm is used by specifying `Q1Opt >0` in the `%mp2` block whereas the disk based algorithm is the default or specified by `Q1Opt = -1`. Gradients are presently only available for the memory based algorithm.

The RI approximation to MP2[96, 258, 881, 882] is fairly easy to use, too. It results in a tremendous speedup of the calculation, while errors in energy differences are very small. For example, consider the same calculation as before:

```
# only the auxiliary basis set def2-TZVP/C is added to
# the keyword line
#
! RI-MP2 def2-TZVP def2-TZVP/C TightSCF
%mp2     MaxCore 100
         end
%paras   rCO  =  1.20
         ACOH =  120
         rCH  =  1.08
         end
* int 0 1
 C  0 0 0 0.00   0.0     0.00
 O  1 0 0 {rCO}  0.0     0.00
 H  1 2 0 {rCH} {ACOH}   0.00
 H  1 2 3 {rCH} {ACOH}   180.00
*
```

Generally, the RI approximation can be switched on by setting `RI true` in the `%mp2` block. Specification of an appropriate auxiliary basis set ("/C") for correlated calculations is required. Note that if the RIJCOSX method (section *Hartree–Fock and Hybrid DFT Calculations with RIJCOSX*) or the RI-JK method (section *Hartree–Fock and Hybrid DFT Calculations with RI-JK*) is used to accelerate the SCF calculation, then two basis sets should be specified: firstly the appropriate Coulomb ("/J") or exchange fitting set ("/JK"), and secondly the correlation fitting set ("/C"), as shown in the example below.

```
# Simple input line for RIJCOSX:
! RHF RI-MP2 RIJCOSX def2-TZVP def2/J def2-TZVP/C TightSCF

# Simple input line for RI-JK:
! RHF RI-MP2 RI-JK def2-TZVP def2/JK def2-TZVP/C TightSCF
```

The MP2 module can also do Grimme's spin-component scaled MP2 [318]. It is a semi-empirical modification of MP2 which applies different scaling factors to same-spin and opposite-spin components of the MP2 energy. Typically it gives fairly bit better results than MP2 itself.

```
#
# Spin-component scaled MP2 example
#
! SCS-MP2 def2-TZVPP TightSCF
%paras   rCO  =  1.20
         ACOH =  120
         rCH  =  1.08
         end
* int 0 1
 C  0 0 0 0.00   0.0     0.00
 O  1 0 0 {rCO}  0.0     0.00
 H  1 2 0 {rCH}  {ACOH}  0.00
 H  1 2 3 {rCH}  {ACOH}  180.00
*
```

Energy differences with SCS-MP2 appear to be much better than with MP2 itself according to Grimme's detailed evaluation study. For the sake of efficiency, it is beneficial to make use of the RI approximation using the `RI-SCS-MP2` keyword. The opposite-spin and same-spin scaling factors can be modified using `PS` and `PT` in the `%mp2` block, respectively. By default, `PS` $= 6/5$ and `PT` $= 1/3$.

NOTE

- In very large RI-MP2 runs you can cut down the amount of main memory used by a factor of two if you use the keyword `! FLOAT`. This is more important in gradient runs than in single point runs. Deviations from double precision values for energies and gradients should be in the μEh and sub-μEh range. However, we have met cases where this option introduced a large and unacceptable error, in particular in transition metal calculations. You are therefore advised to be careful and check things out beforehand.

A word of caution is due regarding MP2 calculations with a linearly dependent basis. This can happen, for example, with very diffuse basis sets (see *Linear Dependence* for more information). If some vectors were removed from the basis in the SCF procedure, those redundant vectors are still present as "virtual" functions with a zero orbital energy in the MP2 calculation. When the number of redundant vectors is small, this is often not critical (and when their number is large, one should probably use a different basis). However, it is better to avoid linearly dependent basis sets in MP2 calculations whenever possible. Moreover, in such a situation the orbitals should not be read with the `MORead` and `NoIter` keywords, as that is going to produce wrong results!

Frozen Core Options

In MP2 energy and gradient runs the Frozen Core (FC) approximation is applied by default. This implies that the core electrons are not included in the perturbation treatment, since the inclusion of dynamic correlation in the core electrons usually effects relative energies or geometry parameters insignificantly.

The frozen core option can be switched on or off with `FrozenCore` or `NoFrozenCore` in the simple input line. Furthermore, frozen orbitals can be selected by means of an energy window:

```
%method FrozenCore FC_EWIN end
%mp2 ewin -1.5, 1.0e3 end
```

More information and the different options can be found in section *Frozen Core Options*

Orbital Optimized MP2 Methods

By making the Hylleraas functional stationary with respect to the orbital rotations one obtains the orbital-optimized MP2 method that is implemented in ORCA in combination with the RI approximation (OO-RI-MP2). One obtains from these calculations orbitals that are adjusted to the dynamic correlation field at the level of second order many-body perturbation theory. Also, the total energy of the OO-RI-MP2 method is lower than that of the RI-MP2 method itself. One might think of this method as a special form of multiconfigurational SCF theory except for the fact that the Hamiltonian is divided into a 0^{th} order term and a perturbation.

The main benefit of the OO-RI-MP2 method is that it "repairs" the poor Hartree–Fock orbitals to some extent which should be particularly beneficial for systems which suffer from the imbalance in the Hartree-Fock treatment of the Coulomb and the Exchange hole. Based on the experience gained so far, the OO-RI-MP2 method is no better than RI-MP2 itself for the thermochemistry of organic molecules. However, for reactions barriers and radicals the benefits of OO-MP2 over MP2 are substantial. This is particularly true with respect to the spin-component scaled variant of OO-RI-MP2 that is OO-RI-SCS-MP2. Furthermore, the OO-RI-MP2 method substantially reduces the spin contamination in UHF calculations on radicals.

Since every iteration of the OO-MP2 method is as expensive as a RI-MP2 relaxed density calculation, the computational cost is much higher than for RI-MP2 itself. One should estimate about a factor of 10 increase in computational time with respect to the RI-MP2 time of a normal calculation. This may still be feasible for calculations in the range of 1000–2000 basis functions (the upper limit, however, implies very significant computational costs). A full assessment of the orbital optimized MP2 method has been published.[620]

OO-RI-MP2 is triggered either with `! OO-RI-MP2` or `! OO-RI-SCS-MP2` (with spin component scaling) in the simple input line or by `OrbOpt true` in the `%mp2` block. The method comes with the following new variables:

```
%mp2 OrbOpt  true   # turns on the orbital optimization
     CalcS2  false  # calculate the S**2 expectation value
                    # in spin-unrestricted calculations
     MaxOrbIter 64  # Max. number of iterations
     MP2Shift  0.1  # Level shift for the procedure
     end
```

The solver is a simple DIIS type scheme with additional level shifting. We have found that it is not really beneficial to first converge the Hartree-Fock equations. Thus it is sensible to additionally use the keyword `! noiter` in order to turn off the standard Hartree-Fock SCF process before entering the orbital optimizations.

The OO-RI-MP2 method is implemented for RHF and UHF reference wavefunctions. Analytic gradients are available.

The density does not need to be requested separately in OO-RI-MP2 calculations because it is automatically calculated. Also, there is no distinction between relaxed and unrelaxed densities because the OO-RI-MP2 energy is fully stationary with respect to all wavefunction parameters and hence the unrelaxed and relaxed densities coincide.

MP2 and RI-MP2 Gradients

Geometry optimization with MP2, RI-MP2, SCS-MP2 and RI-SCS-MP2 proceeds just as with any SCF method. With frozen core orbitals, second derivatives of any kind are currently only available numerically. The RIJCOSX approximation (section *Hartree–Fock and Hybrid DFT Calculations with RIJCOSX*) is supported in RI-MP2 and hence also in double-hybrid DFT gradient runs (it is in fact the default for double-hybrid DFT since ORCA 5.0). This leads to large speedups in larger calculations, particularly if the basis sets are accurate.

```
#
# MP2 optimization example
#
! SCS-MP2 def2-TZVP OPT NoFrozenCore
* int 0 1
 C  0 0 0 0.00   0.0     0.00
 O  1 0 0 1.20   0.0     0.00
 H  1 2 0 1.09 120.0     0.00
```

(continues on next page)

(continued from previous page)

```
 H  1 2 3 1.09 120.0   180.00
*
```

This job results in:

```
-------------------------------------------------------------------------
Redundant Internal Coordinates

--- Optimized Parameters ---
(Angstroem and degrees)

Definition                    OldVal   dE/dq     Step     FinalVal
-------------------------------------------------------------------------
1. B(O   1,C   0)               1.2081  0.000488 -0.0003    1.2078
2. B(H   2,C   0)               1.1027  0.000009 -0.0000    1.1027
3. B(H   3,C   0)               1.1027  0.000009 -0.0000    1.1027
4. A(O   1,C   0,H   3)         121.85  0.000026   -0.00    121.85
5. A(H   2,C   0,H   3)         116.29 -0.000053    0.01    116.30
6. A(O   1,C   0,H   2)         121.85  0.000026   -0.00    121.85
7. I(O   1,H   3,H   2,C   0)     -0.00 -0.000000    0.00      0.00
-------------------------------------------------------------------------
```

Just to demonstrate the accuracy of RI-MP2, here is the result with RI-SCS-MP2 instead of SCS-MP2, with the addition of def2-TZVP/C:

```
-------------------------------------------------------------------------
Redundant Internal Coordinates

--- Optimized Parameters ---
(Angstroem and degrees)

Definition                    OldVal   dE/dq     Step     FinalVal
-------------------------------------------------------------------------
1. B(O   1,C   0)               1.2081  0.000487 -0.0003    1.2078
2. B(H   2,C   0)               1.1027  0.000009 -0.0000    1.1027
3. B(H   3,C   0)               1.1027  0.000009 -0.0000    1.1027
4. A(O   1,C   0,H   3)         121.85  0.000026   -0.00    121.85
5. A(H   2,C   0,H   3)         116.29 -0.000053    0.01    116.30
6. A(O   1,C   0,H   2)         121.85  0.000026   -0.00    121.85
7. I(O   1,H   3,H   2,C   0)     -0.00  0.000000   -0.00     -0.00
-------------------------------------------------------------------------
```

You see that *nothing* is lost in the optimized geometry through the RI approximation thanks to the efficient and accurate RI-auxiliary basis sets of the Karlsruhe group (in general the deviations in the geometries between standard MP2 and RI-MP2 are very small). Thus, RI-MP2 really is a substantial improvement in efficiency over standard MP2.

Geometric gradients can be calculated with RI-MP2 in conjunction with the RIJCOSX method. They are called the same way as with a conventional SCF wave function, for example to perform a geometry optimization with tight convergence parameters: (Please note that you have to switch on NumFreq for the MP2-Hessian, as the analytical (RI-)MP2-Hessians are no longer available).

```
! RI-MP2 def2-TZVPP def2/J def2-TZVPP/C TightSCF RIJCOSX
! TightOpt
...
```

MP2 Properties, Densities and Natural Orbitals

The MP2 method can be used to calculate electric and magnetic properties such as dipole moments, polarizabilities, hyperfine couplings, g-tensors or NMR chemical shielding tensors. For this purpose, the appropriate MP2 density needs to be requested - otherwise the properties are calculated using the SCF density!

Two types of densities can be constructed - an "unrelaxed" density (which basically corresponds to the MP2 expectation value density) and a "relaxed" density which incorporates orbital relaxation. For both sets of densities a population analysis is printed if the SCF calculation also requested this population analysis. These two densities are stored as `JobName.pmp2ur.tmp` and `JobName.pmp2re.tmp`, respectively. For the open shell case case the corresponding spin densities are also constructed and stored as `JobName.rmp2ur.tmp` and `JobName.rmp2re.tmp`.

In addition to the density options, the user has the ability to construct MP2 natural orbitals. If relaxed densities are available, the program uses the relaxed densities and otherwise the unrelaxed ones. The natural orbitals are stored as `JobName.mp2nat` which is a GBW type file that can be read as input for other jobs (for example, it is sensible to start CASSCF calculations from MP2 natural orbitals). The density construction can be controlled separately in the input file (even without running a gradient or optimization) by:

```
#
# MP2 densities and natural orbitals
#
%mp2  Density none         # no density
              unrelaxed    # unrelaxed density
              relaxed      # relaxed density
      NatOrbs true         # Natural orbital construction on or off
      end
```

Below is a calculation of the dipole and quadrupole moments of a water molecule:

```
! RI-MP2 def2-SVP def2-SVP/C
%mp2 density relaxed end
%elprop dipole     true
        quadrupole true
        end
* int 0 1
O 0 0 0 0      0      0
H 1 0 0 0.9584 0      0
H 1 2 0 0.9584 104.45 0
*
```

Another example is a simple g-tensor calculation with MP2:

```
! RI-MP2 def2-SVP def2-SVP/C TightSCF SOMF(1X) NoFrozenCore
%eprnmr gtensor  1
        ori     CenterOfElCharge
end
%mp2 density relaxed end
* int 1 2
O 0 0 0 0      0      0
H 1 0 0 1.1056 0      0
H 1 2 0 1.1056 109.62 0
*
```

NMR chemical shielding as well as g-tensor calculations with GIAOs are only available for RI-MP2. The input for NMR chemical shielding looks as follows:

```
! RIJK RI-MP2 def2-SVP def2/JK def2-SVP/C TightSCF NMR NoFrozenCore
%mp2
  density relaxed # required
end
* int 0 1
```

(continues on next page)

(continued from previous page)

```
O 0 0 0 0      0      0
H 1 0 0 1.1056 0      0
H 1 2 0 1.1056 109.62 0
*
```

Note that by default core electrons are not correlated unless the `NoFrozenCore` keyword is present.

For details, see sections *The Second Order Many Body Pertubation Theory Module (MP2)* and *MP2 level magnetic properties*.

Explicitly correlated MP2 calculations

ORCA features an efficient explicit correlation module that is available for MP2 and coupled-cluster calculations (section *Explicitly Correlated MP2 and CCSD(T) Calculations*). It is described below in the context of coupled-cluster calculations.

Local MP2 calculations

Purely domain-based local MP2 methodology dates back to Pulay and has been developed further by Werner, Schütz and co-workers. ORCA features a local MP2 method (DLPNO-MP2) that combines the ideas of domains and local pair natural orbitals, so that RI-MP2 energies are reproduced efficiently to within chemical accuracy. Due to the intricate connections with other DLPNO methods, reading of the sections *Local Coupled Pair and Coupled-Cluster Calculations* and and *Local correlation* is recommended. A full description of the method for RHF reference wave functions has been published.[684]

Since DLPNO-MP2 employs an auxiliary basis set to evaluate integrals, its energies converge systematically to RI-MP2 as thresholds are tightened. The computational effort of DLPNO-MP2 with default settings is usually comparable with or less than that of a Hartree-Fock calculation. However, for small and medium-sized molecules, RI-MP2 is even faster than DLPNO-MP2.

Calculations on open-shell systems are supported through a UHF treatment. While most approximations are consistent between the RHF and UHF versions, this is not true for the PNO spaces. **DLPNO-MP2 gives different energies for closed-shell molecules in the RHF and UHF formalisms. When calculating reaction energies or other energy differences involving open-shell species, energies of closed-shell species must also be calculated with UHF-DLPNO-MP2, and not with RHF-DLPNO-MP2.** As for canonical MP2, ROHF reference wave functions are subject to an ROMP2 treatment through the UHF machinery. It is not consistent with the RHF version of DLPNO-MP2, unlike in the case of RHF-/ROHF-DLPNO-CCSD.

Input for DLPNO-MP2 requires little specification from the user:

```
# DLPNO-MP2 calculation with standard settings
# sufficient for most purposes
! def2-TZVP def2-TZVP/C DLPNO-MP2 TightSCF

# OR: DLPNO-MP2 with tighter thresholds
# May be interesting for weak interactions, calculations with diffuse basis sets etc.
! def2-TZVP def2-TZVP/C DLPNO-MP2 TightPNO TightSCF

%maxcore 2000

*xyz 0 1
... (coordinates)
*
```

Noteworthy aspects of the DLPNO-MP2 method:

- Both DLPNO-CCSD(T) and DLPNO-MP2 are linear-scaling methods (albeit the former has a larger prefactor). This means that if a DLPNO-MP2 calculation can be performed, DLPNO-CCSD(T) is often going to be within reach, too. However, CCSD(T) is generally much more accurate than MP2 and thus should be given preference.

- A correlation fitting set must be provided, as the method makes use of the RI approximation.
- Canonical RI-MP2 energy differences are typically reproduced to within a fraction of 1 kcal/mol. The default thresholds have been chosen so as to reproduce about 99.9 % of the total RI-MP2 correlation energy.
- The preferred way to control the accuracy of the method is by means of specifying "LoosePNO", "NormalPNO" and "TightPNO" keywords. "NormalPNO" corresponds to default settings and does not need to be given explicitly. More details and an exhaustive list of input parameters are provided in section *Local MP2*. Note that the thresholds differ from DLPNO coupled cluster.
- Results obtained from RI-MP2 and DLPNO-MP2, or from DLPNO-MP2 with different accuracy settings, must never be mixed, such as when computing energy differences. In calculations involving open-shell species, even the closed-shell molecules need to be subject to a UHF treatment.
- Spin-component scaled DLPNO-MP2 calculations are invoked by using the `! DLPNO-SCS-MP2` keyword instead of `! DLPNO-MP2` in the simple input line. Weights for same-spin and opposite-spin contributions can be adjusted as described for the canonical SCS-MP2 method. Likewise, there is a `DLPNO-SOS-MP2` keyword to set the parameters defined by the SOS-MP2 method (but there is no Laplace transformation involved).
- The frozen core approximation is used by default. If core orbitals are involved in the calculation, they are subject to the treatment described in section *Local MP2*.
- Calculations can be performed in parallel.
- It may be beneficial to accelerate the Hartree-Fock calculation by means of the RIJCOSX method (requiring specification of a second auxiliary set).

Explicit correlation has been implemented in the DLPNO-MP2-F12 methodology for RHF reference wave functions.[653] The available approaches are C (keyword `! DLPNO-MP2-F12`) and the somewhat more approximate D (keyword `! DLPNO-MP2-F12/D`). Approach D is generally recommended as it results in a significant speedup while leading only to small errors relative to approach C. In addition to the MO and correlation fitting sets, a CABS basis set is also required for both F12 approaches as shown below.

```
# DLPNO-MP2-F12 calculation using approach C
! cc-pVDZ-F12 aug-cc-pVDZ/C cc-pVDZ-F12-CABS DLPNO-MP2-F12 TightSCF

# OR: DLPNO-MP2-F12 calculation using approach D (recommended)
! cc-pVDZ-F12 aug-cc-pVDZ/C cc-pVDZ-F12-CABS DLPNO-MP2-F12/D TightSCF
```

Local MP2 derivatives

Analytical gradients and the response density are available for the RHF variant of the DLPNO-MP2 method.[685, 686] Usage is as simple as that of RI-MP2. For example, the following input calculates the gradient and the natural orbitals:

```
! DLPNO-MP2 def2-SVP def2-SVP/C TightSCF EnGrad
%MaxCore 512
# With 'EnGrad', specifying 'density relaxed' is unnecessary.
# However, it is needed when calculating properties without the gradient.
%MP2 Density Relaxed
     NatOrbs True
     End
*xyz 0 1
C   0.000   0.000   0.000
O   0.000   0.000   1.162
O   0.000   0.000  -1.162
*
```

The implementation supports spin-component scaling and can be used together with double-hybrid density functionals. The latter are invoked with the name of the functional preceded by "`DLPNO-`". A simple geometry optimization with a double-hybrid density functional is illustrated in the example below:

```
! DLPNO-B2PLYP D3 NormalPNO def2-TZVP def2-TZVP/C Opt
%MaxCore 1000
*xyz 0 1
O   0.000   0.000   0.000
H   0.000   0.000   1.000
H   0.000   1.000   0.000
*
```

For smaller systems, the performance difference between DLPNO-MP2 and RI-MP2 is not particularly large, but very substantial savings in computational time over RI-MP2 can be achieved for systems containing more than approximately 70-80 atoms.

Since MP2 is an expensive method for geometry optimizations, it is generally a good idea to use well-optimized starting structures (calculated, for example, with a dispersion-corrected DFT functional). Moreover, it is highly advisable to employ accurate Grids for RIJCOSX or the exchange-correlation functional (if applicable), as the SCF iterations account only for a fraction of the overall computational cost. If calculating calculating properties without requesting the gradient, `Density Relaxed` needs to be specified in the `%MP2-block`.

Only the Foster-Boys localization scheme is presently supported by the derivatives implementation. The default localizer in DLPNO-MP2 is `AHFB`, and changing this setting is strongly discouraged, since tightly converged localized orbitals are necessary to calculate the gradient.

Analytical second derivatives for closed-shell DLPNO-MP2 are available for the calculation of NMR shielding and static polarizability tensors.[826] The implementation supports spin-component scaling and double-hybrid functionals. Errors in the calculated properties are well below 0.5% when `NormalPNO` thresholds are used. Refer to section *Local MP2 Response Properties* for more information about the DLPNO-MP2 second derivatives implementation, as well as to the sections on electric (*Electric Properties*) and magnetic (*EPR and NMR properties*) properties and CP-SCF settings (*CP-SCF Options*). Below is an example for a simple DLPNO-MP2 NMR shielding calculation:

```
! DLPNO-MP2 def2-TZVP def2-TZVP/C TightSCF NMR
# MP2 relaxed density is requested automatically
*xyz 0 1
  H 0 0 0
  F 0 0 0.9
*
```

6.1.3 Coupled-Cluster and Coupled-Pair Methods

Basics

The coupled-cluster method is presently available for RHF and UHF references. The implementation is fairly efficient and suitable for large-scale calculations. The most elementary use of this module is fairly simple.

```
!  METHOD
# where METHOD is:
#   CCSD CCSD(T) QCISD QCISD(T) CPF/n NCPF/n CEPA/n NCEPA/n
#   (n=1,2,3 for all variants) ACPF NACPF AQCC CISD

!  AOX-METHOD
# computes contributions from integrals with 3- and 4-external
# labels directly from AO integrals that are pre-stored in a
# packed format suitable for efficient processing

!  AO-METHOD
# computes contributions from integrals with 3- and 4-external
# labels directly from AO integrals. Can be done for integral
# direct and conventional runs. In particular, the conventional
# calculations can be very efficient
```

(continues on next page)

(continued from previous page)

```
!  MO-METHOD   (this is the default)
# performs a full four index integral transformation. This is
# also often a good choice

!  RI-METHOD
# selects the RI approximation for all integrals. Rarely advisable

!  RI34-METHOD
# selects the RI approximation for the integrals with 3- and 4-
# external labels
#
# The module has many additional options that are documented
# later in the manual.

! RCSinglesFock
! RIJKSinglesFock
! NoRCSinglesFock
! NoRIJKSinglesFock
# Keywords to select the way the so-called singles Fock calculation
# is evaluated. The first two keywords turn on, the second two turn off
# RIJCOSX or RIJK, respectively.
```

Note

- The same FrozenCore options as for MP2 are applied in the MDCI module.
- Since ORCA 4.2, an additional term, called "4th-order doubles-triples correction" is considered in open-shell CCSD(T). To reproduce previous results, one should use a keyword,

```
%mdci
  Include_4thOrder_DT_in_Triples  false
end
```

The computational effort for these methods is high — $O(N^6)$ for all methods and $O(N^7)$ if the triples correction is to be computed (calculations based on an unrestricted determinant are roughly 3 times more expensive than closed-shell calculations and approximately six times more expensive if triple excitations are to be calculated). This restricts the calculations somewhat: on presently available PCs 300–400 basis functions are feasible and if you are patient and stretch it to the limit it may be possible to go up to 500–600; if not too many electrons are correlated maybe even up to 800–900 basis functions (when using AO-direct methods).

Tip

- For calculations on small molecules and large basis sets the MO-METHOD option is usually the most efficient; say perhaps up to about 300 basis functions. For integral conventional runs, the AO-METHOD may even more efficient.
- For large calculations (>300 basis functions) the AO-METHOD option is a good choice. If, however, you use very deeply contracted basis sets such as ANOs these calculations should be run in the integral conventional mode.
- AOX-METHOD is usually slightly less efficient than MO-METHOD or AO-METHOD.
- RI-METHOD is seldom the most efficient choice. If the integral transformation time is an issue than you can select `%mdci trafotype trafo_ri` or choose RI-METHOD and then `%mdci kcopt kc_ao`.
- Regarding the singles Fock keywords (RCSinglesFock, etc.), the program usually decides which method to use to evaluate the singles Fock term. For more details on the nature of this term, and options related to its evaluation, see *The singles Fock term*.

To put this into perspective, consider a calculation on serine with the cc-pVDZ basis set — a basis on the lower end of what is suitable for a highly correlated calculation. The time required to solve the equations is listed in Table 6.1. We can draw the following conclusions:

- As long as one can store the integrals and the I/O system of the computer is not the bottleneck, the most efficient way to do coupled-cluster type calculations is usually to go via the full transformation (it scales as $O(N^5)$ whereas the later steps scale as $O(N^6)$ and $O(N^7)$ respectively).
- AO-based coupled-cluster calculations are not much inferior. For larger basis sets (i.e. when the ratio of virtual to occupied orbitals is larger), the computation times will be even more favorable for the AO based implementation. The AO direct method uses much less disk space. However, when you use a very expensive basis set the overhead will be larger than what is observed in this example. Hence, conventionally stored integrals — if affordable — are a good choice.
- AOX-based calculations run at essentially the same speed as AO-based calculations. Since AOX-based calculations take four times as much disk space, they are pretty much outdated and the AOX implementation is only kept for historical reasons.
- RI-based coupled-cluster methods are significantly slower. There are some disk space savings, but the computationally dominant steps are executed less efficiently.
- CCSD is at most 10% more expensive than QCISD. With the latest AO implementation the awkward coupled-cluster terms are handled efficiently.
- CEPA is not much more than 20% faster than CCSD. In many cases CEPA results will be better than CCSD and then it is a real saving compared to CCSD(T), which is the most rigorous.
- If triples are included practically the same comments apply for MO versus AO based implementations as in the case of CCSD.

ORCA is quite efficient in this type of calculation, but it is also clear that the range of application of these rigorous methods is limited as long as one uses canonical MOs. ORCA implements novel variants of the so-called local coupled-cluster method which can calculate large, real-life molecules in a linear scaling time. This will be addressed in Sec. *Local Coupled Pair and Coupled-Cluster Calculations*.

Table 6.1: Computer times (minutes) for solving the coupled-cluster/coupled-pair equations for Serine (cc-pVDZ basis set)

Method	SCFMode	Time (min)
MO-CCSD	`Conv`	38.2
AO-CCSD	`Conv`	47.5
AO-CCSD	`Direct`	50.8
AOX-CCSD	`Conv`	48.7
RI-CCSD	`Conv`	64.3
AO-QCISD	`Conv`	44.8
AO-CEPA/1	`Conv`	40.5
MO-CCSD(T)	`Conv`	147.0
AO-CCSD(T)	`Conv`	156.7

All of these methods are designed to cover dynamic correlation in systems where the Hartree-Fock determinant dominates the wavefunctions. The least attractive of these methods is CISD which is not size-consistent and therefore practically useless. The most rigorous are CCSD(T) and QCISD(T). The former is perhaps to be preferred, since it is more stable in difficult situations.[1] One can get highly accurate results from such calculations. However, one only gets this accuracy in conjunction with large basis sets. It is perhaps not very meaningful to perform a CCSD(T) calculation with a double-zeta basis set (see Table 6.2). The very least basis set quality required for meaningful results would perhaps be something like def2-TZVP(-f) or preferably def2-TZVPP (cc-pVTZ, ano-pVTZ). For accurate results quadruple-zeta and even larger basis sets are required and at this stage the method is restricted to rather small systems.

[1] The exponential of the T1 operator serves to essentially fully relax the orbitals of the reference wavefunction. This is not included in the QCISD model that only features at most a linear T1T2 term in the singles residuum. Hence, if the Hartree-Fock wavefunction is a poor starting point but static correlation is not the main problem, CCSD is much preferred over QCISD. This is not uncommon in transition metal complexes.

Let us look at the case of the potential energy surface of the N_2 molecule. We study it with three different basis sets: TZVP, TZVPP and QZVP. The input is the following:

```
! TZVPP CCSD(T)
%paras R= 1.05,1.13,8
       end
* xyz 0 1
N 0 0 0
N 0 0 {R}
*
```

For even higher accuracy we would need to introduce relativistic effects and - in particular - turn the core correlation on.[2]

Table 6.2: Computed spectroscopic constants of N2 with coupled-cluster methods.

Method	Basis set	R_e (pm)	ω_e (cm^{-1})	$\omega_e x_e$ (cm^{-1})
CCSD(T)	SVP	111.2	2397	14.4
	TZVP	110.5	2354	14.9
	TZVPP	110.2	2349	14.1
	QZVP	110.0	2357	14.3
	ano-pVDZ	111.3	2320	14.9
	ano-pVTZ	110.5	2337	14.4
	ano-pVQZ	110.1	2351	14.5
CCSD	QZVP	109.3	2437	13.5
Exp		**109.7**	**2358.57**	**14.32**

One can see from Table 6.2 that for high accuracy - in particular for the vibrational frequency - one needs both - the connected triple-excitations and large basis sets (the TZVP result is fortuitously good). While this is an isolated example, the conclusion holds more generally. If one pushes it, CCSD(T) has an accuracy (for reasonably well-behaved systems) of approximately 0.2 pm in distances, <10 cm^{-1} for harmonic frequencies and a few kcal/mol for atomization energies.[3] It is also astonishing how well the Ahlrichs basis sets do in these calculations — even slightly better than the much more elaborate ANO bases.

Note

The quality of a given calculation is not always high because it carries the label "coupled-cluster". Accurate results are only obtained in conjunction with large basis sets and for systems where the HF approximation is a good 0^{th} order starting point.

[2] Note that core correlation is not simply introduced by including the core orbitals in the correlation problem. In addition, special correlation core-polarization functions are needed. They have been standardized for a few elements in the cc-pCVxZ (X=D,T,Q,5,6) basis sets.

[3] However, in recent years it became more evident that even CCSD(T) achieves its high apparent accuracy through error cancellations. The full CCSDT method (triples fully included) usually performs worse than CCSD(T). The reason is that the (T) correction undershoots the effects of the triples to some extent and thereby compensates for the neglect of connected quadruple excitations. For very high accuracy quantum chemistry, even these must be considered. The prospects for treating chemically more relevant molecules with such methods is not particularly bright for the foreseeable future…

Coupled-Cluster Densities

If one is mainly accustomed to Hartree-Fock or DFT calculations, the calculation of the density matrix is more or less a triviality and is automatically done together with the solution of the self-consistent field equations. Unfortunately, this is not the case in coupled-cluster theory (and also not in MP2 theory). The underlying reason is that in coupled-cluster theory, the expansion of the exponential $e^{\hat{T}}$ in the expectation value

$$D_{pq} = \frac{\langle \Psi | E_p^q | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\langle e^{\hat{T}} \Psi_0 | E_p^q | e^{\hat{T}} \Psi_0 \rangle}{\langle e^{\hat{T}} \Psi_0 | e^{\hat{T}} \Psi_0 \rangle}$$

only terminates if all possible excitation levels are exhausted, i.e., if all electrons in the reference determinant Ψ_0 (typically the HF determinant) are excited from the space of occupied to the space of virtual orbitals (here D_{pq} denotes the first order density matrix, E_p^q are the spin traced second quantized orbital replacement operators, and $\hat{T}$ is the cluster operator). Hence, the straightforward application of these equations is far too expensive. It is, however, possible to expand the exponentials and only keep the linear term. This then defines a linearized density which coincides with the density that one would calculate from linearized coupled-cluster theory (CEPA/0). The difference to the CEPA/0 density is that converged coupled-cluster amplitudes are used for its evaluation. This density is straightforward to compute and the computational effort for the evaluation is very low. Hence, this is a density that can be easily produced in a coupled-cluster run. It is not, however, what coupled-cluster aficionados would accept as a density.

The subject of a density in coupled-cluster theory is approached from the viewpoint of response theory. Imagine one adds a perturbation of the form

$$H^{(\lambda)} = \lambda \sum\nolimits_{pq} h_{pq}^{\lambda} E_p^q$$

to the Hamiltonian. Then it is always possible to cast the first derivative of the total energy in the form:

$$\frac{dE}{d\lambda} = \sum_{pq} D_{pq}^{\text{(response)}} h_{pq}^{\lambda}$$

This is a nice result. The quantity $D_{pq}^{\text{(response)}}$ is the so-called response density. In the case of CC theory where the energy is not obtained by variational optimization of an energy functional, the energy has to be replaced by a Lagrangian reading as follows:

$$L_{CC} = \langle \Phi_0 | \bar{H} | \Phi_0 \rangle + \sum_{\eta} \lambda_{\eta} \langle \Phi_{\eta} | \bar{H} | \Phi_0 \rangle + \sum_{ai} f_{ai} z_{ai}$$

Here $\langle \Phi_{\eta}|$ denotes any excited determinant (singly, doubly, triply, ...). There are two sets of Lagrange multipliers: the quantities z_{ai} that guarantee that the perturbed wavefunction fulfills the Hartree-Fock conditions by making the off-diagonal Fock matrix blocks zero and the quantities λ_{η} that guarantee that the coupled-cluster projection equations for the amplitudes are fulfilled. If both sets of conditions are fulfilled then the coupled-cluster Lagrangian simply evaluates to the coupled-cluster energy. The coupled-cluster Lagrangian can be made stationary with respect to the Lagrangian multipliers z_{ai} and λ_{η}. The response density is then defined through:

$$\frac{dL_{CC}}{d\lambda} = \sum_{pq} D_{pq}^{\text{(response)}} \, h_{pq}^{\lambda}$$

The density D_{pq} appearing in this equation does not have the same properties as the density that would arise from an expectation value. For example, the response density can have eigenvalues lower than 0 or larger than 2. In practice, the response density is, however, the best "density" there is for coupled-cluster theory.

Unfortunately, the calculation of the coupled-cluster response density is quite involved because additional sets of equations need to be solved in order to determine the z_{ai} and λ_{η}. If only the equations for λ_{η} are solved one speaks of an "unrelaxed" coupled-cluster density. If both sets of equations are solved, one speaks of a "relaxed" coupled-cluster density. For most intents and purposes, the orbital relaxation effects incorporated into the relaxed density are small for a coupled-cluster density. This is so, because the coupled-cluster equations contain the exponential of the single excitation operator $e^{\hat{T}_1} = \exp(\sum_{ai} t_a^i E_i^a)$. This brings in most of the effects of orbital relaxation. In fact, replacing the $\hat{T}_1$ operator by the operator $\hat{\kappa} = \sum_{ai} \kappa_a^i (E_i^a - E_a^i)$ would provide all of the orbital relaxation thus leading to "orbital optimized coupled-cluster theory" (OOCC).

Not surprisingly, the equations that determine the coefficients λ_η (the Lambda equations) are as complicated as the coupled-cluster amplitude equations themselves. Hence, the calculation of the unrelaxed coupled-cluster density matrix is about twice as expensive as the calculation of the coupled-cluster energy (but not quite as with proper program organization terms can be reused and the Lambda equations are linear equations that converge somewhat better than the non-linear amplitude equations).

ORCA features the calculation of the unrelaxed coupled-cluster density on the basis of the Lambda equations for closed- and open-shell systems. If a fully relaxed coupled-cluster density is desired then ORCA still features the orbital-optimized coupled-cluster doubles method (OOCCD). This is not exactly equivalent to the fully relaxed CCSD density matrix because of the operator $\hat{\kappa}$ instead of $\hat{T}_1$. However, results are very close and orbital optimized coupled-cluster doubles is the method of choice if orbital relaxation effects are presumed to be large.

In terms of ORCA keywords, the coupled-cluster density is obtained through the following keywords:

```
#
# coupled-cluster density
#
%mdci  density   none
                 linearized
                 unrelaxed
                 orbopt
       end
```

which will work together with CCSD or QCISD (QCISD and CCSD are identical in the case of OOCCD because of the absence of single excitations). Note, that an unrelaxed density for CCSD(T) is NOT available.

Instead of using the density option "orbopt" in the mdci-block, OOCCD can also be invoked by using the keyword:

```
! OOCCD
```

Static versus Dynamic Correlation

Having said that, let us look at an "abuse" of the single reference correlation methods by studying (very superficially) a system which is not well described by a single HF determinant. This already occurs for the twisting of the double bond of C_2H_4. At a 90° twist angle the system behaves like a diradical and should be described by a multireference method (see section *Complete Active Space Self-Consistent Field Method*)

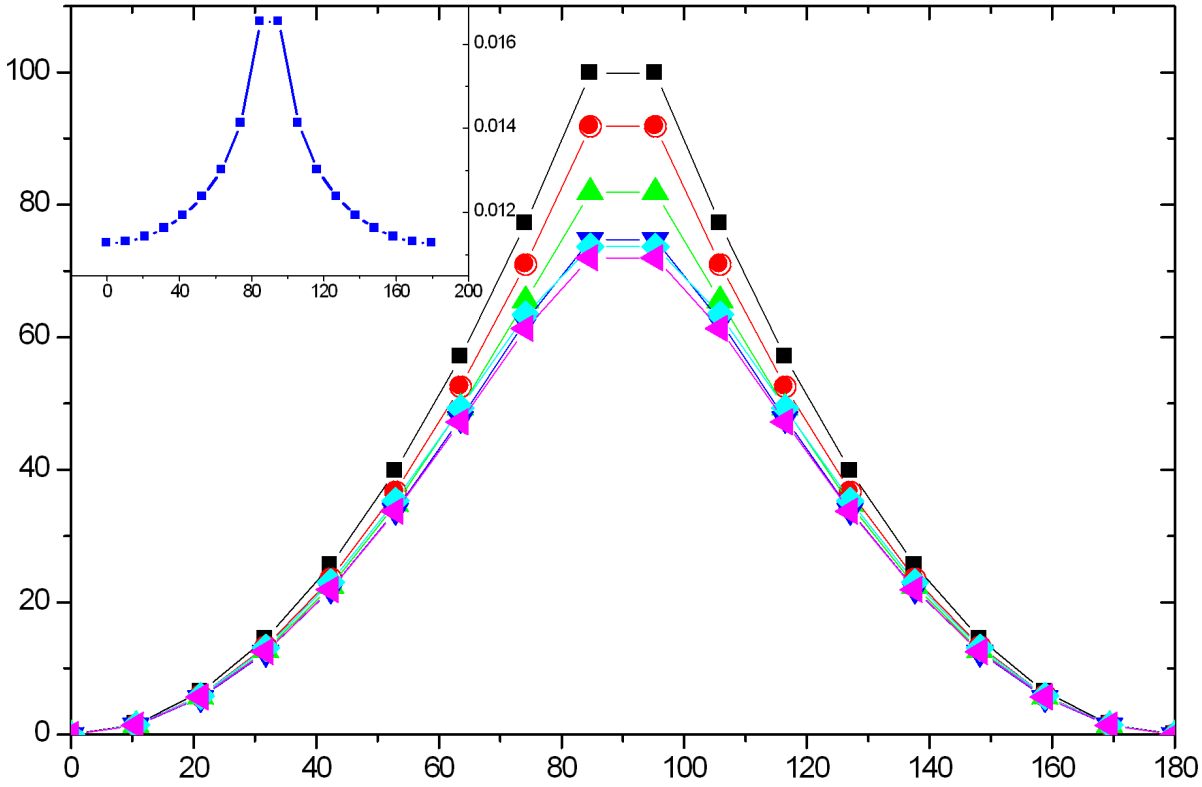


Fig. 6.1: A rigid scan along the twisting coordinate of C_2H_4. The inset shows the T_1 diagnostic for the CCSD calculation.

As can be seen in Fig. 6.1, there is a steep rise in energy as one approaches a 90° twist angle. The HF curve is actually discontinuous and has a cusp at 90°. This is immediately fixed by a simple CASSCF(2,2) calculation which gives a smooth potential energy surface. Dynamic correlation is treated on top of the CASSCF(2,2) method with the MRACPF approach as follows:

```
#
# twisting the double bond of C2H4
#
! SV(P) def2-TZVP/C SmallPrint NoPop MRACPF
%casscf nel         2
        norb        2
        mult        1
        nroots      1
        TrafoStep  RI
        end
%mrci   tsel 1e-10
        tpre 1e-10
        end
%method scanguess pmodel
        end
%paras  R= 1.3385
        Alpha=0,180,18
        end
* int 0 1
C  0  0  0  0      0   0
C  1  0  0 {R}     0   0
H  1  2  0 1.07  120   0
H  1  2  3 1.07  120 180
H  2  1  3 1.07  120 {Alpha}
H  2  1  3 1.07  120 {Alpha+180}
*
```

This is the reference calculation for this problem. One can see that the RHF curve is far from the MRACPF reference but the CASSCF calculation is very close. Thus, dynamic correlation is not important for this problem! It only appears to be important since the RHF determinant is such a poor choice. The MP2 correlation energy is insufficient in order to repair the RHF result. The CCSD method is better but still falls short of quantitative accuracy. Finally, the CCSD(T) curve is very close the MRACPF. This even holds for the total energy (inset of Fig. 6.2) which does not deviate by more than 2–3 mEh from each other. Thus, in this case one uses the powerful CCSD(T) method in an inappropriate way in order to describe a system that has multireference character. Nevertheless, the success of CCSD(T) shows how stable this method is even in tricky situations. The "alarm" bell for CCSD and CCSD(T) is the so-called "T_1-diagnostic"[4] that is also shown in Fig. 6.2. A rule of thumb says, that for a value of the diagnostic of larger than 0.02 the results are not to be trusted. In this calculation we have not quite reached this critical point although the T_1 diagnostic blows up around the 90° twist.

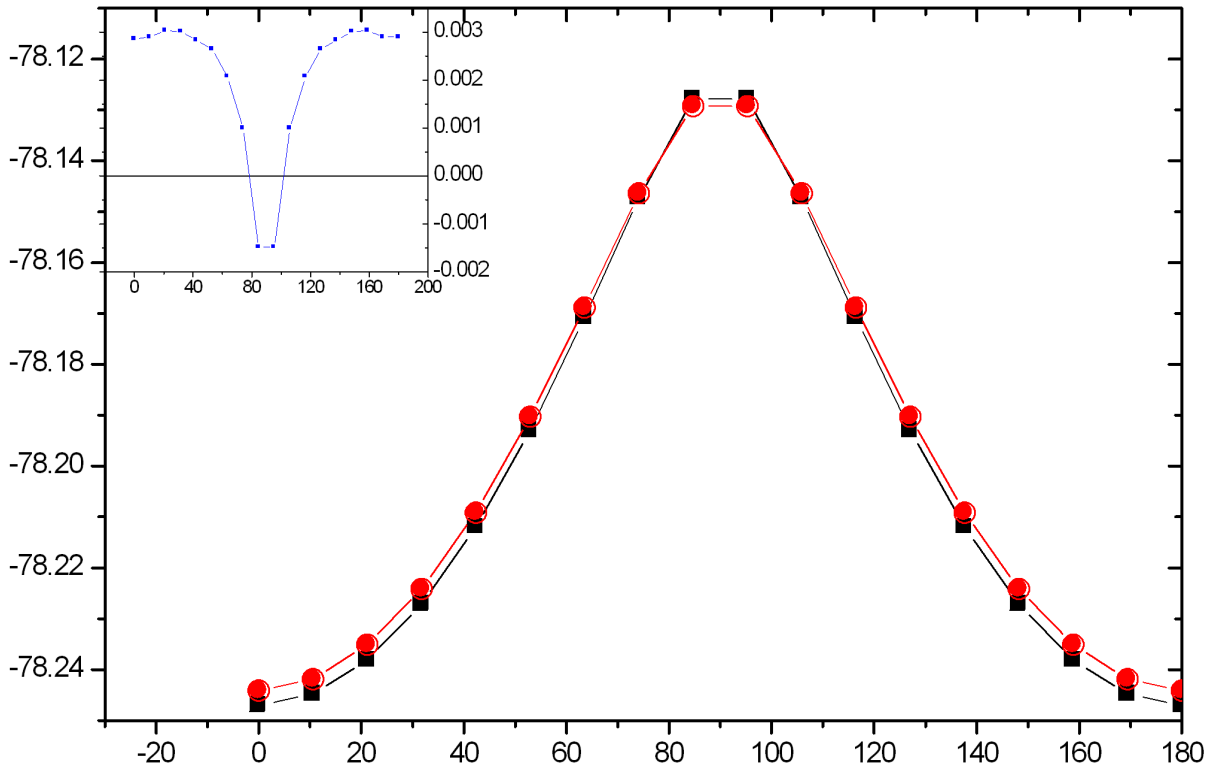


Fig. 6.2: Comparison of the CCSD(T) and MRACPF total energies of the C_2H_4 along the twisting coordinate. The inset shows the difference E(MRACPF)-E(CCSD(T)).

The computational cost (disregarding the triples) is such that the CCSD method is the most expensive followed by QCISD (∼10% cheaper) and all other methods (about 50% to a factor of two cheaper than CCSD). The most accurate method is generally CCSD(T). However, this is not so clear if the triples are omitted and in this regime the coupled pair methods (in particular CPF/1 and NCPF/1[5]) can compete with CCSD.

Let us look at the same type of situation from a slightly different perspective and dissociate the single bond of F_2. As is well known, the RHF approximation fails completely for this molecule and predicts it to be unbound. Again we use a much too small basis set for quantitative results but it is enough to illustrate the principle.

We first generate a "reference" PES with the MRACPF method:

```
! def2-SV def2-SVP/C MRACPF
%casscf nel    2
        norb   2
```

(continues on next page)

[4] It is defined as $\|T_1\| / N^{1/2}$ where T_1 are the singles amplitudes and N the number of correlated electrons. The original reference is [504]

[5] The "N" methods have been suggested by [885] and are exclusive to ORCA. Please note that our NCPF/1 is different from the MCPF method in the literature [173]. The original CPF method — which we prefer — is from [16]; see also [15] for a nice review about the coupled pair approaches and the underlying philosophy.

(continued from previous page)

```
        nroots 1
        mult   1
        end
%mrci   tsel    1e-10
        tpre    1e-10
        end
%paras  R= 3.0,1.3,35
        end
* xyz 0 1
F 0 0 0
F 0 0 {R}
*
```

Note that we scan from outward to inward. This helps the program to find the correct potential energy surface since at large distances the σ and σ^* orbitals are close in energy and fall within the desired 2×2 window for the CASSCF calculation (see section *Complete Active Space Self-Consistent Field Method*). Comparing the MRACPF and CASSCF curves it becomes evident that the dynamic correlation brought in by the MRACPF procedure is very important and changes the asymptote (loosely speaking the binding energy) by almost a factor of two (see Fig. 6.3). Around the minimum (roughly up to 2.0 Å) the CCSD(T) and MRACPF curves agree beautifully and are almost indistinguishable. Beyond this distance the CCSD(T) calculation begins to diverge and shows an unphysical behavior while the multireference method is able to describe the entire PES up to the dissociation limit. The CCSD curve is qualitatively OK but has pronounced quantitative shortcomings: it predicts a minimum that is much too short and a dissociation energy that is much too high. Thus, already for this rather "simple" molecule, the effect of the connected triple excitations is very important. Given this (rather unpleasant) situation, the behavior of the much simpler CEPA method is rather satisfying since it predicts a minimum and dissociation energy that is much closer to the reference MRACPF result than CCSD or CASSCF. It appears that in this particular case CEPA/1 and CEPA/2 predict the correct result.

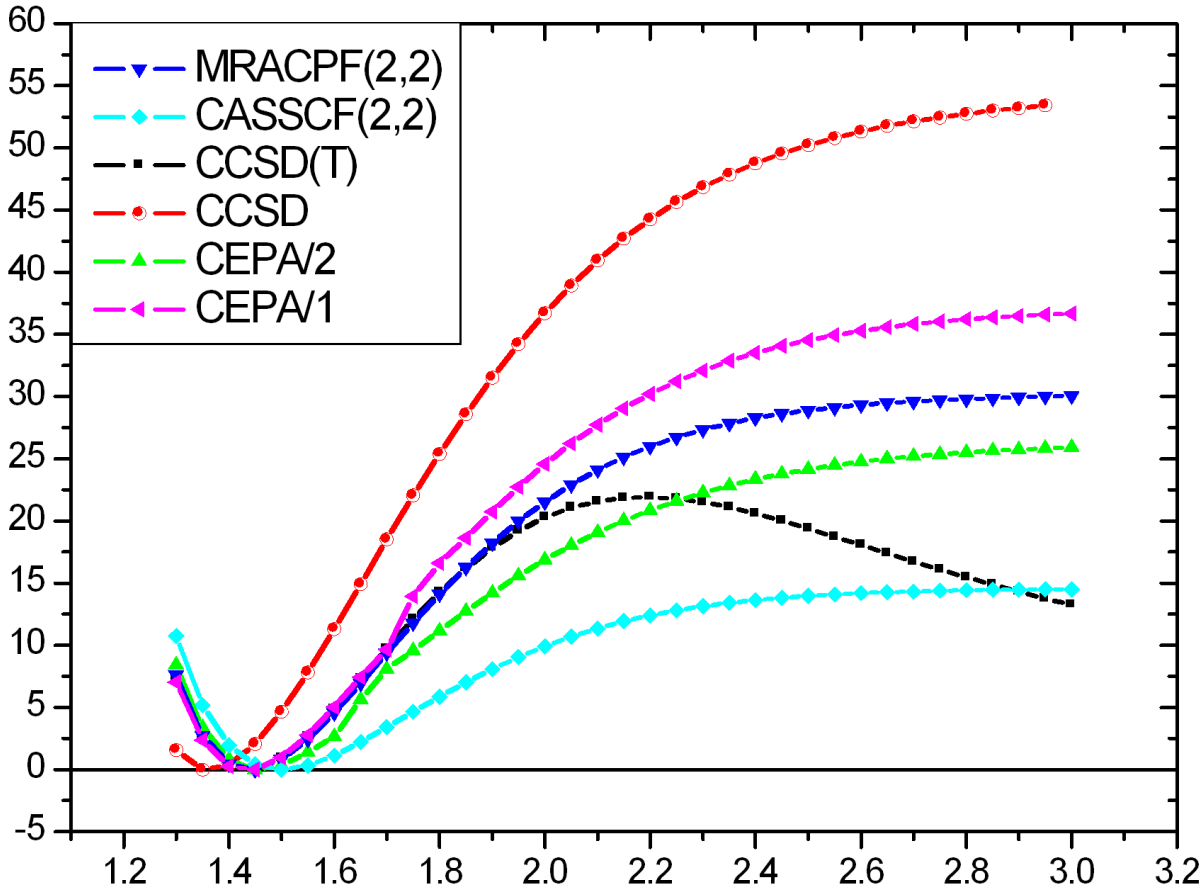


Fig. 6.3: Potential energy surface of the F_2 molecule calculated with some single-reference methods and compared to the MRACPF reference.

Basis Sets for Correlated Calculations. The case of ANOs.

In HF and DFT calculations the generation and digestion of the two-electron repulsion integrals is usually the most expensive step of the entire calculation. Therefore, the most efficient approach is to use loosely contracted basis sets with as few primitives as possible — the Ahlrichs basis sets (SVP, TZVP, TZVPP, QZVP, def2-TZVPP, def2-QZVPP) are probably the best in this respect. Alternatively, the polarization-consistent basis sets pc-1 through pc-4 could be used, but they are only available for H-Ar. For large molecules such basis sets also lead to efficient prescreening and consequently efficient calculations.

This situation is different in highly correlated calculations such as CCSD and CCSD(T) where the effort scales steeply with the number of basis functions. In addition, the calculations are usually only feasible for a limited number of basis functions and are often run in the integral conventional mode, since high angular momentum basis functions are present and these are expensive to recompute all the time. Hence, a different strategy concerning the basis set design seems logical. It would be good to use as few basis functions as possible but make them as accurate as possible. This is compatible with the philosophy of atomic natural orbital (ANO) basis sets. Such basis sets are generated from correlated atomic calculations and replicate the primitives of a given angular momentum for each basis function. Therefore, these basis sets are deeply contracted and expensive but the natural atomic orbitals form a beautiful basis for molecular calculations. In ORCA an accurate and systematic set of ANOs (ano-pVnZ, n = D, T, Q, 5 is incorporated). A related strategy underlies the design of the correlation-consistent basis sets (cc-pVnZ, n = D, T, Q, 5, 6,...) that are also generally contracted except for the outermost primitives of the "principal" orbitals and the polarization functions that are left uncontracted.

Let us study this subject in some detail using the H_2CO molecule at a standard geometry and compute the SCF and correlation energies with various basis sets. In judging the results one should view the total energy in conjunction with the number of basis functions and the total time elapsed. Looking at the data in the Table below, it is obvious that the by far lowest SCF energies for a given cardinal number (2 for double-zeta, 3 for triple zeta

and 4 for quadruple-zeta) are provided by the ANO basis sets. Using specially optimized ANO integrals that are available since ORCA 2.7.0, the calculations are not even much more expensive than those with standard basis sets. Obviously, the correlation energies delivered by the ANO bases are also the best of all 12 basis sets tested. Hence, ANO basis sets are a very good choice for highly correlated calculations. The advantages are particularly large for the early members (DZ/TZ).

Table 6.3: Comparison of various basis sets for highly correlated calculations

Basis set	No. Basis Fcns	E(SCF)	E_C(CCSD(T))	E_{tot}(CCSD(T))	Total Time
cc-pVDZ	38	-113.876184	-0.34117952	-114.217364	2
cc-pVTZ	88	-113.911871	-0.42135475	-114.333226	40
cc-pVQZ	170	-113.920926	-0.44760332	-114.368529	695
def2-SVP	38	-113.778427	-0.34056109	-114.118988	2
def2-TZVPP	90	-113.917271	-0.41990287	-114.337174	46
def2-QZVPP	174	-113.922738	-0.44643753	-114.369175	730
pc-1	38	-113.840092	-0.33918253	-114.179274	2
pc-2	88	-113.914256	-0.41321906	-114.327475	43
pc-3	196	-113.922543	-0.44911659	-114.371660	1176
ano-pVDZ	38	-113.910571	-0.35822337	-114.268795	12
ano-pVTZ	88	-113.920389	-0.42772994	-114.348119	113
ano-pVQZ	170	-113.922788	-0.44995355	-114.372742	960

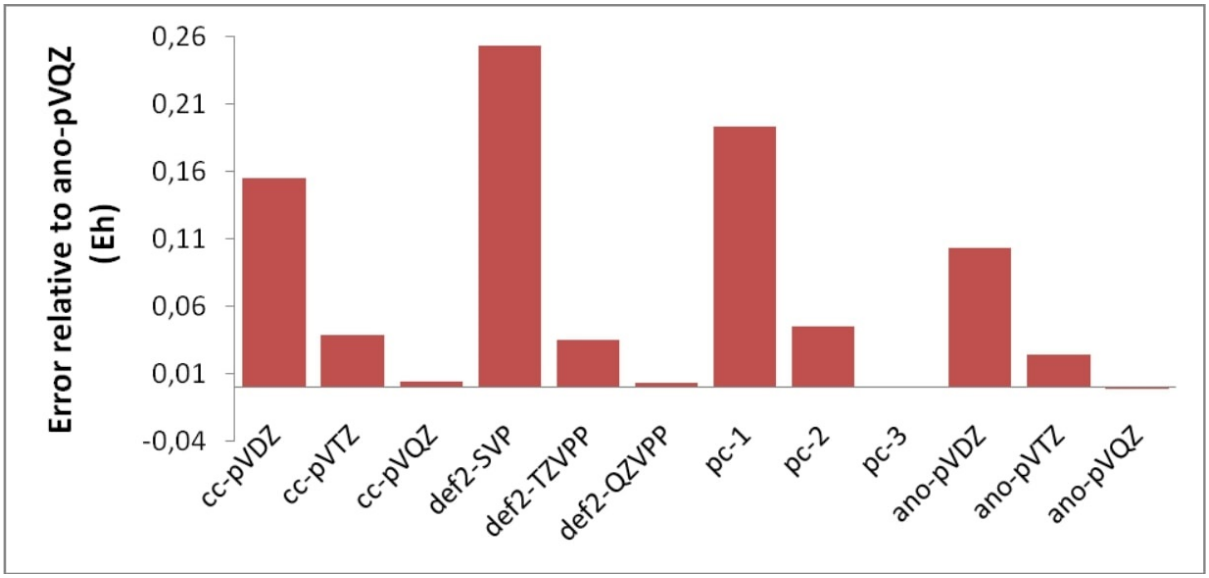


Fig. 6.4: Error in Eh for various basis sets for highly correlated calculations relative to the ano-pVQZ basis set.

Let us look at one more example in Table 6.4: the optimized structure of the N_2 molecule as a function of basis set using the MP2 method *(these calculations are a bit older from the time when the ano-pVnZ basis sets did not yet exist. Today, the ano-pVnZ would be preferred)* .

The highest quality basis set here is QZVP and it also gives the lowest total energy. However, this basis set contains up to g-functions and is very expensive. Not using g-functions and a set of f-functions (as in TZVPP) has a noticeable effect on the outcome of the calculations and leads to an overestimation of the bond distance of 0.2 pm — a small change but for benchmark calculations of this kind still significant. The error made by the TZVP basis set that lacks the second set of d-functions on the bond distance, binding energy and ionization potential is surprisingly small even though the deletion of the second d-set "costs" more than 20 mEh in the total energy as compared to TZV(2d,2p), and even more compared to the larger TZVPP.

A significant error on the order of 1 – 2 pm in the calculated distances is produced by smaller DZP type basis sets, which underlines once more that such basis sets are really too small for correlated molecular calculations — the ANO-DZP basis sets are too strongly biased towards the atom, while the "usual" molecule targeted DZP basis sets

like SVP have the d-set designed to cover polarization but not correlation (the correlating d-functions are steeper than the polarizing ones). The performance of the very economical SVP basis set should be considered as very good, and (a bit surprisingly) slightly better than cc-pVDZ despite that it gives a higher absolute energy.

Essentially the same picture is obtained by looking at the (uncorrected for ZPE) binding energy calculated at the MP2 level – the largest basis set, QZVP, gives the largest binding energy while the smaller basis sets underestimate it. The error of the DZP type basis sets is fairly large ($\approx$ 2 eV) and therefore caution is advisable when using such bases.

Table 6.4: Comparison of various basis sets for correlated calculations.

Basis set	R_{eq} (pm)	E(2N-N_2) (eV)	IP(N/N^+) (eV)	E(MP2) (Eh)
SVP	112.2	9.67	14.45	**-109.1677**
cc-pVDZ	112.9	9.35	14.35	**-109.2672**
TZVP	111.5	10.41	14.37	**-109.3423**
TZV(2d,2p)	111.4	10.61	14.49	**-109.3683**
TZVPP	111.1	10.94	14.56	**-109.3973**
QZVP	110.9	11.52	14.60	**-109.4389**

Automatic extrapolation to the basis set limit

Note

- This functionality is deprecated - it may still be usable but we will not actively maintain this part of code anymore. For basis set extrapolation please use the respective compound scripts (Table *Protocols, known to the simple input line, with short explanation*).

As eluded to in the previous section, one of the biggest problems with correlation calculations is the slow convergence to the basis set limit. One possibility to overcome this problem is the use of explicitly correlated methods. The other possibility is to use basis set extrapolation techniques. Since this involves some fairly repetitive work, some procedures were hardwired into the ORCA program. So far, only energies are supported. For extrapolation, a systematic series of basis sets is required. This is, for example, provided by the cc-pVnZ, aug-cc-pVnZ or the corresponding ANO basis sets. Here n is the "cardinal number" that is 2 for the double-zeta basis sets, 3 for triple-zeta, etc.

The convergence of the HF energy to the basis set limit is assumed to be given by:

$$E_{\mathrm{SCF}}^{(X)} = E_{\mathrm{SCF}}^{(\infty)} + A \exp\left(-\alpha\sqrt{X}\right) \tag{6.1}$$

Here, $E_{\mathrm{SCF}}^{(X)}$ is the SCF energy calculated with the basis set with cardinal number X, $E_{\mathrm{SCF}}^{(\infty)}$ is the basis set limit SCF energy and A and α are constants. The approach taken in ORCA is to do a two-point extrapolation. This means that either A or α have to be known. Here, we take A as to be determined and α as a basis set specific constant.

The correlation energy is supposed to converge as:

$$E_{\mathrm{corr}}^{(\infty)} = \frac{X^\beta E_{\mathrm{corr}}^{(X)} - Y^\beta E_{\mathrm{corr}}^{(Y)}}{X^\beta - Y^\beta} \tag{6.2}$$

The theoretical value for β is 3.0. However, it was found by Truhlar and confirmed by us, that for 2/3 extrapolations $\beta = 2.4$ performs considerably better.

For a number of basis sets, we have determined the optimum values for α and β[606]:

	α_{23}	β_{23}	α_{34}	β_{34}
cc-pV*n*Z	4.42	2.46	5.46	3.05
pc-*n*	7.02	2.01	9.78	4.09
def2	10.39	2.40	7.88	2.97
ano-pV*n*Z	5.41	2.43	4.48	2.97
saug-ano-pV*n*Z	5.48	2.21	4.18	2.83
aug-ano-pV*n*Z	5.12	2.41		

Since the β values for 2/3 are close to 2.4, we always take this value. Likewise, all 3/4 and higher extrapolations are done with $\beta = 3$. However, the optimized values for α are taken throughout.

Using the keyword `! Extrapolate(X/Y,basis)`, where `X` and `Y` are the corresponding successive cardinal numbers and `basis` is the type of basis set requested (= `cc`, `aug-cc`, `cc-core`, `ano`, `saug-ano`, `aug-ano`, `def2`) ORCA will calculate the SCF and optionally the MP2 or MDCI energies with two basis sets and separately extrapolate.

The keyword works also in the following way: `! Extrapolate(n,basis)` where `n` is the is the number of energies to be used. In this way the program will start from a double-zeta basis and perform calculations with `n` cardinal numbers and then extrapolate the different pairs of basis sets. Thus for example the keyword `! Extrapolate(3, CC)` will perform calculations with cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets and then estimate the extrapolation results of both cc-pVDZ/cc-pVTZ and cc-pVTZ/cc-pVQZ combinations.

Let us take the example of the H2O molecule at the B3LYP/TZVP optimized geometry. The reference values have been determined from a HF calculation with the decontracted aug-cc-pV6Z basis set and the correlation energy was obtained from the cc-pV5Z/cc-pV6Z extrapolation. This gives:

```
E(SCF,CBS)           = -76.066958 Eh
EC(CCSD(T),CBS)      =  -0.30866 Eh
Etot(CCSD(T),CBS)    = -76.37561 Eh
```

Now we can see what extrapolation can bring in:

```
!CCSD(T) Extrapolate(2/3) TightSCF Conv Bohrs
* int 0 1
O 0 0 0  0 0 0
H 1 0 0  1.81975 0 0
H 1 2 0  1.81975 105.237 0
*
```

NOTE:

- The RI-JK and RIJCOSX approximations work well together with this option and RI-MP2 is also possible. Auxiliary basis sets are automatically chosen and can not be changed.
- All other basis set choices, externally defined bases etc. will be ignored — the automatic procedure only works with the default basis sets!
- The basis sets with the "core" postfix contain core correlation functions. By default it is assumed that this means that the core electrons are also to be correlated and the frozen core approximation is turned off. However, this can be overridden in the method block by choosing, e.g. `%method frozencore fc_electrons end`!
- So far, the extrapolation is only implemented for single points and not for gradients. Hence, geometry optimizations cannot be done in this way.
- The extrapolation method should only be used with very tight SCF convergence criteria. For open shell methods, additional caution is advised.

This gives:

```
Alpha(2/3)     :  4.420 (SCF Extrapolation)
Beta(2/3)      :  2.460 (correlation extrapolation)

SCF energy with basis cc-pVDZ:                          -76.026430944
SCF energy with basis cc-pVTZ:                          -76.056728252
Extrapolated CBS SCF energy (2/3) :                     -76.066581429 (-0.009853177)

MDCI energy with basis cc-pVDZ:                          -0.214591061
MDCI energy with basis cc-pVTZ:                          -0.275383015
Extrapolated CBS correlation energy (2/3) :              -0.310905962 (-0.035522947)

Estimated CBS total energy (2/3) :                      -76.377487391
```

Thus, the error in the total energy is indeed strongly reduced. Let us look at the more rigorous 3/4 extrapolation:

```
Alpha(3/4)     :  5.460 (SCF Extrapolation)
Beta(3/4)      :  3.050 (correlation extrapolation)

SCF energy with basis cc-pVTZ:                          -76.056728252
SCF energy with basis cc-pVQZ:                          -76.064381269
Extrapolated CBS SCF energy (3/4) :                     -76.066687152 (-0.002305884)

MDCI energy with basis cc-pVTZ:                          -0.275383016
MDCI energy with basis cc-pVQZ:                          -0.295324345
Extrapolated CBS correlation energy (3/4) :              -0.309520368 (-0.014196023)

Estimated CBS total energy (3/4) :                      -76.376207520
```

In our experience, the ANO basis sets extrapolate similarly to the cc-basis sets. Hence, repeating the entire calculation with `Extrapolate(3,ANO)` gives:

```
Estimated CBS total energy (2/3) :                      -76.377652792
Estimated CBS total energy (3/4) :                      -76.376983433
```

Which is within 1 mEh of the estimated CCSD(T) basis set limit energy in the case of the 3/4 extrapolation and within 2 mEh for the 2/3 extrapolation.

For larger molecules, the bottleneck of the calculation will be the CCSD(T) calculation with the larger basis set. In order to avoid this expensive (or prohibitive) calculation, it is possible to estimate the CCSD(T) energy at the basis set limit as:

$$E_{\mathrm{corr}}^{(\mathrm{CCSD(T)};Y)} \approx E_{\mathrm{corr}}^{(\mathrm{CCSD(T)};X)} + E_{\mathrm{corr}}^{(\mathrm{MP2};\infty)} - E_{\mathrm{corr}}^{(\mathrm{MP2};X)} \tag{6.3}$$

This assumes that the basis set dependence of MP2 and CCSD(T) is similar. One can then extrapolate as before. Alternatively, the standard way — as extensively exercised by Hobza and co-workers — is to simply use:

$$E_{\mathrm{total}}^{(\mathrm{CCSD(T)};\mathrm{CBS})} \approx E_{\mathrm{SCF}}^{(Y)} + E_{\mathrm{corr}}^{(\mathrm{CCSD(T)};X)} + E_{\mathrm{corr}}^{(\mathrm{MP2};\infty)} - E_{\mathrm{corr}}^{(\mathrm{MP2};X)} \tag{6.4}$$

The appropriate keyword is:

```
! ExtrapolateEP2(2/3,ANO,MP2) TightSCF Conv Bohrs
* int 0 1
O 0 0 0  0 0 0
H 1 0 0  1.81975 0 0
H 1 2 0  1.81975 105.237 0
*
```

This creates the following output:

```
Alpha      :  5.410 (SCF Extrapolation)
Beta       :  2.430 (correlation extrapolation)
```

(continues on next page)

(continued from previous page)

```
SCF energy with basis ano-pVDZ:                          -76.059178452
SCF energy with basis ano-pVTZ:                          -76.064774379
Extrapolated CBS SCF energy :                            -76.065995735 (-0.001221356)

MP2 energy with basis ano-pVDZ:                           -0.219202871
MP2 energy with basis ano-pVTZ:                           -0.267058634
Extrapolated CBS correlation energy :                     -0.295568604 (-0.028509970)

CCSD(T) correlation energy with basis ano-pVDZ:           -0.229478341
CCSD(T) - MP2 energy with basis ano-pVDZ:                 -0.010275470

Estimated CBS total energy :                             -76.371839809
```

The estimated correlation energy is not really bad — within 3 mEh from the basis set limit.

Using the `ExtrapolateEP2(n/m,bas,[method, method-details])` keyword one can use a generalization of the above method where instead of MP2 any available correlation method can be used as described in Ref. [519]. `method` is optional and can be either MP2 or DLPNO-CCSD(T), the latter being the default. In case the method is DLPNO-CCSD(T) in the `method-details` option one can ask for LoosePNO, NormalPNO or TightPNO.

$$E_{\rm corr}^{({\rm CCSD(T)};CBS)} \approx E_{\rm corr}^{({\rm CCSD(T)};X)} + E_{\rm corr}^{({\rm M};CBS)}(X, X+1) - E_{\rm corr}^{({\rm M};X)} \tag{6.5}$$

Here M represents any correlation method one would like to use. For the previous water molecule the input of a calculation that uses DLPNO-CCSD(T) (which is the default now) instead of MP2 would look like:

```
! ExtrapolateEP2(2/3,cc,DLPNO-CCSD(T)) TightSCF Conv Bohrs
* int 0 1
O 0 0 0  0 0 0
H 1 0 0  1.81975 0 0
H 1 2 0  1.81975 105.237 0
*
```

and it would produce the following output:

```
Alpha     :  4.420 (SCF Extrapolation)
Beta      :  2.460 (correlation extrapolation)

SCF energy with basis cc-pVDZ:                           -76.026430944
SCF energy with basis cc-pVTZ:                           -76.056728252
Extrapolated CBS SCF energy :                            -76.066581429 (-0.009853177)

MDCI energy with basis cc-pVDZ:                           -0.214429497
MDCI energy with basis cc-pVTZ:                           -0.275299699
Extrapolated CBS correlation energy :                     -0.310868368 (-0.035568670)

CCSD(T) correlation energy with basis cc-pVDZ:            -0.214548320
CCSD(T) - MDCI energy with basis cc-pVDZ:                 -0.000118824

Estimated CBS total energy :                             -76.377568621
```

which is less than 2 mEh from the basis set limit. Finally it was shown [519] that instead of extrapolating the cheap method, M, using cardinal numbers X and $X+1$ it is better to use cardinal numbers $X+1$ and $X+2$.

$$E_{\rm corr}^{({\rm CCSD(T)};CBS)} \approx E_{\rm corr}^{({\rm CCSD(T)};X)} + E_{\rm corr}^{({\rm M};CBS)}(X+1, X+2) - E_{\rm corr}^{({\rm M};X)} \tag{6.6}$$

This can be done using the `ExtrapolateEP3(bas,[method,method-details])` keyword:

```
! ExtrapolateEP3(CC) TightSCF Conv Bohrs
```

and the corresponding output would be:

```
Alpha     :  5.460 (SCF Extrapolation)
Beta      :  3.050 (correlation extrapolation)

SCF energy with basis cc-pVDZ:                          -76.026430944
SCF energy with basis cc-pVTZ:                          -76.056728252
SCF energy with basis cc-pVQZ:                          -76.064381269
Extrapolated CBS SCF energy :                           -76.066687152 (-0.002305884)

MDCI energy with basis cc-pVDZ:                          -0.214429497
MDCI energy with basis cc-pVTZ:                          -0.275299699
MDCI energy with basis cc-pVQZ:                          -0.295229871
Extrapolated CBS correlation energy :                    -0.309417951 (-0.014188080)

CCSD(T) correlation energy with basis cc-pVDZ:           -0.214548319
CCSD(T) - MDCI energy with basis cc-pVDZ:                -0.000118822

Estimated CBS total energy :                            -76.376223926
```

For the ExtrapolateEP2, and ExtrapolateEP3 keywords the default cheap method is the DLPNO-CCSD(T) with the NormalPNO thresholds. There also available options with MP2, and DLPNO-CCSD(T) with LoosePNO and TightPNO settings.

Explicitly Correlated MP2 and CCSD(T) Calculations

A physically perhaps somewhat more satisfying alternative to basis set extrapolation is the theory of explicit correlation. In this method terms are added to the wavefunction Ansatz that contain the interelectronic coordinates explicitly (hence the name "explicit correlation"). Initially these terms were linear in the interelectronic distances ("R12-methods"). However, it has later been found that better results can be obtained by using other functions, such as an exponential, of the interelectronic distance ("F12-methods"). These methods are known to yield near basis set limit results for correlation energies in conjunction with much smaller orbital basis sets.

In applying these methods several points are important:

- Special orbital basis sets are at least advantageous. The development of such basis sets is still in its infancy. For a restricted range of elements the basis sets cc-pVnZ-F12 are available (where n = D, T, Q) and are recommended. Note, that other than their names suggest, these are a fair bit larger than regular double, triple or quadruple-zeta basis sets
- In addition to an orbital basis set, a near-complete auxiliary basis set must be specified. This is the so-called "CABS" basis. For the three basis sets mentioned above these are called cc-pVnZ-F12-CABS. If you have elements that are not covered you are on your own to supply a CABS basis set. CABS basis sets can be read into ORCA in a way analogous to RI auxiliary basis sets (replace "AUX" by "CABS" in the input). There are automatic tools for building a CABS basis from an arbitrary orbital basis, e.g. AutoCABS[780]
- if the RI approximation is used in conjunction with F12, a third basis set is required - this can be the regular auxiliary "/C" basis, but we recommend to step one level up in the auxiliary basis set (e.g. use a cc-pVTZ/C fitting basis in conjunction with cc-pVDZ-F12)
- It is perfectly feasible to use RIJCOSX or RI-JK at the same time. In this case, you should provide a fourth basis set for the Coulomb fitting
- RHF and UHF are available, ROHF not. (Although, one can do a ROHF like calculation by using QROs)
- Gradients are not available

Doing explicitly correlated MP2 calculations is straightforward. For example look at the following calculation on the water molecule at a given geometry:

```
#
! F12-MP2 cc-pVDZ-F12 cc-pVDZ-F12-CABS  VeryTightSCF PModel

* xyz 0 1
```

(continues on next page)

(continued from previous page)

```
    O    0.000000000000     0.000000000000     0.369372944000
    H    0.783975899000     0.000000000000    -0.184686472000
    H   -0.783975899000     0.000000000000    -0.184686472000
*
```

and similarly in conjunction with the RI approximation:

```
#
! F12-RI-MP2 cc-pVDZ-F12 cc-pVDZ-F12-CABS cc-pVTZ/C VeryTightSCF PModel

* xyz 0 1
    O    0.000000000000     0.000000000000     0.369372944000
    H    0.783975899000     0.000000000000    -0.184686472000
    H   -0.783975899000     0.000000000000    -0.184686472000
*
```

The output is relatively easy to interpret:

```
-----------------
RI-MP2-F12 ENERGY
-----------------

EMP2 correlation Energy            :      -0.241038994909
F12 correction                     :      -0.054735459470
                                         -----------------
MP2 basis set limit estimate       :      -0.295774454379

Hartree-Fock energy                :     -76.057963800414
(2)_S CABS correction to EHF       :      -0.003475342535
                                         -----------------
HF basis set limit estimate        :     -76.061439142949

MP2 total energy before F12        :     -76.299002795323
Total F12 correction               :      -0.058210802005
                                         -----------------
Final basis set limit MP2 estimate :     -76.357213597328
```

It consists of several parts. The first is the regular (RI-)MP2 correlation energy in the orbitals basis followed by the additive MP2 correction which are combined to provide an MP2 correlation energy basis set limit estimate. The second part consists of an estimate in the error in the underlying SCF energy. This is the "(2)_S CABS" correction. The combination of the SCF energy with this correction yields an estimate of the SCF basis set limit. The correction will typically undershoot somewhat, but the error is very smooth. Finally, the corrected correlation energy and the corrected SCF energy are added to yield the F12 total energy estimate at the basis set limit.

Let's look at some results and compare to extrapolation:

```
#
# Correlation energies of the water molecule: extrapolation versus F12
#
# cc-pVDZ MP2: -0.201380894
#      T     : -0.261263141
#      Q     : -0.282661311
#     T/Q    : -0.298276192
#     Q/5    : -0.300598282
#  F12-DZ    : -0.295775804
#  RI-F12-DZ : -0.295933560 (cc-pVDZ/C)
#              -0.295774489 (cc-pVTZ/C)
#  F12-TZ    : -0.299164006
#  RI-F12-TZ : -0.299163478 (cc-pVQZ/C)
#  F12-QZ    : -0.300130086
```

It is obvious that extrapolated and F12 correlation energies converge to the same number (in this case around 300

mEh). The best extrapolated result is still below the F12 result (this would primarily be meaningful in a variational calculation). However, first of all this was an expensive extrapolation and second, the small residual F12 error is very smooth and cancels in energy differences. In any case, already the F12-double-zeta (where "double zeta" is to be interpreted rather loosely) brings one into within 5 mEh of the basis set limit correlation energy and the F12-triple-zeta calculation to within 1 mEh, which is impressive.

The additional effort for the F12 calculation is rather high, since five types of additional two-electron integrals need to be calculated. Both integrals in CABS space and in the original orbital (OBS) space must be calculated and mixed Fock matrices are also required. Hence, one may wonder, whether a double-zeta F12 calculation actually saves any time over, say, a quadruple-zeta regular calculation. The actual answer to this question is: "NO". Given all possibilities of obtained approximate MP2 and SCF energies, we have investigated the question of how to obtain MP2 basis set limit energies most efficiently in some detail. The results show that in terms of timings, basis set extrapolation in combination with RI-JK is the method of choice for MP2.[521] However, energy differences are more reliable with F12-MP2. In combination with RI-JK or RIJCOSX F12-MP2 becomes also competitive in terms of computational efficiency.

This situation is different in the case of coupled-cluster methods, where F12 methods outperform extrapolation and are the method of choice.

For coupled-cluster theory, everything works in a very similar fashion:

```
# the keywords
! F12-CCSD(T)
# and
! CCSD(T)-F12
# are equivalent
```

A special feature of ORCA that can save large amounts of time, is to use the RI approximation only for the F12-part. The keyword here is:

```
! F12/RI-CCSD(T)
# or
! CCSD(T)-F12/RI
```

Everything else works as described for F12-MP2.

Frozen Core Options

In coupled-cluster calculations the Frozen Core (FC) approximation is applied by default. This implies that the core electrons are not included in the correlation treatment, since the inclusion of dynamic correlation in the core electrons usually affects relative energies insignificantly.

The frozen core option can be switched on or off with `! FrozenCore` or `! NoFrozenCore` in the simple input. More information and further options are given in section *Frozen Core Options* and in section *Including (semi)core orbitals in the correlation treatment*.

Local Coupled Pair and Coupled-Cluster Calculations

ORCA features a special set of local correlation methods. The prevalent local coupled-cluster approaches date back to ideas of Pulay and have been extensively developed by Werner, Schütz and co-workers. They use the concept of correlation domains in order to achieve linear scaling with respect to CPU, disk and main memory. While the central concept of electron pairs is very similar in both approaches, the local correlation methods in ORCA follow a completely different and original philosophy.

In ORCA rather than trying to use sparsity, we exploit data compression. To this end two concepts are used: (a) localization of internal orbitals, which reduces the number of electron pairs to be correlated since the pair correlation energies are known to fall off sharply with distance; (b) use of a truncated pair specific natural orbital basis to span the significant part of the virtual space for each electron pair. This guarantees the fastest convergence of the pair wavefunction and a nearly optimal convergence of the pair correlation energy while not introducing any real space cut-offs or geometrically defined domains. These PNOs have been used previously by the pioneers of

correlation theory. However, as discussed in the original papers, the way in which they have been implemented into ORCA is very different. For a full description of technical details and numerical tests see:

- F. Neese, A. Hansen, D. G. Liakos: Efficient and accurate local approximations to the coupled-cluster singles and doubles method using a truncated pair natural orbital basis.[616]
- F. Neese, A. Hansen, F. Wennmohs, S. Grimme: Accurate Theoretical Chemistry with Coupled Electron Pair Models.[617]
- F. Neese, F. Wennmohs, A. Hansen: Efficient and accurate local approximations to coupled electron pair approaches. An attempt to revive the pair-natural orbital method.[622]
- D. G. Liakos, A. Hansen, F. Neese: Weak molecular interactions studied with parallel implementations of the local pair natural orbital coupled pair and coupled-cluster methods.[520]
- A. Hansen, D. G. Liakos, F. Neese: Efficient and accurate local single reference correlation methods for high-spin open-shell molecules using pair natural orbitals.[361]
- C. Riplinger, F. Neese: An efficient and near linear scaling pair natural orbital based local coupled-cluster method.[720]
- C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese: Natural triple excitations in local coupled-cluster calculations with pair natural orbitals.[722]
- C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, F. Neese: Sparse maps - A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory.[721]
- D. Datta, S. Kossmann, F. Neese: Analytic energy derivatives for the calculation of the first-order molecular properties using the domain-based local pair-natural orbital coupled-cluster theory[191]
- M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, F. Neese: A new linear scaling, efficient and accurate, open-shell domain based pair natural orbital coupled cluster singles and doubles theory.[739]

In 2013, the so-called DLPNO-CCSD method ("domain based local pair natural orbital") was introduced.[720] This method is near linear scaling with system size and allows for giant calculations to be performed. In 2016, significant changes to the algorithm were implemented leading to linear scaling with system size concerning computing time, hard disk and memory consumption.[721] The principal idea behind DLPNO is the following: it became clear early on that the PNO space for a given electron pair (ij) is local and located in the same region of space as the electron pair (ij). In LPNO-CCSD this locality was partially used in the local fitting to the PNOs (controlled by the parameter TCutMKN). However, the PNOs were expanded in canonical virtual orbitals which led to some higher order scaling steps. In DLPNO, the PNOs are expanded in the set of projected atomic orbitals:

$$|\tilde{\mu}\rangle = \left(1 - \sum\nolimits_i |i\rangle \langle i|\right) |\mu\rangle$$

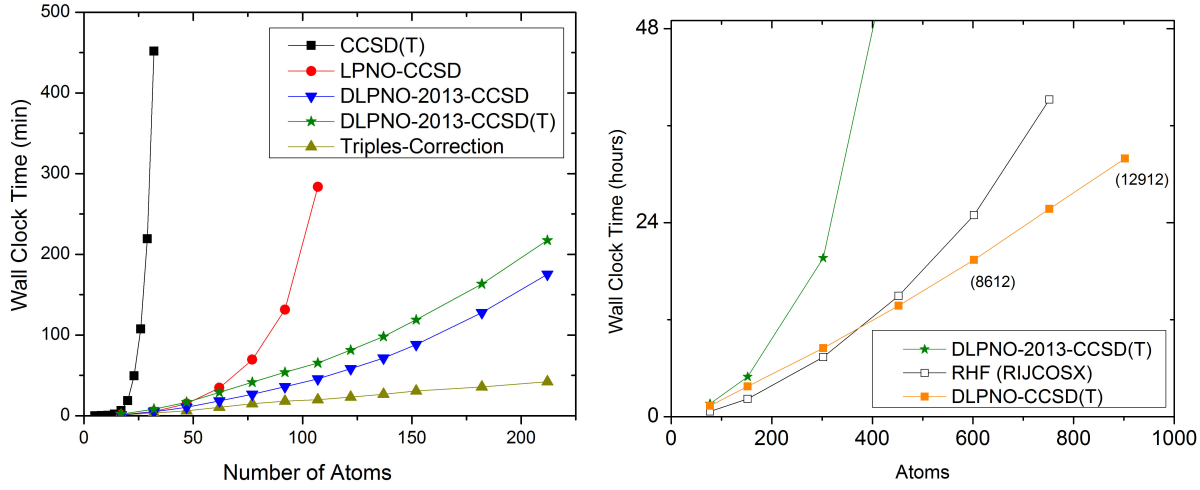
where $|\mu\rangle$ is an atomic orbital and $|i\rangle$ refers to an occupied molecular orbital. Such projected orbitals are an overcomplete representation of the virtual space. The projected orbital $|\tilde{\mu}\rangle$ is located in the same region of space as $|\mu\rangle$ and hence can be assigned to atomic centers. This has first been invented and used by Pulay and Saebo [704] in their pioneering work on local correlation methods and widely exploited by Werner, Schütz and co-workers in their local correlation approaches. [754, 755] DLPNO-CCSD goes one step further in expanding the PNOs $|\tilde{a}_{ij}\rangle$ of a given pair (ij) as:

$$|\tilde{a}_{ij}\rangle = \sum_{\tilde{\mu} \in \{ij\}} d^{ij}_{\tilde{\mu}\tilde{a}} |\tilde{\mu}\rangle$$

where $\tilde{\mu} \in \{ij\}$ is the domain of atoms (range of $\tilde{\mu}$) that is associated with the electron pair ij. The advantage of the PNO method is, that these domains can be chosen to be large (>15-20 atoms) without compromising the efficiency of the method.

The comparison between LPNO-CCSD and DLPNO-CCSD is shown in Fig. 6.5. It is obvious that DLPNO-CCSD is (almost) never slower than LPNO-CCSD. However, its true advantages do become most apparent for molecules with more than approximately 60 atoms. The triples correction, that was added with our second paper from 2013, shows a perfect linear scaling, as is shown in part (a) of Fig. 6.5. For large systems it adds about 10%–20% to the DLPNO-CCSD computation time, hence its addition is possible for all systems for which the latter can still be

obtained. Since 2016, the entire DLPNO-CCSD(T) algorithm is linear scaling. The improvements of the linear-scaling algorithm, compared to DLPNO2013-CCSD(T), start to become significant at system sizes of about 300 atoms, as becomes evident in part (b) of Fig. 6.5.



(a) (a) DLPNO2013 Scaling

(b) (b) DLPNO Scaling

Fig. 6.5: a) Scaling behavior of the canonical CCSD, LPNO-CCSD and DLPNO2013-CCSD(T) methods. It is obvious that only DLPNO2013-CCSD and DLPNO2013-CCSD(T) can be applied to large molecules. The advantages of DLPNO2013-CCSD over LPNO-CCSD do not show before the system has reached a size of about 60 atoms. b) Scaling behavior of DLPNO2013-CCSD(T), DLPNO-CCSD(T) and RHF using RIJCOSX. It is obvious that only DLPNO-CCSD(T) can be applied to truly large molecules, is faster than the DLPNO2013 version, and even has a crossover with RHF at about 400 atoms.

Using the DLPNO-CCSD(T) approach it was possible for the first time (in 2013) to perform a CCSD(T) level calculation on an entire protein (Crambin with more than 650 atoms, Fig. 6.6). While the calculation using a double-zeta basis took about 4 weeks on one CPU with DLPNO2013-CCSD(T), it takes only about 4 days to complete with DLPNO-CCSD(T). With DLPNO-CCSD(T) even the triple-zeta basis calculation can be completed within reasonable time, taking 2 weeks on 4 CPUs.

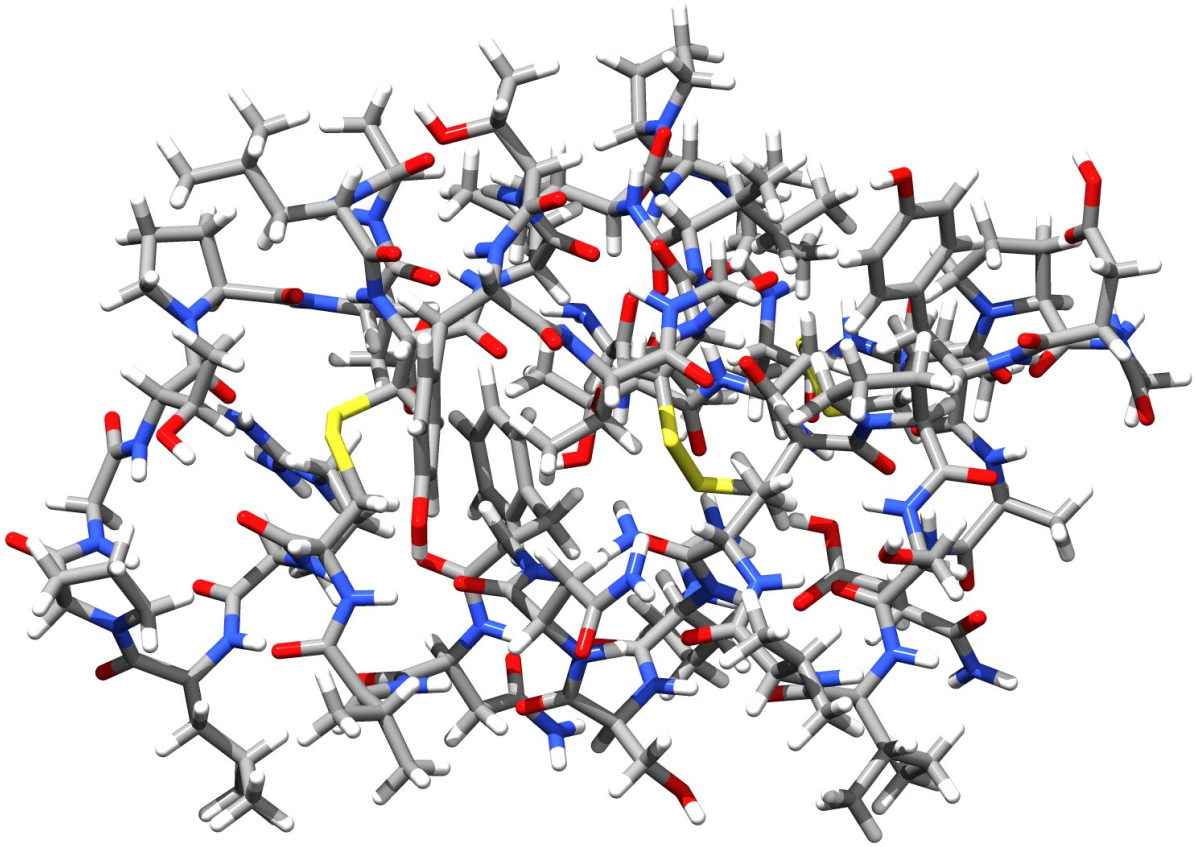


Fig. 6.6: Structure of the Crambin protein - the first protein to be treated with a CCSD(T) level ab initio method

The use of the LPNO (and DLPNO) methods is simple and requires little special attention from the user:

```
# Local Pair Natural Orbital Test
! cc-pVTZ cc-pVTZ/C LPNO-CCSD TightSCF
# or
! cc-pVTZ cc-pVTZ/C DLPNO-CCSD TightSCF
%maxcore 2000

# these are the default values - they need not to be touched!
%mdci  TCutPNO   3.33e-7   # cutoff for PNO occupation numbers. This
                             is the main truncation parameter
       TCutPairs 1e-4  # cut-off for estimated pair correlation energies.
                         This exploits the locality in the internal space
       TCutMKN   1e-3  # this is a technical parameter here that controls the domain
                         size for the local fit to the PNOs. It is conservative.
       end

* xyz 0 1
... (coordinates)
*
```

Using the well tested default settings, the LPNO-CEPA (LPNO-CPF, LPNO-VCEPA), LPNO-QCISD and LPNO-CCSD (LPNO-pCCSD) methods[6] can be run in strict analogy to canonical calculations and should approximate

[6] As a technical detail: The closed-shell LPNO QCISD and CCSD come in two technical variants - LPNO1-CEPA/QCISD/CCSD and LPNO2-CEPA/CCSD/QCISD. The "2" variants consume less disk space but are also slightly less accurate than the "1" variants. This is discussed in the original paper in the case of QCISD and CCSD. For the sake of accuracy, the "1" variants are the default. In those cases, where "1" can still be performed, the computational efficiency of both approaches is not grossly different. For LPNO CCSD there is also a third variant (LPNO3-CCSD, also in the open-shell version) which avoids neglecting the dressing of the external exchange operator. However, the results do not differ significantly from variant 1 but the calculations will become more expensive. Thus it is not recommend to use variant 3. Variant 2 is not available in the open-shell version.

the canonical result very closely. In fact, one should not view the LPNO methods as new model chemistry - they are designed to reproduce the canonical results, including BSSE. This is different from the domain based local correlation methods that do constitute a new model chemistry with properties that are different from the original methods.

In some situations, it may be appropriate to adapt the accuracy of the calculation. Sensible defaults have been determined from extensive benchmark calculations and are accessible via LoosePNO, NormalPNO and TightPNO keywords in the simple input line.[522]

These keywords represent the recommended way to control the accuracy of DLPNO calculations as follows. Manual changing of thresholds beyond these specifying these keywords is usually discouraged.

```
# Tight settings for increased accuracy, e.g. when investigating
# weak interactions or conformational equilibria
! cc-pVTZ cc-pVTZ/C DLPNO-CCSD(T) TightPNO TightSCF

# OR: Default settings (no need to give NormalPNO explicitly)
# Useful for general thermochemistry
! cc-pVTZ cc-pVTZ/C DLPNO-CCSD(T) NormalPNO TightSCF

# OR: Loose settings for rapid estimates
! cc-pVTZ cc-pVTZ/C DLPNO-CCSD(T) LoosePNO TightSCF

%maxcore 2000

* xyz 0 1
... (coordinates)
*
```

Since ORCA 4.0, the linear-scaling DLPNO implementation described in reference [721] is the default DLPNO algorithm. However, for comparison, the first DLPNO implementation from references [720] and [722] can still be called by using the DLPNO2013 prefix instead of the DLPNO- prefix.

```
# DLPNO-CCSD(T) calculation using the 2013 implementation
! cc-pVTZ cc-pVTZ/C DLPNO2013-CCSD(T)

# DLPNO-CCSD(T) calculation using the linear-scaling implementation
! cc-pVTZ cc-pVTZ/C DLPNO-CCSD(T)

* xyz 0 1
... (coordinates)
*
```

Until ORCA 4.0, the "semi-canonical" approximation is used in the perturbative triples correction for DLPNO-CCSD. It was found that the "semi-canonical" approximation is a very good approximation for most systems. However, the "semi-canonical" approximation can introduce large errors in rare cases (particularly when the HOMO-LUMO gap is small), whereas the DLPNO-CCSD is still very accurate. To improve the accuracy of perturbative triples correction, since 4.1, an improved perturbative triples correction for DLPNO-CCSD is available, DLPNO-CCSD(T1)[341]. In DLPNO-CCSD(T1), the triples amplitudes are computed iteratively, which can reproduce more accurately the canonical (T) energies.

It is necessary to clarify the nomenclature used in ORCA input files. The keyword to invoke "semi-canonical" perturbative triples correction approximation is DLPNO-CCSD(T). While, the keyword of improved iterative approximation is DLPNO-CCSD(T1). However, in our recent paper[341], the "semi-canonical" perturbative triples correction approximation is named DLPNO-CCSD(T0), whereas the improved iterative one is called DLPNO-CCSD(T). Thus, the names used in our paper are different from those in ORCA input files. An example input file to perform improved iterative perturbative triples correction for DLPNO-CCSD is given below,

```
# DLPNO-CCSD(T1) calculation using the iterative triples correction
! cc-pVTZ cc-pVTZ/C DLPNO-CCSD(T1)

%mdci
```

(continues on next page)

(continued from previous page)

```
      TNOSCALES    10.0 #TNO truncation scale for strong triples, TNOSCALES*TCutTNO.
                        Default setting is 10.0
      TNOSCALEW   100.0 #TNO truncation scale for weak triples, TNOSCALEW*TCutTNO
                        Default setting is 100.0
      TriTolE      1e-4 #(T) energy convergence tolerance
                        Default setting is 1e-4
%end
* xyz 0 1
... (coordinates)
*
```

Since ORCA 4.2, the improved iterative perturbative triples correction for open-shell DLPNO-CCSD is available as well. The keyword of open-shell DLPNO-CCSD(T) is the same as that of the closed-shell case.

Since ORCA 4.0, the high-spin open-shell version of the DLPNO-CISD/QCISD/CCSD implementations have been made available on top of the same machinery as the 2016 version of the RHF-DLPNO-CCSD code. The present UHF-DLPNO-CCSD is designed to be an heir to the UHF-LPNO-CCSD and serves as a natural extension to the RHF-DLPNO-CCSD. A striking difference between UHF-LPNO and newly developed UHF-DLPNO methods is that the UHF-DLPNO approach gives identical results to that of the RHF variant when applied to closed-shell species while UHF-LPNO does not. Usage of this program is quite straightforward and shown below:

```
# (1) In case of ROHF reference
! ROHF DLPNO-CCSD def2-TZVPP def2-TZVPP/C TightSCF TightPNO

# (2) In case of UHF reference, the QROs are constructed first and used for
#     the open-shell DLPNO-CCSD computations
! UHF DLPNO-CCSD def2-TZVPP def2-TZVPP/C TightSCF TightPNO

# (3) In case that UKS is specified, the QROs are constructed first and used as
#     "unconverged" UHF orbitals for the open-shell DLPNO-CCSD computations.
#     This approach is useful when the converged UHF wavefunction is qualitatively
#     wrong but the UKS wavefunction is not
! UKS CAM-B3LYP DLPNO-CCSD def2-TZVPP def2-TZVPP/C TightSCF TightPNO
```

Note

DLPNO-CISD/QCISD/CCSD methods are dedicated to closed-shell and high-spin open-shell species, but not spin-polarized systems (e.g. open shell singlets or antiferromagnetically coupled transition metal clusters). Performing DLPNO-CISD/QCISD/CCSD calculations upon open shell singlet UHF/UKS wavefunctions will give results resembling the corresponding closed shell singlet calculations, because the DLPNO calculations will be done on the closed-shell determinant composed of the QRO orbitals. Similarly, calculations of spin-polarized systems other than open shell singlets may give qualitatively wrong results. For spin-polarized systems, the UHF-LPNO-CCSD or Mk-LPNO-CCSD methods are available, in addition to DLPNO-NEVPT2.

The same set of truncation parameters as closed-shell DLPNO-CCSD is used also in case of open-shell DLPNO. The open-shell DLPNO-CCSD produces more than 99.9 % of the canonical CCSD correlation energy as in case of the closed-shell variant. This feature is certainly different from the UHF-LPNO methods because the open-shell DLPNO-CCSD is re-designed from scratch on the basis of a new PNO ansatz which makes use of the high-spin open-shell NEVPT framework. The computational timings of the UHF-DLPNO-CCSD and RIJCOSX-UHF for linear alkane chains in triplet state are shown in Fig. 6.7.

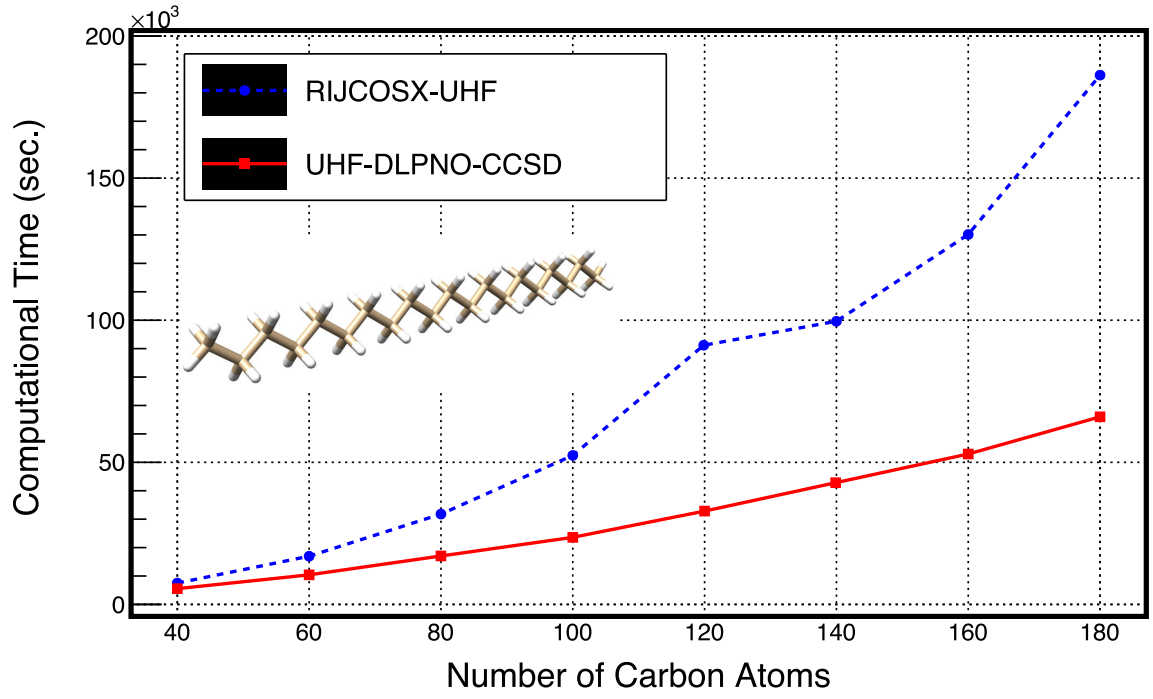


Fig. 6.7: Computational times of RIJCOSX-UHF and UHF-DLPNO-CCSD for the linear alkane chains (C_nH_{2n+2}) in triplet state with def2-TZVPP basis and default frozen core settings. 4 CPU cores and 128 GB of memory were used on a single cluster node.

Although those systems are somewhat idealized for the DLPNO method to best perform, it is clear that the preceding RIJCOSX-UHF is the rate-determining step in the total computational time for large examples. In the open-shell DLPNO implementations, SOMOs are included not only in the occupied space but also in the PNO space in the preceding integral transformation step. This means the presence of more SOMOs may lead to more demanding PNO integral transformation and DLPNO-CCSD iterations. The illustrative examples include active site model of the [NiFe] Hydrogenase in triplet state and the oxygen evolving complex (OEC) in the high-spin state, which are shown in Figures *7* and *8*, respectively. With def2-TZVPP basis set and NormalPNO settings, a single point calculation on [NiFe] Hydrogenase (Fig. 6.8) took approximately 45 hours on a single cluster node by using 4 CPU cores of Xeon E5-2670. A single point calculation on the OEC compound (Fig. 6.9) with the same computational settings finished in 44 hours even though the number of AOs in this system is even fewer than the Hydrogenase: the Hydrogenase active site model and OEC involve 4007 and 2606 AO basis functions, respectively. Special care should be taken if the system possesses more than ten SOMOs, since inclusion of more SOMOs may drastically increase the prefactor of the calculations. In addition, if the SOMOs are distributed over the entire molecular skeleton, each pair domain may not be truncated at all; in this case speedup attributed to the domain truncation will not be achieved at all.

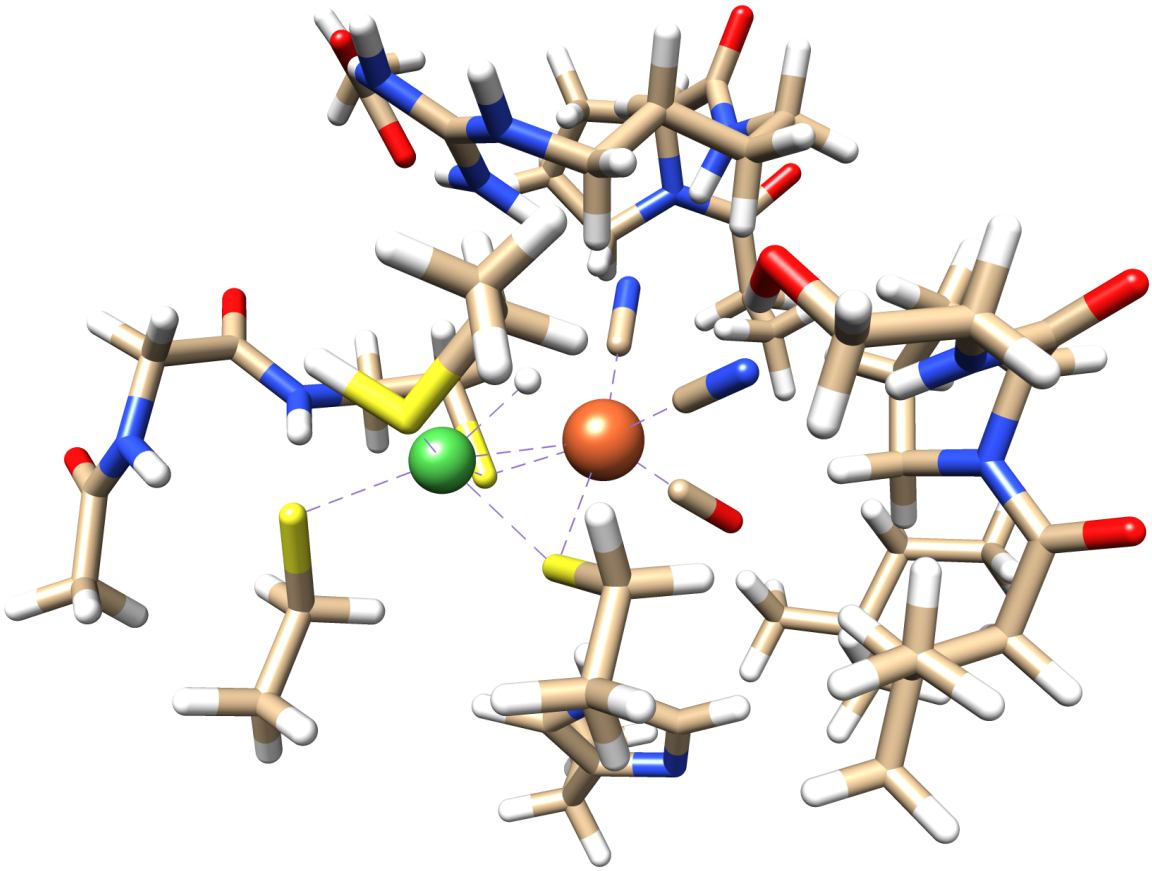


Fig. 6.8: Ni-Fe active center in the [NiFe] Hydrogenase in its second-coordination sphere. The whole model system is composed of 180 atoms.

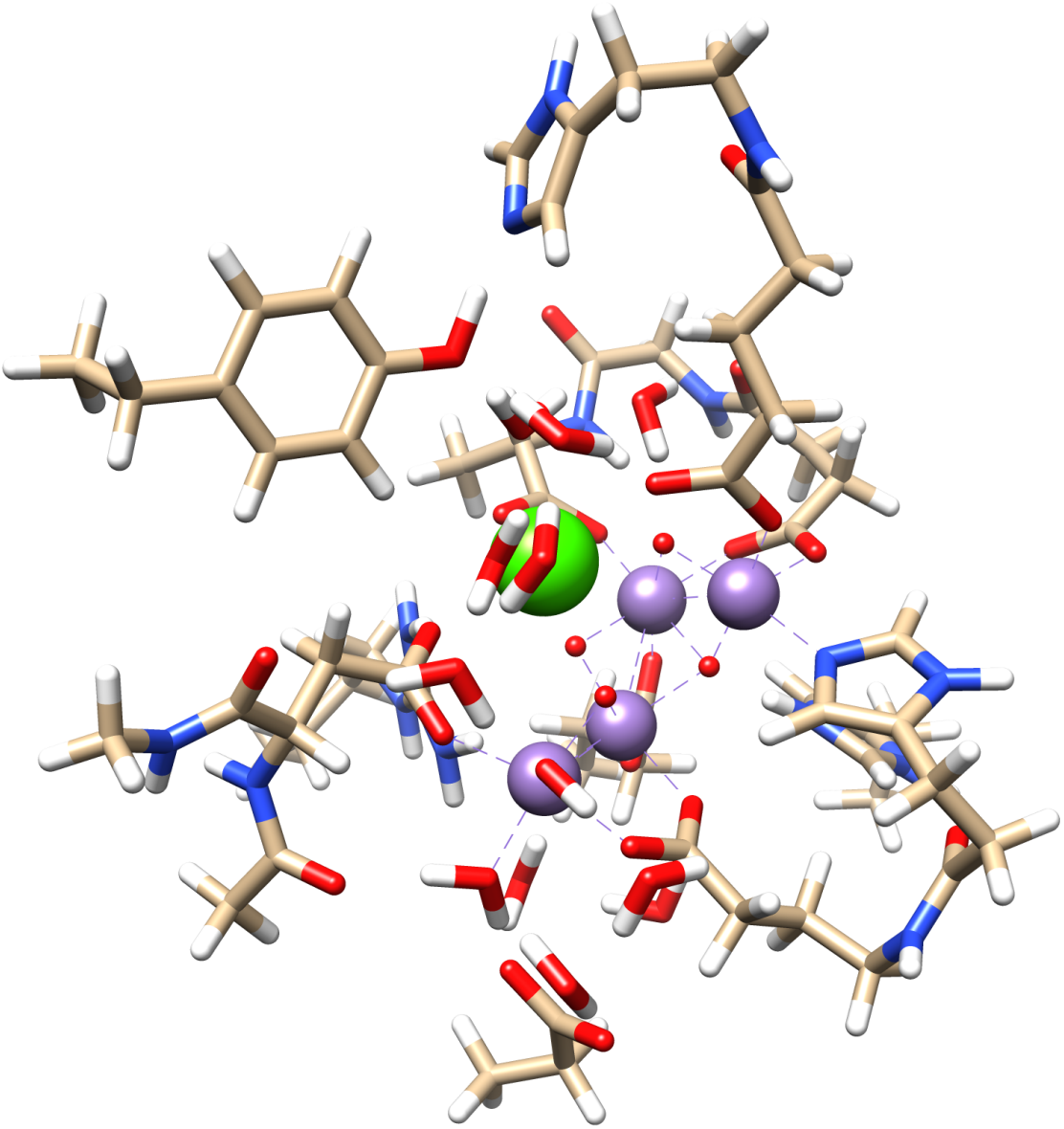


Fig. 6.9: A model compound for the OEC in the S_2 state of photosystem II which is composed of 238 atoms. In its high-spin state, the OEC possesses 13 SOMOs in total.

Calculation of the orbital-unrelaxed density has been implemented for closed-shell DLPNO-CCSD. This permits analytical computation of first-order properties, such as multipole moments or electric field gradients. In order to reproduce conventional unrelaxed CCSD properties to a high degree of accuracy, tighter thresholds may be needed than given by the default settings. Reading of the reference[191] is recommended. Calculation of the unrelaxed density is requested as usual:

```
%MDCI Density Unrelaxed End
```

There are a few things to be noticed about (D)LPNO methods:

- The LPNO methods obligatorily make use of the RI approximation. Hence, a correlation fit set must be provided.
- The DLPNO-CCSD(T) method is applicable to closed-shell or high-spin open-shell species. When performing DLPNO calculations on open-shell species, it is always better to have UCO option: If preceding SCF converges to broken-symmetry solutions, it is not guaranteed that the DLPNO-CCSD gives physically meaningful results.

- Besides the closed-shell version which uses a RHF or RKS reference determinant there is an open-shell version of the LPNO-CCSD for high-spin open-shell molecules (see original paper) using an UHF or UKS reference determinant built from quasi-restricted orbitals (QROs, see section *Open-Shell Equations*). Since the results of the current open-shell version are slightly less accurate than that of the closed-shell version it is mandatory to specify if you want to use the closed-shell or open-shell version for calculations of closed-shell systems, i.e. always put the "RHF" ("RKS") or "UHF" ("UKS") keyword in the simple keyword line. Open-shell systems can be of course only treated by the open-shell version. **Do not mix results of the closed- and open-shell versions of LPNO methods** (e.g. if you calculate reaction energies of a reaction in which both closed- and open-shell molecules take part, you should use the open-shell version throughout). This is because the open-shell LPNO results for the closed-shell species certainly differ from those of closed-shell implementations. This drawback of the open-shell LPNO methods has led to the development of a brand new open-shell DLPNO approach which converges to the RHF-DLPNO in the closed-shell limit. **Importantly, one can mix the results of closed- and open-shell versions of DLPNO approaches.**
- The open-shell version of the DLPNO approach uses a different strategy to the LPNO variant to define the open-shell PNOs. This ensures that, unlike the open-shell LPNO, the PNO space converges to the closed-shell counterpart in the closed-shell limit. Therefore, in the closed-shell limit, the open-shell DLPNO gives identical correlation energy to the RHF variant up to at least the third decimal place. The perturbative triples correction referred to as, (T), is also available for the open-shell species.
- When performing a calculation on the open-shell species with either of canonical/LPNO/DLPNO methods on top of the Slater determinant constructed from the QROs, special attention should be paid on the orbitals energies of those QROs. In some cases, the orbitals energy of the highest SOMO appear to be higher than that of the lowest VMO. Similarly to this, the orbital energy of the highest DOMO may appear to higher than that of the lowest SOMOs. In such cases, the CEPA/QCISD/CCSD iteration may show difficulty in convergence. In the worst case, it just diverges. Most likely, in such cases, one has to suspect the charge and multiplicity might be wrong. If they are correct, you may need much prettier starting orbitals and a bit of good luck! Apart from a careful choice of starting orbitals (in particular, DFT orbitals can be used in place of the default HF orbitals if the latter have qualitative deficiencies, including but not limited to severe spin contamination), changing the maximum DIIS expansion space size (`MaxDIIS`) and the level shift (`LShift`) in the `%mdci` block may alleviate the convergence problems to some extent.
- DLPNO-CCSD(T)-F12 and DLPNO-CCSD(T1)-F12 (iterative triples) are available for both closed- and open-shell cases. These methods employ a perturbative F12 correction on top of the DLPNO-CCSD(T) correlation energy calculation. The F12 part of the code uses the RI approximation in the same spirit as the canonical RI-F12 methods (refer to section *Explicitly Correlated MP2 and CCSD(T) Calculations*). Hence, they should be compared with methods using the RI approximation for both CC and F12 parts. The F12 correction takes only a fraction (usually 10-30%) of the total time (excluding SCF) required to calculate the DLPNO-CCSD(T)-F12 correlation energy. Thus, the F12 correction scales the same (linear or near-linear) as the parent DLPNO method. Furthermore, no new truncation parameters are introduced for the F12 procedure, preserving the black-box nature of the DLPNO method. The F12D approximation is highly recommended as it is computationally cheaper than the F12 approach which involves a double RI summation. Keywords: DLPNO-CCSD(T)-F12D, DLPNO-CCSD(T)-F12, DLPNO-CCSD(T1)-F12D, DLPNO-CCSD(T1)-F12, DLPNO-CCSD-F12D, DLPNO-CCSD-F12.
- Parallelization is done.
- There are three thresholds that can be user controlled that can all be adjusted in the %mdci block: (a) T_{CutPNO} controls the number of PNOs per electron pair. This is the most critical parameter and has a default value of 3.33×10^{-7}. (b) T_{CutPairs} controls a perturbative selection of significant pairs and has a default value of 10^{-4}. (c) T_{CutMKN} is a technical parameter and controls the size of the fit set for each electron pair. It has a default value of 10^{-3}. All of these default values are conservative. Hence, no adjustment of these parameters is necessary. All DLPNO-CCSD truncations are bound to these three truncation parameters and should almost not be touched (Hence they are also not documented :)).
- The preferred way to adjust accuracy when needed is to use the "LoosePNO/NormalPNO/TightPNO" keywords. In addition, "TightPNO" triggers the full iterative (DLPNO-MP2) treatment in the MP2 guess, whereas the other options use a semicanonical MP2 calculation. Table 6.5 and Table 6.6 contain the thresholds used by the current (2016) and old (2013) implementations, respectively.
- LPNO-VCEPA/n (n=1,2,3) methods are only available in the open-shell version yet.

- LPNO variants of the parameterized coupled-cluster methods (pCCSD, see section *Theory*) are also available (e.g. LPNO-pCCSD/1a and LPNO-pC/2a).
- The LPNO methods reproduce the canonical energy differences to typically better than 1 kcal/mol. This accuracy exists over large parts of the potential energy surface. Tightening TCutPairs to 1e-5 gives more accurate results but also leads to significantly longer computation times.
- Potential energy surfaces are virtually but not perfectly smooth (like any method that involves cut-offs). Numerical gradient calculations have been attempted and reported to have been successful.
- The LPNO methods do work together with RIJCOSX, RI-JK and also with ANO basis sets and basis set extrapolation. They also work for conventional integral handling.
- The methods behave excellently with large basis sets. Thus, they stay efficient even when large basis sets are used that are necessary to obtain accurate results with wavefunction based *ab initio* methods. This is a prerequisite for efficient computational chemistry applications.
- For LPNO-CCSD, calculations with about 1000 basis functions are routine, calculations with about 1500 basis functions are possible and calculations with 2000-2500 basis functions are the limit on powerful computers. For DLPNO-CCSD much larger calculations are possible. There is virtually no crossover and DLPNO-CCSD is essentially always more efficient than LPNO-CCSD. Starting from about 50 atoms the differences become large. The largest DLPNO-CCSD calculation to date featured $>$1000 atoms and more than 20000 basis functions!
- Using large main memory is not mandatory but advantageous since it speeds up the initial integral transformation significantly (controlled by "MaxCore" in the %mdci block, see section *Local correlation*).
- The open-shell versions are about twice as expensive as the corresponding closed-shell versions.
- Analytic gradients are not available.
- An unrelaxed density implementation is available for closed-shell DLPNO-CCSD, permitting calculation of first-order properties.

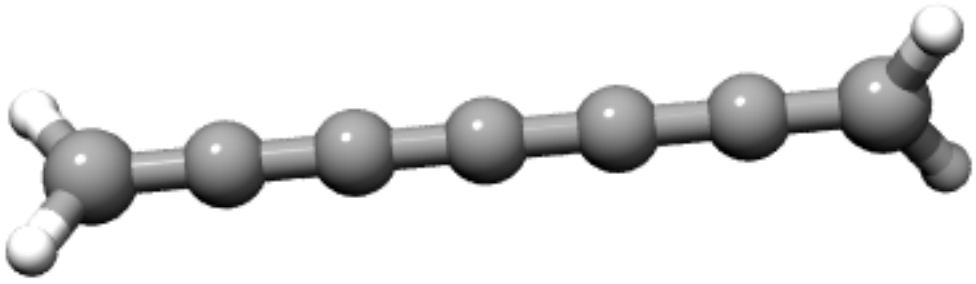
Table 6.5: Accuracy settings for DLPNO coupled cluster (current version).

Setting	T_{CutPairs}	T_{CutDO}	T_{CutPNO}	T_{CutMKN}	MP2 pair treatment
LoosePNO	10^{-3}	2×10^{-2}	1.00×10^{-6}	10^{-3}	semicanonical
NormalPNO	10^{-4}	1×10^{-2}	3.33×10^{-7}	10^{-3}	semicanonical
TightPNO	10^{-5}	5×10^{-3}	1.00×10^{-7}	10^{-3}	full iterative

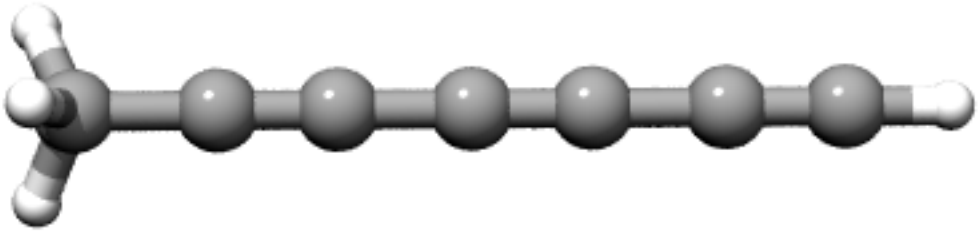
Table 6.6: Accuracy settings for DLPNO coupled cluster (deprecated 2013 version).

Setting	T_{CutPairs}	T_{CutPNO}	T_{CutMKN}	MP2 pair treatment
LoosePNO	10^{-3}	1.00×10^{-6}	10^{-3}	semicanonical
NormalPNO	10^{-4}	3.33×10^{-7}	10^{-3}	semicanonical
TightPNO	10^{-5}	1.00×10^{-7}	10^{-4}	full iterative

As an example, see the following isomerization reaction that appears to be particularly difficult for DFT:



Isomerizes to:



The results of the calculations (closed-shell versions) with the def2-TZVP basis set (about 240 basis functions) are shown below:

Method	Energy Difference (kcal/mol)	Time (min)
CCSD(T)	-14.6	92.4
CCSD	-18.0	55.3
LPNO-CCSD	-18.6	20.0
CEPA/1	-12.4	42.2
LPNO-CEPA/1	-13.5	13.4

The calculations are typical in the sense that: (a) the LPNO methods provide answers that are within 1 kcal/mol of the canonical results, (b) CEPA approximates CCSD(T) more closely than CCSD. The speedups of a factor of 2 – 5 are moderate in this case. However, this is also a fairly small calculation. For larger systems, speedups of the LPNO methods compared to their canonical counterparts are on the order of a factor $>$100–1000.

Cluster in molecules (CIM)

Cluster in molecules (CIM) approach is a linear scaling local correlation method developed by Li and the coworkers in 2002.[514] It was further improved by Li, Piecuch, Kállay and other groups recently.[337, 340, 515, 516, 729] The CIM is inspired by the early local correlation method developed by Förner and coworkers.[250] The total correlation energy of a closed-shell molecule can be considered as a summation of correlation energies of each occupied LMOs.

$$E_{\text{corr}} = \sum_{i}^{occ} E_i = \sum_{i}^{occ} \frac{1}{4} \sum_{j,ab} \langle ij||ab\rangle T_{ab}^{ij} \tag{6.7}$$

For each occupied LMO, it only correlates with its nearby occupied LMOs and virtual MOs. To reproduce the correlation energy of each occupied LMO, only a subset of occupied and virtual LMOs are needed in the correlation calculation. Instead of doing the correlation calculation of the whole molecule, the correlation energies of all LMOs can be obtained within various subsystems.

The CIM approach implemented in ORCA is following an algorithm proposed by Guo and coworkers with a few improvements.[337, 340]

1. To avoid the real space cutoff, the differential overlap integral (DOI) is used instead of distance threshold. There is only one parameter 'CIMTHRESH' in CIM approach, controlling the construction of CIM subsystems. If the DOI between LMO *i* and LMO *j* is larger than CIMTHRESH, LMO *j* will be included into the MO domain of *i*. By including all nearby LMO of *i*, one can construct a subsystem for MO *i*. The default value of CIMTHRESH is 0.001. If accurate results are needed, a tighter CIMTHRESH must be used.
2. Since ORCA 4.1, the neglected correlations between LMO *i* and LMOs outside the MO domain of *i* are considered as well. These weak correlations are approximately evaluated by dipole moment integrals. With this correction, the CIM results of 3 dimensional proteins are significantly improved. About 99.8% of the correlation energies are recovered.

The CIM can invoke different single reference correlation methods for the subsystem calculations. In ORCA the CIM-RI-MP2, CIM-CCSD(T), CIM-DLPNO-MP2 and CIM-DLPNO-CCSD(T) methods are available. The CIM-RI-MP2 and CIM-DLPNO-CCSD(T) have been proved to be very efficient and accurate methods to compute correlation energies of very big molecules, containing a few thousand atoms.[340]
The usage of CIM in ORCA is simple. For CIM-RI-MP2,

```
#
# CIM-RI-MP2 calculation
#
! RI-MP2 cc-pVDZ cc-pVDZ/C CIM
%CIM
  CIMTHRESH 0.0005  # Default value is 0.001
end

* xyzfile 0 1 CIM.xyz
```

For CIM-DLPNO-CCSD(T),

```
#
# CIM-DLPNO-CCSD calculation
#
! DLPNO-CCSD(T) cc-pVDZ  cc-pVDZ/C CIM
* xyzfile 0 1 CIM.xyz
```

The parallel efficiency of CIM has been significantly improved.[340] Except for a few domain construction substeps, the CIM algorithm can achieve very high parallel efficiency. Since ORCA 4.1, the parallel version does not support Windows platform anymore due to the parallelization strategy. The generalization of CIM from closed-shell to open-shell (multi-reference) will also be implemented in the near future.

Arbitrary Order Coupled-Cluster Calculations

ORCA features an interface to Kallay's powerful MRCC program. This program must be obtained separately. The interface is restricted to single point energies but can be used for rigid scan calculations or numerical frequencies.

The use of the interface is simple:

```
#
# Test the MRCC code of Mihael Kallay
#
! cc-pVDZ Conv SCFConv10 UseSym

%mrcc method   "CCSDT"
      ETol     9
      end

* xyz 0 1
F 0 0 0
H 0 0 0.95
*
```

The Method string can be any of:

```
# The excitation level specification can be anything
# like SD, SDT, SDTQ, SDTQP etc.
 %mrcc method   "CCSDT"
                "CCSD(T)"
                "CCSD[T]"
                "CCSD(T)_L"  (the lambda version)
                "CC3"
                "CCSDT-1a"
```

(continues on next page)

(continued from previous page)

```
                "CCSDT-1b"
                "CISDT"
```

It is not a good idea, of course, to use this code for CCSD or CCSD(T) or CISD. Its real power lies in performing the higher order calculations. Open-shell calculations can presently not be done with the interface.

Note also that certain high-order configuration interaction or coupled cluster methods, such as CISDT, CISDTQ, CC3 and CCSDT etc., have now been implemented natively in ORCA in the AUTOCI module. For details please consult section *CI methods using generated code*.

6.1.4 Density Functional Theory

Standard Density Functional Calculations

Density functional calculations are as simple to run as HF calculations. In this case, the RI-J approximation will be the default for LDA, GGA or meta-GGA non-hybrid functionals, and the RIJCOSX for the hybrids. The RI-JK approximation might also offer large speedups for smaller systems.

For example, consider this B3LYP calculation on the cyclohexane molecule.

```
# Test a simple DFT calculation
! B3LYP SVP
* xyz 0 1
C   -0.79263     0.55338    -1.58694
C    0.68078     0.13314    -1.72622
C    1.50034     0.61020    -0.52199
C    1.01517    -0.06749     0.77103
C   -0.49095    -0.38008     0.74228
C   -1.24341     0.64080    -0.11866
H    1.10490     0.53546    -2.67754
H    0.76075    -0.97866    -1.78666
H   -0.95741     1.54560    -2.07170
H   -1.42795    -0.17916    -2.14055
H   -2.34640     0.48232    -0.04725
H   -1.04144     1.66089     0.28731
H   -0.66608    -1.39636     0.31480
H   -0.89815    -0.39708     1.78184
H    1.25353     0.59796     1.63523
H    1.57519    -1.01856     0.93954
H    2.58691     0.40499    -0.67666
H    1.39420     1.71843    -0.44053
*
```

If you want an accurate single point energy then it is wise to choose "`TightSCF`" and select a basis set of at least valence triple-zeta plus polarization quality (e.g. `def2-TZVP`).

DFT Calculations with RI

DFT calculations that do not require the HF exchange to be calculated (non-hybrid DFT) can be *very* efficiently executed with the RI-J approximation. It leads to very large speedups at essentially no loss of accuracy. The use of the RI-J approximation may be illustrated for a medium sized organic molecule - Penicillin:

```
# RI-DFT calculation on the Penicillin molecule
! BP86 SVP TightSCF

* xyz 0 1
N     3.17265     1.15815    -0.09175
C     2.66167     0.72032     1.18601
```

(continues on next page)

(continued from previous page)

```
C     4.31931     0.59242    -0.73003
C     2.02252     1.86922    -0.54680
C     1.37143     1.52404     0.79659
S     2.72625    -1.05563     0.80065
C     4.01305    -0.91195    -0.52441
C     5.58297     1.09423    -0.06535
O     1.80801     2.36292    -1.62137
N     0.15715     0.73759     0.70095
C     5.25122    -1.72918    -0.12001
C     3.41769    -1.50152    -1.81857
O     6.60623     1.14077    -0.91855
O     5.72538     1.40990     1.08931
C    -1.08932     1.35001     0.75816
C    -2.30230     0.45820     0.54941
O    -1.19855     2.53493     0.96288
O    -3.48875     1.21403     0.57063
C    -4.66939     0.59150     0.27339
C    -4.84065    -0.79240     0.11956
C    -5.79523     1.39165     0.03916
C    -6.07568    -1.34753    -0.22401
C    -7.03670     0.85454    -0.30482
C    -7.18253    -0.52580    -0.43612
H     3.24354     1.09074     2.02120
H     4.33865     0.87909    -1.77554
H     1.26605     2.42501     1.39138
H     0.17381    -0.25857     0.47675
H     6.05024    -1.64196    -0.89101
H     5.67754    -1.39089     0.85176
H     5.01118    -2.81229    -0.01401
H     2.50304    -0.95210    -2.14173
H     4.15186    -1.44541    -2.65467
H     3.14138    -2.57427    -1.69700
H     7.29069     1.46408    -0.31004
H    -2.21049    -0.02915    -0.44909
H    -2.34192    -0.28647     1.37775
H    -4.00164    -1.48999     0.26950
H    -5.69703     2.48656     0.12872
H    -6.17811    -2.44045    -0.33185
H    -7.89945     1.51981    -0.47737
H    -8.15811    -0.96111    -0.71027
*
```

The job has 42 atoms and 430 contracted basis functions. Yet, it executes in just a few minutes elapsed time on any reasonable personal computer.

NOTES:

- The RI-J approximation requires an "auxiliary basis set" in addition to a normal orbital basis set. For the Karlsruhe basis sets there is the universal auxiliary basis set of Weigend that is called with the name `def2/J` (all-electron up to Kr). When scalar relativistic Hamiltonians are used (DKH or ZORA) along with all-electron basis sets, then a general-purpose auxiliary basis set is the `SARC/J` that covers most of the periodic table. Other choices are documented in sections *Basis Sets* and *Choice of Basis Set*.
- For "pure" functionals the use of RI-J with the `def2/J` auxiliary basis set is the default.

Since DFT is frequently applied to open-shell transition metals we also show one (more or less trivial) example of a Cu(II) complex treated with DFT.

```
! BP86 SV SlowConv
%base "temp"
* xyz -2 2
  Cu  0     0     0
```

(continues on next page)

(continued from previous page)

```
  Cl  2.25  0     0
  Cl -2.25  0     0
  Cl  0     2.25  0
  Cl  0    -2.25  0
*

$new_job

! B3LYP NoRI TZVP TightSCF MORead
%moinp "temp.gbw"
%scf GuessMode CMatrix
     end
* xyz -2 2
  Cu  0     0     0
  Cl  2.25  0     0
  Cl -2.25  0     0
  Cl  0     2.25  0
  Cl  0    -2.25  0
*
```

Although it would not have been necessary for this example, it shows a possible strategy how to converge such calculations. First a less accurate but fast job is performed using the RI approximation, a GGA functional and a small basis set without polarization functions. Note that a larger damping factor has been used in order to guide the calculation (`SlowConv`). The second job takes the orbitals of the first as input and performs a more accurate hybrid DFT calculation. A subtle point in this calculation on a dianion in the gas phase is the command `GuessMode CMatrix` that causes the corresponding orbital transformation to be used in order to match the orbitals of the small and the large basis set calculation. This is always required when the orbital energies of the small basis set calculation are positive as will be the case for anions.

Hartree–Fock and Hybrid DFT Calculations with RIJCOSX

Frustrated by the large difference in execution times between pure and hybrid functionals, we have been motivated to study approximations to the Hartree-Fock exchange term. The method that we have finally come up with is called the "chain of spheres" COSX approximation and may be thought of as a variant of the pseudo-spectral philosophy. Essentially, in performing two electron integrals, the first integration is done numerically on a grid and the second (involving the Coulomb singularity) is done analytically. For algorithmic and theoretical details see Refs. [623] and [383]. Upon combining this treatment with the Split-RI-J method for the Coulomb term (thus, a Coulomb fitting basis is needed!), we have designed the RIJCOSX approximation that can be used to accelerate Hartree-Fock and hybrid DFT calculations. Note that this introduces another grid on top of the DFT integration grid which is usually significantly smaller.

OBS.: Since ORCA 5, RIJCOSX is the default option for hybrid DFT (can be turned off by using !NOCOSX). However, it is by default NOT turned on for HF.

In particular for large and accurate basis sets, the speedups obtained in this way are very large - we have observed up to a factor of sixty! The procedure is essentially linear scaling such that large and accurate calculations become possible with high efficiency. The RIJCOSX approximation is basically available throughout the program. The default errors are on the order of 0.05 ± 0.1 kcal mol^{-1} or less in the total energies as well as in energy differences and can be made smaller with larger than the default grids or by running the final SCF cycle without this approximation. The impact on bond distances is a fraction of a pm, angles are better than a few tenth of a degree and soft dihedral angles are good to about 1 degree. To the limited extent to which it has been tested, vibrational frequencies are roughly good to 0.1 wavenumbers with the default settings.

The use of RIJCOSX is very simple:

```
! HF def2-TZVPP TightSCF RIJCOSX
...
```

One thing to be mentioned in correlation calculations with RIJCOSX is that the requirements for the SCF and correlation fitting bases are quite different. We therefore support two different auxiliary basis sets in the same run:

```
! RI-MP2 def2-TZVPP def2/J def2-TZVPP/C TightSCF RIJCOSX
...
```

Hartree–Fock and Hybrid DFT Calculations with RI-JK

An alternative algorithm for accelerating the HF exchange in hybrid DFT or HF calculations is to use the RI approximation for both Coulomb and exchange. This is implemented in ORCA for SCF single point energies but not for gradients.

```
! RHF def2-TZVPP def2/JK RI-JK
...
```

The speedups for small molecules are better than for RIJCOSX, for medium sized molecules (e.g. $(\mathrm{gly})_4$) similar, and for larger molecules RI-JK is less efficient than RIJCOSX. The errors of RI-JK are usually below 1 mEh and the error is very smooth (smoother than for RIJCOSX). Hence, for small calculations with large basis sets, RI-JK is a good idea, for large calculations on large molecules RIJCOSX is better.

Note

- For RI-JK you will need a larger auxiliary basis set. For the Karlsruhe basis set, the universal def2/JK and def2/JKsmall basis sets are available. They are large and accurate.
- For UHF RI-JK is roughly twice as expensive as for RHF. This is not true for RIJCOSX.
- RI-JK is available for conventional and direct runs and also for ANO bases. There the conventional mode is recommended.

A comparison of the RIJCOSX and RI-JK methods (taken from Ref. [465]) for the $(\mathrm{gly})_2$, $(\mathrm{gly})_4$ and $(\mathrm{gly})_8$ is shown below (wall clock times in second for performing the entire SCF):

		Def2-SVP	Def2-TZVP(-df)	Def2-TZVPP	Def2-QZVPP
(gly)$_2$	*Default*	105	319	2574	27856
	RI-JK	44	71	326	3072
	RIJCOSX	70	122	527	3659
(gly)$_4$	*Default*	609	1917	13965	161047
	RI-JK	333	678	2746	30398
	RIJCOSX	281	569	2414	15383
(gly)$_8$	*Default*	3317	12505	82774	
	RI-JK	3431	5452	16586	117795
	RIJCOSX	1156	2219	8558	56505

It is obvious from the data that for small molecules the RI-JK approximation is the most efficient choice. For $(\mathrm{gly})_4$ this is already no longer obvious. For up to the def2-TZVPP basis set, RI-JK and RIJCOSX are almost identical and for def2-QZVPP RIJCOSX is already a factor of two faster than RI-JK. For large molecules like $(\mathrm{gly})_8$ with small basis sets RI-JK is not a big improvement but for large basis set it still beats the normal 4-index calculation. RIJCOSX on the other hand is consistently faster. It leads to speedups of around 10 for def2-TZVPP and up to 50-60 for def2-QZVPP. Here it outperforms RI-JK by, again, a factor of two.

DFT Calculations with Second Order Perturbative Correction (Double-Hybrid Functionals)

There is a family of functionals which came up in 2006 and were proposed by Grimme [320]. They consist of a semi-empirical mixture of DFT components and the MP2 correlation energy calculated with the DFT orbitals and their energies. Grimme referred to his functional as B2PLYP (B88 exchange, 2 parameters that were fitted and perturbative mixture of MP2 and LYP) – a version with improved performance (in particular for weak interactions) is mPW2PLYP [772] and is also implemented. From the extensive calibration work, the new functionals appear to give better energetics and a narrower error distribution than B3LYP. Thus, the additional cost of the calculation of the MP2 energy may be well invested (and is quite limited in conjunction with density fitting in the RI part). Martin has reported reparameterizations of B2PLYP (B2GP-PLYP, B2K-PLYP and B2T-PLYP) that are optimized for "general-purpose", "kinetic" and "thermochemistry" applications.[436, 843] In 2011, Goerigk and Grimme published the PWPB95 functional with spin-opposite-scaling and relatively low amounts of Fock exchange, which make it promising for both main-group and transition-metal chemistry. [308]

Among the best performing density functionals[312] are Martin's "DSD"-double-hybrids, which use different combinations of exchange and correlation potentials and spin-component-scaled MP2 mixing. Three of these double-hybrids (DSD-BLYP, DSD-PBEP86 and DSD-PBEB95)[467, 468, 469] are available via simple input keywords. Different sets of parameters for the DSD-double-hybrids are published, e.g. for the use with and without D3. The keywords `DSD-BLYP`, `DSD-PBEP86` and `DSD-PBEB95` request parameters consistent with the GMTKN55[312] benchmark set results. The keywords `DSD-BLYP/2013` and `DSD-PBEP86/2013` request the slightly different parameter sets used in the 2013 paper by Kozuch and Martin.[469] To avoid confusion, the different parameters are presented in Table 6.7.

Table 6.7: DSD-DFT parameters defined in ORCA

Keywords	ScalDFX	ScalHFX	ScalGGAC	PS	PT	D3S6	D3S8	D3A2
DSD-BLYP	0.25	0.75	0.53	0.46	0.60			
DSD-BLYP D3BJ	0.31	0.69	0.54	0.46	0.37	0.50	0.213	6.0519
DSD-BLYP/2013 D3BJ	0.29	0.71	0.54	0.47	0.40	0.57	0	5.4
DSD-PBEP86	0.28	0.72	0.44	0.51	0.36			
DSD-PBEP86 D3BJ	0.30	0.70	0.43	0.53	0.25	0.418	0	5.65
DSD-PBEP86/2013 D3BJ	0.31	0.69	0.44	0.52	0.22	0.48	0	5.6
DSD-PBEB95	0.30	0.70	0.52	0.48	0.22			
DSD-PBEB95 D3BJ	0.34	0.66	0.55	0.46	0.09	0.61	0	6.2

Note that D3A1 is always 0 for these functionals.

Three different variants of MP2 can be used in conjunction with these functionals. Just specifying the functional name leads to the use of RI-MP2 by default. In this case, an appropriate auxiliary basis set for correlation fitting needs to be specified. It is very strongly recommended to use the RI variants instead of conventional MP2, as their performance is vastly better. Indeed, there is hardly ever any reason to use conventional MP2. To turn this option off just use !NORI in the simple input (which also turns off the RIJCOSX approximation) or `%mp2 RI false end`. More information can be found in the relevant sections regarding RI-MP2.

Finally, DLPNO-MP2 can be used as a component of double-hybrid density functionals. In that case, a "`DLPNO-`" prefix needs to be added to the functional name, for example `DLPNO-B2GP-PLYP` or `DLPNO-DSD-PBEP86`. Please refer to the relevant manual sections for more information on the DLPNO-MP2 method.

For each functional, parameters can be specified explicitly in the input file, e.g. for RI-DSD-PBEB95 with D3BJ:

```
! D3BJ
%method
  Method      DFT
  DoMP2       True
  Exchange    X_PBE
  Correlation C_B95
```

(continues on next page)

(continued from previous page)

```
  LDAOpt      C_PWLDA # specific for B95
  ScalDFX     0.34
  ScalHFX     0.66
  ScalGGAC    0.55
  ScalLDAC    0.55 # must be equal to ScalGGAC
  ScalMP2C    1.00 # for all DSD-DFs
  D3S6        0.61
  D3S8        0
  D3A1        0    # for all DSD-DFs
  D3A2        6.2
end
%mp2
  DoSCS       True
  RI          True
  PS          0.46
  PT          0.09
end
```

In this version of ORCA, double-hybrid DFT is available for single points, geometry optimizations [619], dipole moments and other first order properties, magnetic second order properties (chemical shifts, g-tensors), as well as for numerical polarizabilities and frequencies.

There are also double-hybrid functionals, such as XYG3 and ωB97M(2), which must be applied to orbitals converged with a different functional. This can be accomplished with a two-step calculation using `MORead` and `MaxIter=1`. Note that because the orbitals are not obtained self-consistently, only single point energies can be computed in this way, i.e. no density, gradient, or properties! For example, the ωB97M(2) functional must be used with ωB97M-V orbitals,[557] which can be done with the following input:

```
*xyz 0 1
  H 0.0 0.0 0.0
  F 0.0 0.0 0.9
*
%compound
  Variable EwB97MV, EwB97M2; # Output variables
  # Step 1: wB97M-V calculation to obtain the orbitals
  New_Step
    ! wB97M-V SCNL def2-TZVP
  Step_End
  # Step 2: single iteration with the wB97M(2) functional
  #         + MP2 correlation to get the final energy
  ReadMOs(1);
  New_Step
    ! wB97M(2) SCNL NoFrozenCore def2-TZVP def2-TZVP/C
    %scf
      MaxIter    1
      IgnoreConv 1 # prevent the "not converged" error
    end
  Step_End
  Read EwB97MV = DFT_Total_En[1];     # wB97M-V energy
  Read EwB97M2 = MP2_Total_Energy[2]; # wB97M(2) energy
End
```

DFT Calculations with Atom-pairwise Dispersion Correction

It is well known that many DFT functionals do not include dispersion forces. It is possible to use a simple atom-pairwise correction to account for the major parts of this contribution to the energy [131, 321, 324, 326]. We have adopted the code and method developed by Stefan Grimme in this ORCA version. The method is parameterized for many established functionals (e.g. BLYP, BP86, PBE, TPSS, B3LYP, B2PLYP).[7] For the 2010 model the Becke-Johnson damping version (`! D3BJ`) is the default and will automatically be invoked by the simple keyword `! D3`. The charge dependent atom-pairwise dispersion correction (keyword `! D4`) is using the D4(EEQ)-ATM dispersion model[132], other D4 versions, using tight-binding partial charges, are currently only available with the standalone DFT-D4 program.

```
! BLYP D3 def2-QZVPP Opt

%paras R= 2.5,4.0,16
end

%geom Constraints
      { C 0 C }
      { C 1 C }
      end
end

* xyz 0 1
Ar     0.0000000     0.0000000     {R}
H      0.0000000     0.0000000     0.0000000
C      0.0000000     0.0000000    -1.0951073
H      0.5163499     0.8943443    -1.4604101
H      0.5163499    -0.8943443    -1.4604101
H     -1.0326998     0.0000000    -1.4604101
*
```

In this example, a BLYP calculation without dispersion correction will show a repulsive potential between the argon atom and the methane molecule. Using the D3 dispersion correction as shown above, the potential curve shows a minimum at about 3.1−3.2 Å. The atom-pairwise correction is quite successful and Grimme's work suggests that this is more generally true. For many systems like stacked DNA base pairs, hydrogen bond complexes and other weak interactions the atom-pairwise dispersion correction will improve substantially the results of standard functionals at essentially no extra cost.

Note

- Dispersion corrections do not only affect non-covalent complexes, but also affect conformational energies (and conformer structures) which are heavily influenced by intramolecular dispersion. Therefore, for large and/or flexible molecules, including the dispersion correction is almost always recommended or even required (except for a handful of cases where it cannot, should not or need not be used, see below). For small systems, the dipersion correction may result in basically no improvement of the results, but is usually harmless anyway.
- DFT calculations with small basis sets (such as double zeta basis sets) often yield attractive potential energy surfaces even without the dispersion correction. However, this is due to basis set superposition error (BSSE), and the interaction energy brought about by the BSSE frequently does not match the true interaction energy due to dispersion (because they have completely different origins). Therefore, although a DFT double zeta calculation without the dispersion correction may appear to give qualitatively correct results, or occasionally even better results than a double zeta calculation with dispersion corrections (because in the latter case one typically overestimates the total attraction), it is still highly recommended to "get the right answer for the right reason" by reducing the BSSE and turning on the dispersion correction.

[7] For expert users: The keyword `D2`, `D3ZERO`, `D3BJ` and `D4` select the empirical 2006, the atom-pairwise 2010 model, respectively, with either zero-damping or Becke-Johnson damping, or the partial charge dependent atom-pairwise 2018 model. The default is the most accurate `D3BJ` model. The outdated model from 2004 [319] is no longer supported and can only be invoked by setting `DFTDOPT = 1`. The C6-scaling coefficient can be user defined using e.g. "`%method DFTDScaleC6 1.2 end`"

The BSSE can be corrected by a variety of means, for example (1) by using a larger basis set; (2) by using the counterpoise correction (*Counterpoise Correction*); or (3) by using the geometrical counterpoise correction (section *DFT and HF Calculations with the Geometrical Counterpoise Correction: gCP*). Of these, (3) is available at almost no cost (including analytic gradient contributions), and is especially suitable for geometry optimization of large molecules. Otherwise (1) (or its combination with (2)) may be more appropriate due to its higher accuracy.

- Functionals that contain VV10-type non-local dispersion (in general, these are the functionals whose names end with "-V") do not need (and cannot be used together with) dispersion corrections. The same holds for post-HF and multireference methods, like MP2, CCSD(T), CASSCF and NEVPT2. However, one can add a dispersion correction on top of HF.
- Certain functionals, especially the Minnesota family of functionals (e.g. M06-2X), describe medium-range dispersion but miss long-range dispersion. They give reasonable dispersion energies for small to medium systems but may slightly underestimate the dispersion energies for large systems. For them, dispersion corrections are only available in the zero-damping variant, and one should use the `D3ZERO` keyword instead of the `D3` keyword. As the uncorrected functional already accounts for the bulk of the dispersion in this case, the dispersion correction is much less important than e.g. the case of B3LYP, and should in general be considered as beneficial but not mandatory.
- Some density functional developers reparameterize the functional itself while parameterizing the dispersion correction. A famous example is the ωB97X family of functionals, to be detailed in the next section. For these functionals, the `D3`, `D3BJ` or `D4` keywords should be hyphenated with the name of the functional itself, and some quantities that normally would not change when adding dispersion corrections (e.g. orbitals, excitation energies, dipole moments) may change slightly when adding or removing the dispersion correction. Likewise, for these functionals the structural changes that one observe upon adding or removing the dispersion correction cannot be completely attributed to the dispersion correction itself, but may contain contributions due to the change of the functional.

DFT Calculations with Range-Separated Hybrid Functionals

All range-separated functionals in ORCA use the error function based approach according to Hirao and coworkers.[410] This allows the definition of DFT functionals that dominate the short-range part by an adapted exchange functional of LDA, GGA or meta-GGA level and the long-range part by Hartree-Fock exchange.

CAM-B3LYP,[899] LC-BLYP[845], LC-PBE[410, 601] and members of the ωB97-family of functionals have been implemented into ORCA, namely ωB97, ωB97X[151], ωB97X-D3[525], ωB97X-V[554], ωB97M-V[556], ωB97X-D3BJ and ωB97M-D3BJ.[602] (For more information on ωB97X-V[554] and ωB97M-V[556] see section *DFT Calculations with the Non-Local, Density Dependent Dispersion Correction (VV10): DFT-NL*.) Some of them incorporate fixed amounts of Hartree-Fock exchange (EXX) and/or DFT exchange and they differ in the RS-parameter μ. In the case of ωB97X-D3, the proper D3 correction (employing the zero-damping scheme) should be calculated automatically. The D3BJ correction is used automatically for ωB97X-D3BJ and ωB97M-D3BJ (as well as for the meta-GGA B97M-D3BJ). The same is true for the D4-based variants ωB97M-D4 and ωB97X-D4. The D3BJ and D4 variants have also been shown to perform well for geometry optimizations [603].

Several restrictions apply to these functionals at the moment. They have only been implemented and tested for use with the `libint` integral package and for RHF and UHF single-point, ground state nuclear gradient, ground state nuclear Hessian, TDDFT, and TDDFT nuclear gradient calculations. Only the standard integral handling (NORI), RIJONX, and RIJCOSX are supported. **Do not use these functionals with any other options**.

DFT Calculations with Range-Separated Double Hybrid Functionals

For the specifics of the range-separated double-hybrid functionals the user is referred to sections *DFT Calculations with Second Order Perturbative Correction (Double-Hybrid Functionals)*, *DFT Calculations with Range-Separated Hybrid Functionals* and *Doubles Correction*. The first range-separated double hybrids available in ORCA were ωB2PLYP and ωB2GP-PLYP[146]. Both were optimized for the calculation of excitation energies, but they have recently also been tested for ground-state properties[601].

A large variety of range-separated double hybrids with and without spin-component/opposite scaling have become available in ORCA 5. Some have been developed with ground-state properties in mind, most for excitation energies. See Section *Choice of Functional* for more details and citations.

6.1.5 Quadratic Convergence

The standard SCF implementation in ORCA uses the DIIS algorithm[702, 703] for initial and an approximate second-order converger for final convergence[264, 607]. This approach converges quickly for most chemical systems. However, there are many interesting systems with a more complicated electronic structure for which the standard SCF protocol converges either **slowly** ("creeping"), converges to an **excited state**, or **diverges**. In those cases, a newly developed trust-region augmented Hessian (TRAH) SCF approach[64, 349, 381, 734] should be used. The TRAH-SCF method always converges to a local minimum and converges quadratically near the solution.

You can run TRAH from the beginning by adding

```
! TRAH
```

to the simple input line if you expect convergence difficulties. Open-shell molecules notoriously have SCF convergence issues, in particular, if they are composed of many open-shell atoms. In Fig. 6.10, the convergence of a TRAH-SCF calculation is shown for a high-spin Rh cluster for which the standard SCF diverges. The errors of the electronic gradient or residual vector converge almost steadily below the default TRAH accuracy of 10^{-6}.

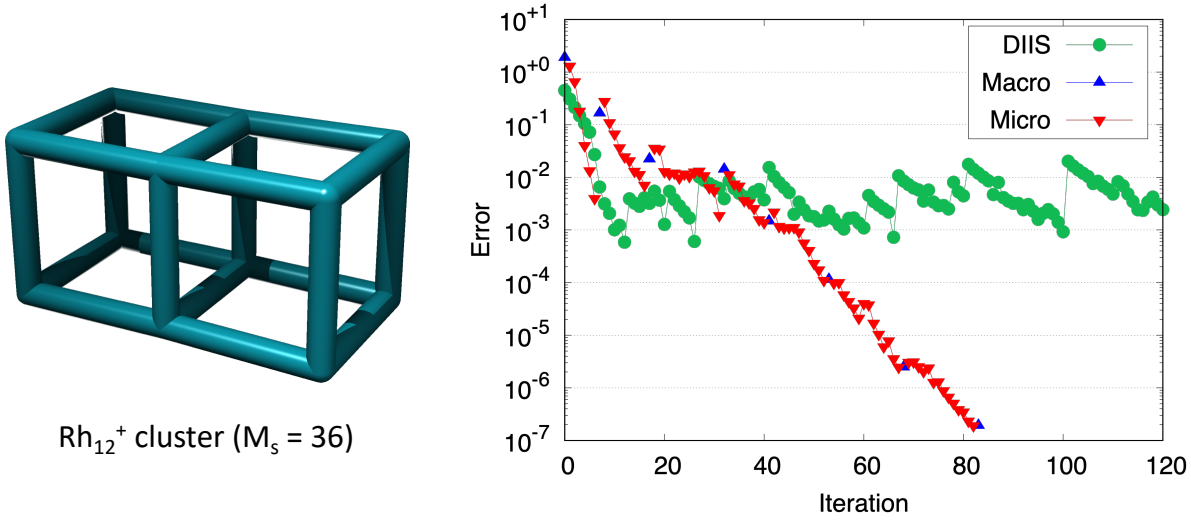


Fig. 6.10: TRAH-SCF gradient norm of a PBE/def2-TZVP calculation for a Rh_{12}^{+} cluster in high-spin configuration ($M_s = 36$). The structure was taken from Ref. [362].

Alternatively, TRAH is launched automatically if standard SCF (DIIS/SOSCF) shows converge problems (default), an approach which is called AutoTRAH.

```
%scf
 AutoTRAH true
end
```

You can switch off the automatic start of TRAH by adding

```
! NOTRAH
```

to the simple input line or

```
%scf
 AutoTRAH false
end
```

Convergence problems are detected by comparing the norm of the electronic gradient at multiple iterations which is explained in more detail in Sec. *Trust-Region Augmented Hessian (TRAH) SCF*.

TRAH-SCF is currently implemented for restricted closed-shell (`RHF` and `RKS`) and unrestricted open-shell determinants (`UHF` and `UKS`) and can be accelerated with `RIJ`, `RIJONX`, `RIJK`, or `RIJCOSX`. Solvation effects can also be accounted for with the `C-PCM` and `SMD` models. Restricted open-shell calculations are not possible yet.

TRAH-SCF can also be applied to large molecules as it is parallelized and works with AO Fock matrices. However, for systems with large HOMO-LUMO gaps that converge well, the default SCF converger is usually faster because the screening in `TRAH` is less effective and more iterations are required.

For a more detailed documentation we refer to Sec. *Trust-Region Augmented Hessian (TRAH) SCF*.

Note

- TRAH is mathematically guaranteed to converge with a sufficient number of iterations, provided that there is no numerical noise (e.g. round-off error, truncation error) in the calculation. Therefore, if TRAH fails to converge, this means that either the default number of iterations is not large enough, or certain numerical thresholds are not tight enough. One can verify whether the former possibility is operative by checking whether the error still decreases steadily towards zero in the last SCF iterations. If yes, one can increase the number of iterations; otherwise it may be worthwhile to try increasing the integration grid, tighten the integral thresholds, etc.
- For some functionals (e.g. `PWPB95`), the native ORCA implementation supports only their XC energies and potentials, but not their XC kernels. In this case one should switch to the LibXC implementation instead, e.g. replace `PWPB95` by `LIBXC(PWPB95)`. Otherwise the calculation aborts upon entering the TRAH procedure.

6.1.6 Counterpoise Correction

In calculating weak molecular interactions the nasty subject of the basis set superposition error (BSSE) arises. It consists of the fact that if one describes a dimer, the basis functions on A help to lower the energy of fragment B and vice versa. Thus, one obtains an energy that is biased towards the dimer formation due to basis set effects. Since this is unwanted, the Boys and Bernardi procedure aims to correct for this deficiency by estimating what the energies of the monomers would be if they had been calculated with the dimer basis set. This will stabilize the monomers relative to the dimers. The effect can be a quite sizable fraction of the interaction energy and should therefore be taken into account. The original Boys and Bernardi formula for the interaction energy between fragments A and B is:

$$\Delta E = E_{AB}^{AB}(AB) - E_{A}^{A}(A) - E_{B}^{B}(B) - \left[E_{A}^{AB}(AB) - E_{A}^{AB}(A) + E_{B}^{AB}(AB) - E_{B}^{AB}(B)\right] \tag{6.8}$$

Here $E_X^Y\,(Z)$ is the energy of fragment X calculated at the optimized geometry of fragment Y with the basis set of fragment Z. Thus, you need to do a total the following series of calculations:

1. optimize the geometry of the dimer and the monomers with some basis set Z. This gives you $E_{AB}^{AB}\,(AB)$, $E_{A}^{A}\,(A)$ and $E_{B}^{B}\,(B)$
2. delete fragment A (B) from the optimized structure of the dimer and re-run the single point calculation with basis set Z. This gives you $E_{B}^{AB}\,(B)$ and $E_{A}^{AB}\,(A)$.
3. Now, the final calculation consists of calculating the energies of A and B at the dimer geometry but with the dimer basis set. This gives you $E_{A}^{AB}\,(AB)$ and $E_{B}^{AB}\,(AB)$.

In order to achieve the last step efficiently, a special notation was put into ORCA which allows you to delete the electrons and nuclear charges that come with certain atoms but retain the assigned basis set. This trick consists of putting a ":" after the symbol of the atom. Here is an example of how to run such a calculation of the water dimer at the MP2 level (with frozen core):

```
#
# BSSE test
#

# ------------------------------------------
# First the monomer. It is a waste of course
# to run the monomer twice ...
# ------------------------------------------
! RHF MP2 TZVPP VeryTightSCF XYZFile PModel
%id "monomer"
* xyz 0 1
 O     7.405639     6.725069     7.710504
 H     7.029206     6.234628     8.442160
 H     8.247948     6.296600     7.554030
*

$new_job
! RHF MP2 TZVPP VeryTightSCF XYZFile PModel
%id "monomer"
* xyz 0 1
 O     7.405639     6.725069     7.710504
 H     7.029206     6.234628     8.442160
 H     8.247948     6.296600     7.554030
*

# ------------------------------------------
# now the dimer
# ------------------------------------------
$new_job
! RHF MP2 TZVPP VeryTightSCF XYZFile PModel
%id "dimer"
* xyz 0 1
O   7.439917   6.726792   7.762120
O   5.752050   6.489306   5.407671
H   7.025510   6.226170   8.467436
H   8.274883   6.280259   7.609894
H   6.313507   6.644667   6.176902
H   5.522285   7.367132   5.103852
*

# ------------------------------------------
# Now the calculations of the monomer at the
# dimer geometry
# ------------------------------------------
$new_job
! RHF MP2 TZVPP VeryTightSCF XYZFile PModel
%id "monomer_1"

* xyz 0 1
O   7.439917   6.726792   7.762120
H   7.025510   6.226170   8.467436
H   8.274883   6.280259   7.609894
*

$new_job
! RHF MP2 TZVPP VeryTightSCF XYZFile PModel
%id "monomer_1"
```

(continues on next page)

(continued from previous page)

```
* xyz 0 1
O   5.752050   6.489306   5.407671
H   6.313507   6.644667   6.176902
H   5.522285   7.367132   5.103852
*

# ------------------------------------------
# Now the calculation of the monomer at the
# dimer geometry but with the dimer basis set
# ------------------------------------------
$new_job
! RHF MP2 TZVPP VeryTightSCF XYZFile PModel
%id "monomer_2"
* xyz 0 1
O   7.439917   6.726792   7.762120
O : 5.752050   6.489306   5.407671
H   7.025510   6.226170   8.467436
H   8.274883   6.280259   7.609894
H : 6.313507   6.644667   6.176902
H : 5.522285   7.367132   5.103852
*

$new_job
! RHF MP2 TZVPP VeryTightSCF XYZFile PModel
%id "monomer_2"
* xyz 0 1
O : 7.439917   6.726792   7.762120
O   5.752050   6.489306   5.407671
H : 7.025510   6.226170   8.467436
H : 8.274883   6.280259   7.609894
H   6.313507   6.644667   6.176902
H   5.522285   7.367132   5.103852
*
```

You obtain the energies:

```
Monomer                        :   -152.647062118 Eh
Dimer                                :   -152.655623625 Eh   -5.372 kcal/mol
Monomer at dimer geometry:   -152.647006948 Eh    0.035 kcal/mol
Same with AB Basis set         :   -152.648364970 Eh   -0.818 kcal/mol

Thus, the corrected interaction energy is:
                -5.372 kcal/mol - (-0.818-0.035)=-4.52 kcal/mol
```

It is also possible to set entire fragments as ghost atoms using the `GhostFrags` keyword as shown below. See section *Fragment Specification* for different ways of defining fragments.

```
! MP2 TZVPP VeryTightSCF XYZFile PModel
* xyz 0 1
O   7.439917   6.726792   7.762120
O   5.752050   6.489306   5.407671
H   7.025510   6.226170   8.467436
H   8.274883   6.280259   7.609894
H   6.313507   6.644667   6.176902
H   5.522285   7.367132   5.103852
*
%geom
  GhostFrags {1} end # space-separated list and X:Y ranges accepted
  fragments
    1 {0 2 3} end
    2 {1 4 5} end
```

(continues on next page)

(continued from previous page)

```
  end
end
```

Starting from ORCA 6.0, we support geometry optimizations with the counterpoise correction, using analytic gradients. This opens up the way of obtaining accurate non-covalent complex geometries (instead of just interaction energies) using modest basis sets. To use this functionality, one should NOT simply add `!Opt` to the above input files, but should instead use the dedicated compound script `BSSEOptimization.cmp` available in the ORCA Compound Script repository (https://github.com/ORCAQuantumChemistry/CompoundScripts/blob/main/GeometryOptimization/BSSEOptimization.cmp). Detailed usage are described in the comments of the compound script.

6.1.7 Complete Active Space Self-Consistent Field Method

Introduction

There are several situations where a complete-active space self-consistent field (CASSCF) treatment is a good idea:

- Wavefunctions with significant multireference character arising from several nearly degenerate configurations (static correlation)
- Wavefunctions which require a multideterminantal treatment (for example multiplets of atoms, ions, transition metal complexes,)
- Situations in which bonds are broken or partially broken.
- Generation of orbitals which are a compromise between the requirements for several states.
- Generation of start orbitals for multireference methods covering dynamic correlation (NEVPT2, MRCI, MREOM, . . .)
- Generation of genuine spin eigenfunctions for multideterminantal/multireference wavefunctions.

In all of these cases the single-determinantal Hartree-Fock method fails badly and in most of these cases DFT methods will also fail. In these cases a CASSCF method is a good starting point. CASSCF is a special case of multiconfigurational SCF (MCSCF) methods which specialize to the situation where the orbitals are divided into three-subspaces: (a) the internal orbitals which are doubly occupied in all configuration state functions (CSFs) (b) partially occupied (active) orbitals (c) virtual (external) orbitals which are empty in all CSFs.

A fixed number of electrons is assigned to the internal subspace and the active subspace. If N-electrons are "active" in M orbitals one speaks of a CASSCF(N,M) wavefunctions. All spin-eigenfunctions for N-electrons in M orbitals are included in the configuration interaction step and the energy is made stationary with respect to variations in the MO and the CI coefficients. Any number of roots of any number of different multiplicities can be calculated and the CASSCF energy may be optimized with respect to a user defined average of these states.

The CASSCF method has the nice advantage that it is fully variational which renders the calculation of analytical gradients relatively easy. Thus, the CASSCF method may be used for geometry optimizations and numerical frequency calculations.

The price to pay for this strongly enhanced flexibility relative to the single-determinantal HF method is that the CASSCF method requires more computational resources and also more insight and planning from the user side. The technical details are explained in section *The Complete Active Space Self-Consistent Field (CASSCF) Module*. Here we explain the use of the CASSCF method by examples. In addition to the description in the manual, there is a separate tutorial for CASSCF with many more examples in the field of coordination chemistry. The tutorial covers the design of the calculation, practical tips on convergence as well as the computation of properties.

A number of properties are available in ORCA (g-tensor, ZFS splitting, CD, MCD, susceptibility, dipoles, . . .). The majority of CASSCF properties such as EPR parameters are computed in the framework of the quasi-degenerate perturbation theory. Some properties such as ZFS splittings can also be computed via perturbation theory or rigorously extracted from an effective Hamiltonian. For a detailed description of the available properties and options see section *CASSCF Properties*. All the aforementioned properties are computed within the CASSCF module. An exception are Mössbauer parameters, which are computed with the usual keywords using the EPRNMR module (*Mössbauer Parameters*).

A simple Example

One standard example of a multireference system is the Be atom. Let us run two calculations, a standard closed-shell calculation ($1s^2 2s^2$) and a CASSCF(2,4) calculation which also includes the ($1s^2 2s^1 2p^1$) and ($1s^2 2s^0 2p^2$) configurations.

```
! TZVPP TightSCF
* xyz 0 1
Be 0 0 0
*
```

This standard closed-shell calculation yields the energy `-14.56213241 Eh`. The CASSCF calculation

```
! TZVPP TightSCF
%casscf nel  2
        norb 4
        end
* xyz 0 1
Be 0 0 0
*
```

yields the energy `-14.605381525 Eh`. Thus, the inclusion of the 2p shell results in an energy lowering of 43 mEh which is considerable. The CASSCF program also prints the composition of the wavefunction:

```
------------------------------------------------
CAS-SCF STATES FOR BLOCK  1 MULT= 1 NROOTS= 1
------------------------------------------------

ROOT   0:  E=     -14.6053815294 Eh
      0.90060 [     0]: 2000
      0.03313 [     4]: 0200
      0.03313 [     9]: 0002
      0.03313 [     7]: 0020
```

This information is to be read as follows: The lowest state is composed of 90% of the configuration which has the active space occupation pattern 2000 which means that the first active orbital is doubly occupied in this configuration while the other three are empty. The MO vector composition tells us what these orbitals are (ORCA uses natural orbitals to canonicalize the active space).

```
                      0        1        2        3        4        5
                  -4.70502  -0.27270   0.11579   0.11579   0.11579   0.16796
                   2.00000   1.80121   0.06626   0.06626   0.06626   0.00000
                  --------  --------  --------  --------  --------  --------
 0 Be s             100.0     100.0       0.0       0.0       0.0     100.0
 0 Be pz              0.0       0.0      13.6       6.1      80.4       0.0
 0 Be px              0.0       0.0       1.5      93.8       4.6       0.0
 0 Be py              0.0       0.0      84.9       0.1      15.0       0.0
```

Thus, the first active space orbital has occupation number 1.80121 and is the Be-2s orbital. The other three orbitals are 2p in character and all have the same occupation number 0.06626. Since they are degenerate in occupation number space, they are arbitrary mixtures of the three 2p orbitals. It is then clear that the other components of the wavefunction (each with 3.31%) are those in which one of the 2p orbitals is doubly occupied.

How did we know how to put the 2s and 2p orbitals in the active space? The answer is – WE DID NOT KNOW! In this case it was "good luck" that the initial guess produced the orbitals in such an order that we had the 2s and 2p orbitals active. **IN GENERAL IT IS YOUR RESPONSIBILITY THAT THE ORBITALS ARE ORDERED SUCH THAT THE ORBITALS THAT YOU WANT IN THE ACTIVE SPACE COME IN THE DESIRED ORDER**. In many cases this will require re-ordering and **CAREFUL INSPECTION** of the starting orbitals.

Attention

If you include orbitals in the active space that are nearly empty or nearly doubly occupied, convegence problems are likely. The SuperCI(PT) [459] and Newton-Raphson method are less prone to these problems.

Starting Orbitals

Tip

In many cases natural orbitals of a simple correlated calculation of some kind provide a good starting point for CASSCF.

Let us illustrate this principle with a calculation on the Benzene molecule where we want to include all six π-orbitals in the active space. After doing a RHF calculation:

```
! RHF SV(P)

* int 0 1
C 0 0 0 0.000000 0.000 0.000
C 1 0 0 1.389437 0.000 0.000
C 2 1 0 1.389437 120.000 0.000
C 3 2 1 1.389437 120.000 0.000
C 4 3 2 1.389437 120.000 0.000
C 5 4 3 1.389437 120.000 0.000
H 1 2 3 1.082921 120.000 180.000
H 2 1 3 1.082921 120.000 180.000
H 3 2 1 1.082921 120.000 180.000
H 4 3 2 1.082921 120.000 180.000
H 5 4 3 1.082921 120.000 180.000
H 6 5 4 1.082921 120.000 180.000
*
%Output
        Print[P_ReducedOrbPopMO_L]      1
End
```

We can look at the orbitals around the HOMO/LUMO gap:

```
12        13        14        15        16        17
                  -0.63810  -0.62613  -0.59153  -0.59153  -0.50570  -0.49833
                   2.00000   2.00000   2.00000   2.00000   2.00000   2.00000
                  --------  --------  --------  --------  --------  --------
 0 C  s              2.9       0.0       0.3       0.1       0.0       0.0
 0 C  pz             0.0       0.0       0.0       0.0      16.5       0.0
 0 C  px             1.4      12.4       5.9       0.3       0.0      11.2
 0 C  py             4.2       4.1      10.1       5.9       0.0       0.1
 0 C  dyz            0.0       0.0       0.0       0.0       0.1       0.0
 0 C  dx2y2          0.1       0.1       0.2       0.2       0.0       0.5
 0 C  dxy            0.4       0.0       0.0       0.2       0.0       0.0
 1 C  s              2.9       0.0       0.3       0.1       0.0       0.0
 1 C  pz             0.0       0.0       0.0       0.0      16.5       0.0
 1 C  px             1.4      12.4       5.9       0.3       0.0      11.2
 1 C  py             4.2       4.1      10.1       5.9       0.0       0.1
 1 C  dyz            0.0       0.0       0.0       0.0       0.1       0.0
 1 C  dx2y2          0.1       0.1       0.2       0.2       0.0       0.5
 1 C  dxy            0.4       0.0       0.0       0.2       0.0       0.0
 2 C  s              2.9       0.0       0.0       0.4       0.0       0.1
 2 C  pz             0.0       0.0       0.0       0.0      16.5       0.0
 2 C  px             5.7       0.0       0.0      20.9       0.0      10.1
 2 C  py             0.0      16.5       1.3       0.0       0.0       0.0
 2 C  dxz            0.0       0.0       0.0       0.0       0.1       0.0
```

(continues on next page)

(continued from previous page)

```
 2 C  dx2y2           0.6       0.0       0.0       0.2       0.0       1.2
 2 C  dxy             0.0       0.1       0.5       0.0       0.0       0.0
 3 C  s               2.9       0.0       0.3       0.1       0.0       0.0
 3 C  pz              0.0       0.0       0.0       0.0      16.5       0.0
 3 C  px              1.4      12.4       5.9       0.3       0.0      11.2
 3 C  py              4.2       4.1      10.1       5.9       0.0       0.1
 3 C  dyz             0.0       0.0       0.0       0.0       0.1       0.0
 3 C  dx2y2           0.1       0.1       0.2       0.2       0.0       0.5
 3 C  dxy             0.4       0.0       0.0       0.2       0.0       0.0
 4 C  s               2.9       0.0       0.3       0.1       0.0       0.0
 4 C  pz              0.0       0.0       0.0       0.0      16.5       0.0
 4 C  px              1.4      12.4       5.9       0.3       0.0      11.2
 4 C  py              4.2       4.1      10.1       5.9       0.0       0.1
 4 C  dyz             0.0       0.0       0.0       0.0       0.1       0.0
 4 C  dx2y2           0.1       0.1       0.2       0.2       0.0       0.5
 4 C  dxy             0.4       0.0       0.0       0.2       0.0       0.0
 5 C  s               2.9       0.0       0.0       0.4       0.0       0.1
 5 C  pz              0.0       0.0       0.0       0.0      16.5       0.0
 5 C  px              5.7       0.0       0.0      20.9       0.0      10.1
 5 C  py              0.0      16.5       1.3       0.0       0.0       0.0
 5 C  dxz             0.0       0.0       0.0       0.0       0.1       0.0
 5 C  dx2y2           0.6       0.0       0.0       0.2       0.0       1.2
 5 C  dxy             0.0       0.1       0.5       0.0       0.0       0.0
 6 H  s               7.5       0.0       7.5       2.5       0.0       2.5
 7 H  s               7.5       0.0       7.5       2.5       0.0       2.5
 8 H  s               7.5       0.0       0.0      10.0       0.0       9.9
 9 H  s               7.5       0.0       7.5       2.5       0.0       2.5
10 H  s               7.5       0.0       7.5       2.5       0.0       2.5
11 H  s               7.5       0.0       0.0      10.0       0.0       9.9

                      18        19        20        21        22        23
                  -0.49833  -0.33937  -0.33937   0.13472   0.13472   0.18198
                   2.00000   2.00000   2.00000   0.00000   0.00000   0.00000
                  --------  --------  --------  --------  --------  --------
 0 C  s               0.1       0.0       0.0       0.0       0.0       2.2
 0 C  pz              0.0       8.1      24.4       7.8      23.4       0.0
 0 C  px              0.1       0.0       0.0       0.0       0.0       0.6
 0 C  py             10.4       0.0       0.0       0.0       0.0       1.7
 0 C  dxz             0.0       0.4       0.2       0.7       0.7       0.0
 0 C  dyz             0.0       0.2       0.0       0.7       0.0       0.0
 0 C  dx2y2           0.0       0.0       0.0       0.0       0.0       0.2
 0 C  dxy             1.0       0.0       0.0       0.0       0.0       0.5
 1 C  s               0.1       0.0       0.0       0.0       0.0       2.2
 1 C  pz              0.0       8.1      24.4       7.8      23.4       0.0
 1 C  px              0.1       0.0       0.0       0.0       0.0       0.6
 1 C  py             10.4       0.0       0.0       0.0       0.0       1.7
 1 C  dxz             0.0       0.4       0.2       0.7       0.7       0.0
 1 C  dyz             0.0       0.2       0.0       0.7       0.0       0.0
 1 C  dx2y2           0.0       0.0       0.0       0.0       0.0       0.2
 1 C  dxy             1.0       0.0       0.0       0.0       0.0       0.5
 2 C  s               0.0       0.0       0.0       0.0       0.0       2.2
 2 C  pz              0.0      32.5       0.0      31.2       0.0       0.0
 2 C  px              0.0       0.0       0.0       0.0       0.0       2.2
 2 C  py             11.6       0.0       0.0       0.0       0.0       0.0
 2 C  dxz             0.0       0.1       0.0       0.3       0.0       0.0
 2 C  dyz             0.0       0.0       0.8       0.0       1.8       0.0
 2 C  dx2y2           0.0       0.0       0.0       0.0       0.0       0.7
 2 C  dxy             0.4       0.0       0.0       0.0       0.0       0.0
 3 C  s               0.1       0.0       0.0       0.0       0.0       2.2
 3 C  pz              0.0       8.1      24.4       7.8      23.4       0.0
 3 C  px              0.1       0.0       0.0       0.0       0.0       0.6
```

(continues on next page)

(continued from previous page)

```
 3 C  py             10.4       0.0       0.0       0.0       0.0       1.7
 3 C  dxz             0.0       0.4       0.2       0.7       0.7       0.0
 3 C  dyz             0.0       0.2       0.0       0.7       0.0       0.0
 3 C  dx2y2           0.0       0.0       0.0       0.0       0.0       0.2
 3 C  dxy             1.0       0.0       0.0       0.0       0.0       0.5
 4 C  s               0.1       0.0       0.0       0.0       0.0       2.2
 4 C  pz              0.0       8.1      24.4       7.8      23.4       0.0
 4 C  px              0.1       0.0       0.0       0.0       0.0       0.6
 4 C  py             10.4       0.0       0.0       0.0       0.0       1.7
 4 C  dxz             0.0       0.4       0.2       0.7       0.7       0.0
 4 C  dyz             0.0       0.2       0.0       0.7       0.0       0.0
 4 C  dx2y2           0.0       0.0       0.0       0.0       0.0       0.2
 4 C  dxy             1.0       0.0       0.0       0.0       0.0       0.5
 5 C  s               0.0       0.0       0.0       0.0       0.0       2.2
 5 C  pz              0.0      32.5       0.0      31.2       0.0       0.0
 5 C  px              0.0       0.0       0.0       0.0       0.0       2.2
 5 C  py             11.6       0.0       0.0       0.0       0.0       0.0
 5 C  dxz             0.0       0.1       0.0       0.3       0.0       0.0
 5 C  dyz             0.0       0.0       0.8       0.0       1.8       0.0
 5 C  dx2y2           0.0       0.0       0.0       0.0       0.0       0.7
 5 C  dxy             0.4       0.0       0.0       0.0       0.0       0.0
 6 H  s               7.4       0.0       0.0       0.0       0.0      11.5
 7 H  s               7.4       0.0       0.0       0.0       0.0      11.5
 8 H  s               0.0       0.0       0.0       0.0       0.0      11.5
 9 H  s               7.4       0.0       0.0       0.0       0.0      11.5
10 H  s               7.4       0.0       0.0       0.0       0.0      11.5
11 H  s               0.0       0.0       0.0       0.0       0.0      11.5
```

We see that the occupied π-orbitals number 16, 19, 20 and the unoccupied ones start with 21 and 22. However, the sixth high-lying π^*-orbital cannot easily be found. Thus, let us run a simple selected CEPA/2 calculation and look at the natural orbitals.

```
! RHF SV(P)
! moread
%moinp "Test-CASSCF-Benzene-1.gbw"

%mrci citype cepa2
      tsel 1e-5
      natorbiters 1
      newblock 1 *
        nroots 1
        refs cas(0,0) end
        end
      end
# ...etc, input of coordinates
```

The calculation prints the occupation numbers:

```
N[  6] =   1.98784765
N[  7] =   1.98513069
N[  8] =   1.98508633
N[  9] =   1.97963799
N[ 10] =   1.97957039
N[ 11] =   1.97737886
N[ 12] =   1.97509724
N[ 13] =   1.97370616
N[ 14] =   1.97360821
N[ 15] =   1.96960145
N[ 16] =   1.96958645
N[ 17] =   1.96958581
N[ 18] =   1.95478929
```

(continues on next page)

(continued from previous page)

```
N[ 19] =   1.91751184
N[ 20] =   1.91747498
N[ 21] =   0.07186879
N[ 22] =   0.07181758
N[ 23] =   0.03203528
N[ 24] =   0.01766832
N[ 25] =   0.01757735
N[ 26] =   0.01708578
N[ 27] =   0.01707675
N[ 28] =   0.01671912
N[ 29] =   0.01526139
N[ 30] =   0.01424982
```

From these occupation number it becomes evident that there are several natural orbitals which are not quite doubly occupied MOs. Those with an occupation number of 1.95 and less should certainly be taken as active. In addition the rather strongly occupied virtual MOs 21-23 should also be active leading to CASSCF(6,6). Let us see what these orbitals are before starting CASSCF:

```
! RHF SV(P)
! moread noiter
%moinp "Test-CASSCF-Benzene-2.mrci.nat"
```

Leading to:

```
                     18        19        20        21        22        23
                   1.00000   1.00000   1.00000   1.00000   1.00000   1.00000
                   1.95479   1.91751   1.91747   0.07187   0.07182   0.03204
                  --------  --------  --------  --------  --------  --------
 0 C  pz             16.5       8.1      24.4      23.4       7.8      16.1
 0 C  dxz             0.0       0.4       0.2       0.6       0.9       0.1
 0 C  dyz             0.1       0.2       0.0       0.0       0.6       0.4
 1 C  pz             16.5       8.1      24.4      23.5       7.8      16.1
 1 C  dxz             0.0       0.4       0.2       0.6       0.9       0.1
 1 C  dyz             0.1       0.2       0.0       0.0       0.6       0.4
 2 C  pz             16.5      32.5       0.0       0.0      31.3      16.3
 2 C  dxz             0.1       0.1       0.0       0.0       0.2       0.5
 2 C  dyz             0.0       0.0       0.8       1.9       0.0       0.0
 3 C  pz             16.5       8.1      24.4      23.4       7.8      16.1
 3 C  dxz             0.0       0.4       0.2       0.6       0.9       0.1
 3 C  dyz             0.1       0.2       0.0       0.0       0.6       0.4
 4 C  pz             16.5       8.1      24.4      23.5       7.8      16.1
 4 C  dxz             0.0       0.4       0.2       0.6       0.9       0.1
 4 C  dyz             0.1       0.2       0.0       0.0       0.6       0.4
 5 C  pz             16.5      32.5       0.0       0.0      31.3      16.3
 5 C  dxz             0.1       0.1       0.0       0.0       0.2       0.5
 5 C  dyz             0.0       0.0       0.8       1.9       0.0       0.0
```

This shows us that these six orbitals are precisely the π/π^* orbitals that we wanted to have active (you can also plot them to get even more insight).

Now we know that the desired orbitals are in the correct order, we can do CASSCF:

```
! SV(P)
! moread
%moinp "Test-CASSCF-Benzene-2.mrci.nat"

%casscf  nel    6
         norb   6
         nroots 1
         mult   1
         switchstep nr # For illustration purpose
```

(continues on next page)

(continued from previous page)

```
        end
```

To highlight the feature `SwitchStep` of the CASSCF program, we employ the Newton-Raphson method (NR) after a certain convergence has been reached (`SwitchStep NR` statement). In general, it is not recommended to change the default convergence settings! The output of the CASSCF program is:

```
------------------
CAS-SCF ITERATIONS
------------------


MACRO-ITERATION   1:
--- Inactive Energy E0 = -224.09725414 Eh
CI-ITERATION   0:
-230.588253032   0.000000000000 (    0.00)
CI-PROBLEM SOLVED
DENSITIES MADE

<<<<<<<<<<<<<<<<<<INITIAL CI STATE CHECK>>>>>>>>>>>>>>>>>>

BLOCK  1 MULT= 1 NROOTS= 1
ROOT   0:  E=    -230.5882530315 Eh
0.89482 [     0]: 222000
0.02897 [    14]: 211110
0.01982 [    29]: 202020
0.01977 [     4]: 220200
0.01177 [    65]: 112011
0.01169 [    50]: 121101

<<<<<<<<<<<<<<<<<<INITIAL CI STATE CHECK>>>>>>>>>>>>>>>>>>

E(CAS)=  -230.588253032 Eh DE=    0.000000e+00
--- Energy gap subspaces: Ext-Act = 0.195   Act-Int = 0.127
--- current l-shift: Up(Ext-Act) = 1.40   Dn(Act-Int) = 1.47
N(occ)=  1.96393 1.90933 1.90956 0.09190 0.09208 0.03319
||g|| =     1.046979e-01 Max(G)=   -4.638985e-02 Rot=53,19
--- Orbital Update [SuperCI(PT)]
--- Canonicalize Internal Space
--- Canonicalize External Space
--- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.063973050 Max(X)(83,23) =     -0.035491133
--- SFit(Active Orbitals)

MACRO-ITERATION   2:
--- Inactive Energy E0 = -224.09299157 Eh
CI-ITERATION   0:
-230.590141151   0.000000000000 (    0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)=  -230.590141151 Eh DE=   -1.888119e-03
--- Energy gap subspaces: Ext-Act = 0.202   Act-Int = 0.126
--- current l-shift: Up(Ext-Act) = 0.90   Dn(Act-Int) = 0.97
N(occ)=  1.96182 1.90357 1.90364 0.09771 0.09777 0.03549
||g|| =     2.971340e-02 Max(G)=   -8.643429e-03 Rot=52,20
--- Orbital Update [SuperCI(PT)]
--- Canonicalize Internal Space
--- Canonicalize External Space
--- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.009811159 Max(X)(67,21) =     -0.003665750
--- SFit(Active Orbitals)

MACRO-ITERATION   3:
===>>> Convergence to 3.0e-02 achieved - switching to Step=NR
```

(continues on next page)

(continued from previous page)

```
--- Inactive Energy E0 = -224.07872151 Eh
CI-ITERATION   0:
-230.590260496   0.000000000000 (    0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)=  -230.590260496 Eh DE=   -1.193453e-04
--- Energy gap subspaces: Ext-Act = 0.203   Act-Int = 0.125
--- current l-shift: Up(Ext-Act) = 0.73   Dn(Act-Int) = 0.81
N(occ)=  1.96145 1.90275 1.90278 0.09856 0.09857 0.03589
||g|| =     8.761362e-03 Max(G)=    4.388664e-03 Rot=43,19
--- Orbital Update [        NR]
AUGHESS-ITER   0: E=    -0.000016434 <r|r>= 2.70127912e-05
AUGHESS-ITER   1: E=    -0.000021148 <r|r>= 2.91399830e-06
AUGHESS-ITER   2: E=    -0.000021780 <r|r>= 4.01336069e-07 => CONVERGED
DE(predicted)=   -0.000010890 First Element= 0.999987718
<X(rot)|X(rot)>=      0.000024564
--- SFit(Active Orbitals)

MACRO-ITERATION   4:
--- Inactive Energy E0 = -224.07787812 Eh
CI-ITERATION   0:
-230.590271490   0.000000000000 (    0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)=  -230.590271490 Eh DE=   -1.099363e-05
--- Energy gap subspaces: Ext-Act = 0.202   Act-Int = 0.125
--- current l-shift: Up(Ext-Act) = 0.40   Dn(Act-Int) = 0.47
N(occ)=  1.96135 1.90267 1.90267 0.09866 0.09866 0.03599
||g|| =     6.216730e-04 Max(G)=    1.417079e-04 Rot=66,13
---- THE CAS-SCF GRADIENT HAS CONVERGED ----
--- FINALIZING ORBITALS ---
---- DOING ONE FINAL ITERATION FOR PRINTING ----
--- Forming Natural Orbitals
--- Canonicalize Internal Space
--- Canonicalize External Space

MACRO-ITERATION   5:
--- Inactive Energy E0 = -224.07787811 Eh
--- All densities will be recomputed
CI-ITERATION   0:
-230.590271485   0.000000000000 (    0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)=  -230.590271485 Eh DE=    5.179942e-09
--- Energy gap subspaces: Ext-Act = -0.242   Act-Int = -0.002
--- current l-shift: Up(Ext-Act) = 0.84   Dn(Act-Int) = 0.60
N(occ)=  1.96135 1.90267 1.90267 0.09866 0.09866 0.03599
||g|| =     6.216710e-04 Max(G)=    1.544017e-04 Rot=29,12
--------------
CASSCF RESULTS
--------------

Final CASSCF energy       : -230.590271485 Eh   -6274.6803 eV
```

First of all you can see how the program cycles between CI-vector optimization and orbital optimization steps (so-called unfolded two-step procedure). After 3 iterations, the program switches to the Newton-Raphson solver which then converges very rapidly. Orbital optimization with the Newton-Raphson solver is limited to smaller sized molecules, as the program produces lengthy integrals and Hessian files. In the majority of situations the default converger (SuperCI(PT)) is the preferred choice.[459]

Atomic Valence Active Space

Very good starting orbitals that are targeted to a specific user-given active space can be generated with the Atomic Valence Active Space (AVAS) procedure. [751, 752] The general idea is that the user provides a set of atomic orbitals (AO) of a minimal basis set that are sufficient to qualitatively represent the final CASSCF active orbitals. Typical examples are

- p_z orbitals of a π system chromophore in a molecule
- five valence (or 10 double-shell) d orbitals of a transition-metal (TM) atom in a molecule
- seven valence (or 14 double-shell) f orbitals of a lanthanide or actinide atom in a molecule

Then, by the help of linear algebra (singular-value decomposition) AVAS rotates the starting molecular orbitals (MOs) such that they have maximum overlap with the target AOs. With those rotated MOs that have a sufficiently large singular value (> 0.4 (default)) are considered as active orbitals. In that manner, AVAS can automatically determine an active space, i.e. the number of active orbitals and electrons, that is now specified by the target AOs.

As a first example, we now consider $CuCl_4^-$ in a minimal active space

```
! cc-pvtz TightSCF Def2/JK PModel NoIter
! AVAS(Valence-D)

%maxcore 3000

%paras
 cucl  = 2.291
end

* int -1 1
Cu  0  0  0   0.0     0.0     0.0
Cl   1  0  0   {cucl}  0.0      0.0
Cl   1  2  0   {cucl}  90.0     0.0
Cl   1  3  2   {cucl}  90.0   180.
Cl   1  4  3   {cucl}  90.0   180.
*
```

The keyword `! AVAS(Valence-D)` seeks for all transition-metal atoms in the molecule and inserts a single minimal d basis function for each TM atom. All five component M_L of the basis function are then considered. The AVAS procedure prints singular / eigen values for the occupied and virtual orbital space and easily finds the desired minimal active space CAS(9,5).

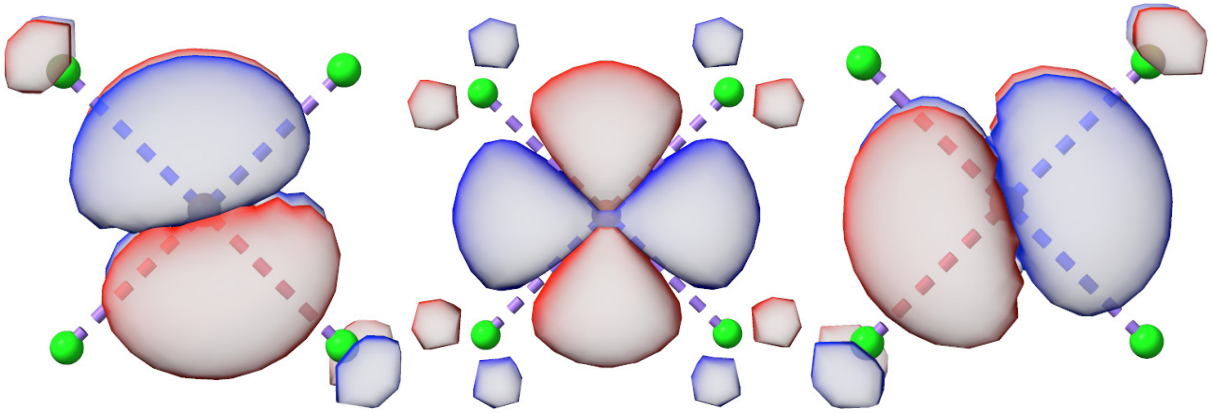
```
-------------------------------------------------
INITIAL GUESS: Atomic Valence Active space (AVAS)
-------------------------------------------------

AVAS threshold         : 0.400000
AVAS minimal basis set : MINAO
AVAS list              : Shell | 3 2  0> at atom 0 (system 0)
 \\\\\: Shell | 3 2  1> at atom 0 (system 0)\\\\\: Shell | 3 2 -1> at atom 0 (system 0)\\\\\:␣
↪Shell | 3 2  2> at atom 0 (system 0)\\\\\: Shell | 3 2 -2> at atom 0 (system 0)

AVAS P matrix eig. val (  Occupied) : 0.966698
AVAS P matrix eig. val (  Occupied) : 0.974913
AVAS P matrix eig. val (  Occupied) : 0.977443
AVAS P matrix eig. val (  Occupied) : 0.977443
AVAS P matrix eig. val (  Occupied) : 0.985233
AVAS P matrix eig. val (   Virtual) : (0.032829)
AVAS P matrix eig. val (   Virtual) : (0.024687)
AVAS P matrix eig. val (   Virtual) : (0.022047)
AVAS P matrix eig. val (   Virtual) : (0.022047)
AVAS P matrix eig. val (   Virtual) : (0.014546)
```

(continues on next page)

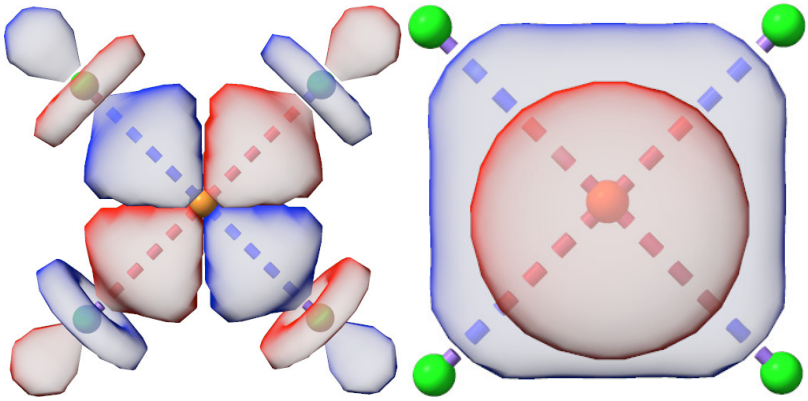
(continued from previous page)

```
AVAS electrons          : 9
AVAS orbitals           : 5
```

The five initial active orbitals after being processed by AVAS indeed look like the desired Cu d-orbitals.



(a) $\sigma^o{}_1 = 0.977$ (b) $\sigma^o{}_3 = 0.975$ (c) $\sigma^o{}_0 = 0.985$



(d) $\sigma^o{}_4 = 0.967$ (e) $\sigma^o{}_2 = 0.977$

Fig. 6.11: Initial Minimal AS orbitals of $CuCl_4^-$ generated by AVAS.

The same calculation can be started also by using the `%scf avas ... end end` block.

```
%scf
 avas
  system
   shell   3,  3,  3,  3,  3
   l       2,  2,  2,  2,  2
   m_l     0,  1, -1,  2, -2
   center  0,  0, 0,  0, 0
  end
 end
end
```

Here, it is also possible to use target basis functions at different atoms (`center`) and to select only a subset of functions in a shell (m_l). Note that if not all functions of a shell (3p, 5d, 7f) are selected, the molecule should be oriented manually to accomplish the desired basis function overlap.

AVAS can be also used very conveniently in the same fashion for double-d shell calculations with transition-metal complexes (`! AVAS(Double-D)`). For each 3d transition-metal center in a molecule all 3d and 4d target functions are considered. Similarly, double-shell active spaces can be also set up for 4d and 5d transition-metal complexes.

There is also a similar keyword for lanthanides and actinides. `! AVAS( Valence-F )` attempts to set up an active space with 7 f functions for each lanthanide or actinide atom in a molecule. There is also the possibility to

run double-f shell calculations using the `! AVAS( Double-F )` keyword.

To avoid this issue for π active space calculation, all three 2p target AOs are considered first but they are weighted by the three component of the principle axis of inertia with the largest moment. [751] For those inertia moment calculations, masses are ignored and only the centers of the desired target p AO are considered.

For a CAS(10,9) π-active space calculation on tryptophan, the AVAS input read

```
! cc-pVTZ PModel NoIter

%scf
 avas
  tol 0.4
  system
   center 0, 1, 2, 3, 4, 5, 10, 11, 12
   type  pz, pz, pz, pz, pz, pz, pz, pz, pz
  end
 end
end

* xyz 0 1
 C          0.4512549872        2.3796411953        0.0577773122
 C          0.1094760583        1.0035547288       -0.1566676092
 C          1.7801675822        2.8072137170        0.2571892289
 C          2.7806901872        1.8262977582        0.2356692574
 C          2.4656511421        0.4546661378        0.0230301052
 C          1.1452272475        0.0339609344       -0.1736410480
 H          2.0259509453        3.8615102004        0.4187388370
 H          3.8237760997        2.1214062744        0.3833595222
 H          3.2752609035       -0.2812140701        0.0109117373
 H          0.9203743659       -1.0244858148       -0.3388820373
 C         -1.3215206968        0.9316755285       -0.3177965838
 C         -1.7965156128        2.2386249398       -0.2022300378
 N         -0.7296902726        3.0958334808        0.0227806512
 H         -0.8107596679        4.0971334562        0.1485860796
 H         -2.8167763088        2.6080109542       -0.2688439980
 C         -2.1029028025       -0.3291635000       -0.5688909937
 C         -3.4238543678       -0.4065989881        0.2267575199
 H         -1.4745479852       -1.1909157350       -0.2884954137
 H         -2.3461010681       -0.4478113906       -1.6421457333
 C         -3.9423325138       -1.8379287141        0.1258142785
 N         -4.3742952299        0.5836598444       -0.2812451173
 H         -3.2051519657       -0.1892488262        1.2846794690
 O         -3.2924970778       -2.6708957465        0.9924074621
 O         -4.8043368378       -2.2232843366       -0.6488988164
 H         -3.6480373076       -3.5631013900        0.8277551234
 H         -5.2270970579        0.5578136152        0.2816027849
 H         -4.6658127461        0.2911757460       -1.2180819802
*
```

and leads to the following output

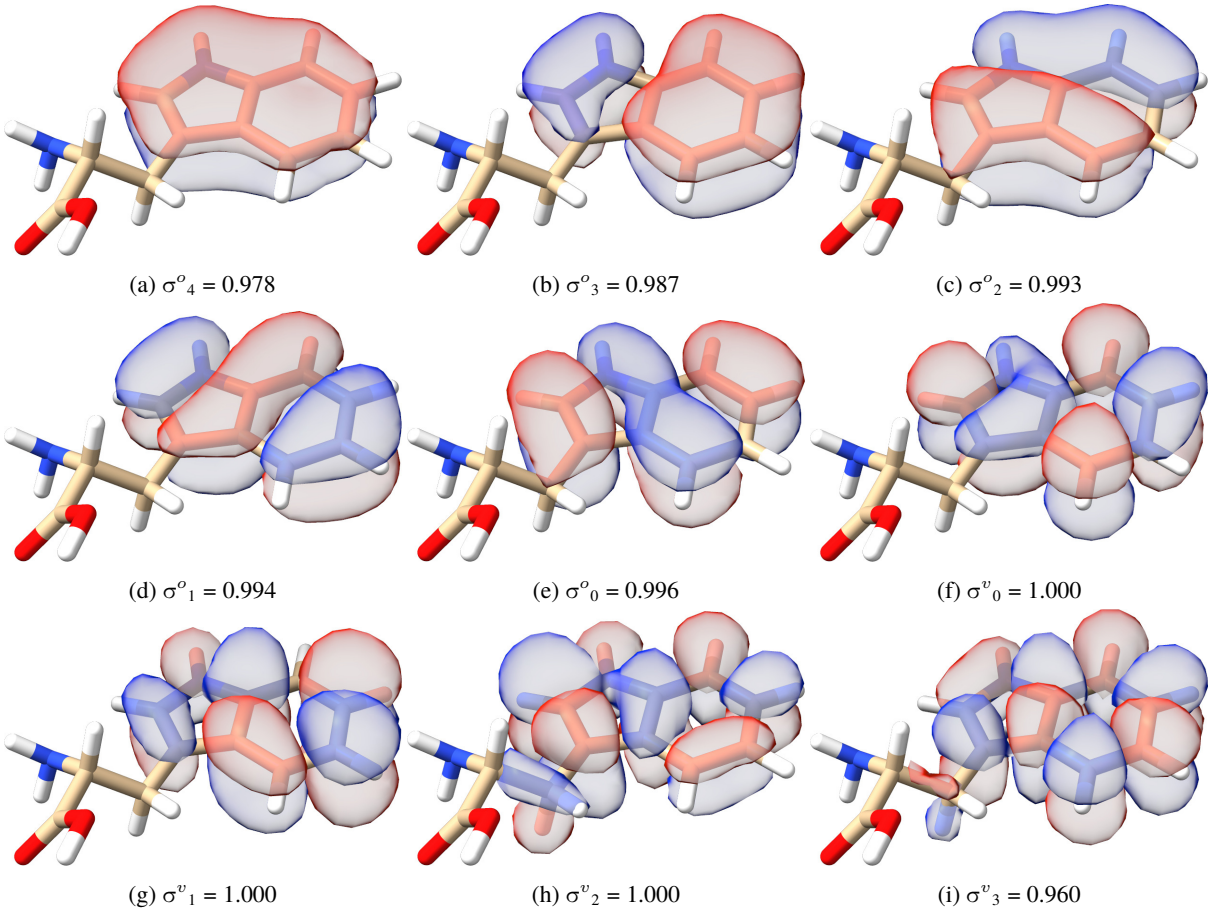
```
------------------------------------------------
INITIAL GUESS: Atomic Valence Active space (AVAS)
------------------------------------------------

   AVAS threshold         : 0.400000
   AVAS minimal basis set : AUTO
   AVAS list              : Shell | 2 1  0> at atom 0 (system 0, type pz)
                          : Shell | 2 1  1> at atom 0 (system 0, type pz)
                          : Shell | 2 1 -1> at atom 0 (system 0, type pz)
                          : Shell | 2 1  0> at atom 1 (system 0, type pz)
                          : Shell | 2 1  1> at atom 1 (system 0, type pz)
```

(continues on next page)

(continued from previous page)

```
                        : Shell | 2 1 -1> at atom 1 (system 0, type pz)
                        : Shell | 2 1  0> at atom 2 (system 0, type pz)
                        : Shell | 2 1  1> at atom 2 (system 0, type pz)
                        : Shell | 2 1 -1> at atom 2 (system 0, type pz)
                        : Shell | 2 1  0> at atom 3 (system 0, type pz)
                        : Shell | 2 1  1> at atom 3 (system 0, type pz)
                        : Shell | 2 1 -1> at atom 3 (system 0, type pz)
                        : Shell | 2 1  0> at atom 4 (system 0, type pz)
                        : Shell | 2 1  1> at atom 4 (system 0, type pz)
                        : Shell | 2 1 -1> at atom 4 (system 0, type pz)
                        : Shell | 2 1  0> at atom 5 (system 0, type pz)
                        : Shell | 2 1  1> at atom 5 (system 0, type pz)
                        : Shell | 2 1 -1> at atom 5 (system 0, type pz)
                        : Shell | 2 1  0> at atom 10 (system 0, type pz)
                        : Shell | 2 1  1> at atom 10 (system 0, type pz)
                        : Shell | 2 1 -1> at atom 10 (system 0, type pz)
                        : Shell | 2 1  0> at atom 11 (system 0, type pz)
                        : Shell | 2 1  1> at atom 11 (system 0, type pz)
                        : Shell | 2 1 -1> at atom 11 (system 0, type pz)
                        : Shell | 2 1  0> at atom 12 (system 0, type pz)
                        : Shell | 2 1  1> at atom 12 (system 0, type pz)
                        : Shell | 2 1 -1> at atom 12 (system 0, type pz)

  AVAS P matrix eig. val (  Occupied) : (0.000004)
  AVAS P matrix eig. val (  Occupied) : (0.000014)
  AVAS P matrix eig. val (  Occupied) : (0.000292)
  AVAS P matrix eig. val (  Occupied) : (0.040014)
  AVAS P matrix eig. val (  Occupied) : 0.978162
  AVAS P matrix eig. val (  Occupied) : 0.986637
  AVAS P matrix eig. val (  Occupied) : 0.993225
  AVAS P matrix eig. val (  Occupied) : 0.994300
  AVAS P matrix eig. val (  Occupied) : 0.996447
  AVAS P matrix eig. val (   Virtual) : 0.999996
  AVAS P matrix eig. val (   Virtual) : 0.999986
  AVAS P matrix eig. val (   Virtual) : 0.999708
  AVAS P matrix eig. val (   Virtual) : 0.959986
  AVAS P matrix eig. val (   Virtual) : (0.021838)
  AVAS P matrix eig. val (   Virtual) : (0.013363)
  AVAS P matrix eig. val (   Virtual) : (0.006775)
  AVAS P matrix eig. val (   Virtual) : (0.005700)
  AVAS P matrix eig. val (   Virtual) : (0.003553)

  AVAS electrons         : 10
  AVAS orbitals          : 9

                     ------------------
                     INITIAL GUESS DONE (   0.3 sec)
                     ------------------
```

and initial active orbitals.



(a) $\sigma^o{}_4 = 0.978$ (b) $\sigma^o{}_3 = 0.987$ (c) $\sigma^o{}_2 = 0.993$

(d) $\sigma^o{}_1 = 0.994$ (e) $\sigma^o{}_0 = 0.996$ (f) $\sigma^v{}_0 = 1.000$

(g) $\sigma^v{}_1 = 1.000$ (h) $\sigma^v{}_2 = 1.000$ (i) $\sigma^v{}_3 = 0.960$

Fig. 6.12: Initial π AS orbitals of tryptophan generated by AVAS.

It is also possible to specify the number of active electrons `nel` and orbitals `norb` directly. For such a calculation, the AVAS singular value decomposition threshold `tol` is ignored. In the following calculation, the strongly occupied orbital from the previous CAS(10,9) (σ^o_4 in Fig. 6.12) calculation is omitted.

```
%scf
 avas
  system
   norb 8
   nel 8
   center 0, 1, 2, 3, 4, 5, 10, 11, 12
   type  pz, pz, pz, pz, pz, pz, pz, pz, pz
  end
 end
end
```

It is also possible to do the AVAS start MO generation for several `systems` independently and then re-orthonormalize all MOs at the end similar to [751]. This becomes interesting for generating starting orbitals for multiple π chromophores like the bridged bithiophene biradical

```
%scf
 avas
  system
   norb 4
   nel  3
   center   0, 1, 2, 3, 4 # C / S atoms system 1
   type    pz, pz, pz, pz, pz
  end
  system
```

(continues on next page)

(continued from previous page)

```
   norb 4
   nel  5
   center  38, 39, 40, 41, 43 # C / S atoms system 2
   type    pz, pz, pz, pz, pz
  end
 end
end
```

or the FeTPP molecule.

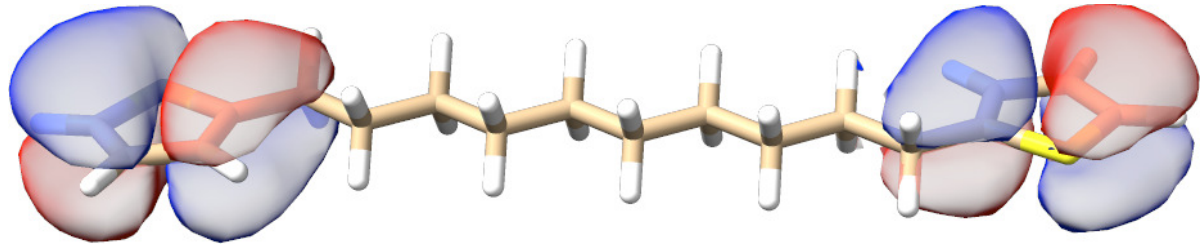
```
! SVP NoIter
! PModel

%scf
 avas
  system
   center 0
   type d
  end
  system
   center  1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 29,␣
↪30, 31, 32
   type    pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz,␣
↪pz, pz, pz
  end
 end
end

*xyz 0 1
Fe 0.0000 0.0000  0.0000
N  1.9764 0.0000  0.0000
N  0.0000 0.0000  1.9884
N -1.9764 0.0000  0.0000
N  0.0000 0.0000 -1.9884
C  2.8182 0.0000 -1.0903
C  2.8182 0.0000  1.0903
C  1.0918 0.0000  2.8249
C -1.0918 0.0000  2.8249
C -2.8182 0.0000  1.0903
C -2.8182 0.0000 -1.0903
C -1.0918 0.0000 -2.8249
C  1.0918 0.0000 -2.8249
C  4.1961 0.0000 -0.6773
C  4.1961 0.0000  0.6773
C  0.6825 0.0000  4.1912
C -0.6825 0.0000  4.1912
C -4.1961 0.0000  0.6773
C -4.1961 0.0000 -0.6773
C -0.6825 0.0000 -4.1912
C  0.6825 0.0000 -4.1912
H  5.0441 0.0000 -1.3538
H  5.0441 0.0000  1.3538
H  1.3558 0.0000  5.0416
H -1.3558 0.0000  5.0416
H -5.0441 0.0000  1.3538
H -5.0441 0.0000 -1.3538
H -1.3558 0.0000 -5.0416
H  1.3558 0.0000 -5.0416
C  2.4150 0.0000  2.4083
C -2.4150 0.0000  2.4083
C -2.4150 0.0000 -2.4083
```

(continues on next page)

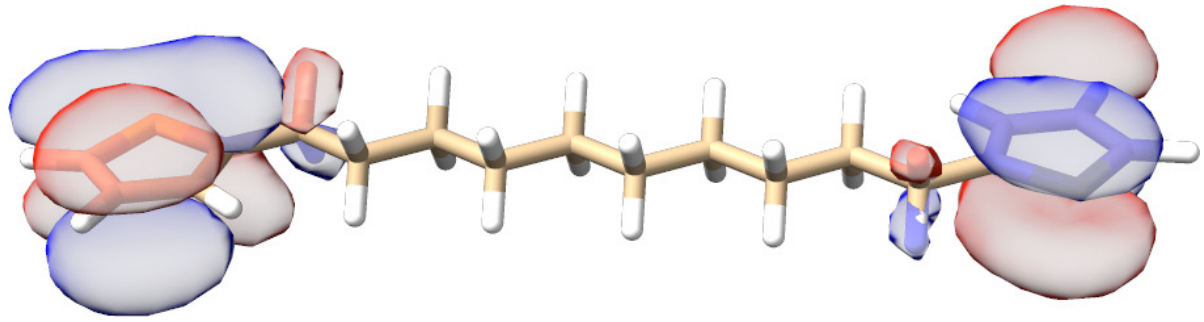
(continued from previous page)

```
C  2.4150 0.0000 -2.4083
H  3.1855 0.0000  3.1752
H -3.1855 0.0000  3.1752
H -3.1855 0.0000 -3.1752
H  3.1855 0.0000 -3.1752
*
```

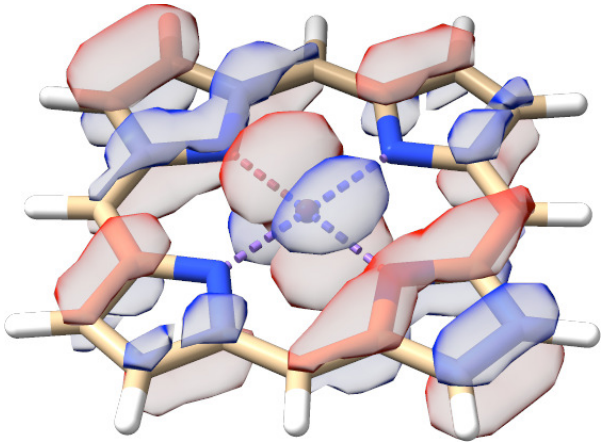
For those systems, all AVAS starting MOs have the desired π or d character as illustrated for the "active frontier orbitals" in Fig. 6.13.



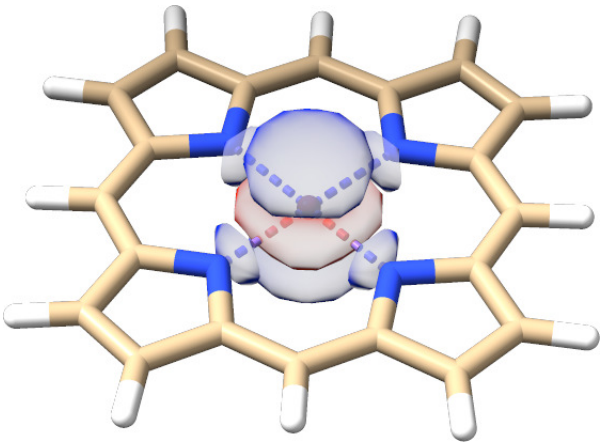
(a) AVAS HOMO



(b) AVAS LUMO



(c) AVAS HOMO



(d) AVAS LUMO

Fig. 6.13: Initial HOMO and LUMO AVAS orbitals of a bridged bithiophene biradical and FeTPP.

CASSCF and Symmetry

The CASSCF program can make some use of symmetry. Thus, it is possible to do the CI calculations separated by irreducible representations. This allows one to calculate electronic states in a more controlled fashion.

Let us look at a simple example: C_2H_4. We first generate symmetry adapted MP2 natural orbitals. Since we opt for initial guess orbitals, the computationally cheaper unrelaxed density suffices:

```
! def2-TZVP def2-TZVP/C UseSym RI-MP2 conv # conventional is faster for small molecules
%mp2
density unrelaxed
natorbs true
end
* int 0 1
C 0 0 0  0 0 0
C 1 0 0  1.35 0 0
H 1 2 0  1.1 120 0
H 1 2 3  1.1 120 180
H 2 1 3  1.1 120 0
H 2 1 3  1.1 120 180
*
```

The program does the following. It first identifies the group correctly as D_{2h} and sets up its irreducible representations. The process detects symmetry within `SymThresh` (10^{-4}) and purifies the geometry thereafter:

```
------------------
SYMMETRY DETECTION
------------------
The point group will now be determined using a tolerance of 1.0000e-04.
Splitting atom subsets according to nuclear charge, mass and basis set.
Splitting atom subsets according to distance from the molecule's center.
Identifying relative distance patterns of the atoms.
Splitting atom subsets according to atoms' relative distance patterns.
Bring atoms of each subset into input order.
The molecule is planar.
The molecule has a center of inversion.
Analyzing the first atom subset for its symmetry.
The atoms in the selected subset form a 4-gon with alternating side lengths.
Testing point group D2h.
Success!
This point group has been found:    D2h
Largest non-degenerate subgroup:    D2h


Symmetry-perfected Cartesians (point group D2h):

Atom          Symmetry-perfected Cartesians (x, y, z; au)
  0      -1.275565140397    0.000000000000    0.000000000000
  1       1.275565140397    0.000000000000    0.000000000000
  2      -2.314914514054    1.800205921988    0.000000000000
  3      -2.314914514054   -1.800205921988    0.000000000000
  4       2.314914514054    1.800205921988    0.000000000000
  5       2.314914514054   -1.800205921988    0.000000000000


------------------------------------------------
SYMMETRY-PERFECTED CARTESIAN COORDINATES (A.U.)
------------------------------------------------
Warning (ORCA_SYM): Coordinates were not cleaned so far!

------------------
SYMMETRY REDUCTION
```

(continues on next page)

(continued from previous page)

```
------------------
ORCA supports only abelian point groups.
It is now checked, if the determined point group is supported:
Point Group ( D2h   ) is          ... supported

(Re)building abelian point group:
Creating Character Table          ... done
Making direct product table       ... done
Constructing symmetry operations  ... done
Creating atom transfer table      ... done
Creating asymmetric unit          ... done

----------------------
ASYMMETRIC UNIT IN D2h
----------------------
  #  AT     MASS              COORDS (A.U.)              BAS
   0 C   12.0110  -1.27556514   0.00000000   0.00000000   0
   2 H    1.0080  -2.31491451   1.80020592   0.00000000   0

----------------------
SYMMETRY ADAPTED BASIS
----------------------
The coefficients for the symmetry adapted linear combinations (SALCS)
of basis functions will now be computed:
Number of basis functions         ...    86
Preparing memory                  ... done
Constructing Gamma(red)           ... done
Reducing Gamma(red)               ... done
Constructing SALCs                ... done
Checking SALC integrity           ... nothing suspicious
Normalizing SALCs                 ... done

Storing the symmetry object:
Symmetry file                     ... Test-SYM-CAS-C2H4-1.sym.tmp
Writing symmetry information      ... done
```

It then performs the SCF calculation and keeps the symmetry in the molecular orbitals.

```
NO   OCC          E(Eh)            E(eV)    Irrep
   0   2.0000     -11.236728      -305.7669    1-Ag
   1   2.0000     -11.235157      -305.7242    1-B3u
   2   2.0000      -1.027144       -27.9500    2-Ag
   3   2.0000      -0.784021       -21.3343    2-B3u
   4   2.0000      -0.641566       -17.4579    1-B2u
   5   2.0000      -0.575842       -15.6694    3-Ag
   6   2.0000      -0.508313       -13.8319    1-B1g
   7   2.0000      -0.373406       -10.1609    1-B1u
   8   0.0000       0.139580         3.7982    1-B2g
   9   0.0000       0.171982         4.6799    4-Ag
  10   0.0000       0.195186         5.3113    3-B3u
  11   0.0000       0.196786         5.3548    2-B2u
  12   0.0000       0.242832         6.6078    2-B1g
  13   0.0000       0.300191         8.1686    5-Ag
  14   0.0000       0.326339         8.8801    4-B3u
... etc
```

The MP2 module does not take any advantage of this information but produces natural orbitals that are symmetry adapted:

```
N[  0](B3u) =   2.00000360
N[  1]( Ag) =   2.00000219
```

(continues on next page)

(continued from previous page)

```
N[  2]( Ag) =   1.98056435
N[  3](B3u) =   1.97195041
N[  4](B2u) =   1.96746753
N[  5](B1g) =   1.96578954
N[  6]( Ag) =   1.95864726
N[  7](B1u) =   1.93107098
N[  8](B2g) =   0.04702701
N[  9](B3u) =   0.02071784
N[ 10](B2u) =   0.01727252
N[ 11]( Ag) =   0.01651489
N[ 12](B1g) =   0.01602695
N[ 13](B3u) =   0.01443373
N[ 14](B1u) =   0.01164204
N[ 15]( Ag) =   0.01008617
N[ 16](B2u) =   0.00999302
N[ 17]( Ag) =   0.00840326
N[ 18](B3g) =   0.00795053
N[ 19](B3u) =   0.00532044
N[ 20]( Au) =   0.00450556
etc.
```

From this information and visual inspection you will know what orbitals you will have in the active space:

These natural orbitals can then be fed into the CASSCF calculation. We perform a simple calculation in which we keep the ground state singlet (A_{1g} symmetry, irrep=0) and the first excited triplet state (B_{3u} symmetry, irrep=7). In general the ordering of irreps follows standard conventions and in case of doubt you will find the relevant number for each irrep in the output.

For example, here (using LargePrint):

```
----------------------------
CHARACTER TABLE OF GROUP D2h
----------------------------
GAMMA  O1   O2   O3   O4   O5   O6   O7   O8
 Ag :  1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0
 B1g:  1.0  1.0 -1.0 -1.0  1.0  1.0 -1.0 -1.0
 B2g:  1.0 -1.0  1.0 -1.0  1.0 -1.0  1.0 -1.0
 B3g:  1.0 -1.0 -1.0  1.0  1.0 -1.0 -1.0  1.0
 Au :  1.0  1.0  1.0  1.0 -1.0 -1.0 -1.0 -1.0
 B1u:  1.0  1.0 -1.0 -1.0 -1.0 -1.0  1.0  1.0
 B2u:  1.0 -1.0  1.0 -1.0 -1.0  1.0 -1.0  1.0
 B3u:  1.0 -1.0 -1.0  1.0 -1.0  1.0  1.0 -1.0

---------------------------------
DIRECT PRODUCT TABLE OF GROUP D2h
---------------------------------
   **   Ag  B1g B2g B3g Au  B1u B2u B3u

 Ag     Ag  B1g B2g B3g Au  B1u B2u B3u
 B1g    B1g Ag  B3g B2g B1u Au  B3u B2u
 B2g    B2g B3g Ag  B1g B2u B3u Au  B1u
 B3g    B3g B2g B1g Ag  B3u B2u B1u Au
 Au     Au  B1u B2u B3u Ag  B1g B2g B3g
 B1u    B1u Au  B3u B2u B1g Ag  B3g B2g
 B2u    B2u B3u Au  B1u B2g B3g Ag  B1g
 B3u    B3u B2u B1u Au  B3g B2g B1g Ag
```

We use the following input for CASSCF, where we tightened the integral cut-offs and the convergence criteria using `!VeryTightSCF`.

```
! def2-TZVP  Conv NormalPrint UseSym
! moread
```

(continues on next page)

(continued from previous page)

```
%moinp "Test-SYM-CAS-C2H4-1.mp2nat"
%casscf nel            4
        norb           4
        # This is only here to show that NR can also be used from
        # the start with orbstep
        orbstep    nr
        switchstep nr
        # the lowest singet and triplet states. The new feature
        # is the array "irrep" that lets you give the irrep for
        # a given block. Thus, now you can have several blocks of
        # the same multiplicity but different spatial symmetry
        irrep       0,7
        mult        1,3
        nroots      1,1
        end

* int 0 1
C 0 0 0  0 0 0
C 1 0 0  1.35 0 0
H 1 2 0  1.1 120 0
H 1 2 3  1.1 120 180
H 2 1 3  1.1 120 0
H 2 1 3  1.1 120 180
*
```

And gives:

```
------------
SCF SETTINGS
------------
Hamiltonian:
 Ab initio Hamiltonian  Method          .... Hartree-Fock(GTOs)


General Settings:
 Integral files         IntName         .... Test-SYM-CAS-C2H4-1
 Hartree-Fock type      HFTyp           .... CASSCF
 Total Charge           Charge          ....    0
 Multiplicity           Mult            ....    1
 Number of Electrons    NEL             ....   16
 Basis Dimension        Dim             ....   86
 Nuclear Repulsion      ENuc            ....     32.9609050695 Eh

 Symmetry handling      UseSym         .... ON
 Point group                           .... D2h
 Used point group                      .... D2h
 Number of irreps                      .... 8
   Irrep   Ag has   19 symmetry adapted basis functions (ofs=   0)
   Irrep  B1g has   12 symmetry adapted basis functions (ofs=  19)
   Irrep  B2g has    8 symmetry adapted basis functions (ofs=  31)
   Irrep  B3g has    4 symmetry adapted basis functions (ofs=  39)
   Irrep   Au has    4 symmetry adapted basis functions (ofs=  43)
   Irrep  B1u has    8 symmetry adapted basis functions (ofs=  47)
   Irrep  B2u has   12 symmetry adapted basis functions (ofs=  55)
   Irrep  B3u has   19 symmetry adapted basis functions (ofs=  67)
```

And further in the CASSCF program:

```
Symmetry handling      UseSym     ... ON
Point group                       ... D2h
Used point group                  ... D2h
```

(continues on next page)

(continued from previous page)

```
Number of irreps                  ... 8
   Irrep   Ag has   19 SALCs (ofs=   0) #(closed)=   2 #(active)=   1
   Irrep  B1g has   12 SALCs (ofs=  19) #(closed)=   1 #(active)=   0
   Irrep  B2g has    8 SALCs (ofs=  31) #(closed)=   0 #(active)=   1
   Irrep  B3g has    4 SALCs (ofs=  39) #(closed)=   0 #(active)=   0
   Irrep   Au has    4 SALCs (ofs=  43) #(closed)=   0 #(active)=   0
   Irrep  B1u has    8 SALCs (ofs=  47) #(closed)=   0 #(active)=   1
   Irrep  B2u has   12 SALCs (ofs=  55) #(closed)=   1 #(active)=   0
   Irrep  B3u has   19 SALCs (ofs=  67) #(closed)=   2 #(active)=   1
 Symmetries of active orbitals:
   MO =    6  IRREP= 0 (Ag)
   MO =    7  IRREP= 5 (B1u)
   MO =    8  IRREP= 2 (B2g)
   MO =    9  IRREP= 7 (B3u)

Setting up the integral package       ... done
Building the CAS space                ... done (7 configurations for Mult=1 Irrep=0)
Building the CAS space                ... done (4 configurations for Mult=3 Irrep=7)
```

Note that the irrep occupations and active space irreps will be frozen to what they are upon entering the CASSCF program. This helps to setup the CI problem.

After which it smoothly converges to give:

```
    6:  1.986258       -0.753012    -20.4905   3-Ag
    7:  1.457849       -0.291201     -7.9240   1-B1u
    8:  0.541977        0.100890      2.7454   1-B2g
    9:  0.013915        0.964186     26.2368   3-B3u
```

As well as:

```
-------------------------------
SA-CASSCF TRANSITION ENERGIES
-------------------------------

LOWEST ROOT =    -78.110314788 Eh -2125.490 eV

STATE   ROOT MULT IRREP DE/a.u.    DE/eV    DE/cm**-1
   1:    0    3  B3u   0.163741     4.456  35937.1
```

RI, RIJCOSX and RIJK approximations for CASSCF

A significant speedup of CASSCF calculations on larger molecules can be achieved with the RI, RI-JK and RIJCOSX approximations. [459] There are two independent integral generation and transformation steps in a CASSCF procedure. In addition to the usual Fock matrix construction, that is central to HF and DFT approaches, more integrals appear in the construction of the orbital gradient and Hessian. The latter are approximated using the keyword `trafostep RI`, where an auxiliary basis (/C or the more accurate /JK auxiliary basis) is required. Note that auxiliary basis sets of the type /J are not sufficient to fit these integrals. If no suitable auxiliary basis set is available, the `AutoAux` feature might be useful (see comment in the input below). [827] We note passing, that there are in principle three distinguished auxiliary basis slots, that can be individually assigned in the `%basis` block (section *Choice of Basis Set*). As an example, we recompute the benzene ground state example from Section *Starting Orbitals* with a CAS(6,6).

```
! SV(P) def2-svp/C
! moread
%moinp "Test-CASSCF-Benzene-2.mrci.nat"

# Commented out: Detailed settings of the auxiliary basis in the %basis block,
#                where the AuxC slot is relevant for the option TrafoStep RI.
```

(continues on next page)

(continued from previous page)

```
# %basis
# auxC "def2-svp/C" # "AutoAux" or "def2/JK"
# end

%casscf  nel    6
         norb   6
         nroots 1
         mult   1
         trafostep ri
         end
```

The energy of this calculation is `-230.590328 Eh` compared to the previous result `-230.590271 Eh`. Thus, the RI error is only 0.06 mEh which is certainly negligible for all intents and purposes. With the larger `/JK` auxiliary basis the error is typically much smaller (0.02 mEh in this example). Even if more accurate results are necessary, it is a good idea to pre-converge the CASSCF with RI. The resulting orbitals should be a much better guess for the subsequent calculation without RI and thus save computation time.

The `TrafoStep RI` only affects the integral transformation in CASSCF calculations while the Fock operators are still calculated in the standard way using four index integrals. In order to fully avoid any four-index integral evaluation, you can significantly speed up the time needed in each iteration by specifying `!RIJCOSX`. The keyword implies `TrafoStep RI` . The COSX approximation is used for the construction of the Fock matrices. In this case, an additional auxiliary basis (**/J** auxiliary basis) is mandatory.

```
! SV(P) def2-svp/C RIJCOSX def2/J
! moread
%moinp "Test-CASSCF-Benzene-2.mrci.nat"

# Commented out: Detailed settings of the auxiliary basis in the %basis block,
#                where the AuxJ and AuxC slot are mandatory.
# %basis
# auxJ "def2/J"     # "AutoAux"
# auxC "def2-svp/C" # "AutoAux", "def2/JK"
# end

%casscf  nel    6
  norb   6
  nroots 1
  mult   1
end
```

The speedup and accuracy is similar to what is observed in RHF and UHF calculations. In this example the `RIJCOSX` leads to an error of 1 mEh. The methodology performs better for the computation of energy differences, where it profits from error cancellation. The RIJCOSX is ideally suited to converge large-scale systems. Note that for large calculations the integral cut-offs and numerical grids should be tightened. See section *Using the RI Approximation for Hartree-Fock and Hybrid DFT (RIJCOSX)* for details. With a floppy numerical grid setting the accuracy as well as the convergence behavior of CASSCF deteriorate. The RIJK approximation offers an alternative ansatz. The latter is set with `!RIJK` and can also be run in conventional mode (`conv`) for additional speed-up. With `conv`, a **single auxiliary basis must be provided** that is sufficiently larger to approximate the Fock matrices as well the gradient/Hessian integrals. In direct mode an additional auxiliary basis set can be set for the `AuxC` slot.

```
! SV(P) RIJK  def2/JK

# Commented out: Detailed settings of the auxiliary basis in the %basis block,
#                where only the auxJK slot must be set.
# %basis
# auxJK "def2/JK" # or "AutoAux"
# end
```

The RIJK methodology is more accurate and robust for CASSCF e.g. here the error is just 0.5 mEH.

Organic molecules with nearly double occupied orbitals can be challenge for the orbital optimization process. We compare calculations done with/without the NR solver:

```
! SV(P)
! moread
%moinp "Test-CASSCF-Benzene-2.mrci.nat"

%casscf  nel    6
  norb   6
  nroots 1
  mult   1
  # overwriting default settings with NR close to convergence
  switchstep NR
end
```

The NR variant takes 5 cycles to converge, whereas the default (`SuperCI_PT`) requires 8 cycles. In general, first order methods, take more iterations compared to the NR method. However, first order methods are much cheaper than the NR and therefore it may pay off to do a few iterations more rather than switching to the expensive second order methods. Moreover, second order methods are less robust and may diverge in certain circumstances (too far from convergence). When playing with the convergence settings, there is always a trade-off between speed versus robustness. The default settings are chosen carefully.[459] Facing convergence problems, it can be useful to use an alternative scheme (`orbstep SuperCI` and `switchstep DIIS`) in conjunction with a level-shifts (`ShiftUp`, `ShiftDn`). Alternatively, changing the guess orbitals may avoid convergence problems as well.

Robust Convergence with TRAH-CASSCF

The restricted-step second-order converger TRAH *Quadratic Convergence* is now also available for both state-specific and state-averaged CASSCF calculations.[382] To activate TRAH for your CASSCF calculation, you just need to add `!TRAH` in one of the simple input lines and add an auxiliary basis.

```
! TRAH Def2-SVP Def2-SVP/C TightSCF

%casscf
  nel 6
  norb 6
  mult 1
  nroots 2
end

*xyz 0 1
  N   0.0   0.0   0.0
  N   0.0   0.0   1.1
end
```

In most cases, there is no need to play with any input parameters. The only exception is the choice of active molecular orbital representations that can have a significant impact on the convergence rate for spin-coupled systems. As can be seen from Fig. Fig. 6.14, for such calculations localized active orbitals perform best. In any other case, the natural orbitals (default) should be employed.

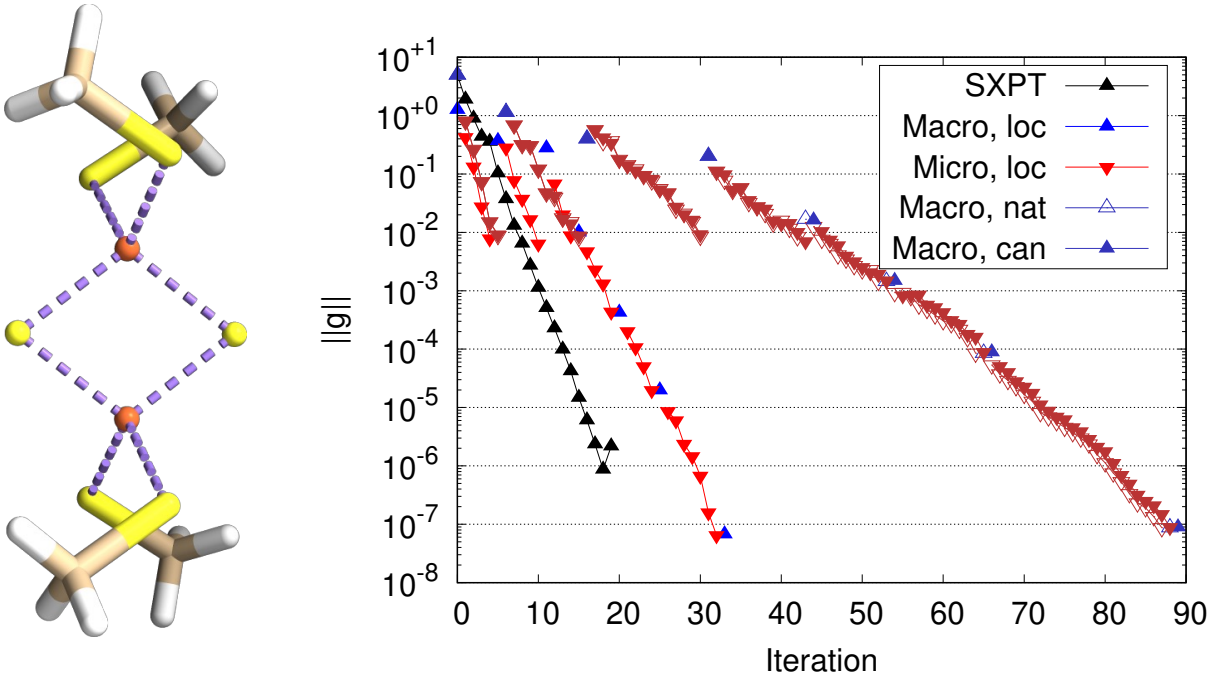


Fig. 6.14: SXPT and TRAH error convergence using different choices for the active-orbital basis.

Possible input options for the active-orbital basis are

```
%casscf
 TRAHCAS
  #ActiveMOs NotSet
  #ActiveMOs Canonical
  #ActiveMOs Localized
  ActiveMOs Natural   # default
 end
end
```

Active Orbitals	Meaning
Natural	Keeps the one-electron density matrix (1-RDM) diagonal. **Default**
Localized	A Foster-Boys localization of active MOs is performed in every macro iteration. This is recommended for **spin-coupled systems**.
NotSet	The active MO basis is not changed. Primarily debug option.
Canonical	Keeps the total active-MO Fock matrix diagonal. Experimental option.

Note that, in contrast to the SCF program, there is **no** `AutoTRAH` feature for CASSCF yet. The TRAH feature has to be requested explicitly in the input.

Breaking Chemical Bonds

Let us turn to the breaking of chemical bonds. As a first example we study the dissociation of the H_2 molecule. Scanning a bond, we have two potential setups for the calculation: a) scan from the inside to the outside or b) from the outside to inside. Of course both setups yield identical results, but they differ in practical aspects i.e. convergence properties. In general, **scanning from the outside to the inside is the recommended procedure**. Using the default guess (PModel), starting orbitals are much easier identified than at shorter distances, where the antibonding orbitals are probably 'impure' and hence would require some additional preparation. To ensure a smooth potential energy surface, in all subsequent geometry steps, ORCA reads the converged CASSCF orbitals from the previous geometry step. In the following, `TightSCF` is used to tighten the convergence settings of CASSCF.

```
!Def2-SVP TightSCF

%casscf
 nel         2
 norb        2
 mult        1
 nroots      1
end

# Scanning from the outside to the inside
%paras
 R [4.1 3.8 3.5 3.2 2.9 2.6 2.4 2.2
    2 1.7 1.5 1.3 1.1 1 0.9 0.8
    0.75 0.7 0.65 0.6]
end

* xyz 0 1
 H   0.0   0.0   0.0
 H   0.0   0.0   {R}
end
```

The resulting potential energy surface (PES) is depicted in Fig. 6.15 together with PESs obtained from RHF and broken-symmetry UHF calculations (input below).

```
! RHF Def2-SVP TightSCF

# etc...
```

And

```
! UHF Def2-SVP TightSCF

%scf
  FlipSpin 1
  FinalMs 0.0
end

# etc...
```

Note

The `FlipSpin` option does not work together with the parameter scan. Only the first structure will undergo a spin flip. Therefore, at the current status, a separate input file (including the coordinates or with a corresponding coordinate file) has to be provided for each structure that is scanned along the PES.

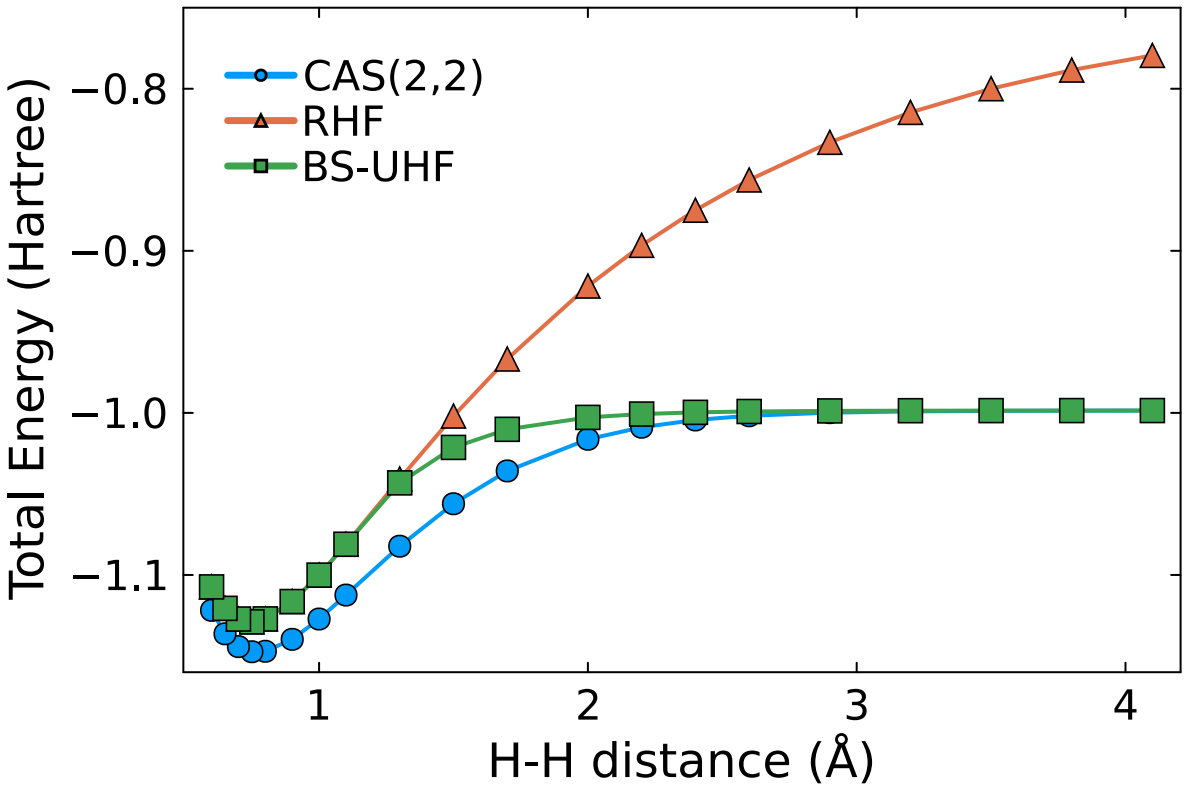


Fig. 6.15: Potential Energy Surface of the H_2 molecule from RHF, UHF and CASSCF(2,2) calculations (Def2-SVP basis).

It is obvious, that the CASSCF surface is concise and yields the correct dissociation behavior. The RHF surface is roughly parallel to the CASSCF surface in the vicinity of the minimum but then starts to fail badly as the H-H bond starts to break. The broken-symmetry UHF solution is identical to RHF in the vicinity of the minimum and dissociates correctly. It is, however, of rather mediocre quality in the intermediate region where it follows the RHF surface.

A more challenging case is to dissociate the N-N bond of the N_2 molecule correctly. Using CASSCF with the six p-orbitals we get a nice potential energy curve (The depth of the minimum is still too shallow compared to experiment by some 1 eV or so. A good dissociation energy requires a dynamic correlation treatment on top of CASSCF and a larger basis set).

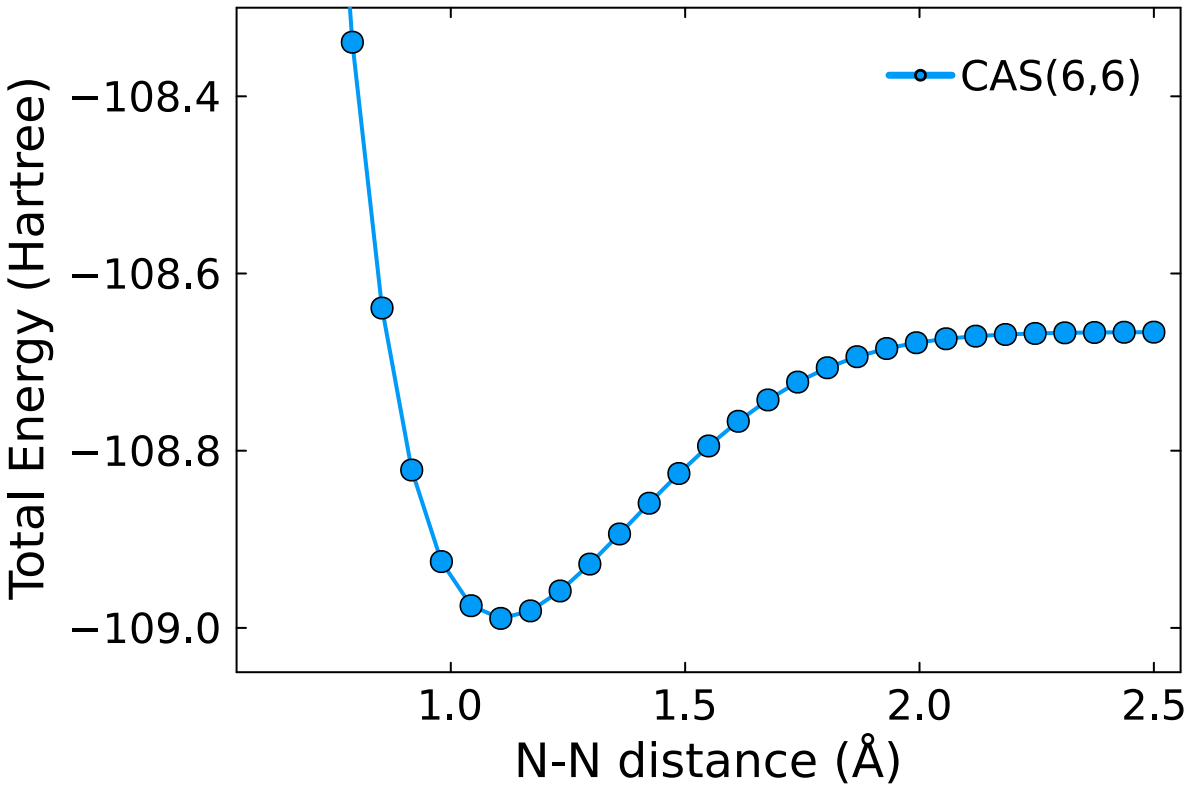


Fig. 6.16: Potential Energy Surface of the N_2 molecule from CASSCF(6,6) calculations (Def2-SVP basis).

One can use the H_2 example to illustrate the state-averaging feature. Since we have two active electrons we have two singlets and one triplet. Let us average the orbitals over these three states (we take equal weights for all multiplicity blocks):

```
!Def2-SVP TightSCF

%casscf
 nel         2
 norb        2
 mult        3,1
 nroots      1,2
end

# Scanning from the outside to the inside
%paras
R [4.1 3.8 3.5 3.2 2.9 2.6 2.4 2.2
   2 1.7 1.5 1.3 1.1 1 0.9 0.8
   0.75 0.7 0.65 0.6]
end

* xyz 0 1
 H 0 0 0
 H 0 0 {R}
end
```

which gives:

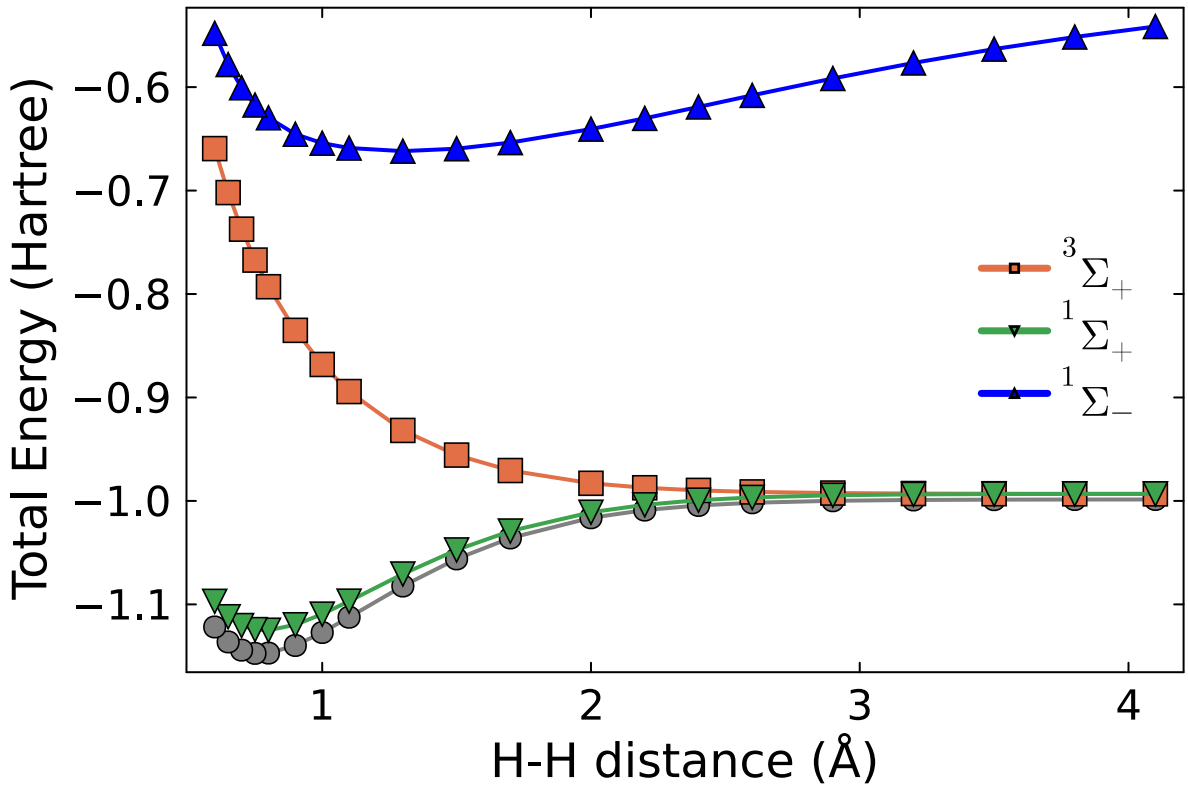


Fig. 6.17: State averaged CASSCF(2,2) calculations on H_2 (two singlets, one triplet; Def2-SVP basis). The grey curve is the ground state CASSCF(2,2) curve

One observes, that the singlet and triplet ground states become degenerate for large distances (as required) while the second singlet becomes the ionic singlet state which is high in energy. If one compares the lowest root of the state-averaged calculation (in green) with the dedicated ground state calculation (in gray) one gets an idea of the energetic penalty that is associated with averaged as opposed to dedicated orbitals.

A more involved example is the rotation around the double bond in C_2H_4. Here, the π-bond is broken as one twists the molecule. The means the proper active space consists of two active electron in two orbitals.

The input is (for fun, we average over the lowest two singlets and the triplet):

```
!def2-SVP TightSCF

%casscf
 nel        2
 norb       2
 mult       3,1
 nroots     1,2
end

%paras
 Alpha = 0,180,37
end

* int 0 1
 C  0  0  0  0      0   0
 C  1  0  0 1.34    0   0
 H  1  2  0 1.07  120   0
 H  1  2  3 1.07  120 180
 H  2  1  3 1.07  120 {Alpha}
 H  2  1  3 1.07  120 {Alpha+180}
edn
```

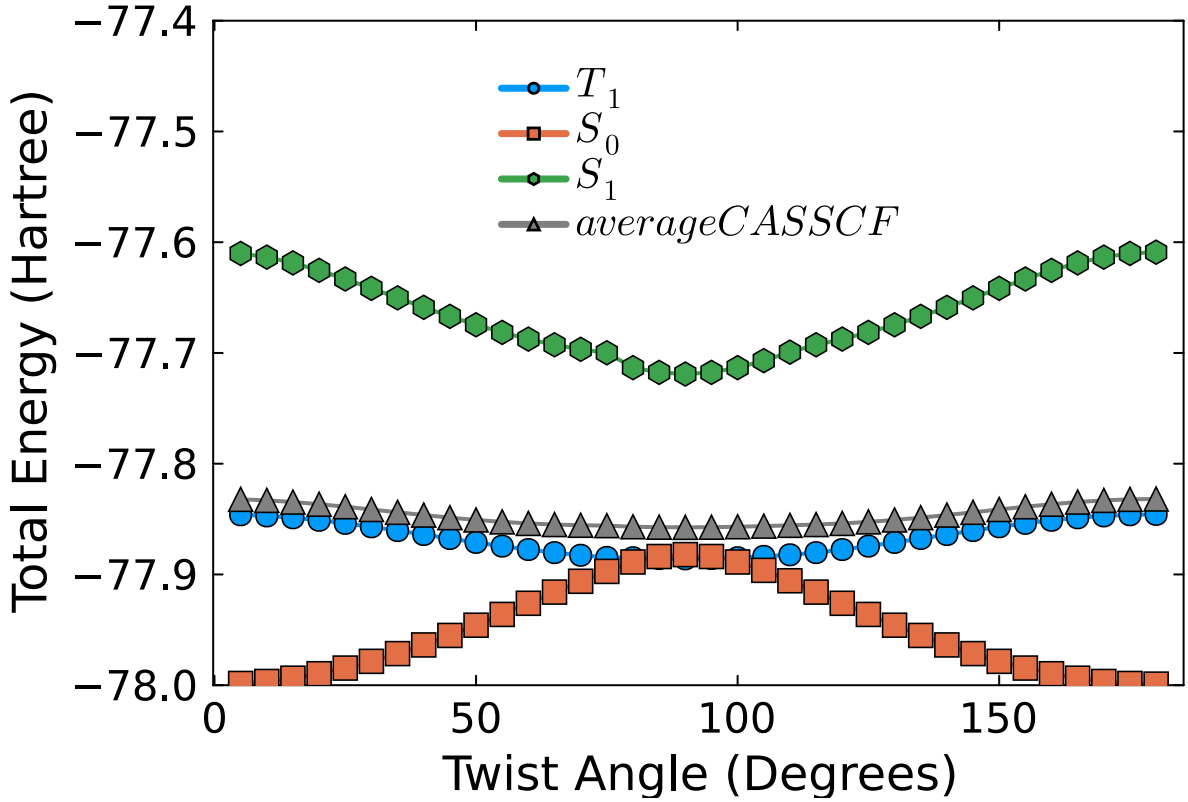


Fig. 6.18: State averaged CASSCF(2,2) calculations on C_2H_4 (two singlets, one triplet; SV(P) basis). The grey curve is the state averaged energy.

We can see from this plot, that the CASSCF method produces a nice ground state surface with the correct periodicity and degeneracy at the end points, which represent the planar ethylene molecule. At 90° one has a weakly coupled diradical and the singlet and triplet states become nearly degenerate, again as expected. Calculations with larger basis sets and inclusion of dynamic correlation would give nice quantitative results.

Excited States

As a final example, we do a state-average calculation on H_2CO in order to illustrate excited state treatments. We expect from the ground state (basically closed-shell) a n $\rightarrow$ π^* and a π $\rightarrow$ π^* excited state which we want to describe. For the n$\rightarrow$ π^* we also want to calculate the triplet since it is well known experimentally. First we take DFT orbitals as starting guess.

```
! BP86 Def2-SVP TightSCF

*int 0 1
 C  0 0 0 0.00   0.0    0.00
 O  1 0 0 1.20   0.0    0.00
 H  1 2 0 1.10 120.0    0.00
 H  1 2 3 1.10 120.0  180.00
end
```

In this example the DFT calculation produces the desired active space (n,π andπ^* orbitals) without further modification (e.g. swapping orbitals). In general it is advised to verify the final converged orbitals.

```
! Def2-SVP TightSCF MOREAD

%moinp "orbs.gbw"

%casscf
```

(continues on next page)

(continued from previous page)

```
 nel         4
 norb        3
 mult        1,3
 nroots      3,1
end

*int 0 1
C  0 0 0 0.00   0.0    0.00
O  1 0 0 1.20   0.0    0.00
H  1 2 0 1.10 120.0    0.00
H  1 2 3 1.10 120.0  180.00
end
```

We get:

```
-----------------------------
SA-CASSCF TRANSITION ENERGIES
-----------------------------

LOWEST ROOT (ROOT 0 ,MULT 1) =   -113.805194041 Eh -3096.797 eV

STATE   ROOT MULT  DE/a.u.    DE/eV    DE/cm**-1
1:    0    3   0.129029     3.511  28318.5
2:    1    1   0.141507     3.851  31057.3
3:    2    1   0.453905    12.351  99620.7
```

The triplet n $\rightarrow \pi^*$ states is spot on with the experiment excitation energy of 3.5 eV.[725] Similarly, the singlet n $\rightarrow \pi^*$ excited state is well reproduced compared to 3.79 eV and 4.07 eV reported in the literature.[725, 878] Only the singlet $\pi \rightarrow \pi^*$ excited state stands out compared to the theoretical estimate of 9.84 eV computed with MR-AQCC.[542]. The good results are very fortuitous given the small basis set, the minimal active space and the complete neglect of dynamical correlation.

The state-average procedure might not do justice to the different nature of the states (n $\rightarrow \pi^*$ versus $\pi \rightarrow \pi^*$). The agreement should be better with the orbitals optimized for each state. In ORCA, state-specific optimization are realized adjusting the weights i.e. for the second singlet excited root:

```
Second-Singlet:
%casscf
  nel         4
  norb        3
  mult        1
  nroots      3
  weights[0] = 0,0,1 # weights for the roots
end
```

Note, that state-specific orbital optimization are challenging to converge and often prone to root-flipping.[511]

To analyze electronic transitions, natural transition orbitals (NTO) are available for state-averaged CASSCF (and also CASCI) calculations. NTOs are switched on for every ground- to excited-state transition by just adding `DoNTO true` to the `%casscf ... end` input block, i.e.

```
%casscf
  nel    4
  norb   3
  mult   1,3
  nroots 3,1
  DoNTO  true
end
```

For each excitation, the most dominant natural occupation numbers (singular values >1.e-4) are printed for each transition. A set of donor orbitals and a set of acceptor orbitals, each of dimension Nbf x (Nocc + Nact), are created

and stored in files with unique names. We obtain for the previous formamide example the following CASSCF NTO output

```
=========================================
CASSCF Natural Transition Orbitals
=========================================

-----------------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE     1 1A
-----------------------------------------------

STATE    1 1A  :  E=   0.141508 au      3.851 eV    31057.3 cm**-1

Threshold for printing occupation numbers 1.0000e-04

     0  : n=  1.30882812
     1  : n=  0.02641080

=> Natural Transition Orbitals (donor   ) were saved in Test-CASSCF.H2CO-1.casscf.1-1A_nto-
↪donor.gbw
=> Natural Transition Orbitals (acceptor) were saved in Test-CASSCF.H2CO-1.casscf.1-1A_nto-
↪acceptor.gbw

-----------------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE     2 1A
-----------------------------------------------

STATE    2 1A  :  E=   0.453905 au     12.351 eV    99620.7 cm**-1

Threshold for printing occupation numbers 1.0000e-04

     0  : n=  1.30519478
     1  : n=  0.24869813
     2  : n=  0.00471742

=> Natural Transition Orbitals (donor   ) were saved in Test-CASSCF.H2CO-1.casscf.2-1A_nto-
↪donor.gbw
=> Natural Transition Orbitals (acceptor) were saved in Test-CASSCF.H2CO-1.casscf.2-1A_nto-
↪acceptor.gbw
```

For each transition, plots of the NTO pairs can be generated with the `orca_plot` program (see Sec. *Orbital and Density Plots* for details), e.g. acceptor orbitals of the 2 1A1 state in interactive mode:

```
orca_plot Test-CASSCF.H2CO-1.casscf.2-1A_nto-acceptor.gbw -i
```

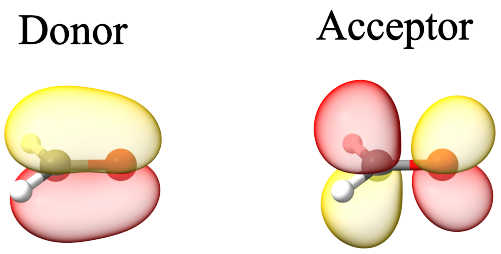


Fig. 6.19: Most dominant natural transition orbital (NTO) pair for the 2 1A1 (S2) transition in formaldehyde.

Alternatively, NTOs can also be computed directly in `orca_plot` from the CASSCF transition density matrices. Those need to be stored and kept in the density container by invoking

```
! KeepTransDensity
```

CASSCF Natural Orbitals as Input for Coupled-Cluster Calculations

Consider the possibility that you are not sure about the orbital occupancy of your system. Hence you carry out some CASSCF calculation for various states of the system in an effort to decide on the ground state. You can of course follow the CASSCF by MR-MP2 or MR-ACPF or SORCI calculations to get a true multireference result for the state ordering. Yet, in some cases you may also want to obtain a coupled-cluster estimate for the state energy difference. Converging coupled-cluster calculation on alternative states in a controlled manner is anything but trivial. Here a feature of ORCA might be helpful. The best single configuration that resembles a given CASSCF state is built from the natural orbitals of this state. These orbitals are also ordered in the right way to be input into the MDCI program. The convergence to excited states is, of course, not without pitfalls and limitations as will become evident in the two examples below.

As an example, consider some ionized states of the water cation:

First, we generate the natural orbitals for each state with the help of the MRCI module. To this end we run a state average CASSCF for the lowest three doublet states and pass that information on to the MRCI module that does a CASCI calculation and produces the natural orbitals:

```
! ano-pVDZ TightSCF

%casscf
 nel     7
 norb    6
 nroots  3
 mult    2
end

%mrci
 tsel       0
 tpre       0
 donatorbs  2
 densities  5,1
 newblock 2 *
  nroots 3
  excitations none
  refs
   cas(7,6)
  end
 end
end

* int 1 2
 O     0   0   0   0.000000     0.000     0.000
 H     1   0   0   1.012277     0.000     0.000
 H     1   2   0   1.012177   109.288     0.000
end
```

This produces the files `Basename.bm_sn.nat` where "m" is the number of the block (m = 0 correspond to the doublet in this case) and "n" stands of the relevant state (n = 0,1,2).

These natural orbitals are then fed into unrestricted QCISD(T) calculations:

```
! ano-pVDZ TightSCF QCISD(T) MOREAD NoIter

%moinp "H2O+.b0_s0.nat"

* int 1 2
 O     0   0   0   0.000000     0.000     0.000
 H     1   0   0   1.012277     0.000     0.000
 H     1   2   0   1.012177   109.288     0.000
*
```

As a reference we also perform a SORCI on the same system

```
! ano-pVDZ TightSCF SORCI

%casscf
 nel     7
 norb    6
 nroots  3
 mult    2
end

* int 1 2
 O     0   0   0   0.000000     0.000     0.000
 H     1   0   0   1.012277     0.000     0.000
 H     1   2   0   1.012177   109.288     0.000
*
```

we obtain the transition energies:

```
      SORCI   QCISD(T) (in cm-1)
  D0      0     0.0
  D1  16269   16293
  D2  50403   50509
```

Thus, in this example the agreement between single- and multireference methods is good and the unrestricted QCISD(T) method is able to describe these excited doublet states. The natural orbitals have been a reliable way to guide the CC equations into the desired solutions. This will work in many cases.

Large Scale CAS-SCF calculations using ICE-CI

The CASSCF procedure can be used for the calculation of spin-state energetics of molecules showing a multi-reference character via the state-averaged CASSCF protocol as described in the CASSCF section *The Complete Active Space Self-Consistent Field (CASSCF) Module*. The main obstacle in getting qualitatively accurate spin-state energetics for molecules with many transition metal centers is the proper treatment of the static-correlation effects between the large number of open-shell electrons. In this section, we describe how one can effectively perform CASSCF calculations on such systems containing a large number of high-spin open-shell transition metal atoms.

As an example, consider the Iron-Sulfur dimer $[\mathrm{Fe(III)}_2\mathrm{SR}_2]^{2-}$ molecule. In this system, the Fe(III) centers can be seen as being made up mostly of S=5/2 local spin states (lower spin states such as 3/2 and 1/2 will have small contributions due to Hunds' rule.) The main hurdle while using the CASSCF protocol on such systems (with increasing number of metal atoms) is the exponential growth of the Hilbert space although the physics can be effectively seen as occurring in a very small set of configuration state functions (CSFs). Therefore, in order to obtain qualitatively correct spin-state energetics, one need not perform a Full-CI on such molecules but rather a CIPSI like procedure using the ICE-CI solver should give chemically accurate results. In the case of the Fe(III) dimer, one can imagine that the ground singlet state is composed almost entirely of the CSF where the two Fe(III) centers are coupled antiferromagnetically. Such a CSF is represented as follows:

$$\left|\Psi_0^{S=0}\right\rangle = [1, 1, 1, 1, 1, -1, -1, -1, -1, -1]$$

In order to make sense of this CSF representation, one needs to clarify a few points which are as follows:

- First, in the above basis the 10 orbitals are localized to 5 on each Fe center (following a high-spin UHF/UKS calculation.)
- Second, the orbitals are ordered (as automatically done in ORCA_LOC) such that the first five orbitals lie on one Fe(III) center and the last five orbitals on the second Fe(III) center.

Using this ordering, one can read the CSF shown above in the following way: The first five *1* represent the five electrons on the first Fe(III) coupled in a parallel fashion to give a S=5/2 spin. The next five *-1* represent two points:

- First, the five consecutive *-1* signify the presence of five ferromagnetically coupled electrons on the second Fe(III) center resulting in a local S=5/2 spin state.

- Second, the second set of spins are coupled to the first *1* via anti-parallel coupling as signified by the sign of the last five *-1* entries.

Therefore, we can see that using the CSF representation, one can obtain an extremely compact representation of the wavefunction for molecules consisting of open-shell transition metal atoms. This protocol of using localized orbitals in a specified order to form compact CSF representations for transition metal systems can be systematically extended for large molecules.

We will use the example of the Iron-Sulfur dimer $[Fe(III)_2SR_2]^{2-}$ to demonstrate how to prepare a reference CSF and perform spin-state energetics using the state-averaged CASSCF protocol. In such systems, often one can obtain an estimate of the energy gap between the singlet-state and the high-spin states from experiment. Ab initio values for this gap be obtained using the state-averaged CASSCF protocol using the input shown below.

```
! def2-SVP MOREAD

%moinp "locorbs.gbw"

%casscf
 nel     10
 norb    10
 mult    11,1
 nroots  1,1
 refs                              # reference for multiplicity 11
  { 1 1 1 1 1  1  1  1  1  1}
 end
 refs                              # reference for multiplicity 1
  { 1 1 1 1 1 -1 -1 -1 -1 -1}
 end
  cistep ice
 ci
  icetype 1
 end
 actorbs unchanged
end

* xyz -2 11
Fe          0.000000000      0.000000000     -1.343567812
Fe          0.000000000      0.000000000      1.343567812
S           1.071733501      1.373366082      0.000000000
S           1.346714284     -1.345901486     -2.651621449
S          -1.346714284      1.345901486     -2.651621449
S          -1.071733501     -1.373366082      0.000000000
S          -1.346714284      1.345901486      2.651621449
S           1.346714284     -1.345901486      2.651621449
C          -2.485663304      0.362543393     -3.600795276
H          -3.319937516      0.596731755     -3.505882795
H          -2.347446507      0.388292903     -4.463380590
H          -2.472404709     -0.485711203     -3.404167343
C           2.485663304     -0.362543393     -3.600795276
H           3.319937516     -0.596731755     -3.505882795
H           2.347446507     -0.388292903     -4.463380590
H           2.472404709      0.485711203     -3.404167343
C           2.485663304     -0.362543393      3.600795276
H           2.347446507     -0.388292903      4.463380590
H           3.319937516     -0.596731755      3.505882795
H           2.472404709      0.485711203      3.404167343
C          -2.485663304      0.362543393      3.600795276
H          -3.319937516      0.596731755      3.505882795
H          -2.472404709     -0.485711203      3.404167343
H          -2.347446507      0.388292903      4.463380590
*
```

The main keyword that needs to be used here (unlike in other CAS-SCF calculations) is the *actorbs* keyword.

Since we are using a local basis with a specific ordering of the orbitals, in order to represent our wavefunction it is imperative to preserve the local nature of the orbitals as well as the orbital ordering. Therefore, we do not calculate natural orbitals at the end of the CASSCF calculation (as is traditionally done) instead we impose the orbitals to be as similar to the input orbitals as possible. This is automatically enabled for intermediate CASSCF macro iterations. The resulting CASSCF calculation provides a chemically intuitive and simple wavefunction and transition energy as shown below:

```
-----------------------------------------
CAS-SCF STATES FOR BLOCK  1 MULT=11 NROOTS= 1
-----------------------------------------

STATE   0 MULT=11: E=  -5066.8462457411 Eh W=  0.5000 DE= 0.000 eV      0.0 cm**-1
      1.00000 ( 1.000000000) CSF = 1+1+1+1+1+1+1+1+1+1+



-----------------------------------------
CAS-SCF STATES FOR BLOCK  2 MULT= 1 NROOTS= 1
-----------------------------------------

STATE   0 MULT= 1: E=  -5066.8548894831 Eh W=  0.5000 DE= 0.000 eV      0.0 cm**-1
      0.98159 (-0.990753235) CSF = 1+1+1+1+1+1-1-1-1-1-



-----------------------------
SA-CASSCF TRANSITION ENERGIES
-----------------------------

LOWEST ROOT (ROOT 0 ,MULT 1) =  -5066.854889483 Eh -137876.131 eV

STATE   ROOT MULT  DE/a.u.    DE/eV    DE/cm**-1
   1:    0   11   0.008644     0.235   1897.1
```

As we can see from the output above, 98% of the wavefunction for the singlet-state is given by a single CSF which we gave as a reference CSF. This CSF has a very simple chemical interpretation representing the anti-parallel coupling between the two high-spin Fe(III) centers. Since Iron-Sulfur molecules show a strong anti-ferromagnetic coupling, we expect the singlet state to be lower in energy than the high-spin (S=5) state. The CASSCF transition energies show essentially this fact. The transition energy is about 2000cm^{-1} which is what one expects from a CASSCF calculation on such sulfide bridged transition-metal molecules.

6.1.8 N-Electron Valence State Perturbation Theory (NEVPT2)

NEVPT2 is an internally contracted multireference perturbation theory, which applies to CASSCF type wavefunctions. The NEVPT2 method, as described in the original papers of Angeli et al, comes in two flavor the strongly contracted NEVPT2 (SC-NEVPT2) and the so called partially contracted NEVPT2 (PC-NEVPT2).[44, 45, 46] In fact, the latter employs a fully internally contracted wavefunction and should more appropriately called FIC-NEVPT2. Both methods produces energies of similar quality as the CASPT2 approach.[366, 756] The strongly and fully internally contracted NEVPT2 are implemented in ORCA together with a number of approximations that makes the methodology very attractive for large scale applications. In conjunction with the RI approximation systems with active space of to 16 active orbitals and 2000 basis functions can be computed. With the newly developed DLPNO version of the FIC-NEVPT2 the size of the molecules does not matter anymore.[344] For a more complete list of keywords and features, we refer to detailed documentation section *N-Electron Valence State Pertubation Theory*.

Besides corrections to the correlation energy, ORCA features UV, IR, CD and MCD spectra as well as EPR parameters for NEVPT2. These properties are computed using the "quasi-degenerate perturbation theory" that is described in section *CASSCF Properties*. The NEVPT2 corrections enter as "improved diagonal energies" in this formalism. ORCA also features the multi-state extension (QD-NEVPT2) for the strongly contracted NEVPT2 variant.[48, 493] Here, the reference wavefunction is revised in the presence of dynamical correlation. For systems,

where such reference relaxation is important, the computed spectroscopic properties will improve.

As a simple example for NEVPT2, consider the ground state of the nitrogen molecule N_2 . After defining the computational details of our CASSCF calculation, we insert "!SC-NEVPT2" as simple input or specify "PTMethod SC_NEVPT2" in the %casscf block. Please note the difference in the two keywords' spelling: Simple input uses hyphen, block input uses underscore for technical reasons. There are more optional settings, which are described in section *N-Electron Valence State Pertubation Theory* of this manual.

```
!def2-svp nofrozencore PAtom
%casscf nel  6
        norb 6
        mult 1
        PTMethod SC_NEVPT2  # SC_NEVPT2   for strongly contracted NEVPT2
                            # FIC_NEVPT2  for the fully internally contracted NEVPT2
                            # DLPNO_NEVPT2 for the FIC-NEVPT2 with DLPNO
                            # DLPNO requires: trafostep RI and an aux basis
end

* xyz 0 1
N        0.0        0.0        0.0
N        0.0        0.0        1.09768
*
```

For better control of the program flow it is advised to split the calculation into two parts. First converge the CASSCF wave function and then in a second step read the converged orbitals and execute the actual NEVPT2.

```
------------------------------------------------------------------
                         ORCA-CASSCF
------------------------------------------------------------------

...
PT2-SETTINGS:
A PT2 calculation will be performed on top of the CASSCF wave function (PT2 = SC-NEVPT2)
...
------------------------------------------------------------------
                           < NEVPT2  >
------------------------------------------------------------------
...
 ================================================================
                       NEVPT2 Results
================================================================
   ********************
    MULT 1, ROOT 0
   ********************

  Class V0_ijab :        dE = -0.017748
  Class Vm1_iab :        dE = -0.023171
  Class Vm2_ab  :        dE = -0.042194
  Class V1_ija  :        dE = -0.006806
  Class V2_ij   :        dE = -0.005056
  Class V0_ia   :        dE = -0.054000
  Class Vm1_a   :        dE = -0.007091
  Class V1_i    :        dE = -0.001963

------------------------------------------------------------------
         Total Energy Correction : dE = -0.15802909
------------------------------------------------------------------
         Zero Order Energy       : E0 = -108.98888640
------------------------------------------------------------------
         Total Energy (E0+dE)    : E  = -109.14691549
------------------------------------------------------------------
```

Introducing dynamic correlation with the SC-NEVPT2 approach lowers the energy by 150 mEh. ORCA also

prints the contribution of each "excitation class V" to the first order wave function. We note that in the case of a single reference wavefunction corresponding to a CAS(0,0) the `V0_ij,ab` excitation class produces the exact MP2 correlation energy. Unlike older versions of ORCA (pre version 4.0), NEVPT2 calculations employ the frozen core approximation by default. Results from previous versions can be obtained with the added keyword `!NoFrozenCore`.

In chapter *Breaking Chemical Bonds* the dissociation of the N_2 molecule has been investigated with the CASSCF method. Inserting `PTMethod SC_NEVPT2` into the `%casscf` block we obtain the NEVPT2 correction as additional information.

```
! def2-svp nofrozencore
%casscf nel  6
        norb 6
        mult 1
        PTMethod SC_NEVPT2
end

# scanning from the outside to the inside
%paras
        R = 2.5,0.7, 30
end

*xyz 0 1
N 0.0 0.0 0.0
N 0.0 0.0 {R}
*
```

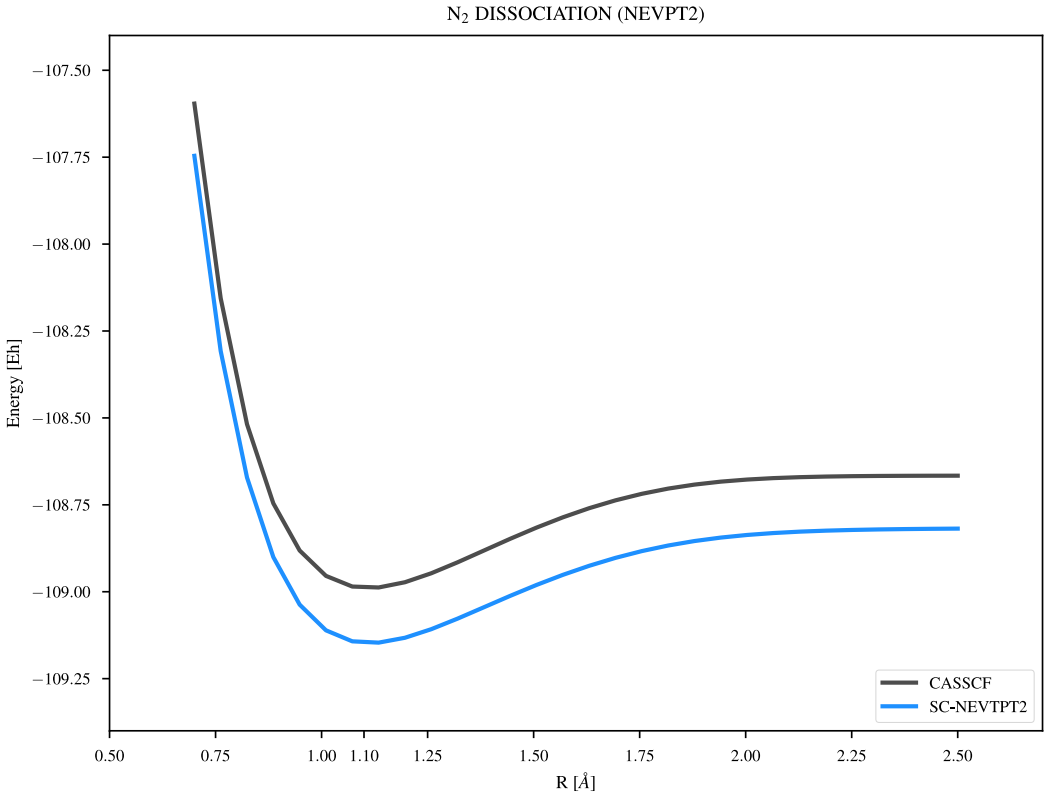


Fig. 6.20: Potential Energy Surface of the N_2 molecule from CASSCF(6,6) and NEVPT2 calculations (def2-SVP).

All of the options available in CASSCF can in principle be applied to NEVPT2. Since NEVPT2 is implemented

as a submodule of CASSCF, it will inherit all settings from CASSCF (`!tightscf, !UseSym, !RIJCOSX, ...`).

NOTE

- NEVPT2 analytic gradients are not available, but numerical gradients are!

6.1.9 Complete Active Space Perturbation Theory: CASPT2 and CASPT2-K

The fully internally contracted CASPT2 (FIC-CASPT2) approach shares its wave function ansatz with the FIC-NEVPT2 approach mentioned in the previous section.[40] The two methods differ in the definition of the zeroth order Hamiltonian. The CASPT2 approach employs the generalized Fock-operator, which may result in intruder states problems (singularities in the perturbation expression). Real and imaginary level shifting techniques are introduced to avoid intruder states.[270, 731] Note that both level shifts are mutually exclusive. Since level shifts in general affect the total energies, they should be avoided or chosen as small as possible. As argued by Roos and coworkers, CASPT2 systematically underestimates open-shell energies, since the Fock operator itself is not suited to describe excitations into and out of partially occupied orbitals. The deficiency can be adjusted with the inclusion of IPEA shifts - an empirical parameter.[295] While the implementation of the canonical CASPT2 with real and imaginary shifts is validated against OpenMOLCAS.[253], the ORCA version differs in the implementation of the IPEA shifts and yields slightly different results. The IPEA shift, λ, is added to the matrix elements of the internally contracted CSFs $\Phi_{qs}^{pr} = E_q^p E_s^r |\Psi^0>$ with the generalized Fock operator

$$< \Phi_{q's'}^{p'r'} | \hat{F} | \Phi_{qs}^{pr} > + = < \Phi_{q's'}^{p'r'} | \Phi_{qs}^{pr} > \cdot \frac{\lambda}{2} \cdot (4 + \gamma_p^p - \gamma_q^q + \gamma_r^r - \gamma_s^s),$$

where $\gamma_q^p = < \Psi^0 | E_q^p | \Psi^0 >$ is the expectation value of the spin-traced excitation operator.[441] The labels p,q,r,s refer to general molecular orbitals (inactive, active and virtual). Irrespective of the ORCA implementation, the validity of the IPEA shift in general remains questionable and is thus by default disabled.[921]

ORCA features an alternative approach, denoted as **CASPT2-K**, that reformulates the zeroth order Hamiltonian itself.[460] Here, two additional Fock matrices are introduced for excitation classes that add or remove electrons from the active space. The new Fock matrices are derived from the generalized Koopmans' matrices corresponding to electron ionization and attachment processes. The resulting method is less prone to intruder states and the same time more accurate compared to the canonical CASPT2 approach. For a more detailed discussion, we refer to the paper by Kollmar et al.[460]

The CASPT2 and CASPT2-K methodologies are called in complete analogy to the NEVPT2 branch in ORCA and can be combined with the resolution of identity (RI) approximation.

```
%casscf
...
       PTMethod FIC_CASPT2   # fully internally contracted CASPT2
                FIC_CASPT2K  # CASPT2-K (revised H0)

       # Optional settings
       PTSettings
       CASPT2_rshift    0.0 # (default) real level shift
       CASPT2_ishift    0.0 # (default) imaginary level shift
       CASPT2_IPEAshift 0.0 # (default) IPEA shift
       end
end
```

The RI approximated results are comparable to the CD-CASPT2 approach presented elsewhere.[50] For a general discussion of the RI and CD approximations, we refer to the literature.[884] Many of the input parameter are shared with the FIC-NEVPT2 approach. A list with the available options is presented in section *Complete Active Space Peturbation Theory : CASPT2 and CASPT2-K*.

In this short section, we add the CASPT2 results to the previously computed NEVPT2 potential energy surface of the N_2 molecule.

```
! def2-svp nofrozencore
%casscf nel  6
```

(continues on next page)

(continued from previous page)

```
        norb 6
        mult 1
        PTMethod FIC_CASPT2 # fully internally contracted CASPT2
end

# scanning from the outside to the inside
%paras
        R = 2.5,0.7, 30
end

*xyz 0 1
N 0.0 0.0 0.0
N 0.0 0.0 {R}
*
```

The CASPT2 output lists the settings prior to the computation. The printed reference weights should be checked. Small **reference weights** indicate intruder states. Along the lines, the program also prints the **smallest denominators** in the perturbation expression (highlighted in the snippet below). Small denominator may lead to intruder states.

```
------------------------------------------------------------
ORCA-CASSCF
------------------------------------------------------------

...
PT2-SETTINGS:
A PT2 calculation will be performed on top of the CASSCF wave function (PT2 = CASPT2)
CASPT2 Real Levelshift            ...  0.00e+00
CASPT2 Im.  Levelshift            ...  0.00e+00
CASPT2 IPEA Levelshift            ...  0.00e+00
...
------------------------------------------------------------
                          < CASPT2  >
------------------------------------------------------------
...
-----------------------------------
CASPT2-D Energy =     -0.171839049
-----------------------------------

Class V0_ijab: dE=     -0.013891923
Class Vm1_iab: dE=     -0.034571085
Class Vm2_ab : dE=     -0.040985427
Class V1_ija : dE=     -0.003511548
Class V2_ij  : dE=     -0.000579508
Class V0_ia  : dE=     -0.075176596
Class Vm1_a  : dE=     -0.002917335
Class V1_i   : dE=     -0.000205627

smallest energy denominator IJAB =      3.237539973
smallest energy denominator ITAB =      2.500295823
smallest energy denominator IJTA =      2.339868413
smallest energy denominator TUAB =      1.664398302
smallest energy denominator IJTU =      1.342421639
smallest energy denominator ITAU =      1.496042538
smallest energy denominator TUVA =      0.706288250
smallest energy denominator ITUV =      0.545304334

...
 Iter           EPT2     EHylleraas  residual norm    Time
    1    -0.17183905    -0.17057203     0.03246225     0.0
    2    -0.17057203    -0.17119523     0.00616509     0.0
```

(continues on next page)

(continued from previous page)

```
   3    -0.17117095    -0.17121211     0.00086389     0.0
   4    -0.17120782    -0.17121281     0.00013292     0.0
   5    -0.17121273    -0.17121282     0.00000990     0.0
   6    -0.17121283    -0.17121282     0.00000159     0.0
   7    -0.17121282    -0.17121282     0.00000020     0.0

 CASPT2 calculation converged in 7 iterations

...
 ================================================================
 CASPT2 Results
 ================================================================
   *********************
   MULT 1, ROOT 0
   *********************

  Class V0_ijab :        dE = -0.013831560889
  Class Vm1_iab :        dE = -0.034124733943
  Class Vm2_ab  :        dE = -0.041334010085
  Class V1_ija  :        dE = -0.003446396316
  Class V2_ij   :        dE = -0.000584401134
  Class V0_ia   :        dE = -0.074688029120
  Class Vm1_a   :        dE = -0.002962355569
  Class V1_i    :        dE = -0.000241331405

 ----------------------------------------------------------------
         Total Energy Correction : dE = -0.17121281846205
 ----------------------------------------------------------------
         Reference  Energy       : E0 = -108.66619981448225
         Reference Weight        : W0 = 0.94765190644139
 ----------------------------------------------------------------
         Total Energy (E0+dE)    : E  = -108.83741263294431
 ----------------------------------------------------------------
```

Note that the program prints CASPT2-D results prior entering the CASPT2 iterations.[40] In case of intruder states, the residual equation may not converge. The program will not abort. Hence, it is important to check convergence for every CASPT2 run. In this particular example with the small basis sets, there are no intruder states.

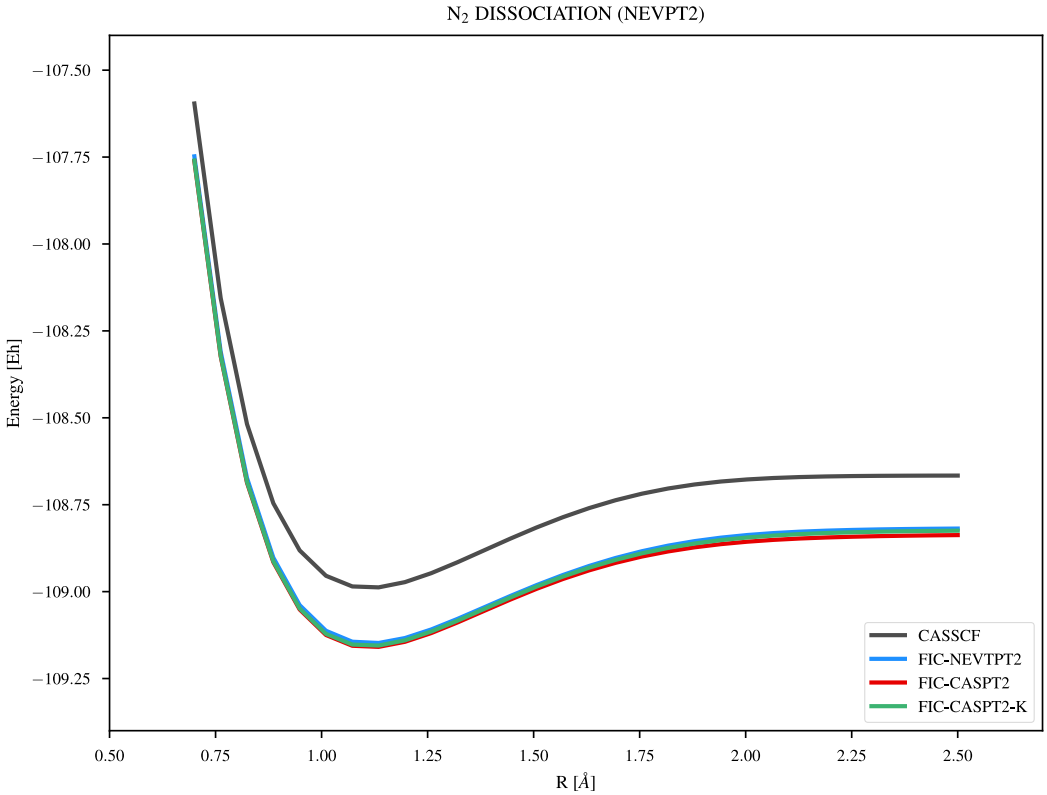


Fig. 6.21: Potential Energy Surface of the N_2 molecule from CASSCF(6,6) and CASPT2 calculations (def2-SVP).

The potential energy surface in Fig. 6.21 is indeed very similar to the FIC-NEVPT2 approach, which is more efficient (no iterations) and robust (absence of intruder states). The figure also shows the CASPT2-K results, which is typically a compromise between the two methods. As expected, the largest deviation from CASPT2 is observed at the dissociation limit, where the open shell character dominates the reference wave function. In this example, the discrepancy between the three methods is rather subtle. However, the results may differ substantially on some challenging systems, such as Chromium dimer studied in the CASPT2-K publication. [460]. Despite its flaws, the CASPT2 method is of historical importance and remains a popular methodology. In the future we might consider further extension such as the (X)MS-CASPT2.[792]

6.1.10 2nd order Dynamic Correlation Dressed Complete Active Space method (DCD-CAS(2))

Non-degenerate multireference perturbation theory (MRPT) methods, such as NEVPT2 or CASPT2, have the 0th order part of the wave function fixed by a preceding CASSCF calculation. The latter can be a problem if the CASSCF states are biased towards a wrong state composition in terms of electron configurations. In these instances, a quasi-degenerate or multi-state formulation is necessary, for example the QD-NEVPT2 described in Section *Quasi-Degenerate SC-NEVPT2*. A drawback of these approaches is that the results depend on the number of included states. The DCD-CAS(2) offers an alternative uncontracted approach, where a dressed CASCI matrix is constructed. Its diagonalization yields correlated energies and 0th order states that are remixed in the CASCI space under the effect of dynamic correlation.[652]

The basic usage is very simple: One just needs a `%casscf` block and the simple input keyword `!DCD-CAS(2)` . The following example is a calculation on the LiF molecule. It possesses two singlet states that can be qualitatively described as ionic (Li^+^ and F^-) and covalent (neutral Li with electron in 2s orbital and neutral F with hole in $2p_z$ orbital). At distances close to the equilibrium geometry, the ground state is ionic, while in the dissociation limit the ground state is neutral. Somewhere in between, there is an avoided crossing of the adiabatic potential

energy curves where the character of the two states quickly changes (see figure Fig. 7.7 for potential energy curves for this system at the (QD)NEVPT2 level). At the CASSCF level, the neutral state is described better than the ionic state, with the result that the latter is too high in energy and the avoided crossing occurs at a too small interatomic distance. In the region where the avoided crossing actually takes place, the CASSCF states are then purely neutral or purely ionic. DCD-CAS(2) allows for a remixing of the states in the CASCI space under the effect of dynamic correlation, which will lower the ionic state more in energy than the neutral one. The input file is as follows:

```
! def2-TZVP DCD-CAS(2)
!moread

%moinp "casorbs.gbw" # guess with active orbitals in place

%casscf
  nel 2
  norb 2
  mult 1
  nroots 2
  actorbs locorbs
end

*xyz 0 1
Li 0.0 0.0 0.0
F 0.0 0.0 5.5
*
```

Since none of the standard guesses (`!PAtom`, `!PModel`, etc.) produces the correct active orbitals (Li 2s and F 2p~z~), we read them from the file casorbs.gbw. We also use the `actorbs locorbs` option to preserve the atomic character of the active orbitals and interpret the states in terms of neutral and ionic components easier. The following is the state composition of LiF at an interatomic distance of 5.5 angstrom at the CASSCF and DCD-CAS(2) levels.

```
-----------------------------------------------
CAS-SCF STATES FOR BLOCK  1 MULT= 1 NROOTS= 2
-----------------------------------------------

ROOT   0:  E=    -106.8043590118 Eh
      0.99395 [     1]: 11
      0.00604 [     2]: 02
ROOT   1:  E=    -106.7485794535 Eh  1.518 eV  12242.2 cm**-1
      0.99396 [     2]: 02
      0.00604 [     1]: 11
```

```
---------------------------------------
      DCD-CAS(2) STATES
---------------------------------------

ROOT   0:  E=    -107.0917611937 Eh
      0.60590 [     2]: 02
      0.39410 [     1]: 11
ROOT   1:  E=    -107.0837717163 Eh  0.217 eV   1753.5 cm**-1
      0.60590 [     1]: 11
      0.39410 [     2]: 02
```

One can clearly see that while the CASSCF states are purely neutral (dominated by CFG 11) or purely ionic (dominated by CFG 02), the DCD-CAS(2) states are mixtures of neutral and ionic contributions. The calculation indicates that the interatomic distance of 5.5 is in the avoided crossing region. Note that the energies that are printed together with the DCD-CAS(2) state composition are the ones that are obtained from diagonalization of the DCD-CAS(2) dressed Hamiltonian. For excited states, these energies suffer from what we call *ground state bias* (see the original paper for a discussion [652]). A perturbative correction has been devised to overcome this problem. Our standard choice is first-order bias correction. The corrected energies are also printed in the output file and those energies should be used in production use of the DCD-CAS(2) method:

```
-------------------------------------------------------
 BIAS-CORRECTED (ORDER 1) STATE AND TRANSITION ENERGIES
=======================================================
 ROOT      Energy/a.u.     DE/a.u.      DE/eV      DE/cm**-1
=======================================================
   0:  -107.093214435   0.000000      0.000        0.0
   1:  -107.084988306   0.008226      0.224     1805.4
```

Last but not least, spin orbit coupling (SOC) and spin spin coupling (SSC) are implemented in conjunction with the DCD-CAS(2) method in a QDPT-like procedure and a variety of different magnetic and spectroscopic properties can be also calculated. We refer to the detailed documentation (Section *Dynamic Correlation Dressed CAS*) for further information.

Warning

Note that the computational cost of a DCD-CAS(2) calculation scales as roughly the 3rd power of the size of the CASCI space. This makes calculations with active spaces containing more than a few hundred CSFs very expensive!

6.1.11 Full Configuration Interaction Energies

ORCA provides several exact and approximate approaches to tackle the full configuration interaction (FCI) problem. These methods are accessible via the CASSCF module (see Section *General Description*) or the ICE module (described in Section *Approximate Full CI Calculations in Subspace: ICE-CI*).

In the following, we compute the FCI energy of the lithium hydride molecule using the CASSCF module, where a typical input requires the declaration of an active space. The latter defines the number of active electron and orbitals, which are evaluated with the FCI ansatz. In the special case that all electrons and orbitals are treated with the FCI ansatz, we can use the keyword `DoFCI` in the `%CASSCF` block and let the program set the active space accordingly. In this example, we focus on the singlet ground state. Note that excited states for arbitrary multiplicities can be computed with the keywords `Mult` and `NRoots`. The FCI approach is invariant to orbital rotations and thus orbital optimization is skipped in the CASSCF module. Nevertheless, it is important to employ a set of meaningful orbitals, e.g. from a converged Hartree-Fock calculation, to reduce the number of FCI iterations.

```
# Hartree-Fock orbitals
!def2-tzvp RHF

*xyz 0 1
  Li 0 0 0
  H  0 0 1.597
*
```

The output of the Hartree-Fock calculation also reports on the total number of electrons and orbitals in your system (see snippet below).

```
Number of Electrons    NEL             ....    4
Basis Dimension        Dim             ....   20
```

In the given example, there are 4 electrons in 20 orbitals, which is a “CAS(4,20)”. Reading the converged RHF orbitals, we can start the FCI calculation.

```
!def2-tzvp extremescf

!moread
%moinp "RHF.gbw"

%maxcore 2000
```

(continues on next page)

(continued from previous page)

```
%casscf
  DoFCI true # sets NEL 4 and NORB 20 in this example.
end

*xyz  0 1
  Li 0 0 0
  H  0 0 1.597
*
```

The output reports on the detailed CI settings, the number of configuration state functions (CSFs) and the CI convergence thresholds.

```
CI-STEP:
        CI strategy                          ... General CI
        Number of multiplicity blocks        ...    1
        BLOCK  1 WEIGHT=   1.0000
        Multiplicity                         ...    1
        #(Configurations)                    ... 8455
        #(CSFs)                              ... 13300
        #(Roots)                             ...    1
        ROOT=0 WEIGHT=    1.000000

        PrintLevel                           ...    1
        N(GuessMat)                          ...      512
        MaxDim(CI)                           ...       10
        MaxIter(CI)                          ...       64
        Energy Tolerance CI                  ...  1.00e-13
        Residual Tolerance CI                ...  1.00e-13
        Shift(CI)                            ...  1.00e-04
        ...
```

The program then prints the actual CI iterations, the final energy, and the composition of the wave function in terms of configurations (CFGs).

```
------------------
CAS-SCF ITERATIONS
------------------


        MACRO-ITERATION   1:
        --- Inactive Energy E0 = 0.99407115 Eh
        --- All densities will be recomputed
        CI-ITERATION   0:
        -8.012799617   0.526896429727 (    0.25)
        CI-ITERATION   1:
        -8.047996328   0.001601312242 (    0.25)
        CI-ITERATION   2:
        -8.048134967   0.000022625293 (    0.25)
        CI-ITERATION   3:
        -8.048137773   0.000000462227 (    0.25)
        CI-ITERATION   4:
        -8.048137841   0.000000035496 (    0.25)
        CI-ITERATION   5:
        -8.048137845   0.000000001357 (    0.25)
        CI-ITERATION   6:
        -8.048137845   0.000000000254 (    0.25)
        CI-ITERATION   7:
        -8.048137845   0.000000000006 (    0.25)
        CI-ITERATION   8:
        -8.048137845   0.000000000001 (    0.25)
        CI-ITERATION   9:
```

(continues on next page)

(continued from previous page)

```
        -8.048137845   0.000000000000 (    0.25)
        CI-PROBLEM SOLVED
        DENSITIES MADE

        <<<<<<<<<<<<<<<<<<INITIAL CI STATE CHECK>>>>>>>>>>>>>>>>>>

        BLOCK  1 MULT= 1 NROOTS= 1
        ROOT   0:  E=      -8.0481378449 Eh
        0.97242 [     0]: 22000000000000000000
        0.00296 [    99]: 20000002000000000000
        0.00258 [    89]: 20000010001000000000
        0.00252 [    85]: 20000020000000000000
```

Aside from energies, the CASSCF module offers a number of properties (g-tensors, ZFS, ...), which are described in Section *CASSCF Properties*.

The exact solution of the FCI problem has very steep scaling and is thus limited to smaller problems (at most active spaces of 16 electrons in 16 orbitals). Larger systems are accessible with approximate solutions, e.g. with the density matrix renormalization group approach (DMRG), described in Section *Density Matrix Renormalization Group*, or the iterative configuration expansion (ICE) reported in Section *Approximate Full CI Calculations in Subspace: ICE-CI*. For fun, we repeat the calculation with the ICE-CI ansatz, which offers a more traditional approach to get an approximate full CI result.

```
!def2-tzvp extremescf

!moread
%moinp "RHF.gbw"

%maxcore 2000

%ice
  Nel 4
  Norb 20
end

*xyz  0 1
  Li 0 0 0
  H  0 0 1.597
*
```

The single most important parameter to control the accuracy is `TGen`. It is printed with the more refined settings in the output. We note passing that the wave function expansion and its truncation can be carried out in the basis of CSFs, configurations, or determinants. The different strategies are discussed in detail by Chilkuri *et al.* [171, 172].

```
ICE-CI:
        General Strategy                ... CONFIGURATIONS (all CSFs to a given CFG, spin␣
↪adapted)
        Max. no of macroiterations      ...   12
        Variational selection threshold ... -1.000e-07
        negative! => TVar will be set to 1.000e-07*Tgen=1.000e-11
        Generator selection threshold   ... 1.000e-04
        Excitation level                ... 2
        Selection on initial CSF list   ... YES
        Selection on later CSFs lists   ... YES

        ...

*****************************
*  ICECI MACROITERATION   3  *
*****************************
```

(continues on next page)

(continued from previous page)

```
# of active configurations =   2808
Initializing the CI                  ...  (CI/Run=3,2 UseCC=0)done (  0.0 sec)
Building coupling coefficients       ... (CI/Run=3,2)Calling BuildCouplings_RI UseCCLib=0␣
↪DoRISX=0
CI_BuildCouplings NCFG= 2808 NORB=20 NEL=4 UseCCLib=0 MaxCore=2000
PASS 1 completed. NCFG= 2808 NCFGK= 8416 MaxNSOMOI=4 MaxNSOMOK=4
PASS 2 completed.
PASS 3 completed.
Memory used for RI tree              =    2.99 MB (av. dim=    35)
Memory used for ONE tree             =    1.32 MB (av. dim=    46)
Memory used for coupling coefficients=    0.01 MB
done (    0 sec)
Now calling CI solver (4095 CSFs)

****Iteration    0****
Maximum residual norm  :     0.000293130557

****Iteration    1****
Maximum residual norm  :     0.000000565920

****Iteration    2****
Maximum residual norm  :     0.000001755176

****Iteration    3****
Maximum residual norm  :     0.000000435942
Rebuilding the expansion space

****Iteration    4****

*** CONVERGENCE OF ENERGIES REACHED ***
CI problem solved in   0.4 sec

CI SOLUTION :
STATE   0 MULT= 1: E=     -8.0481340246 Eh W=  1.0000 DE= 0.000 eV      0.0 cm**-1
0.97249  : 2200000000000000000
Selecting new configurations         ... (CI/Run=3,2)done (  0.0 sec)
# of selected configurations          ...   2747
# of generator configurations         ...     43 (NEW=1 (CREF=43))
Performing single and double excitations relative to generators ... done (  0.0 sec)
# of configurations after S+D         ...   7038
Selecting from the generated configurations  ... done (  0.1 sec)
# of configurations after Selection  ...   2808
Root   0:     -8.048134025       -0.000000023      -8.048134048
==>>> CI space seems to have converged. No new configurations
maximum energy change                ... 1.727e-05 Eh

                    ********* ICECI IS CONVERGED *********
        Initializing the CI                  ...  (CI/Run=3,3 UseCC=0)done (  0.0 sec)
        Building coupling coefficients       ... (CI/Run=3,3)Calling BuildCouplings_RI␣
↪UseCCLib=0 DoRISX=
        CI_BuildCouplings NCFG= 2808 NORB=20 NEL=4 UseCCLib=0 MaxCore=2000
        PASS 1 completed. NCFG= 2808 NCFGK= 8416 MaxNSOMOI=4 MaxNSOMOK=4
        PASS 2 completed.
        PASS 3 completed.
        Memory used for RI tree              =    2.99 MB (av. dim=    35)
        Memory used for ONE tree             =    1.32 MB (av. dim=    46)
        Memory used for coupling coefficients=    0.01 MB
        done (    0 sec)
        Now calling CI solver (4095 CSFs)

        ****Iteration    0****
```

(continues on next page)

(continued from previous page)

```
       Maximum residual norm  :     0.000000471011

       ****Iteration    1****

       *** CONVERGENCE OF ENERGIES REACHED ***
       CI problem solved in   0.1 sec

       CI SOLUTION :
       STATE   0 MULT= 1: E=     -8.0481340245 Eh W=  1.0000 DE= 0.000 eV      0.0 cm**-1
       0.97249  : 22000000000000000000
```

With Hartree-Fock orbitals and the default settings, the ICE converges in 3 macro iterations to an energy of $-8.048134047513\ E_\mathrm{h}$. The deviation from the exact solution is just $3.8 \times 10^{-6}\ E_\mathrm{h}$ in this example.

6.1.12 Efficient Calculations with Atomic Natural Orbitals

Atomic natural orbitals are a special class of basis sets. They are represented by the orthonormal set of orbitals that diagonalizes a spherically symmetric, correlated atomic density. The idea is to put as much information as possible into each basis functions such that one obtains the best possible result with the given number of basis functions. This is particularly important for correlated calculations where the number of primitives is less an issue than the number of basis functions.

Usually, ANO basis sets are "generally contracted" which means that for any given angular momentum all primitives contribute to all basis functions. Since the concept of ANOs only makes sense if the underlying set of primitives is large, the calculations readily become very expensive unless special precaution is taken in the integral evaluation algorithms. ORCA features special algorithms for ANO basis sets together with accurate ANO basis sets for non-relativistic calculations. However, even then the integral evaluation is so expensive that efficiency can only be realized if all integrals are stored on disk and are re-used as needed.

In the first implementation, the use of ANOs is restricted to the built-in ANO basis sets (ano-pVnZ, saug-ano-pVnZ, aug-ano-pVnZ, n = D, T, Q, 5). These are built upon the cc-pV6Z primitives and hence, the calculations take significant time.

Note

- Geometry optimizations with ANOs are discouraged; they will be *very* inefficient.

The use of ANOs is recommended in the following way:

```
! ano-pVTZ Conv TightSCF CCSD(T)
%maxcore 2000
* int 0 1
C  0 0 0    0   0   0
O  1 0 0  1.2   0   0
H  1 2 0  1.1 120   0
H  1 2 3  1.1 120 180
*
```

This yields:

```
ano-pVTZ:
E(SCF) = -113.920388785
E(corr)=   -0.427730189
```

Compare to the cc-pVTZ value of:

```
cc-pVTZ:
E(SCF) = -113.911870901
E(corr)=   -0.421354947
```

Thus, the ANO-based SCF energy is ca. 8–9 mEh lower and the correlation energy almost 2 mEh lower than with the cc-basis set of the same size. Usually, the ANO results are much closer to the basis set limit than the cc-results. Also, ANO values extrapolate very well (see section *Automatic extrapolation to the basis set limit*)

Importantly, the integrals are all stored in this job. Depending on your system and your patience, this may be possible up to 300–500 basis functions. The ORCA correlation modules have been rewritten such that they deal efficiently with these stored integrals. Thus, we might as well have used `! MO-CCSD(T)` or `! AO-CCSD(T)` , both of which would perform well.

Yet, the burden of generating and storing all four-index integrals quickly becomes rather heavy. Hence, the combination of ANO basis sets with the RI-JK technique is particularly powerful and efficient. For example:

```
! ano-pVTZ cc-pVTZ/JK RI-JK Conv TightSCF RI-CCSD(T)
```

For the SCF, this works very well and allows for much larger ANO based calculations to be done efficiently. Also, RI-MP2 can be done very efficiently in this way. However, for higher order correlation methods such as CCSD(T) the logical choice would be RI-CCSD(T) which is distinctly less efficient than the AO or MO based CCSD(T) (roughly a factor of two slower). Hence, ORCA implements a hybrid method where the RI approximation is used to generate all four index integrals. This is done via the "RI-AO" keyword:

```
! ano-pVTZ cc-pVTZ/JK RI-AO Conv TightSCF AO-CCSD(T)
```

In this case either AO-CCSD(T) or MO-CCSD(T) would both work well. This does not solve the storage bottleneck with respect to the four index integrals of course. If this becomes a real issue, then RI-CCSD(T) is mandatory. The error in the total energy is less than 0.1 mEh in the present example.

Note

- **With conventional RI calculations the use of a second fit basis set is not possible and inconsistent results will be obtained. Hence, stick to one auxiliary basis!**

6.1.13 Local-SCF Method

The Local-SCF (LSCF) method developed by X. Assfeld and J.-L. Rivail ([55]) allows to optimize a single determinant wave function under the constraint of keeping frozen (i.e. unmodified) a subset of orbitals. Also, optimized orbitals fulfill the condition of orthogonality with the frozen ones. The LSCF method can be applied to restricted/unrestricted Hartree-Fock or DFT Kohn-Sham wavefunctions.

To use the LSCF method, one chooses the spin-up and spin-down frozen orbitals with the "LSCFalpha" and "LSCFbeta" keywords, respectively. Frozen orbitals are specified using **intervals** of orbital indexes. In the following example, the selection "0,4,5,6,10,10" for the alpha frozen orbitals means that the orbitals ranging from 0 to 4 (**0,4**,5,6,10,10), 5 and 6 (0,4,**5,6**,10,10) and the orbital 10 (0,4,5,6,**10,10**) will be frozen. In the case of the beta orbitals, the orbitals with indexes 0, 1, 2, 3 and 5 will be frozen. Up to 5 intervals (2*5 numbers) are allowed.

```
#
# Example of LSCF Calculation
#
! UKS B3LYP/G SVP TightSCF
%scf
LSCFalpha 0,4,5,6,10,10
LSCFbeta  0,3,5,5
end
```

For the sake of user-friendliness, two other keywords are available within the LSCF method. They can be used to modify the orbital first guess, as read from the gbw file with the same name or another gbw file with the "MOInp" keyword.

The "LSCFCopyOrbs" keyword allows to copy one orbital into another one. The input works by intervals like the LSCFalpha/LSCFbeta selections. However, be aware that spin-up orbital indexes range from 0 to M-1 (where M is the size of the basis set), while spin-down orbital indexes range from M to 2M-1. In the following example, with M=11, the user copies the fifth spin-up orbital in the fifth spin-down orbital.

```
%scf
  LSCFalpha 0,4,5,6,10,10
  LSCFbeta  0,3,5,5
  LSCFCopyOrbs 4,15
end
```

The second keyword is "LSCFSwapOrbs" and allows to swap the indexes of subsets made of two orbitals. In the following example, still with M=11, the user swaps the fifth spin-up orbital with the fifth spin-down orbital.

```
%scf
  LSCFalpha 0,4,5,6,10,10
  LSCFbeta  0,3,5,5
  LSCFSwapOrbs 4,15
end
```

Caution

During the LSCF procedure, frozen occupied orbitals energies are fixed at -1000 Hartrees and frozen virtual orbitals energies at 1000 Hartrees. This means that the frozen occupied orbitals and the frozen virtual orbitals are placed respectively at the beginning and at the end of the indexation.

6.1.14 Adding finite electric field

Electric fields can have significant influences on the electronic structure of molecules. In general, when an electric field is applied to a molecule, the electron cloud of the molecule will polarize along the direction of the field. The redistribution of charges across the molecule will then influence the wavefunction of the molecule. Even when polarization effects are not significant, the electric field still exerts a drag on the negatively and positively charged atoms of the molecule in opposite directions, and therefore affect the orientation and structure of the molecule. The combination of electrostatic and polarization effects make electric fields a useful degree of freedom in tuning e.g. reactivities, molecular structures and spectra [785]. Meanwhile, the energy/dipole moment/quadrupole moment changes of the system in the presence of small dipolar or quadrupolar electric fields are useful for calculating many electric properties of the system via numerical differentiation, including the dipole moment, quadrupole moment, dipole-dipole polarizability, quadrupole-quadrupole polarizability, etc. Such finite difference property calculations can be conveniently done using compound scripts in the ORCA Compound Scripts Repository (https://github.com/ORCAQuantumChemistry/CompoundScripts/tree/main/Polarizabilities).

In ORCA, a uniform (or equivalently speaking, dipolar) electric field can be added to a calculation via the following keyword:

```
%scf
  EField 0.1, 0.0, 0.0 # x, y, z components (in au) of the electric field
end
```

Although the keyword is in the %scf block, it applies the electric field to all other methods (post-HF methods, multireference methods, TDDFT, etc.) as well, except XTB and force field methods (as well as any method that involves XTB or force fields, e.g. QM/XTB and QM/MM) for which the electric field contributions are not implemented and will result in an abort. Analytic gradient contributions of the electric field are available for all methods (except XTB and MM) that already support analytic gradients, but analytic Hessian contributions are not.

The sign convention of the electric field is chosen in the following way: suppose that the electric field is generated by a positive charge in the negative z direction, and a negative charge in the positive z direction, then the z component of the electric field is positive. This convention is consistent with most but not all other programs [785], so care must be taken when comparing the results of ORCA with other programs.

Another important aspect is the gauge origin of the electric field. The gauge origin of the electric field is the point (or more accurately speaking, one of the points - as there are infinitely many such points) where the electric potential due to the electric field is zero. Different choices of the gauge origin do not affect the geometry and wavefunction of the molecule, as they do not change the electric field felt by the molecule, but they do change the energy of the molecule. The default gauge origin is the (0,0,0) point of the Cartesian coordinate system, but it is possible to choose other gauge origins:

```
%scf
 EFieldOrigin CenterOfMass      # use center of mass
              CenterOfNucCharge # use center of nuclear charge
              0.0, 0.0, 0.0     # use given X,Y,Z as origin (default: 0,0,0)
                                # in the units chosen for the coordinates (Angstrom/Bohr)
end
```

Note the default gauge origin of the electric field is different from the default gauge origin of the ELPROP module, which is the center of mass. If the user chooses the center of mass/nuclear charge as the gauge origin of the electric field, the gauge origin will move as the molecule translates; this has important consequences. For example, in an MD simulation of a charged molecule in an electric field, the molecule will not accelerate, unlike when `EFieldOrigin` is fixed at a given set of coordinates, where the molecule will accelerate forever. In general, `CenterOfMass` and `CenterOfNucCharge` are mostly suited for the finite difference calculation of electric properties, where one frequently wants to choose the center of mass or nuclear charge as the gauge origin of the resulting multipole moment or polarizability tensor. Instead, a fixed origin is expected to be more useful for simulating the changes of wavefunction, geometry, reactivity, spectra etc. under an externally applied electric field, as experimentally the electric field is usually applied in the lab frame, rather than the comoving frame of the molecule.

Similar to `EField`, one can also add a quadrupolar field:

```
%scf
  QField 0.1, 0.0, 0.0, 0.05, 0.0, 0.0 # xx, yy, zz, xy, xz, yz components (in au)
                                        # of the quadrupolar field
end
```

The gauge origin of the quadrupolar field is the same as that of the dipolar electric field. Moreover, the `QField` can be used together with the `EField` keyword. This allows one to simulate a gradually varying electric field, for example the following input specifies an electric field that has a strength of 0.01 au at the gauge origin ((0,0,0) by default), pointing to the positive z direction, and increases by 0.001 au for every Bohr as one goes in the positive z direction:

```
%scf
  EField 0.0, 0.0, 0.01
  QField 0.0, 0.0, 0.001, 0.0, 0.0, 0.0
end
```

As a second example, one can also simulate an ion trap:

```
%scf
  QField -0.01, -0.01, -0.01, 0.0, 0.0, 0.0
end
```

Under this quadrupolar field setting, a particle will feel an electric field that points towards the gauge origin, whose strength (in au) is 0.01 times the distance to the gauge origin (in Bohr). This will keep cations close to the origin, but pushes anions away from the origin. Unfortunately, there is no analytic gradient available for quadrupolar fields.

NOTE

- An au (atomic unit) is a fairly large unit for electric fields: 1 au = 51.4 V/Angstrom. By comparison, charged residues in proteins, as well as scanning tunneling microscope (STM) tips, typically generate electric fields

within about 1 V/Angstrom; electrode surfaces usually generate electric fields within 0.1 V/Angstrom under typical electrolysis conditions [785]. If the molecule is not close to the source of the electric field, it is even harder to generate strong electric fields: for example, a 100 V voltage across two metal plates that are 1 mm apart generates an electric field of merely 10^{-5} V/Angstrom. Therefore, if experimentally a certain strength of homogeneous electric field seems to promote a reaction, but no such effect is found in calculation, please consider the possibility that the experimentally observed reactivity is due to a strong local electric field near the electrode surface (that is much higher than the average field strength in the system), or due to other effects such as electrolysis. Conversely, if you predict a certain molecular property change at an electric field strength of, e.g. > 0.1 au, it may be a non-trivial question whether such an electric field can be easily realized experimentally.

- The electric field breaks the rotational symmetry of the molecule, in the sense that rotating the molecule can change its energy. Therefore, geometry optimizations in electric fields cannot be done with internal coordinates. When the user requests geometry optimization, the program automatically switches to Cartesian coordinates if it detects an electric field. While Cartesian coordinates allow the correct treatment of molecular rotation, they generally lead to poor convergence, so a large number of iterations is frequently necessary.
- Similarly, when the molecule is charged, its energy is not invariant with respect to translations. However, when there is only a dipolar electric field but no other translational symmetry-breaking forces (quadrupolar field, point charges, wall potentials), a charged molecule will accelerate forever in the field, and its position will never converge. Therefore, for geometry optimizations within a purely dipolar electric field and no wall potentials, we do not allow global translations of the molecule, even when translation can reduce its energy. For MD simulations we however do allow the global translations of the molecule by default. If this is not desired, one can fix the center of mass in the MD run using the `CenterCOM` keyword (section *Run*).
- For frequency calculations in electric fields, we do not project out the translational and rotational contributions of the Hessian (equivalent to setting `ProjectTR false` in `%freq`; see *Frequency calculations - numerical and analytical* for details). Therefore, the frequencies of translational and rotational modes can be different from zero, and can mix with the vibrational modes. When the electric field is extremely small but not zero, the "true" translational/rotational symmetry breaking of the Hessian may be smaller than the symmetry breaking due to numerical error; this must be kept in mind when comparing the frequency results under small electric fields versus under zero electric field (in the latter case `ProjectTR` is by default true). Besides, when the translational and rotational frequencies exceed CutOffFreq (which is 1 cm^{-1} by default; see section *Frequency calculations - numerical and analytical*), their thermochemical contributions are calculated as if they are vibrations.
- While the program allows the combination of electric fields with an implicit solvation model, the results must be interpreted with caution, because the solvent medium does not feel the electric field. The results may therefore differ substantially from those given by experimental setups where both the solute and the solvent are subjected to the electric field. If the solvent's response to the electric field is important, one should use an explicit solvation model instead. Alternatively, one can also simulate the electric field in the implicit solvent by adding inert ions (e.g. Na^+, Cl^-) to the system. Similarly, implicit solvation models cannot describe the formation of electrical double layers in the electric field and their influence on solute properties, so in case electrical double layers are important, MD simulations with explicit treatment of the ions must be carried out.
- The electric field not only contributes to the core Hamiltonian, but has extra contributions in GIAO calculations, due to the magnetic field derivatives of dipole integrals. In the case of a dipolar electric field, the GIAO contributions have been implemented, making it possible to study e.g. the effect of electric fields on NMR shieldings, and as a special case, nucleus independent chemical shieldings (NICSs), which are useful tools for analyzing aromaticity. Quadrupolar fields cannot be used in GIAO calculations at the moment.

6.2 SCF Stability Analysis

The SCF stability will give an indication whether the SCF solution is at a local minimum or a saddle point.[80, 779] It is available for RHF/RKS and UHF/UKS. In the latter case, the SCF is restarted by default using new unrestricted start orbitals if an instability was detected. For a demonstration, consider the following input:

```
! BHLYP def2-SVP NORI

%scf
 guess hcore
 HFTyp UHF
 STABPerform true
end

* xyz 0 1
h 0.0 0.0 0.0
h 0.0 0.0 1.4
*
```

The HCORE guess leads to a symmetric/restricted guess, which does not yield the unrestricted solution. The same is often true for other guess options. For more details on the stability analysis, see Section *SCF Stability Analysis*.

6.3 Geometry Optimizations, Surface Scans, Transition States, MECPs, Conical Intersections, IRC, NEB

The usage of analytic gradients is necessary for efficient geometry optimization. In ORCA 5.0, the following methods provide analytic first derivatives

- Hartree-Fock (HF) and DFT (including the RI, RIJK and RIJCOSX approximations)
- MP2, RI-MP2 and DLPNO-MP2
- TD-DFT for excited states
- CAS-SCF

When the analytic gradients are not available, it is possible to evaluate the first derivatives numerically by finite displacements. This is available for all methods.

The coordinate system chosen for geometry optimization affects the convergence rate, with redundant internal coordinates being usually the best choice.

Some methods for locating transition states (TS) require second derivative matrices (Hessian), implemented analytically for HF, DFT and MP2. Additionally, several approaches to construct an initial approximate Hessian for TS optimization are available. A very useful feature for locating complicated TSs is the Nudged-Elastic Band method in combination with the TS finding algorithm (NEB-TS, ZOOM-NEB-TS). An essential feature for chemical processes involving excited states is the conical intersection optimizer. Another interesting feature are MECP (Minimum Energy Crossing Point) optimizations.

For very large systems ORCA provides a very efficient L-BFGS optimizer, which makes use of the orca_md module. It can also be invoked via simple keywords described at the end of this section.

6.3.1 Geometry Optimizations

Optimizations are fairly easy as in the following example:

```
! B3LYP/G SV(P) Opt
* int 0 1
    C  0 0 0 0.0000   0.000    0.00
    O  1 0 0 1.2029   0.000    0.00
    H  1 2 0 1.1075 122.016    0.00
    H  1 2 3 1.1075 122.016  180.00
*
```

An optimization with the RI method (the BP functional is recommend) would simply look like:

```
! BP SV(P) OPT
* int 0 1
    C  0 0 0 0.0000   0.000    0.00
    O  1 0 0 1.2029   0.000    0.00
    H  1 2 0 1.1075 122.016    0.00
    H  1 2 3 1.1075 122.016  180.00
*
```

An optimization of the first excited state of ethylene:

```
! BLYP SVP OPT

%tddft
   IRoot 1
end

* xyz 0 1
 C     0.000000     0.000000     0.666723
 C     0.000000     0.000000    -0.666723
 H     0.000000    -0.928802     1.141480
 H    -0.804366    -0.464401    -1.341480
 H     0.000000     0.928802     1.241480
 H     0.804366     0.464401    -1.241480
*
```

6.3.2 Numerical Gradients

If the analytic gradient is not available, the numerical gradient can simply be requested via:

```
! NumGrad
```

as in the following example:

```
!CCSD(T) TZVPP
!Opt NumGrad
* int 0 1
C 0 0 0  0   0   0
O 1 0 0  1.2 0   0
H 1 2 0  1.1 120 0
H 1 2 3  1.1 120 180
*
```

NOTE

- Be aware that the numerical gradient is quite expensive. The time for one gradient calculation is equal to 6 $\times$ (number of atoms) $\times$ (time for one single point calculation).

- The numerical gradient can be calculated in a multi-process run, using a maximum of three times the number of atoms (see section *Calling the Program with Multiple Processes*).

More details on various options, geometry convergence criteria and the like are found in section *Geometry Optimization*.

6.3.3 Some Notes and Tricks

Note

- `TightSCF` in the SCF part is set as default to avoid the buildup of too much numerical noise in the gradients.
- Even if the optimization does not converge, the ORCA output may still end with "****ORCA TERMINATED NORMALLY****". Therefore do not rely on the presence of this line as an indicator of whether the geometry optimization is converged! Rather, one should instead rely on the fact that, an optimization job that terminates because the maximum number of iterations has been reached, will generate the following output message:

Warning

```
The optimization did not converge but reached the maximum number of
optimization cycles. Please check your results very carefully.
```

While a successfully converged job will generate the following message instead:

```
**********************HURRAY********************
***         THE OPTIMIZATION HAS CONVERGED      ***
************************************************
```

Tip

- In rare cases the redundant internal coordinate optimization fails. In this case, you may try to use `COPT` (optimization in Cartesian coordinates). This will likely take many more steps to converge but should be stable.
- For optimizations in Cartesian coordinates the initial guess Hessian is constructed in internal coordinates and thus these optimizations should converge only slightly slower than those in internal coordinates. Nevertheless, if you observe a slow convergence behaviour, it may be a good idea to compute a Hessian initially (perhaps at a lower level of theory) and use `InHess read` in order to improve convergence.
- At the beginning of a TS optimization more information on the curvature of the PES is needed than a model Hessian can give. The best choice is analytic Hessian, available for HF, DFT and MP2. In other cases (e.g. CAS-SCF), the numerical evaluation is necessary. Nevertheless you do not need to calculate the full Hessian when starting such a calculation. With ORCA we have good experience with approximations to the exact Hessian. Here it is recommended to either directly combine the TS optimization with the results of a relaxed surface scan or to use the Hybrid Hessian as the initial Hessian, depending on the nature of the TS mode. Note that these approximate Hessians do never replace the exact Hessian at the end of the optimization, which is always needed to verify the minimum or first order saddle point nature of the obtained structure.

6.3.4 Initial Hessian for Minimization

The convergence of a geometry optimization crucially depends on the quality of the initial Hessian. In the simplest case it is taken as a unit matrix (in redundant internal coordinates we use 0.5 for bonds, 0.2 for angles and 0.1 for dihedrals and improper torsions). However, simple model force-fields like the ones proposed by Schlegel, Lindh, Swart or Almlöf are available and lead to much better convergence. The different guess Hessians can be set via the `InHess` option which can be either `unit`, `Almloef`, `Lindh`, `Swart` or `Schlegel` in redundant internal coordinates. Since version 2.5.30, these model force-fields (built up in internal coordinates) can also be used in optimizations in Cartesian coordinates.

For minimizations we recommend the `Almloef` Hessian, which is the default for minimizations. The `Lindh` and `Schlegel` Hessian yield a similar convergence behaviour. From version 4.1?, there is also the option for the `Swart` model Hessian, which is less parametrized and should improve for weakly interacting and/or unusual structures. Of course the best Hessian is the exact one. `Read` may be used to input an exact Hessian or one that has been calculated at a lower level of theory (or a "faster" level of theory). From version 2.5.30 on this option is also available in redundant internal coordinates. But we have to point out that the use of the exact Hessian as initial one is only of little help, since in these cases the convergence is usually only slightly faster, while at the same time much more time is spent in the calculation of the initial Hessian.

To sum it up: we advise to use one of the simple model force-fields for minimizations.

6.3.5 Coordinate Systems for Optimizations

The coordinate system for the optimization can be chosen by the `coordsys` variable that can be set to `cartesian` or `redundant`within the `%geom` block. The default is the redundant internal coordinate system. If the optimization with `redundant` fails, you can still try `cartesian`. If the optimization is then carried out in Cartesian displacement coordinates with a simple model force-field Hessian, the convergence will be only slightly slower. With a unit matrix initial Hessian very slow convergence will result.

A compound job `two_step_opt.inp` that first computes a semi-empirical Hessian to start from is shown below:

```
* int 0 1
C 0 0 0 0 0 0
O 1 0 0 1.3 0 0
H 1 2 0 1.1 110 0
H 1 2 3 1.1 110 180
*

%compound
# Step 1: semiempirical calculation of the Hessian
New_Step
! AM1 NumFreq
Step_End

# Step 2: optimization starting from previous Hessian
New_Step
!B3LYP def2-svp def2/J Opt
%geom
InHess
 Read
InHessName "two_step_opt_Compound_1.hess"
# this file must be either a .hess file from a
# frequency run or, a .opt/.carthess file left over from a
# previous geometry optimization
end
Step_End

End
```

Tip

- For transition metal complexes MNDO, AM1 or PM3 Hessians are not available. You can use ZINDO/1 or NDDO/1 Hessians instead. They are of lower quality than MNDO, AM1 or PM3 for organic molecules but they are still far better than the standard unit matrix choice.
- If the quality of the initial semi-empirical Hessian is not sufficient you may use a "quick" RI-DFT job (e.g. `BP def2-sv(p) defgrid1`)
- In semi-empirical geometry optimizations on larger molecules or in general when the molecules become larger the redundant internal space may become large and the relaxation step may take a significant fraction of the total computing time.

For condensed molecular systems and folded molecules (e.g. a U-shaped carbon chain) atoms can get very close in space, while they are distant in terms of number of bonds connecting them. As damping of optimization steps in internal coordinates might not work well for these cases, convergence can slow down. ORCA's automatic internal coordinate generation takes care of this problem by assigning bonds to atom pairs that are close in real space, but distant in terms of number of bonds connecting them.

```
%geom
 AddExtraBonds true         # switch on/off assigning bonds to atom pairs that are
                            #  connected by more than <Max_Length> bonds and are less
                            #  than <MaxDist> Ang. apart (default true)
 AddExtraBonds_MaxLength 10 # cutoff for number of bonds connecting the two
                            #  atoms (default 10)
 AddExtraBonds_MaxDist 5    # cutoff for distance between two atoms (default 5 Ang.)
end
```

For solid systems modeled as embedded solids the automatically generated set of internal coordinates might become very large, rendering the computing time spent in the optimization routine unnecessarily large. Usually, in such calculations the cartesian positions of outer atoms, coreless ECPs and point charges are constrained during the optimization - thus most of their internal coordinates are not needed. By requesting:

```
%geom
 ReduceRedInts true
 end
```

only the required needed internal coordinates (of the constrained atoms) are generated.

OBS: If the step in `redundant` fails badly and only Cartesian constrains are set (or no constrains), ORCA will fallback to a `cartesian` step automatically. This can be turned off by setting CARTFALLBACK to FALSE.

6.3.6 Constrained Optimizations

You can perform constrained optimizations which can, at times, be extremely helpful. This works as shown in the following example:

```
! RKS B3LYP/G SV(P) Opt
%geom Constraints
        { B 0 1 1.25 C }
        { A 2 0 3 120.0 C }
        end
      end

* int 0 1
    C  0 0 0 0.0000   0.000    0.00
    O  1 0 0 1.2500   0.000    0.00
    H  1 2 0 1.1075 122.016    0.00
    H  1 2 3 1.1075 122.016  180.00
*
```

```
Constraining bond distances        :   { B N1 N2 value C }
Constraining bond angles           :   { A N1 N2 N1 value C }
Constraining dihedral angles       :   { D N1 N2 N3 N4 value C }
Constraining cartesian coordinates :   { C N1 C }
```

Note

- Like for normal optimizations you can use numerical gradients (see *Numerical Gradients*.) for constrained optimizations. In this case the numerical gradient will be evaluated only for non-constrained coordinates, saving a lot of computational effort, if a large part of the structure is constrained.
- "value" in the constraint input is optional. If you do not give a value, the present value in the structure is constrained. For cartesian constraints you can't give a value, but always the initial position is constrained.
- It is recommended to use a value not too far away from your initial structure.
- It is possible to constrain whole sets of coordinates:

```
all bond lengths where N1 is involved       : { B N1 * C}
all bond lengths                            : { B * * C}
all bond angles where N2 is the central atom: { A * N2 * C }
all bond angles                             : { A * * * C }
all dihedral angles with central bond N2-N3 : { D * N2 N3 * C }
all dihedral angles                         : { D * * * * C }
```

- For Cartesian constraints lists of atoms can be defined:

```
a list of atoms (10 to 17) with Cartesian constraints  : { C 10:17 C}
```

- Coordinates along a single Cartesian direction can be frozen as described in section *Special definitions*.
- If there are only a few coordinates that have to be optimized you can use the `invertConstraints` option:

```
%geom Constraints
        { B 0 1 C }
        end
      invertConstraints true # only the C-O distance is optimized
                             # does not affect Cartesian coordinates
      end
```

- In some cases it is advantageous to optimize only the positions of the hydrogen atoms and let the remaining molecule skeleton fixed:

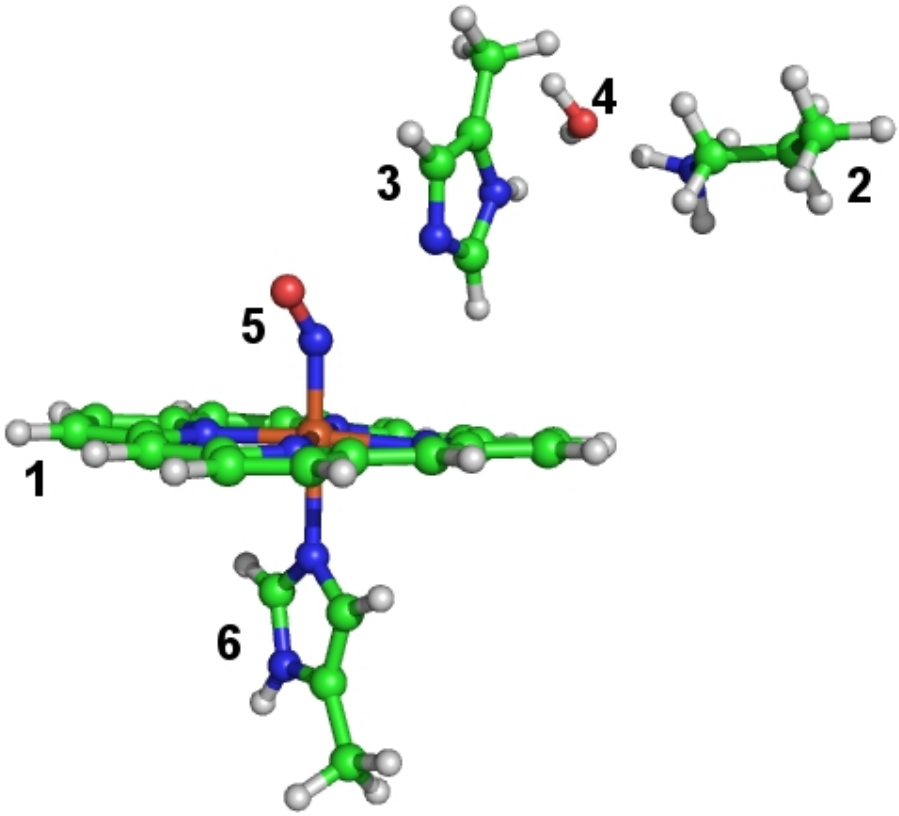
```
%geom optimizehydrogens true
      end
```

Note

- In the special case of a fragment optimization (see next point) the `optimizehydrogens` keyword does not fix the heteroatoms, but ensures that all hydrogen positions are relaxed.
- In Cartesian optimization, only Cartesian constraints are allowed.

6.3.7 Constrained Optimizations for Molecular Clusters (Fragment Optimization)

If you want to study systems, which consist of several molecules (e.g. the active site of a protein) with constraints, then you can either use cartesian constraints (see above) or use ORCA's fragment constraint option. ORCA allows the user to define fragments in the system. For each fragment one can then choose separately whether it should be optimized or constrained. Furthermore, it is possible to choose fragment pairs whose distance and orientation with respect to each other should be constrained. Here, the user can either define the atoms which make up the connection between the fragments, or the program chooses the atom pair automatically via a closest distance criterion. ORCA then chooses the respective constrained coordinates automatically. An example for this procedure is shown below.



The coordinates are taken from a crystal structure [PDB-code 2FRJ]. In our gas phase model we choose only a small part of the protein, which is important for its spectroscopic properties. Our selection consists of a heme-group (fragment 1), important residues around the reaction site (lysine (fragment 2) and histidine (fragment 3)), an important water molecule (fragment 4), the NO-ligand (fragment 5) and part of a histidine (fragment 6) coordinated to the heme-iron. In this constrained optimization we want to maintain the position of the heteroatoms of the heme group. Since the protein backbone is missing, we have to constrain the orientation of lysine and histidine (fragments 2 and 3) side chains to the heme group. All other fragments (the ones which are directly bound to the heme-iron and the water molecule) are fully optimized internally and with respect to the other fragments. Since the crystal structure does not reliably resolve the hydrogen positions, we relax also the hydrogen positions of the heme group.

```
# !! If you want to run this optimization: be aware
# !!  that it will take some time!
! BP86 SV(P) Opt
%geom
    ConstrainFragments { 1 } end # constrain all internal
                                 # coordinates of fragment 1
    ConnectFragments
     {1 2 C 12 28}  # connect the fragments via the atom pair 12/28 and 15/28 and
```

(continues on next page)

(continued from previous page)

```
     {1 3 C 15 28}  #  constrain the internal coordinates connecting
                    #  fragments 1/2 and 1/3
      {1 5 O}
      {1 6 O}
      {2 4 O}
      {3 4 O}
    end
    optimizeHydrogens true  # do not constrain any hydrogen position
  end
* xyz 1 2
Fe(1)     -0.847213     -1.548312     -1.216237 newgto "TZVP" end
 N(5)     -0.712253     -2.291076      0.352054 newgto "TZVP" end
 O(5)     -0.521243     -3.342329      0.855804 newgto "TZVP" end
 N(6)     -0.953604     -0.686422     -3.215231 newgto "TZVP" end
 N(3)     -0.338154     -0.678533      3.030265 newgto "TZVP" end
 N(3)     -0.868050      0.768738      4.605152 newgto "TZVP" end
 N(6)     -1.770675      0.099480     -5.112455 newgto "TZVP" end
 N(1)     -2.216029     -0.133298     -0.614782 newgto "TZVP" end
 N(1)     -2.371465     -2.775999     -1.706931 newgto "TZVP" end
 N(1)      0.489683     -2.865714     -1.944343 newgto "TZVP" end
 N(1)      0.690468     -0.243375     -0.860813 newgto "TZVP" end
 N(2)      1.284320      3.558259      6.254287
 C(2)      5.049207      2.620412      6.377683
 C(2)      3.776069      3.471320      6.499073
 C(2)      2.526618      2.691959      6.084652
 C(3)     -0.599599     -0.564699      6.760567
 C(3)     -0.526122     -0.400630      5.274274
 C(3)     -0.194880     -1.277967      4.253789
 C(3)     -0.746348      0.566081      3.234394
 C(6)      0.292699      0.510431     -6.539061
 C(6)     -0.388964      0.079551     -5.279555
 C(6)      0.092848     -0.416283     -4.078708
 C(6)     -2.067764     -0.368729     -3.863111
 C(1)     -0.663232      1.693332     -0.100834
 C(1)     -4.293109     -1.414165     -0.956846
 C(1)     -1.066190     -4.647587     -2.644424
 C(1)      2.597468     -1.667470     -1.451465
 C(1)     -1.953033      1.169088     -0.235289
 C(1)     -3.187993      1.886468      0.015415
 C(1)     -4.209406      0.988964     -0.187584
 C(1)     -3.589675     -0.259849     -0.590758
 C(1)     -3.721903     -2.580894     -1.476315
 C(1)     -4.480120     -3.742821     -1.900939
 C(1)     -3.573258     -4.645939     -2.395341
 C(1)     -2.264047     -4.035699     -2.263491
 C(1)      0.211734     -4.103525     -2.488426
 C(1)      1.439292     -4.787113     -2.850669
 C(1)      2.470808     -3.954284     -2.499593
 C(1)      1.869913     -2.761303     -1.932055
 C(1)      2.037681     -0.489452     -0.943105
 C(1)      2.779195      0.652885     -0.459645
 C(1)      1.856237      1.597800     -0.084165
 C(1)      0.535175      1.024425     -0.348298
 O(4)     -1.208602      2.657534      6.962748
 H(3)     -0.347830     -1.611062      7.033565
 H(3)     -1.627274     -0.387020      7.166806
 H(3)      0.121698      0.079621      7.324626
 H(3)      0.134234     -2.323398      4.336203
 H(3)     -1.286646      1.590976      5.066768
 H(3)     -0.990234      1.312025      2.466155
 H(4)     -2.043444      3.171674      7.047572
```

(continues on next page)

(continued from previous page)

```
 H(2)      1.364935      4.120133      7.126900
 H(2)      0.354760      3.035674      6.348933
 H(2)      1.194590      4.240746      5.475280
 H(2)      2.545448      2.356268      5.027434
 H(2)      2.371622      1.797317      6.723020
 H(2)      3.874443      4.385720      5.867972
 H(2)      3.657837      3.815973      7.554224
 H(2)      5.217429      2.283681      5.331496
 H(2)      5.001815      1.718797      7.026903
 H(6)     -3.086380     -0.461543     -3.469767
 H(6)     -2.456569      0.406212     -5.813597
 H(6)      1.132150     -0.595619     -3.782287
 H(6)      0.040799      1.559730     -6.816417
 H(6)      0.026444     -0.139572     -7.404408
 H(6)      1.392925      0.454387     -6.407850
 H(1)      2.033677      2.608809      0.310182
 H(1)      3.875944      0.716790     -0.424466
 H(1)      3.695978     -1.736841     -1.485681
 H(1)      3.551716     -4.118236     -2.608239
 H(1)      1.487995     -5.784645     -3.308145
 H(1)     -1.133703     -5.654603     -3.084826
 H(1)     -3.758074     -5.644867     -2.813441
 H(1)     -5.572112     -3.838210     -1.826943
 H(1)     -0.580615      2.741869      0.231737
 H(1)     -3.255623      2.942818      0.312508
 H(1)     -5.292444      1.151326     -0.096157
 H(1)     -5.390011     -1.391441     -0.858996
 H(4)     -1.370815      1.780473      7.384747
 H(2)      5.936602      3.211249      6.686961
*
```

Note

- You have to connect the fragments in such a way that the whole system is connected.
- You can divide a molecule into several fragments.
- Since the initial Hessian for the optimization is based upon the internal coordinates: Connect the fragments in a way that their real interaction is reflected.
- This option can be combined with the definition of constraints, scan coordinates and the `optimizeHydrogens` option (but: its meaning in this context is different to its meaning in a normal optimization run, relatively straightforward see section *Geometry Optimization*).
- Can be helpful in the location of complicated transition states (with relaxed surface scans).

6.3.8 Adding Arbitrary Wall Potentials

For some applications, it might be interesting to add arbitrary wall potentials during the geometry optimization. For example, if you want to optimize an intermolecular complex and need that both structures stick together, without one flying away during the optimization, or when using microsolvation.

In ORCA you can add three kinds of arbitrary "wall potentials": an ellipsoid or spherical of the form

$$V = \left(\frac{|\mathbf{R} - \mathbf{O}|}{radius} \right)^{30}$$

or a rectangular box potential with 6 walls of the form

$$V = e^{5(\mathbf{R} - wall)}$$

These can be given in two ways: by explicitly defining the origin of the potential and its limits, e.g:

```
%GEOM ELLIPSEPOT 0,0,0,5,3,4 # the last are the a,b and c radii
```

or:

```
%GEOM SPHEREPOT 0,0,0,5 # the last is the radius
```

or:

```
%GEOM BOXPOT 0,0,0,4,-4,3,-3,6,-6 # maxx, minx, maxy, miny, maxz and minz last
```

where the first three numbers are the center and the last is the radius for the sphere (or a,b and c for the ellipsoid) and the max and min x,y and z dimensions of the box. All numbers should be given in Ångström.

In case a single number is given instead, the walls will be automatically centered around the centroid of the molecule and that number will be added to the minimum sphere or box that is necessary to contain the molecule. For example:

```
%GEOM SPHEREPOT 2
```

or:

```
%GEOM BOXPOT 2
```

will build a minimum wall centered on the centroid that encloses the molecule and add 2 Ångström on top of it. Still on the sphere case, a negative number like

```
%GEOM SPHEREPOT -2
```

will make the total radius of the sphere to be Ångström.

OBS: This will apply to regular geometry optimizations, as well as to the Global Optimizer (GOAT).

6.3.9 Relaxed Surface Scans

A final thing that comes in really handy are relaxed surface scans, i.e. you can scan through one coordinate while all others are relaxed. It works as shown in the following example:

```
! B3LYP/G SV(P) Opt
%geom Scan
        B 0 1 = 1.35, 1.10, 12 # C-O distance that will be scanned
        end
      end

* int 0 1
    C  0 0 0 0.0000   0.000    0.00
    O  1 0 0 1.3500   0.000    0.00
    H  1 2 0 1.1075 122.016    0.00
    H  1 2 3 1.1075 122.016  180.00
*
```

In the example above the value of the bond length between C and O will be changed in 12 equidistant steps from 1.35 down to 1.10 Ångströms and at each point a constrained geometry optimization will be carried out.

Note

- If you want to perform a geometry optimization at a series of values with non-equidistant steps you can give this series in square brackets, [ ]. The general syntax is as follows:

```
B N1 N2 = initial-value, final-value, NSteps
or:
B N1 N2 [value1 value2 value3 ... valueN]
```

- In addition to bond lengths you can also scan bond angles and dihedral angles:

```
B N1 N2 = ...              # bond length
A N1 N2 N3 = ...           # bond angle
D N1 N2 N3 N4 = ...        # dihedral angle
```

Tip

- As in constrained geometry optimizations it is possible to start the relaxed surface scan with a different scan parameter than the value present in your molecule. But keep in mind that this value should not be too far away from your initial structure.

A more challenging example is shown below. Here, the H-atom abstraction step from CH_4 to OH-radical is computed with a relaxed surface scan (*vide supra*). The job was run as follows:

```
! B3LYP SV(P) Opt SlowConv NoTRAH
%geom scan B 1 0 = 2.0, 1.0, 15 end end
* int 0 2
 C     0   0   0   0.000000     0.000     0.000
 H     1   0   0   1.999962     0.000     0.000
 H     1   2   0   1.095870   100.445     0.000
 H     1   2   3   1.095971    90.180   119.467
 H     1   2   3   1.095530    95.161   238.880
 O     2   1   3   0.984205   164.404    27.073
 H     6   2   1   0.972562   103.807    10.843
*
```

It is obvious that the reaction is exothermic and passes through an early transition state in which the hydrogen jumps from the carbon to the oxygen. The structure at the maximum of the curve is probably a very good guess for the true transition state that might be located by a transition state finder.

You will probably find that such relaxed surface scans are incredibly useful but also time consuming. Even the simple job shown below required several hundred single point and gradient evaluations (convergence problems appear for the SCF close to the transition state and for the geometry once the reaction partners actually dissociate – this is to be expected). Yet, when you search for a transition state or you want to get insight into the shapes of the potential energy surfaces involved in a reaction it might be a good idea to use this feature. One possibility to ease the burden somewhat is to perform the relaxed surface scan with a "fast" method and a smaller basis set and then do single point calculations on all optimized geometries with a larger basis set and/or higher level of theory. At least you can hope that this should give a reasonable approximation to the desired surface at the higher level of theory – this is the case if the geometries at the lower level are reasonable.

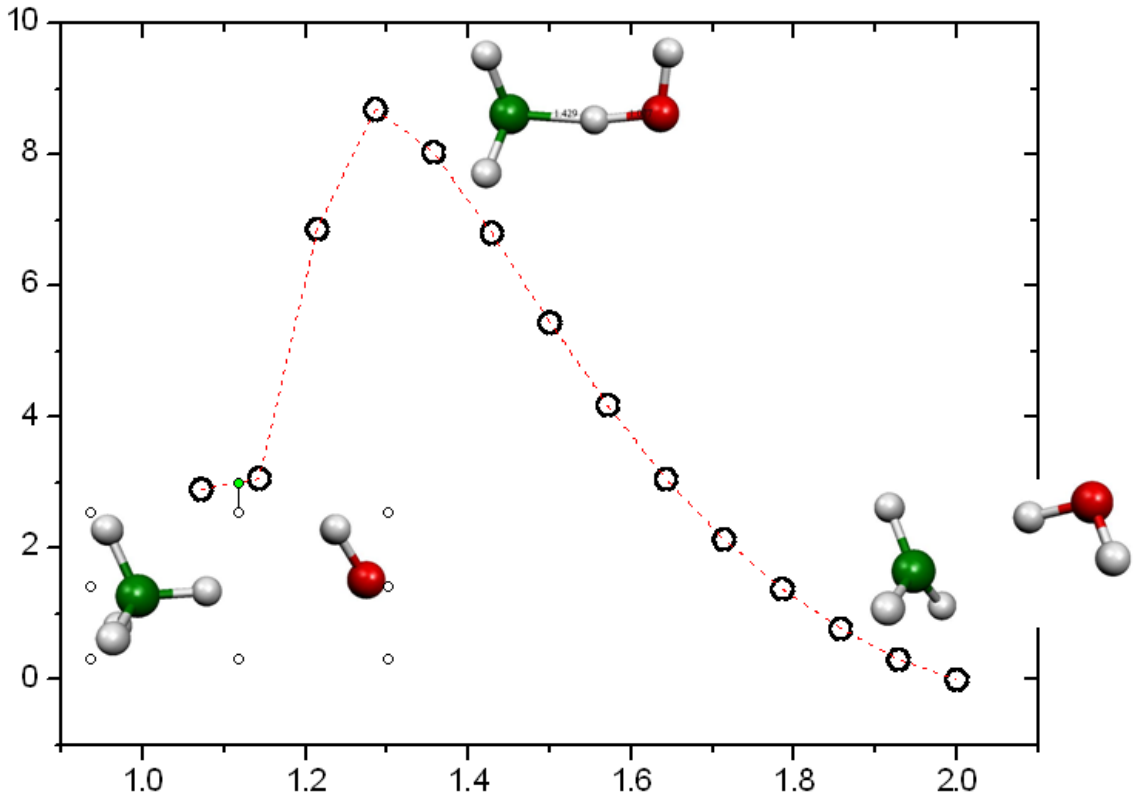


Fig. 6.22: Relaxed surface scan for the H-atom abstraction from CH_4 by OH-radical (B3LYP/SV(P)).

Multidimensional Scans

After several requests from our users ORCA now allows up to three coordinates to be scanned within one calculation.

```
! B3LYP/G SV(P) Opt
%geom Scan
        B 0 1 = 1.35, 1.10, 12 # C-O distance that will be scanned
        B 0 2 = 1.20, 1.00, 5  # C-H distance that will be scanned
        A 2 0 1 = 140, 100, 5  # H-C-O angle that will be scanned
        end
      end

* int 0 1
    C  0 0 0 0.0000   0.000    0.00
    O  1 0 0 1.3500   0.000    0.00
    H  1 2 0 1.1075 122.016    0.00
    H  1 2 3 1.1075 122.016  180.00
*
```

Note

- For finding transition state structures of more complicated reaction paths ORCA now offers its very efficient NEB-TS implementation (see section *Nudged Elastic Band Method*).
- 2-dimensional or even 3-dimensional relaxed surface scans can become very expensive - e.g. requesting 10 steps per scan, ORCA has to carry out 1000 constrained optimizations for a 3-D scan.
- The results can depend on the direction of the individual scans and the ordering of the scans.

Simultaneous multidimensional scans, in which all scan coordinates are changed at the same time, can be requested via the following keyword (which brings the cost of a multidimensional relaxed surface scan down to the cost of a single relaxed surface scan):

```
%geom
  Scan B 0 1 = 3, 1, 15 end
  Scan B 1 2 = 1, 3, 15 end
  Simul_Scan true
end
```

6.3.10 Multiple XYZ File Scans

A different type of scan is implemented in ORCA in conjunction with relaxed surface scans. Such scans produce a series of structures that are typically calculated using some ground state method. Afterwards one may want to do additional or different calculations along the generated pathway such as excited state calculations or special property calculations. In this instance, the "multiple XYZ scan" feature is useful. If you request reading from a XYZ file via:

```
* xyzfile Charge Multiplicity FileName
```

this file could contain a number of structures. The format of the file is:

```
Number of atoms M
  Comment line
AtomName1 X Y Z
AtomName2 X Y Z
...
AtomNameM X Y Z
>
Number of atoms M
  Comment line
AtomName1 X Y Z
...
```

Thus, the structures are simply of the standard XYZ format, separated by a ">" sign. After the last structure no ">" should be given but a blank line instead. The program then automatically recognizes that a multiple XYZ scan run is to be performed. Thus, single point calculations are performed on each structure in sequence and the results are collected at the end of the run in the same kind of trajectory `.dat` files as produced from trajectory calculations.

In order to aid in using this feature, the relaxed surface scans produce a file called `MyJob.allxyz` that is of the correct format to be re-read in a subsequent run.

6.3.11 Transition States

Introduction to Transition State Searches

If you provide a good estimate for the structure of the transition state (TS) structure, then you can find the respective transition state with the following keywords (in this example we take the structure with highest energy of the above relaxed surface scan):

```
! B3LYP SV(P) TightSCF SlowConv OptTS
# performs a TS optimization with the EF-algorithm
# Transition state: H-atom abstraction from CH4 to OH-radical

%geom
   Calc_Hess true  # calculation of the exact Hessian
                   #  before the first optimization step
end
```

(continues on next page)

(continued from previous page)

```
* int 0 2
    C  0 0 0 0.000000   0.000   0.000
    H  1 0 0 1.285714   0.000   0.000
    H  1 2 0 1.100174 107.375   0.000
    H  1 2 3 1.100975 103.353 119.612
    H  1 2 3 1.100756 105.481 238.889
    O  2 1 3 1.244156 169.257  17.024
    H  6 2 1 0.980342 100.836  10.515
*
```

Note

- You need a good guess of the TS structure. Relaxed surface scans can help in almost all cases (see also example above).
- For TS optimization (in contrast to geometry optimization) an exact Hessian, a Hybrid Hessian or a modification of selected second derivatives is necessary.
- Analytic Hessian evaluation is available for HF and SCF methods, including the RI and RIJCOSX approximations and canonical MP2.
- Check the eigenmodes of the optimized structure for the eigenmode with a single imaginary frequency. You can also visualize this eigenmode with `orca_pltvib` (section *Animation of Vibrational Modes*) or any other visualization program that reads ORCA output files.
- If the Hessian is calculated during the TS optimization, it is stored as basename.001.hess, if it is recalculated several times, then the subsequently calculated Hessians are stored as basename.002.hess, basename.003.hess, ...
- If you are using the Hybrid Hessian, then you have to check carefully at the beginning of the TS optimization (after the first three to five cycles) whether the algorithm is following the correct mode (see TIP below). If this is not the case you can use the same Hybrid Hessian again via the inhess read keyword and try to target a different mode (via the `TS_Mode` keyword, see below).

In the example above the TS mode is of local nature. In such a case you can directly combine the relaxed surface scan with the TS optimization with the

```
! ScanTS
```

command, as used in the following example:

```
! B3LYP SV(P) TightSCF SlowConv
! ScanTS # perform a relaxed surface scan and TS optimization
         #  in one calculation
%geom scan B 1 0 = 2.0, 1.0, 15 end end
* int 0 2
 C     0   0   0   0.000000     0.000     0.000
 H     1   0   0   1.999962     0.000     0.000
 H     1   2   0   1.095870   100.445     0.000
 H     1   2   3   1.095971    90.180   119.467
 H     1   2   3   1.095530    95.161   238.880
 O     2   1   3   0.984205   164.404    27.073
 H     6   2   1   0.972562   103.807    10.843
*
```

Note

- The algorithm performs the relaxed surface scan, aborts the Scan after the maximum is surmounted,

chooses the optimized structure with highest energy, calculates the second derivative of the scanned coordinate and finally performs a TS optimization.

- If you do not want the scan to be aborted after the highest point has been reached but be carried out up to the last point, then you have to type:

```
%geom
  fullScan true # do not abort the scan with !ScanTS
end
```

As transition state finder we implemented the quasi-Newton like Hessian mode following algorithm.[67, 241, 365, 389, 399, 513, 762, 763, 764] This algorithm maximizes the energy with respect to one (usually the lowest) eigenmode and minimizes with respect to the remaining $3N - 7(6)$ eigenmodes of the Hessian.

Tip

- You can check at an early stage if the optimization will lead to the "correct" transition state. After the first optimization step you find the following output for the redundant internal coordinates:

```
-------------------------------------------------------------------------
                    Redundant Internal Coordinates
                       (Angstroem and degrees)
   Definition                      Value    dE/dq     Step    New-Value  comp.(TS mode)
-------------------------------------------------------------------------
 1. B(H   1,C   0)                1.2857  0.013136  0.0286    1.3143          0.58
 2. B(H   2,C   0)                1.1002  0.014201 -0.0220    1.0782
 3. B(H   3,C   0)                1.1010  0.014753 -0.0230    1.0779
 4. B(H   4,C   0)                1.1008  0.014842 -0.0229    1.0779
 5. B(O   5,H   1)                1.2442 -0.015421 -0.0488    1.1954          0.80
 6. B(H   6,O   5)                0.9803  0.025828 -0.0289    0.9514
 7. A(H   1,C   0,H   2)          107.38 -0.001418   -0.88    106.49
 8. A(H   1,C   0,H   4)          105.48 -0.002209   -0.46    105.02
 9. A(H   1,C   0,H   3)          103.35 -0.003406    0.08    103.43
10. A(H   2,C   0,H   4)          113.30  0.001833    0.35    113.65
11. A(H   3,C   0,H   4)          113.38  0.002116    0.26    113.64
12. A(H   2,C   0,H   3)          112.95  0.001923    0.45    113.40
13. A(C   0,H   1,O   5)          169.26 -0.002089    4.30    173.56
14. A(H   1,O   5,H   6)          100.84  0.003097   -1.41     99.43
15. D(O   5,H   1,C   0,H   2)     17.02  0.000135    0.24     17.26
16. D(O   5,H   1,C   0,H   4)   -104.09 -0.000100    0.52   -103.57
17. D(O   5,H   1,C   0,H   3)    136.64  0.000004    0.39    137.03
18. D(H   6,O   5,H   1,C   0)     10.52  0.000078   -0.72      9.79
-------------------------------------------------------------------------
```

Every Hessian eigenmode can be represented by a linear combination of the redundant internal coordinates. In the last column of this list the internal coordinates, that represent a big part of the mode which is followed uphill, are labelled. The numbers reflect their magnitude in the TS eigenvector (fraction of this internal coordinate in the linear combination of the eigenvector of the TS mode). Thus at this point you can already check whether your TS optimization is following the right mode (which is the case in our example, since we are interested in the abstraction of H1 from C0 by O5.

- If you want the algorithm to follow a different mode than the one with lowest eigenvalue, you can either choose the number of the mode:

```
%geom
   TS_Mode {M 1} # {M 1} mode with second lowest eigenvalue
     end         #  (default: {M 0}, mode with lowest eigenvalue)
  end
```

or you can give an internal coordinate that should be strongly involved in this mode:

```
%geom
   TS_Mode {B 1 5} # bond between atoms 1 and 5,
     end           # you can also choose an angle: {A N1 N2 N1}
                   #  or a dihedral: {D N1 N2 N3 N4}
  end
```

Tip

- If you look for a TS of a breaking bond the respective internal coordinate might not be included in the list of redundant internal coordinates due to the bond distance being slightly too large, leading to slow or even no convergence at all. In order to prevent that behavior a region of atoms that are active in the TS search can be defined, consisting of e.g. the two atoms of the breaking bond. During the automatic generation of the internal coordinates the bond radii of these atoms (and their neighbouring atoms) are increased, making it more probable that breaking or forming bonds in the TS are detected as bonds.

```
%geom
   TS_Active_Atoms  { 1 2 3 } # atoms that are involved in TS, e.g. for proton
     end                      # transfer the proton, its acceptor and its donor
   TS_Active_Atoms_Factor 1.5 # factor by which the cutoff for bonds is increased for
                              # the above defined atoms.
                              # (Default 1.5, i.e. increased by 50%)
  end
```

Hessians for Transition State Calculations

For transition state (TS) optimization a simple initial Hessian, which is used for minimization, is not sufficient. In a TS optimization we are looking for a first order saddle point, and thus for a point on the PES where the curvature is negative in the direction of the TS mode (the TS mode is also called transition state vector, the only eigenvector of the Hessian at the TS geometry with a negative eigenvalue). Starting from an initial guess structure the algorithm used in the ORCA TS optimization has to climb uphill with respect to the TS mode, which means that the starting structure has to be near the TS and the initial Hessian has to account for the negative curvature of the PES at that point. The simple force-field Hessians cannot account for this, since they only know harmonic potentials and thus positive curvature.

The most straightforward option in this case would be (after having looked for a promising initial guess structure with the help of a relaxed surface scan) to calculate the exact Hessian before starting the TS optimization. With this Hessian (depending on the quality of the initial guess structure) we know the TS eigenvector with its negative eigenvalue and we have also calculated the exact force constants for all other eigenmodes (which should have positive force constants). For the HF, DFT methods and MP2, the analytic Hessian evaluation is available and is the best choice, for details see section Frequencies (*Vibrational Frequencies*).

When only the gradients are available (most notably the CASSCF), the numerical calculation of the exact Hessian is very time consuming, and one could ask if it is really necessary to calculate the full exact Hessian since the only special thing (compared to the simple force-field Hessians) that we need is the TS mode with a negative eigenvalue.

Here ORCA provides two different possibilities to speed up the Hessian calculation, depending on the nature of the TS mode: the Hybrid Hessian and the calculation of the Hessian value of an internal coordinate. For both possibilities the initial Hessian is based on a force-field Hessian and only parts of it are calculated exactly. If the TS mode is of very local nature, which would be the case when e.g. cleaving or forming a bond, then the exactly calculated part of the Hessian can be the second derivative of only one internal coordinate, the one which is supposed to make up the TS mode (the formed or cleaved bond). If the TS mode is more complicated and more delocalized, as e.g. in a concerted proton transfer reaction, then the hybrid Hessian, a Hessian matrix in which the numerical second derivatives are calculated only for those atoms, which are involved in the TS mode (for more details, see section *Geometry Optimization*), should be sufficient. If you are dealing with more complicated cases where these two approaches do not succeed, then you still have the possibility to start the TS optimization with a full exact Hessian.

Numerical Frequency calculations are quite expensive. You can first calculate the Hessian at a lower level of theory

or with a smaller basis set and use this Hessian as input for a subsequent TS optimization:

```
%geom
  inhess   Read                   # this command comes with:
  InHessName "yourHessian.hess"   # filename of Hessian input file
end
```

Another possibility to save computational time is to calculate exact Hessian values only for those atoms which are crucial for the TS optimization and to use approximate Hessian values for the rest. This option is very useful for big systems, where only a small part of the molecule changes its geometry during the transition and hence the information of the full exact Hessian is not necessary. With this option the coupling of the selected atoms are calculated exactly and the remaining Hessian matrix is filled up with a model initial Hessian:

```
%geom
  Calc_Hess true
  Hybrid_Hess {0 1 5 6} end   # calculates a Hybrid Hessian with
                              # exact calculation for atoms 0, 1, 5 and 6
end
```

For some molecules the PES near the TS can be very far from ideal for a Newton-Raphson step. In such a case ORCA can recalculate the Hessian after a number of steps:

```
%geom
  Recalc_Hess  5   # calculate the Hessian at the beginning
                   # and recalculate it after 5,10,15,... steps
end
```

Another solution in that case is to switch on the trust radius update, which reduces the step size if the Newton-Raphson steps behave unexpected and ensures bigger step size if the PES seems to be quite quadratic:

```
%geom
  Trust 0.3 # Trust <0 - use fixed trust radius (default: -0.3 au)
            # Trust >0 - use trust radius update, i.e. 0.3 means:
            # start with trust radius 0.3 and use trust radius update
end
```

Special Coordinates for Transition State Optimizations

- If you look for a TS of a breaking bond the respective internal coordinate might not be included in the list of redundant internal coordinates (but this would accelerate the convergence). In such a case (and of course in others) you can add coordinates to or remove them from the set of autogenerated redundant internal coordinates (alternatively check the TS_Active_Atoms keyword):

```
# add ( A ) or remove ( R ) internal coordinates
%geom
  modify_internal
      { B 10 0 A }      # add a bond between atoms 0 and 10
      { A 8 9 10 R }    # remove the angle defined
                        # by atoms 8, 9 and 10
      { D 7 8 9 10 R }  # remove the dihedral angle defined
  end                   # by atoms 7, 8, 9 and 10
end
```

6.3.12 MECP Optimization

There are reactions where the analysis of only one spin state of a system is not sufficient, but where the reactivity is determined by two or more different spin states (Two- or Multi-state reactivity). The analysis of such reactions reveals that the different PESs cross each other while moving from one stationary point to the other. In such a case you might want to use the ORCA optimizer to locate the point of lowest energy of the crossing surfaces (called the minimum energy crossing point, MECP).

As an example for such an analysis we show the MECP optimization of the quartet and sextet state of $[FeO]^+$.

```
!B3LYP TZVP Opt SurfCrossOpt SlowConv
%mecp
  Mult 4
 end
* xyz +1 6
Fe 0.000000     0.000000     0.000000
O  0.000000     0.000000     1.670000
*
```

- For further options for the MECP calculation, see section *Minimum Energy Crossing Points*.

Tip

You can often use a minimum or TS structure of one of the two spin states as initial guess for your MECP-optimization. If this doesn't work, you might try a scan to get a better initial guess.

The results of the MECP optimization are given in the following output. The distance where both surfaces cross is at 1.994 Å. In this simple example there is only one degree of freedom and we can also locate the MECP via a parameter scan. The results of the scan are given in Fig. 6.23 for comparison. Here we see that the crossing occurs at a Fe-O-distance of around 2 Å.

For systems with more than two atoms a scan is not sufficient any more and you have to use the MECP optimization.

```
**********************HURRAY********************
***         THE OPTIMIZATION HAS CONVERGED     ***
************************************************


---------------------------------------------------------------------
                  Redundant Internal Coordinates
                  --- Optimized Parameters ---
                  (Angstroem and degrees)
  Definition                    OldVal   dE/dq     Step     FinalVal
   ----------------------------------------------------------------------
1. B(O   1,Fe  0)               1.9939 -0.000001  0.0000    1.9939
   ----------------------------------------------------------


      ********************************************************
      *** FINAL ENERGY EVALUATION AT THE STATIONARY POINT ***
      ***               (AFTER    8 CYCLES)               ***
      ********************************************************

    -------------------------------------   ----------------
    Energy difference between both states       -0.000000061
    -------------------------------------   ----------------
```

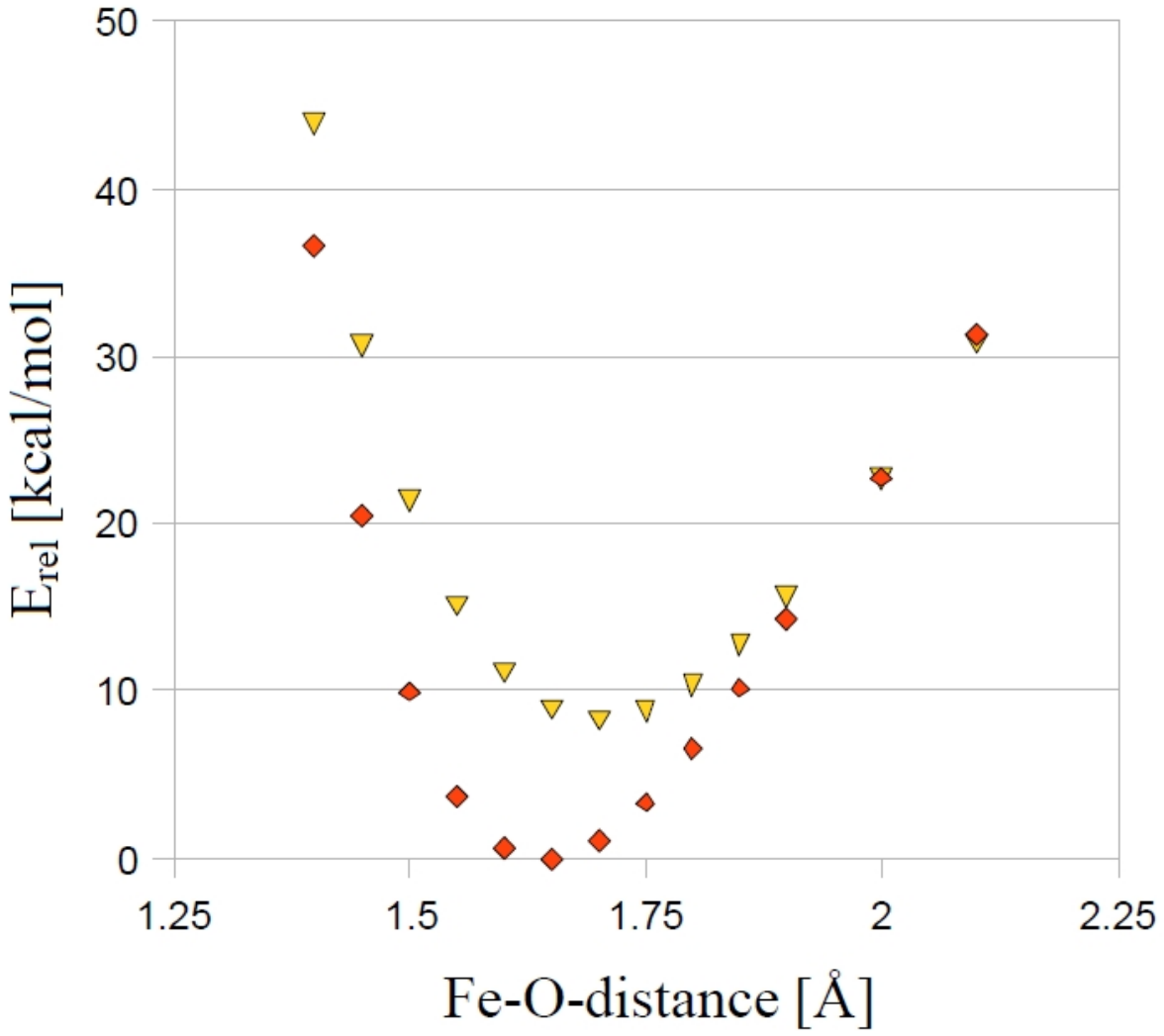


Fig. 6.23: Parameter scan for the quartet and sextet state of $[FeO]^+$ (B3LYP/SV(P)).

A more realistic example with more than one degree of freedom is the MECP optimization of a structure along the reaction path of the $CH_3O \leftrightarrow CH_2OH$ isomerization.

```
!B3LYP SV SurfCrossOpt SurfCrossNumFreq
%mecp Mult 1
  end
*xyz 1 3
 C     0.000000     0.000000     0.000000
 H     0.000000     0.000000     1.300000
 H     1.026719     0.000000    -0.363000
 O    -0.879955     0.000000    -1.088889
 H    -0.119662    -0.866667     0.961546
*
```

Note

- To verify that a stationary point in a MECP optimization is a minimum, you have to use an adapted frequency analysis, called by `SurfCrossNumFreq` (see section *Minimum Energy Crossing Points*).

6.3.13 Conical Intersection Optimization

OBS.: It is currently only available using TD-DFT, will be expanded in future versions. More details about the specific options on *Conical Intersections*.

A conical intersection (CI) is a complicated 3N-8 dimensional space, where two potential energy surfaces cross and the energy difference between these two states is zero. Inside this so-called "seam-space" minima and transition states can exist. Locating these minima is essential to understand photo-chemical processes, that are governed by non-adiabatic events, as e.g. photoisomerization, photostability - similar to locating transition states for chemical reactions.

As an example for such an analysis we show the conical intersection optimization of the ground and first excited state of singlet ethylene.

Note

Even though locating the CI of a TD-DFT excited state and the reference state is supported, it is not the recommended way of finding the ground state-excited state CI, because such CIs are not described properly by TD-DFT (in particular, TD-DFT even predicts the wrong dimensionality for the intersection space). Instead, it is advised to use SF-TD-DFT for this purpose, e.g. use the T_1 state as the reference state, and calculate both the S_0 and S_1 states as excited states. (*vide infra*)

```
!B3LYP DEF2-SVP CI-OPT
%TDDFT IROOT 1 END
* xyz 0 1
C        0.595560237     -0.010483480     -0.000284187
C       -0.831313750      0.167231832      0.001482505
H       -1.381857976      0.227877089      0.963419721
H        1.265119434      0.874806815      0.006897459
H       -1.382258208      0.243775568     -0.959090898
H        1.027489724     -1.032962768     -0.008829646
*
```

Tip

You can often use a structure between the optimized structures of both states for your CI-optimization. If this doesn't work, you might try a scan to get a better initial guess.

The results of the CI-optimization are given in the following output. The energy difference between the ground and excited state is printed as E diff. (CI), being reasonabley close for a conical intersection. For a description of the calculation of the non-adiabatic couplings at this geometry, see section *Numerical non-adiabatic coupling matrix elements*.

```
                        .--------------------.
----------------------|Geometry convergence|-------------------------
Item                 value                   Tolerance       Converged
-----------------------------------------------------------------------
Energy change        0.0000164283            0.0000050000      NO
E diff. (CI)         0.0000025162            0.0001000000      YES
RMS gradient         0.0000068173            0.0001000000      YES
MAX gradient         0.0000136891            0.0003000000      YES
RMS step             0.0000358228            0.0020000000      YES
MAX step             0.0000821130            0.0040000000      YES
.......................................................
Max(Bonds)      0.0000      Max(Angles)    0.00
Max(Dihed)        0.00      Max(Improp)    0.00
-----------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
             Everything but the energy has converged. However, the energy
             appears to be close enough to convergence to make sure that the
             final evaluation at the new geometry represents the equilibrium energy.
             Convergence will therefore be signaled now


                          **********************HURRAY********************
                          ***        THE OPTIMIZATION HAS CONVERGED     ***
                          ************************************************


    ---------------------------------------------------------------------------
                         Redundant Internal Coordinates

                           --- Optimized Parameters ---
                             (Angstroem and degrees)

       Definition                    OldVal   dE/dq     Step     FinalVal
    ---------------------------------------------------------------------------
     1. B(C   1,C   0)                1.3254 -0.000005 -0.0000    1.3254
     2. B(H   2,C   1)                1.1270  0.000004 -0.0000    1.1270
     3. B(H   3,C   0)                1.1271 -0.000002  0.0000    1.1271
     4. B(H   4,C   1)                1.1271  0.000000 -0.0000    1.1271
     5. B(H   5,C   0)                1.1271 -0.000002  0.0000    1.1271
     6. A(H   3,C   0,H   5)          106.00  0.000001   -0.00    106.00
     7. A(C   1,C   0,H   5)          126.97 -0.000013    0.00    126.97
     8. A(C   1,C   0,H   3)          127.03  0.000011   -0.00    127.03
     9. A(C   0,C   1,H   4)          127.03  0.000013   -0.00    127.03
     10. A(H   2,C   1,H   4)         106.01  0.000001   -0.00    106.01
     11. A(C   0,C   1,H   2)         126.96 -0.000014    0.00    126.96
     12. D(H   2,C   1,C   0,H   5)    73.60  0.000000   -0.00     73.59
     13. D(H   4,C   1,C   0,H   3)    72.78 -0.000001    0.00     72.78
     14. D(H   4,C   1,C   0,H   5)  -106.81  0.000001   -0.00   -106.82
     15. D(H   2,C   1,C   0,H   3)  -106.82 -0.000002    0.00   -106.81
    ---------------------------------------------------------------------------
                 ********************************************************
                 *** FINAL ENERGY EVALUATION AT THE STATIONARY POINT ***
                 ***               (AFTER   12 CYCLES)                ***
                 ********************************************************
```

CI minima between excited states In an analogous way, the conical intersection minima between two excited states can be requested by selection both an IROOT and a JROOT, shown below.

```
!B3LYP DEF2-SVP CI-OPT
%TDDFT IROOT 2
       JROOT 1
       #IROOTMULT TRIPLET would search in the triplet PES
       #SF TRUE would search for the S0-S1 CI from a T1 reference, using SF-TD-DFT
       # (but remember to set the multiplicity as 3 instead of 1)
END
* xyz 0 1
C        0.595560237     -0.010483480     -0.000284187
C       -0.831313750      0.167231832      0.001482505
H       -1.381857976      0.227877089      0.963419721
H        1.265119434      0.874806815      0.006897459
H       -1.382258208      0.243775568     -0.959090898
H        1.027489724     -1.032962768     -0.008829646
*
```

6.3.14 Constant External Force - Mechanochemistry

Constant external force can be applied on the molecule within the EFEI formalism[718] by pulling on the two defined atoms. To apply the external force, use the POTENTIALS in the geom block. The potential type is C for Constant force, indexes of two atoms (zero-based) and the value of force in nN.

```
! def2-svp  OPT
%geom
  POTENTIALS
    { C 2 3 4.0 }
  end
end

 * xyz 0 1
O          0.73020       -0.07940       -0.00000
O         -0.73020        0.07940       -0.00000
H          1.21670        0.75630        0.00000
H         -1.21670       -0.75630        0.00000
 *
```

The results are seen in the output of the SCF procedure, where the total energy already contains the force term.

```
----------------
TOTAL SCF ENERGY
----------------

Total Energy       :         -150.89704913 Eh           -4106.11746 eV

Components:
Nuclear Repulsion  :           36.90074715 Eh            1004.12038 eV
External potential :           -0.25613618 Eh              -6.96982 eV
Electronic Energy  :         -187.54166010 Eh           -5103.26802 eV
```

6.3.15 Intrinsic Reaction Coordinate

The Intrinsic Reaction Coordinate (IRC) is a special form of a minimum energy path, connecting a transition state (TS) with its downhill-nearest intermediates. A method determining the IRC is thus useful to determine whether a transition state is directly connected to a given reactant and/or a product.

ORCA features its own implementation of Morokuma and coworkers' popular method.[412] The IRC method can be simply invoked by adding the IRC keyword as in the following example.

```
! B3LYP SV(P) TightSCF KDIIS SOSCF Freq IRC
* xyz 0 2
  C  -0.000   0.001   -0.000
  H   1.290   0.005   -0.006
  H  -0.330   1.050   -0.002
  H  -0.252  -0.532   -0.929
  H  -0.286  -0.545    0.911
  O   2.499   0.220    0.065
  H   2.509   1.085    0.525
*
```

For more information and further options see section *Intrinsic Reaction Coordinate*.

Note

- The same method and basis set as used for optimization and frequency calculation should be used for the IRC run.

- The IRC keyword can be requested without, but also together with OptTS, ScanTS, NEB-TS, AnFreq and NumFreq keywords.
- In its default settings the IRC code checks whether a Hessian was computed before the IRC run. If that is not the case, and if no Hessian is defined via the %irc block, a new Hessian is computed at the beginning of the IRC run.
- A final trajectory (_IRC_Full_trj.xyz) is generated which contains both directions, forward and backward, by starting from one endpoint and going to the other endpoint, visualizing the entire IRC. Forward (_IRC_F_trj.xyz and _IRC_F.xyz) and backward (_IRC_B_trj.xyz and _IRC_B.xyz) trajectories and xyz files contain the IRC and the last geometry of that respective run.

6.3.16 Printing Hessian in Internal Coordinates

When a Hessian is available, it can be printed out in redundant internal coordinates as in the following example:

```
! opt
%geom inhess read
      inhessname "h2o.hess"
      PrintInternalHess true
end
*xyz 0 1
  O      0.000000    0.000000    0.000000
  H      0.968700    0.000000    0.000000
  H     -0.233013    0.940258    0.000000
*
```

Note

- The Hessian in internal coordinates is (for the input `printHess.inp`) stored in the file `printHess_internal.hess`.
- The corresponding lists of redundant internals is stored in `printHess.opt`.
- Although the `!Opt` keyword is necessary, an optimization is not carried out. ORCA exits after storing the Hessian in internal coordinates.

6.3.17 Using model Hessian from previous calculations

If you had a geometry optimization interrupted, or for some reason want to use the model Hessian updated from a previous calculation, you can do that by passing a `basename.opt` file, a `basename.carthess` file or the initial Hessian on a new calculation.

```
%GEOM InHess     READ
      InHessName "basename.carthess"
END
```

6.3.18 Geometry Optimizations using the L-BFGS optimizer

Optimizations using the L-BFGS optimizer are done in Cartesian coordinates. They can be invoked quite simple as in the following example:

```
! L-Opt
! MM
%mm
  ORCAFFFILENAME "CHMH.ORCAFF.prms"
end
*pdbfile 0 1 CHMH.pdb
```

Using this optimizer systems with 100s of thousands of atoms can be optimized. Of course, the energy and gradient calculations should not become the bottleneck for such calculations, thus MM or QM/MM methods should be used for such large systems.

The default maximum number of iterations is 200, and can be increased as follows:

```
! L-Opt
%geom
  maxIter 500 # default 200
 end
*pdbfile 0 1 CHMH.pdb
```

Only the hydrogen positions can be optimized with the following command:

```
! L-OptH
```

But also other elements can be exclusively optimized with the following command:

```
! L-OptH
%geom
 OptElement F  # optimize fluorine only when L-OptH is invoked.
               #  Does not work with the regular optimizer.
end
```

When fragments are defined for the system, each fragment can be optimized differently (similar to the fragment optimization described above). The following options are available:

FixFrags
: Freeze the coordinates of all atoms of the specified fragments.

RelaxHFrags
: Relax the hydrogen atoms of the specified fragments. Default for all atoms if !L-OptH is defined.

RelaxFrags
: Relax all atoms of the specified fragments. Default for all atoms if !L-Opt is defined.

RigidFrags
: Treat each specified fragment as a rigid body, but relax the position and orientation of these rigid bodies.

> **Note**
>
> - The fragments have to be defined after the coordinate input.

A more complex example is depicted in the following:

```
! L-OptH
%mm
  ORCAFFFILENAME "CHMH.ORCAFF.prms"
end
*pdbfile 0 1 CHMH.pdb
```

(continues on next page)

(continued from previous page)

```
%geom
 Frags
  2 {8168:8614} end     # First the fragments need to be defined
  3 {8615:8699} end     #  Note that all other atoms belong to
  4 {8700:8772} end     #  fragment 1 by default
  5 {8773:8791} end     #
 RelaxFrags {2} end     # Fragment 2 is fully relaxed
 RigidFrags {3 4 5} end # Fragments 3, 4 and 5 are treated as rigid bodies each.
end
```

6.3.19 Nudged Elastic Band Method

The Nudged Elastic Band (NEB) method is used to find a minimum energy path (MEP) connecting given reactant and product state minima on the energy surface. An initial path is generated and represented by a discrete set of configurations of the atoms, referred to as images of the system. The number of images is specified by the user and has to be large enough to obtain sufficient resolution of the path. The implementation in ORCA is described in detail in the article by Ásgeirsson et. al.[4] and in section *Nudged Elastic Band Method* along with the input options. The most common use of the NEB method is to find the highest energy saddle point on the potential energy surface specifying the transition state for a given initial and final state. Rigorous convergence to a first order saddle point can be obtained with the climbing image NEB (CI-NEB), where the highest energy image is pushed uphill in energy along the tangent to the path while relaxing downhill in orthogonal directions. Another method for finding a first order saddle point is the NEB-TS which uses the CI-NEB method with a loose tolerance to begin with and then switches over to the OptTS method to converge on the saddle point. This combination can be a good choice for calculations of complex reactions where the ScanTS method fails or where 2D relaxed surface scans are necessary to find a good initial guess structure for the OptTS method. The zoomNEB variants are a good choice in case of very complex transition states with long tails. For more and detailed information on the various NEB variants implemented in ORCA please consult section *Nudged Elastic Band Method*.

In their simplest form NEB, NEB-CI and NEB-TS only require the reactant and product state configurations (one via the xyz block, and the other one via the keyword neb_end_xyzfile):

```
!NEB-TS # or !NEB or !NEB-CI or !ZOOM-NEB-TS or !ZOOM-NEB-CI
        # or !Fast-NEB-TS  (corresponds to IDPP-TS defined in the NEB-TS manuscript)
        # or !Loose-NEB-TS (corresponds to default NEB-TS in the NEB-TS manuscript)
%neb
 neb_end_xyzfile "final.xyz"
end
```

Below is an example of an NEB-TS run involving an intramolecular proton transfer within acetic acid. The simplest input is

```
!XTB NEB-TS
%neb
 neb_end_xyzfile "final.xyz"
end

*xyz 0 1
C      0.416168     0.038758    -0.014077
C      0.041816     0.011798     1.439610
O      1.524458     0.176600    -0.453888
O     -0.654209    -0.127881    -0.803857
H     -0.391037    -0.126036    -1.737478
H     -0.913438     0.507022     1.585301
H     -0.057787    -1.026455     1.750845
H      0.819515     0.485425     2.030252
*
```

Where the final.xyz structure contains the corresponding structure with the proton on the other oxygen.

The initial path is reasonable and the CI calculation can be switched on after five NEB iterations.

```
Starting iterations:

Optim.  Iteration  HEI  E(HEI)-E(0)  max(|Fp|)   RMS(Fp)    dS
Switch-on CI threshold                0.020000
   LBFGS     0      4    0.081017    0.073897   0.018915  3.2882
   LBFGS     1      5    0.070244    0.056668   0.013913  3.2770
   LBFGS     2      5    0.062934    0.038972   0.008763  3.3376
   LBFGS     3      5    0.057358    0.032076   0.006535  3.3950
   LBFGS     4      4    0.053260    0.019015   0.003599  3.4826

Image  4 will be converted to a climbing image in the next iteration (max(|Fp|) < 0.0200)

Optim.  Iteration  CI   E(CI)-E(0)   max(|Fp|)   RMS(Fp)    dS     max(|FCI|)   RMS(FCI)
Convergence thresholds                0.020000   0.010000           0.002000    0.001000
```

The CI run converges after another couple of iterations:

```
*********************H U R R A Y*********************
***         THE NEB OPTIMIZATION HAS CONVERGED      ***
*****************************************************
```

Subsequently a summary of the MEP is printed:

```
-----------------------------------------------------------------
                        PATH SUMMARY
-----------------------------------------------------------------
All forces in Eh/Bohr.

Image Dist.(Ang.)    E(Eh)   dE(kcal/mol)  max(|Fp|)  RMS(Fp)
  0     0.000     -14.45993      0.00       0.00011   0.00004
  1     0.426     -14.44891      6.91       0.00092   0.00033
  2     0.652     -14.42864     19.63       0.00084   0.00038
  3     0.805     -14.41132     30.50       0.00075   0.00027
  4     0.932     -14.40562     34.08       0.00057   0.00018 <= CI
  5     1.044     -14.41047     31.03       0.00057   0.00024
  6     1.153     -14.42200     23.80       0.00103   0.00034
  7     1.280     -14.43666     14.60       0.00098   0.00037
  8     1.476     -14.45106      5.56       0.00106   0.00033
  9     1.869     -14.45988      0.03       0.00013   0.00006
```

Additionally, detailed information on the highest energy image (or the CI) is printed:

```
-----------------------------------------------------------------
               INFORMATION ABOUT SADDLE POINT
-----------------------------------------------------------------

Climbing image                          ....  4
Energy                                  ....  -14.40561577 Eh
Max. abs. force                         ....  9.5976e-04 Eh/Bohr

----------------------------------------
  SADDLE POINT (ANGSTROEM)
----------------------------------------
C      0.040867     0.007347    -0.497635
C     -0.075595     0.017879     0.979075
O      1.122340     0.126074    -1.145534
O     -0.928470    -0.137946    -1.298318
H      0.165808    -0.021676    -2.055704
H     -0.996979     0.514720     1.271668
H     -0.116377    -1.013504     1.327873
H      0.788406     0.507105     1.418575
```

(continues on next page)

(continued from previous page)

```
----------------------------------------
  FORCES (Eh/Bohr)
----------------------------------------
C      -0.000646    -0.000111     0.000086
...

----------------------------------------
  UNIT TANGENT
----------------------------------------
C      -0.246569    -0.031821    -0.019359
...

 => Unit tangent is an approximation to the TS mode at the saddle point

Next a TS optimization is performed on the CI from the NEB run.
```

Finally, a TS optimization is started, after which the MEP information is updated by including the TS structure:

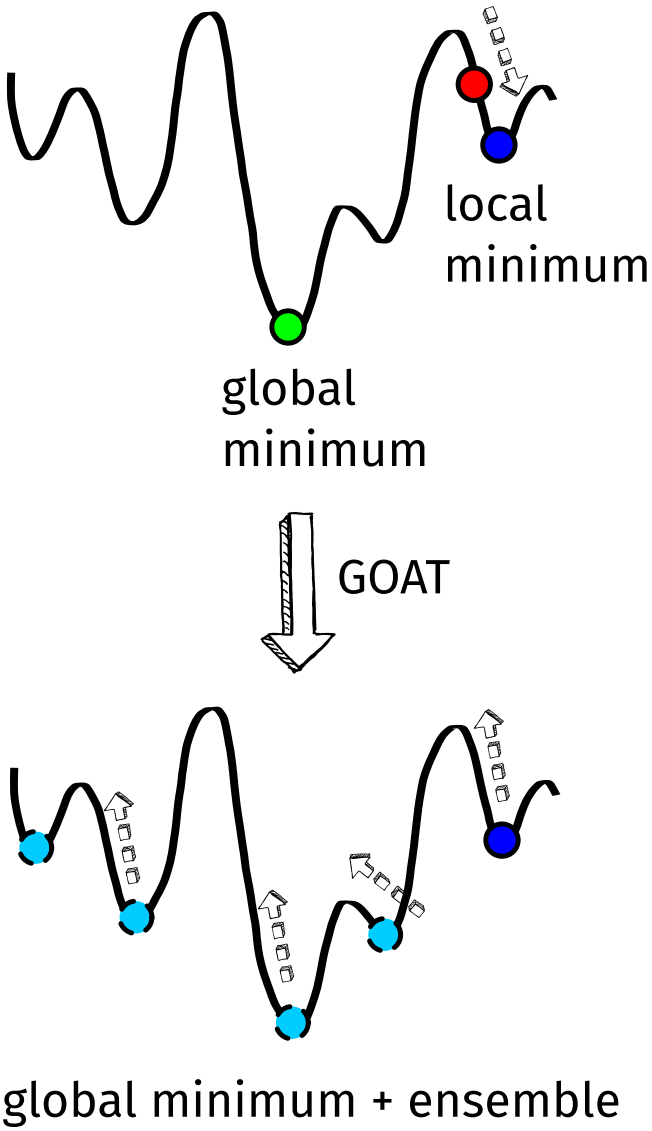
```
-----------------------------------------------------------------
                     PATH SUMMARY FOR NEB-TS
-----------------------------------------------------------------
All forces in Eh/Bohr. Global forces for TS.

Image     E(Eh)   dE(kcal/mol)  max(|Fp|)  RMS(Fp)
  0     -14.45993     0.00        0.00011   0.00004
  1     -14.44891     6.91        0.00092   0.00033
  2     -14.42864    19.63        0.00084   0.00038
  3     -14.41132    30.50        0.00075   0.00027
  4     -14.40562    34.08        0.00057   0.00018 <= CI
 TS     -14.40562    34.08        0.00033   0.00013 <= TS
  5     -14.41047    31.03        0.00057   0.00024
  6     -14.42200    23.80        0.00103   0.00034
  7     -14.43666    14.60        0.00098   0.00037
  8     -14.45106     5.56        0.00106   0.00033
  9     -14.45988     0.03        0.00013   0.00006
```

Note that here both TS and CI are printed for comparison.

6.4 GOAT: global geometry optimization and ensemble generator

If instead of trying to optimize a single structure, starting from a given guess geometry, you want to find the **global** minimum or the ensemble around it, ORCA features a Global Optimizer Algorithm (GOAT) inspired by Wales and Doye's basin-hopping [877], Goedecker's minima hopping [306], Simulated Annealing and Taboo Search.

The idea is to start from somewhere on the potential energy surface (PES; red ball on Fig. 6.24), go first to the nearest local minimum (blue ball), and from there start pushing "uphill" on a random direction until a barrier is crossed. Then a new minimum is found and the process is restarted, with another uphill push followed by an optimization. After several of these GOAT iterations (uphill + downhill), if no new global minimum was found between the two last global iterations.



global minimum + ensemble

Fig. 6.24: A simple depiction of the difference between a regular geometry optimization (above), and the GOAT global optimizer (below). By using the latter, one finds not only one local minimum but the global one and the conformational ensemble around it.

Since structures are collected along the way to the global minimum, we have in the end not only the global minimum, but also the conformational ensemble for that molecule, meaning all the conformations it **can** have and their relative energies. This is also useful later to compute Boltzmann-averaged spectra and properties.

The idea is similar to what is done with CREST from the group of Prof. Grimme [697], except that no metadynamics is required and thus much less gradient runs are needed. It is thus suitable not only for super fast methods such as XTB and force-fields, but can also be used directly using DFT or with any method available in ORCA.

Please note that there is no *ab initio* way to find global minima for arbitrary unknown functions, and stochastic methods are the most efficient on finding these. The drawback is that it is based on random choices, so that many geometry optimizations are needed - here in the order of $100\times$ the number of atoms. Good news is: these can be efficiently parallelized (even multinode) and this number can be brought down to less than $3\times$ the number of atoms (see below)!

6.4.1 GOAT simple usage example - Histidine

Let's start with a simple example, the amino acid histidine:

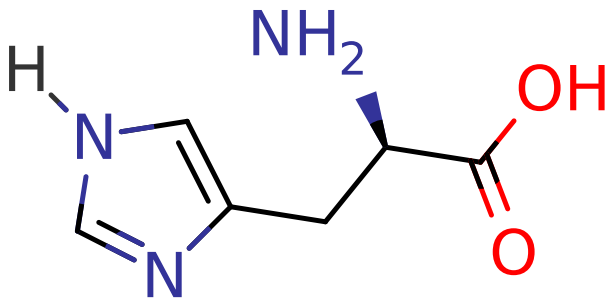


Fig. 6.25: A histidine molecule.

By simply looking at its Lewis structure, it is not at all evident that there are actually at least 20 conformers in the 3 kcal/mol range from the global minimum on the XTB PES! In order to find them, one can run:

```
!XTB GOAT #XTB version 6.4.0
* xyz 0 1
N          -0.13033       -0.28496       -0.67901
C           1.30551       -0.36383       -0.41824
C           1.51611       -1.04435        0.94169
O           0.58926       -1.43597        1.64771
C           1.97932        1.02323       -0.44331
H           2.41593        1.27074        0.53237
H           1.24598        1.81337       -0.65132
C           3.04894        1.09545       -1.48009
N           2.77779        1.01389       -2.82650
C           3.97051        1.07618       -3.48654
H           4.04071        1.01588       -4.56429
N           4.97724        1.21154       -2.65295
C           4.41545        1.24248       -1.40412
H           5.02774        1.35929       -0.51882
H           1.85722        0.88822       -3.24287
H          -0.75420       -0.54264        0.08211
H           1.72153       -1.03742       -1.17913
H          -0.47113        0.20468       -1.50142
O           2.79298       -1.20515        1.35271
H           3.59528       -0.86615        0.74156
*
```

The command to call the global optimizer is simply `!GOAT`, like you would with `!OPT`, and its options can be given under the `%GOAT` block as usual. You can give it together with any other method available in ORCA, but it needs to be a fast one because a lot of geometry optimizations need to be done. Here we will just use `GFN2` (or `!XTB`). What will happen next is:

1. First a regular geometry optimization will be done to find the minimum closest to the input structure.
2. With that information in hand, the number of necessary GOAT iterations will be computed and divided among `NWorkers` (8 by default).
3. Each `Worker` has its own parameters and will run a certain number of geometry optimizations.
4. After all workers in a global cycle are done, data will be collected and a new cycle will begin. There will be at least a "Minimum global steps" number of global cycles like this.
5. Once the difference between two global steps is negligible, it stops, collects everything and prints the ensemble energies and a file with all structures.

6.4.2 Understanding the output

After the usual geometry optimization, the output looks like:

```
Global parameters
-----------------
GOAT version                              ... default
Minimum global steps                      ... 3
Number of base workers                    ... 4
Split workers by                          ... 2
Final number of workers                   ... 8
Number of available CPUs                  ... 16
Parameter list (worker : temperature)     ...  0 : 2903.97,  1 : 1451.98,
                                               2 :  725.99,  3 :  363.00,
                                               4 : 2903.97,  5 : 1451.98,
                                               6 :  725.99,  7 :  363.00
GradComp (mean : sigma)                   ... 1.00 : 0.50
Number of atoms                           ... 20
Number of fragments                       ... 1
Flexibility parameter                     ... 0.45
Optimizations per global step             ... 160
Optimizations per worker                  ... 20


Filtering criteria
------------------
RMSD                                      ... 0.125 Angs (atom. pos.)
EnDiff                                    ... 0.100 kcal/mol
RotConst                                  ... 1.00-2.50 %
Maximum Conf. Energy                      ... 6.000 kcal/mol

Thermodynamics
--------------
Ensemble temperature                      ... 298.15 K
Degeneracy of conformers                  ... 1
*No rotamers will be included in Gconf
```

On the top there is some general information. Most important here is that we have 8 workers and 1 CPU, meaning each worker will run only after the other is done. If you want to speed up, just add more CPUs via e.g., `!PAL8` and workers will run in parallel making it 8x faster. GOAT can also run multinode, so feel free to use any number of processors via `%PAL`. In the end the filtering criteria used to differentiate conformers and rotamers is printed.

The default filtering is precisely the same as that of CREST: RMSD of atomic positions together with the rotational constant, considering its anisotropy. For `GOAT-EXPLORE` (see below), the default RMSD metric is based on the eigenvalues of the distance matrix instead, since it is invariant to translations, rotations and the atom ordering, which changes quite often in these cases.

After that, the algorithm starts:

```
-------------------------------------------------------------------------------
Iter  MinTemp     MaxEn       GradComp    NOpt   NProcs    Output
-------------------------------------------------------------------------------
0     2903.97     60.00    1.00 : 0.50      20        2    HIS.goat.0.0.out
1     1451.98     60.00    1.00 : 0.50      20        2    HIS.goat.0.1.out
2      725.99     60.00    1.00 : 0.50      20        2    HIS.goat.0.2.out
3      363.00     60.00    1.00 : 0.50      20        2    HIS.goat.0.3.out
4     2903.97     60.00    1.00 : 0.50      20        2    HIS.goat.0.4.out
5     1451.98     60.00    1.00 : 0.50      20        2    HIS.goat.0.5.out
6      725.99     60.00    1.00 : 0.50      20        2    HIS.goat.0.6.out
7      363.00     60.00    1.00 : 0.50      20        2    HIS.goat.0.7.out

                            GOAT Global Iter 1
                Iter    Min En       Sconf       Gconf
```

(continues on next page)

(continued from previous page)

```
                        Hartree      cal/(molK)  kcal/mol
                  ========================================

                    1    -34.346656   4.432       -0.551

0     2903.97     60.00    1.00 : 0.50      20        2    HIS.goat.0.0.out
1     1451.98     60.00    1.00 : 0.50      20        2    HIS.goat.0.1.out
2      725.99     60.00    1.00 : 0.50      20        2    HIS.goat.0.2.out
3      363.00     60.00    1.00 : 0.50      20        2    HIS.goat.0.3.out
4     2903.97     60.00    1.00 : 0.50      20        2    HIS.goat.0.4.out
5     1451.98     60.00    1.00 : 0.50      20        2    HIS.goat.0.5.out
6      725.99     60.00    1.00 : 0.50      20        2    HIS.goat.0.6.out
7      363.00     60.00    1.00 : 0.50      20        2    HIS.goat.0.7.out

                           GOAT Global Iter 2
                  Iter    Min En       Sconf       Gconf
                          Hartree      cal/(molK)  kcal/mol
                  ========================================

                    1    -34.346656   4.432       -0.551
                    2    -34.346656   4.528       -0.559

0     2903.97     60.00    1.00 : 0.50      20        2    HIS.goat.0.0.out
1     1451.98     60.00    1.00 : 0.50      20        2    HIS.goat.0.1.out
2      725.99     60.00    1.00 : 0.50      20        2    HIS.goat.0.2.out
3      363.00     60.00    1.00 : 0.50      20        2    HIS.goat.0.3.out
4     2903.97     60.00    1.00 : 0.50      20        2    HIS.goat.0.4.out
5     1451.98     60.00    1.00 : 0.50      20        2    HIS.goat.0.5.out
6      725.99     60.00    1.00 : 0.50      20        2    HIS.goat.0.6.out
7      363.00     60.00    1.00 : 0.50      20        2    HIS.goat.0.7.out

                           GOAT Global Iter 3
                  Iter    Min En       Sconf       Gconf
                          Hartree      cal/(molK)  kcal/mol
                  ========================================

                    1    -34.346656   4.432       -0.551
                    2    -34.346656   4.528       -0.559
                    3    -34.346656   4.541       -0.560

                                Global minimum found!
                         Writing structure to HIS.globalminimum.xyz
```

On the top header one can see what are the temperatures used, the maximum energy allowed during an uphill step, the maximum coefficient for gradient reflection, the number of optimizations done per worker, the number of processors used for each and the local output file name.

The names of the output files are chosen as `BaseName.goat.globaliteration.workernumber.out`. These are deleted after the run by default, but can be kept by setting `KEEPWORKERDATA TRUE` under `%GOAT`.

During each global iteration, the minimum energy so far, the conformational entropy S_{conf} and the conformational Gibbs free energy G_{conf} are printed. Since **there are no rotamers here**, the entropy is calculated only on the basis of the conformer energies and its convergence hints to the completeness of the ensemble created.

6.4.3 The final ensemble

In this case, as you can see, it already found the global minimum after the first global cycle with E = -34.346656 Hartree, but it keeps running for at least 3 cycles, following the defaults. This happens because it is a small molecule, but it is not necessarily so and more cycles will be done if needed.

The final relative energies of the ensemble are printed afterwards, together with a `BaseName.finalensemble.xyz` file:

```
                # Final ensemble info #
                Conformer     Energy     Degen.   % total   % cumul.
                              (kcal/mol)
                -----------------------------------------------------
                        0     0.000          1      37.96      37.96
                        1     0.503          1      16.25      54.20
                        2     0.914          1       8.11      62.32
                        3     1.159          1       5.37      67.68
                        4     1.297          1       4.25      71.94
                        5     1.320          1       4.09      76.03
                                           (...)
                       41     5.180          1       0.01      99.98
                       42     5.199          1       0.01      99.99
                       43     5.491          1       0.00      99.99
                       44     5.547          1       0.00      99.99
                       45     5.861          1       0.00     100.00

Conformers below 3 kcal/mol: 22
Lowest energy conformer    : -34.346656 Eh
Sconf at 298.15 K          :  4.54 cal/(molK)
Gconf at 298.15 K          : -0.56 kcal/mol

Writing final ensemble to HIS.finalensemble.xyz
```

Just for the record, here is how the four lowest lying conformers look like:

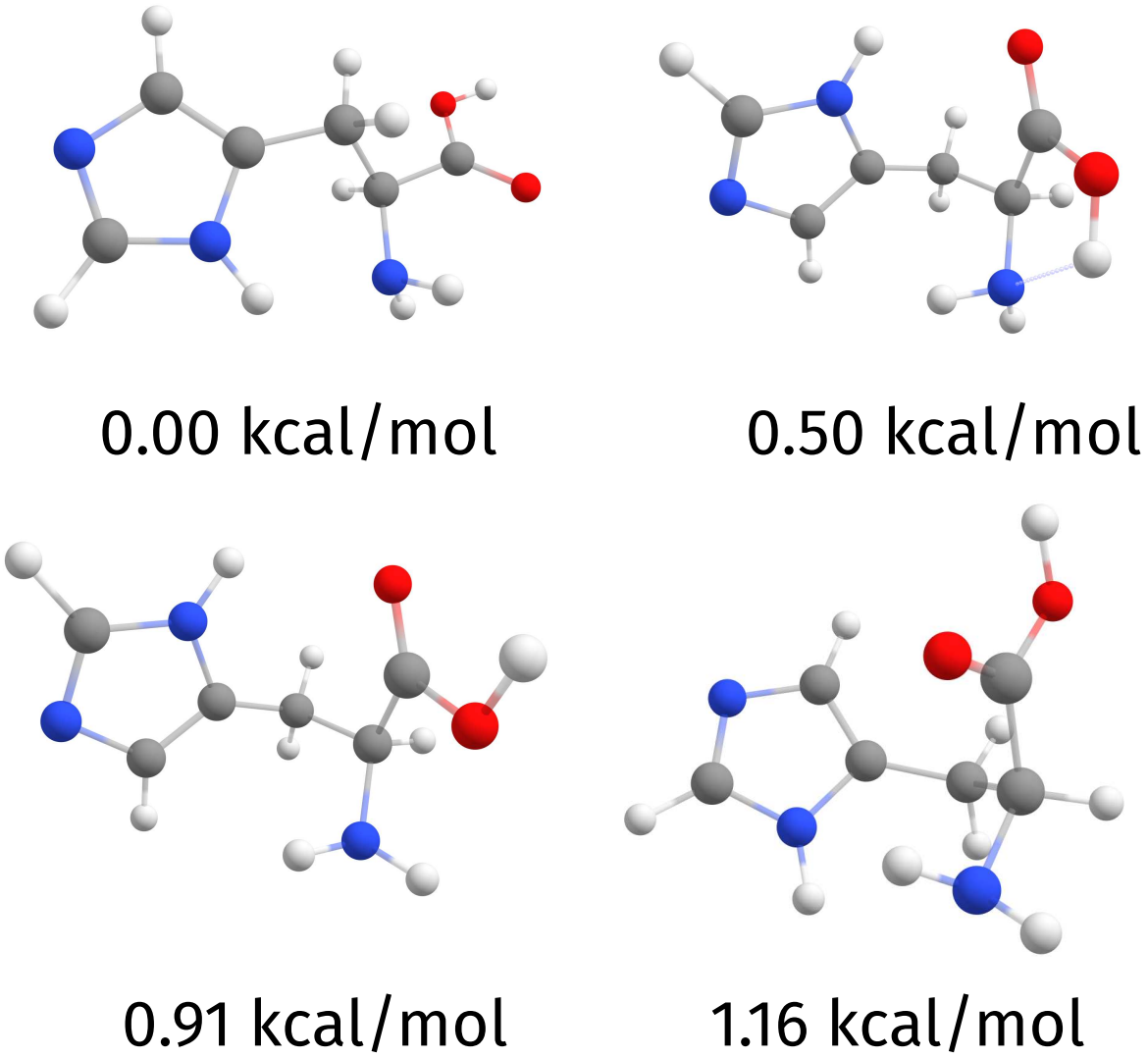


Fig. 6.26: The four lowest conformers found for histidine on the XTB PES.

Whenever using `GOAT`, the `EnforceStrictConvergence` policy from ORCA's optimizer is set to `TRUE`. This is recommended here the ensure equal criteria for different molecules in the ensemble. It can be turned off by setting `%GEOM EnforceStrictConvergence FALSE END`.

The regular optimization thresholds are already good, but it might also be a good idea to use `!TIGHTOPT` to make sure all your ensemble molecules are well converged, specially for the `GOAT-EXPLORE`!

6.4.4 GOAT-ENTROPY: expanding ensemble completeness by maximizing entropy

If you want to be as complete as possible in terms of the ensemble, you can use the `!GOAT-ENTROPY` keyword instead. This will not only try to find the global minimum until the energy is converged, but will actually only stop when the ΔS_{conf} also converges to less than 0.1 cal/(molK), which is equivalent to maximizing the conformational entropy (the threshold can be altered – see keyword list below).

This will push the algorithm so that all conformers around the global minimum should be found together with it. Both temperature and ΔS_{conf} can be changed via specific keywords shown at the list below. A higher temperature will make the ΔS_{conf} more sensitive to changes in high energy conformational regions and should make the search even more complete.

Being more explicit, the conformational entropy, enthalpy and Gibbs free energy are calculated according to:

$$S_{\text{conf}} = R\left[\ln \sum g_i' e^{-E_i\beta} + \frac{\sum g_i'(E_i\beta)e^{-E_i\beta}}{\sum g_i'(E_i\beta)e^{-E_i\beta}}\right]$$

$$\left[H(T) - H(0)\right]_{\text{conf}} = RT\frac{\sum g_i(E_i\beta)e^{-E_i\beta}}{\sum g_i(E_i\beta)e^{-E_i\beta}}$$

where $\beta = \frac{1}{k_B T}$ and g_i is the "degeneracy" of conformer i, i.e. its number of rotamers. This is the correct approach to deal with degenerate states [699]. The only difference from the reference above is that g_i' is always one and we don't discriminate any factor for "geometrical enantiomers".

6.4.5 More on the ΔS_{conf}

It is important to say that, by default, the ΔS_{conf} **is not the same as that found by a default CREST run**. There it includes also the rotamer degeneracy on the calculation of the entropy, while here that is 1. The reasons for that are:

1. There are formal arguments for using only one, assuming that rotamers are indistinguishable. Please check Grimme's reference [699] for details.
2. For systems with many rotamers, e.g. molecules with 3 *tert*-butyl groups which give rise to at least (27^3) 19683 rotamers per conformer, the algorithm will never find all of these anyway.
3. When calling `GOAT-ENTROPY`, it is the entropy of the **conformers** that will be maximized, not of the ensemble. That guarantees the maximal distribution of different conformers during these searches.

Once the final ensemble is found and if you know how many rotamers per conformer you have (assuming a constant number, like the *tert*-butyl case), one can reset that number by using `READENSEMBLE "ensemble.xyz"` to read it and `CONFDEGEN` to set a degeneracy. In the previous example that would be `CONFDEGEN 19683` and would give you the desired ΔG_{conf} for that given ensemble.

Finding rotamers automatically

It is always possible to switch on the automatic search for rotamers and their degeneracy by setting `CONFDEGEN AUTO` if that is what you want. They will be added to the ensemble instead of being filtered out and the full ensemble will be saved in files named `.confrot.xyz`.

Another approach would be to assume that, since the algorithm might find all rotamers for some conformers but not for all, one might set the degeneracy of all equal to the maximum value found so far (`CONFDEGEN AUTOMAX`). Let's take as an example a system with a *tert*-butyl + a methyl group with 81 rotamers per conformer. GOAT will hardly find 81 rotamers for every single conformer, but if it finds them for one conformer, all the others will also have that same number.

Please be aware that there are cases where different conformers might have different numbers of rotamers (e.g. decane or long alkyl chains), and these cases should be treated with care.

6.4.6 GOAT-EXPLORE: global minima of atomic clusters or topology-free free PES searches

In case you want to find the lowest energy conformer for a cluster or don't want to keep the initial topology at all, you can use the `!GOAT-EXPLORE` option instead. This will possibly break all bonds and find the lowest energy structure for that given set of atoms, be that a nanoparticle or an organic molecule.

For instance, let's find the minimum of an Au_8 nanoparticle on the GFN1 PES, starting from just a random agglomeration of gold atoms:

```
!XTB1 GOAT-EXPLORE PAL16 #XTB version 6.4.0
%GOAT NWORKERS 16 END
* xyz 0 1
Au        -1.39858        2.62611       -0.79278
Au        -2.50552       -0.07122        0.67538
Au        -0.52174       -2.57892       -0.02415
Au         0.78881        0.39733        0.21816
Au         1.21116        1.90621       -2.64617
Au        -1.19205       -0.30099       -2.30429
Au        -0.33808       -1.07129        2.90822
Au        -0.85565        2.15342        2.41956
*
```

Please note that there is a minimum number of optimizations per worker that must be respected in order for the algorithm to make sense. Otherwise, on the limit, one optimization per worker would mean almost nothing happens. This minimum number is `max(N, 15)` for the regular GOAT and `max(3N, 45)` for GOAT-EXPLORE and GOAT-REACT (see below), where N is the number of atoms. The searches using free topology are more demanding because there are many more degrees of freedom.

When the minimum number of optimizations per worker is reached, the information is printed on the bottom of the header as:

```
GOAT version                              ... explore
Minimum global steps                      ... 3
Number of base workers                    ... 4
Split workers by                          ... 4
Final number of workers                   ... 16
Number of available CPUs                  ... 1
Parameter list (worker : temperature)     ...  0 : 2903.97,  1 : 1451.98,
                                               2 :  725.99,  3 :  363.00,
                                               4 : 2903.97,  5 : 1451.98,
                                               6 :  725.99,  7 :  363.00,
                                               8 : 2903.97,  9 : 1451.98,
                                               10 :  725.99, 11 :  363.00,
                                               12 : 2903.97, 13 : 1451.98,
                                               14 :  725.99, 15 :  363.00
GradComp (mean : sigma)                   ... 1.00 : 0.50
Number of atoms                           ... 8
Number of fragments                       ... 1
Flexibility parameter                     ... 1.00
Optimizations per global step             ... 528
Optimizations per worker                  ... 66
*Reached the minimum optimizations per worker [fmax(3 * NAtoms, 45)]!
```

The global minimum found is a D_{4h} planar structure, the same as found on the literature for the Au_8 cluster using other DFT methods [54]:

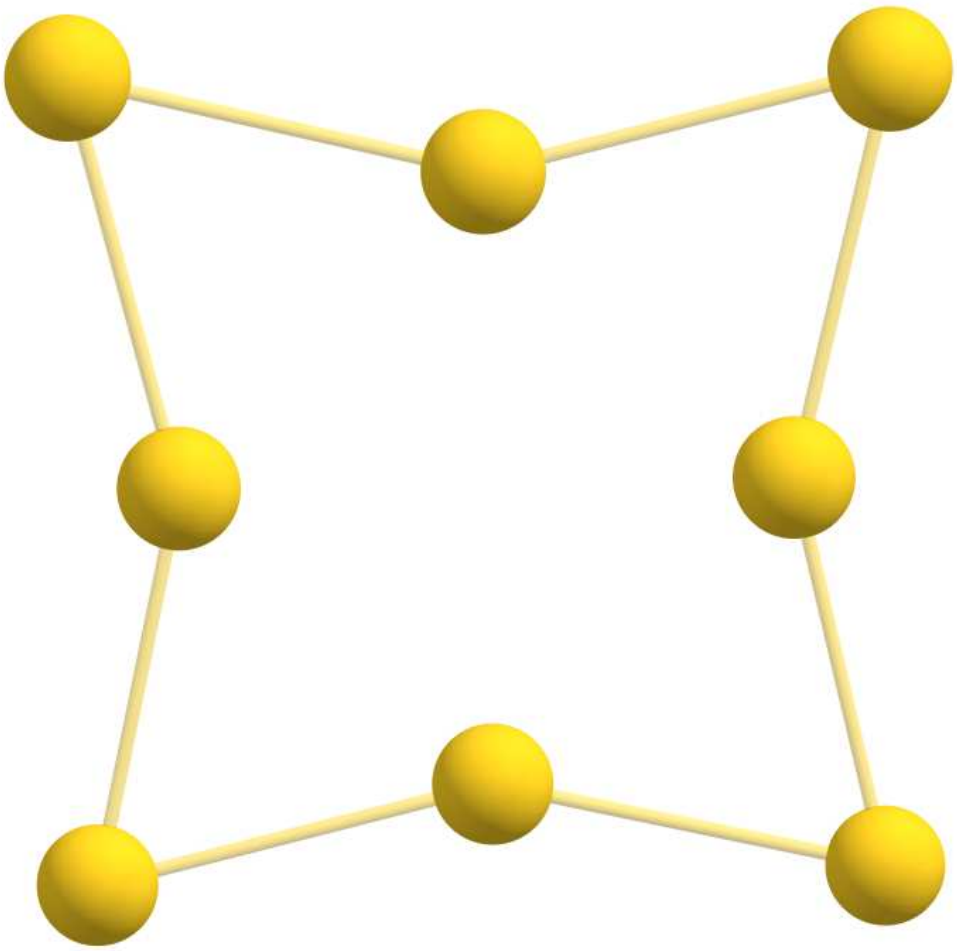


Fig. 6.27: The lowest energy conformer of the Au_8 cluster on the GFN1 PES.

6.4.7 GOAT-REACT: an algorithm for automatic reaction pathway exploration

Another variant of the GOAT algorithm was created to allow for automatic reaction exploration, which in the end is nothing more than an exploration on the collective PES of reactant and product.

Here the user can be really creative and there are many different ways to explore this algorithm, but let us start with a simple reaction: the gas phase reaction of ethylene and singlet oxygen.

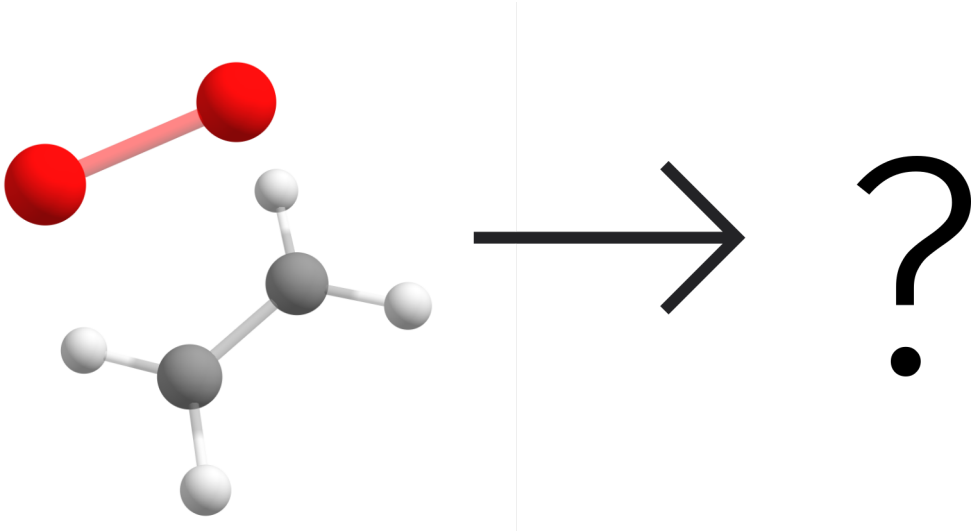


Fig. 6.28: What could be the products of the reaction between ethylene and singlet oxygen?

After running the following input:

```
!XTB GOAT-REACT
* XYZ 0 1
C         -3.26482       -0.47497        0.33191
C         -2.16518        0.24269        0.35382
H         -4.23539       -0.01923        0.27823
H         -3.23979       -1.54754        0.37118
H         -2.19035        1.31540        0.31866
H         -1.19481       -0.21295        0.41157
O         -3.42426       -0.30941        2.30779
O         -2.17088        0.05188        2.32517
*
```

the output looks more or less similar to the regular `GOAT-EXPLORE`, except for a few differences:

1. The maximum barriers GOAT is allowed to cross are higher.
2. The very initial geometry optimization is skipped by default.
3. `AUTOWALL` is set to `TRUE` which means that an ellipsoid wall potential is added using the maximum x,y,z dimensions of the molecule + 5 Angs as radii.

Another important factor is the maximum topological difference (`MAXTOPODIFF`), which is set to 8 by default, as printed on the output:

```
Global parameters
-----------------
  GOAT version                          ... react
  Max. topological diff.                ... 8
  Minimum global steps                  ... 3
  Number of base workers                ... 4
```

`MAXTOPODIFF` is a key concept here. If one simply looks for all possible topological permutations between reactants and products, even simple systems such as this could lead to an enormous number of combinations.

We defined the topological difference simply as the sum of broken bonds + formed bonds from some reference structure, which is taken from the structure obtained after the first geometry optimization inside the GOAT iterations (before any uphill step).

There will be more files printed than usual, the most important ones are:

1. `Basename.products.xyz` – contains all *reactomers* and all their conformers for the reaction. It is usually a very large file.

2. `Basename.products.topodiff2.xyz` – contains only those separated with topodiff 2 from the reference structure, and so on.
3. `Basename.products.unique.xyz` – contains a list of all topologically **unique** products, without their conformers or rotamers, only the lowest energy conformer is printed. Here the reference structure will come first and the others will be shown with their relative energy difference in kcal/mol.

Some of the products found as an example for this reaction are:

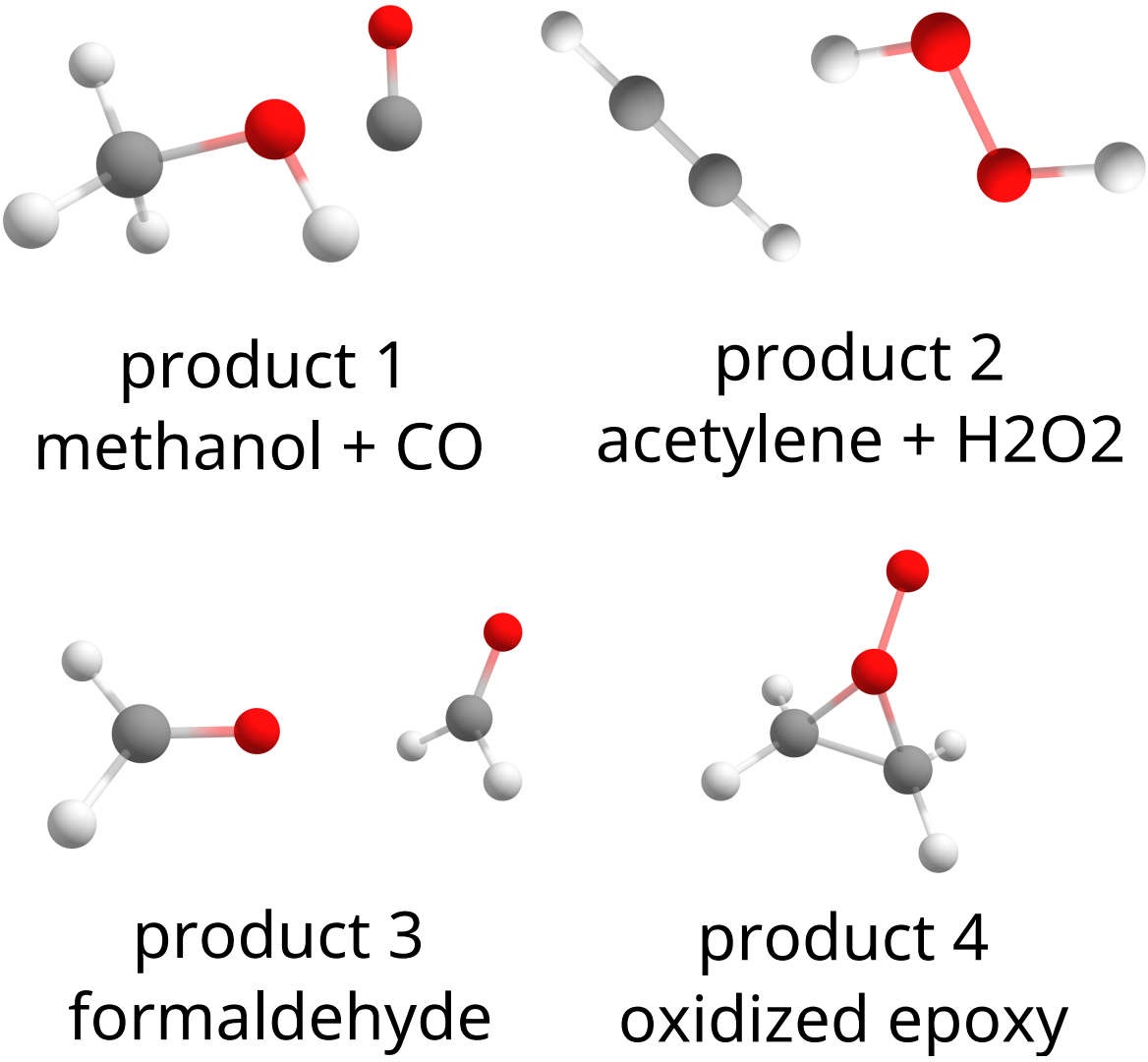


Fig. 6.29: Some products automatically found by GOAT-REACT using the input given above.

Note

Please be aware that singlet oxygen is so reactive that even the first optimization leads to a cyclization reaction and the reaction product is then taken as the reference structure.

Important

The use of force-fields like the GFN-FF is not recommended here, because it is not supposed to break bonds.

6.4.8 Some general observations

Default frozen coordinates during uphill step

During the uphill phase only, by default GOAT will freeze:

1. all bonds,
2. all angles involving two sp2 atoms within the same ring,
3. all dihedrals around a strong bond (d(B,C) < [0.9 x (sum of covalent radii)]).

"sp2 atoms" are here loosely defined only for C, N and O with less than 4,3 and 2 bonds respectively.

The first freeze is to avoid change of topology by bond breaking. The second and third are to avoid going over very high energy barriers on changing these angles, which in practice, unless under very special circumstances will never flip anyway!

These constraints are automatically lifted for `GOAT-EXPLORE` and can also be set to `FALSE` with their specific keywords.

Parallelization of GOAT

GOAT will profit from a large number of cores in a different way than most ORCA jobs, because it distributes the necessary work along different workers. It can also work multidone and distribute these workers through different nodes.

Since there is usually a regular first optimization step before starting GOAT, which will **not** profit from a large number of cores, these are limited by the flag MAXCORESOPT and set to a maximum of 32. After that, GOAT will switch back to use all cores provided. We do not recommend changing that maximum number, because it will probably only make things slower, but it can be controlled inside the %GOAT block.

Tips and extra details

Note

- GOAT will work with any method in ORCA, all you need is the gradient. That includes using DFT, QM/MM, ONIOM, broken-symmetry states, excited states etc.
- Be aware that DFT is **much** costlier than XTB. It is perfectly possible to run GOAT with R2SCAN-3C, but be prepared to use many cores or wait for a few days :D. We recommend at least `%PAL NPROCS 32 END`, to have 8 workers with 4 cores each. Hybrid DFT is even heavier, so if you want to use B3LYP,go with at least `NPROCS 64` - and don't hurry. The aim is to do a global search here, it does not come for free!
- In many cases, it might be useful to use `GFNUPHILL GFNFF` to use the GFN-FF force-field PES during the uphill steps. There, an exact potential is not really needed as the main objective is to take structures out of their current minimum and GOAT will run much faster, only using the chosen method for the actual optimizations. GFN2XTB, GFN1XTB or GFN0XTB are also valid options.
- For methods that need bond breaking, such as `GOAT-EXPLORE` or `GOAT-REACT`, GFNUPHILL GFNFF cannot be used because the GFN-FF will not allow for bond breaking. Choose GFN2XTB, GFN1XTB or GFN0XTB.
- You can always check what the workers are doing by looking into the `Basename.goat.x.x.out` files. The first number refers to the global iteration and the second to the specific worker. This is an ORCA output (with some suppressed printing to save space) that can be opened in most GUIs.
- GOAT will automatically detect fragments at the very beginning, even before the first geometry optimization. It will also respect fragments given via the geometry blocks. You can turn this off by setting `AUTOFRAG` to `FALSE` under `%GOAT`.

- Amide bond chirality is **not** frozen by default, which means the input topology you gave for amides (cis or trans) may change. If you want to freeze it, set FREEZEAMIDES to TRUE.
- Similarly double bonds outside rings can also change their topology. Choose FREEZECISTRANS TRUE in order to freeze those dihedrals.
- For certain molecules, it might be interesting to limit the coordination number of certain atoms, in that case use MAXCOORDNUMBER.
- GOAT will respect the choices from the %GEOM block for the geometries so you can use all kinds of constraints you need for other types of coordinate freeze. It can also be combined with all kinds of arbitrary wall potentials available (see Section 6.3.8).
- If you want to push only certain atoms uphill, you can give a list to UPHILLATOMS. In that case the uphill force shown in Fig. 6.24 will be applied only to the coordinates involving those atoms and the rest of the molecule will only react to that. This is useful for conformational searches on parts of a bigger system.
- By default conformers up to 12.0 kcal/mol from the global minimum are included, this can be changed by setting MAXEN.

6.4.9 Basic keyword list

Here we present a basic list of options to be given under %GOAT:

```
%GOAT

   #
   # general options
   #

   MAXITER        128  # defines an arbitrary number of max GOAT geom opt iters per worker.
   MAXOPTITER     256  # maximum number of geometry optimizations per GOAT iter.

   SKIPINITIALOPT TRUE # if you want to skip the initial optimization (default FALSE).

   RANDOMSEED     TRUE # set it to FALSE to have a deterministic GOAT run. since the
                       # geometry optimization can change due to numerical differences
                       # it might not be fully deterministic in some cases.

   READENSEMBLE "name.xyz" # an ensemble file to be read. the comment line should
                           # have the format "Energy (float)", as generated by GOAT.
                           # nothing will be done, except that the filters
                           # will be reapplied.

   AUTOWALL       TRUE  # automatically create an ellipsoid wall potential
                        # around the structure (+5 Angs)? (default is FALSE).

   TEMPLIST       3000, 2000, 750, 500 # a list of temperatures, defines the number
                                       # of basic workers, in Kelvin. do not change
                                       # unless you know what to do.

   MAXCORESOPT    32    # the max. number of cores used during the very first opt

   #
   # worker options
   #

   NWORKERS       AUTO # define the number of workers (default AUTO).
                       # AUTO for an automatic ideal assignment,
                       # or give any number multiple of 4 (number of temperatures).
   MAXITERMULT    3 # a simple keyword to multiply the number of geometry
```

(continues on next page)

(continued from previous page)

```
                # optimizations per worker. will quickly
                # increase MAXITER.
KEEPWORKERDATA FALSE # set to TRUE to keep the worker outputs
                     # (might be a lot of data!).
WORKERRANDOMSTART TRUE  # after the first cycle, each worker starts with a random
                        # structure from the previous set up to 3 kcal/mol
                        # instead of the lowest energy only.
                        # at least one starts from the lowest (default TRUE).

#
# uphill step
#

UPHILLATOMS {0:2 5 14:29} END # if given, only those atoms listed will be pushed,
                              # uphill others will just respond to it.
GFNUPHILL    GFNFF      # use GFN-FF only during the uphill steps? GFN2XTB, GFN1XTB or
                        # GFN0XTB are also valid options for the respective methods.


#
# filtering and screening
#

ALIGN          FALSE # align all final conformers with respect to the
                     # lowest energy one?
ENDIFF         0.1   # minimum energy difference needed to differentiate
                     # conformers, in kcal/mol.
MAXEN          6.0   # the maximum relative energy of a conformer to
                     # be taken, in kcal/mol. 6 kcal/mol by default.
RMSD           0.125 # minimum RMSD to differentiate conformers, in Angstroem.
ROTCONSTDIFF   0.01  # maximum difference for the  rotational constant, in %.
RMSDMETRIC     EIGENVALUE # use eigenvalues of distance matrix for RMSD?
                          # default is RMSD in general
                          # and EIGENVALUE for GOAT-EXPLORE.

#
# entropy mode
#

MAXENTROPY     FALSE  # add delta Gconf as convergence criteria (default FALSE)?
CONFTEMP       298.15 # temperature used to compute the free energy, in Kelvin.
MINDELS        0.1    # the minimum entropy difference between two iterations
                      # to signal convergence, in cal/(molK).
CONFDEGEN      2        # set an arbitrary degeneracy per conformer?
               AUTO     # find that automatically based on the RMSD.
               AUTOMAX  # same as AUTO, but take the largest value as reference
                        # for all conformers.

#
# free topology
#

FREEHETEROATOMS FALSE # free all atoms besides H and C.
FREENONHATOMS   FALSE # self explained.
FREEFRAGMENTS   FALSE # free interfragment topology, i.e., bonds between fragments
                        might be formed or broken during the search but bonds
                        within the same fragment will be kept.

# we don't recommend changing these unless you really need to!
FREEZEBONDS    FALSE # freeze bonds uphill (default TRUE)?
FREEZEANGLES   FALSE # freeze sp2 angles and dihedrals uphill (default TRUE)?
```

(continues on next page)

(continued from previous page)

```
   FREEZECISTRANS FALSE # freeze cis-trans isomers outside rings (default FALSE)?
   FREEZEAMIDES   FALSE # freeze amide cis/trans chirality (default FALSE)?

   MAXCOORDNUMBER 10, 4, 11, 6 # a list of "atom number, coordination number" that
                               # will define the maximum coordination number for
                               # those listed atoms, taken from distance-based criteria.
                               # in this case atom 10 will have a maximum of 4 and
                               # atom 11 a maximum of 6. others follow defaults.

   #
   # goat react
   #

   MAXTOPODIFF 8  # the maximum topological difference that is allow. Topodiff
                  # is simply defined based on number of broken + number of formed bonds
                  # using ORCA's regular distance based criteria [1.3 * (sum of Cov. Radii)]

END
```

6.5 Vibrational Frequencies

Vibrational frequency calculations are available through analytical differentiation of the SCF energy as well as one- or two-sided numerical differentiation of analytical gradients, i.e. for Hartree-Fock and DFT models. For methods without analytical gradient a numerically calculated gradient can be used (keyword `NumGrad`) for numerical frequencies. Please note, that this will be a very time consuming calculation.

The use of vibrational frequency calculations is fairly simple:

```
# any Hartree-Fock or DFT model can be used here
! BP def2-TZVP

# Tight SCF convergence is advisable to minimize the numerical
# noise in the frequencies.
! TightSCF

# perform a geometry optimization first
! Opt

# Run an analytical or numerical frequency calculation afterwards
! AnFreq   # or just ``! Freq''
# numerical:
! NumFreq

# details of the numerical frequency calculation
%freq  CentralDiff true   # use central-differences (this is the default)
       Increment   0.005  # increment in bohr for the
                          # differentiation (default 0.005)
       end

! bohrs
* xyz 0 1
 O          -1.396288    -0.075107     0.052125
 O           1.396289    -0.016261    -0.089970
 H          -1.775703     1.309756    -1.111179
 H           1.775687     0.140443     1.711854
*
```

At the end of the frequency job you get an output like this:

```
-----------------------
VIBRATIONAL FREQUENCIES
-----------------------

   0:         0.00 cm**-1
   1:         0.00 cm**-1
   2:         0.00 cm**-1
   3:         0.00 cm**-1
   4:         0.00 cm**-1
   5:         0.00 cm**-1
   6:       311.78 cm**-1
   7:       887.65 cm**-1
   8:      1225.38 cm**-1
   9:      1394.81 cm**-1
  10:      3624.88 cm**-1
  11:      3635.73 cm**-1
```

This output consists of the calculated vibrational frequencies, the vibrational modes and the thermochemical properties at 298.15 K. In the example above there are six frequencies which are identically zero. These frequencies correspond to the rotations and translations of the molecule. They have been projected out of the Hessian before the calculation of the frequencies and thus, the zero values do not tell you anything about the quality of the Hessian that has been diagonalized. The projection can be turned off by `PROJECTTR FALSE` under %FREQ, so that the frequencies of the translations and rotations can deviate from zero and the deviations represent a metric of the numerical error of the Hessian calculation. This is done automatically when there is e.g. an external electric field that makes the exact translational and/or rotational modes have non-zero frequencies (see section *Adding finite electric field*). However, in normal cases where the molecule is expected to obey both translational and rotational invariance, it is **strongly discouraged** to turn off `PROJECTTR` when calculating thermochemical quantities (especially entropies and Gibbs free energies). This is because when the frequencies of translational and rotational modes exceed `CutOffFreq` (which is 1 cm^{-1} by default), their contributions to the partition function will be calculated using the formulas for vibrations. As a result, the calculated entropy is inaccurate (due to treating translations and rotations as vibrations), is sensitive to numerical noise, and in particular exhibits a finite jump when the (theoretically zero) frequencies of the translational and rotational modes cross `CutOffFreq`. Therefore, the only case where the user needs to turn off `PROJECTTR` manually is when the exact Hessian is expected to have zero translational and rotational frequencies, and one wants to check how much the translational and rotational eigenvalues of the actually computed Hessian deviate from zero. The thermochemical quantities from such a calculation are less reliable and should not be used; even if they differ considerably from the results with `PROJECTTR TRUE`, this does not necessarily mean that the latter are unreliable.

Without `PROJECTTR FALSE`, the reliability of the calculated frequencies has to be judged by comparison of calculations with different convergence criteria, increments, integration grids etc. The numerical error in the frequencies may reach 50 cm^{-1} but should be considerably smaller in most cases. Significant negative frequencies indicate saddle points of the energy hypersurface and prove that the optimization has not resulted in an energy minimum.

OBS: By default, the Hessian is made translation invariant by applying the "acoustic sum rule" ([787]), which reduces the effect of noise from numerical integration coming from DFT or COSX, except for the Partial and Hybrid Hessians where it does not make sense. It can be set to false by using TRANSINVAR FALSE under %FREQ.

6.5.1 Mass dependencies

Of course the calculated frequencies depend on the masses used for each atom. While this can be influenced later through the `orca_vib` routine (see Section *Isotope Shifts* for more detail) and individually for each atom in the geometry input, one might prefer using a set of precise atomic masses rather than the set of atomic weights (which are set as default). This can be achieved through the `!Mass2016` keyword, which triggers Orca to use those atomic masses representing either the most abundant isotope or the most stable isotope (if all isotopes are unstable) of a certain element (e.g. the mass of ^{35}Cl for chlorine or the mass of ^{98}Tc).

Note

The calculation of numerical frequencies puts rather high demands on both computer time and accuracy. In order to get reliable frequencies make sure that:

- Your SCF is tightly converged. A convergence accuracy of at least 10^{-7} Eh in the total energy and 10^{-6} in the density is desirable.
- Grids of at least `DEFGRID2` (default) are used.
- The use of two-sided (i.e. central) differences increases the computation time by a factor of two but gives more accurate and reliable results.
- Small auxiliary basis sets like DGauss/J or DeMon/J may not result in fully converged frequencies (up to 40 cm^{-1} difference compared to frequencies calculated without RI). The def2/J universal auxiliary basis sets of Weigend that are now the default in ORCA (or the SARC/J for scalar relativistic calculations) are thought to give sufficiently reliable results.
- Possibly, the convergence criteria of the geometry optimization need to be tightened in order to get fully converged results.
- If you can afford it, decrease the numerical increment to 0.001 Bohr or so. This puts even higher demands on the convergence characteristics of the SCF calculation but should also give more accurate numerical second derivatives. If the increment is too small or too high inaccurate results are expected.

The calculation of analytical frequencies is memory consuming. To control memory consumption the `%maxcore` parameter must be set. For example `%maxcore 8192` - use 8 Gb of memory per processor for the calculation. The user should provide the value according to the computer available memory. The batching based on `%maxcore` parameter will be introduced automatically to overcome probable memory shortage.

Numerical frequency calculations are restartable (but analytical frequency calculations are not). If the numerical frequencies job died for one reason or another you can simply continue from where it stopped as in the following example:

```
! STO-3G NumFreq
%freq  Restart true  # restart an old calculation
                     # this requires .res.* files to be present
       end
* int 0 1
    C  0 0 0 0.0000   0 0
    C  1 0 0 1.2160   0 0
    H  1 2 0 1.083  180 0
    H  2 1 3 1.083  180 0
*
```

Note

- You must not change the level of theory, basis set or any other detail of the calculation. Any change will produce an inconsistent, essentially meaningless Hessian.
- The geometry at which the Hessian is calculated must be identical. If you followed a geometry optimization by a frequency run then you must restart the numerical frequency calculation from the optimized geometry.
- Numerical frequencies can be performed in multi-process mode. Please see section *Calling the Program with Multiple Processes* ("Hints on the use of parallel ORCA") for more information.
- The restart of Numerical frequencies will take off from the result files produced during the preceding run (`BaseName.res.%5d.Type`, whith `Type` being `Dipoles`, `Gradients` - and if requested `Ramans` or `Nacmes`). Please make sure that all these local result files get copied to your compute directory. If restart is set and no local files to be found, ORCA will restart from scratch. If ORCA finds a Hessian file on disk, it will only repeat the subsequent analysis.

- The Hessian can be transformed to redundant internal coordinates. More information can be found in section *Printing Hessian in Internal Coordinates*.

6.6 Excited States Calculations

A plethora of methods to compute excited states exists in ORCA . In the following section, we illustrate typical single-reference approaches. Multi-reference methods, such as NEVPT2 or MRCI, are described elsewhere in the manual.

6.6.1 Excited States with RPA, CIS, CIS(D), ROCIS and TD-DFT

ORCA features a module to perform TD-DFT, single-excitation CI (CIS) and RPA. The module works with either closed-shell (RHF or RKS) or unrestricted (UHF or UKS) reference wavefunctions. For DFT models the module automatically chooses TD-DFT and for HF wavefunctions the CIS model. If the RI approximation is used in the SCF part it will also be used in the excited states calculation. A detailed documentation is provided in section *Excited States via RPA, CIS, TD-DFT and SF-TDA*.

General Use

In its simplest form it is only necessary to provide the number of roots sought:

```
! B3LYP DEF2-SVP

%TDDFT NROOTS    10
       TRIPLETS  TRUE
END

* int 0 1
    C  0 0 0 0.00 0.0     0.00
    O  1 0 0 1.20 0.0     0.00
    H  1 2 0 1.08 120     0.00
    H  1 2 3 1.08 120   180.00
*
```

The triplets parameter is only valid for closed-shell references. If chosen as true the program will also determine the triplet excitation energies in addition to the singlets. We will discuss many more options in the following sections.

Spin-Flip

The collinear spin-flip version of CIS/TDA (always starting from an open-shell reference!) can be invoked in a similar manner, using:

```
%tddft
  nroots 5
  sf     true
end
```

Please check Sec. *Collinear Spin-Flip TDA (SF-TD-DFT)* for more details on how to use it, and how to understand its results.

Population analysis

If you want to print excited-state charges and bond orders, you can use UPOP TRUE under %TDDFT to get the analysis from the unrelaxed density and !ENGRAD if you want to use the relaxed density. Multiple states can be indicated by the IROOTLIST and TROOTLIST keywords. For more details please check Sec. *Population Analysis of Excited States*.

Use of TD-DFT for the Calculation of X-ray Absorption Spectra

In principle X-ray absorption spectra are "normal" absorption spectra that are just taken in a special high-energy wavelength range. Due to the high energy of the radiation employed (several thousand eV), core-electrons rather than valence electrons are excited. This has two consequences: a) the method becomes element specific because the core-level energies divide rather cleanly into regions that are specific for a given element. b) the wavelength of the radiation is so short that higher-order terms in the expansion of the light-matter interaction become important. Most noticeably, quadrupole intensity becomes important.

X-ray absorption spectra can be generally divided into three regions: a) the pre-edge that corresponds to transitions of core electrons into low lying virtual orbitals that lead to bound states. b) the rising edge that corresponds to excitations to high-lying states that are barely bound, and c) the extended X-ray absorption fine structure region (EXAFS) that corresponds to electrons being ejected from the absorber atom and scattered at neighbouring atoms.

With the simple TD-DFT calculations described here, one focuses the attention on the pre-edge region. Neither the rising edge nor the EXAFS region are reasonably described with standard electronic structure methods and no comparison should be attempted. In addition, these calculations are restricted to K-edges as the calculation of L-edges is much more laborious and requires a detailed treatment of the core hole spin orbit coupling.

It is clearly hopeless to try to calculate enough states to cover all transitions from the valence to the pre-edge region. Hence, instead one hand-selects the appropriate donor core orbitals and only allows excitations out of these orbitals into the entire virtual space. This approximation has been shown to be justified.[200] One should distinguish two situations: First, the core orbital in question may be well isolated and unambiguously defined. This is usually the case for metal 1s orbitals if there is only one metal of the given type in the molecule. Secondly, there may be several atoms of the same kind in the molecule and their core orbitals form the appropriate symmetry adapted linear combinations dictated by group theory. In this latter case special treatment is necessary: The sudden approximation dictates that the excitations occurs from a local core orbital. In previous versions of the program you had to manually localize the core holes. In the present version there is an automatic procedure that is described below.

A typical example is $TiCl_4$. If we want to calculate the titanium K-edge, the following input is appropriate:

```
! BP86 ZORA ZORA-def2-TZVP(-f) SARC/J TightSCF

%tddft   OrbWin[0] = 0,0,-1,-1
         NRoots    25
         DoHigherMoments  true
         DoFullSemiclassical true
         end

* int 0 1
Ti 0 0 0  0 0 0
Cl 1 2 3 2.15 0 0
Cl 1 2 3 2.15 109.4712 0
Cl 1 2 3 2.15 109.4712 120
Cl 1 2 3 2.15 109.4712 240
*
```

Note

- The absolute transition energies from such calculations are off by a few hundred electron volts due to the shortcomings of DFT. The shift is constant and very systematic for a given element. Hence, calibration

is possible and has been done for a number of edges already. Calibration depends on the basis set!

- Electric quadrupole contributions and magnetic dipole contributions have been invoked with `DoHigherMoments true` (check section *One Photon Spectroscopy* for more information), which is essential for metal edges. For ligand edges, the contributions are much smaller.
- `OrbWin` is used to select the single donor orbital (in this case the metal 1s). The LUMO (45) and last orbital in the set (174) are selected automatically if "-1" is given. This is different from previous program versions where the numbers had to be given manually.

The output contains standard TD-DFT output but also:

```
-------------------------------------------------------------------------------------------------
-----------------------------------
                      ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM
-------------------------------------------------------------------------------------------------
-----------------------------------
     Transition      Energy     Energy  Wavelength fosc(D2)  fosc(M2)  fosc(Q2)  ␣
↪fosc(D2+M2+Q2)     D2/TOT    M2/TOT    Q2/TOT
                      (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-------------------------------------------------------------------------------------------------
-----------------------------------
```

This section contains the relevant output since it combines electric dipole, electric quadrupole and magnetic dipole transition intensities into the final spectrum. Importantly, there is a gauge issue with the quadrupole intensity: the results depend on the where the origin is placed. We have proposed a minimization procedure that guarantees the fastest possible convergence of the multipole expansion.[201]

The spectra are plotted by calling

```
orca_mapspc  MyOutput.out ABSQ -eV -x04890 -x14915 -w1.3
```

Starting from ORCA version 4.1 one may obtain origin independent transition moments formulations which can be combined with the multipole moments up to 2nd order to regenerate the electric dipole, electric quadrupole and magnetic dipole contributions in either length or the velocity representations. This requires in addition to the electric dipole (D), electric quadrupole (Q) and magnetic dipole (m) intensities the corresponding electric dipole - magnetic quadrupole (DM) and the electric dipole - electric octupole (DO) intensities.[810][95]. See also section *General Use*.

These spectra are requested by (check section *One Photon Spectroscopy* for more information)

```
DoHigherMoments true
DecomposeFoscLength true
DecomposeFoscVelocity true
```

Resulting in:

```
-------------------------------------------------------------------------------------------------
-----------------------------------
                      ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM (Origin Independent, Length)
-------------------------------------------------------------------------------------------------
-----------------------------------
     Transition      Energy     Energy  Wavelength fosc(D2)  fosc(M2)  fosc(Q2)␣
↪fosc(D2+M2+Q2+DM+DO) D2/TOT    M2/TOT    Q2/TOT
                      (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-------------------------------------------------------------------------------------------------
-----------------------------------
...
```

(continues on next page)

(continued from previous page)

```
-------------------------------------------------------------------------------------------------
↪---------------------------------
                      ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM (Origin Independent, Velocity)
-------------------------------------------------------------------------------------------------
↪---------------------------------
     Transition      Energy     Energy  Wavelength fosc(P2)  fosc(M2)  fosc(Q2)␣
↪fosc(P2+M2+Q2+PM+PO) P2/TOT    M2/TOT    Q2/TOT
                       (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-------------------------------------------------------------------------------------------------
↪---------------------------------
...
```

The Origin Independent transition moments spectra are plotted by calling:

```
orca_mapspc  MyOutput.out ABSOI/ABSVOI -eV -x04890 -x14915 -w1.3
```

Although the multipole moments up to 2nd order:

- Only approximate origin independence is achieved by using the length approximation for distances from the excited atom up to about 5 Angstrom.
- Can form negative intensities which can be partly cured by using larger basis sets.

Starting from ORCA version 6.0 the full semi-classical ligth-matter interaction[95][398][528] can be computed by including the keyword:

```
DoFullSemiclassical true
```

Resulting in:

```
-------------------------------------------------------------------------------
                      ABSORPTION SPECTRUM VIA FULL SEMI-CLASSICAL FORMULATION
-------------------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength   fosc(FFMIO)
                       (eV)      (cm-1)    (nm)
-------------------------------------------------------------------------------
```

The full-semiclassical transition moments:

- Behave like the multipole expansion in the velocity representation.
- They are by definition origin independent they do not suffer from artificial negative values like the multipole moments beyond 1st order.

Now, let us turn to the Cl K-edge. Looking at the output of the first calculation, we have:

```
----------------
ORBITAL ENERGIES
----------------

  NO   OCC          E(Eh)            E(eV)
   0   2.0000    -180.132624     -4901.6579
   1   2.0000    -101.520058     -2762.5012
   2   2.0000    -101.520052     -2762.5010
   3   2.0000    -101.520048     -2762.5010
   4   2.0000    -101.520048     -2762.5010
   5   2.0000     -19.823233      -539.4176
   6   2.0000     -16.411730      -446.5859
   7   2.0000     -16.411729      -446.5858
   8   2.0000     -16.411729      -446.5858
   9   2.0000      -9.280963      -252.5478
  10   2.0000      -9.280957      -252.5477
```

(continues on next page)

(continued from previous page)

```
  11   2.0000      -9.280953      -252.5476
  12   2.0000      -9.280953      -252.5476
  13   2.0000      -7.037815      -191.5087
  14   2.0000      -7.037805      -191.5084
  15   2.0000      -7.037791      -191.5080
  16   2.0000      -7.037791      -191.5080
  17   2.0000      -7.035288      -191.4399
  18   2.0000      -7.035287      -191.4399
....
```

And looking at the energy range or the orbital composition, we find that orbitals 1 through 4 are Cl 1s-orbitals. They all have the same energy since they are essentially non-interacting. Hence, we can localize them without invalidating the calculation. To this end, you can invoke the automatic localization for XAS which modifies the input to:

```
! BP86 ZORA ZORA-def2-TZVP(-f) SARC/J TightSCF

%tddft   XASLoc[0] = 1,4
         OrbWin[0] = 1,1,-1,-1
         NRoots    25
         DoHigherMoments  true
         DoFullSemiclassical true
         end

* int 0 1
Ti 0 0 0  0 0 0
Cl 1 2 3 2.15 0 0
Cl 1 2 3 2.15 109.4712 0
Cl 1 2 3 2.15 109.4712 120
Cl 1 2 3 2.15 109.4712 240
*
```

- This localizes the orbitals 1 through 4 of operator 0 (the closed-shell) and then allows excitations (arbitrarily) from core hole 1 only. You could choose any of the three other localized 1s orbitals instead without changing the result. You could even do all four core holes simultaneously (they produce identical spectra) in which case you have the entire ligand K-edge intensity and not just the one normalized to a single chlorine (this would be achieved with `OrbWin[0] = 1,4,-1,-1`).
- If you have a spin unrestricted calculation, you need to give the same `XASLoc` and `OrbWin` information for the spin-down orbitals as well.

Quite nice results have been obtained for a number of systems in this way.[712]

Excited State Geometry Optimization

For RPA, CIS, TDA and TD-DFT the program can calculate analytic gradients. With the help of the `IRoot` keyword, a given state can be selected for geometry optimization. Note however, that if two states cross during the optimization it may fail to converge or fail to converge to the desired excited state (see section *Root Following Scheme for Difficult Cases* below)! If you want to follow a triplet state instead of the singlet, please set IROOTMULT to TRIPLET.

```
! HF DEF2-SVP Opt

%CIS   NRoots    1
       IRoot     1
       end

* int 0 1
    C  0 0 0 0.00 0.0     0.00
    O  1 0 0 1.20 0.0     0.00
```

(continues on next page)

(continued from previous page)

```
    H  1 2 0 1.08 120     0.00
    H  1 2 3 1.08 120   180.00
*
```

Note that this example converges to a saddle point as can be verified through a numerical frequency calculation (which is also possible with the methods mentioned above). The excited state relaxed density matrix is available from such gradient runs (`MyJob.cisp` when using the `KeepDens` keyword) and can be used for various types of analysis. Note that the frozen core option is available starting from version 2.8.0.

Root Following Scheme for Difficult Cases

In case there is a root flipping after a step during the geometry optimization, it might be impossible to converge an excited state geometry using the regular methods. To help in those cases, the flag FOLLOWIROOT might be set to TRUE. Then, excited state wavefunction will be analyzed and compared with the reference one (more below), and the IROOT will be automatically adjusted to keep homing the target state.

One example of such a calculation is:

```
! wB97X OPT
%TDDFT
  NROOTS         5
  IROOT          3
  FOLLOWIROOT    TRUE
END
* xyz 0 1
N 0.0  0.0  0.0
H 0.0  0.0  1.0
H 0.0 -0.9  0.5
H 0.0  0.9  0.5
*
```

This will ask for an optimization of the third excited state of ammonia. At some point, there is a state crossing and what was state 3 now becomes state 2. The algorithm will recognize this and automatically change the IROOT flag, to keep following the same state. FOLLOWIROOT also works with spin-adapted triplets and spin-flip states.

In cases where you want to keep the comparison only with the density from **the very first** computed excited state, e.g. the one you get on the first cycle of a geometry optimization, you can use FIRKEEPFIRSTREF, as in:

```
%TDDFT
  NROOTS            5
  IROOT             3
  FOLLOWIROOT       TRUE
  FIRKEEPFIRSTREF   TRUE # default false
END
```

Criteria to Follow IROOTs - starting from ORCA6

Starting from ORCA6 we have a much more robust algorithm to follow these excited states, inspired by some of the recent literature [819] [135]. The algorithm now works as follows, after each excited state calculation using CIS/TDDFT:

1. Given a reference state, take all states within an energy difference of up to 1 eV to it. We don't want to check states that are too far apart in energy. Controlled by `%TDDFT FIRENTHRESH 1.0 END`, number in eV.
2. Now take all states with a difference of $\hat{S}^2$ not larger than 0.5. We don't want to compare singlets to triplets. Controlled by `%TDDFT FIRS2THRESH 0.5 END`.
3. Calculate the overlap between the transition densities of all states with the reference - this is the core part.

4. In case there is ambiguity - that is if two states have overlaps differing by only 0.05 - take the one with the closer transition dipole angle. Controlled by `%TDDFT FIRSTHRESH 0.05 END`.
5. Update the IROOT to the state that went best on all these tests.

> **Note**
>
> These `FIR` keywords are specific to **F**ollow **IR**oot.

Now by default we might also update the reference state from time to time in case the separation of states is very clear. The way it work is:

1. If the best overlap is larger than `FIRMINOVERLAP`, which is 0.5 by default and `FIRDYNOVERLAP` is `FALSE`, we will assume that the overlap is good enough and we will always update the reference. However, the default is `FIRDYNOVERLAP TRUE`, which means also have a second check for robustness.
2. If `FIRDYNOVERLAP TRUE` and the best overlap is larger than 0.5 (or `FIRMINOVERLAP`), we will check for the `ratio` between the best and the second best states. If this `ratio` is between 0.3 and 0.6 (controlled by `%TDDFT FIRDYNOVERRATIO 0.3,0.6 END`), it means that there is a clear separation between the best and the second best and the reference can be updated safely. If the ratio is too close to 1, both states are too similar and it would be dangerous to update the reference state. If it is too close to zero, they are easy to distinguish and we don't need to update the reference yet [819].

> **Important**
>
> It is important to stress that this will not necessarily solve all problems (root flipping can be particularly bad if the system is highly symmetric), for the excited states may change too much during the optimization. If that happens, it is advisable to restart the calculation after some steps and check which IROOT you still want. This can also be used when calculating numerical gradients and Hessians, in case you suspect of root flipping after the displacements.

> **Important**
>
> This algorithm is completely general and should work for any excited state method, as long as there are transition densities. We will include more methods in the future when possible.

Doubles Correction

For CIS (and also for perturbatively corrected time-dependent double-hybrid functionals) the program can calculate a doubles correction to the singles-only excited states. The theory is due to Head-Gordon and co-workers [371].

```
%cis dcorr n  # n=1,2,3,4 are four different algorithms that
              # lead to (essentially) the same result but differ
              # in the way the rate-limiting steps are handled
```

Spin-component scaling versions of CIS(D) can be evoked in the %cis block by setting DOSCS TRUE and the four scaling parameters, as defined by Head-Gordon and co-workers [717], in the following order: same-spin indirect term (CTss), opposite-spin indirect term(CTos), same-spin direct term(CUss), and opposite-spin direct term(CUos). Note that this implementation only works for the version with the parameter $\lambda = 1$ as defined in Ref. [717]. The example below shows how to apply the SCS-CIS(D) version with $\lambda = 1$ whose usage has been advocated in Ref. [311]. The user is able to specify other scaling parameters.

```
%cis
dcorr     # n=1,2,3,4
doscs     true                    # set SCS-CIS(D) to true (default: false)
```

(continues on next page)

(continued from previous page)

```
scspar    0.333, 1.2, 0.43, 1.24  #SCS-CIS(D) scaling parameters in this order
                                  CTss, CTos, CUss, CUos
end
```

Note the use of commas to separate the parameters. These parameters do not communicate with the SCS/SOS parameters set for ground-state SCS/SOS-MP2 in the %mp2 block.

Note

- CIS(D) is often a quite big improvement over CIS.
- The cost of the (D) correction is O(N^5) and therefore comparable to RI-MP2. Since there are quite a few things more to be done for (D) compared to RI-MP2, expect the calculations to take longer. In the most elementary implementation the cost is about two times the time for RI-MP2 for each root.
- The (D) correction is compatible with the philosophy of the double-hybrid density functionals and should be used if these functionals are combined with TD-DFT. The program takes this as the default but will not enforce it. The (D) correction can be used both in a TD-DFT and TDA-DFT context.
- In our implementation it is only implemented together with the RI approximation and therefore you need to supply an appropriate ("/C") fitting basis.
- The program will automatically put the RI-MP2 module into operation together with the (D) correction. This will result in the necessary integrals becoming available to the CIS module.
- Singlet-triplet excitations can be calculated by setting TRIPLETS TRUE in the %cis or %tddft blocks, respectively. The implementation has been tested for double hybrids in Ref. [145].
- For spin-adapted triplets (TRIPLETS TRUE), the only option available currently is DCORR 1.
- Spin-component and spin-opposite scaling techniques for double-hybrids within the TD- and TDA-DFT frameworks, as defined by Schwabe and Goerigk [771], can be evoked in the same way in the %tddft block as described for SCS-CIS(D) above. While user-defined parameters can be entered in such a way, a series of new functionals are available through normal keywords, which use the herein presented SCS/SOS-CIS(D) implementation. [147] See Sec. *Choice of Functional* for a list of those functionals.

Spin-orbit coupling

It also possible to include spin-orbit coupling between singlets and triplets calculated from TD-DFT by using quasi-degenerate perturbation theory (please refer to the relevant publication [198]), similarly to what is done in ROCIS. In order to do that, the flag DOSOC must be set to TRUE. The reduced matrix elements are printed and the new transition dipoles between all SOC coupled states are also printed after the regular ones. This option is currently still not compatible with double hybrids, but works for all other cases including CPCM. All the options regarding the SOC integrals can be altered in the %rel block, as usual.

```
%CIS DOSOC  TRUE END
```

Please have in mind that, as it is, you can only calculate the SOC between excited singlets and the spin-adapted triplets. There is no SOC starting from a UHF/UKS wavefunction. If you want more information printed such as the full SOC matrix or triplet-triplet couplings, please set a higher PRINTLEVEL.

SOC and ECPs

ORCA currently does not have SOC integrals for ECPs, and these are by default ignored in the SOC module. If you try to use ORCA together with ECPs, an abort message will be printed. If you absolutely need to use ECPs, for instance for embedded potentials, please use:

```
%TDDFT FORCEECP TRUE END
```

OBS.: Do not use ECPs in atoms where SOC might be important. In that case, always use all-electron basis functions or the results will not make sense.

Geometry Optimization of SOC States

If you want to compute geometries for the SOC states, just choose SOCGRAD TRUE and a given IROOT. The weigthed "unrelaxed" gradient will then be calculated after selecting the CIS/TD-DFT states with contribution larger than 0.01%. Each gradient will be calculated separately and, after that, the final SOC gradient will be computed as a weighted sum. Setting IROOT 0 in this case corresponds to ask for the SOC ground state, which is NOT necessarily equal to the ground state from HF/DFT.

Transient spectra

If one wants to compute transient spectra, or transition dipoles starting from a given excited state, the option DOTRANS must be set to TRUE and an IROOT should be given for the initial state (the default is 1). If `DOTRANS ALL` is requested instead, the transition dipoles between all states are computed. The transient transition dipoles will then be printed after the normal spectra. This option is currently only available for CIS/TDA and is done usng the expectation value formalism, as the other transition dipole moments in ORCA.

```
%cis
  DOTRANS  TRUE
           #or
  DOTRANS  ALL
end
```

Non-adiabatic coupling matrix elements

The CIS module can compute the non-adiabatic coupling matrix elements (NACME) between ground and an excited state given by an IROOT, $\langle \Psi_{GS} | \frac{\partial}{\partial R_x} | \Psi_{IROOT} \rangle$ [782]. These can also include LR-CPCM effects if !CPCM(solvent) is chosen in the main input, ZORA effects and will make use of RIJ and COSX, if they are chosen for the SCF. The usage is simple, e.g.:

```
!PBE0 DEF2-SVP TIGHTSCF
%TDDFT  NROOTS  5
        IROOT   2
        NACME   TRUE
END
* xyz 0 1
O         0.000000000      0.000000000       0.611403292
C         0.000000000      0.000000000      -0.613232096
H         0.931880792      0.000000000      -1.200880848
H        -0.931880792      0.000000000      -1.200880848
*
```

By choosing NACME TRUE under %TDDFT, a regular gradient calculation will be done, and the NACMEs will be computed together with it. After the usual gradient output, the NACMEs will be printed as:

```
---------------------------------
CARTESIAN NON-ADIABATIC COUPLINGS
       <GS|d/dx|ES>
---------------------------------

1   O   :   -0.161958900   -0.000000135    0.000001029
2   C   :   -0.088213027   -0.000000024    0.000000167
3   H   :    0.226241398    0.000000001   -0.109102154
4   H   :    0.226241406   -0.000000000    0.109102161

Difference to translation invariance:
:    0.2023108777   -0.0000001585    0.0000012017

Norm of the NACs                   ...    0.4002363416
RMS NACs                           ...    0.1155382798
MAX NAC                            ...    0.2262414063
```

NACMEs with built-in electron-translation factor

As you can see, the calculation above does not have full translation invariance! That is a feature of NACs calculated from CI wavefunctions, due to the Born-Oppenheimer approximation. It can be somehow fixed by including the so-called "electron-translation factors" (ETFs) [252], and those are added with ETF TRUE under %TDDFT. By now using the input:

```
!PBE0 DEF2-SVP
%TDDFT  NROOTS  5
     IROOT   2
     NACME   TRUE
     ETF     TRUE
END
* xyz 0 1
O        0.000000000      0.000000000      0.611403292
C        0.000000000      0.000000000     -0.613232096
H        0.931880792      0.000000000     -1.200880848
H       -0.931880792      0.000000000     -1.200880848
*
```

one gets the following output:

```
---------------------------------
CARTESIAN NON-ADIABATIC COUPLINGS
          <GS|d/dx|ES>
        with built-in ETFs
---------------------------------

1   O   :   -0.071334028   -0.000001941    0.000003727
2   C   :   -0.362514525   -0.000000130   -0.000000776
3   H   :    0.217014763    0.000000003   -0.128968922
4   H   :    0.217014813   -0.000000002    0.128968939

Difference to translation invariance:
:    0.0001810232   -0.0000020693    0.0000029689

Norm of the NACs                   ...    0.5137724505
RMS NACs                           ...    0.1483133313
MAX NAC                            ...    0.3625145251
```

where the residual translation variance is due to the DFT and COSX grids only.

> **Warning**
>
> These are the recommended NACs to be used with any kind of dynamics or conical intersection optimization, otherwise moving the center of mass of you system would already change the couplings!

Numerical non-adiabatic coupling matrix elements

The numerical non-adiabatic coupling matrix elements between ground and excited states from CIS/TD-DFT can be calculated in a numerical fashion, by setting the NumNACME flag on the main input line:

```
! NumNACME
```

ORCA will then calculate both the NACMEs and the numerical gradient for a given IROOT at the same cost. Please be careful with the SCF options and GRID sizes since there are displacements involved, for more information check *Numerical Gradients*. All options regarding step size and so on can be changed from %NUMGRAD.

These are current implemented in both RHF/RKS and UHF/UKS, but only for CIS/TDA and RPA/TD-DFT, no multireference methods yet. For the latter case, the overlap of the $|X - Y\rangle$ vector is used [517].

Restricted Open-shell CIS

In addition to the CIS/TD-DFT description of excited states, ORCA features the `orca_rocis` module to perform configuration interaction with single excitations calculations using a restricted open-shell Hartree–Fock (ROHF) reference. It can be used to calculate excitation energies, absorption intensities and CD UHFintensities. In general, ROCIS calculations work on restricted open-shell HF reference functions but in this implementation it is possible to enter the calculations with RHF (only for closed-shell molecules) or UHF reference functions as well. If the calculation starts with an UHF/UKS calculation, it will automatically produce the quasi-restricted orbitals which will then be used for the subsequent ROCIS calculations. Note that if the reference function is a RHF/RKS function the method produces the CIS results. The module is invoked by providing the number of roots sought in the `%rocis` block of the input file:

```
! SVP TightSCF

%rocis NRoots 2
MaxDim 5 #Davidson expansion space = MaxDim * NRoots
       end

* xyz -2 2
Cu  0.00    0.00    0.00
Cl  2.25    0.00    0.00
Cl -2.25    0.00    0.00
Cl  0.00    2.25    0.00
Cl  0.00   -2.25    0.00
*
```

In this example the `MaxDim` parameter is given in addition to the number of roots to be calculated. It controls the maximum dimension of the expansion space in the Davidson procedure that is used to solve the CI problem.

The use of ROCIS is explained in greater detail in section *Excited States via ROCIS and DFT/ROCIS*.

Starting from ORCA 6.0, the General-Spin ROCIS (`GS-ROCIS`) implementation is available. This new implementation can handle arbitrary CSFs as references. For this, one would use the `CSF-ROHF` method to obtain the reference wavefunction for which ROCIS will be performed. The `GS-ROCIS` calculation can be invoked as follows:

```
%scf
  HFTyp ROHF
  ROHF_CASE USER_CSF or AF_CSF
  etc.
```

(continues on next page)

(continued from previous page)

```
end

%rocis
  DoGenROCIS true  # Turns the General-Spin ROCIS procedure on
  ReferenceMult 1  # The reference wavefunction multiplicity (it needs to agree with the ROHF␣
↪solution)
  etc.
end
```

Currently, there is no DFT/ROCIS implemented for the General-Spin procedure. Spin-Orbit coupling is also not available in the present version.

6.6.2 Excited States for Open-Shell Molecules with CASSCF Linear Response (MC-RPA)

ORCA has the possibility to calculate excitation energies, oscillator and rotatory strengths for CASSCF wave functions within the response theory (MC-RPA) formalism.[380, 417, 901] The main scope of MC-RPA is to simiulate UV/Vis and ECD absorption spectra of open-shell molecules like transition metal complexes and organic radicals. MC-RPA absorption spectra are usually more accurate than those obtained from the state-averaged CASSCF ansatz as orbital relaxation effects for excited states are taken into account. The computational costs are ususally larger than those of SA-CASSCF and should be comparable to a TD-DFT calculation for feasible active space sizes.

General Use

MC-RPA needs a converged state-specific CASSCF calculation of the electronic ground state. The only necessary information that the user has to provide is the desired number of excited states (roots). All other keywords are just needed to control the Davidson algorithm or post process the results. A minimal input for calculating the four lowest singlet excited states of ethylene could like the following:

```
#
# CASSCF + MCRPA for C2H4
#
! DEF2-SVP DEF2-TZVP/C VeryTightSCF

%casscf
 nel          2
 norb         2
 mult         1
 nroots       1
 gtol 1e-6
 etol 1e-10
end

%mcrpa
 nroots       8
end

* int 0 1
C  0  0  0 0         0    0
C  1  0  0 1.3385    0    0
H  1  2  0 1.07    120    0
H  1  2  3 1.07    120  180
H  2  1  3 1.07    120    0
H  2  1  3 1.07    120  180
*
```

After the residual norm is below a user-given threshold `TolR` we get the following information

```
Final Eigenvalues

 State       Eigenvalue     RMSD error        Converged
   0          0.3352792890      2.4181038930e-07 T
   1          0.3484190806      9.8077823429e-07 T
   2          0.3514832140      2.7908735363e-07 T
   3          0.3741119713      2.9210937348e-07 T

 4 roots were CONVERGED within 19 iterations!
 64 Sigma vectors were computed in total!
```

and the absorption and ECD spectrum

```
-------------------------------------------------------------------------------------------
-----
                    ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------
-----
     Transition      Energy     Energy  Wavelength fosc(D2)      D2        DX        DY
DZ
                      (eV)      (cm-1)    (nm)                (au**2)    (au)      (au)
(au)
-------------------------------------------------------------------------------------------
-----
  0-1A  ->  1-1A    9.123912   73589.3   135.9   0.430770278   1.92711  -1.38820   0.00000  -0.
00000
  0-1A  ->  2-1A    9.483952   76493.2   130.7   0.009915149   0.04267  -0.00000   0.00000  -0.
20657
  0-1A  ->  3-1A    9.564384   77142.0   129.6   0.000000000   0.00000  -0.00000   0.00000  -0.
00000
  0-1A  ->  4-1A   10.180358   82110.1   121.8   0.000000000   0.00000   0.00000   0.00000   0.
00000
  0-1A  ->  5-1A   10.187869   82170.7   121.7   0.000000000   0.00000   0.00000  -0.00000   0.
00000
  0-1A  ->  6-1A   10.995304   88683.1   112.8   0.000000000   0.00000  -0.00000  -0.00000  -0.
00000
  0-1A  ->  7-1A   12.188654   98308.1   101.7   0.000000000   0.00000   0.00000  -0.00000  -0.
00000
  0-1A  ->  8-1A   12.543751  101172.2    98.8   0.000000000   0.00000  -0.00000   0.00000   0.
00000

...
-------------------------------------------------------------------------------------------
                    CD SPECTRUM VIA TRANSITION VELOCITY DIPOLE MOMENTS
-------------------------------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength    R        MX        MY        MZ
                      (eV)      (cm-1)    (nm)   (1e40*cgs)   (au)      (au)      (au)
-------------------------------------------------------------------------------------------
  0-1A  ->  1-1A    9.123912   73589.3   135.9   -0.00000  -0.00000  -0.00000  -0.00000
  0-1A  ->  2-1A    9.483952   76493.2   130.7    0.00000  -0.00000  -0.00000   0.00000
  0-1A  ->  3-1A    9.564384   77142.0   129.6    0.00000   0.69943  -0.00000   0.00000
  0-1A  ->  4-1A   10.180358   82110.1   121.8   -0.00000   0.15776   0.00000  -0.00000
  0-1A  ->  5-1A   10.187869   82170.7   121.7    0.00000   0.00000  -0.73302   0.00000
  0-1A  ->  6-1A   10.995304   88683.1   112.8   -0.00000   0.00000  -0.54038  -0.00000
  0-1A  ->  7-1A   12.188654   98308.1   101.7   -0.00000   0.00000   0.00000  -0.00000
  0-1A  ->  8-1A   12.543751  101172.2    98.8    0.00000   0.00000   0.00000  -0.90854
```

Capabilities

At the moment, we can simulate UV/Vis and ECD absorption spectra by computing excitation energies, oscillator and rotatory strengths. The code is parallelized and the computational bottleneck is the integral direct AO-Fock matrix construction. All intermediates that depend on the number of states are stored on disk, which makes the MC-RPA implementation suitable for computing many low-lying electronic states of larger molecules. Abelian point-group symmetry can be exploited in the calculation (up to D_{2h}). But there are no calculations of spin-flip excitations possible at the moment. That means all excited states will have the same spin as the reference state, which is specified in the `%casscf` input block.

It is also possible to analyze and visualize the ground-to-excited-state transitions by means of natural transition orbitals[561] (NTO), which is explained in more detail in section *Excited States via MC-RPA*.

For further details, please study our recent publications[379, 380].

6.6.3 Excited States with EOM-CCSD

The methods described in the previous section are all based on the single excitation framework. For a more accurate treatment, double excitations should also be considered. The equation of motion (EOM) CCSD method (and the closely related family of linear response CC methods) provides an accurate way of describing excited, ionized and electron attached states based on singles and doubles excitations within the coupled-cluster framework. In this chapter, the typical usage of the EOM-CCSD routine will be described, along with a short list of its present capabilities. A detailed description will be given in Section *Excited States via EOM-CCSD*.

General Use

The simplest way to perform an EOM calculation is via the usage of the `EOM-CCSD` keyword, together with the specification of the desired number of roots:

```
! RHF EOM-CCSD cc-pVDZ TightSCF

%mdci
  nroots 9
end

*xyz 0 1
  C     0.016227   -0.000000    0.000000
  O     1.236847    0.000000   -0.000000
  H    -0.576537    0.951580   -0.000000
  H    -0.576537   -0.951580   -0.000000
*
```

The above input will call the EOM routine with default settings. The main output is a list of excitation energies, augmented with some further state specific data. For the above input, the following output is obtained:

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT=  1:  0.147823 au     4.022 eV   32443.5 cm**-1
  Amplitude    Excitation
   0.107945     4 ->   8
   0.665496     7 ->   8
   0.104633     7 ->   8     6 ->   8
  Ground state amplitude:  0.000000
Percentage singles character=     92.32

IROOT=  2:  0.314133 au     8.548 eV   68944.3 cm**-1
  Amplitude    Excitation
```

(continues on next page)

(continued from previous page)

```
   0.671246     7 ->   9
  Ground state amplitude: -0.000000
Percentage singles character=     90.42

IROOT=  3:  0.343833 au     9.356 eV   75462.6 cm**-1
  Amplitude    Excitation
  -0.670633     5 ->   8
  -0.112538     6 ->   8     5 ->   8
  Ground state amplitude:  0.000000
Percentage singles character=     92.00

IROOT=  4:  0.364199 au     9.910 eV   79932.5 cm**-1
  Amplitude    Excitation
   0.102777     4 ->  10
  -0.484661     6 ->   8
   0.438311     7 ->  10
  -0.167512     6 ->   8     6 ->   8
  Ground state amplitude: -0.021060
Percentage singles character=     87.22

IROOT=  5:  0.389398 au    10.596 eV   85463.0 cm**-1
  Amplitude    Excitation
   0.646812     4 ->   8
  -0.122387     7 ->   8
   0.171366     7 ->   8     6 ->   8
  Ground state amplitude:  0.000000
Percentage singles character=     87.47

IROOT=  6:  0.414587 au    11.281 eV   90991.4 cm**-1
  Amplitude    Excitation
  -0.378418     6 ->   8
  -0.537292     7 ->  10
  -0.124246     6 ->   8     6 ->   8
  Ground state amplitude: -0.061047
Percentage singles character=     89.13

IROOT=  7:  0.423861 au    11.534 eV   93026.7 cm**-1
  Amplitude    Excitation
   0.673806     7 ->  11
  Ground state amplitude:  0.000000
Percentage singles character=     93.14

IROOT=  8:  0.444201 au    12.087 eV   97490.8 cm**-1
  Amplitude    Excitation
   0.664877     6 ->   9
   0.130475     6 ->   9     6 ->   8
  Ground state amplitude: -0.000000
Percentage singles character=     87.17

IROOT=  9:  0.510514 au    13.892 eV  112044.8 cm**-1
  Amplitude    Excitation
  -0.665791     6 ->  10
   0.114259     6 ->  15
  -0.124374     6 ->  10     6 ->   8
  Ground state amplitude: -0.000000
```

The IP and EA versions can be called using the keywords `IP-EOM-CCSD` and `EA-EOM-CCSD` respectively. For open-shell systems (UHF reference wavefunction), IP/EA-EOM-CCSD calculations require an additional keywords. Namely, an IP/EA calculation involving the removal/attachment of an α electron is requested by setting the `DoAlpha` keyword to true in the %mdci block, while setting the `DoBeta` keyword to true selects an IP/EA calculation for the removal/attachment of a β electron. Note that `DoAlpha` and `DoBeta` cannot simultaneously be

true and that the calculation defaults to one in which DoAlpha is true if no keyword is specified on input. A simple example of the input for a UHF IP-EOM-CCSD calculation for the removal of an α electron is given below.

```
! IP-EOM-CCSD cc-pVDZ
%mdci
DoAlpha true
NRoots 7
end

*xyz 0 3
  O      0.0 0.0 0.0
  O      0.0 0.0 1.207
*
```

Capabilities

At present, the EOM routine is able to perform excited, ionized and electron attached state calculations, for both closed- or open-shell systems, using RHF or UHF reference wavefunctions, respectively. It can be used for serial and parallel calculations. The method is available in the back-transformed PNO and DLPNO framework enabling the calculation of large molecules - see Section *Excited States with PNO based coupled cluster methods* and Section *Excited States with DLPNO based coupled cluster methods*. In the closed-shell case (RHF), a lower scaling version can be invoked by setting the CCSD2 keyword to true in the %mdci section. The latter is a second order approximation to the conventional EOM-CCSD. For the time being, the most useful information provided is the list of the excitation energies, the ionization potentials or the electron affinities. The ground to excited state transition moments are also available for the closed-shell implementation of EE-EOM-CCSD.

6.6.4 Excited States with ADC2

Among the various approximate correlation methods available for excited states, one of the most popular one is algebraic diagrammatic construction(ADC) method. The ADC has it origin in the Green's function theory. It expands the energy and wave-function in perturbation order and can directly calculate the excitation energy, ionization potential and electron affinity, similar to that in the EOM-CCSD method. Because of the symmetric eigenvalue problem in ADC, the calculation of properties are more straight forward to calculate than EOM-CCSD. In ORCA, only the second-order approximation to ADC(ADC2) is implemented. It scales as O(N^5) power of the basis set.

General Use

The simplest way to perform an ADC2 calculation is via the usage of the ADC2 keyword, together with the specification of the desired number of roots:

```
! ADC2 cc-pVDZ cc-pVDZ/C TightSCF

%mdci
  nroots 9
end

*xyz 0 1
  C     0.016227   -0.000000    0.000000
  O     1.236847    0.000000   -0.000000
  H    -0.576537    0.951580   -0.000000
  H    -0.576537   -0.951580   -0.000000
*
```

The above input will call the ADC2 routine with default settings. The main output is a list of excitation energies, augmented with some further state specific data. The integral transformation in the ADC2 implementation of ORCA is done using the density-fitting approximation. Therefore, one need to specify an auxiliary basis. For the above input, the following output is obtained:

```
---------------------
ADC(2) RESULTS (RHS)
---------------------

IROOT=  1:  0.146914 au     3.998 eV   32243.8 cm**-1
  Amplitude    Excitation
  -0.116970     4 ->   8
  -0.672069     7 ->   8
IROOT=  2:  0.286012 au     7.783 eV   62772.3 cm**-1
  Amplitude    Excitation
  -0.659777     7 ->   9
IROOT=  3:  0.341919 au     9.304 eV   75042.4 cm**-1
  Amplitude    Excitation
  -0.676913     5 ->   8
IROOT=  4:  0.352206 au     9.584 eV   77300.2 cm**-1
  Amplitude    Excitation
  -0.126824     4 ->  10
   0.360690     6 ->   8
  -0.547669     7 ->  10
IROOT=  5:  0.393965 au    10.720 eV   86465.3 cm**-1
  Amplitude    Excitation
  -0.551344     6 ->   8
  -0.363451     7 ->  10
  -0.109270     6 ->   8     6 ->   8
IROOT=  6:  0.404946 au    11.019 eV   88875.5 cm**-1
  Amplitude    Excitation
   0.669682     4 ->   8
  -0.126557     7 ->   8
IROOT=  7:  0.412800 au    11.233 eV   90599.2 cm**-1
  Amplitude    Excitation
   0.100274     4 ->  11
   0.671884     7 ->  11
IROOT=  8:  0.439251 au    11.953 eV   96404.6 cm**-1
  Amplitude    Excitation
  -0.674114     6 ->   9
  -0.104541     6 ->   9     6 ->   8
IROOT=  9:  0.486582 au    13.241 eV  106792.5 cm**-1
  Amplitude    Excitation
  -0.654624     5 ->   9
```

The transition moment for ADC2 in ORCA is calculated using an EOM-like expectation value approach, unlike the traditionally used intermediate state representation. However, the two approaches gives almost identical result.

```
-------------------------------------------------------------------
                    SPECTRUM FOR LEFT-RIGHT TRANSITION MOMENTS
-------------------------------------------------------------------



-------------------------------------------------------------------------------------------------
↪-----
                     ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
↪-----
     Transition      Energy     Energy  Wavelength fosc(D2)      D2        DX        DY       ␣
↪DZ
                      (eV)       (cm-1)    (nm)               (au**2)    (au)      (au)     ␣
↪(au)
-------------------------------------------------------------------------------------------------
↪-----
  0-1A  ->  1-1A    3.997726   32243.8   310.1   0.000000000   0.00000   0.00000  -0.00000   0.
↪00000
  0-1A  ->  2-1A    7.782776   62772.3   159.3   0.096710371   0.50720   0.00000  -0.70536   0.
```

(continues on next page)

(continued from previous page)

```
↪00000
  0-1A  ->  3-1A    9.304078   75042.5   133.3   0.002261744   0.00992  -0.00000  -0.00000   0.
↪09835
  0-1A  ->  4-1A    9.584003   77300.2   129.4   0.007937829   0.03381   0.18502   0.00000  -0.
↪00000
  0-1A  ->  5-1A   10.720332   86465.3   115.7   0.465055079   1.77067   1.32377   0.00000   0.
↪00000
  0-1A  ->  6-1A   11.019150   88875.4   112.5   0.000000000   0.00000   0.00000   0.00000   0.
↪00000
  0-1A  ->  7-1A   11.232869   90599.2   110.4   0.022236623   0.08080  -0.00000   0.28105   0.
↪00000
  0-1A  ->  8-1A   11.952640   96404.5   103.7   0.009103120   0.03109  -0.00000   0.00000  -0.
↪17328
  0-1A  ->  9-1A   13.240575  106792.5    93.6   0.071433742   0.22021  -0.46692   0.00000   0.
↪00000
```

The IP and EA versions can be called using the keywords IP-ADC2 and EA-ADC2, respectively.

Capabilities

At present, the ADC2 module is able to perform excited, ionized and electron attached state calculations, only for closed-shell systems. No open-shell version of the ADC2 is currently available. Below are all the parameters that influence the ADC2 module.

```
%mdci
#ADC2 parameters - defaults displayed
  NDav 20               # maximum size of reduced space (i.e. 20*NRoots)
  CheckEachRoot true    # check convergence for each root separately
  RootHoming true       # apply root homing
  DoLanczos false       # use the Lanczos procedure rather than Davidson
  UseCISUpdate true     # use diagonal CIS for updating
  NInitS 0              # number of roots in the initial guess, if 0, use preset value
  DoRootwise false      # solves for each root separately,
                        # more stable for large number of roots
  FOLLOWCIS false       # follows the initial singles guess
end
```

One can notice that features available in the ADC2 module is quite limited as compared to the EOM module and the option to specifically target the core-orbitals are yet not available. A word of caution, **The 'second order black magic' of ADC2 can fail in many of the cases.** The readers are encouraged to try the DLPNO based EOM-CCSD methods(*Excited States with DLPNO based coupled cluster methods*) which are much more accurate and computationally efficient.

6.6.5 Excited States with STEOM-CCSD

The STEOM-CCSD method provides an efficient way to calculate excitation energies, with an accuracy comparable to the EOM-CCSD approach, at a nominal cost. A detailed description will be given in Section *Excited States via STEOM-CCSD*.

General Use

The simplest way to perform a STEOM calculation is using the `STEOM-CCSD` keyword, together with the specification of the desired number of roots (`NRoots`):

```
! STEOM-CCSD cc-pVDZ TightSCF

%mdci
  NRoots 9        # Number of excited states
  DoDbfilter true # Remove doubly excited states
end

*xyz 0 1
  C     0.016227   -0.000000    0.000000
  O     1.236847    0.000000   -0.000000
  H    -0.576537    0.951580   -0.000000
  H    -0.576537   -0.951580   -0.000000
*
```

The above input calls the STEOM routine with default settings, where, for instance, the doubly excited states are eliminated (`DoDbFilter true`). The main output is a list of excitation energies, augmented with some further state specific data. The STEOMCC approach in ORCA uses state-averaged CIS natural transition orbitals (NTO) for the selection of the active space. For the above input, the following output is obtained:

```
------------------
STEOM-CCSD RESULTS
------------------

IROOT=  1:  0.146552 au     3.988 eV   32164.5 cm**-1
  Amplitude    Excitation
   0.196225     4 ->   8
  -0.979974     7 ->   8

  Amplitude    Excitation in Canonical Basis
  -0.153212     4 ->   8
   0.977931     7 ->   8
  -0.121980     7 ->  13

IROOT=  2:  0.308608 au     8.398 eV   67731.7 cm**-1
  Amplitude    Excitation
  -0.141414     4 ->   9
   0.988498     7 ->   9

  Amplitude    Excitation in Canonical Basis
  -0.989700     7 ->   9

IROOT=  3:  0.336979 au     9.170 eV   73958.3 cm**-1
  Amplitude    Excitation
  -0.994070     5 ->   8

  Amplitude    Excitation in Canonical Basis
   0.983934     5 ->   8
  -0.137018     5 ->  13

IROOT=  4:  0.362974 au     9.877 eV   79663.6 cm**-1
```

(continues on next page)

(continued from previous page)

```
  Amplitude    Excitation
   0.177265     4 ->  10
   0.825223     6 ->   8
  -0.500412     7 ->  10
  -0.118642     7 ->  12

  Amplitude    Excitation in Canonical Basis
  -0.152751     4 ->  10
  -0.821991     6 ->   8
   0.506004     7 ->  10

IROOT=  5:  0.402096 au    10.942 eV   88249.9 cm**-1
  Amplitude    Excitation
   0.100684     5 ->  11
   0.617781     6 ->   8
   0.761064     7 ->  10

  Amplitude    Excitation in Canonical Basis
  -0.612814     6 ->   8
  -0.754151     7 ->  10

IROOT=  6:  0.421001 au    11.456 eV   92399.1 cm**-1
  Amplitude    Excitation
  -0.165095     4 ->  11
   0.983905     7 ->  11

  Amplitude    Excitation in Canonical Basis
   0.121348     4 ->  11
  -0.983982     7 ->  11

IROOT=  7:  0.445178 au    12.114 eV   97705.3 cm**-1
  Amplitude    Excitation
   0.995471     6 ->   9

  Amplitude    Excitation in Canonical Basis
  -0.989647     6 ->   9

IROOT=  8:  0.462852 au    12.595 eV  101584.3 cm**-1
  Amplitude    Excitation
  -0.985707     4 ->   8
  -0.130220     6 ->  10

  Amplitude    Excitation in Canonical Basis
   0.975461     4 ->   8
  -0.147945     4 ->  13
   0.128680     6 ->  10

IROOT=  9:  0.512757 au    13.953 eV  112537.1 cm**-1
  Amplitude    Excitation
   0.121760     4 ->   8
  -0.989185     6 ->  10

  Amplitude    Excitation in Canonical Basis
  -0.121079     4 ->   8
   0.979589     6 ->  10
  -0.154643     6 ->  15
```

The first set of excitation amplitudes, printed for each root, have been calculated in the CIS NTO (Natural Transition Orbitals) basis. The second set of amplitudes have been evaluated in the RHF canonical basis.

Capabilities

At present, the STEOM routine is able to calculate excitation energies, for both closed- or open-shell systems, using an RHF or UHF reference function, respectively. It can be used for both serial and parallel calculations. The method is available in the back-tranformed PNO and DLPNO framework allowing the calculation of large molecules (Section *Capabilities* and *Excited States with DLPNO based coupled cluster methods*). In the closed-shell case (RHF), a lower scaling version can be invoked by setting the CCSD2 keyword to true in the %mdci section, which sets a second order approximation to the exact parent approach. The transition moments can also be obtained for closed- and open-shell systems. For more details see Section *Excited States via STEOM-CCSD*.

6.6.6 Excited States with IH-FSMR-CCSD

The intermediate Hamiltonian Fock-space coupled cluster method (IH-FSMR-CCSD) provides an alternate way to calculate excitation energies, with an accuracy comparable to the STEOM-CCSD approach. A detailed description is given in Section *General Description*.

General Use

The IH-FSMR-CCSD calculation is called using the simple input keyword IH-FSMR-CCSD and specifying the desired number of excited states (NRoots) in the %mdci block.:

```
! IH-FSMR-CCSD cc-pVDZ TightSCF

%mdci
  nroots 6
end

*xyz 0 1
  C     0.016227   -0.000000    0.000000
  O     1.236847    0.000000   -0.000000
  H    -0.576537    0.951580   -0.000000
  H    -0.576537   -0.951580   -0.000000
*
```

The above input will call the IH-FSMR-CCSD routine with default settings. The main output is a list of excitation energies, augmented with some further state specific data. The IH-FSMR-CCSD approach in ORCA uses state-averaged CIS natural transition orbitals(NTO) for the selection of the active space - similar to STEOM-CCSD. For the above input, the following output is obtained:

```
------------------
IH-FSMR-CCSD RESULTS
------------------

IROOT=  1:  0.144300 au      3.927 eV   31670.2 cm**-1
  Amplitude    Excitation
  -0.173154     4 ->   8
  -0.984515     7 ->   8
  Ground state amplitude:  0.000000

Percentage Active Character     99.93

  Amplitude    Excitation in Canonical Basis
  -0.170951     4 ->   8
   0.976572     7 ->   8
  -0.111271     7 ->  13

IROOT=  2:  0.309445 au      8.420 eV   67915.3 cm**-1
  Amplitude    Excitation
   0.993733     7 ->   9
```

(continues on next page)

(continued from previous page)

```
  Ground state amplitude:  0.000000

Percentage Active Character     99.65

  Amplitude    Excitation in Canonical Basis
  -0.991663     7 ->   9

IROOT=  3:  0.335928 au     9.141 eV   73727.6 cm**-1
  Amplitude    Excitation
   0.994414     5 ->   8
  Ground state amplitude:  0.000000

Percentage Active Character     98.98

  Amplitude    Excitation in Canonical Basis
  -0.986238     5 ->   8
   0.122237     5 ->  13

IROOT=  4:  0.358174 au     9.746 eV   78610.1 cm**-1
  Amplitude    Excitation
  -0.176281     4 ->  10
   0.736812     6 ->   8
  -0.594366     7 ->  10
  -0.213482     7 ->  12
  Ground state amplitude:  0.000000

Percentage Active Character     92.76

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
  -0.184685     4 ->  10
   0.734266     6 ->   8
   0.630467     7 ->  10

IROOT=  5:  0.385852 au    10.500 eV   84684.8 cm**-1
  Amplitude    Excitation
  -0.981051     4 ->   8
   0.179230     7 ->   8
  Ground state amplitude:  0.000000

Percentage Active Character     99.86

  Amplitude    Excitation in Canonical Basis
  -0.973509     4 ->   8
   0.112468     4 ->  13
  -0.178795     7 ->   8

IROOT=  6:  0.445155 au    12.113 eV   97700.1 cm**-1
  Amplitude    Excitation
  -0.996250     6 ->   9
  Ground state amplitude:  0.000000

Percentage Active Character     99.38

  Amplitude    Excitation in Canonical Basis
  -0.992457     6 ->   9
```

The first set of excitation amplitudes, printed for each root, have been calculated in the CIS NTO (Natural Transition Orbitals) basis. The second set of amplitudes have been evaluated in the RHF canonical basis.

Capabilities

At present, the IH-FSMR-CCSD routine is able to calculate excitation energies, for only closed shell systems using an RHF reference. It can be used for both serial and parallel calculations. In the closed-shell case (RHF), a lower scaling version can be invoked by using bt-PNO approximation. The transition moments and solvation correction can be obtained using the CIS approximation.

6.6.7 Excited States with PNO based coupled cluster methods

The methods described in the previous section are performed over a canonical CCSD or MP2 ground state. The use of canonical CCSD amplitudes restricts the use of EOM-CC and STEOM-CC methods to small molecules. The use of MP2 amplitudes is possible (e.g. the EOM-CCSD(2) or STEOM-CCSD(2) approaches), but it seriously compromises the accuracy of the method.

The bt-PNO-EOM-CCSD methods gives an economical compromise between accuracy and computational cost by replacing the most expensive ground state CCSD calculation with a DLPNO based CCSD calculation. The typical deviation of the results from the canonical EOM-CCSD results is around 0.01 eV. A detailed description will be given in *Excited States using PNO-based coupled cluster*.

General Use

The simplest way to perform a PNO based EOM calculation is via the usage of the bt-PNO-EOM-CCSD keyword, together with the specification of the desired number of roots. The specification of an auxilary basis set is also required, just as for ground state DLPNO-CCSD calculations.

```
! bt-PNO-EOM-CCSD def2-TZVP def2-TZVP/C def2/J TightSCF

%mdci
  nroots 9
end

*xyz 0 1
  C     0.016227   -0.000000    0.000000
  O     1.236847    0.000000   -0.000000
  H    -0.576537    0.951580   -0.000000
  H    -0.576537   -0.951580   -0.000000
*
```

The output is similar to that from a canonical EOM-CCSD calculation:

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT=  1:  0.145339 au      3.955 eV   31898.3 cm**-1
Amplitude    Excitation
-0.402736     2 ->   8
-0.101455     2 ->  13
0.402595      3 ->   8
0.101420      3 ->  13
0.231140      6 ->   8
-0.231142     7 ->   8
Ground state amplitude:  0.000000
IROOT=  2:  0.311159 au      8.467 eV   68291.5 cm**-1
Amplitude    Excitation
-0.382967     2 ->   9
0.382816      3 ->   9
0.257265      6 ->   9
-0.257276     7 ->   9
```

(continues on next page)

(continued from previous page)

```
Ground state amplitude:  0.000000
IROOT=  3:  0.337350 au      9.180 eV   74039.8 cm**-1
Amplitude    Excitation
0.342418     2 ->   8
0.342586     3 ->   8
-0.257991     4 ->   8
0.257936     5 ->   8
0.172202     6 ->   8
0.172230     7 ->   8
Ground state amplitude:  0.000010
IROOT=  4:  0.348181 au      9.474 eV   76416.9 cm**-1
Amplitude    Excitation
0.393166     2 ->  11
-0.393020     3 ->  11
-0.246227     6 ->  11
0.246232     7 ->  11
Ground state amplitude:  0.000001
IROOT=  5:  0.354611 au      9.649 eV   77828.2 cm**-1
Amplitude    Excitation
0.226219     2 ->  10
-0.226139     3 ->  10
-0.385817     4 ->   8
-0.385755     5 ->   8
-0.100298     6 ->  10
0.100300     7 ->  10
Ground state amplitude:  0.032619
IROOT=  6:  0.379574 au     10.329 eV   83307.0 cm**-1
Amplitude    Excitation
0.214487     2 ->   8
-0.214423     3 ->   8
0.402942     6 ->   8
-0.402947     7 ->   8
Ground state amplitude: -0.000001
IROOT=  7:  0.386805 au     10.525 eV   84893.8 cm**-1
Amplitude    Excitation
-0.337735     2 ->  10
-0.113836     2 ->  14
0.337611     3 ->  10
0.113798     3 ->  14
-0.182472     4 ->   8
-0.182457     5 ->   8
0.239131     6 ->  10
-0.239136     7 ->  10
Ground state amplitude:  0.038944
IROOT=  8:  0.440569 au     11.989 eV   96693.8 cm**-1
Amplitude    Excitation
-0.463727     4 ->   9
-0.463700     5 ->   9
Ground state amplitude: -0.000004
IROOT=  9:  0.447197 au     12.169 eV   98148.3 cm**-1
Amplitude    Excitation
-0.107379     2 ->   8
0.385138     2 ->  13
0.107343     3 ->   8
-0.385019     3 ->  13
-0.254544     6 ->  13
0.254548     7 ->  13
Ground state amplitude:  0.000000
```

The IP and EA versions can be called by using the keywords bt-PNO-IP-EOM-CCSD and bt-PNO-EA-EOM-CCSD, respectively. Furthermore, the STEOM version can be invoked by using the keywords bt-PNO-STEOM-CCSD.

Capabilities

All of the features of canonical EOM-CC and STEOM-CC are available in the PNO based approaches for both closed- and open-shell systems.

6.6.8 Excited States with DLPNO based coupled cluster methods

The DLPNO-STEOM-CCSD method uses the full potential of DLPNO to reduce the computational scaling while keeping the accuracy of STEOM-CCSD.

Important: DLPNO-STEOM-CCSD is currently only available for closed-shell systems!

General Use

The simplest way to perform a DLPNO based STEOM calculation is via the usage of the `STEOM-DLPNO-CCSD` keyword, together with the specification of the desired number of roots. The specification of an auxiliary basis set is also required, just as for ground state DLPNO-CCSD calculations.

As any CCSD methods, it is important to allow ORCA to access a significant amount of memory. In term of scaling the limiting factor of the method is the size of temporary files and thus the disk space. For molecules above 1500 basis functions it starts to increase exponentially up to several teraoctets.

Here is the standard input we would recommend for STEOM-DLPNO-CCSD calculations. More information on the different keywords and other capabilities are available in the detailed part of the manual *Excited States via STEOM-CCSD*, *Excited States via DLPNO-STEOM-CCSD*. The following publications referenced some applications for this method either in organic molecules [100], [804] or for Semiconductors [213].

```
! STEOM-DLPNO-CCSD def2-TZVP def2-TZVP/C def2/J TightSCF

%mdci
  NRoots 6
  DoRootWise true
  OThresh 0.005
  VThresh 0.005
  TCutPNOSingles 1e-11
  NDAV 400
  DoStoreSTEOM true
  DoSimpleDens false
  AddL2Term True
  DTol 1e-5
end

* xyz 0 1
  C     0.016227   -0.000000    0.000000
  O     1.236847    0.000000   -0.000000
  H    -0.576537    0.951580   -0.000000
  H    -0.576537   -0.951580   -0.000000
*
```

The output is similar to that from a canonical DLPNO-STEOM-CCSD calculation:

```
------------------
STEOM-CCSD RESULTS
------------------

IROOT=  1:  0.144275 au      3.926 eV   31664.7 cm**-1
  Amplitude    Excitation
  -0.142146     4 ->   8
  -0.988793     7 ->   8
  Ground state amplitude: -0.000000
```

(continues on next page)

(continued from previous page)

```
Percentage Active Character     99.79

  Amplitude    Excitation in Canonical Basis
  -0.134936     4 ->   8
  -0.955031     7 ->   8
   0.236745     7 ->  13

IROOT=  2:  0.308093 au     8.384 eV   67618.5 cm**-1
  Amplitude    Excitation
  -0.971471     7 ->   9
  -0.214898     7 ->  10
  Ground state amplitude: -0.000000

Percentage Active Character     99.67

  Amplitude    Excitation in Canonical Basis
  -0.956930     7 ->   9
   0.236567     7 ->  11
  -0.102574     7 ->  16

IROOT=  3:  0.331796 au     9.029 eV   72820.8 cm**-1
  Amplitude    Excitation
   0.993677     5 ->   8
  Ground state amplitude: -0.000000

Percentage Active Character     98.87

  Amplitude    Excitation in Canonical Basis
  -0.957218     5 ->   8
   0.250144     5 ->  13
   0.105963     5 ->  18

IROOT=  4:  0.346876 au     9.439 eV   76130.5 cm**-1
  Amplitude    Excitation
  -0.104900     4 ->  10
   0.198181     7 ->   9
  -0.972571     7 ->  10
  Ground state amplitude:  0.000000

Percentage Active Character     99.65

  Amplitude    Excitation in Canonical Basis
   0.100880     4 ->  11
   0.218876     7 ->   9
   0.956922     7 ->  11
  -0.113898     7 ->  19

IROOT=  5:  0.347460 au     9.455 eV   76258.7 cm**-1
  Amplitude    Excitation
  -0.139550     4 ->  11
  -0.106648     4 ->  12
  -0.801181     6 ->   8
  -0.455618     7 ->  11
  -0.302466     7 ->  12
  Ground state amplitude:  0.027266

Percentage Active Character     87.08

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------
```

(continues on next page)

(continued from previous page)

```
  Amplitude    Excitation in Canonical Basis
  -0.163789     4 ->  10
  -0.785695     6 ->   8
   0.159147     6 ->  13
  -0.527842     7 ->  10
   0.133087     7 ->  17

IROOT=  6:  0.379059 au    10.315 eV   83193.9 cm**-1
  Amplitude    Excitation
  -0.983700     4 ->   8
   0.155238     7 ->   8
  Ground state amplitude: -0.000000

Percentage Active Character     99.48

  Amplitude    Excitation in Canonical Basis
  -0.951092     4 ->   8
   0.235048     4 ->  13
   0.157713     7 ->   8


STEOM-CCSD done (    2.4 sec)
Transforming integrals                             ... done

-------------------------------------------------------------------
               UNRELAXED EXCITED STATE DIPOLE MOMENTS
-------------------------------------------------------------------
               E(eV)     DX(au)      DY(au)      DZ(au)      |D|(D)
IROOT=  0:     0.000   -0.928848   -0.000000   -0.000000    2.360944
IROOT=  1:     3.926   -0.627710   -0.000000   -0.000002    1.595512
IROOT=  2:     8.384    1.034480   -0.000000   -0.000000    2.629438
IROOT=  3:     9.029   -0.401280   -0.000000    0.000000    1.019972
IROOT=  4:     9.439   -0.250433    0.000000    0.000002    0.636550
IROOT=  5:     9.455    0.304050    0.000000   -0.000000    0.772833
IROOT=  6:    10.315   -1.244475    0.000000    0.000000    3.163205
-------------------------------------------------------------------

...

-------------------------------------------------------------------------------------------------
-----
                     ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
-----
     Transition      Energy     Energy  Wavelength fosc(D2)      D2        DX        DY        
DZ
                      (eV)      (cm-1)    (nm)                 (au**2)    (au)      (au)      
(au)
-------------------------------------------------------------------------------------------------
-----
  0-1A  ->  1-1A    3.925923   31664.7   315.8   0.000000000   0.00000  -0.00000  -0.00000   0.
00000
  0-1A  ->  2-1A    8.383625   67618.5   147.9   0.088173876   0.42929  -0.00000   0.65546  -0.
00000
  0-1A  ->  3-1A    9.028624   72820.8   137.3   0.000908615   0.00411   0.00000  -0.00000  -0.
06033
  0-1A  ->  4-1A    9.438972   76130.5   131.4   0.057997877   0.25080   0.00000  -0.49380  -0.
00000
  0-1A  ->  5-1A    9.454876   76258.7   131.1   0.029389253   0.12687   0.35160  -0.00000  -0.
00000
```

(continues on next page)

(continued from previous page)

```
  0-1A  ->  6-1A   10.314723   83193.9   120.2   0.000000000   0.00000  -0.00000   0.00000   0.
↪00000

...

 STEOM-CCSD done in (    1.3)
```

The IP and EA versions can be called by using the keywords IP-EOM-DLPNO-CCSD and EA-EOM-DLPNO-CCSD, respectively. As in canonical STEOM-CCSD, the first set of excitation amplitudes, printed for each root, are calculated in the CIS NTO (Natural Transition Orbitals) basis, while the second set is evaluated in the RHF canonical basis.

6.6.9 Excited States with DeltaSCF

The `DeltaSCF` approach can converge the SCF directly to excited states. Since this method involves a few more details, it is more thoroughly described on its specific section, please check *DeltaSCF: Converging to Arbitrary Single-Reference Wavefunctions*.

6.7 Multireference Configuration Interaction and Pertubation Theory

6.7.1 Introductory Remarks

ORCA contains a multireference correlation module designed for traditional (uncontracted) approaches (configuration interaction, MR-CI, and perturbation theory, MR-PT). For clarification, these approaches have in common that they consider excitations from each and every configuration state function (CSF) of the reference wavefunction. Hence, the computational cost of such approaches grows rapidly with the size of the reference space (e.g. CAS-CI). Internally contracted on the other hand define excitations with respect to the entire reference wavefunction and hence do not share the same bottlenecks. ORCA also features internally contracted approaches (perturbation theory, *NEVPT2* and configuration interaction, *FIC-MRCI*), which are described elsewhere in the manual.

> **Note**
>
> NEVPT2 is typically the method of choice as it is fast and easy to use. It is highly recommended to check the respective section, when new to the field. The following chapter focuses on the traditional multi-reference approaches as part of the `orca_mrci` module.

Although there has been quite a bit of experience with it, this part of the program is still somewhat hard to use and requires patience and careful testing before the results should be accepted. While we try to make your life as easy as possible, you have to be aware that ultimately any meaningful multireference *ab initio* calculation requires more insight and planning from the user side than standard SCF or DFT calculation or single reference correlation approaches like MP2 – so don't be fainthearted! You should also be aware that with multireference methods it is very easy to let a large computer run for a long time and still to not produce a meaningful result – your insight is a key ingredient to a successful application! Below a few examples illustrate some basic uses of the `orca_mrci` module.

RI-approximation

First of all, it is important to understand that the default mode of the MR-CI module in its present implementation performs a full integral transformation from the AO to the MO basis. This becomes very laborious and extremely memory intensive beyond approximately 200 MOs that are included in the CI. Alternatively, one can construct molecular electron-electron repulsion integrals from the resolution of the identity (RI) approximation. *Thus a meaningful auxiliary basis set must be provided if this option is chosen*. We recommend the fitting bases developed by the TurboMole developers for MP2 calculations. These give accurate transition energies; however, the error in the total energies is somewhat higher and may be on the order of 1 mEh or so. Check `IntMode` to change the default mode for the integral transformation. Note that in either way, the individually selecting MRCI module requires to have all integrals in memory which sets a limit on the size of the molecule that can be studied.

Individual Selection

Secondly, it is important to understand that the MR-CI module is of the *individually selecting* type. Thus, only those excited configuration state functions (CSFs) which interact more strongly than a given threshold ($\mathbf{T}_{\mathbf{sel}}$) with the 0^{th} order approximations to the target states will be included in the variational procedure. The effect of the rejected CSFs is estimated using second order perturbation theory. The 0^{th} order approximations to the target states are obtained from the diagonalization of the reference space configurations. A further approximation is to reduce the size of this reference space through another selection – all initial references which contribute less than a second threshold ($\mathbf{T}_{\mathbf{pre}}$) to the 0^{th} order states are rejected from the reference space.

Single excitations

One important aspect concerns the single excitations. If the reference orbitals come from a CASSCF calculation the matrix elements between the reference state and the single excitations vanishes and the singles will not be selected. However, they contribute to fourth and higher orders in perturbation theory and may be important for obtaining smooth potential energy surfaces and accurate molecular properties. Hence, the default mode of the MRCI module requires to include all of the single excitations via the flag `AllSingles =true`. This may lead to lengthy computations if the reference spaces becomes large!

Reference Spaces

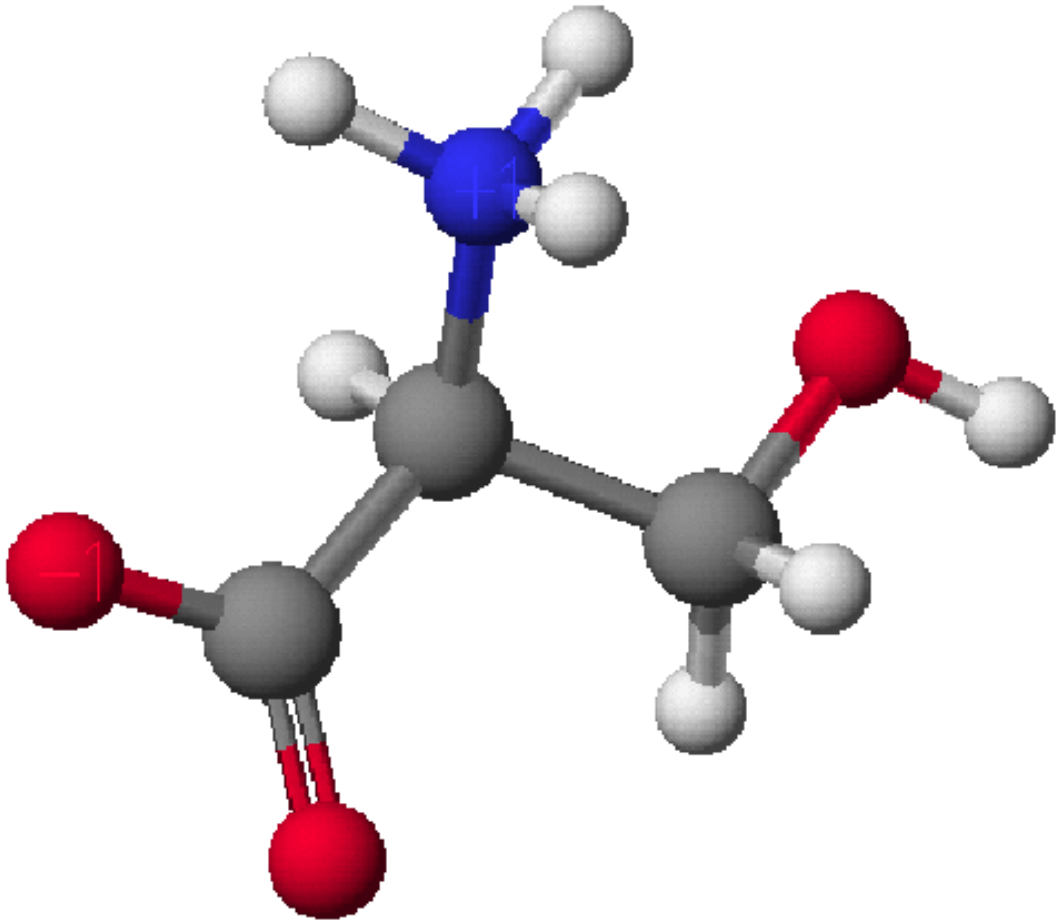
Third, the reference spaces in the MR-CI module can be of the complete active space (**CAS(n-electrons,m-orbitals)**) or restricted active space (**RAS**, explained later) type. It is important to understand that the program uses the orbitals around the HOMO-LUMO gap as provided by the user to build up the reference space! Thus, if the orbitals that you want to put in the active space are not coming "naturally" from your SCF calculation in the right place you have to reorder them using the "`moread`" and "`rotate`" features together with the `NoIter` directive. To select the most meaningful and economic reference space is the most important step in a multireference calculation. It *always* requires insight from the user side and also care and, perhaps, a little trial and error.

Size Consistency

Fourth, it is important to understand that CI type methods are *not* size consistent. Practically speaking the energy of the supermolecule A-B with noninteracting A and B fragments is not equal to the energies of isolated A and isolated B. There are approximate ways to account for this (**ACPF, AQCC** and **CEPA** methods) but the effect will be present in the energies, the more so the more electrons are included in the treatment. The same is *not* true for the perturbation theory based methods which are size consistent as long as the reference wavefunction is.

Performance

There are many flags that control the performance of the MR-CI program. Please refer to chapter 0 for a description of possible flags, thresholds and cut-offs. The most important thresholds are T_{sel} and T_{pre}, and for SORCI also T_{nat}.



For some methods, like ACPF, it is possible to compare the performance of the MRCI module with the performance of the MDCI module. The MDCI module has been written to provide optimum performance if no approximations are introduced. The MRCI module has ben written more with the idea of flexibility rather than the idea of performance. Let us compare the performance of the two programs in a slightly nontrivial calculation – the zwitter-ionic form of serine. We compare the selecting MRCI approach with the approximation free MDCI module. The molecular size is such that still all four index integrals can be stored in memory.

Table 6.8: Comparison of the performance of the MRCI and MDCI modules for a single reference calculation with the bn-ANO-DZP basis set on the zwitter-ionic form of serine (14 atoms, 133 basis functions).

Module	Method	T_{sel}(Eh)	Time (sec)	Energy (Eh)
MRCI	ACPF	10^{-6}	3277	-397.943250
MDCI	ACPF	0	1530	-397.946429
MDCI	CCSD	0	2995	-397.934824
MDCI	CCSD(T)	0	5146	-397.974239

The selecting ACPF calculation selects about 15% of the possible double excitations and solves a secular problem of size $\approx$ 360,000 CSFs. The MDCI module ACPF calculation optimizes approximately 2.5 million wavefunction

amplitudes — and this is not a large molecule or a large basis set! Despite the fact that the MDCI module makes no approximation, it runs twice as fast as the *selected* MRCI module and an estimated 50 times faster than the *unselected* MRCI module! This will become even more pronounced for the larger and more accurate basis sets that one should use in such calculations anyways. The error of the selection is on the order of 3 mEh or 2 kcal/mol in the total energy. One can hope that at least part of this error cancels upon taking energy differences.[1] The more rigorous CCSD calculation takes about a factor of two longer than the ACPF calculation which seems reasonable. The triples add another factor of roughly 2 in this example but this will increase for larger calculations since it has a steeper scaling with the system size. The ACPF energy is intermediate between CCSD and CCSD(T) which is typical — ACPF overshoots the effects of disconnected quadruples which partially compensates for the neglect of triples.

These timings will strongly depend on the system that you run the calculation on. Nevertheless, what you should take from this example are the message that if you can use the MDCI module, do it.

The MDCI module can avoid a full integral transformation for larger systems while the MRCI module can use selection and the RI approximation for larger systems. Both types of calculation will become very expensive very quickly! Approximate MDCI calculations are under development.

Symmetry

The MRCI program really takes advantage of symmetry adapted orbitals. In this case the MRCI matrix can be blocked according to irreducible representations and be diagonalized irrep by irrep. This is a big computational advantage and allows one to converge on specific excited states much more readily than if symmetry is not taken into account.

The syntax is relatively easy. If you specify:

```
newblock 1 *
        nroots 8
        refs cas(4,4) end
        end
```

Then the "*" indicates that this is to be repeated in each irrep of the point group. Thus, in C_{2v} the program would calculate 8 singlet roots in each of the four irreps of the C_{2v} point group thus leading to a total of 32 states.

Alternatively, you can calculate just a few roots in the desired irreps:

```
newblock 1 0
        nroots 3
        refs cas(4,4) end
        end
      newblock 1 2
        nroots 5
        refs cas(4,4) end
        end
      newblock 3 1
        nroots 1
        refs cas(4,4) end
        end
```

In this example, we would calculate 3 singlet roots in the irrep "0" (which is A_1), then five roots in irrep "2" (which is B_1) and then 1 triplet root in irrep 1 (which is B_2).

Obviously, the results with and without symmetry will differ slightly. This is due to the fact that without symmetry the reference space will contain references that belong to "wrong" symmetry but will carry with them excited configurations of "right" symmetry. Hence, the calculation without use of symmetry will have more selected CSFs and hence a slightly lower energy. This appears to be unavoidable. However, the effects should not be very large for well designed reference spaces since the additional CSFs do not belong to the first order interacing space.

[1] Depending on whether one wants to take a pessimistic or an optimistic view one could either say that this result shows what can be achieved with a code that is dedicated to a single determinant reference. Alternatively one could (and perhaps should) complain about the high price one pays for the generality of the MRCI approach. In any case, the name of the game would be to develop MR approaches that are equally efficient to single reference approaches. See FIC-MRCI chapter for more information.

6.7.2 A Tutorial Type Example of a MR Calculation

Perhaps, the most important use of the MR-CI module is for the calculation of transition energies and optical spectra. Let us first calculate the first excited singlet and triplet state of the formaldehyde molecule using the MR-CI method together with the Davidson correction to approximately account for the effect of unlinked quadruple substitutions. We deliberately choose a somewhat small basis set for this calculation which is already reasonable since we only look at a valence excited state and want to demonstrate the principle.

Suppose that we already know from a ground state calculation that the HOMO of H_2CO is an oxygen lone pair orbitals and the LUMO the π^* MO. Thus, we want to calculate the singlet and triplet n$\rightarrow \pi^*$ transitions and nothing else. Consequently, we only need to correlate two electrons in two orbitals suggesting a CAS(2,2) reference space.

```
# A simple MRCI example
! def2-SVP def2-SVP/C UseSym

%method frozencore fc_ewin
                 end

%mrci  ewin          -3,1000
       CIType           MRCI
       EUnselOpt     FullMP2
       DavidsonOpt Davidson1
       UseIVOs          true
       tsel             1e-6
       tpre             1e-2
       MaxMemInt         256
       MaxMemVec          32
       IntMode     FullTrafo
       AllSingles       true
       Solver           Diag
       # ground state 1A1
       NewBlock 1 0
         NRoots 1
         Excitations cisd
         Refs CAS(2,2) end
         End
       # HOMO LUMO transition 1A2
       NewBlock 1 1
         NRoots 1
         Excitations cisd
         Refs CAS(2,2) end
         End
       # HOMO LUMO triplet transition 3A2
       NewBlock 3 1
         NRoots 1
         Excitations cisd
         Refs CAS(2,2) end
         end
       end

* int 0 1
 C     0   0   0   0.000000     0.000     0.000
 O     1   0   0   1.200371     0.000     0.000
 H     1   2   0   1.107372   121.941     0.000
 H     1   2   3   1.107372   121.941   180.000
*
```

This input – which is much more than what is really required - needs some explanations: First of all, we choose a standard RHF calculation with the SVP basis set and we assign the SV/C fitting basis although it is not used in the SCF procedure at all. In the `%mrci` block all details of the MR-CI procedure are specified. First, `EWin` (`%method frozencore fc_ewin`) selects the MOs within the given orbital energy range to be included in the correlation treatment. The `CIType` variable selects the type of multireference treatment. Numerous choices are possible and

MRCI is just the one selected for this application.

> **Note**
>
> The CIType statement selects several default values for other variables. So it is a very good idea to place this statement at the beginning of the MR-CI block and possibly overwrite the program selected defaults later. If you place the `CIType` statement after one of the values which it selects by default your input will simply be overwritten!

The variables `EUnselOpt` and `DavidsonOpt` control the corrections to the MR-CI energies. `EUnselOpt` specifies the way in which the MR-CI energies are extrapolated to zero threshold T_{Sel}. Here we choose a full MR-MP2 calculation of the missing contributions to be done *after* the variational step, i.e. using the relaxed part of the reference wavefunction as a 0^{th} order state for MR-PT. The `DavidsonOpt` controls the type of estimate made for the effect of higher substitutions. Again, multiple choices are possible but the most commonly used one (despite some real shortcomings) is certainly the choice `Davidson1`. The flag `UseIVOs` instructs the program to use "improved virtual orbitals". These are virtual orbitals obtained from a diagonalization of the Fock operator from which one electron has been removed in an averaged way from the valence orbitals. Thus, these orbitals "see" only a $N-1$ electron potential (as required) and are not as diffuse as the standard virtual orbitals from Hartree-Fock calculations. If you input DFT orbitals in the MR-CI moldule (which is perfectly admittable and also recommended in some cases, for example for transition metal complexes) then it is recommended to turn that flag off since the DFT orbitals are already o.k. in this respect. The two thresholds `Tsel` and `Tpre` are already explained above and represent the selection criteria for the first order interacting space and the reference space respectively. `Tsel` is given in units of Eh and refers to the second order MR-MP2 energy contribution from a given excited CSF. 10^{-6} Eh is a pretty good value. Reliable results for transition energies start with $\approx 10^{-5}$; however, the total energy is converging pretty slowly with this parameter and this is one of the greatest drawbacks of individually selecting CI procedures! (see below). `Tpre` is dimensionless and refers to the *weight* of a given initial reference after diagonalization of the given initial reference space (10^{-4} is a pretty good value and there is little need to go much lower. Aggressive values such as 10^{-2} only select the truly leading configurations for a given target state which can be time saving. Intermediate values are not really recommended). The parameters `MaxMemInt` and `MaxMemVec` tell the program how much memory (in MB) it is allowed to allocate for integrals and for trial and sigma-vectors respectively.

The flag `IntMode` tells the program to perform a full integral transformation. This is possible for small cases with less than, say, 100–200 MOs. In this case that it is possible it speeds up the calculations considerably. For larger molecules you *have to* set this flag to `RITrafo` which means that integrals are recomputed on the fly using the RI approximation which is more expensive but the only way to do the calculation. To switch between the possible modes use:

```
%mrci  IntMode FullTrafo # exact 4 index transformation
               RITrafo   # use auxiliary basis sets
```

For small molecules or if high accuracy in the total energies is required it is much better to use the exact four index transformation. The limitations are that you will run out of disk space or main memory with more than ca. 200–300 MOs.

The variable `Solver` can be `diag` (for Davidson type diagonalization) or `DIIS` for multirrot DIIS type treatments.

```
%mrci  Solver Diag   # Davidson solver
              DIIS   # Multiroot DIIS
```

For CI methods, the diag solver is usually preferable. For methods like ACPF that contain nonlinear terms, DIIS is imperative.

Next in the input comes the definition of what CI matrices are to be constructed and diagonalized. Each multiplicity defines a *block* of the CI matrix which is separately specified. Here we ask for two blocks – singlet and triplet. The general syntax is:

```
NewBlock Multiplicity Irrep
         NRoots 1           # Number of roots to determine
         Excitations cisd   # Type of excitations
```

(continues on next page)

(continued from previous page)

```
        Refs CAS(NEl,NOrb) end  # Reference space def.
        end                # Finalize the block
```

Now that all input is understood let us look at the outcome of this calculation:

The first thing that happens after the SCF calculation is the preparation of the frozen core Fock matrix and the improved virtual orbitals by the program `orca_ciprep`. From the output the energies of the IVOs can be seen. In this case the LUMO comes down to –8.2 eV which is much more reasonable than the SCF value of +3.... eV. Concomitantly, the shape of this MO will be much more realistic and this important since this orbital is in the reference space!

```
-------------------------------------------------------------------------------
                           ORCA CI-PREPARATION
-------------------------------------------------------------------------------

Reading the GBW file               ... done
Symmetry usage                     ... ON

One-Electron Matrix                ... test1.H.tmp
GBW-File                           ... test1.gbw
Improved virtual orbitals          ... test1.ivo
First MO in the CI                 ... 2
Internal Fock matrix               ... test1.fi.tmp
LastInternal Orbital               ... 6
Integral package used              ... LIBINT
Reading the GBW file               ... done
Symmetry usage                     ... ON

Reading the one-electron matrix    ... done
Forming inactive density           ... done
Forming averaged valence density   ...
Scaling the occupied orbital occupation numbers
First MO                            ... 2
Last MO                             ... 7
Number of electrons in the range    ...    12
Scaling factor                      ... 0.917

done
Forming internal density           ... done
Forming Fock matrix/matrices       ...
Nuclear repulsion    ...    31.371502
Core repulsion       ...    31.371502
One-electron energy  ...  -114.942080
Fock-energy          ...   -94.993431
Final value          ...   -73.596254
done
Modifying virtual orbitals         ...
Last occupied MO               ...    7
Total number of MOs            ...   38
Number of virtual MOs          ...   30
Doing diagonalization with symmetry
The improved virtual eigenvalues:
    0: -0.2955 au   -8.041 eV    2- B2
    1: -0.0701 au   -1.907 eV    6- A1
    2: -0.0176 au   -0.479 eV    3- B1
    3:  0.0064 au    0.175 eV    7- A1
    4:  0.2922 au    7.951 eV    8- A1
    5:  0.2948 au    8.021 eV    3- B2
    6:  0.3836 au   10.439 eV    4- B1
    7:  0.4333 au   11.790 eV    9- A1
    8:  0.4825 au   13.128 eV    5- B1
```

(continues on next page)

(continued from previous page)

```
    9:  0.5027 au   13.680 eV   10- A1
   10:  0.7218 au   19.642 eV   11- A1
   11:  0.8351 au   22.723 eV    4- B2
   12:  0.9371 au   25.501 eV    6- B1
   13:  1.0265 au   27.933 eV    1- A2
   14:  1.1141 au   30.317 eV   12- A1
   15:  1.2869 au   35.017 eV    5- B2
   16:  1.4605 au   39.743 eV    7- B1
...

done
Transforming integrals             ... done
Storing passive energy             ... done (    -73.59625384 Eh)
Transforming internal FI           ... done
             .... done with the Frozen Core Fock matrices
```

The next step is to transform the electron-electron repulsion integrals into the MO basis:

```
-------------------------
SHARK HALF TRANSFORMATION
-------------------------

Number of basis functions               ...     38
Number of operators                     ...      1
   Operator  0:    2-  37

Integral generator used                 ... SHARK
Contraction scheme used                 ... SEGMENTED CONTRACTION
MaxCore in resort                       ...    256 MB

Half transformed integrals for op=   0   ... test1.SHARK_MNPQ0.tmp
Resorted half transformed integrals     ... test1.JAO.tmp
Starting integral generation + half trafo...
Half trafo (segmented) done. Total time =  0.1 sec. integrals=  0.0 sec trafo=  0.0 sec
Starting integral resorting             ... done (  0.0 sec)

SHARK half integral transformation done. Total time =  0.1 sec.


-------------------
FULL TRANSFORMATION
-------------------

Processing MO  10
Processing MO  20
Processing MO  30
Full transformation done
Number of integrals made            ...       222111
Number of integrals stored          ...        59070
Timings:
Time for first half transformation  ...     0.068 sec
Time for second half transformation ...     0.014 sec
Total time                          ...     0.086 sec
```

This will result in a few additional disk files required by `orca_mrci`. The program then tells you which multiplicities will be treated in this MRCI run:

```
------------------
CI-BLOCK STRUCTURE
------------------

```

(continues on next page)

(continued from previous page)

```
Number of CI-blocks                ... 3

===========
CI BLOCK 1
===========
Multiplicity             ...   1
Irrep                    ...   0
Number of reference defs ...   1
  Reference   1: CAS(2,2)

Excitation type          ... CISD
Excitation flags for singles:
    1 1 1 1
Excitation flags for doubles:
    1 1 1 / 1 1 1 / 1 1 1

===========
CI BLOCK 2
===========
Multiplicity             ...   1
Irrep                    ...   1
Number of reference defs ...   1
  Reference   1: CAS(2,2)

Excitation type          ... CISD
Excitation flags for singles:
    1 1 1 1
Excitation flags for doubles:
    1 1 1 / 1 1 1 / 1 1 1

===========
CI BLOCK 3
===========
Multiplicity             ...   3
Irrep                    ...   1
Number of reference defs ...   1
  Reference   1: CAS(2,2)

Excitation type          ... CISD
Excitation flags for singles:
    1 1 1 1
Excitation flags for doubles:
    1 1 1 / 1 1 1 / 1 1 1

        ------------------------------------------------------------------------
        -------------------- ALL SETUP TASKS ACCOMPLISHED --------------------
        --------------------      (     0.139 sec)        --------------------
        ------------------------------------------------------------------------
```

Now that all the setup tasks have been accomplished the MRCI calculation itself begins.

```
#####################################################
#                                                   #
#                       M R C I                     #
#                                                   #
# TSel    =   1.000e-06 Eh                          #
# TPre    =   1.000e-02                             #
# TIntCut =   1.000e-10 Eh                          #
# Extrapolation to unselected MR-CI by full MP2     #
# DAVIDSON-1 Correction to full CI                  #
#                                                   #
```

(continues on next page)

(continued from previous page)

```
####################################################


---------------------
INTEGRAL ORGANIZATION
---------------------

Reading the one-Electron matrix             ... done
E0 read was -73.596253835266
Reading the internal Fock matrix            ... assuming it to be equal to the one-electron␣
↪matrix!!!
done
Preparing the integral list                 ... done
Loading the full integral                   ... done
Making the simple integrals                 ... done

                  ************************************
                  *              CI-BLOCK  1           *
                  ************************************

Configurations with insufficient # of SOMOs WILL be rejected
Building a CAS(2,2) for multiplicity 1 and irrep=A1
Reference Space:
 Initial Number of Configurations :      2
 Internal Orbitals :    2 -    6
 Active Orbitals   :    7 -    8
 External Orbitals :    9 -   37
The number of CSFs in the reference is 2
Calling MRPT_Selection with N(ref)=2
```

In the first step, the reference space is diagonalized. From this CI, the most important configurations are selected with `Tpre`:

```
------------------
REFERENCE SPACE CI
------------------

 Pre-diagonalization threshold                : 1.000e-02
Warning: Setting NGuessMat to 512
N(ref-CFG)=2 N(ref-CSF)=2

                       ****Iteration    0****
   Lowest Energy          :  -113.779221580786
   Maximum Energy change  :   113.779221580786 (vector 0)
   Maximum residual norm  :     0.000000000000

      *** CONVERGENCE OF RESIDUAL NORM REACHED ***
Reference space selection using TPre= 1.00e-02

    ... found 1 reference configurations (1 CSFs)
        ... now redoing the reference space CI ...

Warning: Setting NGuessMat to 512
N(ref-CFG)=1 N(ref-CSF)=1

                       ****Iteration    0****
   Lowest Energy          :  -113.778810013503
   Maximum Energy change  :   113.778810013503 (vector 0)
   Maximum residual norm  :     0.000000000000

      *** CONVERGENCE OF RESIDUAL NORM REACHED ***
```

In this case, the CAS space only has 2 correctly symmetry adapted CSFs one of which (the closed-shell determinant) is selected. In general, larger CAS spaces usually carry around a lot of unnecessary CSFs which are not needed for anything and then the selection is important to reduce the computational effort. The result of the second reference space CI is printed:

```
----------
CI-RESULTS
----------

The threshold for printing is  0.30 percent
The weights of configurations will be printed. The weights are summed over
all CSFs that belong to a given configuration before printing

STATE   0:  Energy=   -113.778810014 Eh RefWeight=  1.0000  0.00 eV      0.0 cm**-1
      1.0000  : h---h---[20]
```

Energy is the total energy in Eh. In the present case we can compare to the SCF energy `-113.778810014 Eh` and find that the reference space CI energy is identical, as it has to be since the lowest state coincides with the reference space. **RefWeight** gives the weight of the reference configurations in a CI state. This is 1.0 in the present case since there were only reference configurations. The number `1.000` is the weight of the following configuration in the CI vector. The description of the configuration **h-h-[20]p-p-** is understood as follows:[2] The occupation of the active orbitals is explicitly given in square brackets. Since the HOMO orbitals is number 7 from the SCF procedure, this refers to MOs 7 and 8 in the present example since we have two active orbitals. The `2` means doubly occupied, the `0` means empty. Any number (instead of `-`) appearing after an `h` gives the index of an internal orbital in which a hole is located. Simarly, any number after a `p` gives the index of an virtual (external) MO where a particle is located. Thus `h-h-[20]` is a closed shell configuration and it coincides with the SCF configuration—this was of course to be expected. The second root (in CI-Block 2) `h-h-[11]` by comparison refers to the configuration in which one electron has been promoted from the HOMO to the LUMO and is therefore the desired state that we wanted to calculate. Things are happy therefore and we can proceed to look at the output.

The next step is the generation of excited configurations and their selection based on `Tsel`:

```
------------------------------
MR-PT SELECTION TSel= 1.00e-06
------------------------------


Setting reference configurations WITH use of symmetry
Building active patterns WITH use of symmetry

Selection will be done from 1 spatial configurations
Selection will make use of spatial symmetry
 ( 0) Refs        : Sel:         1CFGs/         1CSFs Gen:          1CFGs/          1CSFs  (␣
↪      0.000 sec)
Building active space densities            ...        0.002 sec
Building active space Fock operators       ...        0.000 sec
 ( 1) (p,q)->(r,s): Sel:         1CFGs/         1CSFs Gen:          1CFGs/          1CSFs  (␣
↪      0.000 sec)
 ( 2) (i,-)->(p,-): Sel:         1CFGs/         1CSFs Gen:          1CFGs/          1CSFs  (␣
↪      0.000 sec)
 ( 3) (i,j)->(p,q): Sel:         8CFGs/         8CSFs Gen:          8CFGs/          8CSFs  (␣
↪      0.000 sec)
 ( 4) (i,p)->(q,r): Sel:         0CFGs/         0CSFs Gen:          1CFGs/          1CSFs  (␣
↪      0.000 sec)
 ( 5) (p,-)->(a,-): Sel:         8CFGs/         8CSFs Gen:          8CFGs/          8CSFs  (␣
↪      0.000 sec)
 ( 6) (i,-)->(a,-): Sel:        52CFGs/        52CSFs Gen:         52CFGs/         52CSFs  (␣
↪      0.000 sec)
 ( 7) (i,j)->(p,a): Sel:        95CFGs/       166CSFs Gen:         96CFGs/        167CSFs  (␣
```

(continues on next page)

[2] Note that for printing we always sum over all linearly independent spin couplings of a given spatial configuration and only print the summed up weight for the configuration rather than for each individual CSF of the configuration.

(continued from previous page)

```
↪      0.000 sec)
 ( 8) (i,p)->(q,a): Sel:        21CFGs/        42CSFs Gen:          22CFGs/          44CSFs  (␣
↪      0.000 sec)
 ( 9) (p,q)->(r,a): Sel:         3CFGs/         3CSFs Gen:           5CFGs/           5CSFs  (␣
↪      0.000 sec)
 (10) (i,p)->(a,b): Sel:       555CFGs/      1082CSFs Gen:         584CFGs/        1139CSFs  (␣
↪      0.001 sec)
 (11) (p,q)->(a,b): Sel:       124CFGs/       124CSFs Gen:         148CFGs/         148CSFs  (␣
↪      0.000 sec)
 (12) (i,j)->(a,b): Sel:      1688CFGs/      2685CSFs Gen:        1887CFGs/        2947CSFs  (␣
↪      0.001 sec)

Selection results:
Total number of generated configurations:       2814
Number of selected configurations       :       2557 ( 90.9%)
Total number of generated CSFs          :       4522
Number of selected CSFS                 :       4173 ( 92.3%)

The selected tree structure:
Number of selected Internal Portions       :     11
Number of selected Singly External Portions:     27
   average number of VMOs/Portion          :   6.39
   percentage of selected singly externals :  22.83
Number of selected Doubly External Portions:     21
   average number of VMOs/Portion          : 107.59
   percentage of selected doubly externals :  27.76
```

Here, the program loops through classes of excitations. For each excitation it produces the excited configurations (CFGs) and from it the linearly independent spin functions (CSFs) which are possible within the configuration. It then calculates the interaction with the contracted 0^{th} order roots and includes all CSFs belonging to a given CFG in the variational space if the largest second order perturbation energy is larger or equal to `Tsel`. In the present case ≈136,000 CSFs are produced of which 25% are selected. For larger molecules and basis sets it is not uncommon to produce 10^9–10^{10} configurations and then there is no choice but to select a much smaller fraction than 20%. For your enjoyment, the program also prints the total energies of each state after selection:

```
Diagonal second order perturbation results:
State      E(tot)           E(0)+E(1)     E2(sel)      E2(unsel)
             Eh               Eh           Eh            Eh
-------------------------------------------------------------------
 0    -114.108350270   -113.778810014    -0.329433    -0.000107
```

You can ignore this output if you want. In cases that the perturbation procedure is divergent (not that uncommon!) the total energies look strange—don't worry—the following variational calculation is still OK. The second order perturbation energy is here divided into a selected part `E2(sel)` and the part procedure by the unselected configurations `E2(unsel)`. Depending on the mode of `EUnselOpt` this value may already be used later as an estimate of the energetic contribution of the unselected CSFs.[3]

Now we have ≈ 4,200 CSFs in the variational space of CI block 1 and proceed to diagonalize the Hamiltonian over these CSFs using a Davidson or DIIS type procedure:

```
------------------------
DAVIDSON-DIAGONALIZATION
------------------------

Dimension of the eigenvalue problem            ...   4173
Number of roots to be determined               ...      1
Maximum size of the expansion space            ...      4
```

(continues on next page)

[3] In this case the maximum overlap of the 0^{th} order states with the final CI vectors is computed and the perturbation energy is added to the "most similar root". This is of course a rather crude approximation and a better choice is to recomputed the second order energy of the unselected configurations rigorously as is done with `EUnselOpt = FullMP2`.

(continued from previous page)

```
Maximum number of iterations                    ...     35
Convergence tolerance for the residual          ...    1.000e-06
Convergence tolerance for the energies          ...    1.000e-06
Orthogonality tolerance                         ...    1.000e-14
Level Shift                                     ...    0.000e+00
Constructing the preconditioner                 ... o.k.
Building the initial guess                      ... o.k.
Number of trial vectors determined              ...      4


                      ****Iteration    0****
   Size of expansion space: 3
   Lowest Energy          :  -113.937028067251
   Maximum Energy change  :   113.937028067251 (vector 0)
   Maximum residual norm  :     0.741727830968

                      ****Iteration    1****
   Size of expansion space: 4
   Lowest Energy          :  -114.082265676116
   Maximum Energy change  :     0.145237608865 (vector 0)
   Maximum residual norm  :     0.012707561344
   Rebuilding the expansion space

                      ****Iteration    2****
   Size of expansion space: 2
   Lowest Energy          :  -114.085350429118
   Maximum Energy change  :     0.003084753001 (vector 0)
   Maximum residual norm  :     0.002880697397

                      ****Iteration    3****
   Size of expansion space: 3
   Lowest Energy          :  -114.086043274125
   Maximum Energy change  :     0.000692845007 (vector 0)
   Maximum residual norm  :     0.000098595378

                      ****Iteration    4****
   Size of expansion space: 4
   Lowest Energy          :  -114.086074300143
   Maximum Energy change  :     0.000031026018 (vector 0)
   Maximum residual norm  :     0.000004959126
   Rebuilding the expansion space

                      ****Iteration    5****
   Size of expansion space: 2
   Lowest Energy          :  -114.086076038587
   Maximum Energy change  :     0.000001738444 (vector 0)
   Maximum residual norm  :     0.000000572348

      *** CONVERGENCE OF RESIDUAL NORM REACHED ***

Storing the converged CI vectors                ... test1.mrci.vec

                *** DAVIDSON DONE ***
Returned from DIAG section
```

The procedure converges on all roots simultaneously and finishes after six iterations which is reasonable. Now the program calculates the Davidson correction (`DavidsonOpt`) which is printed for each root.

```
Davidson type correction:
Root= 0 W=   0.912 E0=  -113.778810014 ECI=  -114.086076039 DE=-0.026913
```

Already in this small example the correction is pretty large, ca. 27 mEh for the ground state (and $\approx$ 36 mEh for the

excited state, later in the output). Thus, a contribution of ≈ 9 mEh = 0.25 eV is obtained for the transition energy which is certainly significant. Unfortunately, the correction becomes unreliable as the reference space weight drops or the number of correlated electrons becomes large. Here 0.912 and 0.888 are still OK and the system is small enough to expect good results from the Davidson correction.

The next step is to estimate the correction for the unselected configurations:

```
Unselected CSF estimate:
Full relaxed MR-MP2 calculation           ...

Selection will be done from 1 spatial configurations
Selection will make use of spatial symmetry
Selection will make use of spatial symmetry
Selection will make use of spatial symmetry
done
Selected MR-MP2 energies                  ...

 Root=  0  E(unsel)=     -0.000106931
```

In the present case this is below 1 mEh and also very similar for all three states such that it is not important for the transition energy.

```
----------
CI-RESULTS
----------

The threshold for printing is  0.30 percent
The weights of configurations will be printed. The weights are summed over
all CSFs that belong to a given configuration before printing

STATE   0:  Energy=   -114.113096211 Eh RefWeight=  0.9124  0.00 eV      0.0 cm**-1
      0.9124  : h---h---[20]
      0.0114  : h  6h  6[22]
```

The final ground state energy is `-114.113096211` which is an estimate of the full CI energy in this basis set. The leading configuration is still the closed-shell configuration with a weight of $\approx 91\%$. However, a double excitation *outside* the reference space contributes some 1%. This is the excitation MO6,MO6 $\rightarrow$LUMO,LUMO. This indicates that more accurate results are expected once MO6 is also included in the reference space (this is the HOMO-1). The excited state is dominated by the HOMO-LUMO transition (as desired) but a few other single- and double- excitations also show up in the final CI vector.

Now that all CI vectors are known we can order the states according to increasing energy and print (vertical) transition energies:

```
-------------------
TRANSITION ENERGIES
-------------------

The lowest energy is   -114.113096211 Eh

State Mult Irrep Root Block    mEh          eV      1/cm
  0    1    A1    0    0      0.000      0.000       0.0
  1    3    A2    0    2    134.086      3.649   29428.4
  2    1    A2    0    1    148.499      4.041   32591.8
```

This result is already pretty good and the transition energies are within ≈ 0.1 eV of their experimental gas phase values (≈ 3.50 and ≈ 4.00 eV) and may be compared to the CIS values of 3.8 and 4.6 eV which are considerably in error.

In the next step the densities and transition densities are evaluated and the absorption and CD spectra are calculated (in the dipole length formalism) for the spin-allowed transitions together with state dipole moments:

```
-------------------------------------------------------------------------------------------------
↪-----
                    ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
↪-----
     Transition      Energy     Energy  Wavelength fosc(D2)      D2        DX        DY       ␣
↪DZ
                      (eV)      (cm-1)    (nm)                (au**2)    (au)      (au)     ␣
↪(au)
-------------------------------------------------------------------------------------------------
↪-----
  0-1A1 ->  0-1A2   4.040866   32591.8   306.8   0.000000000   0.00000  -0.00000   0.00000   0.
↪00000

-------------------------------------------------------------------------------
                    CD SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength    R        MX        MY        MZ
                      (eV)      (cm-1)    (nm)   (1e40*cgs)   (au)      (au)      (au)
-------------------------------------------------------------------------------
  0-1A1 ->  0-1A2   4.040866   32591.8   306.8   -0.00000  -0.00000  -0.00000   0.59348


-----------------------------------------------------------------------
                           STATE DIPOLE MOMENTS
-----------------------------------------------------------------------
  Root  Block          TX         TY         TZ          |T|
                     (Debye)    (Debye)    (Debye)      (Debye)
-----------------------------------------------------------------------
    0     0          0.00000   -0.00000    2.33244      2.33244
    0     2          0.00000   -0.00000    1.45831      1.45831
    0     1          0.00000   -0.00000    1.58658      1.58658
```

Here the transition is symmetry forbidden and therefore has no oscillator strength. The state dipole moment for the ground state is 2.33 Debye which is somewhat lower than 2.87 Debye from the SCF calculation. Thus, the effect of correlation is to reduce the polarity consistent with the interpretation that the ionicity of the bonds, which is always overestimated by HF theory, is reduced by the correlation. Finally, you also get a detailed population analysis for each generated state density which may be compared to the corresponding SCF analysis in the preceding part of the output.

This concludes the initial example on the use of the MR-CI module. The module leaves several files on disk most of which are not yet needed but in the future will allow more analysis and restart and the like. The `.ivo` file is a standard `.gbw` type file and the orbitals therein can be used for visualization. This is important in order to figure out the identity of the generated IVOs. Perhaps they are not the ones you wanted and then you need to re-run the MR-CI with the IVOs as input, `NoIter` and the IVO feature in the new run turned off! We could use the IVOs as input for a state averaged CASSCF calculation:

```
! moread UseSym KDIIS
%moinp "Test-SYM-MRCI-H2CO.ivo"

%casscf nel    2
        norb   2
        irrep  0,1,1
        mult   1,1,3
        nroots 1,1,1
        end
```

If we based a MR-ACPF calculation on this reference space we will find that the calculated transition energies are slightly poorer than in the MRCI+Q calculation. This is typical of approximate cluster methods that usually require somewhat larger reference spaces for accurate results. A similar result is obtained with SORCI.

```
%mrci  CIType          SORCI
        tsel             1e-6
        tpre             1e-4
        tnat             1e-5
        AllSingles       true
        doNatOrbs        true
        IntMode     FullTrafo
        # ground state 1A1
        NewBlock 1 0
          NRoots 1
          Excitations cisd
          Refs CAS(2,2) end
          End
        # HOMO LUMO transition 1A2
        NewBlock 1 1
          NRoots 1
          Excitations cisd
          Refs CAS(2,2) end
          End
        # HOMO LUMO triplet transition 3A2
        NewBlock 3 1
          NRoots 1
          Excitations cisd
          Refs CAS(2,2) end
          End
end
```

This gives:

```
State Mult Irrep Root Block    mEh          eV      1/cm
  0    1    A1    0    0      0.000       0.000       0.0
  1    3    A2    0    2    144.563       3.934   31728.0
  2    1    A2    0    1    161.179       4.386   35374.7
```

This is systematically 0.4 eV too high. But let us look at the approximate average natural orbital (AANOs) occupation numbers:

```
------------------------
AVERAGE NATURAL ORBITALS
------------------------

Trace of the density to be diagonalized = 12.000000
Sum of eigenvalues = 12.000000
Natural Orbital Occupation Numbers:
N[  2] ( A1)=   1.99831062
N[  3] ( A1)=   1.99761604
N[  4] ( A1)=   1.99479313
N[  5] ( B1)=   1.99016881
N[  6] ( B2)=   1.95818285
N[  7] ( B1)=   1.33014178
N[  8] ( B2)=   0.70688423
N[  9] ( B1)=   0.00988561
N[ 10] ( A1)=   0.00436843
```

This shows that there is a low-occupancy orbital (MO6) that has not been part of the reference space. Thus, we try the same calculation again but now with one more active orbital and two more active electrons:

```
! moread
%moinp "Test-SYM-MRCI-H2CO.gbw"

%casscf nel 4
        norb 3
```

(continues on next page)

(continued from previous page)

```
        irrep  0,1,1
        mult   1,1,3
        nroots 1,1,1
        end

%mrci CIType          SORCI
      tsel             1e-6
      tpre             1e-4
      tnat             1e-5
      AllSingles       true
      doNatOrbs        true
      IntMode     FullTrafo
      # ground state 1A1
      NewBlock 1 0
        NRoots 1
        Excitations cisd
        Refs CAS(4,3) end
        End
      # HOMO LUMO transition 1A2
      NewBlock 1 1
        NRoots 1
        Excitations cisd
        Refs CAS(4,3) end
        End
      # HOMO LUMO triplet transition 3A2
      NewBlock 3 1
        NRoots 1
        Excitations cisd
        Refs CAS(4,3) end
        End
end
```

This gives:

```
State Mult Irrep Root Block    mEh          eV      1/cm
  0    1    A1    0    0      0.000      0.000       0.0
  1    3    A2    0    2    145.494      3.959   31932.3
  2    1    A2    0    1    162.222      4.414   35603.6
```

Which is now fine since all essential physics has been in the reference space. Inspection of the occupation numbers show that there is no suspicious orbital any more. Note that this is still a much more compact calculation that the MRCI+Q.

Likewise, we get an accurate result from MRACPF with the extended reference space.

```
State Mult Irrep Root Block    mEh          eV      1/cm
  0    1    A1    0    0      0.000      0.000       0.0
  1    3    A2    0    2    134.985      3.673   29625.8
  2    1    A2    0    1    148.330      4.036   32554.6
```

However, the SORCI calculation is *much* more compact. For larger molecules the difference becomes more and more pronounced and SORCI or even MRDDCI2 (with or without +Q) maybe the only feasible methods—if at all.

6.7.3 Excitation Energies between Different Multiplicities

As an example for a relatively accurate MRCI+Q calculation consider the following job which calculates the triplet-ground and as the first excited singlet states of O_2.

```
! ano-pVQZ RI-AO cc-pVQZ/JK VeryTightSCF NoPop Conv UseSym RI-MP2 PModel
%mp2 density relaxed natorbs true end
%base "O2"
* xyz 0 3
O  0 0 0
O  0 0 1.2
*

$new_job
! ano-pVQZ RI-AO cc-pVQZ/JK VeryTightSCF NoPop Conv UseSym KDIIS
! moread
%moinp "O2.mp2nat"
%casscf nel     8
        norb    6
        irrep   1,0,1
        nroots  1,2,1
        mult    3,1,1
        trafostep    ri
        switchstep   nr
        end

%mrci   citype mrci
        tsel 1e-7
        tpre 1e-5
        newblock 3 1 nroots 1 refs cas(8,6) end end
        newblock 1 0 nroots 2 refs cas(8,6) end end
        newblock 1 1 nroots 1 refs cas(8,6) end end
        end

* xyz 0 3
O  0 0 0
O  0 0 1.2
*
```

Note that the linear molecule is run in D_{2h}. This creates a slight problem as the CASSCF procedure necessarily breaks the symmetry of the $^1\Delta$ state.

```
LOWEST ROOT (ROOT 0, MULT 3, IRREP B1g) =   -149.765383866 Eh -4075.323 eV

STATE   ROOT MULT IRREP DE/a.u.    DE/eV    DE/cm**-1
   1:    0    1  B1g   0.033334     0.907   7316.0
   2:    0    1   Ag   0.033650     0.916   7385.3
   3:    1    1   Ag   0.062381     1.697  13691.1
```

The result of the MRCI+Q is:

```
-------------------
TRANSITION ENERGIES
-------------------

The lowest energy is   -150.176905551 Eh

State Mult Irrep Root Block    mEh          eV      1/cm
  0    3   B1g    0    0      0.000      0.000       0.0
  1    1   B1g    0    2     36.971      1.006    8114.2
  2    1    Ag    0    1     38.021      1.035    8344.7
  3    1    Ag    1    1     62.765      1.708   13775.2
```

These excitation energies are accurate to within a few hundred wavenumbers. Note that the ≈ 200 wavenumber splitting in the degenerate $^1\Delta$ state is due to the symmetry breaking of the CAS and the individual selection. Repeating the calculation with the MP2 natural orbitals gives an almost indistinguishable result and a ground state energy that is even lower than what was found with the CASSCF orbitals. Thus, such natural orbitals (that might often be easier to get) are a good substitute for CASSCF orbitals and at the same time the symmetry breaking due to the use of symmetry appears to be difficult to avoid.

```
-------------------
TRANSITION ENERGIES
-------------------

The lowest energy is   -150.177743426 Eh

State Mult Irrep Root Block    mEh          eV       1/cm
  0    3   B1g    0    0       0.000      0.000       0.0
  1    1   B1g    0    2      37.369      1.017    8201.5
  2    1    Ag    0    1      38.237      1.040    8392.1
  3    1    Ag    1    1      62.731      1.707   13767.9
```

6.7.4 Correlation Energies

The logic we are following here is the following: CID minus SCF gives the effect of the doubles; going to CISD gives the effect of the singles; QCISD(=CCD) minus CID gives the effect of the disconnected quadruples. QCISD minus QCID gives simultaneously the effect of the singles and the disconnected triples. They are a bit difficult to separate but if one looks at the singles alone and compares with singles + disconnected triples, a fair estimate is probably obtained. Finally, QCISD(T) minus QCISD gives the effect of the connected triples. One could of course also use CCSD instead of QCISD but I felt that the higher powers of T_1 obscure the picture a little bit—but this is open to discussion of course.

First H_2O/TZVPP at its MP2/TZVPP equilibrium geometry ($T_{\text{pre}} = 10^{-6}$ and $T_{\text{sel}} = 10^{-9}$ Eh for the MRCI and MRACPF calculations):

Excitation class	Energy (Eh)	Delta-Energy (mEh)
None (RHF)	-76.0624	
Doubles (CID)	-76.3174	255
+Singles (CISD)	-76.3186	1
+Disconnected Quadruples (QCID)	-76.3282	11
+Disconnected Triples (QCISD)	-76.3298	2
+Connected Triples (QCISD(T))	**-76.3372**	7
CASSCF(8,6)	-76.1160	
CASSCF(8,6) + MRCI	-76.3264	210
CASSCF(8,6) + MRCI+Q	**-76.3359**	10
CASSCF(8,6) + MRACPF	**-76.3341**	218

One observes quite good agreement between single- and multireference approaches. In particular, the contribution of the disconnected triples and singles is very small. The estimate for the disconnected quadruples is fairly good from either the multireference Davidson correction or the ACPF and the agreement between CCSD(T) and these MR methods is 2-3 mEh in the total energy which is roughly within chemical accuracy.

In order to also have an open-shell molecule let us look at NH with a N-H distance of 1.0 Å using the TZVPP basis set.

Excitation class	Energy (Eh)	Delta-Energy (mEh)
None (UHF)	-54.9835	
Doubles (CID)	-55.1333	150
+Singles (CISD)	-55.1344	1
+Disconnected Quadruples (QCID)	-55.1366	3
+Disconnected Triples (QCISD)	-55.1378	1
+Connected Triples (QCISD(T))	**-55.1414**	4
CASSCF(6,5)	-55.0004	
CASSCF(6,5) + MRCI	-55.1373	137
CASSCF(6,5) + MRCI+Q	**-55.1429**	6
CASSCF(6,5) + MRACPF	**-55.1413**	141

Again, the agreement is fairly good and show that both single- and multiple reference approaches converge to the same limit.

6.7.5 Thresholds

Now we choose the CO molecule (1.128 Ångström) with the SVP basis set and study the convergence of the results with respect to the selection threshold. Comparison to high level single-reference approaches is feasible (The SCF energy is `-112.645 946 Eh`).

Reference Values for Total Energies

The single-reference values are:

```
BD: -112.938 48002
              CCSD: -112.939 79145
             QCISD: -112.941 95700
              BD(T): -112.950 17278
           CCSD(T): -112.950 63889
          QCISD(T): -112.951 37425
         MP4(SDTQ): -112.954 80113
```

The calculations without connected triples (BD, CCSD, QCISD) are about the best what can be achieved without explicitly considering triple excitations. The CCSD is probably the best in this class. As soon as connected triples are included the CCSD(T), QCISD(T) and BD(T) values are close and from experience they are also close to the full CI values which is then expected somewhere between –112.950 and –112.952 Eh.

Convergence of Single Reference Approaches with Respect to T_{sel}

Next it is studied how these single reference methods converge with T_{sel}:

```
Closed-Shell ACPF:
   Tsel        Energy      (NCSF)    Energy       (NCSF)
   (Eh)       AllSingles=true      AllSingles=false
  TSel=0      -112.943 387 (5671)
  TSel=1e-14  -112.943 387 (2543)  -112.943 387 (2478)
  TSel=1e-10  -112.943 387 (2543)  -112.941 023 (2453)
  TSel=1e-08  -112.943 387 (2451)  -112.937 087 (2346)
  TSel=1e-06  -112.943 350 (2283)  -112.937 046 (2178)
  TSel=1e-05  -112.943 176 (1660)  -112.936 821 (1555)
  TSel=1e-04  -112.944 039 ( 782)  -112.938 381 ( 677)
```

It is clear that the convergence is erratic if the singles are not automatically included. This is the reason for making this the default from release 2.6.35 on. In the present case singles will only be selected due to round-off errors since by Brillouin's theorem the singles have zero-interaction with the ground state determinant. Thus, for individually selecting single-reference methods it is a good idea to automatically include all single-excitations in order to get converged results. The alternative would be a different singles selection procedure which has not yet been developed however. The selection of doubles appear to converge the total energies reasonably well. It is seen that the selection selects most CSFs between 10^{-5} and 10^{-7} Eh. Already a threshold of 10^{-6} Eh yields an error of less than 0.1 mEh which is negligible in relation to reaction energies and the like. Even 10^{-5} Eh gives an error of less than 0.1 kcal/mol.

Convergence of Multireference Approaches with Respect to T_{pre}

We next turn to multireference treatments. Here we want to correlate all valence electrons in all valence orbitals and therefore a CAS(10,8) is the appropriate choice. We first ask for the converged value of T_{pre} by using T_{sel} $=10^{-14}$ and obtain for MRCI+Q:

```
TPre = 1e-1: -112.943 964
         1e-2: -112.952 963
         1e-3: -112.953 786
         1e-4: -112.954 019
         1e-5: -112.954 336
         1e-6: -112.954 416
         1e-7: -112.954 440
```

Thus, pretty good convergence is obtained for $T_{pre} = 10^{-4} - 10^{-6}$. Hence 10^{-4} is the default.

To show a convenient input consider the following:

```
\
#
# Here we calculate the CO ground state correlation energy with several methods
#
! Def2-SVP Def2-SV/C RI-MP2 CCSD(T)
%base "1"

%mp2  density relaxed
      donatorbs  true
      end

* int 0 1
 C     0   0   0   0.000000     0.000     0.000
 O     1   0   0   1.128        0.000     0.000
*

$new_job

! aug-SVP MRACPF
! moread
%moinp "1.mp2nat"
# the CASSCF is done with MP2 natural orbitals which is a good idea and
# secondly we use a large level shift in order to help convergence
%casscf nel     10
        norb     8
        mult     1
        nroots   1
        shiftup  2
        shiftdn  2
        end

%mrci   tsel 1e-8
        tpre 1e-6
```

(continues on next page)

(continued from previous page)

```
        end

* int 0 1
 C     0   0   0   0.000  0.000  0.000
 O     1   0   0   1.128  0.000  0.000
*
```

This job computes at the same time all of the below and demonstrates once more the agreement between consequent single- and multireference correlation methods

```
SCF          =  -112.6459
RI-MP2       =  -112.9330
CCSD         =  -112.9398
CCSD(T)      =  -112.9506
CASSCF(10,8) =  -112.7769
MRACPF       =  -112.9514
```

6.7.6 Energy Differences - Bond Breaking

For the calculation of energy differences we start again with the reference CCSD(T) calculation; this method is one of the few which can claim chemical accuracy in practical applications:

```
Reference Total Energies for N2 at 1.0977 Angstr\"{o}m with
  The SVP basis
    E(CCSD)   = -109.163 497
    E(CCSD(T))= -109.175 625
  Nitrogen Atom (4S), SVP basis, unrestricted
    E(CCSD)   =  -54.421 004
    E(CCSD(T))=  -54.421 7183
  Energy Difference:
  Delta-E(CCSD)   = -0.321 489 = 8.75 eV
  Delta-E(CCSD(T))= -0.332 188 = 9.04 eV
```

The basis set is of course not suitable for quantitative comparison to experimental values. However, this is not the point here in these calculations which are illustrative in nature. The SVP basis is just good enough to allow for a method assessment without leading to excessively expensive calculations.

This is now to be compared with the corresponding energy differences computed with some single-reference approaches. A typical input is (this is a somewhat old-fashioned example – in the present program version you would do a full valence CASSCF(10,8) or CASSCF(6,6) and invoke the MR-methods with a single keyword):

```
! HF def2-SVP def2-TZVPP/C VeryTightSCF NoPop

%base "1"

* xyz 0 1
N 0 0 0
N 0 0 1.0977
*
%method
     frozencore fc_ewin
end

%mrci
     EWin          -3,1000
     CIType       MRACPF2a
     Solver           DIIS
     IntMode      FullTrafo
     UseIVOs          true
```

(continues on next page)

(continued from previous page)

```
     AllSingles       true
     TSel            1e-14
     TPre            1e-05
     TNat              0.0
     ETol            1e-10
     RTol            1e-10
     NewBlock 1 *
          NRoots 1
          Excitations CISD
          refs CAS(0,0) end
     end
end

$new_job

! ROHF def2-SVP def2-TZVPP/C VeryTightSCF NoPop PModel

%base "2"

* xyz 0 4
N 0 0 0
*
%method
       frozencore fc_ewin
end

%mrci
      EWin          -3,1000
      CIType       MRACPF2a
      IntMode      FullTrafo
      UseIVOs          true
      AllSingles       true
      TSel            1e-14
      TPre            1e-05
      TNat              0.0
      ETol            1e-10
      RTol            1e-10
      NewBlock 4 *
          NRoots 1
          Excitations CISD
          refs CAS(3,3) end
      end
end
```

The results are:

```
Single reference approaches:
  Method   N2-Molecule         N-Atom        Delta-E
  CISD+Q : -109.167 904        -54.422 769    8.77 eV
  ACPF   : -109.166 926        -54.421 783    8.80 eV
  ACPF2  : -109.166 751        -54.421 333    8.82 eV
  ACPF2a : -109.166 730        -54.421 186    8.83 eV
  CEPA1  : -109.159 721        -54.422 564    8.56 eV
  CEPA2  : -109.172 888        -54.422 732    8.91 eV
  CEPA3  : -109.161 034        -54.422 589    8.59 eV
  AQCC   : -109.160 574        -54.420 948    8.67 eV
  CEPA-0 : -109.174 924        -54.422 951    8.95 eV
```

With exception is CEPA1 and CEPA3, the results are OK. The reason for the poor performance of these methods is simply that the formalism implemented is only correct for closed shells – open shells require a different formalism which we do not have available in the MRCI module (but in the single reference MDCI module). Due to the simple approximations made in CEPA2 it should also be valid for open shells and the numerical results are in support of

that.

Next we turn to the multireference methods and take a CAS(10,8) reference as for CO in order to correlate all valence electrons.[4]

```
Multi reference approaches:
  Method     N2-Molecule         N-Atom        Delta-E
  MRCISD+Q: -109.180 089        -54.422 667    9.11 eV
  MRACPF  : -109.178 708        -54.421 685    9.12 eV
  MRACPF2 : -109.177 140        -54.421 236    9.11 eV
  MRAQCC  : -109.175 947        -54.420 851    9.10 eV
  SORCI   : -109.179 101        -54.422 703    9.08 eV
```

This test calculation pleasingly shows the high consistency of multireference approaches which all converge more or less to the same result which must be accurate.

6.7.7 Energy Differences - Spin Flipping

There are a number of interesting situations in which one is interested in a small energy difference which arises from two states of different multiplicity but same orbital configuration. This is the phenomenon met in diradicals or in magnetic coupling in transition metal complexes. As a primitive model for such cases one may consider the hypothetical molecule H-Ne-H in a linear configuration which will be used as a model in this section.

The reference value is obtained by a MR-ACPF calculation with all valence electrons active (again, this example is somewhat old fashioned – in the present program version you would do a CASSCF calculation followed by MR methods with a single keyword):

```
! ROHF def2-SVP def2-TZVPP/C VeryTightSCF NoPop
%basis
     NewAuxCGTO Ne "AutoAux" end
end
* xyz 0 3
H  0 0 0
Ne 0 0 2.0
H  0 0 4.0
*
%method frozencore fc_ewin
        end

%mrci  EWin       -3,1000
       CIType     MRACPF2a
       IntMode   FullTrafo
       Solver         DIIS
       UseIVOs        true
       TSel              0
       TPre          1e-10
       ETol          1e-09
       RTol          1e-09
       DoDDCIMP2      true
       NewBlock 1 *
          NRoots 1
          Excitations CISD
          refs CAS(10,6) end
          end
       NewBlock 3 *
          NRoots 1
          Excitations CISD
          refs CAS(10,6) end
```

(continues on next page)

[4] Most of these results have been obtained with a slightly earlier version for which the MR energies are a little different from that what the present version gives. The energy differences will not be affected.

(continued from previous page)

```
         end
      end
```

which gives the reference value 108 cm^{-1}. We now compare that to several other methods which only have the two "magnetic" orbitals (the 1s's on the hydrogens) in the active space:

```
... same as above
%mrci  EWin       -10,1000
       CIType      MRDDCI3
       ... same as previously
       NewBlock 1 *
          NRoots 1
          refs CAS(2,2) end
          end
       NewBlock 3 *
          NRoots 1
          refs CAS(2,2) end
          end
       end
```

This gives the result:

```
Method      S-T gap
MR-CI+Q  :  98 cm-1
MR-CI    :  93 cm-1
MR-ACPF  :  98 cm-1
MR-ACPF2 :  98 cm-1
MR-ACPF2a:  97 cm-1
MR-AQCC  :  95 cm-1
SORCI    : 131 cm-1
MR-DDCI2 :  85 cm-1
MR-DDCI3 : 130 cm-1
```

All these methods give good results with SORCI leading to a somewhat larger error than the others. The (difference dedicated CI) DDCI2 method slightly underestimates the coupling which is characteristic of this method. It is nice in a way that DDCI3 gives the same result as SORCI since SORCI is supposed to approximate the DDCI3 (or better the IDDCI3) result which it obviously does.

This splitting can also be studied using broken symmetry HF and DFT methods as explained elsewhere in this manual:

```
Method    S-T gap
 UHF     : 70 cm-1
 B3LYP/G : 240 cm-1
 BP86    : 354 cm-1
 PW91    : 234 cm-1
 PBE     : 234 cm-1
 PBE0    : 162 cm-1
 RPBE    : 242 cm-1
```

This confirms the usual notions; UHF underestimates the coupling and DFT overestimates it, less so for hybrid functionals than for GGAs. The BP86 is worse than PW91 or PBE. The PBE0 hybrid may be the best of the DFT methods. For some reason most of the DFT methods give the best results if the BS state is simply taken as an approximation for the true open-shell singlet. This is, in our opinion, not backed up by theory but has been observed by other authors too.

Now let us study the dependence on T_{sel} as this is supposed to be critical (we use the DDCI3 method):

```
Tsel     S-T gap
    1e-04     121
    1e-05     128
```

(continues on next page)

(continued from previous page)

```
    1e-06     132
    1e-07     131
    1e-08     131
    1e-10     131
    1e-12     131
    0         131
```

The convergence is excellent once `AllSingles` are included.

6.7.8 Potential Energy Surfaces

Another situation where multireference approaches are necessary is when bond breaking is studied and one wants to calculate a full potential energy surface. Say we want to compute the potential energy surface of the CH molecule. First we have to figure out which states to include. Hence, let us first determine a significant number of roots for the full valence CASSCF reference state (we use a small basis set in order to make the job fast).

```
! ANO-pVDZ VeryTightSCF NoPop Conv

%casscf nel      5
        norb     5
        nroots   2
        mult     2
        end

%mrci  CIType      MRCI
       NewBlock 2 *
         excitations none
         NRoots 15
         refs CAS(5,5) end
         end
       NewBlock 4 *
         excitations none
         NRoots 15
         refs CAS(5,5) end
         end
       end

* xyz 0 2
C  0 0 0
H  0 0 1.15
*
```

This yields:

```
-------------------
TRANSITION ENERGIES
-------------------

The lowest energy is    -38.308119994 Eh

State Mult Irrep Root Block    mEh          eV      1/cm
  0    2    -1    0    0      0.000      0.000       0.0
  1    2    -1    1    0      0.000      0.000       0.0
  2    4    -1    0    1     14.679      0.399    3221.6
  3    2    -1    2    0    126.464      3.441   27755.7
  4    2    -1    3    0    126.464      3.441   27755.7
  5    2    -1    4    0    132.689      3.611   29121.8
  6    2    -1    5    0    164.261      4.470   36051.2
  7    2    -1    6    0    305.087      8.302   66958.9
  8    2    -1    7    0    305.087      8.302   66958.9
```

(continues on next page)

(continued from previous page)

```
  9    4    -1    1    1    328.911      8.950   72187.7
 10    4    -1    2    1    452.676     12.318   99350.8
 11    4    -1    3    1    452.676     12.318   99350.8
 12    2    -1    8    0    460.116     12.520  100983.9
 13    2    -1    9    0    463.438     12.611  101712.9
 14    2    -1   10    0    463.438     12.611  101712.9
 ...
```

Thus, if we want to focus on the low-lying states we should include five doublet and one quartet root. Now we run a second job with these roots and scan the internuclear distance.

```
! ano-pVDZ VeryTightSCF NoPop Conv MRCI+Q

%casscf nel     5
        norb    5
        nroots  5,1
        mult    2,4
        shiftup 2
        end

%paras R = 0.8,2.5,25
       end

* xyz 0 2
C  0 0 0
H  0 0 {R}
*
```

The surfaces obtained in this run are shown in Fig. 6.30. You can nicely see the crossing of the $^2\Sigma$ and $^2\Delta$ states fairly close to the equilibrium distance and also the merging of the $^4\Sigma$ state with $^2\Pi$ and $^2\Sigma$ towards the asymptote that where C-H dissociates in a neutral C-atom in its ^{3}P ground state and a neutral hydrogen atom in its ^{2}S ground state. You can observe that once `AllSingles` is set to true (the default), the default settings of the MRCI module yield fairly smooth potential energy surfaces.

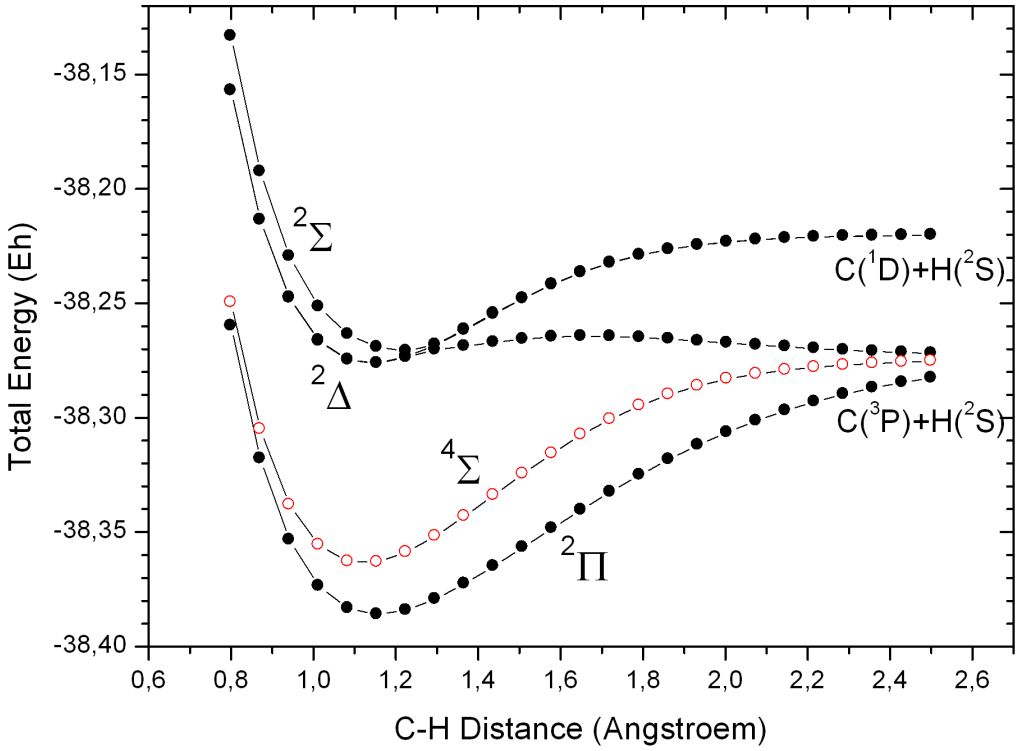


Fig. 6.30: Potential energy surfaces for some low-lying states of CH using the MRCI+Q method

In many cases one will focus on the region around the minimum where the surface is nearly quadratic. In this case one can still perform a few (2, 3, 5,) point polynomial fitting from which the important parameters can be determined. The numerical accuracy and the behavior with respect to T_{sel} has to be studied in these cases since the selection produces some noise in the procedure. We illustrate this with a calculation on the HF molecule:

```
! ano-pVDZ VeryTightSCF NoPop Conv MRCI+Q

%paras  R = 0.85,1.1,7
        end

%casscf nel     8
        norb    5
        nroots  1  mult    1
        shiftup 2.5 shiftdn 2.5 switchstep nr gtol    1e-5
        end

%mrci   tsel 1e-8
        tpre 1e-5
        end

* xyz 0 1
F  0 0 0
H  0 0 {R}
*
```

The output contains the result of a Morse fit:

```
Morse-Fit Results:
 Re               =   0.93014 Angstroem
 we               =   4111.2 cm**-1
 wexe             =     79.5 cm**-1
```

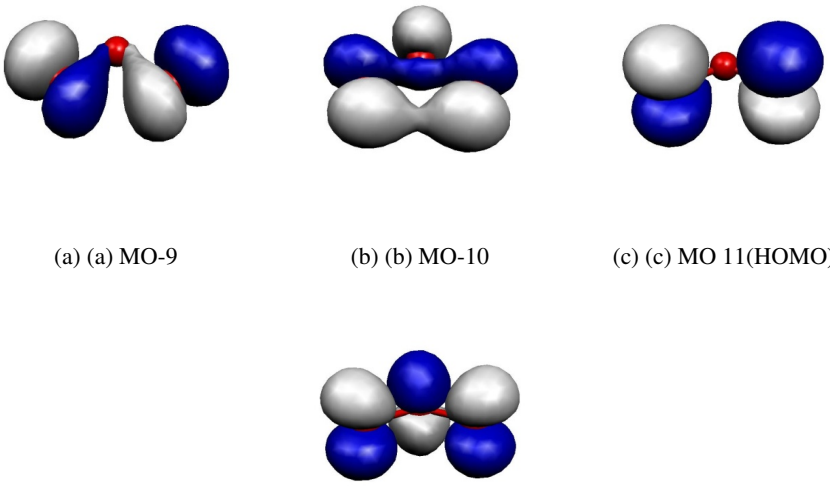
Which may be compared with the CCSD(T) values calculated with the same basis set:

```
Morse-Fit Results:
 Re               =   0.92246 Angstroem
 we               =   4209.8 cm**-1
 wexe             =     97.6 cm**-1
```

The agreement between MRCI+Q and CCSD(T) results is fairly good.

6.7.9 Multireference Systems - Ozone

The ozone molecule is a rather classical multireference system due to its diradical character. Let us look at the three highest occupied and lowest unoccupied MO (the next occupied MO is some 6 eV lower in energy and the next virtual MO some 10 eV higher in energy):



(a) (a) MO-9 (b) (b) MO-10 (c) (c) MO 11(HOMO)

(d) (d) MO 12(LUMO)

Fig. 6.31: Frontier MOs of the Ozone Molecule.

These MOs are two σ lone pairs which are high in energy and then the symmetric and antisymmetric combinations of the oxygen π lone pairs. In particular, the LUMO is low lying and will lead to strong correlation effects since the $(\text{HOMO})^2 \rightarrow (\text{LUMO})^2$ excitation will show up with a large coefficient. Physically speaking this is testimony of the large diradical character of this molecule which is roughly represented by the structure ↑O-O-O↓. Thus, the minimal active space to treat this molecule correctly is a CAS(2,2) space which includes the HOMO and the LUMO. We illustrate the calculation by looking at the RHF, MP2 MRACPF calculations of the two-dimensional potential energy surface along the O–O bond distance and the O-O-O angle (experimental values are 1.2717 Å and 116.78°).

```
! ano-pVDZ VeryTightSCF NoPop MRCI+Q Conv

%paras R     = 1.20,1.40,21
```

(continues on next page)

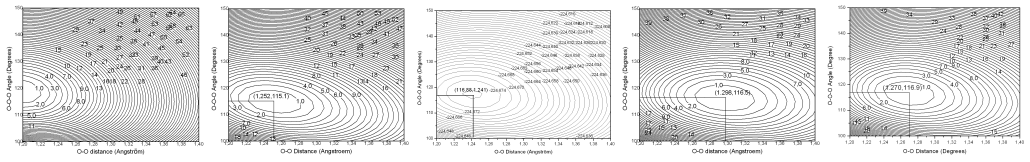
(continued from previous page)

```
       Theta = 100,150,21
       end

%casscf nel     2
        norb    2
        mult    1
        nroots  1
        end

%mrci   tsel    1e-8
        tpre    1e-5
        end

* int 0 1
O  0 0 0  0   0         0
O  1 0 0 {R}  0         0
O  1 2 0 {R}  {Theta}   0
*
```

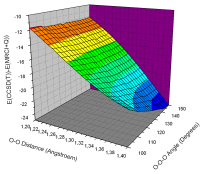
This is a slightly lengthy calculation due to the 441 energy evaluations required. RHF does not find any meaningful minimum within the range of examined geometries. MP2 is much better and comes close to the desired minimum but underestimates the O–O distance by some 0.03 Å. CCSD(T) gives a very good angle but a O–O distance that is too long. In fact, the largest doubles amplitude is ≈ 0.2 in these calculations (the HOMO–LUMO double excitation) which indicates a near degeneracy calculation that even CCSD(T) has problems to deal with. Already the CAS(2,2) calculation is in qualitative agreement with experiment and the MRCI+Q calculation then gives almost perfect agreement.

The difference between the CCSD(T) and MRCI+Q surfaces shows that the CCSD(T) is a bit lower than the MRCI+Q one suggesting that it treats more correlation. However, CCSD(T) does it in an unbalanced way. The MRCI calculation employs single and double excitations on top of the HOMO-LUMO double excitation, which results in triples and quadruples that apparently play an important role in balancing the MR calculation. These excitations are treated to all orders explicitly in the MRCI calculation but only approximately (quadruples as simultaneous pair excitations and triples perturbatively) in the coupled-cluster approach. Thus, despite the considerable robustness of CC theory in electronically difficult situations it is not applicable to genuine multireference problems.

This is a nice result despite the too small basis set used and shows how important it can be to go to a multireference treatment with a physically reasonable active space (even if is only 2×2) in order to get qualitatively and quantitatively correct results.



(a) (a) RHF (b) (b) CASSCF(2,2) (c) (c) MP2 (d) (d) CCSD(T) (e) (e) MRCI+Q



(f) (f) Difference CCSD(T)/MRCI+Q

Fig. 6.32: 2D potential energy surface for the O_3 molecule calculated with different methods

6.7.10 Size Consistency

Finally, we want to study the size consistency errors of the methods. For this we study two non-interacting HF molecules at the single reference level and compare to the energy of a single HF molecule. This should give a reasonably fair idea of the typical performance of each method (energies in Eh)[5]:

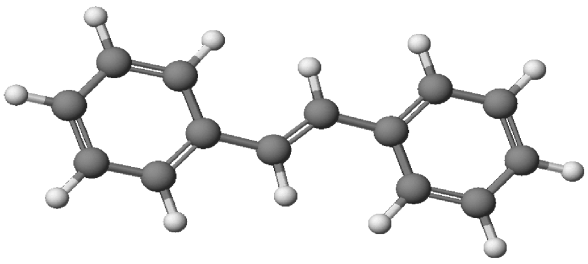
```
E(HF)          E(HF+HF)   |Difference|
  CISD+Q    -100.138 475   -200.273 599      0.00335
  ACPF      -100.137 050   -200.274 010      0.00000
  ACPF2     -100.136 913   -200.273 823      0.00000
  AQCC      -100.135 059   -200.269 792      0.00032
```

The results are roughly as expected – CISD+Q has a relatively large error, ACPF and ACPF/2 are perfect for this type of example; AQCC is not expected to be size consistent and is (only) about a factor of 10 better than CISD+Q in this respect. CEPA-0 is also size consistent.

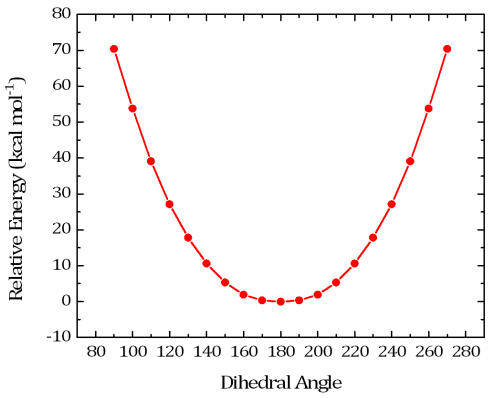
6.7.11 Efficient MR-MP2 Calculations for Larger Molecules

Uncontracted MR-MP2 approaches are nowadays outdated. They are much more expensive than internally contracted e.g. the NEVPT2 method described in section *N-Electron Valence State Pertubation Theory*. Moreover, MR-MP2 is prone to intruder states, which is a major obstacle for practical applications. For historical reasons, this section is dedicated to the traditional MR-MP2 approach that is available since version 2.7.0 ORCA. The implementation avoids the full integral transformation for MR-MP2 which leads to significant savings in terms of time and memory. Thus, relatively large RI-MR-MP2 calculations can be done with fairly high efficiency. However, the program still uses an uncontracted first order wavefunction which means that for very large reference space, the calculations still become untractable.

Consider for example the rotation of the stilbene molecule around the central double bond



(a) (a)



(b) (b)

Fig. 6.33: Rotation of stilbene around the central double bond using a CASSCF(2,2) reference and correlating the reference with MR-MP2.

The input for this calculation is shown below. The calculation has more than 500 basis functions and still runs through in less than one hour per step (CASSCF-MR-MP2). The program takes care of the reduced number of two-electron integrals relative to the parent MRCI method and hence can be applied to larger molecules as well. Note that we have taken a "JK" fitting basis in order to fit the Coulomb and the dynamic correlation contributions both with sufficient accuracy. Thus, this example demonstrates that MR-MP2 calculations for not too large reference spaces can be done efficiently with ORCA (as a minor detail note that the calculations were started at a dihedral angle of 90 degrees in order to make sure that the correct two orbitals are in the active space, namely the central carbon p-orbitals that would make up the pi-bond in the coplanar structure).

[5] Most of these numbers were obtained with a slightly older version but will not change too much in the present version.

```
#
# Stilbene rotation using MRMP2
#
! def2-TZVP def2/JK RIJCOSX RI-MRMP2

%casscf nel            2
        norb           2
        end

%mrci   maxmemint 2000
        tsel 1e-8
        end


%paras  DIHED = 90,270, 19
        end

* int 0 1
 C     0   0   0   0.000000     0.000     0.000
 C     1   0   0   1.343827     0.000     0.000
 C     2   1   0   1.490606   125.126     0.000
 C     1   2   3   1.489535   125.829   \{DIHED\}
 C     4   1   2   1.400473   118.696   180.000
 C     4   1   2   1.400488   122.999     0.000
 C     6   4   1   1.395945   120.752   180.000
 C     5   4   1   1.394580   121.061   180.000
 C     8   5   4   1.392286   120.004     0.000
 C     3   2   1   1.400587   118.959   180.000
 C     3   2   1   1.401106   122.779     0.000
 C    11   3   2   1.395422   120.840   180.001
 C    12  11   3   1.392546   120.181     0.000
 C    13  12  11   1.392464   119.663     0.000
 H     1   2   3   1.099419   118.266     0.000
 H     2   1   3   1.100264   118.477   179.999
 H     5   4   1   1.102119   119.965     0.000
 H     6   4   1   1.100393   121.065     0.000
 H     7   6   4   1.102835   119.956   180.000
 H     8   5   4   1.102774   119.989   180.000
 H     9   8   5   1.102847   120.145   180.000
 H    10   3   2   1.102271   120.003     0.000
 H    11   3   2   1.100185   121.130     0.000
 H    12  11   3   1.103001   119.889   180.000
 H    13  12  11   1.102704   120.113   180.000
 H    14  13  12   1.102746   119.941   180.000
*
```

6.7.12 Keywords

Here is a reasonably complete list of Keywords and their meaning. Note that the MRCI pogram is considered legacy and we can neither guarantuee that the keywords still work as intended, nor is it likely that somebody will be willing or able to fix a problem with any of them. Additional information is found in section 9.

```
%mrci
  CIType        MRCI
                  MRDDCI1
                      MRDDCI2
                      MRDDCI3
                SORCI
                 SORCP
                MRACPF
```

(continues on next page)

(continued from previous page)

```
              MRACPF2
             MRACPF2a
             MRAQCC
             MRCEPA_R
             MRCEPA_0
             MRMP2
              MRMP3
             MRRE2
             MRRE3
             MRRE4
             CEPA1
             CEPA2
             CEPA3

  # CSF selection and convergence thresholds
  TSel            1e-14    # selection threshold
  TPre            1e-05    # pre-diagonalization threshold
  TNat              0.0      #
  ETol            1e-10
  RTol            1e-10

  # Size consistency corrections and the like
  EUnselOpt    MaxOverlap
                       FullMP2
  DavidsonOpt Davidson1
                       Davidson2
                       Siegbahn
                       Pople
    NELCORR   15   # number of electrons correlated for MRACPF and
    the like


  # MRPT stuff
  UsePartialTrafo   true/false    # speedups MRMP2
  UseDiagonalContraction true/false # legacy
  Partitioning  EN    # Epstein Nesbet
                       MP    # Moeller Plesset
                       RE     # Fink's partitioning
   FOpt           Standard   # choice of Fock operators to be
   used in MRPT
                     G0
                     G3
   H0Opt        Diagonal
                     Projected
                     Full
 MRPT_b            0.2   # intruder state fudge factor
 MRPT_SHIFT     1.0   # level shift

    # Integral handling
    IntMode     FullTrafo   # exact transformation (lots of memory)
                     RITRafo     # RI integrals (slow!)
   UseIVOs      true/false  # use improved virtual orbitals?

   # Try at your own risk
   CIMode    Auto
                 Conv
                 Semidirect
                 Direct
                 Direct2
                 Direct3
```

(continues on next page)

(continued from previous page)

```
     # orbital selection
     EWin          epsilon_min,epsilon_max   # orbital energy window
     MORanges  First_internal, Last_Internal, First_active,
     Last_Active, First-Virtual,Last_virtual  # alternative MO
     definition
     XASMOs  x1,x2,x3,...   # List of XAS donor MOs (see above)

     #density generation
    Densities StateDens, TransitionDens
        # StateDens= GS, GS_EL, GS_EL_SPIN, ALL_LOWEST,
        ALL_LOWEST_EL, ALL_LOWEST_EL_SPIN, ALL, ALL_EL, ALL_EL_SPIN
       # TransitionDens= FROM_GS_EL, FROMGS_EL_SPIN,FROM_LOWEST_EL, \\ FROM_LOWEST_EL_SPIN,
↪FROM_ALL_EL,FROM_ALL_EL_SPIN
     # Memory
    MaxMemVec  1024 # in MB
    MaxMemInt   1024 #  in MB


   # Diagonalizer
   Solver DIIS
             DIAG
             NEWDVD
    MaxDIIS
    RelaxRefs true/false
    LevelShift 0.0
    MaxDim  15
    NGuessMat 10
    MaxIter 25
    NGuessMatRefCI  100
    DVDShift  1.0


     # Bells and whistles
     KeepFiles    true/false
     AllSingles       true/false    # Force all singles to be included
     RejectInvalidRefs true/false  # reject references with wrong
     number of unpaired electrons or symmetry
     DoDDCIMP2   true/false # do a MP2 correction for the missing
     DDCI excitation class ijab
     NatOrbIters    5  # number of natural orbital iterations
     DoNatOrbs    0,1,2  # 0=not, 1=only average density, >=2= each
     density
     PrintLevel None, MINI, Normal, Large
     PrintWFN  1
     TPrintWFN 1e-3



    # MREOM stuff (expert territory!)
    DoMREOM  true/false

     # Definition of CI blocks
     NewBlock multiplicity irrep
         NRoots 1
         Excitations none
                           CIS
                           CID
                           CISD
         # active space definition
         refs CAS(nel, norb) end
         # or
```

(continues on next page)

(continued from previous page)

```
          refs RAS(nel: ras1norb ras1nel / ras2norb / ras3norb ras3nel
          ) end
          # or individual definition. Must yield the corret number of
          electrons!
          refs
             { 2 0 1 0 1 1 }
             { 2 0 1 1 0 1 }
             end
      end
end
```

6.8 MR-EOM-CC: Multireference Equation of Motion Coupled-Cluster

The Multireference Equation of Motion Coupled-Cluster (MR-EOM-CC) methodology [193, 194, 203, 406, 407, 639] has been implemented in ORCA. The strength of the MR-EOM-CC methodology lies in its ability to calculate many excited states from a single state-averaged CASSCF solution, for which only a single set of amplitudes needs to be solved and the final transformed Hamiltonian is diagonalized over a small manifold of excited states only through an uncontracted MRCI problem. Hence, a given MR-EOM calculation involves three steps, performed by three separate modules in ORCA :

1. a state-averaged CASSCF calculation (CASSCF module),
2. the solution of amplitude equations and the calculation of the elements of the similarity transformed Hamiltonians (MDCI module),
3. and the uncontracted MRCI diagonalization of the final similarity transformed Hamiltonian (MRCI module).

The current implementation allows for MR-EOM-T|$T^\dagger$-h-v, MR-EOM-T|$T^\dagger$|SXD-h-v and MR-EOM-T|$T^\dagger$|SXD|U-h-v calculations. A more detailed description of these methods and the available input parameters will be given in Sec. *Multireference Equation of Motion Coupled-Cluster (MR-EOM-CC) Theory*. We also note that the theoretical details underlying these methods can be found in Ref. [407]. In Sec. *Multireference Equation of Motion Coupled-Cluster (MR-EOM-CC) Theory*, we will discuss a strategy for the selection of the state-averaged CAS and other steps for setting up an MR-EOM calculation in detail. Furthermore, we will discuss how spin-orbit coupling effects can be included in MR-EOM calculations, a projection scheme to aid with convergence difficulties in the iteration of the T amplitude equations, an orbital selection scheme to reduce the size of the inactive core and virtual subspaces in the calculation of excitation energies and a strategy for obtaining nearly size-consistent results in MR-EOM. The purpose of this section is simply to provide a simple example which illustrates the most basic usage of the MR-EOM implementation in ORCA.

6.8.1 A Simple MR-EOM Calculation

Let us consider an MR-EOM-T|$T^\dagger$|SXD|U-h-v calculation on formaldehyde. An MR-EOM-T|$T^\dagger$|SXD|U-h-v calculation is specified via the `MR-EOM` keyword along with the specification of a state-averaged CASSCF calculation (i.e. CASSCF(nel, norb) calculation with the number of roots of each multiplicity to be included in the state-averaging for the reference state) and the number of desired roots in each multiplicity block for the final MRCI diagonalization. We note that the CASSCF module is described in sections *Complete Active Space Self-Consistent Field Method* and *The Complete Active Space Self-Consistent Field (CASSCF) Module* and that a description of the MRCI module is given in sections *Multireference Configuration Interaction and Pertubation Theory* and *The Multireference Correlation Module*. Here, we have a state-averaged CAS(6,4) calculation, comprised of 3 singlets and 3 triplets and we request 6 singlet roots and 6 triplet roots in our final MRCI diagonalization (i.e. the roots to be computed in the MR-EOM-T|$T^\dagger$|SXD|U-h-v calculation):

```
!MR-EOM def2-TZVP VeryTightSCF

%casscf  # reference state
 nel 6
 norb 4
 mult 1,3
 nroots 3,3
end

%mdci
  STol 1e-7
end

%mrci  # final roots
  newblock 1 *
    nroots 6
    refs cas(6,4) end
  end
  newblock 3 *
    nroots 6
    refs cas(6,4) end
  end
end


* xyz 0 1
 H        0.000000        0.934473       -0.588078
 H        0.000000       -0.934473       -0.588078
 C        0.000000        0.000000        0.000000
 O        0.000000        0.000000        1.221104
*
```

One can alternatively perform an MR-EOM-T|T†-h-v or MR-EOM-T|T†|SXD-h-v calculation by replacing the `MR-EOM` keyword, in the first line of the input above, by `MR-EOM-T|Td` or `MR-EOM-T|Td|SXD`, respectively. Namely, replacing the first line of the input above with

```
!MR-EOM-T|Td def2-TZVP VeryTightSCF
```

runs the MR-EOM-T|T†-h-v calculation, while

```
!MR-EOM-T|Td|SXD def2-TZVP VeryTightSCF
```

runs the MR-EOM-T|T†|SXD-h-v calculation.

The final MRCI diagonalization manifold includes 2h1p, 1h1p, 2h, 1h and 1p excitations in MR-EOM-T|T†-h-v calculations, 2h, 1p and 1h excitations in MR-EOM-T|T†|SXD-h-v calculations and 1h and 1p excitations in MR-EOM-T|T†|SXD|U-h-v calculations. Note that in the `%mdci` block, we have set the convergence tolerance (`STol`) for the residual equations for the amplitudes to 10^{-7}, as this default value is overwritten with the usage of the `TightSCF`, `VeryTightSCF`, etc. keywords. It is always important to inspect the values of the largest T, S (here, we use S to denote the entire set of S, X and D amplitudes) and U amplitudes. If there are amplitudes that are large (absolute values > 0.15), the calculated results should be regarded with suspicion. For the above calculation, we obtain:

```
--------------------
LARGEST T AMPLITUDES
--------------------
   8-> 13   8-> 13       0.060331
   4-> 17   4-> 17       0.029905
   8->  9   8->  9       0.028160
   8-> 16   8-> 16       0.027266
   6-> 20   6-> 20       0.025885
```

(continues on next page)

(continued from previous page)

```
   8-> 21   8-> 21       0.025308
   4-> 16   4-> 16       0.024803
   8-> 12   8-> 12       0.023915
   5-> 18   5-> 18       0.023553
   8-> 23   8-> 23       0.023384
   3-> 16   3-> 16       0.023182
   7-> 19   7-> 19       0.023043
   8-> 13   4-> 11       0.022010
   3-> 19   3-> 19       0.021987
   8-> 16   8->  9       0.021230
   8->  9   8-> 16       0.021230
```

for the T amplitudes,

```
--------------------
LARGEST S AMPLITUDES
--------------------
   4->  8   8-> 11       0.074048
   3->  8   8->  9       0.064886
   4->  5   5-> 11       0.045479
   3->  8   8-> 16       0.042657
   4->  7   7-> 11       0.042598
   4->  5   5-> 17       0.042076
   4->  5   8-> 11       0.039958
   4->  8   8-> 17       0.037532
   3->  5   8->  9       0.035907
   4->  7   7-> 17       0.035767
   2->  6   6-> 19       0.034148
   3->  5   5-> 10       0.033339
   2->  6   6-> 10       0.032691
   4->  6   6-> 11       0.032181
   8->  8   3-> 16       0.031775
   2->  7   7-> 22       0.031238
```

for the S amplitudes, and

```
--------------------
LARGEST U AMPLITUDES
--------------------
   3->  8   3->  8       0.026128
   3->  8   3->  5       0.007683
   2->  8   2->  8       0.006182
   3->  8   2->  5       0.006154
   2->  8   3->  5       0.004954
   3->  5   3->  5       0.004677
   4->  8   4->  8       0.003989
   3->  8   2->  8       0.002040
   2->  8   3->  8       0.002040
   2->  8   2->  5       0.001818
   4->  8   4->  5       0.001173
   2->  5   2->  5       0.001107
   4->  5   4->  5       0.000714
   3->  7   3->  7       0.000607
   3->  6   3->  6       0.000521
   2->  5   3->  5       0.000365
```

for the U amplitudes. Hence, one can see that there are no unusually large amplitudes for this calculation. **We note that there can be convergence issues with the T amplitude iterations and that in such cases, the flag:**

```
DoSingularPT true
```

should be added to the `%mdci` block. The convergence issues are caused by the presence of nearly singular T_2

amplitudes and setting the DoSingularPT flag to true activates a procedure which projects out the offending amplitudes (in each iteration) and replaces them by suitable perturbative amplitudes. For more information, see the examples in section *A Projection/Singular PT Scheme to Overcome Convergence Issues in the T Amplitude Iterations*.

After the computation of the amplitudes and the elements of the similarity transformed Hamiltonians, within the MDCI module, the calculation enters the MRCI module. For a complete, step by step description of the output of an MRCI calculation, we refer the reader to the example described in section *A Tutorial Type Example of a MR Calculation*. Let us first focus on the results for the singlet states (CI-BLOCK 1). Following the convergence of the Davidson diagonalization (default) or DIIS procedure, the following results of the MRCI calculation for the singlet states are printed:

```
----------
CI-RESULTS
----------

The threshold for printing is  0.30 percent
The weights of configurations will be printed. The weights are summed over
all CSFs that belong to a given configuration before printing

STATE   0:  Energy=   -114.321368425 Eh RefWeight=  0.9781  0.00 eV      0.0 cm**-1
      0.0137  : h---h---[0222]
      0.0756  : h---h---[1221]
      0.8879  : h---h---[2220]
STATE   1:  Energy=   -114.176866027 Eh RefWeight=  0.9765  3.93 eV  31714.6 cm**-1
      0.0039  : h---h---[1122]
      0.9726  : h---h---[2121]
      0.0071  : h---h  4[1222]
      0.0085  : h---h  4[2221]
STATE   2:  Energy=   -113.988050555 Eh RefWeight=  0.9774  9.07 eV  73154.8 cm**-1
      0.0044  : h---h---[1212]
      0.9730  : h---h---[2211]
      0.0063  : h---h  3[1222]
      0.0041  : h---h  3[2221]
STATE   3:  Energy=   -113.963862283 Eh RefWeight=  0.8810  9.73 eV  78463.5 cm**-1
      0.7459  : h---h---[1221]
      0.0807  : h---h---[2022]
      0.0533  : h---h---[2220]
      0.0228  : h---h  4[2122]
      0.0034  : h---h---[1220]p13
      0.0072  : h---h---[1220]p18
      0.0236  : h---h---[2120]p11
      0.0148  : h---h---[2120]p14
      0.0069  : h---h---[2120]p17
      0.0056  : h---h---[2120]p20
      0.0098  : h---h---[2210]p19
STATE   4:  Energy=   -113.931144468 Eh RefWeight=  0.0003 10.62 eV  85644.3 cm**-1
      0.0045  : h---h---[0122]p9
      0.0089  : h---h---[1121]p9
      0.9333  : h---h---[2120]p9
      0.0243  : h---h---[2120]p10
      0.0080  : h---h---[2120]p12
      0.0113  : h---h---[2120]p16
STATE   5:  Energy=   -113.929056780 Eh RefWeight=  0.6857 10.68 eV  86102.5 cm**-1
      0.0061  : h---h---[0222]
      0.0918  : h---h---[1221]
      0.5784  : h---h---[2022]
      0.0048  : h---h---[2202]
      0.0047  : h---h---[2220]
      0.2905  : h---h  4[2122]
      0.0045  : h---h---[2021]p13
```

For each state, the total energy is given in E_h; the weight of the reference configurations (RefWeight) in the given

state is provided, and the energy differences from the lowest lying state are given in eV and cm^{-1}. Also, in each case, the weights and a description of the configurations which contribute most strongly to the given state are also provided. See section *A Tutorial Type Example of a MR Calculation* for a discussion of the notation that is used for the description of the various configurations. To avoid confusion, we note that in the literature concerning the MR-EOM methodology [194, 203, 406, 407, 529, 530, 639], the term "%active" is used to denote the reference weight multiplied by 100%. In general, `RefWeight` should be > 0.9, such that the states are dominated by reference space configurations. This criterion is satisfied for the first three states and the reference weight of the fourth state is sufficiently close to 0.9. **However, the reference weights of the two higher lying states (especially state 4) are too small and these states should be discarded as the resulting energies will be inaccurate (i.e. states with significant contributions from configurations outside the reference space cannot be treated accurately)** .

In the case of the triplet states (`CI-BLOCK 2`), we obtain the following results:

```
----------
CI-RESULTS
----------

The threshold for printing is  0.30 percent
The weights of configurations will be printed. The weights are summed over
all CSFs that belong to a given configuration before printing

STATE   0:  Energy=   -114.190840989 Eh RefWeight=  0.9693  0.00 eV      0.0 cm**-1
      0.9691  : h---h---[2121]
      0.0079  : h---h  4[1222]
      0.0115  : h---h  4[2221]
STATE   1:  Energy=   -114.106733017 Eh RefWeight=  0.9941  2.29 eV  18459.6 cm**-1
      0.9941  : h---h---[1221]
STATE   2:  Energy=   -114.015150051 Eh RefWeight=  0.9787  4.78 eV  38559.7 cm**-1
      0.9786  : h---h---[2211]
      0.0050  : h---h  3[1222]
STATE   3:  Energy=   -113.939299674 Eh RefWeight=  0.0006  6.84 eV  55206.9 cm**-1
      0.0044  : h---h---[0122]p9
      0.0084  : h---h---[1121]p9
      0.9419  : h---h---[2120]p9
      0.0131  : h---h---[2120]p10
      0.0043  : h---h---[2120]p12
      0.0173  : h---h---[2120]p16
STATE   4:  Energy=   -113.925571130 Eh RefWeight=  0.4017  7.22 eV  58220.0 cm**-1
      0.3863  : h---h---[1122]
      0.0154  : h---h---[2121]
      0.1722  : h---h  4[1222]
      0.4098  : h---h  4[2221]
      0.0045  : h---h---[2120]p13
STATE   5:  Energy=   -113.910479339 Eh RefWeight=  0.0009  7.63 eV  61532.3 cm**-1
      0.0088  : h---h---[0122]p10
      0.0030  : h---h---[1121]p10
      0.0120  : h---h---[2120]p9
      0.9408  : h---h---[2120]p10
      0.0106  : h---h---[2120]p16
      0.0112  : h---h---[2120]p19
```

Here, we see that the first three states have reference weights which are > 0.9, while the reference weights of the final three states are well below that threshold. Hence, the latter three states should be discarded from any meaningful analysis.

Following the printing of the CI results for the final CI block, the states are ordered according to increasing energy and the vertical transition energies are printed:

```
--------------------
TRANSITION ENERGIES
--------------------
```

(continues on next page)

(continued from previous page)

```
The lowest energy is   -114.321368425 Eh

State Mult Irrep Root Block    mEh          eV      1/cm
  0    1    -1    0    0      0.000       0.000       0.0
  1    3    -1    0    1    130.527       3.552   28647.5
  2    1    -1    1    0    144.502       3.932   31714.6
  3    3    -1    1    1    214.635       5.841   47107.0
  4    3    -1    2    1    306.218       8.333   67207.2
  5    1    -1    2    0    333.318       9.070   73154.8
  6    1    -1    3    0    357.506       9.728   78463.5
  7    3    -1    3    1    382.069      10.397   83854.4
  8    1    -1    4    0    390.224      10.619   85644.3
  9    1    -1    5    0    392.312      10.675   86102.5
 10    3    -1    4    1    395.797      10.770   86867.5
 11    3    -1    5    1    410.889      11.181   90179.7
```

Furthermore, following the generation of the (approximate) densities, the absorption and CD spectra are printed:

```
========================================
MR-EOM Non Relativistic Properties
========================================

-------------------------------------------------------------------------------------------
↪-----
                    ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------
↪-----
     Transition      Energy     Energy  Wavelength fosc(D2)      D2        DX        DY       ␣
↪DZ
                      (eV)      (cm-1)    (nm)                (au**2)    (au)      (au)      ␣
↪(au)
-------------------------------------------------------------------------------------------
↪-----
  0-1A  ->  1-1A    3.932110   31714.6   315.3   0.000000000   0.00000  -0.00000  -0.00000   0.
↪00000
  0-1A  ->  2-1A    9.070040   73154.8   136.7   0.002137450   0.00962   0.09808  -0.00000  -0.
↪00000
  0-1A  ->  3-1A    9.728237   78463.5   127.4   0.157495738   0.66081  -0.00000   0.00000  -0.
↪81290
  0-1A  ->  4-1A   10.618534   85644.3   116.8   0.025353906   0.09746  -0.00000  -0.31218  -0.
↪00000
  0-1A  ->  5-1A   10.675343   86102.5   116.1   0.024673667   0.09434  -0.00000  -0.00000   0.
↪30715

-------------------------------------------------------------------------------------------
                    CD SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength    R         MX        MY        MZ
                      (eV)      (cm-1)    (nm)   (1e40*cgs)   (au)      (au)      (au)
-------------------------------------------------------------------------------------------
  0-1A  ->  1-1A    3.932110   31714.6   315.3   -0.00000    0.00000   0.00000   0.56273
  0-1A  ->  2-1A    9.070040   73154.8   136.7   -0.00000    0.00000  -0.74486   0.00000
  0-1A  ->  3-1A    9.728237   78463.5   127.4   -0.00000   -0.00000  -0.00000  -0.00000
  0-1A  ->  4-1A   10.618534   85644.3   116.8    0.00000    0.35898   0.00000  -0.00000
  0-1A  ->  5-1A   10.675343   86102.5   116.1    0.00000    0.00000  -0.00000  -0.00000
```

Warning

- It is important to note that the transition moments and oscillator strengths (and state dipole moments) have been blindly computed by the MRCI module and currently, no effort has been made to include

the effects of the various similarity transformations in the evaluation of these quantities. Hence these quantities are only approximate and should only be used as a qualitative aid to determine which states are dipole allowed or forbidden. Furthermore, since the calculated densities are approximate, so are the results of the population analysis that are printed before the absorption and CD spectra.

- While both the CASSCF and MRCI modules can make use of spatial point-group symmetry to some extent, the MR-EOM implementation is currently limited to calculations in C_1 symmetry.

6.8.2 Capabilities

The MR-EOM methodology can be used to calculate a desired number of states for both closed- and open-shell systems from a single state-averaged CASSCF solution. Currently, the approach is limited to serial calculations and to smaller systems in smaller active spaces. One should be aware that in the most cost-effective MR-EOM-T|T†|SXD|U-h-v approach (i.e. the smallest diagonalization manifold), an MRCI diagonalization is performed over all 1h and 1p excited configurations out of the CAS, which will inevitably limit the size of the initial CAS which can be used. We have also implemented an orbital selection scheme which can be used to reduce the size of the inactive core and virtual subspaces in the calculation of excitation energies, and this can be employed to extend the applicability of the approach to larger systems. The current implementation can also be used in conjunction with the spin-orbit coupling submodule (*General Description*) of the MRCI module to calculate spin-orbit coupling effects in MR-EOM calculations to first order. These and other features of the current implementation will be discussed in *Multireference Equation of Motion Coupled-Cluster (MR-EOM-CC) Theory*.

6.8.3 Perturbative MR-EOM-PT

The MR-EOM family of methods now also features an almost fully perturbative approach called MR-EOMPT [501]. This method shares the features of the MR-EOMCC parent method while using non-iterative perturbative estimates for the $\hat{T}$ and $\hat{S}, \hat{X}, \hat{D}$ amplitudes. This slightly reduces the accuracy compared to iterative MR-EOMCC while reducing runtime. Furthermore, convergence issues due to nearly singular $\hat{T}$ and $\hat{S}, \hat{X}, \hat{D}$ amplitudes cannot occur anymore.

This method can be invoked by adding the keyword `DoMREOM_MRPT True` to the `%mdci` block.

6.9 Solvation

ORCA features several implicit solvation models, including the fully integrated "conductor-like polarizable continuum (C-PCM)" and "Minnesota SMD" solvation models, which are available in all its components. With these models, various types of calculations can be performed using a polarizable continuum with a realistic van der Waals cavity as summarized below:

- Energies of molecules in solution with a finite dielectric constant ε using HF or any DFT method.
- Optimization of molecular structures in solution using HF or any DFT method with analytic gradients.
- Calculation of vibrational frequencies using the analytic Hessian for HF or any DFT method, provided that the same calculation is available in vacuum.
- Calculation of solvent effects on response properties like polarizabilities through coupled-perturbed SCF theory. For magnetic response properties, such as the g-tensor, the C-PCM response vanishes.
- Calculations of solvent shifts on transition energies using the time-dependent DFT or CIS method. The refractive index of the solvent needs to be provided in addition to the dielectric constant.
- First order perturbation estimate of solvent effects on state and transition energies in multireference perturbation and configuration-interaction calculations.

Other implicit solvation strategies are available in ORCA. In particular, an interface to the open source implementation of the COSMO-RS model (openCOSMO-RS), as well as different solvation models that can be used

in XTB (ALPB, ddCOSMO, and CPCM-X). A detailed overview of the available implicit solvation methods and their usage is provided in Sections *ONIOM Methods*, and *Implicit Solvation Models*.

As a simple example, let us compute the solvent effect on the $n \rightarrow \pi^*$ transition energy in formaldehyde with the C-PCM model. This effect can be obtained by subtracting the solution-phase and gas-phase transition energies. The gas-phase transition energy (4.633 eV) can be computed by using the following input:

```
! def2-TZVP

%cis nroots 1 end

*int 0 1
 C     0   0   0   0.000000     0.000     0.000
 O     1   0   0   1.200371     0.000     0.000
 H     1   2   0   1.107372   121.941     0.000
 H     1   2   3   1.107372   121.941   180.000
*
```

By adding the `CPCM(water)` flag to the input used for the gas-phase calculation, the transition energy can now be computed using the C-PCM model with water as the solvent:

```
! def2-TZVP CPCM(water)

%cis nroots 1 end

*int 0 1
 C     0   0   0   0.000000     0.000     0.000
 O     1   0   0   1.200371     0.000     0.000
 H     1   2   0   1.107372   121.941     0.000
 H     1   2   3   1.107372   121.941   180.000
*
```

This C-PCM calculation yields a transition energy of 4.857 eV:

```
-----------------------------
CIS-EXCITED STATES (SINGLETS)
-----------------------------

the weights of the individual excitations are printed if larger than 1.0e-02

STATE  1:  E=   0.178499 au      4.857 eV    39176.0 cm**-1 <S**2> =   0.000000
7a ->   8a  :     0.929287 (c= -0.96399514)
7a ->  13a  :     0.039268 (c=  0.19816055)
7a ->  18a  :     0.016344 (c=  0.12784298)
```

Hence, water environment increases the transition energy by 0.224 eV. This increase can be attributed to the stabilization of lone pair orbitals by the presence of water molecules.

6.10 ORCA SOLVATOR: Automatic Placement of Explicit Solvent Molecules

From ORCA6, we also have a tool that can automatically place explicit solvent molecules to a given system. It can be done using two different approaches: a `STOCHASTIC` method which is very fast but less accurate, or a `DOCKING` approach which makes use of the *DOCKER*. The later is slower, but more accurate and is the default.

6.10.1 First Example: Adding Water to a Histidine

As a very simple initial example, let's take a Histidine aminoacid and add three explicit water molecules at the best positions using the DOCKER and GFN2-XTB. The input to get this is as simple as:

```
!XTB ALPB(WATER) PAL16
%SOLVATOR NSOLV 3 END
* XYZ 0 1
  N     0.885996   -0.961304   -0.120339
  C     1.798313    0.104987    0.275069
  C     1.249714    0.744242    1.567548
  O     1.573032    1.831447    1.951866
  C     2.049102    1.187937   -0.781297
  C     2.714441    0.645674   -1.999072
  N     2.728606    1.335092   -3.185194
  C     3.401683    0.571600   -4.081842
  N     3.805612   -0.545068   -3.552995
  C     3.389082   -0.516790   -2.258890
  O     0.397674   -0.041862    2.212628
  H     0.272440   -0.853698    1.671197
  H     1.339440   -1.612664   -0.750009
  H     0.086389   -0.572348   -0.612909
  H     2.756495   -0.353470    0.548654
  H     2.661849    1.969568   -0.321790
  H     1.092545    1.646672   -1.057454
  H     2.328734    2.246496   -3.338113
  H     3.566873    0.866218   -5.098496
  H     3.616501   -1.333139   -1.602802
*
```

That is as simple as a regular input with the line %SOLVATOR NSOLV 3 END added. The solvent structure will be automatically taken from the implicit solvation method (in this case ALPB(WATER)), and the three water molecules will be added. The output will look like:

```
                     *****************
                     * ORCA Solvator *
                     *****************

Solvent chosen:                              WATER
Solvent radius:                              .... 1.69 Angs
Solvent max dimensions (x,y,z):              .... 2.73, 2.32, 1.52 Angs
Number of solvent molecules to be added:     .... 3 molecules
Method used to add the solvent:              .... docking
Number of atoms of solvent molecule:         .... 3 atoms
Coordinates of solvent in Angstroem:

   O    0.000014   0.401429   0.000000
   H    0.765192  -0.200729   0.000000
   H   -0.765206  -0.200700   0.000000

Solute radius:                               .... 5.27 Angs
Ellipsoid potential radii:                   .... 8.37, 6.28, 5.76 Angs
```

where the solvent chosen is printed, together with some details about its dimensions, the number of molecules to be added and the method. The structure from the internal database is also always printed.

The process is then monitored per solvent molecule:

```
Adding solvent molecules to the solute        ....


      Iter       Energy      Einter          dE          Time
```

(continues on next page)

(continued from previous page)

```
                  (Eh)      (kcal/mol)    (kcal/mol)      (min)
      ------------------------------------------------------------
         1    -39.452446    -4.331670     -4.331670       0.26
         2    -44.544347    -4.316639      0.015031       0.34
         3    -49.636018    -4.171990      0.144649       0.40

Final radius after microsolvation:             .... 4.88 Angs
Time needed for microsolvation   :             .... 60.54 s
Final structured saved to        :             HIS.solvator.xyz

                ****ORCA-SOLVATOR TERMINATED NORMALLY****
```

and the final result is printed to the file `Basename.solvator.xyz`. There will be also an intermediate file named `Basename.solvator.solventbuild.xyz` with the solvent molecules added one by one.

Note

In contrast to the DOCKER, the solute is **always** frozen by default. Set `FIXSOLUTE FALSE` under the `%SOLVATOR` block to change that.

On the output `Einter` is the interaction energy obtained from the DOCKER and `dE` is the different between the current and the previous `Einter`.

In this case, the result looks like:

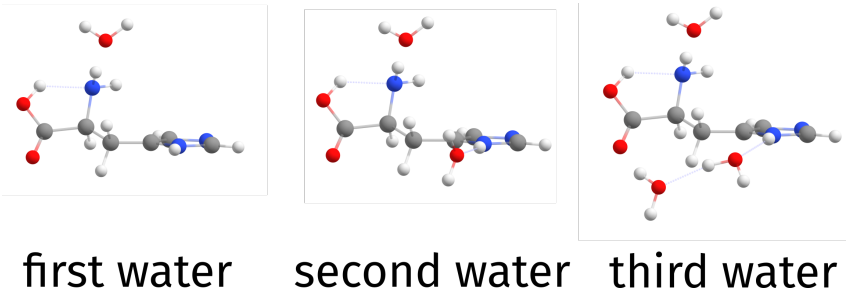


Fig. 6.34: Three water molecules added by the solvator.

Note

Currently the SOLVATOR is **only** working with the GFN-XTB and GFN-FF methods and the ALPB solvation model. It will be expanded later to others.

6.10.2 Other Solvents

The method itself is agnostic to the solvent, and any other could have been used. The example above with DMSO would be:

```
!XTB ALPB(DMSO) PAL16
%SOLVATOR NSOLV 3 END
* XYZFILE 0 1 HIS.xyz
```

and results in:

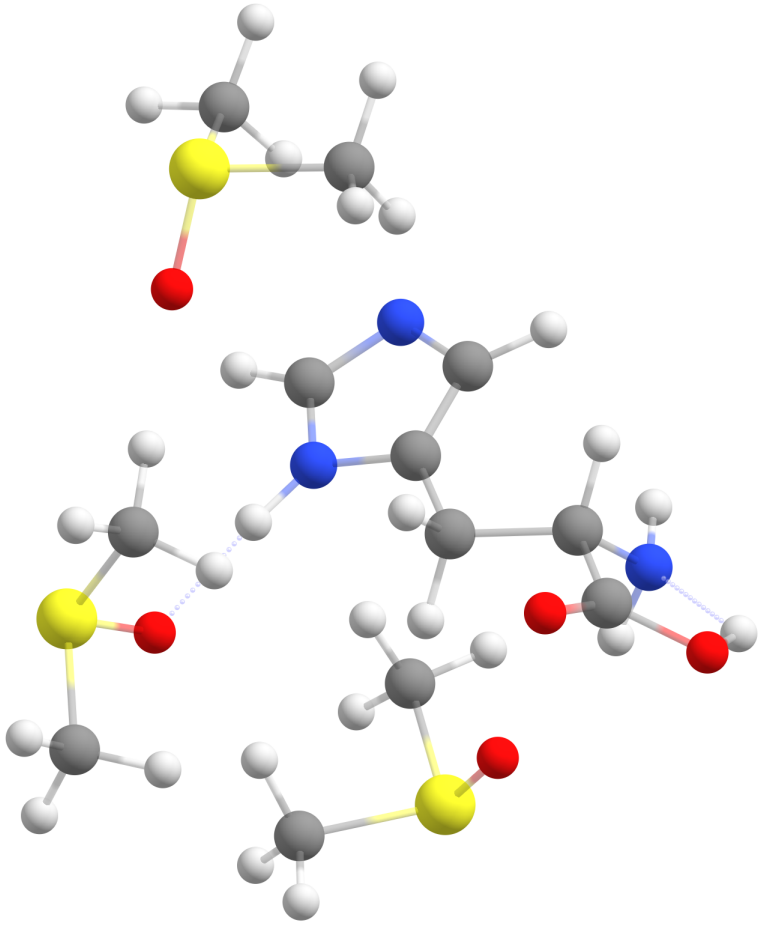


Fig. 6.35: Three DMSO molecules added by the solvator.

and even a custom solvent can be given in the form of a `.xyz` file with:

```
%SOLVATOR SOLVENTFILE "solvent_file_name.xyz" END
```

As with the docker, the `charge` and `multiplicty` can be given to the solvent if given as two integers on the comment line (default is neutral closed-shell). Since the `ALPB` method has a fixed number of solvents, right now one can still not give a custom epsilon value for the custom solvents, but it needs to be approximated to the next closest solvent.

As an example, let's create a file named `isopropanol.xyz` with a solvent which is not on the `ALPB` list:

```
12
0 1
C         -3.79410        2.24670       -0.09622
C         -3.45574        0.76660       -0.18820
H         -2.94382        2.85306       -0.42645
H         -4.00575        2.53923        0.93817
H         -4.66172        2.49512       -0.71492
C         -4.60559       -0.10847        0.28691
H         -4.84802        0.09429        1.33594
H         -4.32886       -1.16657        0.22730
H         -5.50341        0.05251       -0.31744
O         -2.30376        0.49878        0.60542
H         -3.20686        0.51239       -1.22373
H         -2.51694        0.72259        1.52758
```

and run:

```
!XTB ALPB(ETHANOL) PAL16
%SOLVATOR SOLVENTFILE "isopropanol.xyz" NSOLV 3 END
* XYZFILE 0 1 HIS.xyz
```

approximating the dielectric constant of isopropanol to that of ethanol. That might look not too accurate, but the ALPB implicit solvation is also not a very good implicit solvation model anyway, so results will be quite similar.

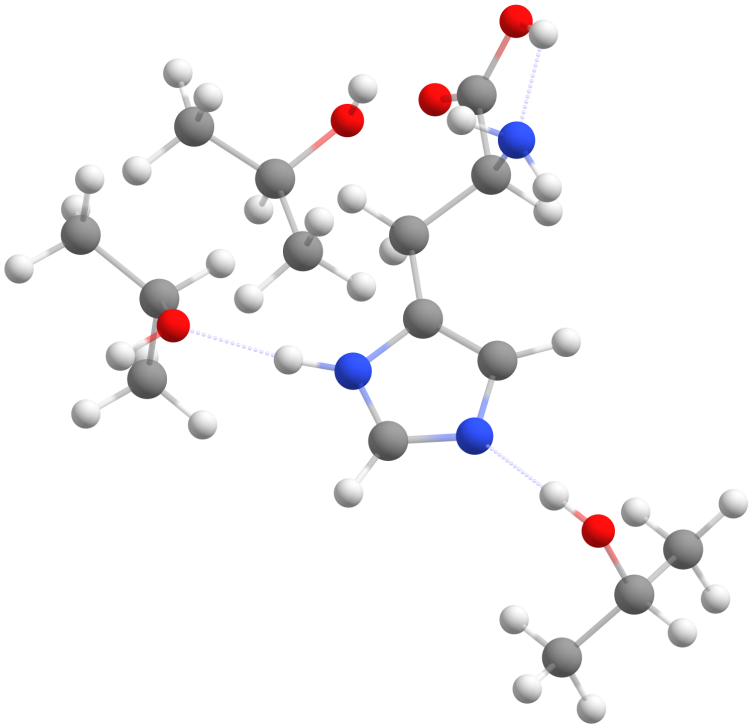


Fig. 6.36: Three iso-propanol molecules added by the solvator.

6.10.3 The Stochastic Method for Multiple Solvents

In case you want to add a really large number of explicit solvent molecules, the `STOCHASTIC` mode will be significantly faster. Let's add 100 water molecules on our target Histidine, now using `CLUSTERMODE STOCHASTIC`:

```
!XTB ALPB(WATER) PAL16
%SOLVATOR
    NSOLV       100
    CLUSTERMODE STOCHASTIC
END
* XYZFILE 0 1 HIS.xyz
```

The output will be somewhat different:

```
Solvent chosen:                                WATER
Solvent radius:                                .... 1.69 Angs
Solvent max dimensions (x,y,z):                .... 2.73, 2.32, 1.52 Angs
Number of solvent molecules to be added:       .... 100 molecules
Method used to add the solvent:                .... stochastic
Number of atoms of solvent molecule:           .... 3 atoms
Coordinates of solvent in Angstroem:

   O    0.000014    0.401429    0.000000
```

(continues on next page)

(continued from previous page)

```
   H    0.765192  -0.200729   0.000000
   H   -0.765206  -0.200700   0.000000

Solute radius:                                  .... 5.27 Angs
Adding solvent molecules to the solute          ....

      Iter     Target function   Time
                  (Coulomb)     (min)
      --------------------------------
         1     -4.659435e-07     0.00
         2     -4.604606e-07     0.00
         3     -2.519684e-07     0.00
      (...)

Final radius after microsolvation:              .... 9.79 Angs
Time needed for microsolvation   :              .... 5.11 s
Final structured saved to        :              HIS_STOCHASTIC.solvator.xyz

                ****ORCA-SOLVATOR TERMINATED NORMALLY****
```

and in a few seconds all solvent molecules are added as can be seen from the result:

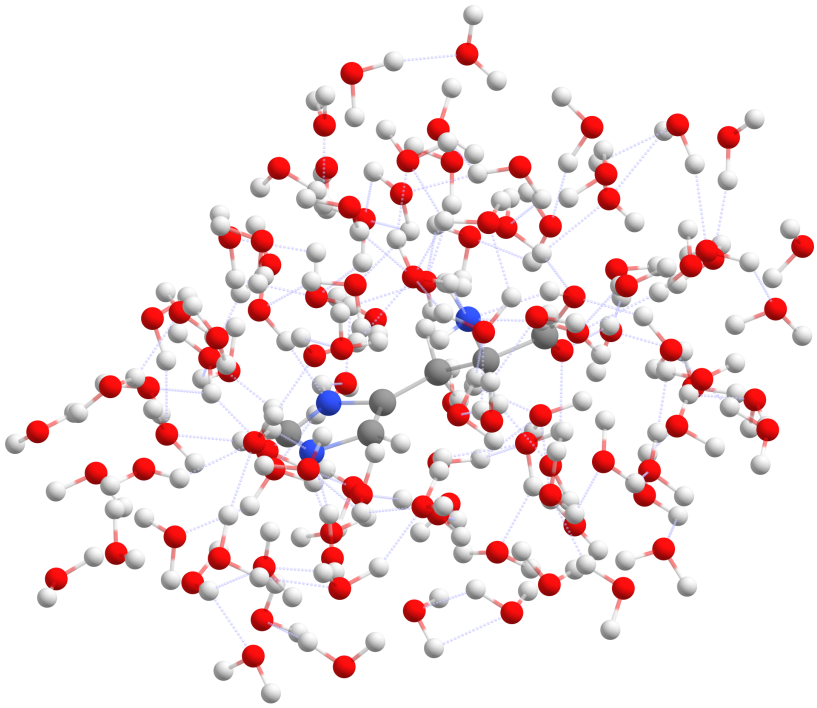


Fig. 6.37: A hundred water molecules added by the solvator.

Important

As the name says, the `CLUSTERMODE STOCHASTIC` is a probabilistic approach and is not nearly as accurate as the `DOCKING` mode! Nonetheless it is useful for quite a few applications.

6.10.4 Creating a Droplet

The regular stochastic method will create a solvation sphere around the solute following it shape and topology. In case you want to create a solvent distribution with spherical symmetry, you have to use set the DROPLET TRUE keyword, such as:

```
!XTB ALPB(WATER) PAL16
%SOLVATOR
    NSOLV       100
    CLUSTERMODE STOCHASTIC
    DROPLET     TRUE
END
* XYZFILE 0 1 HIS.xyz
```

And the result will look like:

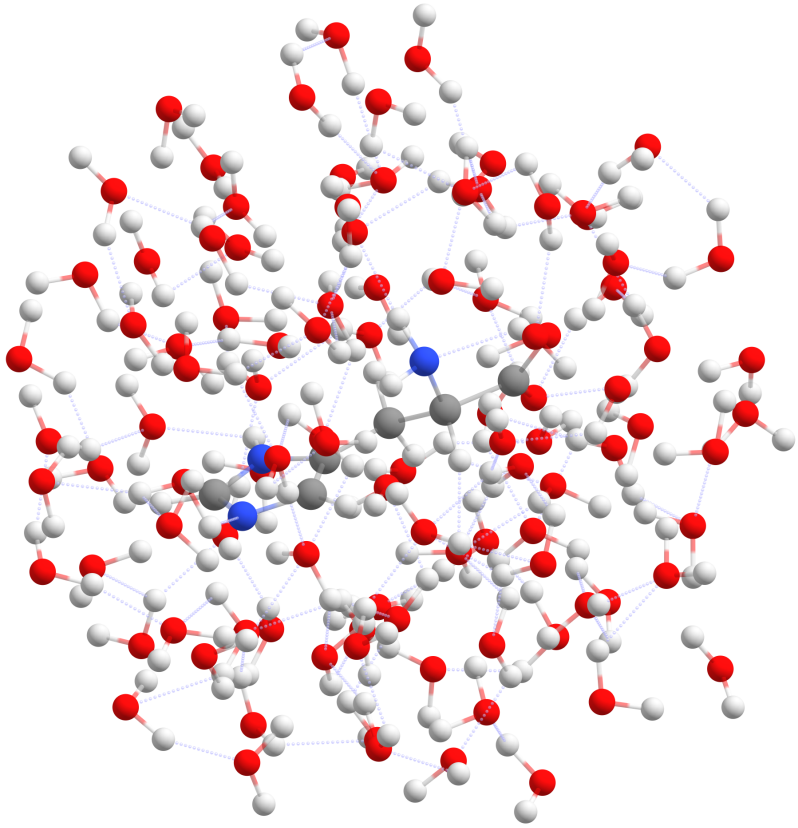


Fig. 6.38: A hundred water molecules added by the solvator by enforcing spherical symmetry.

6.10.5 Creating a Droplet with a Defined Radius

Instead of defining the number of solvent molecules, one can also defined a maximum radius and the SOLVATOR will add as many molecules as necessary until the radius is reached. This is only compatible with CLUSTERMODE STOCHASTIC!

```
!XTB ALPB(WATER) PAL16
%SOLVATOR
    RADIUS      15 # in angstroem
    CLUSTERMODE STOCHASTIC
END
* XYZFILE 0 1 HIS.xyz
```

The output in the end shows:

```
Desired sovent radius:                          .... 15.00 Angs
Actual sovent radius:                           .... 15.18 Angs
Final number of solvent molecules:              .... 668 molecules
```

Note

The radius is taken from the centroid of the solute!

and a radius close to 15 Angstroem was achieved when 668 molecules are added:

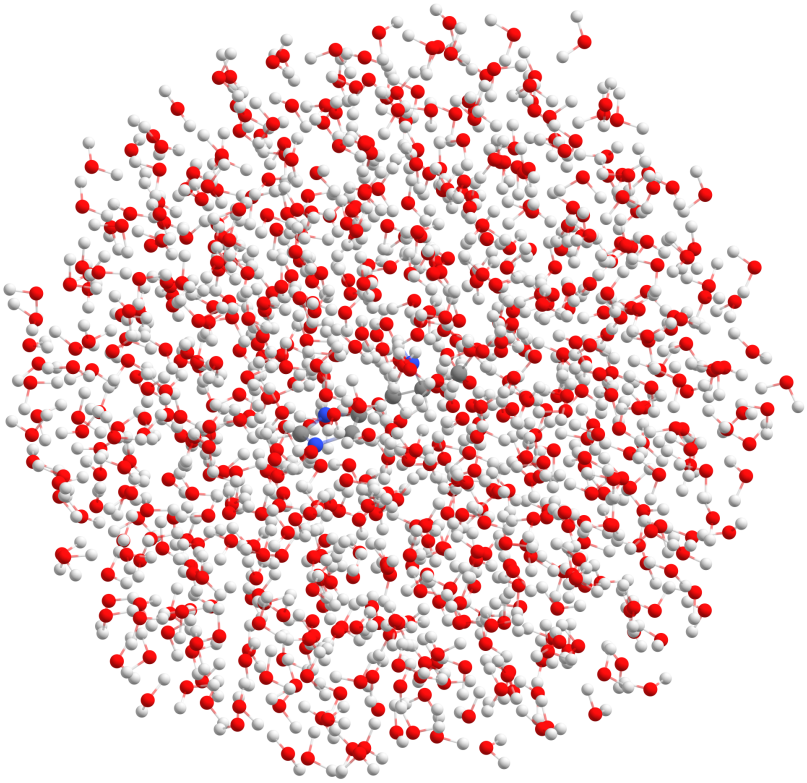


Fig. 6.39: A droplet create with 668 water molecules to achieve a radius of approximately 15 Angstroem.

Important

All examples described above work for any other solvents, including custom ones.

Note

The default DOCKER settings for the solvator are equivalent to `!QUICKDOCK`. For more accurate methods use `!NORMALDOCK` or even `!COMPLETEDOCK`, however they will be much slower.

A complete list of keywords and more discussions on the topic can be found at the later section *More on the ORCA SOLVATOR*

6.11 Relativistic Calculations

ORCA features three different approximations to cover relativistic effects:

1. The „Exact 2 component“ (X2C) Hamiltonian
2. The Douglas-Kroll-Hess (DKH) Hamiltonian to second order
3. The 0th order regular approximation (ZORA) with a model potential

Earlier versions of ORCA supported a number of additional approximations, which are no longer supported.

The main relativistic Hamiltonian that will be pursued in further development is the X2C Hamiltonian. Of the three alternatives, we believe that X2C has the best feature set and we recommend to all of our users to preferentially use this method.

All three relativistic model Hamiltonians are implemented for scalar relativistic energy calculations and these are carried through consistently through the entire program. Scalar relativity shows up as an additional effective potential that is added to the one-electron matrix. Scalar relativistic corrections to the two-electron interaction are not available in ORCA. Furthermore, self-consistent field calculations (HF, DFT, CASSCF) with inclusion of spin-orbit-coupling (SOC) are also not available in ORCA but we will not exclude the possibility to add this feature in a future version of the program.

A general overview and some practical recommendations are given in the next sections. For detailed documentation and all available options see *Relativistic Options*.

6.11.1 Basis sets for relativistic calculations

The different scalar relativistic potentials have different shapes in the core region. Consequently, each one of them requires specialized all electron basis sets that are optimized for the Hamiltonian at hand. The most common choices are listed in the sections *Relativistically recontracted Karlsruhe basis sets* and *SARC basis sets* with all available options listed in *Built-in Basis Sets*. An uncontracted basis set of sufficient size will always work. Likewise, uncontracted fitting basis sets in all forms of RI calculations are always appropriate.

Hint

Use the `!Decontract` keyword to decontract the chosen (all-electron) basis set and make it suitable for any relativistic Hamiltonian, as well as comparisons between them.

If large, uncontracted basis sets are used in scalar relativistic calculations, there is a distinct danger of variational collapse. This behavior is related to the fact that the relativistic orbitals will diverge for a point nucleus. ORCA features the Gaussian finite nucleus model of Dyall and Visscher for DKH and X2C. We recommend to always use this feature (`FiniteNuc`) in relativistic calculations.

Given the fact relativistic all-electron calculations on heavy element compounds feature very steep core basis functions, numeric integration, such as in DFT and COSX, may be challenging. ORCA features automatic procedures that adapt the integration grids for the presence of steep basis functions. However, in case you experience strange results, the numeric integration is one potential source of problem. The cure is to go to larger integration grids and, in particular, increase the radial integration accuracy (`IntAcc`).

6.11.2 Scalar-relativistic gradients and properties

Of the three model Hamiltonians, only X2C features analytic gradients. Hence, for geometry optimizations this is also the preferred methods. For DKH and ZORA, the program automatically switches to the one-center approximation. This requires some attention by the users since final single point energies obtained with the one-center approximation are inconsistent with energies obtained without it. The one-center approximation is usually of sufficient accuracy but we have observed cases in actual applications where it leads to clearly wrong geometries. Hence, we strongly recommend to use the X2C Hamiltonian in this realm.

Caution

Geometry optimizations with DKH and ZORA (but not X2C) automatically use the one-center approximation. When computing relative energies, do not mix energies from single-point calculations without the one-center approximation with those from geometry optimizations that do make use of this feature.

If relativistic calculations are used for molecular properties there is a potential mismatch between non-relativistically calculated property integrals and the relativistic Hamiltonian. The procedure to remove these inconsistencies is referred to as „picture change“. The picture change is usually carried through to the same level of approximation as the decoupling of the relativistic Hamiltonian into two-component and eventually to one-component form. We strongly recommend to use picture change in all relativistic property calculations and consequently, this is also the default. Relativistic property calculations without picture change are wildly inaccurate, in particular if operators are involved that carry inverse powers of the electorn-nucleus distance. Picture change effects are implemented for DKH and X2C and to some extent also for ZORA. However, they are not implemented for all properties that ORCA can calculate. Please pay attention to the output of the property integral and property programs. Both programs will explicitly state which picture change effects are included in the molecular integrals.

```
%rel
  FiniteNuc     true   # Invoke the Gaussian finite nucleus model.
  PictureChange 1 or 2 # First or second order picture change effects.
                       # Second order is potentially more accurate and more expensive.
end
```

6.11.3 Exact two-component method (X2C)

Despite the name, the X2C method is implemented in ORCA only as a scalar-relativistic, effective one-component method. The theory and implementation are discussed in *Exact Two-Component Theory (X2C)*, together with appropriate references to cite in your work. In the simplest case, it is sufficient to add the `X2C` simple keyword to the input and choose an appropriate basis set:

```
! X2C X2C-TZVPall X2C/J
```

The DLU approximation,[659] discussed in *DLU approximation*, is the recommended way to reduce the cost of the X2C transformation, particularly for gradient/Hessian calculations, with minor loss of accuracy. It is available via the simple input keyword `DLU-X2C`.

6.11.4 Douglas-Kroll-Hess (DKH)

The first- or second-order DKH method be requested via the simple input keywords `DKH1` or `DKH2`, respectively (`DKH` is an alias for the latter), together with appropriate basis sets:

```
! DKH DKH-def2-TZVP SARC/J
```

For most calculations, no other settings are needed. See *The Douglas-Kroll-Hess Method* for an overview of the underlying theory.

6.11.5 ZORA and IORA

The 0^{th} order regular approximation (ZORA; pioneered by van Lenthe et al., see Ref. [863] and many follow up papers by the Amsterdam group) implementation in ORCA essentially follows van Wüllen [866] and solves the ZORA equations with a suitable model potential and a model density derived from accurate atomic ZORA calculations. See *Relativistic Options* for explanation of the `ModelPot` and `ModelDens` keywords used to control these models. If the relevant precautions are taken (see below), the use of the ZORA or IORA methods is as easy as in the DKH/X2C case. For example:

```
! ZORA ZORA-def2-TZVP SARC/J
# for more detail use
%rel
     ModelPot 1,1,1,1
     ModelDens rhoZORA
end
```

Attention

The ZORA method is highly dependent on numerical integration and it is very important to pay attention to the subject of radial integration accuracy! By default, from ORCA 5.0 we consider that during the grid construction and the defaults should work very well. Only for very problematic cases, consider using a higher `IntAcc` parameter or at least to increase the radial integration accuracy around the heavy atoms using `SpecialGridAtoms` and `SpecialGridIntAcc`.

6.12 Calculation of Properties

6.12.1 Population Analysis and Related Things

Atomic population related things are not real molecular properties since they are not observables. They are nevertheless highly useful for interpreting experimental and computational findings. By default, ORCA provides very detailed information about calculated molecular orbitals and bonds through Mulliken, Löwdin, and Mayer population analyses. However, as it is easy to become overwhelmed by the extensive population analysis section of the output, ORCA allows users to turn off most features.

```
! HF def2-SVP Mulliken Loewdin Mayer ReducedPOP

*xyz 0 1
 C  0  0  0
 O  0  0  1.13
*
```

The "`ReducedPOP`" keyword reduces the information printed out in the population analysis section, providing orbital population of each atom with percent contribution per basis function type. This is highly useful in figuring out the character of the MOs. Furthermore, one can request a printout of the MO coefficients through the output block of the input file (see section *Population Analyses and Control of Output*) or using the keyword "`PrintMOs`"

The distribution of the frontier molecular orbitals (FMOs) over the system can be requested with the "`FMOPop`" keyword:

```
! HF def2-SVP FMOPop

*xyz 0 1
 C  0  0  0
 O  0  0  1.13
*
```

This provides Mulliken and Loewdin population analyses on HOMO and LUMO:

```
-------------------------------------------
FRONTIER MOLECULAR ORBITAL POPULATION ANALYSIS
-------------------------------------------

ANALYZING ORBITALS: HOMO=   6 LUMO=   7

------------------------------------------------------------------------
   Atom        Q(Mulliken)     Q(Loewdin)     Q(Mulliken)     Q(Loewdin)
              <<<<<<<<<<<<HOMO>>>>>>>>>>>>   <<<<<<<<<<<<LUMO>>>>>>>>>>>>
------------------------------------------------------------------------
   0-C         0.937186        0.906827       0.804044        0.755610
   1-O         0.062814        0.093173       0.195956        0.244390
------------------------------------------------------------------------
```

Visualization of three-dimensional representation of MOs, natural orbitals, electron densities, and spin densities is usually more intuitive than examining MO coefficients and it is is described in detail in section *Orbital and Density Plots*. The files necessary for such visualizations can be readily generated with ORCA in various ways and then opened in visualization software such as `gOpenMol` and `Molekel`.[1] . In the following example, we briefly describe visualization of MOs.

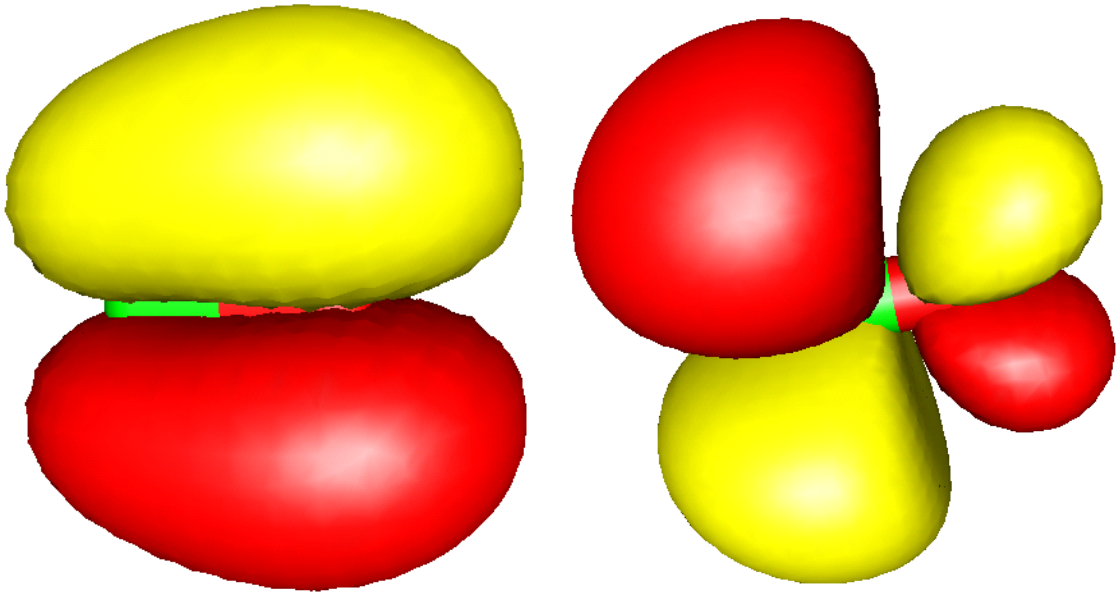
To visualize MOs with`gOpenMol`, the `plt` file of MOs can be generated in the `gOpenMol_bin` format from the `gbw` file using `orca_plot` utility program or directly from the ORCA run through the `%plots` block of the input file:

```
! HF def2-SVP XYZFile

%plots Format gOpenMol_bin
       MO("CO-4.plt",4,0);
       MO("CO-8.plt",8,0);
       end

*xyz 0 1
 C  0  0  0
 O  0  0  1.13
*
```

In this input file, the `MO("CO-4.plt",4,0);` command is used to evaluare MO labeled as 4 for operator 0 and then to strore it in the "CO-4.plt" file. For RHF and ROHF, one should always use operator 0. For UHF, operators 0 and 1 correspond to spin-up and spin-down orbitals, respectively.

When the produced `plt` files are opened with `gOpenMol` (see section *Surface Plots* for details), the textbook-like π and π^* MOs of the CO molecule are visualized as in Figure Fig. 6.40.

[1] The `Molekel` developers ask for the following citation – please do as they ask:
MOLEKEL 4.2, P. Flukiger, H.P. Lüthi, S. Portmann, J. Weber, Swiss Center for Scientific Computing, Manno (Switzerland), 2000-2006.
S. Portmann, H.P.Lüthi. MOLEKEL: An Interactive Molecular Graphics Tool. CHIMIA (2000), 54, 766-770. The program appears to be maintained by Ugo Varetto at this time.



(a) (a) (b) (b)

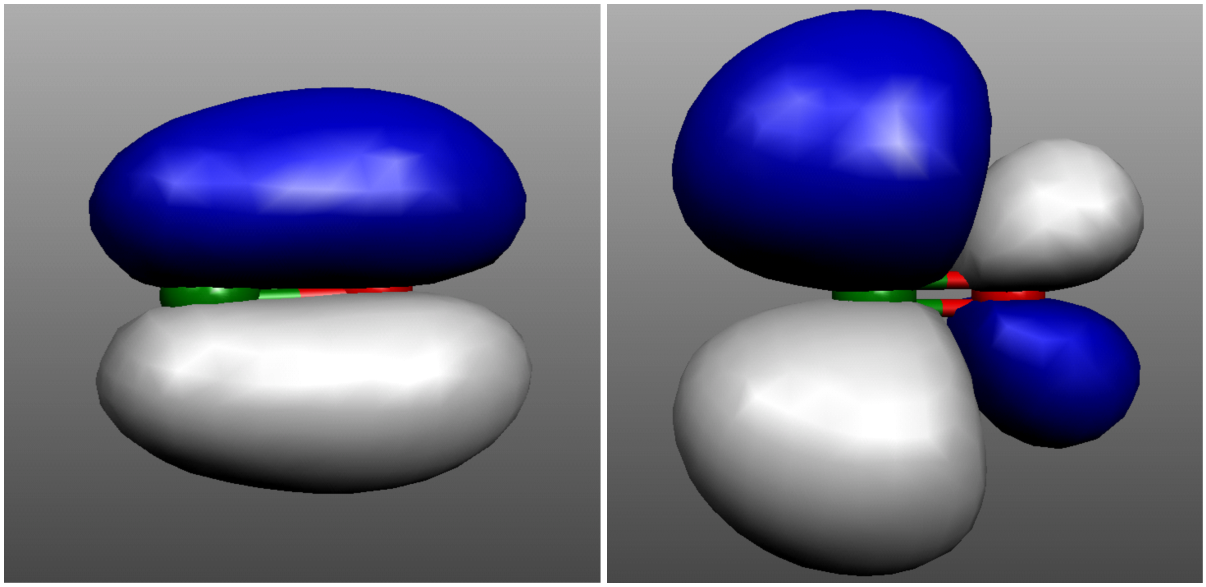
Fig. 6.40: (a) π and (b) π^* MOs of the CO molecule obtained from the interface of ORCA to gOpenMol.

If the gOpenMol_ascii file format was requested, gOpenMol conversion utility or some other tools might then be needed to convert this human-readable file to the machine-readable gOpenMol_bin format.

In order to use the interface to Molekel, an ASCII file in the Cube or Gaussian_Cube format needs to be generated. Such ASCII files can be actually transferred between platforms. The Cube format can be requested in the %plots block as:

```
! HF def2-SVP XYZFile

%plots Format Cube
       MO("CO-4.cube",4,0);
       MO("CO-8.cube",8,0);
       end

*xyz 0 1
 C  0  0  0
 O  0  0  1.13
*
```

To visualize MOs strored in the *.cube file, start Molekel and, via a right mouse click, load the *.xyz file and/or the *.cube file. lternatively, navigate to the surface menu, select the "gaussian-cube" format, and load the surface. For orbitals, click the "both signs" button and select a countour value in the "cutoff" field. Then, click "create surface". The colour schemes and other fine details of the plots can be easily adjusted as desired. Finally, create files via the "snapshot" feature of Molekel. Figure Fig. 6.41 demonstrates a Molekel variant of Figure Fig. 6.40.



(a) (a) (b) (b)

Fig. 6.41: (a) π and (b) π^* MOs of the CO molecule obtained from the interface of `ORCA` to `Molekel`.

It is worth noting that there are several other freeware programs, such as `UCSF CHIMERA`, that can read `Gaussian_Cube` files and provide high-quality plots.

In some situations, visualization of the electronic structure in terms of localized molecular orbitals might be quite helpful. As unitary transformations among occupied orbitals do not change the total wavefunction, such transformations can be applied to the canonical SCF orbitals with no change of the physical content of the SCF wavefunction. The localized orbitals correspond more closely to the pictures of orbitals that chemists often enjoy to think about. Localized orbitals according to the Pipek-Mezey population-localization scheme are quite easy to compute. For example, the following run reproduces the calculations reported by Pipek and Mezey in their original paper for the N_2O_4 molecule.

```
! HF STO-3G Bohrs

%loc
     LocMet  PipekMezey # localization method. Choices:
                            # PipekMezey (=PM)
                            # FosterBoys (=FB)
     T_Core  -1000      # cutoff for core orbitals
     Tol     1e-8       # conv. Tolerance (default=1e-6)
     MaxIter 20         # max. no of iterations (def. 128)
     end

*xyz 0 1
 N    0.000000   -1.653532   0.000000
 N    0.000000    1.653532   0.000000
 O   -2.050381   -2.530377   0.000000
 O    2.050381   -2.530377   0.000000
 O   -2.050381    2.530377   0.000000
 O    2.050381    2.530377   0.000000
*
```

Based on the output file of this job, localized MOs consist of six core like orbitals (one for each N and one for each O), two distinct lone pairs on each oxygen, a σ- and a π-bonding orbital for each N-O bond and one N-N σ-bonding orbital which corresponds to the dominant resonance structure of this molecule. You will also find a file with the extension `.loc` in the directory where you run the calculation. Like the standard `gbw` file, it can used to extract files for plotting or as input for another calculation (warning! The localized orbitals have no well defined orbital energy. If you do use them as input for another calculation use `GuessMode=CMatrix` in the `%scf` block).

If you have access to a version of the gennbo program from Weinhold's group[2], you can also request natural population analysis and natural bond orbital analysis. The interface is elementary and is invoked through the keywords NPA and NBO, respectively:

```
! HF def2-SVP NPA XYZFile

* xyz 0 1
  C  0  0  0
  O  0  0  1.13
*
```

If you choose simple NPA, then you will only obtain a natural population analysis. When NBO is chosen instead, the natural bond orbital analysis will also be carried out. ORCA leaves a FILE.47 file on disk. This file can be edited to use all of the features of the gennbo program in the stand-alone mode. Please refer to the NBO manual for further details.

6.12.2 Absorption and Fluorescence Bandshapes using ORCA_ASA

Please also consider using the more recent ORCA_ESD, described in Section *Excited State Dynamics*, to compute bandshapes.

Bandshape calculations are nontrivial but can be achieved with ORCA using the procedures described in section *Simulation and Fit of Vibronic Structure in Electronic Spectra, Resonance Raman Excitation Profiles and Spectra with the orca_asa Program*. Starting from version 2.80, analytical TD-DFT gradients are available, which make these calculations quite fast and applicable without expert knowledge to larger molecules.

Note

- Functionals with somewhat more HF exchange produce better results and are not as prone to "ghost states" as GGA functionals unfortunately are!
- Calculations can be greatly sped up by the RI or RIJCOSX approximations!
- Analytic gradients for the (D) correction and hence for double-hybrid functionals are NOT available.

In a nutshell, let us look into the H_2CO molecule. First we generate some Hessian (e.g. BP86/SV(P)). Then we run the job that makes the input for the orca_asa program. For example, let us calculate the five lowest excited states:

```
! aug-cc-pVDZ BHandHLYP TightSCF NMGrad

%tddft nroots 5 end

# this is ASA-specific input
%rr    states 1,2,3,4,5
       HessName "Test-ASA-H2CO-freq.hess"
       ASAInput True
       end

*int 0 1
 C 0 0 0  0   0     0
 O 1 0 0  1.2 0     0
 H 1 2 0  1.1 120   0
 H 1 2 3  1.1 120 180
*
```

The ORCA run will produce a file Test-ASA-H2CO.asa.inp that is an input file for the program that generates various spectra. It is an ASCII file that is very similar in appearance to an ORCA input file:

[2] Information about the NBO program can be found at http://nbo7.chem.wisc.edu

```
#
# ASA input
#
%sim     model IMDHO
         method Heller

         AbsRange     25000.0,  100000.0
         NAbsPoints   1024

         FlRange      25000.0,  100000.0
         NFlPoints    1024

         RRPRange     5000.0,  100000.0
         NRRPPoints   1024

         RRSRange        0.0,    4000.0
         NRRSPoints   4000

         # Excitation energies (cm**-1) for which rR spectra will
         # be calculated. Here we choose all allowed transitions
         # and the position of the 0-0 band
         RRSE 58960, 66884, 66602

         # full width half maximum of Raman bands in rR spectra
         # (cm**-1):
         RRS_FWHM 10.0

         AbsScaleMode Ext
         FlScaleMode  Rel
         # RamanOrder=1 means only fundamentals. For 2 combination
         # bands and first overtones are also considered, for 3
         # one has second overtones etc.
         RamanOrder   1

         # E0 means the adiabatic excitation energy
         # EV would mean the vertical one. sprints vertical
         # excitations in the TD-DFT output but for the input into
         # the ASA program the adiabatic excitation energies are
         # estimated. A rigorous calculation would of course in-
         # volve excited state geometry optimization
         EnInput E0

         CAR   0.800
         end

# These are the calculated electronic states and transition moments
# Note that this is in the Franck-Condon approximation and thus
# the transition moments have been calculated vertically
$el_states
5
   1 32200.79   100.00    0.00    -0.0000   0.0000  -0.0000
   2 58960.05   100.00    0.00     0.0000  -0.4219   0.0000
   3 66884.30   100.00    0.00    -0.0000   0.4405   0.0000
   4 66602.64   100.00    0.00    -0.5217  -0.0000   0.0000
   5 72245.42   100.00    0.00     0.0000   0.0000   0.0000

# These are the calculated vibrational frequencies for the totally
# symmetric modes. These are the only ones that contribute. They
# correspond to x, H-C-H bending, C=O stretching and C-H stretching
# respectively
$vib_freq_gs
3
```

(continues on next page)

(continued from previous page)

```
    1     1462.948534
    2     1759.538581
    3     2812.815170

# These are the calculated dimensional displacements for all
# electronic states along all of the totally symmetric modes.
$sdnc
3 5
                  1          2          3          4          5
      1      -0.326244   0.241082  -0.132239   0.559635   0.292190
      2      -1.356209   0.529823   0.438703   0.416161   0.602301
      3      -0.183845   0.418242   0.267520   0.278880   0.231340
```

After setting `NAbsPoints` variable and spectral ranges in this file to the desired values, we invoke `orca_asa` as:

```
orca_asa Test-ASA-H2CO.asa.inp
```

This produces the following output:

```
******************
                                  * O R C A  A S A *
                                  ******************

                  --- A program for analysis of electronic spectra ---


Reading file: Test-ASA-H2CO.asa.inp ... done


         ***************************************************************
         *         GENERAL CHARACTERISTICS OF ELECTRONIC SPECTRA        *
         ***************************************************************

-------------------------------------------------------------------------------
  State      E0           EV       fosc    Stokes shift  Effective Stokes shift
         (cm**-1)     (cm**-1)              (cm**-1)           (cm**-1)
-------------------------------------------------------------------------------
    1:     30457.24   32200.79   0.000000      0.00               0.00
    2:     58424.56   58960.05   0.031879      0.00               0.00
    3:     66601.54   66884.30   0.039422      0.00               0.00
    4:     66111.80   66602.64   0.055063      0.00               0.00
    5:     71788.55   72245.42   0.000000      0.00               0.00


-------------------------------------------------------------------------------
↪---
                                BROADENING PARAMETETRS (cm**-1)
-------------------------------------------------------------------------------
↪---
               Intrinsic                              Effective
   State   -------------------------  ---------------------------------------------
                                                Sigma                      FWHM
            Gamma     Sigma     FWHM      ---------------------------  -------------------------
↪--
                                           0K       77K     298.15K      0K       77K     298.
↪15K
-------------------------------------------------------------------------------
↪---
    1:     100.00     0.00   200.00        0.00      0.00      0.00    200.00    200.00    200.
↪00
    2:     100.00     0.00   200.00        0.00      0.00      0.00    200.00    200.00    200.
```

(continues on next page)

(continued from previous page)

```
↪00
    3:     100.00     0.00   200.00        0.00      0.00      0.00    200.00    200.00    200.
↪00
    4:     100.00     0.00   200.00        0.00      0.00      0.00    200.00    200.00    200.
↪00
    5:     100.00     0.00   200.00        0.00      0.00      0.00    200.00    200.00    200.
↪00


Calculating absorption spectrum   ...
The maximum number of grid points ... 5840
Time for absorption               ...       9.569 sec (=   0.159 min)
Writing file: Test-ASA-H2CO.asa.abs.dat ... done
Writing file: Test-ASA-H2CO.asa.abs.as.dat ... done

Generating vibrational states up to the  1-th(st) order  ...   done
Total number of vibrational states                       ...      3

Calculating rR profiles for all vibrational states up to the 1-th order
State 1 ...
The maximum number of grid points ... 6820
Resonance Raman profile is done
State 2 ...
The maximum number of grid points ... 6820
Resonance Raman profile is done
State 3 ...
The maximum number of grid points ... 6820
Resonance Raman profile is done
Writing file: Test-ASA-H2CO.asa.o1.dat... done
Writing file: Test-ASA-H2CO.asa.o1.info... done

Calculating rR spectra involving vibrational states up to the 1-th(st) order
State    1 ... done
State    2 ... done
State    3 ... done

Writing file: Test-ASA-H2CO.asa.o1.rrs.58960.dat      ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.58960.stk      ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.66884.dat      ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.66884.stk      ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.66602.dat      ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.66602.stk      ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.as.58960.dat   ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.as.58960.stk   ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.as.66884.dat   ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.as.66884.stk   ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.as.66602.dat   ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.as.66602.stk   ... done
Writing file: Test-ASA-H2CO.asa.o1.rrs.all.xyz.dat    ... done

TOTAL RUN TIME: 0 days 0 hours 1 minutes 17 seconds 850 msec
```

The computed vibrationally resolved absorption spectrum is plotted as shown in Figure Fig. 6.42.

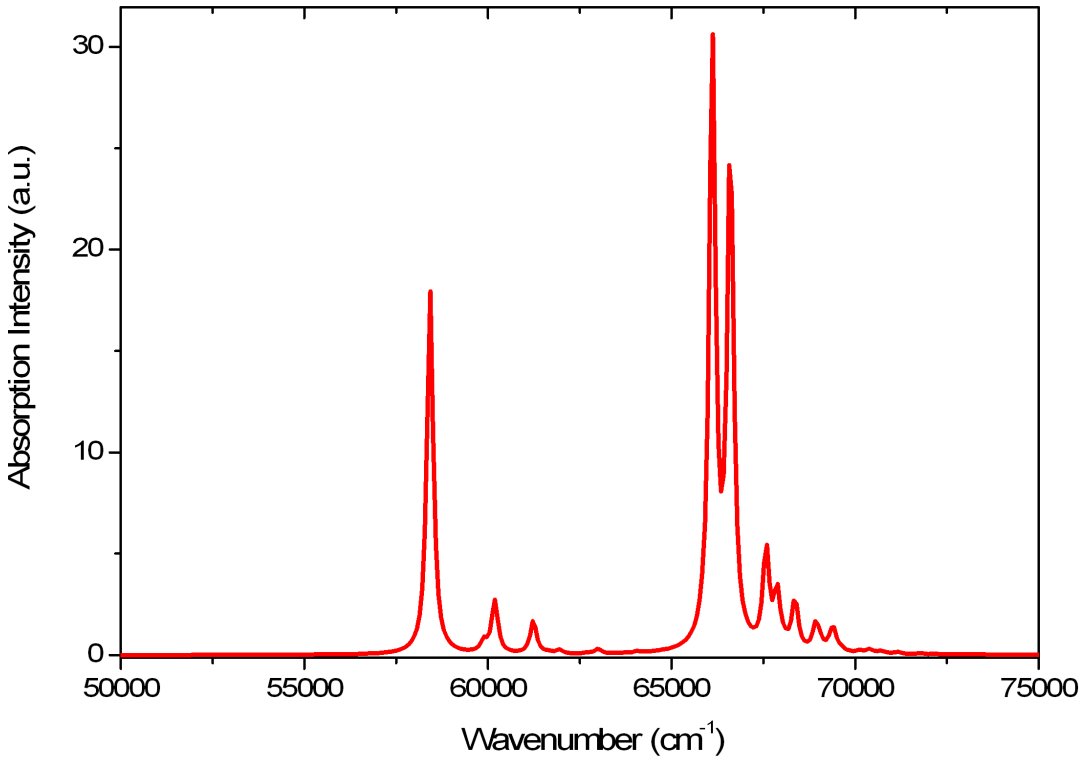


Fig. 6.42: The computed vibrationally resolved absorption spectrum of the H_2CO molecule

The computed fluorescence spectrum of the lowest energy peak is plotted as shown in Figure Fig. 6.43. This peak corresponds to S2. Although it is not realistic, it is sufficient for illustrative purposes.

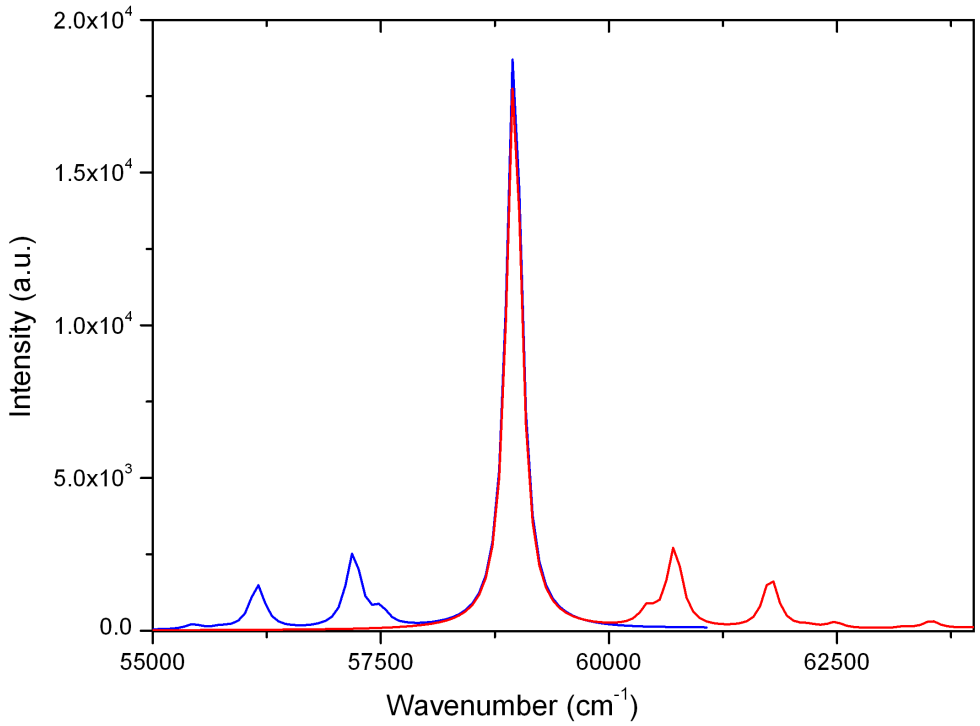


Fig. 6.43: The computed fluorescence spectrum of the lowest energy peak of the H_2CO molecule

The computed Resonance Raman (rR) excitation profiles of the three totally symmetric vibrational modes are plotted as shown in Figure Fig. 6.44.

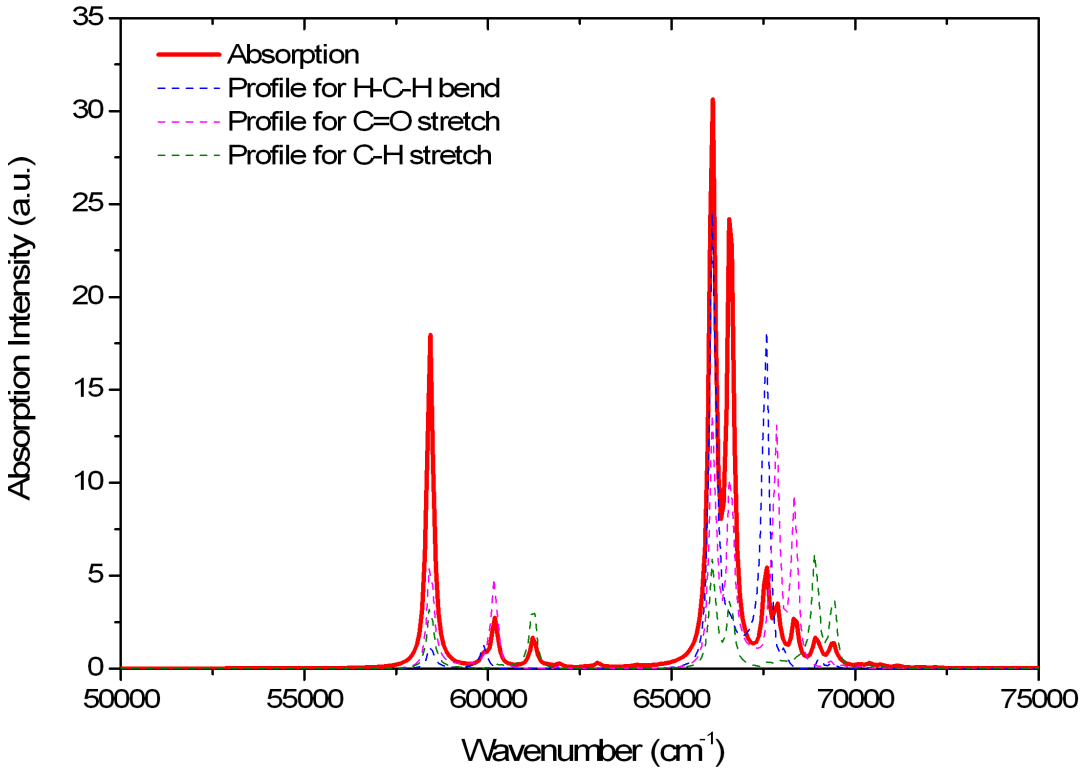


Fig. 6.44: The computed Resonance Raman excitation profiles of the three totally symmetric vibrational modes of the H_2CO molecule

As might be expected, the dominant enhancement occurs under the main peaks for the C=O stretching vibration. Higher energy excitations particularly enhance the C-H vibrations. The computed rR spectra at the vertical excitation energies are provided in Figure Fig. 6.45.

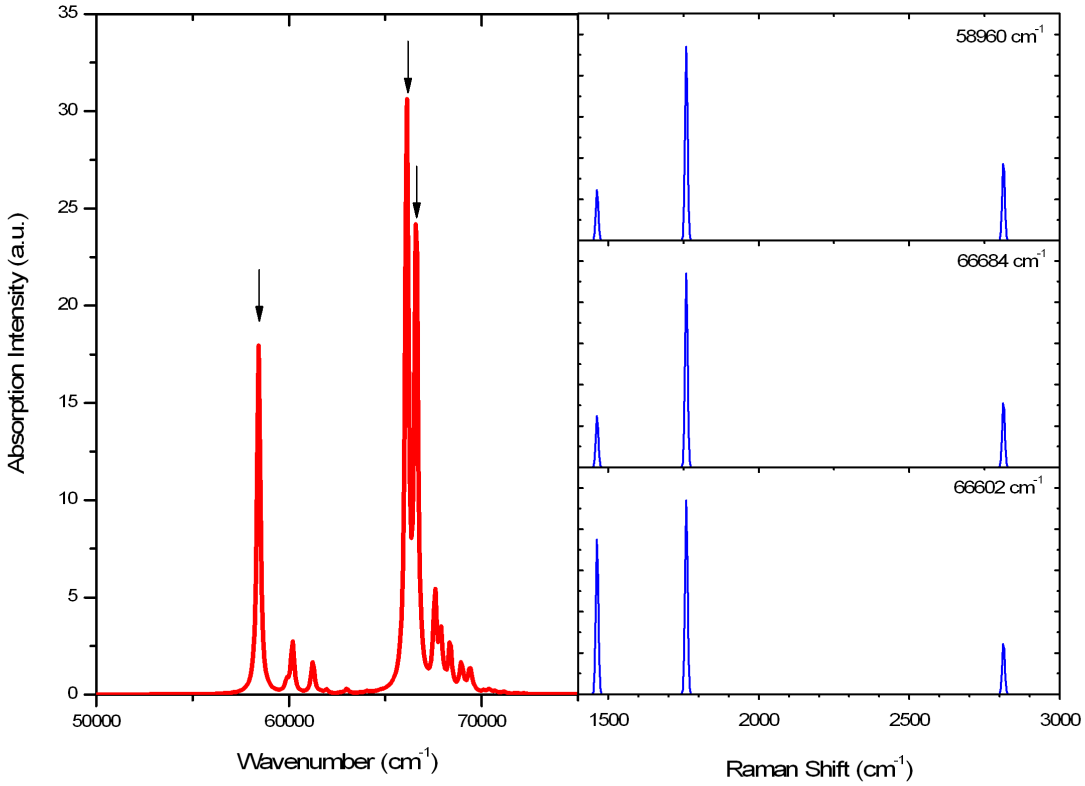


Fig. 6.45: The computed Resonance Raman spectra at the vertical excitation energies of the H_2CO molecule

In this toy example, the dominant mode is the C=O stretching, and the spectra look similar for all excitation wavelengths. However, electronically excited states are mostly of different natures, yielding drastically different rR spectra. Thus, rR spectra serve as powerful fingerprints of the electronic excitation being studied. This is also true even if the vibrational structure of the absorption band is not resolved, which is usually the case for large molecules.

The `orca_asa` program is much more powerful than described in this section. Please refer to section *Simulation and Fit of Vibronic Structure in Electronic Spectra, Resonance Raman Excitation Profiles and Spectra with the orca_asa Program* for a full description of its features. The `orca_asa` program can also be interfaced to other electronic structure codes that deliver excited state gradients and can be used to fit experimental data. It is thus a tool for experimentalists and theoreticians at the same time!

6.12.3 IR/Raman Spectra, Vibrational Modes and Isotope Shifts

IR Spectra

****There were significant changes in the IR printing after ORCA 4.2.1!****

IR spectral intensities are calculated automatically in frequency runs. Thus, there is nothing to control by the user. Consider the following job:

```
! OPT FREQ BP86 def2-SVP

*xyz 0 1
 O    0.000000    0.000000    0.611880
 C    0.000000    0.000000   -0.596849
```

(continues on next page)

(continued from previous page)

```
 H    0.952616    0.000000   -1.209311
 H   -0.952616    0.000000   -1.209311
*
```

which gives the following output:

```
-----------
IR SPECTRUM
-----------

 Mode   freq       eps      Int      T**2         TX        TY        TZ
       cm**-1   L/(mol*cm) km/mol    a.u.
-------------------------------------------------------------------------
  6:   1146.68   0.000341    1.73  0.000093  (-0.000000 -0.009640  0.000000)
  7:   1224.67   0.002004   10.13  0.000511  ( 0.022596  0.000000  0.000000)
  8:   1485.77   0.001002    5.07  0.000211  ( 0.000000 -0.000000  0.014510)
  9:   1806.49   0.020286  102.51  0.003504  ( 0.000000 -0.000000  0.059197)
 10:   2769.13   0.014010   70.80  0.001579  ( 0.000000  0.000000  0.039734)
 11:   2812.52   0.039321  198.71  0.004363  ( 0.066052 -0.000000 -0.000000)
```

The first column ('Mode') labels vibrational modes that increase in frequency from top to bottom." The next column provides vibrational frequencies. The molar absorption coefficient ε of each mode is listed in the "eps" column. This quantity is directly proportional to the intensity of a given fundamental in an IR spectrum, and thus it is used by the `orca_mapspc` utility program as the IR intensity.

The values under "Int" are the integrated absorption coefficient[3], and the "T**2" column lists the norm of the transition dipole derivatives, already including the vibrational part.

To obtain a plot of the spectrum, the `orca_mapspc` utility can be run calling the output file as:

```
orca_mapspc Test-Freq-H2CO.out ir -w25
```

or calling the Hessian file as:

```
orca_mapspc Test-Freq-H2CO.hess ir -w25
```

The basic options of `orca_mapspc` are listed below:

```
-w  : a value for the linewidth (gaussian shape, fwhm)
-x0 : start value of the spectrum in cm**-1
-x1 : end value of the spectrum in cm**-1
-n  : number of points to use
```

To see its options in detail, call `orca_mapspc` without any input. The above `orca_mapspc` runs of the H_2CO molecule provide `Test-NumFreq-H2CO.out.ir.dat` file that contains intensity and wavenumber columns. Therefore, this file can serve as input for any graph plotting program. The plot of the computed IR spectrum of the H_2CO molecule obtained with the above ORCA run is as given in Figure Fig. 6.46.

[3] Explained in more detail by Neugbauer [630]

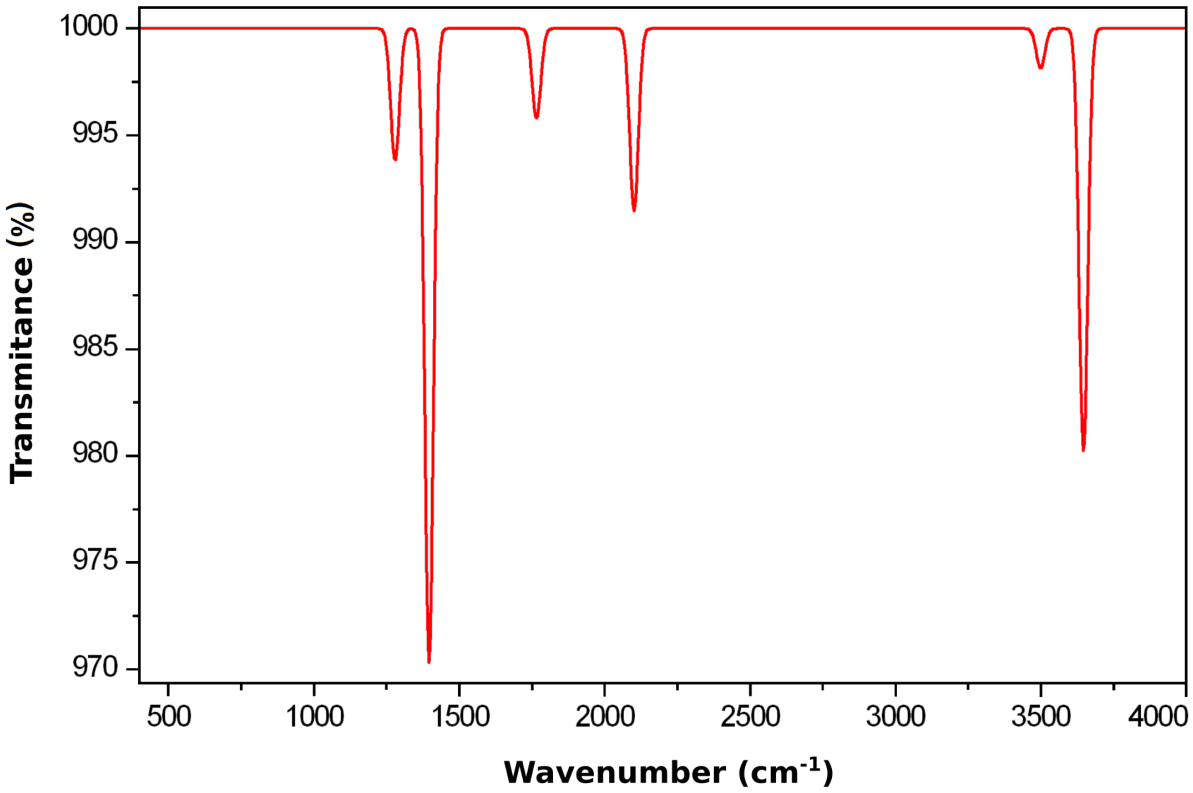


Fig. 6.46: The predicted IR spectrum of the H_2CO molecule plotted using the file generated by the `orca_mapspc` tool.

Overtones, Combination bands and Near IR spectra via NEARIR

Overtones and combination bands can also be incorporated to the computed IR or Near IR spectrum for completeness. The intensities of these bands are strongly dependent on anharmonic effects. ORCA can include these effects by means of the VPT2 approach [77]. The full cubic force field, anharmonic corrections to overtones and combination bands, and a broad range of methods are available in the `orca_vpt2` module (see section *Anharmonic Analysis and Vibrational Corrections using VPT2/GVPT2 and orca_vpt2*).

In particular, the `NEARIR` keyword calls a simpler semidiagonal approach, including only two modes (i and j, also refered as 2MR-QFF in [74, 895]) and force constants up to cubic order (k_{iij}, k_{iji} and k_{iii}). For now, only the intensities are corrected for anharmonic effects - **frequencies are not.**

Overtones and Combination bands

Since the calculation of these terms scale with N^2_{modes}, it can quickly become too expensive, thus we use by default the semiempirical GFN2-xTB [332] to compute the energies and dipole moments necessary to the higher order derivatives (which can be changed later). To request this, simply add `!NEARIR` in the main input. An example input for computing the fundamentals of toluene using B2PLYP double-hybrid functional and for computing the anharmonics using XTB is as follows:

```
! TightOPT NumFREQ RI-B2PLYP def2-TZVP def2-TZVP/C RIJCOSX NEARIR

*xyzfile 0 1 toluene.xyz
```

Note

These anharmonic corrections are very sensitive to the geometry. Therefore, perform a conservative geometry optimization (at least `TightOPT`) whenever possible.

In the output, the characteristics of the regular IR spectrum are printed first. Then, the characteristics of overtones and combination bands are provided similarly to the fundamentals, as follows:

```
-------------------------------
OVERTONES AND COMBINATION BANDS
-------------------------------

Mode      freq       eps       Int      T**2        TX        TY        TZ
         cm**-1   L/(mol*cm)  km/mol    a.u.
-----------------------------------------------------------------------------
6+6:      64.71    0.000994    5.02   0.004792  (-0.009428 -0.066232  0.017796)
6+7:      241.83   0.000022    0.11   0.000028  (-0.005268  0.000255  0.000638)
6+8:      375.36   0.000048    0.24   0.000040  (-0.000740  0.001917  0.006007)
6+9:      442.49   0.000000    0.00   0.000000  ( 0.000010  0.000001  0.000001)
6+10:     506.37   0.000003    0.01   0.000002  ( 0.001078 -0.000061  0.000799)
(...)
```

The "Mode" column shows the overtones, such as 6+6, and combination bands, such as 6+7 and 6+8. These new quantities are automatically detected and incorporated in the IR spectrum when the output file is called with the `orca_mapspc` utility as follows:

```
orca_mapspc toluene-nearir.out ir -w25
```

From the file `orca_mapspc` provided, the IR spectrum can be plotted as shown in Figure Fig. 6.47.

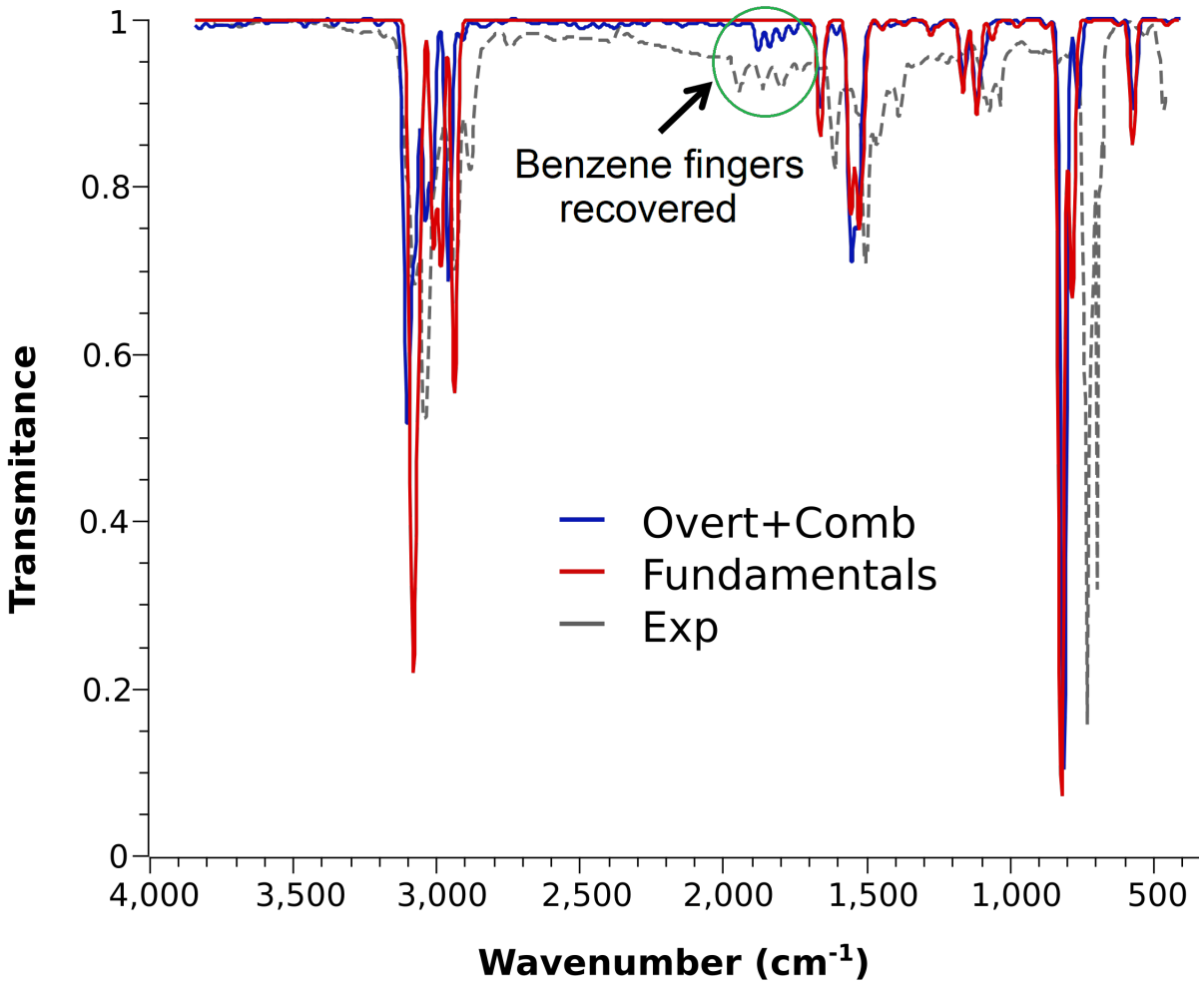


Fig. 6.47: Calculated and experimental infrared spectrum of toluene in gas phase. While the red plot includes only the fundamentals, the blue plot includes also overtones and combination bands. The grey dashed plot is the experimental gas-phase spectrum obtained from the NIST database. The theoretical frequencies were scaled following literature values [442]

"Benzene fingers", i.e., overtones and combination bands of the ring, are recovered in the computed spectrum. Note that the frequencies were scaled using literature values [442], and are not yet corrected using VPT2.

Near IR spectra

Let us simulate near IR spectrum of methanol in CCl_4, as published by Bec and Huck [82], using B3LYP for fundamentals, XTB for overtones, and CPCM for solvation. The input is as follows:

```
! TightOPT FREQ B3LYP def2-TZVP RIJCOSX NEARIR CPCM(CCl4)

*xyz 0 1
 O      0.39517      4.38840     -0.00683
 C     -0.50818      3.29837      0.00221
 H     -0.11943      5.18771      0.19752
 H      0.03977      2.38083     -0.22470
 H     -1.27919      3.45664     -0.75583
 H     -0.96616      3.21170      0.99058
*
```

Calling the output with `orca_mapspc` by setting final point to about $8000 cm^{-1}$ in order to extend the spectrum to the near IR region, i.e.,

```
orca_mapspc toluene-nearir.out ir -w25 -x18000
```

one can simulate the spectrum from the generated "toluene-nearir.dat" file. As seen in Figure Fig. 6.48 the computed spectrum plotted with scaled computational frequencies (not yet corrected using VPT2) according to [442] agrees reasonably well with the experimental spectrum.

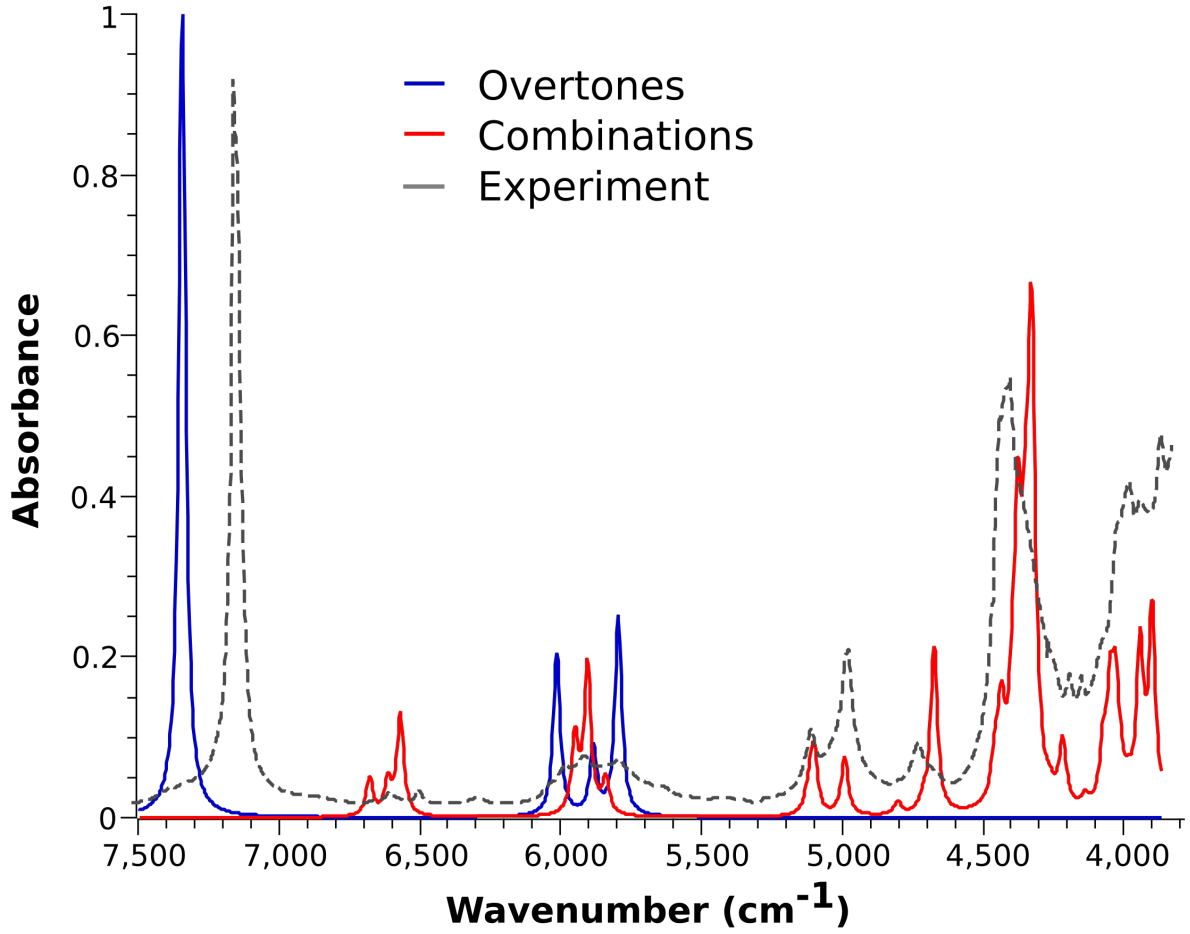


Fig. 6.48: Calculated and experimental near IR spectrum of methanol in CCl_4. The blue plot is for overtones; the red plot is for combination bands; and the grey dashed plot is the experimental spectrum. Theoretical frequencies were scaled according to literature values [442].

Using other methods for the VPT2 correction

To compute overtones with the method chosen for the calculation of the fundamentals, one needs only to set `XTBVPT2` option in the `%freq` block to false, i.e.,

```
%freq XTBVPT2 False end
```

To set a different method for the calculation of overtones and combinations than used for the calculation of fundamentals, one needs first to perform a frequency calculation, then call the resulting Hessian file in `%geom` block, and activate the `PRINTTHERMOCHEM` flag (see section *Thermochemistry* for details), i.e.,

```
! BP86 def2-TZVP NEARIR CPCM(CCl4) PRINTTHERMOCHEM

%geom INHESSNAME "methanol.hess" end
%freq XTBVPT2 False end

*xyzfile 0 1 methanol_opt.xyz
```

In this example, the fundamental modes are read from the "methanol.hess" file, but the anharmonics and intensities of the overtones and combinations are computed using BP86. Any combination of methods, such as B3LYP/BP86 and B2PLYP/AM1, is allowed. Note that this description is an approximation to full VPT2 or GVPT2. For a more complete treatment, see the VPT2 module described in section *Anharmonic Analysis and Vibrational Corrections using VPT2/GVPT2 and orca_vpt2*.

By default, a step size of 0.5 in dimensionless normal mode unit is used during the numerical calculations. This can be changed by setting `DELQ` in the `%freq` block:

```
%freq
    XTBVPT2 False
    DELQ    0.1
end
```

The complete list of options related to VPT2 and in general frequency calculations can be found in Sec. *Frequency calculations - numerical and analytical*.

Vibrational Circular Dichroism (VCD) Spectra

Vibrational circular dichroism spectrum calculations are implemented analytically at the SCF (HF or DFT) level following the derivation of Weigend and coworkers. [715] The basic usage is shown in the following example:

```
# AnFreq + doVCD triggers the VCD calculation
! AnFreq B3LYP def2-SVP

%freq
  doVCD true
end

*xyz 0 1
  C      1.231429   -0.226472   -0.084960
  C     -0.061893    0.507641    0.134338
  C     -1.358912   -0.147897    0.084831
  O     -0.902881    0.641038   -0.969176
  H      1.070541   -1.118875   -0.689778
  H      1.672013   -0.522768    0.869009
  H      1.946503    0.413187   -0.605194
  H      0.017832    1.411161    0.734623
  H     -1.417896   -1.212878   -0.118068
  H     -2.196737    0.255864    0.644375
*
```

Note that in addition to the Hessian, the VCD calculation requires the magnetic field response using GIAOs and the electric field response with the field origin placed at (0,0,0). The latter matches the hard-coded magnetic field gauge origin in the GIAO case and is necessary to ensure gauge-invariance of the results. ORCA does all of this automatically but it means that if VCD is requested together with electric and/or magnetic properties in the same job, the field origins cannot be changed.

Other keywords that influence the VCD calculation include `GIAO_1el` and `GIAO_2el` in `%eprnmr` and `CutOffFreq` in `%freq`. Note also that VCD cannot be computed with `NumFreq`.

Raman Spectra

In order to predict Raman spectrum of a compound, derivatives of the polarizability with respect to the normal modes must be computed. Thus, if a numerical frequency run (`!NumFreq`) is combined with a polarizability calculation, the Raman characteristics will be automatically calculated.

Consider the following example:

```
! OPT NumFreq RHF STO-3G TightSCF SmallPrint

%elprop Polar 1 end

*xyz 0 1
 C    0.000000    0.000000   -0.533905
 O    0.000000    0.000000    0.682807
 H    0.000000    0.926563   -1.129511
 H    0.000000   -0.926563   -1.129511
*
```

The output provides the Raman scattering activity
(in $\AA^4$/AMU)[630] and the Raman depolarization ratio of each mode:

```
--------------
RAMAN SPECTRUM
--------------

 Mode    freq (cm**-1)   Activity   Depolarization
------------------------------------------------------------------
   6:       1277.66      0.010363      0.750000
   7:       1397.45      3.059009      0.750000
   8:       1767.01     16.386535      0.707349
   9:       2099.21      6.701894      0.075708
  10:       3499.49     38.643829      0.186526
  11:       3645.45     24.496534      0.750000
```

The ORCA run generates also a `.hess` file that includes polarizability derivatives and Raman activities. The effect of isotope substitution on the Raman activities can be computed using the `.hess` file.

As in the IR spectrum case, `orca_mapspc` provides a `.dat` file for plotting the computed Raman spectrum:

```
orca_mapspc Test-NumFreq-H2CO.out raman -w50
```

The Raman spectrum of H_2CO plotted by using the corresponding `.dat` file is as given in Figure Fig. 6.49.

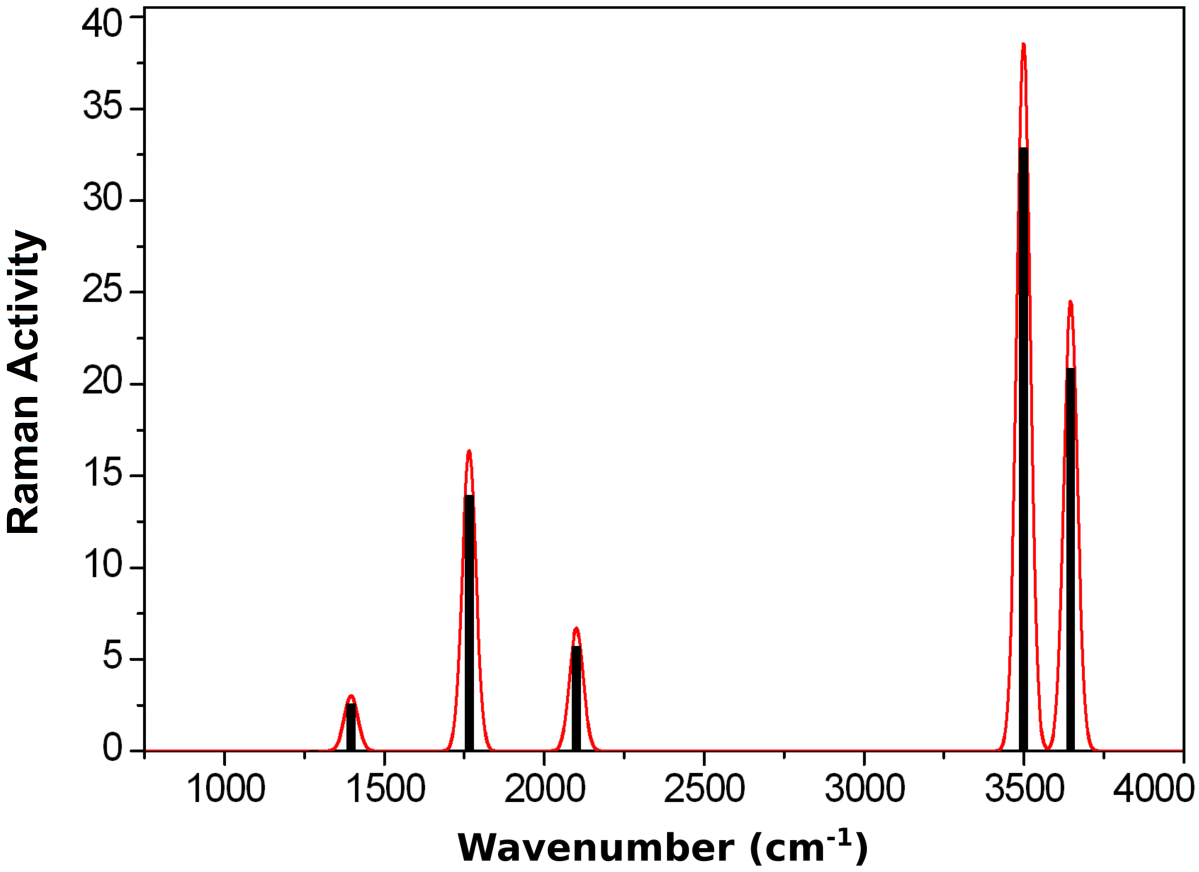


Fig. 6.49: Calculated Raman spectrum of H_2CO at the STO-3G level plotted using the `.dat` generated by the `orca_mapspc` utility from numerical frequencies and Raman activities.

It is worth noting that Raman scattering activity S_i of each mode i is related to but not directly equal to the Raman intensity I_i of the corresponding mode, which is dependent on the excitation line ν_0 of the laser used in the Raman measurement(for Nd:YAG laser: ν_0 = 1064 nm = 9398.5 cm^{-1}). To obtain significantly better agreement between experimental and simulated Raman spectra, I_i of each mode needs to be computed with the following formula:

$$I_i = \frac{f(\nu_0 - \nu_i)^4 S_i}{\nu_i[1 - \exp(-hc\nu_i/kT)]}$$

where f is a normalization constant common for all modes; h, c, k, and T are Planck's constant, speed of light, Boltzmann's constant, and temperature, respectively.

Note

- The Raman module works only when the polarizabilities are calculated analytically. Hence, only the methods, for which the analytical derivatives w.r.t. to external fields are implemeted, can be used.
- Raman calculations take significantly longer than IR calculations due to the extra effort of calculating the polarizabilities at all displaced geometries. Since the latter step is computationally as expensive as the solution of the SCF equations you have to accept an increase in computer time by a factor of ≈ 2.

Resonance Raman Spectra

Resonance Raman spectra (NRVS) and excitation profiles can be predicted or fitted using the procedures described in section *Simulation and Fit of Vibronic Structure in Electronic Spectra, Resonance Raman Excitation Profiles and Spectra with the orca_asa Program*. An example for obtaining the necessary orca_asa input is described in section *Absorption and Fluorescence Bandshapes using ORCA_ASA*.

NRVS Spectra

The details of the theory and implementation of NRVS spectrum are as described in ref. [676, 679]. The NRVS spectrum of $iron-containing\ molecules$ can be simply calculated calling `.hess` file of a previous frequency calculation with the `orca_vib` utility. The output file of this utility can then be called with `orca_mapspc` utility to produce a `.dat` file for plotting the spectrum:

```
orca_vib MyJob.hess > MyJob.vib.out
orca_mapspc MyJob.vib.out NRVS
```

For a the ferric-azide complex [679], the computed and experimental NRVS spectra are provided in Figure Fig. 6.50.

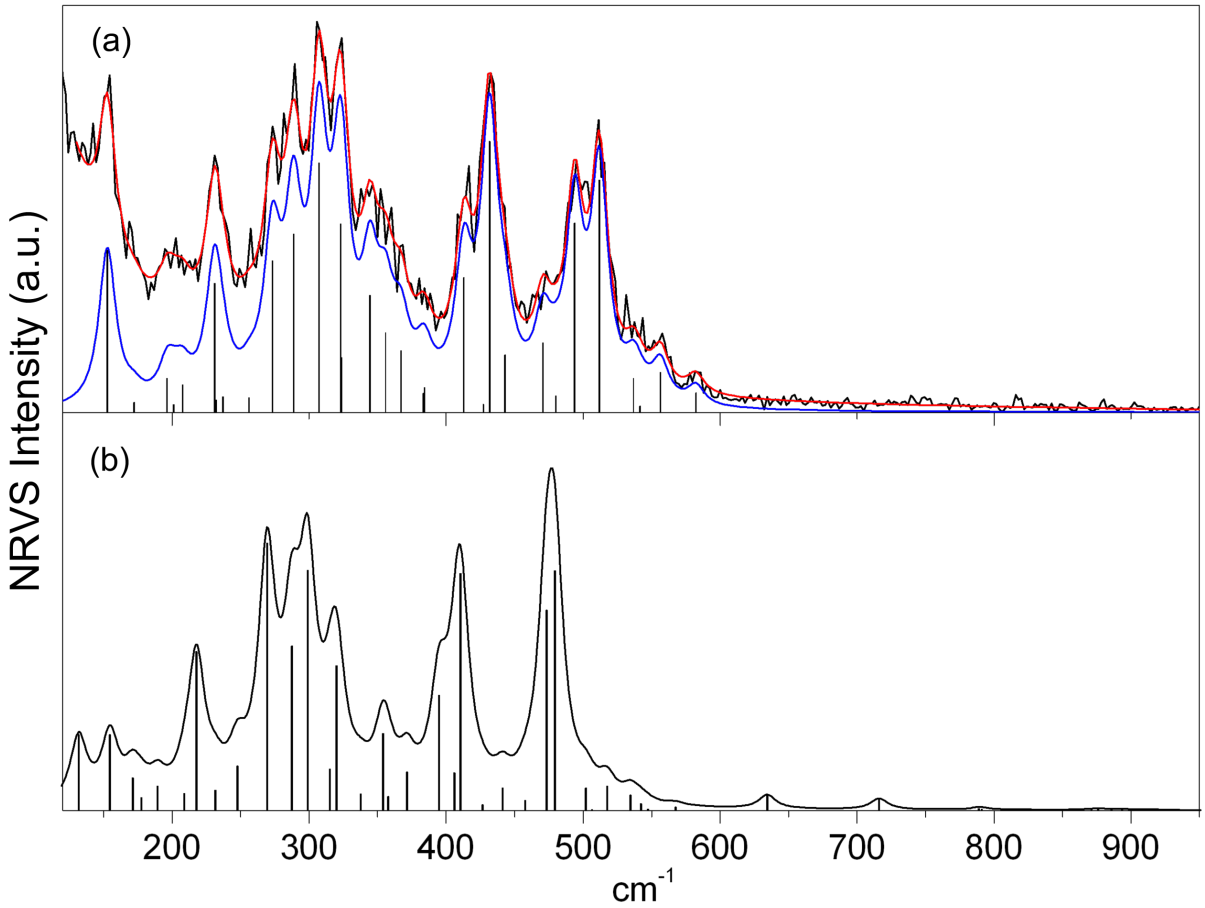


Fig. 6.50: Experimental (a, black curve), fitted (a, red) and simulated (b) NRVS spectrum of the Fe(III)-azide complex obtained at the BP86/TZVP level (T = 20 K). Bar graphs represent the corresponding intensities of the individual vibrational transitions. The blue curve represents the fitted spectrum with a background line removed.

As for the calculation of resonance Raman spectra described in section *Simulation and Fit of Vibronic Structure in Electronic Spectra, Resonance Raman Excitation Profiles and Spectra with the orca_asa Program*, the DFT estimations are usually excellent starting points for least-square refinements.

Below we describe the procedure for computing such NRVS spectra on the $Fe(SH)_4^{1-}$ complex with the BP86 functional, which typically provides good NRVS spectra. One needs first to optimize the geometry of the complex and compute its vibrational structure:

```
! OPT FREQ BP86 def2-TZVP TightSCF SmallPrint

*xyz -1 6
 Fe     -0.115452      0.019090     -0.059506
 S      -0.115452      1.781846      1.465006
 S      -0.115452     -1.743665      1.462801
 S      -1.908178     -0.072782     -1.518702
 S       1.560523      0.154286     -1.656664
 H       0.410700      2.760449      0.687716
 H      -0.674147     -2.708278      0.690223
 H      -2.905212      0.345589     -0.699907
 H       2.647892     -0.211681     -0.932926
*
```

Now run the `orca_vib` utility on the `.hess` file generated by this job to obtain an output file that can be used with `orca_mapspc` utility:

```
orca_vib Test-FeIIISH4-NumFreq.hess > Test-FeIIISH4-NumFreq.out
orca_mapspc Test-FeIIISH4-NumFreq.out NRVS
```

This `orca_mapspc` run generates `Test-FeIIISH4-NumFreq.nrvs.dat` file in the xy-format. This file contains phonon energy (x, in cm^{-1}) and NRVS intensity (y, in atomic units) and thus can be directly used for visualizing the spectrum.

The corresponding NRVS spectrum is given in Figure Fig. 6.51 together with the computational IR spectrum on the same frequency scale. NRVS reports the Doppler broadening of the Moessbauer signal due to resonant scattering of phonons (vibrations) dominated by the movements of Fe nuclei. This is a valuable addition to IR spectrum where the modes have very small intensities.

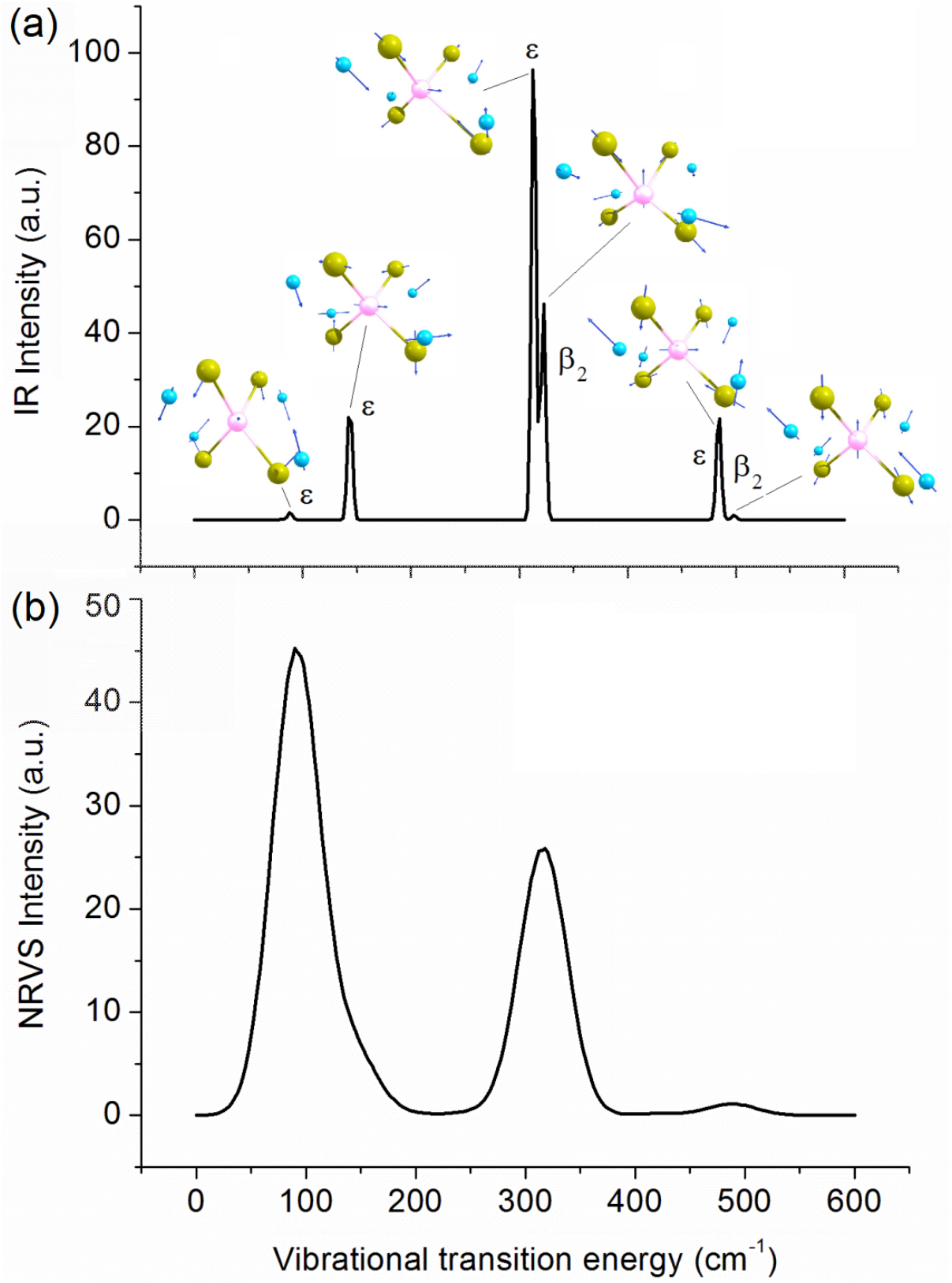


Fig. 6.51: (a) Theoretical IR spectrum of $Fe(SH)_4^{1-}$ and arrow-pictures of the highest intensity modes around the peak maxima. (b) The corresponding NRVS scattering spectrum.

Animation of Vibrational Modes

For describing how to animate vibrational modes and generate their "arrow-pictures", let us perform a frequency calculation on H_2CO:

```
! OPT FREQ RHF STO-3G

*xyz 0 1
 C    0.000000    0.000000   -0.533905
 O    0.000000    0.000000    0.682807
 H    0.000000    0.926563   -1.129511
 H    0.000000   -0.926563   -1.129511
*
```

The output of this job provides vibrational characteristics:

```
Mode      freq       eps       Int    T**2         TX        TY        TZ
         cm**-1   L/(mol*cm) km/mol    a.u.
-------------------------------------------------------------------------
6:       1278.37   0.001222    6.18  0.000298  (-0.017272  0.000000  0.000000)
7:       1397.26   0.005844   29.53  0.001305  ( 0.000000  0.036128  0.000000)
8:       1767.02   0.000828    4.18  0.000146  (-0.000000  0.000000 -0.012089)
9:       2099.24   0.001668    8.43  0.000248  ( 0.000000 -0.000000  0.015749)
10:      3498.54   0.000356    1.80  0.000032  ( 0.000000 -0.000000 -0.005636)
11:      3645.47   0.003922   19.82  0.000336  (-0.000000  0.018322  0.000000)
```

This output can be directly opened with `ChemCraft` to visualize normal modes of H_2CO and to extract their arrow-pictures representing the direction of nuclear movements as shown in Figure Fig. 6.52. As an example, one can infer from this figure that the 1397 cm^{-1} mode corresponds to a rocking vibration.

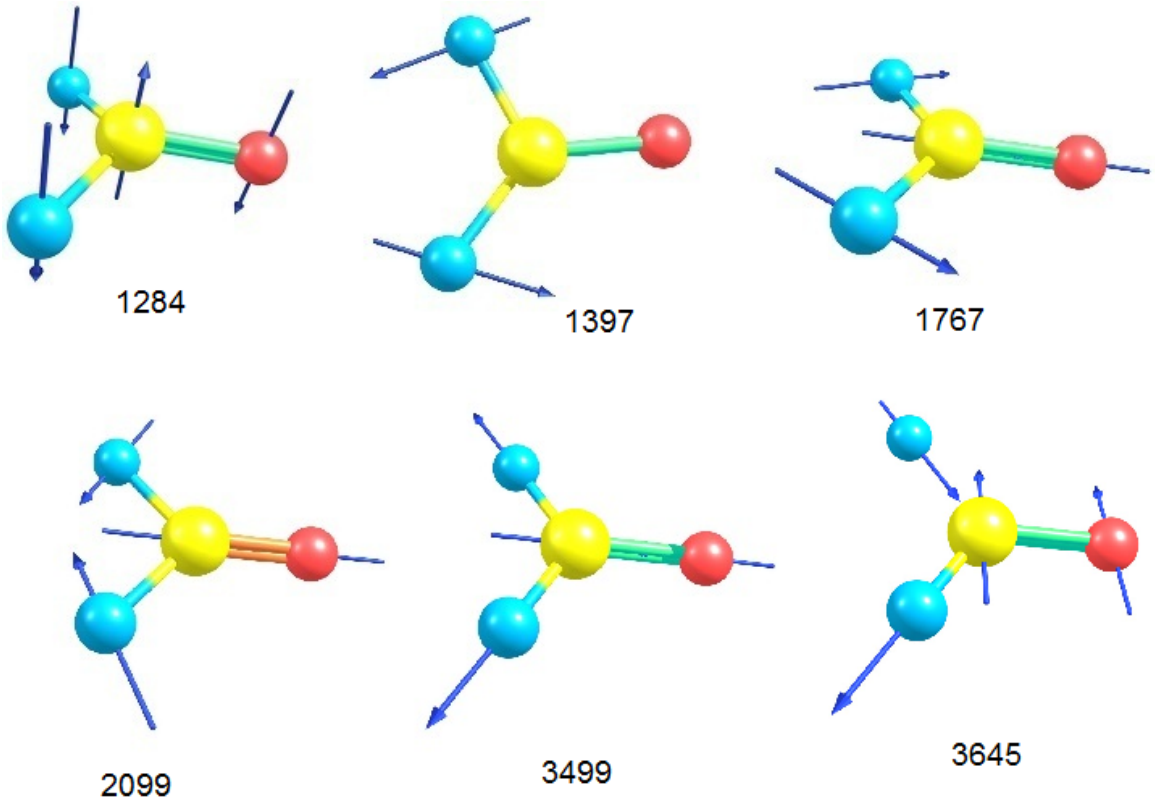


Fig. 6.52: Normal modes of H_2CO with arrows indicating magnitude and direction of nuclear motions and the associated vibrational frequencies in cm^{-1} obtained from `ORCA` output file through the use of `ChemCraft`

In order to animate vibrational modes and to create their "arrow-pictures" by using free program packages like gOpenMol, the small utility program orca_pltvib can be used. This utility program generates a series of files from an ORCA output file of a frequency run, which can be openned with molecular visualization programs. The usage of orca_pltvib is as follows:

```
orca_pltvib Test-FREQ-H2CO.out [list of vibrations or all]
```

For example, let us want to animate the 1397 cm^{-1} mode labeled as 7:

```
orca_pltvib Test-FREQ-H2CO.out 7
```

This call will generate the Test-FREQ-H2CO.out.v007.xyz file. Open gOpenMol and read this file (Import->coords) in Xmol format. Then, go to the Trajectory->Main menu and import again the file in Xmol format. Now you are able to animate the mode. In order to generate a printable picture, press Dismiss and then type lulVectorDemo {4 0.1 black} into the gOpenMol command line window. The generated picture (see Figure Fig. 6.53) demonstrates that this mode corresponds to a rocking vibration.

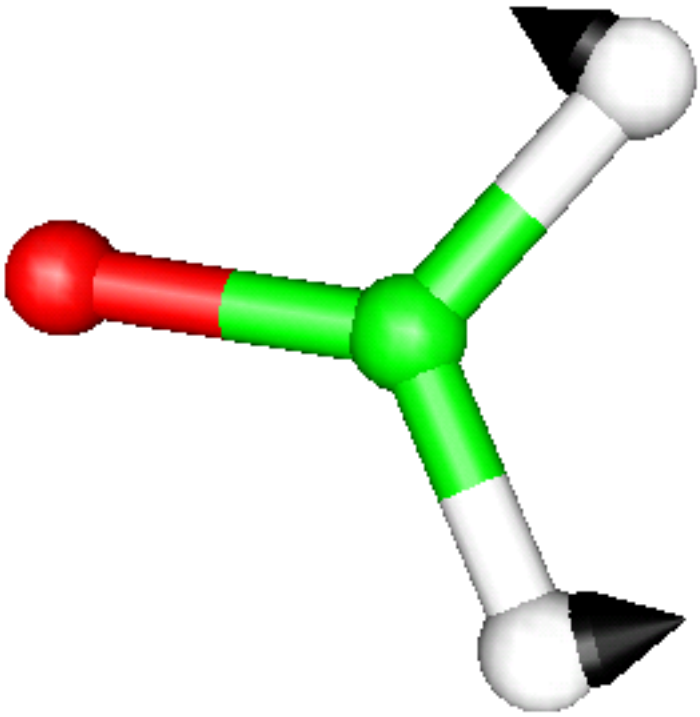


Fig. 6.53: The 1397 cm^{-1} mode of the H_2CO molecule as obtained from the interface of ORCA to gOpenMol and the orca_pltvib tool to create the animation file.

The appearence of the arrows can also be modified as described in the web tutorial of gOpenMol.

Isotope Shifts

The calculated isotope shifts greatly aid in the identification of vibrations, the interpretation of experiments, and the assessment of the reliability of the calculated vibrational normal modes. It would be a very bad practice to recalculate the Hessian for investigating isotope shift since Hessian calculations are typically expensive, and the Hessian itself is independent of the masses. Below we describe how to find the isotope effect without recomputing the Hessian.

Let us suppose that you have calculated a Hessian as in the example discussed above, and you want to predict the effect of ^{18}O substitution. In this case you can use the small utility program `orca_vib`. First of all you need to edit the masses given in the `.hess` file by hand. For the example given above, the `.hess` file is as follows:

```
$orca_hessian_file
.....................
$hessian
12
... the cartesian Hessian in Eh/bohr**"

$vibrational_frequencies
12

...the vibrational frequencies (in cm-1) as in the output

$normal_modes
12 12
... the vibrational normal modes in Cartesian displacements
#
# The atoms: label  mass x  y  z
# !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
#   Here we have changed 15.999 for oxygen into
#   18.0  in order to see the oxygen 18 effects
# !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
$atoms
4
 C     12.0110      0.000000     0.000000    -1.149571
 O     18.0000     -0.000000    -0.000000     1.149695
 H      1.0080     -0.000000     1.750696    -2.275041
 H      1.0080     -0.000000    -1.750696    -2.275041

$actual_temperature
  0.000000

$dipole_derivatives
12

... the dipole derivatives (Cartesian displacements)
#
# The IR spectrum
#  wavenumber T**2 TX TY  TY
#
$ir_spectrum
12
... the IR intensities
```

After changing the mass of O from 15.999 to 18.0 as shown above, let us call:

```
orca_vib Test-FREQ-H2CO.hess
```

This will recompute vibrational frequencies in the presence of ^{18}O. Let us compare vibrational frequencies in the output of this run with the original frequencies in cm^{-1}:

```
  Mode       H2C16O    H2C18O    Shift
-----------------------------------------
   6:        1284.36   1282.82    -1.54
   7:        1397.40   1391.74    -5.66
   8:        1766.60   1751.62   -14.98
   9:        2099.20   2061.49   -37.71
  10:        3499.11   3499.02    -0.09
  11:        3645.24   3645.24     0.00
```

Another way to analyze isotope shifts is to plot the two predicted spectra and then subtract one from the other. This will produce derivative-shaped peaks with zero crossings at the positions of the isotope-sensitive modes.

Note

In the presence of point charges and/or an external electric field, the translational and rotational symmetries of the system may be broken. In such cases, you may prefer NOT to project out the translational and rotational degrees of freedom of the Hessian. This can be achieved as:

```
orca_vib Test-FREQ-H2CO.hess -noproj
```

6.12.4 Thermochemistry

The second thing that you get automatically as the result of a frequency calculation is a thermochemical analysis based on ideal gas statistical mechanics. This can be used to study heats of formation, dissociation energies and similar thermochemical properties. To correct for the breakdown of the harmonic oscillator approximation for low frequencies, entropic contributions to the free energies are computed, by default, using the Quasi-RRHO approach of Grimme.[322] To switch-off the Quasi-RRHO method and use the RRHO method, use:

```
%freq  QuasiRRHO false
       CutOffFreq 35 # in cm-1
       end
```

Where the CutOffFreq parameter controls the cut-off for the low frequencies mode (excluded from the calculation of the thermochemical properties). Note that the default CutOffFreq is 1 (cm^{-1}) when Quasi-RRHO is on, since Quasi-RRHO behaves much more reasonably for low frequencies than RRHO does. In particular, the entropy contribution calculated by Quasi-RRHO approaches a constant value when the vibrational frequency approaches zero, while the RRHO contribution diverges.

The Quasi-RRHO method smoothly interpolates between the entropy formulas of a harmonic oscillator and a rigid rotor, such that high frequency vibrations behave like harmonic vibrations, and low frequency vibrations behave like rotations with the same frequency. The frequency at which the entropy contribution is a half-half mixture of rotation and vibration is called the "reference frequency" ω_0 of the Quasi-RRHO method, accessible via the `QRRHORefFreq` keyword in `%freq` (see *Frequency calculations - numerical and analytical*). The default value (100 cm^{-1}) is consistent with the original Quasi-RRHO paper[322], but other papers may choose different values, such as 50 cm^{-1}. Meanwhile, ORCA's Quasi-RRHO implementation deviates from the original paper in the choice of "average molecular moment of inertia" B_{av}; while in the original paper it is chosen as a fixed value $10^{-44}\mathrm{kg}\cdot\mathrm{m}^2$, in ORCA it is given as the isotropically averaged moment of inertia of the actual molecule at hand. This is theoretically more justified than using the same moment of inertia for molecules of different sizes, although the resulting difference in the Gibbs free energies is rather small, usually within 0.1 kcal/mol.

Note that the rotational contribution to the entropy is calculated using the expressions given by Herzberg [388] including the symmetry number obtained from the order of the point group.[4] While this is a good approximation,

[4] the corresponding equation for the partition function (assuming sufficiently high temperatures) of a linear molecule is $Q_{int} = \frac{kT}{\sigma hcB}$ and for non-linear molecules $Q_{int} = \frac{1}{\sigma}\sqrt{\frac{\pi}{ABC}(\frac{kT}{hc})^3}$. A, B and C are the corresponding rotational constants, σ is the symmetry number. If you want to choose a different symmetry number, ORCA also provides a table with the values for this entropy contribution for other symmetry numbers. Herzberg reports the following symmetry numbers for the point groups C_1,C_i,C_s: 1; C_2,C_{2v}, C_{2h}: 2; C_3,C_{3v},C_{3h}: 3; C_4,C_{4v},C_{4h}: 4;C_6, C_{6v}, C_{6h}: 6; D_2, D_{2d}, D_{2h} =V_h: 4; D_3, D_{3d}, D_{3h}: 6; D_4, D_{4d}, D_{4h}: 8; D_6, D_{6d}, D_{6h}: 12; S_6: 3; $C_{\infty v}$: 1; $D_{\infty h}$: 2;T,T_d: 12; O_h: 24.

one might want to modify the symmetry number or use a different expression [302]. For this purpose, the rotational constants (in cm^{-1}) of the molecule are also given in the thermochemistry output.

For example let us calculate a number for the oxygen-oxygen dissociation energy in the H_2O_2 molecule. First run the following jobs:

```
# Calculate a value for the O-O bond strength in H2O2
! B3LYP DEF2-TZVP OPT FREQ BOHRS
* xyz 0 1
 O          -1.396288    -0.075107     0.052125
 O           1.396289    -0.016261    -0.089970
 H          -1.775703     1.309756    -1.111179
 H           1.775687     0.140443     1.711854
*
```

```
# Now the OH radical job
! B3LYP DEF2-TZVP OPT FREQ BOHRS
* xyz 0 2
 O          -1.396288    -0.075107     0.052125
 H          -1.775703     1.309756    -1.111179
*
```

The first job gives you the following output following the frequency calculation:

```
---------------------------
THERMOCHEMISTRY AT 298.15K
---------------------------

Temperature         ... 298.15 K
Pressure            ... 1.00 atm
Total Mass          ... 34.01 AMU

Throughout the following assumptions are being made:
  (1) The electronic state is orbitally nondegenerate
  (2) There are no thermally accessible electronically excited states
  (3) Hindered rotations indicated by low frequency modes are not
      treated as such but are treated as vibrations and this may
      cause some error
  (4) All equations used are the standard statistical mechanics
      equations for an ideal gas
  (5) All vibrations are strictly harmonic

freq.     370.67  E(vib)   ...       0.21
freq.     947.27  E(vib)   ...       0.03
freq.    1313.46  E(vib)   ...       0.01
freq.    1440.24  E(vib)   ...       0.00
freq.    3739.49  E(vib)   ...       0.00
freq.    3739.86  E(vib)   ...       0.00

------------
INNER ENERGY
------------

The inner energy is: U= E(el) + E(ZPE) + E(vib) + E(rot) + E(trans)
    E(el)   - is the total energy from the electronic structure calculation
              = E(kin-el) + E(nuc-el) + E(el-el) + E(nuc-nuc)
    E(ZPE)  - the zero temperature vibrational energy from the frequency calculation
    E(vib)  - the finite temperature correction to E(ZPE) due to population
              of excited vibrational states
    E(rot)  - is the rotational thermal energy
    E(trans)- is the translational thermal energy
```

(continues on next page)

(continued from previous page)

```
Summary of contributions to the inner energy U:
Electronic energy                ...   -151.55083691 Eh
Zero point energy                ...      0.02631509 Eh      16.51 kcal/mol
Thermal vibrational correction   ...      0.00040105 Eh       0.25 kcal/mol
Thermal rotational correction    ...      0.00141627 Eh       0.89 kcal/mol
Thermal translational correction ...      0.00141627 Eh       0.89 kcal/mol
-----------------------------------------------------------------------
Total thermal energy                    -151.52128823 Eh


Summary of corrections to the electronic energy:
(perhaps to be used in another calculation)
Total thermal correction                 0.00323359 Eh       2.03 kcal/mol
Non-thermal (ZPE) correction             0.02631509 Eh      16.51 kcal/mol
-----------------------------------------------------------------------
Total correction                         0.02954868 Eh      18.54 kcal/mol


--------
ENTHALPY
--------

The enthalpy is H = U + kB*T
               kB is Boltzmann's constant
Total free energy                ...   -151.52129054 Eh
Thermal Enthalpy correction      ...      0.00094421 Eh       0.59 kcal/mol
-----------------------------------------------------------------------
Total Enthalpy                   ...   -151.52034633 Eh


Note: Rotational entropy computed according to Herzberg
Infrared and Raman Spectra, Chapter V,1, Van Nostrand Reinhold, 1945
Point Group:  C2, Symmetry Number:   2
Rotational constants in cm-1:    10.087644     0.882994     0.851333

Vibrational entropy computed according to the QRRHO of S. Grimme
Chem.Eur.J. 2012 18 9955


-------
ENTROPY
-------

The entropy contributions are T*S = T*(S(el)+S(vib)+S(rot)+S(trans))
     S(el)   - electronic entropy
     S(vib)  - vibrational entropy
     S(rot)  - rotational entropy
     S(trans)- translational entropy
The entropies will be listed as multiplied by the temperature to get
units of energy

Electronic entropy               ...      0.00000000 Eh       0.00 kcal/mol
Vibrational entropy              ...      0.00059250 Eh       0.37 kcal/mol
Rotational entropy               ...      0.00789993 Eh       4.96 kcal/mol
Translational entropy            ...      0.01734394 Eh      10.88 kcal/mol
-----------------------------------------------------------------------
Final entropy term               ...      0.02583637 Eh      16.21 kcal/mol

In case the symmetry of your molecule has not been determined correctly
or in case you have a reason to use a different symmetry number we print
out the resulting rotational entropy values for sn=1,12 :
```

(continues on next page)

(continued from previous page)

```
 ---------------------------------------------------
|  sn= 1 | S(rot)=       0.00855439 Eh      5.37 kcal/mol|
|  sn= 2 | S(rot)=       0.00789993 Eh      4.96 kcal/mol|
|  sn= 3 | S(rot)=       0.00751710 Eh      4.72 kcal/mol|
|  sn= 4 | S(rot)=       0.00724548 Eh      4.55 kcal/mol|
|  sn= 5 | S(rot)=       0.00703479 Eh      4.41 kcal/mol|
|  sn= 6 | S(rot)=       0.00686265 Eh      4.31 kcal/mol|
|  sn= 7 | S(rot)=       0.00671710 Eh      4.22 kcal/mol|
|  sn= 8 | S(rot)=       0.00659102 Eh      4.14 kcal/mol|
|  sn= 9 | S(rot)=       0.00647981 Eh      4.07 kcal/mol|
|  sn=10 | S(rot)=       0.00638033 Eh      4.00 kcal/mol|
|  sn=11 | S(rot)=       0.00629034 Eh      3.95 kcal/mol|
|  sn=12 | S(rot)=       0.00620819 Eh      3.90 kcal/mol|
 ---------------------------------------------------



-------------------
GIBBS FREE ENERGY
-------------------

The Gibbs free energy is G = H - T*S

Total enthalpy                    ...   -151.52034633 Eh
Total entropy correction          ...     -0.02583637 Eh    -16.21 kcal/mol
-----------------------------------------------------------------------
Final Gibbs free energy         ...   -151.54618270 Eh

For completeness - the Gibbs free energy minus the electronic energy
G-E(el)                           ...      0.00465413 Eh      2.92 kcal/mol
```

And similarly for the OH-radical job.

```
------------
INNER ENERGY
------------

The inner energy is: U= E(el) + E(ZPE) + E(vib) + E(rot) + E(trans)
    E(el)   - is the total energy from the electronic structure calculation
              = E(kin-el) + E(nuc-el) + E(el-el) + E(nuc-nuc)
    E(ZPE)  - the zero temperature vibrational energy from the frequency calculation
    E(vib)  - the finite temperature correction to E(ZPE) due to population
              of excited vibrational states
    E(rot)  - is the rotational thermal energy
    E(trans)- is the translational thermal energy

Summary of contributions to the inner energy U:
Electronic energy                 ...    -75.73492538 Eh
Zero point energy                 ...      0.00837287 Eh       5.25 kcal/mol
Thermal vibrational correction    ...      0.00000000 Eh       0.00 kcal/mol
Thermal rotational correction     ...      0.00094418 Eh       0.59 kcal/mol
Thermal translational correction ...      0.00141627 Eh       0.89 kcal/mol
-----------------------------------------------------------------------
Total thermal energy                   -75.72419205 Eh


Summary of corrections to the electronic energy:
(perhaps to be used in another calculation)
Total thermal correction                0.00236045 Eh       1.48 kcal/mol
Non-thermal (ZPE) correction            0.00837287 Eh       5.25 kcal/mol
-----------------------------------------------------------------------
Total correction                        0.01073332 Eh       6.74 kcal/mol
```

(continues on next page)

(continued from previous page)

```
--------
ENTHALPY
--------

The enthalpy is H = U + kB*T
               kB is Boltzmann's constant
Total free energy                  ...    -75.72419205 Eh
Thermal Enthalpy correction        ...      0.00094421 Eh       0.59 kcal/mol
-------------------------------------------------------------------------
Total Enthalpy                     ...    -75.72324785 Eh


Note: Rotational entropy computed according to Herzberg
Infrared and Raman Spectra, Chapter V,1, Van Nostrand Reinhold, 1945
Point Group:  C2v, Symmetry Number:   1
Rotational constants in cm-1:     0.000000    18.628159    18.628159

Vibrational entropy computed according to the QRRHO of S. Grimme
Chem.Eur.J. 2012 18 9955


-------
ENTROPY
-------

The entropy contributions are T*S = T*(S(el)+S(vib)+S(rot)+S(trans))
     S(el)   - electronic entropy
     S(vib)  - vibrational entropy
     S(rot)  - rotational entropy
     S(trans)- translational entropy
The entropies will be listed as multiplied by the temperature to get
units of energy


Note: Rotational entropy computed according to Herzberg
Infrared and Raman Spectra, Chapter V,1, Van Nostrand Reinhold, 1945
Point Group:  C2v, Symmetry Number:   1
Rotational constants in cm-1:     0.000000    18.628159    18.628159

Vibrational entropy computed according to the QRRHO of S. Grimme
Chem.Eur.J. 2012 18 9955


-------
ENTROPY
-------

The entropy contributions are T*S = T*(S(el)+S(vib)+S(rot)+S(trans))
     S(el)   - electronic entropy
     S(vib)  - vibrational entropy
     S(rot)  - rotational entropy
     S(trans)- translational entropy
The entropies will be listed as multiplied by the temperature to get
units of energy

Electronic entropy                 ...      0.00065446 Eh      0.41 kcal/mol
Vibrational entropy                ...      0.00000000 Eh      0.00 kcal/mol
Rotational entropy                 ...      0.00321884 Eh      2.02 kcal/mol
Translational entropy              ...      0.01636225 Eh     10.27 kcal/mol
```

(continues on next page)

(continued from previous page)

```
-------------------------------------------------------------------------
Final entropy term                 ...      0.02023555 Eh     12.70 kcal/mol


-------------------
GIBBS FREE ENERGY
-------------------

The Gibbs free energy is G = H - T*S

Total enthalpy                     ...    -75.72324785 Eh
Total entropy correction           ...     -0.02023555 Eh    -12.70 kcal/mol
-------------------------------------------------------------------------
Final Gibbs free energy         ...    -75.74348340 Eh

For completeness - the Gibbs free energy minus the electronic energy
G-E(el)                            ...     -0.00855802 Eh     -5.37 kcal/mol
```

Let us calculate the free energy change for the reaction: $H_2O_2 \rightarrow 2OH$

The individual energy terms are:

Electronic Energy: -151.46985 a.u. -(-151.55084) a.u. = 0.08099 a.u. (50.82 kcal/mol)

Zero-point Energy: 0.01675 a.u. - 0.02631 a.u. = -0.00956 a.u. (-6.00 kcal/mol)

Thermal Correction(translation/rotation): 0.00472 a.u. - 0.00283 a.u. = 0.00189 a.u. (1.19 kcal/mol)

Thermal Enthalpy Correction: 0.00189 a.u. - 0.00094 a.u. = 0.00095 a.u. (0.60 kcal/mol)

Entropy: -0.04047 a.u. -(-0.02584) a.u. = -0.01463 a.u. (-9.18 kcal/mol)

Final ΔG: 37.43 kcal/mol

Thus, both the zero-point energy and the entropy terms contribute significantly to the total free energy change of the reaction. The entropy term is favoring the reaction due to the emergence of new translational and rotational degrees of freedom. The zero-point correction is also favoring the reaction since the zero-point vibrational energy of the O-O bond is lost. The thermal correction and the enthalpy correction are both small.

Tip

- You can run the thermochemistry calculations at several user defined temperatures and pressure by providing the program with a list of temperatures / pressures:

```
%freq  Temp  290, 295, 300     # in Kelvin
        Pressure  1.0, 2.0, 3.0 # in atm
       end
```

- Once a Hessian is available you can rerun the thermochemistry analysis at several user defined temperatures / pressures by providing the keyword PrintThermoChem and providing the name of the Hessian file:

```
! PrintThermoChem
%geom
  inhessname "FreqJob.hess" # default: job-basename.hess
end
%freq  Temp  290, 295, 300     # in Kelvin
        Pressure  1.0, 2.0, 3.0 # in atm
       end
```

6.12.5 Anharmonic Analysis and Vibrational Corrections using VPT2/GVPT2 and orca_vpt2

Building upon (analytical) harmonic calculations of the Hessian, it is possible to calculate the cubic plus semi-quartic force field as well as higher-order property derivatives. For this purpose, the VPT2 module will compute the Hessian and then generate two displaced geometries for each degree of freedom and for each displacement another Hessian (and another property in case of vibrational corrections) will be computed. These are required for an anharmonic analysis according to second-order vibrational perturbation theory. So overall, using VPT2 is costly due to the number of calculations required for the numerical derivatives and is very sensitive to numerical noise due to convergence, approximations and other settings. The VPT2 calculation can be initiated either via the simple input command `!VPT2` or via the VPT2 keyword in the `%vpt2` block. Finer control can be achieved through the `%VPT2` block, as exemplified in this analysis of water.

```
# VPT2 Analysis of H2O
!RHF def2-SVP ExtremeSCF VPT2

%vpt2
    VPT2                On      # do a VPT2 analysis, same as !VPT2 (see above)
    AnharmDisp          0.05    # anharmonic displacement factor, default is 0.05
    HessianCutoff       1e-12   # cut-off for Hessian elements, default is $10^{-10}$
    PrintLevel          1       # VPT2 print level [1, 2, 3, 4]
    MinimiseOrcaPrint   True    # Minimises the remaining orca output
end

%method
    Z_Tol               1e-14
end

* xyz 0 1
  O   0.00000000000000      0.06256176106279      0.06256176106280
  H   0.00000000000000     -0.06185639479702      0.99929463373422
  H   0.00000000000000      0.99929463373424     -0.06185639479703
*
```

After the analysis, a `<basename>.vpt2` file should be present in the working directory. Within that file all the force field and property derivatives are saved. It is used as an input for the `orca_vpt2` programme which is called automatically after the initial displacement calculations. The programme can also be called separately with the command `orca_vpt2 <basename>.vpt2`.

Note

A few remarks about VPT2 calculations:

- A VPT2 starting geometry **should always be tightly converged**. For small molecules the `!TightOPT` option is not good enough ! Depending on your structure, you might want to experiment with the `TolE`, `TolRMSG` and `TolMaxG` keywords of the `%geom` block.
- Similarly, a well converged SCF is required. The use of the `ExtremeSCF` keyword or at least `VeryTightSCF` is recommended.
- The CP-SCF equations should be converged to at least 10^{-12} (modified via the `Z_Tol` setting in the `%method` block.
- For DFT calculations, tight grids like `DEFGRID3` are strongly recommended.
- **Linear molecules are not supported yet**
- Currently, only methods for which analytical Hessians are available are supported. Furthermore, **VPT2 calculations with DFT functionals which do not provide analytical Hessians cannot be carried out**.
- By default, updated atomic masses are used to generate the semi-quartic force field (see *Mass dependencies*). The masses are also printed in the `<basename>.vpt2` file

- A VPT2 analysis can be repeated on a previous calculation by running `orca_vpt2 <basename>.vpt2`.
- VPT2 does have limited restart capabilities. If the directory in which the VPT2 run is carried out already contains `<basename>.hess` or `<basename>_eprnmr.property.txt` files, the program will skip these points and use the information provided in the files.

VPT2 provides a vibrational analysis and thus access to :

- mean and mean square displacement expectation values
- centrifugal distortion constants
- Watson's symmetrically and asymmetrically reduced Hamiltonian parameters
- anharmonic constants
- Fermi resonance analysis
- rotational and vibrational-rotational constants
- fundamental transition (anharmonic frequencies)
- zero-point ro-vibrational energies
- overtones and combination bands with intensities (in contrast to NEARIR with full VPT2/GVBT2 treatment)
- dedicated file interface for codes like SPCAT

If the computed data should be used for the simulation of spectra with codes like SPCAT, ORCA can provide a dedicated file that can serve as a basis for an input. This is triggered in the `%output` block when a VPT2 calculation is run:

```
%output
 Pickettname "pickett.txt"
end
```

This way, ORCA will generate a file called `pickett.txt` that contains the computed data and templates for `.var` which can be modified to serve as input for codes like SPCAT. Please note that this feature is still being refined and extended.

Vibrational corrections of molecular properties using VPT2

Using VPT2 it is also possible to compute zero-point vibrational corrections to molecular properties. Currently, this is available for NMR chemical shieldings, spin-spin coupling constants, g- and A-tensors and requires two successive calculations. The first calculation is a VPT2 calculation just as shown above (`<basename>.inp`) that contains the `VPT2` command and the level of theory at which the Hessians are computed. The second calculation (let's call it `<basename>_Prop.inp` will compute the property derivatives with a final call to VPT2. In order for this to work, the property derivative calculation needs to read the `<basename>.hess` and `<basename>.vpt2` file from the forcefield calculation. This is done by specifying the `%geom inhess read` with the command `inhessname "<basename>.hess"`. This scheme is necessary as properties other than energies, geometries or frequencies often require specialized methods and basis sets. For the numerical calculation of the force field and property derivatives different stepsizes can be used by specifying `AnharmDisp` and `PropDisp` in the VPT2 input block. The defaults are 0.05, and after the calculation, the displaced geometries are stored in files named `myjob_DH001.xyz` and `myjob_DP001.xyz` etc.

A typical example for calculating the vibrational correction to the ^{13}C NMR chemical shifts of methanol with a B3LYP/def2-TZVP anharmonic forcefield and TPSS/pcSseg-2 shielding tensors would look like the following: the standard input file, in our case `vpt2_methanol_FF.inp` with the level of theory for the Hessian and the VPT2 input block :

```
!B3LYP D3 def2-TZVP def2/J def2/JK DEFGRID3 ExtremeSCF VPT2

%method
```

(continues on next page)

(continued from previous page)

```
    Z_Tol  1e-12
end

* xyz 0 1
C  -1.09849212248373      0.14540972773089     -0.00000275092982
O   0.32138758531316      0.08706714755687     -0.00001212477411
H   0.66732439683790      0.98510769198508      0.00001819506998
H  -1.45583606337199     -0.88374271593276      0.00000595999622
H  -1.49206267729630      0.64725244577978      0.89143349761200
H  -1.49208273899904      0.64724452288014     -0.89144277697426
*
```

and the next input file, say `vpt2_methanol_NMR.inp` with the same geometry and the level of theory for the shielding tensor will look like this:

```
!TPSS pcSseg-2 Autoaux ExtremeSCF NMR

%geom inhess read
 inhessname "vpt2_methanol_FF.hess"
end

%vpt2
  VPT2          on
  AvgProp       NMR
  HessianCutoff 1e-12
end

%method
  Z_Tol         1e-12
end

* xyz 0 1
C  -1.09849212248373      0.14540972773089     -0.00000275092982
O   0.32138758531316      0.08706714755687     -0.00001212477411
H   0.66732439683790      0.98510769198508      0.00001819506998
H  -1.45583606337199     -0.88374271593276      0.00000595999622
H  -1.49206267729630      0.64725244577978      0.89143349761200
H  -1.49208273899904      0.64724452288014     -0.89144277697426
*
```

Running ORCA successively on both of these input files in the same directory will yield an output that contains the zero-point vibrational corrections to the shielding tensors for each atom. For Atom 0, which is the carbon in methanol, it looks like this:

```
-----
 Vibrationally averaged isotropic shieldings
 ----

Atom 0 :

 mode    <q>       <q2>      dS/dQ    d2S/dQ2
-----------------------------------------------
   0  -0.000017   0.202578  -0.000089  -5.922644
   1  -0.034052   0.057707   8.269988  -5.666515
   2  -0.036827   0.055687   5.667278 -13.843941
   3   0.000002   0.051446   0.000073  -7.353936
   4   0.027471   0.043993   0.423409  -6.207061
   5  -0.009357   0.040649 -12.736464   3.762324
   6  -0.000001   0.040278  -0.001621  -2.224536
   7   0.001277   0.039898  -1.266298  -3.916647
   8  -0.031609   0.020149  51.647411 -21.635780
```

(continues on next page)

(continued from previous page)

```
   9  -0.000021   0.019859   0.035760 -61.239749
  10  -0.010397   0.019376  18.573156 -50.591165
  11  -0.026641   0.015808  -8.871055  -6.654795
-----------------------------------------------

zpv correction to isotropic shift :  -4.840215 ppm
-----------------------------------------------
```

So the absolute shielding constant of carbon in methanol needs to be corrected by -4.8 ppm due to zero-point vibration. From the mean and mean square displacements and the first and second derivatives of the shieldings with respect to the normal modes, one can also identify degrees of freedom which give rise to larger contributions of the vibrational correction.

A similar input for the HH spin-spin coupling constants would look like this :

```
!TPSS pcJ-2 Autoaux ExtremeSCF NMR

%geom inhess read
 inhessname "vpt2_methanol_FF.hess"
end

%maxcore 4096

%vpt2
  VPT2          on
  AvgProp       JCOUPLING
  AnharmDisp    0.05
  HessianCutoff 1e-12
end

%method
  Z_Tol         1e-12
end

%eprnmr
  Tol           1e-10
end

* xyz 0 1
C  -1.09849212248373      0.14540972773089     -0.00000275092982
O   0.32138758531316      0.08706714755687     -0.00001212477411
H   0.66732439683790      0.98510769198508      0.00001819506998
H  -1.45583606337199     -0.88374271593276      0.00000595999622
H  -1.49206267729630      0.64725244577978      0.89143349761200
H  -1.49208273899904      0.64724452288014     -0.89144277697426
*

%eprnmr
Nuclei = all H {ssfc}
end
```

As mentioned above, the exact same procedure is also available for A-tensors. Here is an example for the NH_2 radical with the VPT2 input file `vpt2_NH2_FF.inp` :

```
!UKS BP86 def2-svp def2/J ExtremeSCF defgrid3

%vpt2
    VPT2                On
end

%method
```

(continues on next page)

(continued from previous page)

```
    Z_Tol               1e-12
end

* xyz 0 2
N   -0.01498947828047     -0.01894387811818      0.00000000000000
H   1.03197835263254      0.00908678452370      0.00000000000000
H   -0.22855980523269      1.00639225931822      0.00000000000000
*
```

and the input file - something like `vpt2_NH2_A.inp` - for the level of theory that will be used in the A-tensor computation:

```
!UKS BP86 def2-SVP TightSCF

%geom inhess read
 inhessname "vpt2_NH2_FF.hess"
end

%vpt2
    VPT2                On
    AvgProp             Atensor
end

*xyz 0 2
N   -0.01498947828047     -0.01894387811818      0.00000000000000
H   1.03197835263254      0.00908678452370      0.00000000000000
H   -0.22855980523269      1.00639225931822      0.00000000000000
*

%eprnmr
Nuclei = all N { aiso, adip }
Nuclei = all H { aiso, adip }
end
```

and similarly for the g-tensor if `Atensor` is replaced by `Gtensor` in the `%vpt2` block (the `%eprnmr` block can be omitted then).

Note that a convenient way to compute vibrational corrections is the usage of a compound script. With an input file called `NH2.inp` :

```
* xyz 0 2
N   0.00312611577632      0.00395297373474      0.00000000000000
H   1.01930353842041      0.00049997276783      0.00000000000000
H   -0.23400058507735      0.99208221922117      0.00000000000000
*

%Compound "NH2.cmp"
```

and the corresponding compound script `NH2.cmp`:

```
New_Step
!UHF def2-SVP VeryTightSCF

%vpt2
    VPT2                On
end

%method
    Z_Tol               1e-12
end
```

(continues on next page)

(continued from previous page)

```
* xyz 0 2
N   0.00312611577632      0.00395297373474      0.00000000000000
H   1.01930353842041      0.00049997276783      0.00000000000000
H   -0.23400058507735      0.99208221922117      0.00000000000000
*
Step_End

New_Step
!UHF def2-SVP VeryTightSCF

%geom inhess read
 inhessname "NH2_Compound_1.hess"
end

%vpt2
    VPT2                On
    AvgProp             Atensor
end

*xyz 0 2
N   0.00312611577632      0.00395297373474      0.00000000000000
H   1.01930353842041      0.00049997276783      0.00000000000000
H   -0.23400058507735      0.99208221922117      0.00000000000000
*

%eprnmr
Nuclei = all N { aiso, adip }
Nuclei = all H { aiso, adip }
end

Step_End

END
```

a similar result can be obtained in one calculation.

> **Note**
>
> Make sure the correct hessian file names are given and the input files MUST not contain a compound block. You can also rerun the VPT2 analysis using `orca_vpt2` directly. If you have an anharmonic force field calculation named `myjob_ff` and a property derivative calculation named `myjob_prop` just call `orca_vpt myjob_ff.vpt2 myjob_prop.vpt2`.

6.12.6 Electrical Properties

A few basic electric properties can be calculated in ORCA although this has never been a focal point of development. The properties can be accessed straightforwardly through the `%elprop` block:

```
! B3LYP DEF2-SVP TightSCF

%elprop Dipole        true # dipole moment
        Quadrupole    true # quadrupole moment

        Polar         true # dipole-dipole polarizability
                         1 # equivalent to true (for backward compatibility)
                           # Note: the flags "polar 2" and "polar 3" for seminumeric
                           # and fully numeric polarizabilities are not supported
                           # anymore! For numerical polarizability calculations
```

(continues on next page)

(continued from previous page)

```
                          # please use the respective compound scripts

       PolarVelocity true # polarizability w.r.t. velocity perturbations
       PolarDipQuad  true # dipole-quadrupole polarizability
       PolarQuadQuad true # quadrupole-quadrupole polarizability
       end
* int 0 1
   C  0  0  0  0       0         0
   H  1  0  0  1.09  109.4712    0
   H  1  2  0  1.09  109.4712    0
   H  1  2  3  1.09  109.4712  120
   H  1  2  3  1.09  109.4712  240
*
```

The polarizability (dipole-dipole, dipole-quadrupole, quadrupole-quadrupole) is calculated analytically through solution of the coupled-perturbed (CP-)SCF equations for HF and DFT runs (see *CP-SCF Options*) and through the CP-CASSCF equations for CASSCF runs (see *CASSCF Linear Response*). Analytic polarizabilities are also available for CCSD (via AUTOCI-CCSD, see *AUTOCI Response Properties via Analytic Derivatives*), RI-MP2 and DLPNO-MP2, as well as double-hybrid DFT methods. For canonical MP2 one can use AUTOCI for analytic calculations (see *AUTOCI Response Properties via Analytic Derivatives*) or differentiate the analytical dipole moment calculated with relaxed densities. For other correlation methods only a fully numeric approach is possible.

```
-------------------------------------------------
STATIC POLARIZABILITY TENSOR (Dipole/Dipole)
-------------------------------------------------

Method             : SCF
Type of density    : Electron Density
Type of derivative : Electric Field (Direction=X)
Multiplicity       :   1
Irrep              :   0
Relativity type    :
Basis              : AO

The raw cartesian tensor (atomic units):
      12.852429555        -0.002199911         0.000000170
      -0.002199911        12.860507003        -0.000000346
       0.000000170        -0.000000346        12.868107945
diagonalized tensor:
      12.851869269        12.861067290        12.868107945

Orientation:
       0.969064588        -0.246807263         0.000017958
       0.246807263         0.969064586        -0.000050696
      -0.000004890         0.000053560         0.999999999

Isotropic polarizability :  12.86035
```

As expected the polarizability tensor is isotropic.

Dipole-quadrupole polarizability tensors are printed as a list of 18 different components, with the first index running over x,y,z and the second index running over xx,yy,zz,xy,xz,yz. This is known as the "pure Cartesian" version of the tensor, although other conventions may exist in the literature that differ from the ORCA values by a constant factor.

```
-------------------------------------------------
STATIC POLARIZABILITY TENSOR (Dipole/Quadrupole)
-------------------------------------------------

Method             : SCF
Type of density    : Electron Density
```

(continues on next page)

(continued from previous page)

```
Type of derivative : Electric Field (Direction=X)
Multiplicity       :   1
Irrep              :   0
Relativity type    :
Basis              : AO

The raw cartesian tensor (atomic units):
  X- X X :     11.577165985
  X- Y Y :     -5.795339382
  X- Z Z :     -5.797320742
  X- X Y :      0.001285565
  X- X Z :      0.000000155
  X- Y Z :     -0.000000077
  Y- X X :      0.001386387
  Y- Y Y :      8.200445841
  Y- Z Z :     -8.198375727
  Y- X Y :     -5.794687548
  Y- X Z :      0.000000228
  Y- Y Z :     -0.000000121
  Z- X X :     -0.000000151
  Z- Y Y :      0.000000627
  Z- Z Z :     -0.000000812
  Z- X Y :     -0.000000312
  Z- X Z :     -5.798359323
  Z- Y Z :     -8.205110537
```

After this, the "traceless" version of the tensor is printed, which is usually denoted by $A_{x,xx}, A_{x,xy}$ etc. in the literature[164, 247, 573]. This is the preferred format for reporting dipole-quadrupole polarizability tensors. Certain references use the opposite sign convention than reported here, but generally the conventions of traceless polarizability tensors are more unified than those of pure Cartesian polarizability tensors.

```
-------------------------------------------------------------
STATIC TRACELESS POLARIZABILITY TENSOR (Dipole/Quadrupole)
-------------------------------------------------------------

Method             : SCF
Type of density    : Electron Density
Type of derivative : Electric Field (Direction=X)
Multiplicity       :   1
Irrep              :   0
Relativity type    :
Basis              : AO

The raw cartesian tensor (atomic units):
  X- X X :     17.373496046
  X- Y Y :     -8.685262003
  X- Z Z :     -8.688234043
  X- X Y :      0.001928347
  X- X Z :      0.000000232
  X- Y Z :     -0.000000116
  Y- X X :      0.000351329
  Y- Y Y :     12.298940512
  Y- Z Z :    -12.299291841
  Y- X Y :     -8.692031322
  Y- X Z :      0.000000342
  Y- Y Z :     -0.000000181
  Z- X X :     -0.000000058
  Z- Y Y :      0.000001109
  Z- Z Z :     -0.000001050
  Z- X Y :     -0.000000468
  Z- X Z :     -8.697538984
```

(continues on next page)

(continued from previous page)

```
  Z- Y Z :    -12.307665806
```

The quadrupole-quadrupole polarizability tensor is similarly printed in both the pure Cartesian and traceless forms. Again, the traceless form (usually denoted as $C_{xx,xx}, C_{xx,xy}$ etc.[164, 247, 573]) is the preferred format for reporting.

```
-------------------------------------------------
STATIC POLARIZABILITY TENSOR (Quadrupole/Quadrupole)
-------------------------------------------------

The order in each direction is XX, YY, ZZ, XY, XZ, YZ

Method             : SCF
Type of density    : Electron Density
Type of derivative : Quadrupolar Field (Direction=X)
Multiplicity       :   1
Irrep              :   0
Relativity type    :
Basis              : AO

The raw cartesian tensor (atomic units):
      60.656194448         8.024072323         8.017351959        -0.002591466         0.
↪000000801        -0.000000184
       8.024072323        55.906127614        12.837825709        -6.821368242        -0.
↪000000954        -0.000000529
       8.017351959        12.837825709        55.938851507         6.815300773         0.
↪000000232         0.000000422
      -0.002591466        -6.821368242         6.815300773        16.716647772         0.
↪000000169        -0.000000030
       0.000000801        -0.000000954         0.000000232         0.000000169        16.
↪715850196         6.818791255
      -0.000000184        -0.000000529         0.000000422        -0.000000030         6.
↪818791255        21.534628724
diagonalized tensor:
      11.893291534        13.566719080        26.357187387        46.234564137        52.
↪663246003        76.753292120

Orientation:
      -0.000000018        -0.000019986        -0.000000013        -0.001433194         0.
↪817436692         0.576016666
       0.000000006         0.219691224         0.000000038         0.673107563        -0.
↪405967809         0.577799371
       0.000000008        -0.219450737        -0.000000021        -0.671194117        -0.
↪408640606         0.578232381
      -0.000000035         0.950566746        -0.000000034        -0.310519906        -0.
↪000497156        -0.000034103
       0.816443231         0.000000038         0.577425709        -0.000000037         0.
↪000000027         0.000000000
      -0.577425709        -0.000000003         0.816443231        -0.000000036         0.
↪000000002        -0.000000003

Isotropic polarizability :  37.91138



-------------------------------------------------------------
STATIC TRACELESS POLARIZABILITY TENSOR (Quadrupole/Quadrupole)
-------------------------------------------------------------

The order in each direction is XX, YY, ZZ, XY, XZ, YZ
```

(continues on next page)

(continued from previous page)

```
Method             : SCF
Type of density    : Electron Density
Type of derivative : Quadrupolar Field (Direction=X)
Multiplicity       :   1
Irrep              :   0
Relativity type    :
Basis              : AO

The raw cartesian tensor (atomic units):
      26.331642600       -13.160050722       -13.171591878         0.000221134         0.
↪000000581        -0.000000065
     -13.160050722        22.733889017        -9.573838294        -5.113861448        -0.
↪000000735        -0.000000324
     -13.171591878        -9.573838294        22.745430172         5.113640314         0.
↪000000154         0.000000389
       0.000221134        -5.113861448         5.113640314        12.537485829         0.
↪000000127        -0.000000022
       0.000000581        -0.000000735         0.000000154         0.000000127        12.
↪536887647         5.114093442
      -0.000000065        -0.000000324         0.000000389        -0.000000022         5.
↪114093442        16.150971543
```

Note

- **Like the quadrupole moments themselves, the dipole-quadrupole and quadrupole-quadrupole polarizabilities depend on the gauge origin of the `%elprop` module. The latter can be changed using the `Origin` keyword in `%elprop`; see section *Electric Properties*.**

At the SCF level, dynamic (frequency-dependent) dipole polarizabilities can be computed using either purely real or purely imaginary frequencies.

```
%elprop
 polar 1
 freq_r 0.08  # purely real frequencies
 #freq_i 0.08 # purely imaginary frequencies
end
```

In the example above, the dynamic dipole polarizability tensor for a single real frequency of 0.8 a.u. is computed. For every frequency, linear response equations must be solved for all component of the perturbation operator (3 Cartesian components of the electric dipole). Note that imaginary-frequency polarizabilities are computed with the same costs as real-frequency polarizabilities. By a simple contour integration they can be used to compute dispersion coefficients.

For methods that do not support analytic polarizabilities, one can calculate numeric dipole-dipole and quadrupole-quadrupole polarizabilities, either by finite differentiation of dipole/quadrupole moments with respect to a finite dipole/quadrupole electric field, or by double finite differentiation of the total energy with respect to a finite dipole/quadrupole electric field. The latter can be done using compound scripts (see *Compound Methods*, *Compound Examples*).

At the MP2 level, polarizabilities can currently be calculated analytically using the RI (*RI-MP2 and Double-Hybrid DFT Response Properties*) or DLPNO (*Local MP2 Response Properties*) approximations or in all-electron canonical calculations, but for canonical MP2 with frozen core orbitals the dipole moment has to be differentiated numerically in order to obtain the polarizability tensor. In general in such cases, you should keep in mind that tight SCF convergence is necessary in order to not get too much numerical noise in the second derivative. Also, you should experiment with the finite field increment in the numerical differentiation process.

As an example, the following Compound job can be used to calculate the seminumeric polarizability at the MP2 level (replacing the now deprecated usage of `Polar 2`):

```
*xyz 0 1
  O   0.00000000000000      0.05591162058341      0.05591162058342
  H   0.00000000000000     -0.06629333722358      1.01038171664016
  H   0.00000000000000      1.01038171664017     -0.06629333722358
*

%Compound "semiNumericalPolarizability.cmp"
  with
    method      = "MP2";
    basis       = "aug-cc-pVDZ cc-pVDZ/C";
    restOfInput = "VeryTightSCF PModel NoFrozenCore";
  end
```

with the file `semiNumericalPolarizability.cmp` containing:

```
# ---------------------------------------------------------------------
# Authors: Dimitrios G. Liakos and Frank Neese
# Date   : March-May of 2024
#
# This is a compound script that calculates the
#  raw porarizability tensor semi-numerically
#  using the dipole moments.
#
# The idea is the following:
#
# 1. Calculate dipole moment  in the field free case
#
# 2. Loop over directions I=X,Y,Z
#    - put a small E-field in direction I+Delta
#    - Solve equations to get the dipole moment D+
#    - put s small E-field in direction I-Delta
#    - Solve equations to get the dipole moment D-
#    - Polarizability alpha(I,J). (D+(I)-D-(I))/(2Delta)
#    - Diagonalize to get eigenValues, eigenVectors, a_iso
#
# 3. Print polarisability
#
# NOTE: We use the last dipole_moment found in the file. If a specific
#       one is needed the 'myProperty' option should be accordingly
#       adjusted and remove the 'property_Base = true' option.
#
# NOTE: This is not the most general version. It is adjusted for testsuite
#       with 'method' and 'mp2' blocks.
# ---------------------------------------------------------------------
# ---------------------------------------------------------------------
# ----------------------    Variables    ------------------------------
#
# --- Variables to be adjusted (e.g. using 'with' ---------------------
Variable molecule    = "h2o.xyz";        # xyz file with coordinates
Variable charge      = 0;
Variable mult        = 1;
Variable method      = "HF";
Variable basis       = " " ;
Variable restOfInput = "NoFrozenCore VeryTightSCF";
Variable E_Field     = 0.0001;           # Size of Electric field
Variable myProperty  = "Dipole_Moment_Total";
Variable removeFiles = true;             # Remove files in the end of the calculation
# ---------------------------------------------------------------------
# -------------- Rest of the variables --------------------------------
Variable D_Free, D_Minus, D_Plus, a[3][3];    #dipole moment and polarizability
Variable aEigenValues[3], aEigenVectors[3][3], a_iso;
Variable FFieldStringPlus, FFieldStringMinus;
```

(continues on next page)

(continued from previous page)

```
Variable res = -1;

# ------------------------------------------------------------------------
# Field Free
# ------------------------------------------------------------------------
New_Step
  !&{method} &{basis} &{restOfInput}
  %Method
    z_tol 1e-8
  End
  %MP2
    Density Relaxed
  End
  #*xyzFile &{charge} &{mult} &{molecule}
Step_End
res = D_Free.readProperty(propertyName=myProperty, property_Base=true);

# ------------------------------------------------------------------
# Loop over the x, y, z directions and create the appropriate string
# ------------------------------------------------------------------
for direction from 0 to 2 Do
  #Create the appropriate direction oriented field string
  if (direction = 0) then          #( X direction)
    write2String(FFieldStringPlus,  " %lf, 0.0, 0.0", E_Field);
    write2String(FFieldStringMinus, "-%lf, 0.0, 0.0", E_Field);
  else if (direction = 1) then     #( Y direction)
    write2String(FFieldStringPlus,  " 0.0,  %lf, 0.0", E_Field);
    write2String(FFieldStringMinus, " 0.0, -%lf, 0.0", E_Field);
  else                             #( Z direction)
    write2String(FFieldStringPlus,  " 0.0,  0.0,  %lf", E_Field);
    write2String(FFieldStringMinus, " 0.0,  0.0, -%lf", E_Field);
  EndIf
  # ---------------------------------------
  # Perform the calculations.
  # First the plus (+) one
  # ---------------------------------------
  ReadMOs(1);
  New_Step
    !&{method} &{basis} &{restOfInput}
    %SCF
      EField = &{FFieldStringPlus}
    End
    %Method
      z_tol 1e-8
    End
    %MP2
      Density Relaxed
    End
  Step_End
  res = D_Plus.readProperty(propertyName=myProperty, property_Base=true);

  # ---------------------------------------
  # And the minus (-) one
  # ---------------------------------------
  ReadMOs(1);
  New_Step
    !&{method} &{basis} &{restOfInput}
    %SCF
      EField = &{FFieldStringMinus}
    End
    %Method
```

(continues on next page)

(continued from previous page)

```
      z_tol 1e-8
    End
    %MP2
      Density Relaxed
    End
  Step_End
  res = D_Minus.readProperty(propertyName=myProperty, property_Base=true);

  # ----------------------------------------
  # Construct and store SCF polarizability
  # ----------------------------------------
  a[direction][0] = (D_Plus[0]-D_Minus[0])/(2*E_Field);
  a[direction][1] = (D_Plus[1]-D_Minus[1])/(2*E_Field);
  a[direction][2] = (D_Plus[2]-D_Minus[2])/(2*E_Field);

EndFor
# ----------------------------------------
# Diagonalize
# ----------------------------------------
a.Diagonalize(aEigenValues, aEigenVectors);

# ----------------------------------------
# Do some printing
# ----------------------------------------
print( "\n\n");
print( " ------------------------------------------------------\n");
print( "                   COMPOUND                            \n");
print( " Semi analytical calculation of polarizability\n");
print( " ------------------------------------------------------\n");
print( " Method:  %s\n", method);
print( " Basis :  %s\n", basis);
print( " The electric field perturbation used was:    %.5lf a.u.\n", E_Field);
print( " \n\n");

print( " ------------------------------------------------------\n");
print( " Raw electric semi-analytical polarizability tensor\n");
print( " ------------------------------------------------------\n");
For i from 0 to 2 Do
  print("%13.8lf  %13.8lf  %13.8lf\n", a[i][0], a[i][1], a[i][2]);
EndFor
print( " ------------------------------------------------------\n");
print("\n");

print( " ------------------------------------------------------\n");
print( " Raw electric semi-analytical polarizability Eigenvalues\n");
print( " ------------------------------------------------------\n");
print("%13.8lf  %13.8lf  %13.8lf\n", aEigenValues[0], aEigenValues[1], aEigenValues[2]);
print( " ------------------------------------------------------\n");
print("\n");

print( " ------------------------------------------------------\n");
print( " Raw electric semi-analytical polarizability Eigenvectors\n");
print( " ------------------------------------------------------\n");
For i from 0 to 2 Do
  print("%13.8lf  %13.8lf  %13.8lf\n", aEigenVectors[i][0], aEigenVectors[i][1],␣
↪aEigenVectors[i][2]);
EndFor

print( "\n a isotropic value : %.5lf\n", (aEigenValues[0]+aEigenValues[1]+aEigenValues[2])/3.
↪0);
#
```

(continues on next page)

(continued from previous page)

```
#
# ----------------------------------------------------
#  Maybe remove unneccesary files
# ----------------------------------------------------
if (removeFiles) then
  sys_cmd("rm *_Compound_* *.bas* ");
EndIf

End
```

For more details on Compound jobs in general, see *Compound Methods*.

For other correlation methods, where not even relaxed densities are available, only a fully numeric approach (using compounds scripts) is possible and requires obnoxiously tight convergence.

Note that polarizability calculations have higher demands on basis sets. A rather nice basis set for this property is the Sadlej one (see *Built-in Basis Sets*). In relation to electric properties, it might be interesting to know that it is possible to carry out finite electric field calculations in ORCA. See section *Adding finite electric field* for more information on such calculations.

6.12.7 NMR Chemical Shifts

NMR chemical shifts at the HF, DFT (standard GGA and hybrid functionals), CASSCF, as well as the RI- and DLPNO-MP2 and double-hybrid DFT levels (see section *MP2 level magnetic properties* and references therein) can be obtained from the EPR/NMR module of ORCA. For the calculation of the NMR shielding tensor the program utilizes Gauge Including Atomic Orbitals (GIAOs - sometimes also referred to as London orbitals). [211, 375, 532] In this approach, field dependent basis functions are introduced, which minimizes the gauge origin dependence and ensures rapid convergence of the results with the one electron basis set. [289] Note that GIAOs are **NOT** currently available with CASSCF linear response and a gauge origin must be provided in the `%eprnmr` block (see *CASSCF Linear Response*). GIAOs for CASSCF response are coming soon to ORCA!

The use of the chemical shift module is simple:

```
# Ethanol - Calculation of the NMR chemical shieldings at the HF/SVP level of theory
! RHF SVP Bohrs NMR

* xyz 0 1
C        -1.22692181     0.24709455    -0.00000000
C        -0.01354839    -0.54677253     0.00000000
H        -2.09280406    -0.41333631     0.00000000
H        -1.24962478     0.87541936    -0.88916500
H        -1.24962478     0.87541936     0.88916500
O         1.09961824     0.30226226    -0.00000000
H         0.00915178    -1.17509696     0.88916500
H         0.00915178    -1.17509696    -0.88916500
H         1.89207683    -0.21621566     0.00000000
*
```

The output for the shieldings contains detailed information about the para- and diamagnetic contribution, the orientation of the tensor, the eigenvalues, its isotropic part etc. For each atom, an output block with this information is given :

```
--------------
 Nucleus   0C :
 --------------

Diamagnetic contribution to the shielding tensor (ppm) :
           209.647        -10.519          0.000
           -26.601        215.858          0.000
            -0.000          0.000        200.382
```

(continues on next page)

(continued from previous page)

```
Paramagnetic contribution to the shielding tensor (ppm):
            59.273         18.302         -0.000
            13.380          6.063         -0.000
             0.000         -0.000         -2.770

Total shielding tensor (ppm):
           268.920          7.783         -0.000
           -13.220        221.921         -0.000
             0.000          0.000        197.611


 Diagonalized sT*s matrix:

 sDSO          200.382          214.507          210.998  iso=     208.629
 sPSO           -2.770            7.279           58.057  iso=      20.855
        ---------------  ---------------  ---------------
 Total          197.611          221.786          269.055  iso=     229.484
```

Note that all units are given in ppm and the chemical shieldings given are *absolute* shieldings (see below). At the end of the atom blocks, a summary is given with the isotropic shieldings and the anisotropy [564] for each nucleus:

```
Nucleus  Element    Isotropic     Anisotropy
  -------  -------  ------------   ------------
      0       C          229.484         59.356
      1       C          227.642         62.878
      2       H           56.015         12.469
      3       H           55.460         15.284
      4       H           55.460         15.284
      5       O          334.125        110.616
      6       H           47.337         27.101
      7       H           47.337         27.101
      8       H           64.252         32.114
```

Thus, the absolute, isotropic shielding for the ^{13}C nuclei are predicted to be 229.5 and 227.6 ppm and for ^{17}O it is 334.1 ppm. While basis set convergence using GIAOs is rapid and smooth, it is still recommended to do NMR calculations with basis sets including tight exponents, such as the purpose-built pcSseg-n. However, TZVPP or QZVP should be sufficient in most cases. [59, 265]

An important thing to note is that in order to compare to experiment, a standard molecule for the type of nucleus of interest has to be chosen. In experiment, NMR chemical shifts are usually determined relative to a standard, for example either CH_4 or TMS for proton shifts. Hence, the shieldings for the molecule of interest and a given standard molecule are calculated, and the relative shieldigs are obtained by subtraction of the reference value from the computed value. It is of course important that the reference and target calculations have been done with the same basis set and functional. This also helps to benefit from error cancellation if the standard is chosen appropriately (one option is even to consider an "internal standard" - that is to use for example the ^{13}C shielding of a methyl group inside the compound of interest as reference).

Let us consider an example - propionic acid (CH_3-CH_2COOH). In databases like the AIST (http://sdbs.db.aist.go.jp) the ^{13}C spectrum in $CDCl_3$ can be found. The chemical shifts are given as δ_1 = 8.9 ppm, δ_2 = 27.6 ppm, δ_3 = 181.5 ppm. While intuition already tells us that the carbon of the carboxylic acid group should be shielded the least and hence shifted to lower fields (larger δ values), let's look at what calculations at the HF, BP86 and B3LYP level of theory using the SVP and the TZVPP basis sets yield:

method	σ_1	σ_2	σ_3
HF/SVP	191.7	176.6	23.7
HF/TZVPP	183.5	167.1	9.7
B86/SVP	181.9	165.8	26.5
B86/TZVPP	174.7	155.5	7.6
B3LYP/SVP	181.8	165.8	22.9
B3LYP/TZVPP	173.9	155.0	2.9

Looking at these results, we can observe several things - first of all, the dramatic effect of using too small basis sets, which yields differences of more than 10 ppm. Second, the results obviously change a lot upon inclusion of electron correlation by DFT and are functional dependent. Last but not least, these values have nothing in common with the experimental ones (they change in the wrong order), as the calculation yields *absolute shieldings* like in the table above, but the experimental ones are *relative shifts*, in this case relative to TMS.

In order to obtain the relative shifts, we calculate the shieldings σ_{TMS} of the standard molecule (TMS HF/TZVPP: 194.1 ppm, BP86/TZVPP: 184.8 ppm, B3LYP/TZVPP: 184.3 ppm) and by using $\delta_{mol} = \sigma_{ref} - \sigma_{mol}$ we can evaluate the chemical shifts (in ppm) and directly compare to experiment:

method	δ_1	δ_2	δ_3
HF/TZVPP	10.6	27.0	184.4
B86/TZVPP	10.1	29.3	177.2
B3LYP/TZVPP	10.4	29.3	181.4
Exp.	8.9	27.6	181.5

A few notes on the GIAO implementation in ORCA are in order. The use of GIAOs lead to some fairly complex molecular one- and two-electron integrals and a number of extra terms on the right hand side of the coupled-perturbed SCF equations. The two-electron contributions in particular can be time consuming to calculate. Thus, the RIJK, RIJDX, and RIJCOSX approximations were implemented and tested.[825] By default, the approximation used for the SCF is also applied to the GIAO integrals, but the latter can be changed using the `GIAO_2el` keyword in the eprnmr input block (see section *EPR and NMR properties*). Note that, while the default COSX grids are typically sufficiently accurate for chemical shifts, the use of `defgrid3` is recommended for special cases or post-HF calculations.

The user can finely control for which nuclei the shifts are calculated (although this will not reduce the computational cost very much, which is dominated by the CP-SCF equations for the magnetic field). This works in exactly the same way as for the hyperfine and quadrupole couplings described in the next section. For example:

```
! B3LYP def2-TZVP TightSCF

* int 0 1
C  0 0 0  0 0 0
C  1 0 0  1.35 0 0
H  1 2 0  1.1 120 0
H  1 2 3  1.1 120 180
H  2 1 3  1.1 120 0
H  2 1 3  1.1 120 180
*

%eprnmr
  Ori    = GIAO
  Nuclei = all C { shift }
  Nuclei = all H { shift }
  end
```

NMR chemical shifts are also implemented in combination with implicit solvent models, hence the NMR keyword can be combined with the CPCM input block. Note that for calculations including implicit solvent, it is highly recommended to also optimize the geometries using the same model. Regardless, explicit solvent–solute interactions

observable in NMR (e.g. H-bonds), cannot be modelled with such a model: solvent molecules must be included explicitly in the calculation.

6.12.8 NMR Spin-Spin Coupling Constants

The indirect spin-spin coupling constants observed in NMR spectra of molecules in solution consist of four contributions: The diamagnetic spin orbit term:

$$\hat{H}_{DSO} = \frac{1}{2} \sum_{ikl} \frac{(\vec{M}_k \times \vec{r}_{ik})(\vec{M}_l \times \vec{r}_{il})}{r_{ik}^3 \, r_{il}^3} \tag{6.9}$$

The paramagnetic spin orbit term:

$$\hat{H}_{PSO} = \sum_{ik} \frac{\vec{M}_k \, \vec{l}_{ik}}{r_{ik}^3} \tag{6.10}$$

The Fermi contact term:

$$\hat{H}_{FC} = \frac{8\,\pi}{3} \sum_{ik} \delta(r_i - r_k) \, \mathbf{m}_k \tag{6.11}$$

And the spin dipole term:

$$\hat{H}_{SD} = \sum_{ik} \mathbf{m}_k^T \, \frac{3 \, \mathbf{r}_{ik} \, \mathbf{r}_{ik}^T - r_{ik}^2}{r_{ik}^5} \, \mathbf{s}_i \tag{6.12}$$

While the Fermi contact term is usually the most significant, all contributions can be computed at the HF and DFT level of theory using ORCA. For this purpose, the keyword `ssall` has to be invoked in the eprnmr input block, while each of the four terms can be requested using `ssdso`, `sspso`, `ssfc`, and `sssd`, respectively. For example:

```
! PBE0 pcJ-1

*xyz 0 1
O 0.00000        0.00000        0.11779
H 0.00000        0.75545       -0.47116
H 0.00000       -0.75545       -0.47116
*

%eprnmr
 Nuclei = all O { ssall }
 Nuclei = all H { ssfc, sssd }
end
```

Results will be given in Hz. Note that the default isotopes used might not be the ones desired for the calculation of NMR properties, so it is recommended to define the corresponding isotopes using the `ist` flag. It is possible to also print the reduced coupling constants K (in units of $10^{19} \cdot \mathrm{T} \cdot \mathrm{J}^{-2}$), which are independent of the nuclear isotopes, using the flag `PrintReducedCoupling=True`.

The CP-SCF equations must be solved for one of the nuclei in each pair and are the bottleneck of the computation. Therefore, spin-spin coupling constants are calculated only between nuclei within a certain distance of eachother (5 Ångstrom by default). The latter can be changed using the `SpinSpinRThresh` keyword.

If mulitple nuclides are requested, it is also possible to select only certain element pairs (e.g. only C–H and H–H, without C–C) using the `SpinSpinElemPairs` keyword. Analogously, the `SpinSpinAtomPairs` keyword selects the actual pairs of nuclei to consider. The union of the latter two options is used to *reduce* the selection made using the `Nuclei` input, after which `SpinSpinRThresh` is applied.

Here is another example illustrating these additional options:

```
! B3LYP EPR-II

* xyz 0 1
C     -1.226922    0.247095   -0.000000
C     -0.013548   -0.546773    0.000000
H     -2.092804   -0.413336    0.000000
H     -1.249625    0.875419   -0.889165
H     -1.249625    0.875419    0.889165
O      1.099618    0.302262   -0.000000
H      0.009152   -1.175097    0.889165
H      0.009152   -1.175097   -0.889165
H      1.892077   -0.216216    0.000000
*

%eprnmr
  nuclei = all C { ssall, ist = 13 };
  nuclei = all H { ssall, ist = 1  };
  nuclei = all O { ssall, ist = 17 };
  PrintReducedCoupling true
  SpinSpinRThresh 3.0 # Angstrom
  SpinSpinElemPairs {C C} {O *} # "*" means any element
  SpinSpinAtomPairs {8 *}       # indices start from 0

  # Final selection:
  # C 0 - C 1
  # C 0 - O 5
  # C 1 - O 5
  # C 1 - H 8
  # H 3 - O 5
  # H 4 - O 5
  # O 5 - H 6
  # O 5 - H 7
  # O 5 - H 8
  # H 6 - H 8
  # H 7 - H 8
end
```

NMR Spectra

From the computed NMR shieldings and spin-spin coupling constants, the coupled NMR spectra can be simulated. The most typical NMR experiments are decoupled ^{13}C and coupled ^{1}H spectra, so we will focus on these here. For simulating a full NMR spectrum, we will use ethyl crotonate as an example, and three steps are required: 1) computation of the spin-spin coupling constants, 2) computation of the chemcial shieldings and 3) simulation of the spectrum using a spin-Hamiltonian formalism. Note that these steps can be carried out independently and different levels of theory can be used for shieldings and couplings and the order of steps 1 and 2 doesn't matter.

Furthermore, if the spectra are to be simulated with TMS as reference, the shieldings for TMS are required (the hydrogen and carbon values in this case are 31.77 and 188.10 ppm, respectively). Here is the input for the calculation for the coupling constants, which is named `ethylcrotonate_sscc.inp`:

```
! PBE pcJ-3 autoaux tightscf

*xyzfile 0 1 ethylcrotonate.xyz

%eprnmr
 Nuclei = all H {ssall}
end
```

The spin-spin coupling constants will be stored in the file `ethylcrotonate_sscc.property.txt`, which the simulation of the NMR spectrum will need to read. The NMR shieldings and will be computed in the next step, for which the input `ethylcrotonate_nmr.inp` looks like this:

```
!TPSS pcSseg-3 autoaux tightscf NMR

*xyzfile 0 1 ethylcrotonate.xyz
```

The final NMR spectrum simulation is carried out using `orca_nmrspectrum`, which requires a separate input file `ethylcrotonate.nmrspec` which looks like this (note the required final END statement):

```
NMRShieldingFile = "ethylcrotonate_nmr"   #property file for shieldings
NMRCouplingFile  = "ethylcrotonate_sscc"  #property file for couplings
NMRSpecFreq = 80.00                       #spectrometer freq [MHz] (default 400)
PrintLevel = 0                            #PrintLevel for debugging info
NMRCoal = 1.0                             #threshold for merged lines [Hz] (default 1)
NMRREF[1] 31.77                           #shielding for 1H reference [ppm]
NMRREF[6] 188.10                          #shielding for 13C reference [ppm]
NMREquiv                                  #lists of NMR-equivalent nuclei
1 {13 14 15} end                          #H 13,14,15 are equivalent (methyl)
2 {16 17} end                             #H 16 and 17 equivalent (ethyl)
3 {8 10 11} end                           #H 8,10,11 again equivalent methyl
end                                       #end equiv nucl block
END                                       #essential end of input
```

and contains the following keywords:
`NMRShieldingFile` and `NMRCouplingFile` denote the `.property.txt` files from which the shielding tensors and coupling constants will be read by the NMR spectrum module. If this line is not given, the program will exepect the shieldings or couplings in the property file of the current calculation.

`NMRSpecFreq` The NMR spectrometer frequency in MHz is decisive for the looks of the spectrum as shieldings are given in ppm and couplings are given in Hz. Default is 400 MHz.

`NMRCoal` If two lines are closer than this threshold (given in Hz) then the module will coalesce the lines to one line with double intensity. Default it 1 Hz.

`NMRREF[X]` Reference value for the absolute shielding of element X used in the relative shifts of the simulated spectrum. Typically, these are the absolute shielding values from a separate calculation of TMS, for example, and will be zero ppm in the simulated spectrum.

`NMREquiv` The user has to specify groups of NMR equivalent nuclei. These are typically atoms which interchange on the NMR timescale, like methyl group protons. The shieldings and couplings will be averaged for each group specified by the user.

with this input, `orca_nmrspectrum` is called with two arguments, the first one is a gbw file which contains all informations about the molecule, typically this is the gbw file of the nmr or the sscc calculation, and the name of the input file given above:

```
orca_nmrspectrum ethylcrotonate_nmr.gbw ethylcrotonate.nmrspec > output
```

If this calculation is carried out, the NMR spectrum module will read the files `ethylcrotonate_sscc.property.txt` and `ethylcrotonate_nmr.property.txt`, extract the shieldings and couplings, average the equivalent nuclei and set up an effective NMR spin Hamiltonian for each nucleus:

$$H_{NMR}(p,q) = \sigma_p \delta_{pq} + J_{pq} I_p I_q. \tag{6.13}$$

Caution

This includes all nuclei this nuclear spin couples to and should not contain too many spins (see `SpinSpinRThres`), as the spin Hamiltonian for each atom and the nuclei it couples to is diagonalized brute force. Afterwards, all spin excitations are accumulated and merged into the final spectrum for each element. For ethyl crotonate the NMR spectrum output looks like this:

```
-----------------------------------------------------
 NMR Peaks for atom type 1, ref value  31.7700 ppm :
```

```
-----------------------------------------------------
Atom   shift[ppm]   rel.intensity
   8        2.34    9.00
   8        2.36    9.00
   8        2.25    9.00
   8        2.27    9.00
   9        6.34    1.00
   9        6.36    3.00
   9        6.38    3.00
   9        6.41    1.00
   9        6.14    1.00
   9        6.16    3.00
   9        6.19    3.00
   9        6.21    1.00
  12        7.95    1.00
  12        7.85    3.00
  12        7.75    4.00
  12        7.65    4.00
  12        7.56    3.00
  12        7.47    1.00
  13        1.71    9.00
  13        1.61   18.00
  13        1.52    9.00
  16        4.56    4.00
  16        4.46   12.00
  16        4.37   12.00
  16        4.27    4.00

-----------------------------------------------------
 NMR Peaks for atom type 6, ref value 188.1000 ppm :
-----------------------------------------------------
Atom   shift[ppm]   rel.intensity
   2       25.70    1.00
   3      155.15    1.00
   4       19.96    1.00
   5       68.91    1.00
   6      174.39    1.00
   7      130.29    1.00

-----------------------------------------------------
 NMR Peaks for atom type 8, ref value 104.8826 ppm :
-----------------------------------------------------
Atom   shift[ppm]   rel.intensity
   0        0.00    5.00
   1      149.74    5.00
```

The first column denotes the atom number of the nucleus the signal arises from, the second column denotes the line position in ppm and the third line denotes the relative intensity of the signal. For oxygen, no reference value was given, so the program will autmatically set the lowest value obtained in the calculation as reference value.

Using gnuplot, for example, to plot the spectrum, the following plots for ^{13}C and ^{1}H are obtained[5] :

[5] The basic plot options for using gnuplot are `plot "mydata" using 2:3 w i, "mydata" using 2:3:1 with labels`

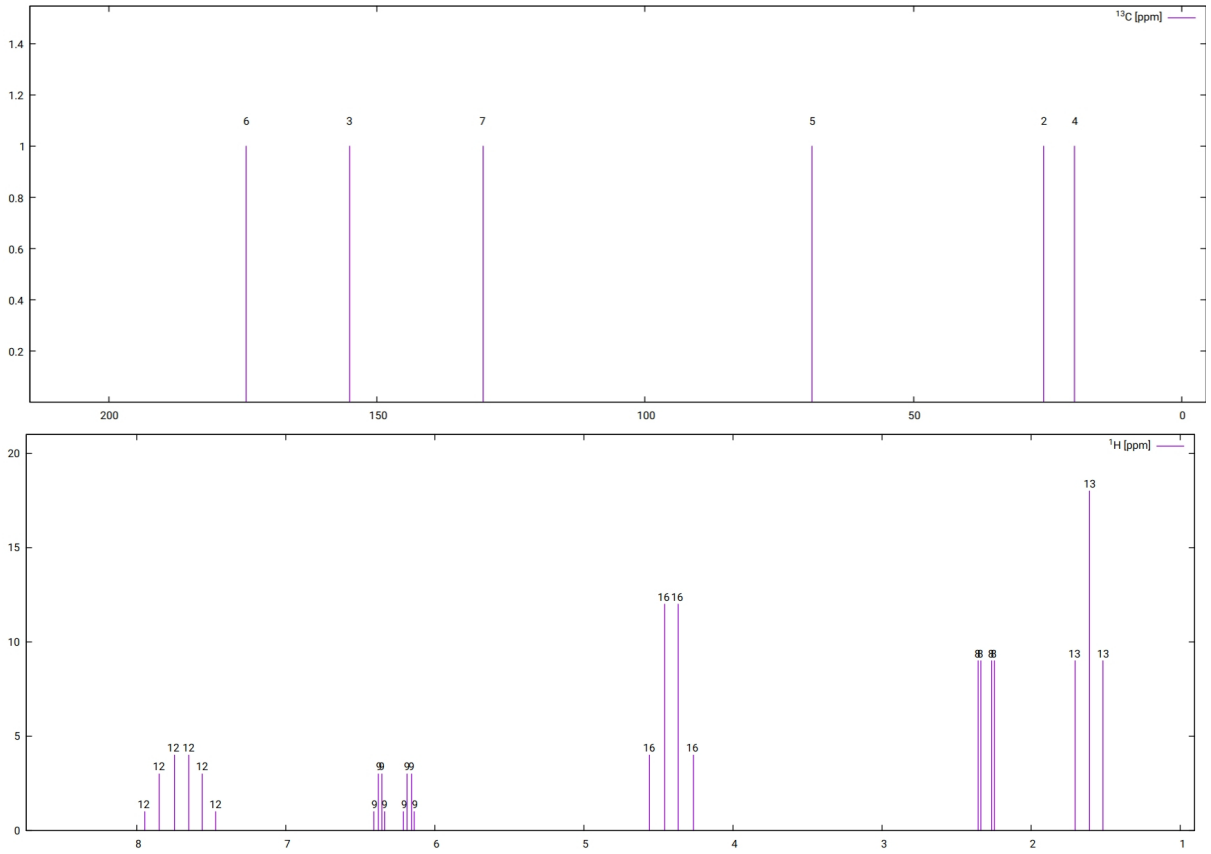


Fig. 6.54: Simulated $^{1}3C$ (top) and 1H (bottom) NMR spectra. Note that as only HH couplings have been computed, the spectra do not include any CH couplings and the carbon spectrum is also uncoupled.

This makes comparison to experiment and assessment of the computed parameters much easier, however, it is not as advanced as other codes and does not, for example, take conformational degrees of freedom etc. into account. Note that the corresponding property files can also be modified to tinker with the computed shieldings and couplings.

Visualizing shielding tensors using `orca_plot`

For the visualization it is recommended to perform an ORCA NMR calculation such that the corresponding `gbw` and `density` files required by `orca_plot` are generted by using the `!keepdens` keyword along with `!NMR`. If `orca_plot` is called in the interactive mode by specifying `orca_plot myjob.gbw -i` (note that `myjob.gbw`, `myjob.densities` and `myjob.property.txt` have to be in this directory), then following `1 - type of plot` and choosing `17 - shielding tensor`, confirming the name of the property file and then choosing `11 - Generate the plot` will generate a `.cube` file with shielding tensors depicted as ellipsoids at the corrsponding nuclei. These can be plotted for example using Avogadro, isosurface values of around 1.0 and somewhat denser grids for the cube file (like 100x100x100) are recommended. A typical plot for CF_3SCH_3 generated with Avogadro looks like this[6]:

[6] the same scheme can be applied to visualize polarisability tensors in the molecular framework using `orca_plot`.

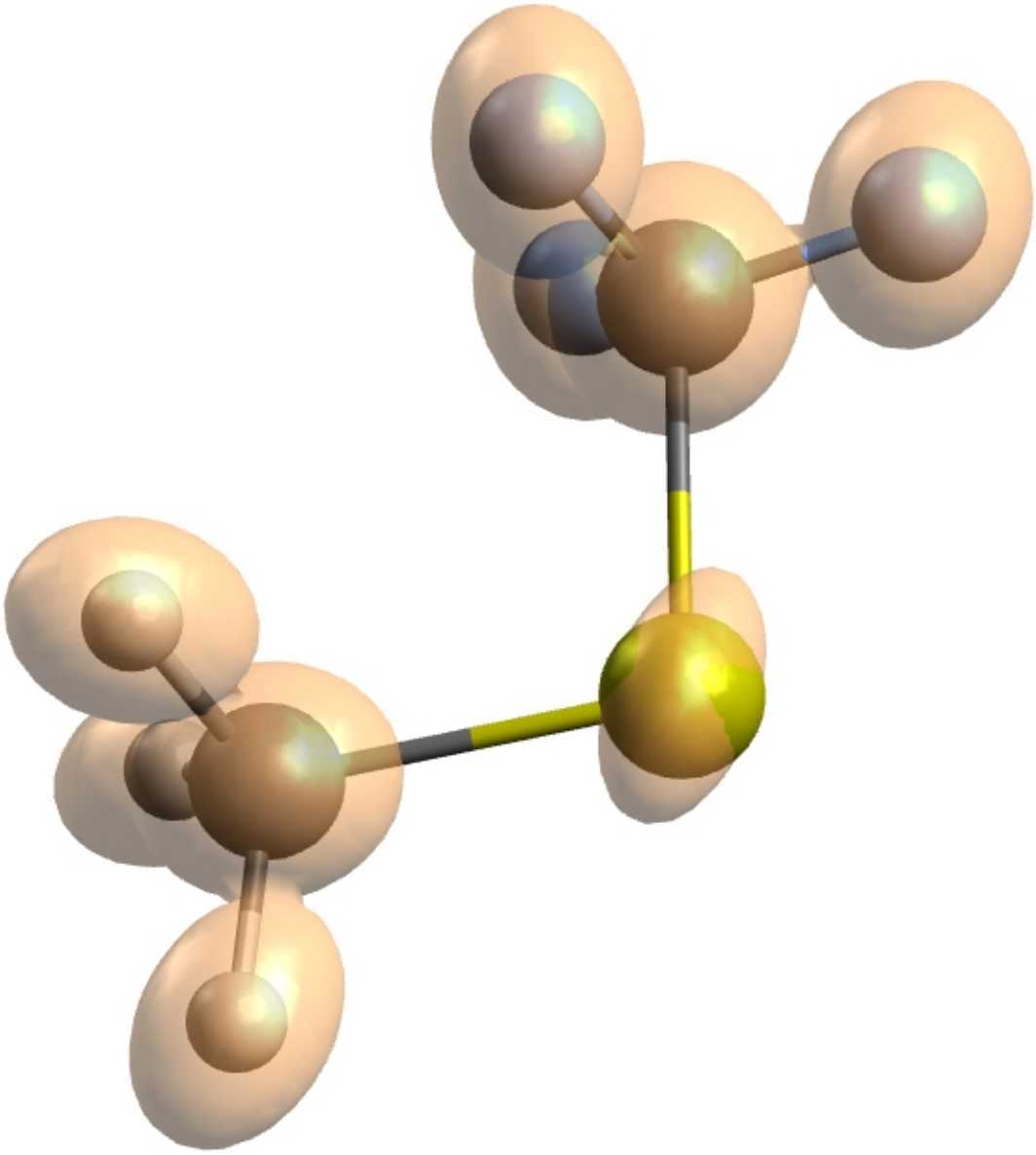


Fig. 6.55: The shielding tensors of each atom in CF3SCH3 have been plotted as ellipsoids (a,b and c axis equivalent to the normalized principle axes of the shielding tensors) at the given nuclei.

6.12.9 Hyperfine and Quadrupole Couplings

Hyperfine and quadrupole couplings can be obtained from the EPR/NMR module of ORCA. Since there may be several nuclei that you might be interested in the input is relatively sophisticated.

An example how to calculate the hyperfine and field gradient tensors for the CN radical is given below:

```
! PBE0 def2-MSVP TightSCF
* int 0 2
   C  0  0  0  0     0  0
   N  1  0  0  1.170 0  0
*
%eprnmr  Nuclei = all C { aiso, adip }
```

(continues on next page)

(continued from previous page)

```
        Nuclei = 2 { aiso, adip, fgrad }
        end
```

In this example the hyperfine tensor is calculated for all carbon atoms and atom 2, which is nitrogen in this case.

Note

- counting of atom numbers starts from 1
- All nuclei mentioned in one line will be assigned the same isotopic mass, i.e. if several nuclei are calculated, there has to be a new line for each of them.
- You have to specify the `Nuclei` statement *after* the definition of the atomic coordinates or the program will not figure out what is meant by "`all`".

The output looks like the following. It contains detailed information about the individual contributions to the hyperfine couplings, its orientation, its eigenvalues, the isotropic part and (if requested) also the quadrupole coupling tensor.

```
----------------------------------------
ELECTRIC AND MAGNETIC HYPERFINE STRUCTURE
----------------------------------------

 ------------------------------------------------------------
 Nucleus   0C : A:ISTP=   13 I=  0.5 P=134.1903 MHz/au**3
                Q:ISTP=   13 I=  0.5 Q=  0.0000 barn
 ------------------------------------------------------------
 Raw HFC matrix (all values in MHz):
 ---------------------------------
               695.8952               0.0000              -0.0000
                 0.0000             543.0617              -0.0000
                -0.0000              -0.0000             543.0617

 A(FC)         594.0062             594.0062             594.0062
 A(SD)         -50.9445             -50.9445             101.8890
              ----------           ----------           ----------
 A(Tot)        543.0617             543.0617             695.8952    A(iso)=  594.0062
 Orientation:
  X           0.0000000            0.0000000           -1.0000000
  Y          -0.8111216           -0.5848776           -0.0000000
  Z          -0.5848776            0.8111216            0.0000000

Notes:  (1) The A matrix conforms to the "SAI" spin Hamiltonian convention.
        (2) Tensor is right-handed.

 ------------------------------------------------------------
 Nucleus   1N : A:ISTP=   14 I=  1.0 P= 38.5677 MHz/au**3
                Q:ISTP=   14 I=  1.0 Q=  0.0204 barn
 ------------------------------------------------------------
 Raw HFC matrix (all values in MHz):
 ---------------------------------
                13.2095              -0.0000               0.0000
                -0.0000             -45.6036              -0.0000
                 0.0000              -0.0000             -45.6036

 A(FC)         -25.9993             -25.9993             -25.9993
 A(SD)          39.2088             -19.6044             -19.6044
              ----------           ----------           ----------
 A(Tot)         13.2095             -45.6036             -45.6036    A(iso)=  -25.9993
 Orientation:
```

(continues on next page)

(continued from previous page)

```
  X           1.0000000            0.0000000           -0.0000000
  Y          -0.0000000            0.9996462            0.0265986
  Z           0.0000000           -0.0265986            0.9996462

Notes:  (1) The A matrix conforms to the "SAI" spin Hamiltonian convention.
        (2) Tensor is right-handed.

 Raw EFG matrix (all values in a.u.**-3):
 ------------------------------------
            -0.1832      -0.0000       0.0000
            -0.0000       0.0916       0.0000
             0.0000       0.0000       0.0916

 V(El)       0.6468       0.6468      -1.2935
 V(Nuc)     -0.5551      -0.5551       1.1103
          ----------   ----------   ----------
 V(Tot)      0.0916       0.0916      -0.1832
 Orientation:
  X      -0.0000003    0.0000002    1.0000000
  Y       0.9878165    0.1556229    0.0000003
  Z      -0.1556229    0.9878165   -0.0000002

Note: Tensor is right-handed


 Quadrupole tensor eigenvalues (in MHz;Q= 0.0204 I=  1.0)
  e**2qQ             =    -0.880 MHz
  e**2qQ/(4I*(2I-1))=    -0.220 MHz
  eta                =     0.000
  NOTE: the diagonal representation of the SH term I*Q*I = e**2qQ/(4I(2I-1))*[-(1-eta),-
↪(1+eta),2]
```

Another important point to consider for hyperfine calculations concerns the choice of basis sets. You should normally use basis sets that have more flexibility in the core region. In the present example a double-zeta basis set was used. For accurate calculations this is not sufficient. There are several dedicated basis set for hyperfine calculations:

- EPR-II basis of Barone and co-workers: It is only available for a few light atoms (H, B, C, N, O, F) and is essentially of double-zeta plus polarization quality with added flexibility in the core region, which should give reasonable results.
- IGLO-II and IGLO-III bases of Kutzelnigg and co-workers: They are fairly accurate but also only available for some first and second row elements.
- CP basis: They are accurate for first row transition metals as well.
- uncontracted Partridge basis: They are general purpose HF-limit basis sets and will probably be too expensive for routine use, but are very useful for calibration purposes.

For other elements ORCA does not yet have dedicated default basis sets for this situation it is very likely that you have to tailor the basis set to your needs. If you use the statement `Print[p_basis] 2` in the `%output` block (or `PrintBasis` in the simple input line) the program will print the actual basis set in input format (for the basis block). You can then add or remove primitives, uncontract core bases etc. For example, here is a printout of the carbon basis DZP in input format:

```
# Basis set for element : C
  NewGTO 6
  s 5
   1  3623.8613000000      0.0022633312
   2   544.0462100000      0.0173452633
   3   123.7433800000      0.0860412011
   4    34.7632090000      0.3022227208
```

(continues on next page)

(continued from previous page)

```
  5    10.9333330000      0.6898436475
 s 1
  1     3.5744765000      1.0000000000
 s 1
  1     0.5748324500      1.0000000000
 s 1
  1     0.1730364000      1.0000000000
 p 3
  1     9.4432819000      0.0570590790
  2     2.0017986000      0.3134587330
  3     0.5462971800      0.7599881644
 p 1
  1     0.1520268400      1.0000000000
 d 1
  1     0.8000000000      1.0000000000
 end;
```

The "`s 5`", for example, stands for the angular momentum and the number of primitives in the first basis function. Then there follow five lines that have the number of the primitive, the exponent and the contraction coefficient (unnormalized) in it. **Remember also that when you add very steep functions you *must* increase the size of the integration grid if you do DFT calculations! *If you do not do that your results will be inaccurate*.** You can increase the radial grid size by using `IntAcc` in the `Method` block or for individual atoms (section *Other details and options* explains how to do this in detail). In the present example the changes caused by larger basis sets in the core region and more accurate integration are relatively modest – on the order of 3%, which is, however, still significant if you are a little puristic.

The program can also calculate the spin-orbit coupling contribution to the hyperfine coupling tensor as described in section *EPR and NMR properties*.To extract the A tensor from a oligonuclear transition metal complex, the `A(iso)` value in the output is to be processed according to the method described in ref. [644].

For the calculation of HFCCs using DLPNO-CCSD it is recommended to use the tailored truncation settings `! DLPNO-HFC1` or `!DLPNO-HFC2` in the simple keyword line which correspond to the "Default1" and "Default2" setting in Ref. [742].

If also EPR g-tensor or D-tensor calculations (see next section) are carried out in the same job, ORCA automatically prints the orientation between the hyperfine/quadrupole couplings and the molecular g- or D-tensor. For more information on this see section *orca_euler*.

6.12.10 The EPR g-Tensor and the Zero-Field Splitting Tensor

The EPR g-tensor is a property that can be well calculated at the SCF level with ORCA through solution of the coupled-perturbed (CP-)SCF equations (see *CP-SCF Options*) and at the CASSCF level via the CP-CASSCF equations (see *CASSCF Linear Response*). As an example, consider the following simple g-tensor job:

```
! BP86 Def2-SVP TightSCF g-tensor SOMF(1X)
* int 1 2
   O  0  0  0  0      0  0
   H  1  0  0  1.1056 0  0
   H  1  2  0  1.1056 109.62 0
*
```

The simplest way is to call the g-tensor property in the simple input line as shown above, but it can also be specified in the `%eprnmr` block with `gtensor true`. Starting from ORCA 5.0 the default treatment of the gauge is the GIAO approach, but this can be modified by the keyword `ori`. Other options are defined in section *EPR and NMR properties*. `SOMF(1X)` defines the chosen spin-orbit coupling (SOC) operator. The details of the SOC operator are defined in section *The Spin-Orbit Coupling Operator*. Other choices and additional variables exist and can be set as explained in detail in section *The Spin-Orbit Coupling Operator*.

The output looks like the following. It contains information on the contributions to the g-tensor (relativistic mass correction, diamagnetic spin-orbit term (= gauge-correction), paramagnetic spin-orbit term (= OZ/SOC)),

the isotropic g-value and the orientation of the total tensor.

```
-------------------
ELECTRONIC G-MATRIX
-------------------

The g-matrix:
              2.0104321   -0.0031354   -0.0000000
             -0.0031354    2.0081968   -0.0000000
             -0.0000000   -0.0000000    2.0021275

Breakdown of the contributions
 gel          2.0023193    2.0023193    2.0023193
 gRMC        -0.0003174   -0.0003174   -0.0003174
 gDSO(tot)    0.0000808    0.0001539    0.0001515
 gPSO(tot)    0.0000449    0.0038301    0.0104898
             ----------   ----------   ----------
 g(tot)       2.0021275    2.0059858    2.0126431 iso=  2.0069188
 Delta-g     -0.0001917    0.0036665    0.0103238 iso=  0.0045995

 Orientation:
  X           0.0000000    0.5762906   -0.8172448
  Y           0.0000000    0.8172448    0.5762906
  Z           1.0000000   -0.0000000    0.0000000
```

G-tensor calculations using GIAOs are now available at SCF and the RI-MP2 level. Note that GIAOs are **NOT** currently available with CASSCF linear response and a gauge origin must be provided in the `%eprnmr` block (see *CASSCF Linear Response*). GIAOs for CASSCF response are coming soon to ORCA!

The GIAO one-electron integrals are done analytically by default whereas the treatment of the GIAO two-electron integrals is chosen to be same as for the SCF. The available options which can be set with `giao_1el / giao_2el` in the `%eprnmr` block can be found in section *EPR and NMR properties*.

Concerning the computational time, for small systems, e.g. phenyl radical (41 electrons), the `rijk`-approximation is good to use for the SCF-procedures as well as the GIAO two-electron integrals. Going to larger systems, e.g. chlorophyll radical (473 electrons), the `rijcosx`-approximation reduces the computational time enormously. While the new default grid settings in ORCA 5.0 (`defgrid2`) should be sufficient in most cases, certain cases might need the use of `defgrid3`. The output looks just the same as for the case without GIAOs, but with additional information on the GIAO-parts.

If the total spin of the system is S >1/2 then the zero-field-splitting tensor can also be calculated and printed. For example consider the following job on a hypothetical Mn(III)-complex.

```
! BP86 def2-SVP SOMF(1X)

%eprnmr DTensor  ssandso
        DSOC     cp  # qro, pk, cvw
        DSS      uno # direct
        end

* int  1 5
Mn 0 0 0  0 0 0
 O 1 0 0  2.05   0   0
 O 1 2 0  2.05  90   0
 O 1 2 3  2.05  90 180
 O 1 2 3  2.05 180   0
 F 1 2 3  1.90  90  90
 F 1 2 3  1.90  90 270
 H 2 1 6  1.00 127   0
 H 2 1 6  1.00 127 180
 H 3 1 6  1.00 127   0
 H 3 1 6  1.00 127 180
 H 4 1 6  1.00 127   0
```

(continues on next page)

(continued from previous page)

```
 H 4 1 6  1.00 127 180
 H 5 1 6  1.00 127   0
 H 5 1 6  1.00 127 180
*
```

The output documents the individual contributions to the D-tensor which also contains (unlike the g-tensor) contributions from spin-flip terms.

Some explanation must be provided:

- The present implementation in ORCA is valid for HF, DFT and hybrid DFT.
- There are four different variants of the SOC-contribution, which shows that this is a difficult property. We will briefly discuss the various choices.
- The QRO method is fully documented[611] and is based on a theory developed earlier.[621] The QRO method is reasonable but somewhat simplistic and is superseded by the CP method described below.
- The Pederson-Khanna model was brought forward in 1999 from qualitative reasoning.[656] It also contains incorrect prefactors for the spin-flip terms. We have nevertheless implemented the method for comparison. In the original form it is only valid for local functionals. In ORCA it is extended to hybrid functionals and HF.
- The coupled-perturbed method is a generalization of the DFT method for ZFSs; it uses revised prefactors for the spin-flip terms and solves a set of coupled-perturbed equations for the SOC perturbation. Therefore it is valid for hybrid functionals. It has been described in detail.[613]
- The DSS part is an expectation value that involves the spin density of the system. In detailed calibration work[799] it was found that the spin-unrestricted DFT methods behave somewhat erratically and that much more accurate values were obtained from open-shell spin-restricted DFT. Therefore the "UNO" option allows the calculation of the SS term with a restricted spin density obtained from the singly occupied unrestricted natural orbitals.
- The DSS part contains an erratic self-interaction term for UKS/UHF wavefunction and canonical orbitals. Thus, `UNO` is recommended for these types of calculations.[723] If the option `DIRECT` is used nevertheless, ORCA will print a warning in the respective part of the output.
- In case that D-tensor is calculated using the correlated wave function methods such as (DLPNO-/LPNO-)CCSD, one should not use `DSS=UNO` option.

More information about the D-tensor can be found in section *Zero-Field-Splitting*.

6.12.11 Mössbauer Parameters

^{57}Fe Mössbauer spectroscopy probes the transitions of the nucleus between the $I = \frac{1}{2}$ ground state and the $I = \frac{3}{2}$ excited state at 14.4 keV above the ground state. The important features of the Mössbauer spectrum are the isomer shift (δ) and the quadrupole splitting (ΔE_{Q}). An idealized spectrum is shown in Fig. 6.56.

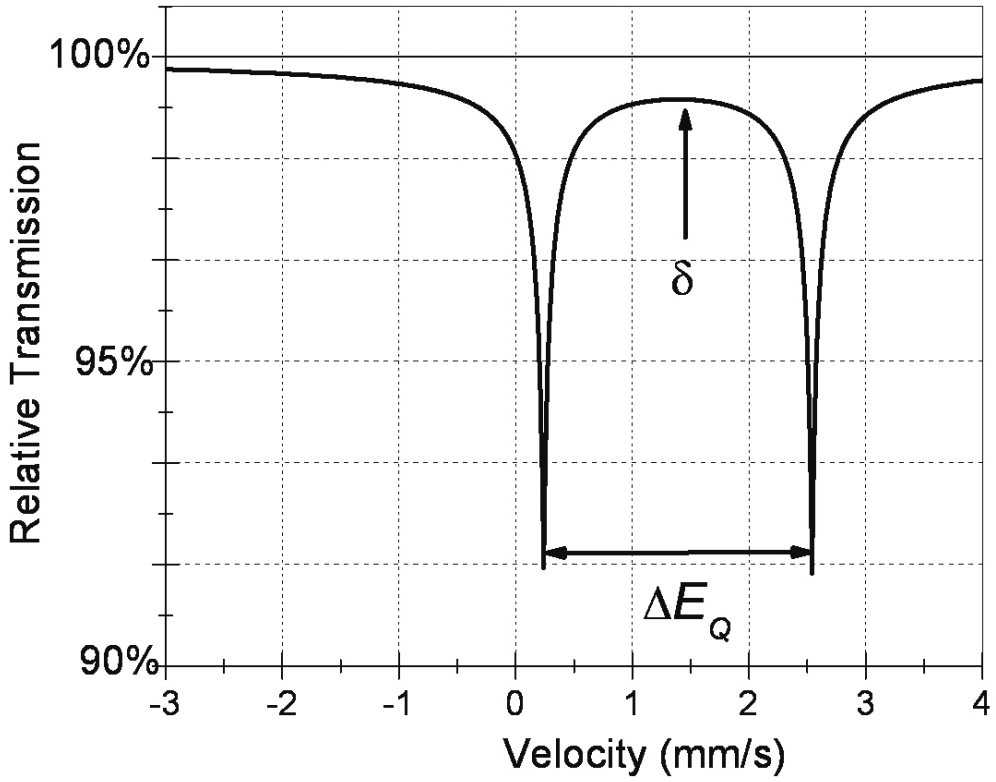


Fig. 6.56: An idealized Mössbauer spectrum showing both the isomer shift, δ, and the quadrupole splitting, ΔE_{Q}.

The isomer shift measures the shift in the energy of the γ-ray absorption relative to a standard, usually Fe foil. The isomer shift is sensitive to the electron density at the nucleus, and indirectly probes changes in the bonding of the valence orbitals due to variations in covalency and 3d shielding. Thus, it can be used to probe oxidation and spin states, and the coordination environment of the iron.

The quadrupole splitting arises from the interaction of the nuclear quadrupole moment of the excited state with the electric field gradient at the nucleus. The former is related to the non-spherical charge distribution in the excited state. As such it is extremely sensitive to the coordination environment and the geometry of the complex.

Both the isomer shift and quadrupole splitting can be successfully predicted using DFT methods. The isomer shift is directly related to the s electron density at the nucleus and can be calculated using the formula

$$\delta = \alpha(\rho_0 - C) + \beta \tag{6.14}$$

where α is a constant that depends on the change in the distribution of the nuclear charge upon absorption, and ρ_0 is the electron density at the nucleus [608]. The constants α and β are usually determined via linear regression analysis of the experimental isomer shifts versus the theoretically calculated electron density for a series of iron compounds with various oxidation and spin states. Since the electron density depends on the functional and basis set employed, fitting must be carried out for each combination used. A compilation of calibration constants (α, β and C) for various methods was assembled.[707] Usually an accuracy of better than 0.10 mm s^{-1} can be achieved for DFT with reasonably sized basis sets.

The quadrupole splitting is proportional to the largest component of the electric field gradient (EFG) tensor at the iron nucleus and can be calculated using the formula:

$$\Delta E_{\mathrm{Q}} = \frac{1}{2} eQV_{zz} \left(1 + \frac{\eta^2}{3}\right)^{\frac{1}{2}} \tag{6.15}$$

where e is the electrical charge of an electron and Q is the nuclear quadrupole moment of Fe (approximately 0.16 barns). V_{xx}, V_{yy} and V_{zz} are the electric field gradient tensors and η, defined as

$$\eta = \left| \frac{V_{xx} - V_{yy}}{V_{zz}} \right| \tag{6.16}$$

is the asymmetry parameter in a coordinate system chosen such that $|V_{zz}| \geq |V_{yy}| \geq |V_{xx}|$.

An example of how to calculate the electron density and quadrupole splitting of an iron center is as follows:

```
%eprnmr
  nuclei = all Fe {fgrad, rho}
end
```

If the core properties basis set CP(PPP) is employed, one might have to increase the radial integration accuracy for the iron atom. From ORCA 5.0 this is considered during grid construction and the defaults should work very well. However for very problematic cases it can be increased by controlling the `SPECIALGRIDINTACC` flag under `%METHOD` (see Sec. *Other details and options* for details).

The output file should contain the following lines, where you obtain the calculated quadrupole splitting directly and the RHO(0) value (the electron density at the iron nucleus). To obtain the isomer shift one has to insert the RHO(0) value into the appropriate linear equation (Eq. (6.14)).

```
Moessbauer quadrupole splitting parameter (proper coordinate system)
e**2qQ  = -0.406 MHz = -0.035 mm/s
eta     = 0.871
Delta-EQ=(1/2{e**2qQ}*sqrt(1+1/3*eta**2) = -0.227 MHz = -0.020 mm/s
RHO(0)=  11581.352209571 a.u.**-3   # the electron density at the Fe nucleus.
```

Note

- Following the same procedure, Mössbauer parameters can be computed with the CASSCF wavefunction. In case of a state-averaged CASSCF calculation, the averaged density is used in the subsequent Mössbauer calculation.

6.12.12 Broken-Symmetry Wavefunctions and Exchange Couplings

A popular way to estimate the phenomenological parameter J_{AB} that enters the Heisenberg–Dirac–van Vleck Hamiltonian which parameterizes the interaction between two spin systems is the "broken-symmetry" formalism. The phenomenological Hamiltonian is:

$$H_{\mathrm{HDvV}} = -2J_{\mathrm{AB}}\vec{S}_{\mathrm{A}}\vec{S}_{\mathrm{B}} \tag{6.17}$$

It is easy to show that such a Hamiltonian leads to a "ladder" of spin states from $S = S_{\mathrm{A}} + S_{\mathrm{B}}$ down to $S = |S_{\mathrm{A}} - S_{\mathrm{B}}|$. If the parameter J_{AB} is positive the systems "A" and "B" are said to be *ferromagnetically* coupled because the highest spin-state is lowest in energy while in the case that J_{AB} is negative the coupling is *antiferromagnetic* and the lowest spin state is lowest in energy.

In the broken symmetry formalism one tries to obtain a wavefunction that breaks spatial (and spin) symmetry. It may be thought of as a "poor man's MC-SCF" that simulates a multideterminantal character within a single determinant framework. Much could be said about the theoretical advantages, disadvantages, problems and assumptions that underly this approach. Here, we only want to show how this formalism is applied within ORCA.

For N_{A} unpaired electrons localized on "site A" and N_{B} unpaired electrons localized on a "site B" one can calculate the parameter J_{AB} from two separate spin-unrestricted SCF calculations: (a) the calculation for the high-spin state

with $S = \frac{(N_A+N_B)}{2}$ and (b) the "broken symmetry" calculation with $M_S = \frac{(N_A-N_B)}{2}$ that features N_A spin-up electrons that are quasi-localized on "site A" and N_B spin-down electrons that are quasi-localized on "site B". Several formalisms exist to extract J_{AB}: [91, 303, 636, 637, 808, 897].

$$J_{AB} = -\frac{(E_{HS} - E_{BS})}{(S_A + S_B)^2} \tag{6.18}$$

$$J_{AB} = -\frac{(E_{HS} - E_{BS})}{(S_A + S_B)(S_A + S_B + 1)} \tag{6.19}$$

$$J_{AB} = -\frac{(E_{HS} - E_{BS})}{\langle S^2\rangle_{HS} - \langle S^2\rangle_{BS}} \tag{6.20}$$

We prefer the last definition (Eq. (6.20)) because it is approximately valid over the whole coupling strength regime while the first equation implies the weak coupling limit and the second the strong coupling limit.

In order to apply the broken symmetry formalism use:

```
%scf BrokenSym NA,NB
     end
```

The program will then go through a number of steps. Essentially it computes an energy and wavefunction for the high-spin state, localizes the orbitals and reconverges to the broken symmetry state.

Caution

Make sure that in your input coordinates "site A" is the site that contains the larger number of unpaired electrons!

Most often the formalism is applied to spin coupling in transition metal complexes or metal-radical coupling or to the calculation of the potential energy surfaces in the case of homolytic bond cleavage. In general, hybrid DFT methods appear to give reasonable semiquantitative results for the experimentally observed splittings.

As an example consider the following simple calculation of the singlet–triplet splitting in a stretched Li_2 molecule:

```
#
# Example of a broken symmetry calculation
#
! B3LYP DEF2-SVP TightSCF
%scf BrokenSym 1,1
     end
* xyz 0 3
  Li   0   0   0
  Li   0   0   4
*
```

There is a second mechanism for generating broken-symmetry solutions in ORCA. This mechanism uses the individual spin densities and is invoked with the keywords `FlipSpin` and `FinalMs`. The strategy is to exchange the α and β spin blocks of the density on certain user-defined centers after converging the high-spin wavefunction. With this method arbitrary spin topologies should be accessible. The use is simple:

```
#
# Example of a broken symmetry calculation using the "FlipSpin" feature
#
! B3LYP DEF2-SVP TightSCF
%scf
     FlipSpin 1
       # Flip spin is a vector and you can give a list of atoms
       # on which you want to have the spin flipped. For example
       #  FlipSpin 17,38,56
       # REMEMBER: counting starts at atom 0!
     FinalMs  0
```

(continues on next page)

(continued from previous page)

```
      # The desired Ms value of the broken symmetry determinant.
      # This value MUST be given since the program cannot determine it by itself.
    end
* xyz 0 3
  Li  0  0  0
  Li  0  0  4
*
```

Finally, you may attempt to break the symmetry by using the SCF stability analysis feature (see Section *SCF Stability Analysis*). The ground spin state can be obtained by diagonalizing the above spin Hamiltonian through **ORCA-ECA** utility (see *orca_eca*).

Approximate Spin Projection Method

The approximate spin projection (AP) method, proposed by Yamaguchi and co-workers, is a technique to remove the spin contamination from exchange coupling constants.[737, 896, 897] In this scheme, the energy of a system is given by

$$E_{\mathrm{AP}} = \alpha E_{\mathrm{BS}} - (\alpha - 1)\, E_{\mathrm{HS}} \tag{6.21}$$

The parameter α is calculated as

$$\alpha = \frac{S^{\mathrm{HS}}\left(S^{\mathrm{HS}}+1\right) - S_Z^{\mathrm{BS}}\left(S_Z^{\mathrm{BS}}+1\right)}{\langle \hat{\mathbf{S}}^2 \rangle^{\mathrm{HS}} - \langle \hat{\mathbf{S}}^2 \rangle^{\mathrm{BS}}} \tag{6.22}$$

Here, S_Z^{BS} is the z-component of the total spin for the BS state ($S_Z^{\mathrm{BS}} = (N_\alpha - N_\beta)/2$). Alternatively, one can adopt Noodleman's scheme,[637] where α is calculated as follows

$$\alpha = \frac{S^{\mathrm{HS}}\left(S^{\mathrm{HS}}+1\right) - S_Z^{\mathrm{BS}}\left(S_Z^{\mathrm{BS}}+1\right)}{\left(S^{\mathrm{HS}}\right)^2} \tag{6.23}$$

or Ruiz's scheme,[733] with α equal to

$$\alpha = \frac{S^{\mathrm{HS}}\left(S^{\mathrm{HS}}+1\right) - S_Z^{\mathrm{BS}}\left(S_Z^{\mathrm{BS}}+1\right)}{S^{\mathrm{HS}}\left(S^{\mathrm{HS}}+1\right)} \tag{6.24}$$

The AP method is requested via the tag `APMethod` in the `%scf` block:

```
%scf BrokenSym NA,NB
     APMethod 3  # 1 = Noodleman
                 # 2 = Ruiz
                 # 3 = Yamaguchi
end
```

The default is `APMethod 0`, which corresponds to a normal BS calculation. Yamaguchi's AP method is available for single point energy calculations and geometry optimizations. If we run a geometry optimization in the context of Yamaguchi's AP method, then, the gradient of equation (6.21) w.r.t a nuclear displacement X reads as

$$\frac{\partial E_{\mathrm{AP}}}{\partial X} = \alpha \frac{\partial E_{\mathrm{BS}}}{\partial X} - (\alpha - 1)\frac{\partial E_{\mathrm{HS}}}{\partial X} + \frac{\partial \alpha}{\partial X}\left(E_{\mathrm{BS}} - E_{\mathrm{HS}}\right) \tag{6.25}$$

The last term contains the derivative $\frac{\partial \alpha}{\partial X}$. ORCA uses the formalism proposed by Saito and Thiel, which involves solving the CP-SCF equations in each geometry optimization cycle.[738] The cost of such a calculation is higher than using Noodleman's or Ruiz's schemes, where $\frac{\partial \alpha}{\partial X} = 0$.

6.13 Local Energy Decomposition

"Local Energy Decomposition" (LED) analysis[33, 105, 767] is a tool for obtaining insights into the nature of intermolecular interactions by decomposing the DLPNO-CCSD(T) energy into physically meaningful contributions. For instance, this approach can be used to decompose the DLPNO-CCSD(T) interaction energy between a pair of interacting fragments, as detailed in Section *Local Energy Decomposition*. A useful comparison of this scheme with alternative ways of decomposing interaction energies is reported in Ref. [27, 28, 29].

6.13.1 Closed shell LED

All that is required to obtain this decomposition in ORCA is to define the fragments and specify the `!LED` keyword in the simple input line.

LED decomposes separately the reference (Hartree-Fock) and correlation parts of the DLPNO-CCSD(T) energy. By default, the decomposition of the reference energy makes use of the RI-JK approximation. `An RIJCOSX variant, which is much more efficient and has a much more favorable scaling for large systems, is also available, as detailed in section` *Additional Features, Defaults and List of Keywords*, `and in` [27].

Note that, for weakly interacting systems, TightPNO settings are typically recommended. As an example, the interaction of H_2O with the carbene CH_2 at the PBE0-D3/def2-TZVP-optimized geometry can be analyzed within the LED framework using the following input file:

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO LED

*xyz 0 1
  C(1)      0.16044643459993      0.10093183177121      0.22603351276210
  H(1)      1.04516129053973     -0.06834886965353     -0.41865951747519
  H(1)     -0.12579332868173      1.14737893892114      0.00305818518771
  O(2)     -1.48285705560792     -1.31933824653169      2.29891474420047
  H(2)     -0.91417368674145     -0.93085192992263      1.60917234463506
  H(2)     -1.15648922489703     -2.21246650333085      2.42094328175662
*
```

The corresponding output file is reported below. The DLPNO-CCSD(T) energy components are printed out in different parts of the output as follows:

```
E(0)                                    ...   -114.913309038
E(CORR)(corrected)                      ...     -0.350582526
Triples Correction (T)                  ...     -0.006098691
E(CCSD(T))                              ...   -115.269990255
```

At the beginning of the LED part of the output, information on the fragments and the assignment of localized MOs to fragments are provided.

```
 ============================================================
         LOCAL ENERGY DECOMPOSITION FOR DLPNO-CC METHODS
 ============================================================

Maximum Iterations for the Localization .. 600
Tolerance for the Localization          ..  1.00e-06
Number of fragments = 2
Fragment  1:  0 1 2
Fragment  2:  3 4 5

Populations of internal orbitals onto Fragments:
   0:      1.000    0.000  assigned to fragment   1
   1:      0.000    1.000  assigned to fragment   2
   2:      1.022    0.008  assigned to fragment   1
   3:      0.001    0.999  assigned to fragment   2
```

(continues on next page)

(continued from previous page)

```
   4:      0.001    0.999  assigned to fragment   2
   5:      1.018    0.000  assigned to fragment   1
   6:      1.019    0.000  assigned to fragment   1
   7:      0.006    1.013  assigned to fragment   2
   8:      0.000    1.016  assigned to fragment   2
```

The decomposition of the Hatree-Fock energy into intra- and inter-fragment contributions follows. It is based on the localization of the occupied orbitals.

```
----------------------------------------
REFERENCE ENERGY E(0) DECOMPOSITION (Eh)
----------------------------------------

Nuclear repulsion     =        28.952049689006
One electron energy   =      -214.430545074583 (T=      114.825132245389, V=     -329.
↪255677319972)
Two electron energy   =        70.565186347093 (J=       71.658914661909, K=       -1.
↪093728314816)
                        ----------------------
Total energy          =      -114.913309038483
Consistency check     =      -114.913309038483 (sum of intra- and inter-fragment energies)

Kinetic energy        =       114.825132245389
Potential energy      =      -229.738441283873
                        ----------------------
Virial ratio          =         2.000767922417

-----------------------------------------
INTRA-FRAGMENT REF. ENERGY FOR FRAGMENT   1
-----------------------------------------

Nuclear repulsion     =         6.037208782874
One electron energy   =       -63.553431032444 (T=       38.870491681225, V=     -102.
↪423922713669)
Two electron energy   =        18.675214766985 (J=       18.935443192480, K=       -0.
↪260228425495)
                        ----------------------
Total energy          =       -38.841007482585

Kinetic energy        =        38.870491681225
Potential energy      =       -77.711499163811
                        ----------------------
Virial ratio          =         1.999241476056


-----------------------------------------
INTRA-FRAGMENT REF. ENERGY FOR FRAGMENT   2
-----------------------------------------

Nuclear repulsion     =         9.103529882464
One electron energy   =      -123.025684625357 (T=       75.954640564164, V=     -198.
↪980325189521)
Two electron energy   =        37.916781954190 (J=       38.739989208810, K=       -0.
↪823207254620)
                        ----------------------
Total energy          =       -76.005372788703

Kinetic energy        =        75.954640564164
Potential energy      =      -151.960013352867
                        ----------------------
Virial ratio          =         2.000667927913
```

(continues on next page)

(continued from previous page)

```
-------------------------------------------------
INTER-FRAGMENT REF. ENERGY FOR FRAGMENTs   2 AND   1
-------------------------------------------------

Nuclear repulsion     =        13.811311023669
Nuclear attraction    =       -27.851429416782
Coulomb repulsion     =        13.983482260618
                       ----------------------
Sum of electrostatics =        -0.056636132494 (  -35.540 kcal/mol)

Two electron exchange =        -0.010292634701 (   -6.459 kcal/mol)
                       ----------------------
Total REF. interaction =       -0.066928767195 (  -41.998 kcal/mol)



Sum of INTRA-fragment REF. energies  =      -114.846380271288
Sum of INTER-fragment REF. energies  =        -0.066928767195
                                      ----------------------
Total REF. energy                    =      -114.913309038483
```

Afterwards, a first decomposition of the correlation energy is carried out. The different energy contributions to the correlation energy (strong pairs, weak pairs and triples correction) are decomposed into intra- and inter-fragment contributions. This decomposition is carried out based on the localization of the occupied orbitals.

```
--------------------------------
CORRELATION ENERGY DECOMPOSITION
--------------------------------



-------------------------------------------------
INTER- vs INTRA-FRAGMENT CORRELATION ENERGIES (Eh)
-------------------------------------------------

                                Fragment   1             Fragment   2
                       ----------------------  ----------------------
Intra strong pairs              -0.136594658271        -0.209970193798  sum=       -0.
↪346564852069
Intra triples                   -0.002692277706        -0.002842791265  sum=       -0.
↪005535068971
Intra weak pairs                -0.000001573694        -0.000002311734  sum=       -0.
↪000003885429
Singles contribution             0.000000000611         0.000000001058  sum=        0.
↪000000001669
                       ----------------------  ----------------------
                                -0.139288509061        -0.212815295738  sum=       -0.
↪352103804799


Interaction correlation for Fragments   2 and   1:
-------------------------------------------------
Inter strong pairs               -0.003998018810 (   -2.509 kcal/mol)
Inter triples                    -0.000563621667 (   -0.354 kcal/mol)
Inter weak pairs                 -0.000015771468 (   -0.010 kcal/mol)
                                 ----------------------
Total interaction                -0.004577411945 (   -2.872 kcal/mol)
```

(continues on next page)

(continued from previous page)

```
Sum of INTRA-fragment correlation energies  =        -0.352103804799
Sum of INTER-fragment correlation energies  =        -0.004577411945
                                              ---------------------
Total correlation energy                    =        -0.356681216744
```

Afterwards, a summary with the decomposition of the total energy (reference energy + correlation) into intra- and inter-fragment contributions is printed.

```
-----------------------------------------
INTER- vs INTRA-FRAGMENT TOTAL ENERGIES (Eh)
-----------------------------------------

                               Fragment   1            Fragment   2
                      ----------------------  ----------------------
Intra REF. energy            -38.841007482585       -76.005372788703  sum=     -114.
↪846380271288
Intra Correlation energy      -0.139288509061        -0.212815295738  sum=       -0.
↪352103804799
                      ----------------------  ----------------------
                             -38.980295991646       -76.218188084441  sum=     -115.
↪198484076087


Interaction of Fragments   2 and   1:
-------------------------------------
Interfragment reference          -0.066928767195 (  -41.998 kcal/mol)
Interfragment correlation        -0.004577411945 (   -2.872 kcal/mol)
                        ----------------------
Total interaction                -0.071506179140 (  -44.871 kcal/mol)


Sum of INTRA-fragment total energies        =      -115.198484076087
Sum of INTER-fragment total energies        =        -0.071506179140
                                              ---------------------
Total energy                                =      -115.269990255228
```

Hence, the decomposition reported above allows one to decompose all the components of the DLPNO-CCSD(T) energy into intrafragment and interfragment contributions simply based on the localization of the occupied orbitals. In order to convert the intra-fragment energy components into contributions to the binding energy, single point energy calculations must be carried out on the isolated monomers, frozen in the geometry they have in the adduct, and the corresponding terms must be subtracted. Note that one can also include the geometrical deformation energy (also called "strain") by simply computing the energy of the geometrically relaxed fragments (see Section *Local Energy Decomposition* for further information).

For the DLPNO-CCSD strong pairs, which typically dominate the correlation energy, a more sophisticated decomposition, based on the localization of both occupied orbitals and PNOs, is also carried out and printed. Accordingly, the correlation energy from the strong pairs is divided into intra-fragment, dispersion and charge transfer components. Note that, due to the charge transfer excitations, the resulting intra-fragment contributions (shown below) differ from the ones obtained above.

```
-----------------------------------------
DECOMPOSITION OF CCSD STRONG PAIRS INTO
    DOUBLE EXCITATION TYPES (Eh)
-----------------------------------------

Foster-Boys localization is used for localizing PNOs

Intra fragment contributions:
INTRA Fragment 1                      -0.132849855
INTRA Fragment 2                      -0.209493798
```

(continues on next page)

(continued from previous page)

```
Charge transfer contributions:
Charge Transfer 1 to 2                  -0.005725404
Charge Transfer 2 to 1                  -0.000899609

Dispersion contributions:
Dispersion  2,1                         -0.001594204

Singles contributions:
Singles energy                           0.000000002
```

More detailed information into the terms reported above can be found in Section *Local Energy Decomposition* and in Ref.[767] All the individual double excitations contributions constituting the terms reported above can be printed by specifying "printlevel 3" in the `%mdci` block. Finally, a summary with the most important terms of the DLPNO-CCSD energy, which are typically discussed in standard applications, is printed.

```
-------------------------------------------------
FINAL SUMMARY DLPNO-CCSD ENERGY DECOMPOSITION (Eh)
-------------------------------------------------

Intrafragment REF. energy:
Intra fragment   1 (REF.)               -38.841007483
Intra fragment   2 (REF.)               -76.005372789

Interaction of fragments  2 and  1:
Electrostatics (REF.)                   -0.056636132
Exchange (REF.)                         -0.010292635
Dispersion (strong pairs)               -0.001594204
Dispersion (weak pairs)                 -0.000015771

Sum of non dispersive correlation terms:
Non dispersion (strong pairs)           -0.348968665
Non dispersion (weak pairs)             -0.000003885
```

Note that the "Non dispersion" terms include all the components of the correlation energy except London dispersion.[28, 106]. For the strong pairs, "Non dispersion" includes charge transfer, intrafragment double excitations and singles contributions. For the weak pairs, it corresponds to the intrafragment correlation contribution. In order to convert the non dispersion correlation components into contributions to the binding energy, single point energy calculations must be carried out on the isolated monomers.

6.13.2 Example: LED analysis of intermolecular interactions

The water-carbene example from the previous section will be used to demonstrate how to analyze intermolecular interactions between two fragments using the LED decomposition (note that all energies are given in a.u. if not denoted otherwise). As often done in practical applications, we will be using geometries optimized at the DFT (PBE0-D3/def2-TZVP) level of theory on which DLPNO-CCSD(T) (cc-pVDZ,TightPNO) single point energies are computed. Note that in practice, basis sets of triple-zeta quality or larger are recommended. In the first step, the geometries of the dimer, H_2O and CH_2 are optimized and DLPNO-CCSD(T) energies are computed to yield E^{opt}_{dimer} and $E^{opt}_{monomers}$. The input examples for the single-point DLPNO-CCSD(T) energies of the monomers at their optimized geometries and the necessary energies from the output files of these runs are as follows:

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO

# H2O optimized at the PBE0-D3/def2-TZVP level

*xyz 0 1
  O     -1.47291471015599     -1.29006364761118      2.29452038079177
  H     -0.88264582939506     -0.99404999457575      1.59835337186103
  H     -1.22136730983407     -2.20010680974562      2.46533021449572
*
```

```
E(0)                                    ...    -76.026656692
E(CORR)(corrected)                      ...     -0.211428886
Triples Correction (T)                  ...     -0.002932804
E(CCSD(T))                              ...    -76.241018382
```

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO

# CH2 optimized at the PBE0-D3/def2-TZVP level

*xyz 0 1
  C      0.16044643459993      0.10093183177121      0.22603351276210
  H      1.04516129053973     -0.06834886965353     -0.41865951747519
  H     -0.12579332868173      1.14737893892114      0.00305818518771
*
```

```
E(0)                                    ...    -38.881042677
E(CORR)(corrected)                      ...     -0.138447953
Triples Correction (T)                  ...     -0.002873032
E(CCSD(T))                              ...    -39.022363662
```

Single-point DLPNO-CCSD(T) energies of the monomers at their in-adduct geometries are also necessary. The corresponding inputs and the necessary output parts for these calculations are as follows:

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO

# H2O at the CH2-H2O geometry optimized at the PBE0-D3/def2-TZVP level

*xyz 0 1
  O     -1.48285705560792     -1.31933824653169      2.29891474420047
  H     -0.91417368674145     -0.93085192992263      1.60917234463506
  H     -1.15648922489703     -2.21246650333085      2.42094328175662
```

```
E(0)                                    ...    -76.026011663
E(CORR)(corrected)                      ...     -0.211931843
Triples Correction (T)                  ...     -0.002963338
E(CCSD(T))                              ...    -76.240906844
```

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO

# CH2 at the CH2-H2O geometry optimized at the PBE0-D3/def2-TZVP level

*xyz 0 1
  C      0.16044643459993      0.10093183177121      0.22603351276210
  H      1.04516129053973     -0.06834886965353     -0.41865951747519
  H     -0.12579332868173      1.14737893892114      0.00305818518771
*
```

```
E(0)                                    ...    -38.881085139
E(CORR)(corrected)                      ...     -0.138097323
Triples Correction (T)                  ...     -0.002869022
E(CCSD(T))                              ...    -39.022051484
```

These energies are summarized in Table Table 6.9).

Table 6.9: Energies of the H_2O-CH_2 example for illustrating how the different LED contributions are valuated. The superscript ^{opt} denotes energies of optimized structures, ^{fixed} denotes energies of isolated fragments in the dimer structure. In the last column the energy of the dimer is reported.

Energy [a.u.]	H_2O^{opt}	H_2O^{fixed}	CH_2^{opt}	CH_2^{fixed}	H_2O - CH_2
E_{HF}	-76.026656692	-76.026011663	-38.881042677	-38.881085139	-114.913309038
E_c CCSD	-0.211428886	-0.211931843	-0.138447953	-0.138097323	-0.350582526
E_c (T)	-0.002932804	-0.002963338	-0.002873032	-0.002869022	-0.006098691
E_{tot}	-76.241018382	-76.240906844	-39.022363662	-39.022051484	-115.269990255

Note that for this example, we do not include any BSSE correction. For this system we obtain a binding energy of

$$E_{int} = E^{opt}_{dimer} - E^{opt}_{monomers} = -115.269990255 \; - \; (-76.241018382 - 39.022363662) = -0.006608211$$

which is -4.147 kcal/mol.

The basic principles and the details of the LED are discussed in section *Local Energy Decomposition*. The first contribution to the binding energy is the energy penalty for the monomers to distort into the geometry of the dimer

$$\Delta E_{geo-prep} = E^{fixed}_{monomers} - E^{opt}_{monomers}$$

(see in equation (7.406)). This contribution is computed as the difference of the DLPNO-CCSD(T) energy of the monomers in the structure of the dimer ($E^{fixed}_{monomers}$) and of the relaxed monomers ($E^{opt}_{monomers}$). The following energies are obtained:

$$\Delta E_{geo-prep} = (-76.240906844 + 76.241018382) \; + \; (-39.022051484 + 39.022363662) = 0.000423716$$

which amounts to 0.266 kcal/mol. Typically, the triples correction is evaluated separately:

$$\Delta E^{C-(T)}_{int} = -0.006098691 - (-0.002963338 - 0.002869022) = -0.000266331$$

This contributes -0.167 kcal/mol. The next terms in equation (7.406) concern the reference energy contributions. The first one is the electronic preparation in the reference, which is evaluated as the difference of the `Intra REF. energy` of the fragments (see previous section) and the reference energy of the separate molecules at the dimer geometry:

$$\Delta E^{ref.}_{el-prep}(H_2O) = -76.005372788703 + 76.026011663 = 0.020638874297$$

$$\Delta E^{ref.}_{el-prep}(CH_2) = -38.841007482585 + 38.881085139 = 0.040077656415$$

which amounts to 0.060716530712 a.u. or 38.100 kcal/mol. The next two contributions stem from the decomposition of the reference inter-fragment contributions $E^{ref.}_{elstat} = -0.056636132$ (-35.540 kcal/mol) and $E^{ref.}_{exch} = -0.010292635$ (-6.459 kcal/mol) and can be found in directly in the LED output (`Electrostatics (REF.)` and `Exchange (REF.)`). The correlation energy also contains an electronic preparation contribution, but it is typically included in the correlation contribution $\Delta E^{C-CCSD}_{non-dispersion}$. Adding the non-dispersive strong and weak pairs contributions from the LED output (`Non dispersion (strong pairs)` and `Non dispersion (weak pairs)`) one obtains :

$$-0.348968665 - 0.000003885 = -0.34897255$$

from which we have to subtract the sum of the correlation contributions of the monomers at the dimer geometry

$$\Delta E^{C-CCSD}_{non-dispersion} = -0.34897255 - (-0.211931843 - 0.138097323) = 0.001056616$$

which is 0.663 kcal/mol. The dispersive contribution can be directly found in the LED output (`Dispersion (strong pairs)` and `Dispersion (weak pairs)`) and amounts to $E^{C-CCSD}_{dispersion} = -0.001609975$ which is -1.010 kcal/mol. So all terms that we have evaluated so far are:

$$\Delta E = \Delta E_{geo-prep} + \Delta E^{ref.}_{el-prep} + E^{ref.}_{elstat} + E^{ref.}_{exch} + \Delta E^{C-CCSD}_{non-dispersion} + E^{C-CCSD}_{dispersion} + \Delta E^{C-(T)}_{int}$$

ΔE	$\Delta E_{geo-prep}$	$\Delta E^{ref.}_{el-prep}$	$E^{ref.}_{elstat}$	$E^{ref.}_{exch}$	$\Delta E^{C-CCSI}_{non-disp}$	$E^{C-CCSD}_{disp.}$	$\Delta E^{C-(T)}_{int}$
a.u.	0.00042371(	0.06071653071	- 0.056636132	- 0.010292635	0.00105661(	- 0.001609975	- 0.000266331
kcal/mol	0.266	38.100	-35.540	-6.459	0.663	-1.010	-0.167

which sum to the total binding energies of -0.006608211 a.u. or -4.147 kcal/mol that we have evaluated at the beginning of this section. A detailed discussion of the underlying physics and chemistry can be found in [33].

6.13.3 Open shell LED

The decomposition of the DLPNO-CCSD(T) energy in the open shell case is carried out similarly to the closed shell case. [33] An example of input file is shown below.

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO LED

*xyz 0 3
C(1)      0.32786304018834      0.25137292674595      0.32985672433579
H(1)      0.78308855251826     -0.37244139824620     -0.42204823336026
H(1)     -0.19639272865450      1.19309490346756      0.33713773666060
O(2)     -1.47005964014997     -1.60804001777555      1.84974416203666
H(2)     -0.89305417808014     -1.00736849071095      1.35216686141176
H(2)     -1.02515061661047     -1.73931270222718      2.69260529998224
*
```

The corresponding output is entirely equivalent to the one just discussed for the closed shell case. However, the open shell variant of the LED scheme relies on a different implementation than the closed shell one. A few important differences exist between the two implementations, which are listed below.

- In the closed shell LED the reference energy is typically the HF energy. Hence, the total energy equals the sum of HF and correlation energies. In the open shell variant, the reference energy is the energy of the QRO determinant. Hence, the total energy in this case equals the sum of the energy of the QRO determinant and the correlation energy.
- The singles contribution is typically negligible in the closed shell case due to the Brillouin's theorem. In the open shell variant, this is not necessarily the case. In both cases, the singles contribution is included in the "Non dispersion" part of the strong pairs.
- In the UHF DLPNO-CCSD(T) framework we have $\alpha\alpha$, $\beta\beta$ and $\alpha\beta$ pairs. Hence, in the open shell LED, all correlation terms (e.g. London dispersion) have $\alpha\alpha$, $\beta\beta$ and $\alpha\beta$ contributions. By adding "printlevel 3" in the `%mdci` block one can obtain information on the relative importance of the different spin terms.
- The open shell DLPNO-CCSD(T) algorithm can also be used for computing the energy of closed shell systems by adding the "UHF" keyword in the simple input line of a DLPNO-CCSD(T) calculation.

6.13.4 Dispersion Interaction Density plot

The Dispersion Interaction Density (DID) plot provides a simple yet powerful tool for the spatial analysis of the London dispersion interaction between a pair of fragments extracted from the LED analysis in the DLPNO-CCSD(T) framework. [29] A similar scheme was developed for the closed shell local MP2 method. [893] The "dispersion energy density", which is necessary for generating the DID plot, can be obtained from a simple LED calculation by adding "DoDIDplot true" in the `%mdci` block.

```
!DLPNO-CCSD(T) ... LED
%mdci DoDIDplot true end
```

These can be converted to a format readable by standard visualization programs, e.g. a cube file, through `orca_plot`. To do that, call `orca_plot` with the command:

```
orca_plot gbwfilename -i
```

and follow the instructions that will appear on your screen, i.e.:

```
1 - Enter type of plot
2 - Enter no of orbital to plot
3 - Enter operator of orbital (0=alpha,1=beta)
4 - Enter number of grid intervals
5 - Select output file format
6 - Plot CIS/TD-DFT difference densities
7 - Plot CIS/TD-DFT transition densities
```

(continues on next page)

(continued from previous page)

```
8 - Set AO(=1) vs MO(=0) to plot
9 - List all available densities

10 - Generate the plot
11 - exit this program
```

Type "1" for selecting the plot type. A few options are possible:

```
1 -   molecular orbitals
2 -   (scf) electron density      ......  (scfp  )  - available
3 -   (scf) spin density          ......  (scfr  )  - available
4 -   natural orbitals
5 -   corresponding orbitals
6 -   atomic orbitals
7 -   mdci electron density       ......  (mdcip )  - NOT available
8 -   mdci spin density           ......  (mdcir )  - NOT available
9 -   OO-RI-MP2 density           ......  (pmp2re)  - NOT available
10 -   OO-RI-MP2 spin density      ......  (pmp2ur)  - NOT available
11 -   MP2 relaxed density         ......  (pmp2re)  - NOT available
12 -   MP2 unrelaxed density       ......  (pmp2ur)  - NOT available
13 -   MP2 relaxed spin density    ......  (rmp2re)  - NOT available
14 -   MP2 unrelaxed spin density  ......  (rmp2ur)  - NOT available
15 -   LED dispersion interaction density  (ded21 )  - available
16 -   Atom pair density
17 -   Shielding Tensors
18 -   Polarisability Tensor
```

Select "LED dispersion interaction density" from the list by typing "15". Afterwards, choose your favorite format and generate the file.

6.13.5 Automatic Fragmentation

Starting from ORCA 4.2 it is possible to let the program define the fragments to be used in the LED analysis. In this case, the program will try to identify all monomers in the system that are not connected through a covalent bond and assign a fragment to each of them. The XYZ coordinates of the fragments are reported in the beginning of the output file. For instance, given the input:

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO LED

*xyz 0 1
C    0.18726407287156      0.08210467619149      0.19811955853244
H    1.07120465088431     -0.00229078749404     -0.46002874025040
H   -0.15524644515923      1.12171178448874      0.04316776563623
O   -1.47509614629583     -1.29358571885374      2.29818864036820
H   -0.87783948760158     -0.98540169212890      1.58987042714267
H   -1.22399221548771     -2.20523304094991      2.47014489963764
*
```

The program will automatically identify the H_2O and the CH_2 fragments. Note that this procedure works for an arbitrary number of interacting molecules. It is also possible to assign only certain atoms to a fragment and let the program define the other ones:

```
! dlpno-ccsd(t) cc-pvdz cc-pvdz/c cc-pvtz/jk rijk verytightscf TightPNO LED

*xyz 0 1
C(1)    0.18726407287156      0.08210467619149      0.19811955853244
H(1)    1.07120465088431     -0.00229078749404     -0.46002874025040
H(1)   -0.15524644515923      1.12171178448874      0.04316776563623
O   -1.47509614629583     -1.29358571885374      2.29818864036820
```

(continues on next page)

(continued from previous page)

```
H   -0.87783948760158     -0.98540169212890      1.58987042714267
H   -1.22399221548771     -2.20523304094991      2.47014489963764
*
```

6.13.6 Additional Features, Defaults and List of Keywords

Note

Starting from ORCA 4.2 the default localization scheme for the PNOs has changed from PM (Pipek Mezey) to FB (Foster Boys). This might cause slight numerical differences in the LED terms with respect to that obtained from previous ORCA versions. To obtain results that are fully consistent with previous ORCA versions, PM must be specified (see below).

The following options can be used in accordance with LED.

```
! DLPNO-CCSD(T) cc-pVDZ cc-pVDZ/C cc-pVTZ/JK RIJK TightPNO LED TightSCF

%mdci
  LED 1               # localization method for the PNOs. Choices:
                      # 1 = PipekMezey
                      # 2 = FosterBoys (default starting from ORCA 4.2)
  PrintLevel 3        # Selects large output for LED and prints the
                      # detailed contribution
                      # of each DLPNO-CCSD strong pair
  LocMaxIterLed 600   # Maximum number of localization iterations for PNOs
  LocTolLed     1e-6  # Absolute threshold for the localization procedure for PNOs
  Maxiter 0           # Skips the CCSD iterations and
                      # the decomposition of the correlation energy
  DoLEDHF true        # Decomposes the reference energy in the LED part.
                      # By default, it is set to true.
end
```

Note

Starting from ORCA 4.2 an RIJCOSX implementation of the LED scheme for the decomposition of the reference energy is also available. This is extremely efficent for large systems. For consistency, the RIJCOSX variant of the LED is used only if the underlying SCF treatment is performed using the RIJCOSX approximation, i.e., if RIJCOSX is specified in the simple input line. An example of input follows.

```
!  dlpno-ccsd(t) def2-TZVP def2-TZVP/C def2/j rijcosx verytightscf TightPNO LED

*xyz 0 1
C(1)    0.18726407287156      0.08210467619149      0.19811955853244
H(1)    1.07120465088431     -0.00229078749404     -0.46002874025040
H(1)   -0.15524644515923      1.12171178448874      0.04316776563623
O(2)   -1.47509614629583     -1.29358571885374      2.29818864036820
H(2)   -0.87783948760158     -0.98540169212890      1.58987042714267
H(2)   -1.22399221548771     -2.20523304094991      2.47014489963764
*
```

Fianlly, here are some tips for advanced users.

- The LED scheme can be used in conjuction with an arbitrary number of fragments.
- The LED scheme can be used to decompose DLPNO-CCSD and DLPNO-CCSD(T) energies. At the moment, it is not possible to use this scheme to decompose DLPNO-MP2 energies directly. However, for closed

shell systems, one can obtain DLPNO-MP2 energies from a DLPNO-CCSD calculation by adding a series of keywords in the `%mdci` block: (i) `TScalePairsMP2PreScr 0` ; (ii) `UseFullLMP2Guess true`; (iii) `TCutPairs 10` (or any large value). The LED can be used as usual to decompose the resulting energy.

- For a closed shell system of two fragments (say A and B), the LED scheme can be used to further decompose the LED components of the reference HF energy (intrafragment, electrostatics and exchange) into a sum of frozen state and orbital relaxation correction contributions. More information can be found in Ref. [29].
- To obtain the frozen state terms one has to: (i) generate a .gbw file containing the orbitals of both fragments (AB.gbw) using `orca_mergefrag A.gbw B.gbw AB.gbw`, where A.gbw and B.gbw are the orbital files of isolated fragments at the aduct geometry; (ii) run the LED as usual by using `MORead` to read the orbitals in the AB.gbw file in conjunction with `Maxiter 0` in both the `%scf` block (to skip the SCF iterations) and the `%mdci` block (to skip the unnecessary CCSD iterations).

6.14 The Hartree-Fock plus London Dispersion (HFLD) method for the study of Noncovalent Interactions

Starting from ORCA 4.2, the efficient and accurate HFLD method[30] can be used for the quantification and analysis of noncovalent interactions between a pair of user-defined fragments. Starting from ORCA 5.0, an open shell variant of the HFLD method is also available.[24]

The leading idea here is to solve the DLPNO coupled cluster equations while neglecting intramonomer correlation. The LED scheme is then used to extract the London dispersion (LD) energy from the intermolecular part of the correlation. Finally, the resulting LD energy is used to correct interaction energies computed at the HF level. Hence, HFLD can be considered as a dispersion-corrected HF approach in which the dispersion interaction between the fragments is added at the DLPNO-CC level. As such, it is particulartly accurate for the quantification of noncovalent interactions such as those found in H-bonded systems, pre-reactive intermediates (e.g., Frustrated Lewis Pairs), dispersion and electrostatically bound systems. Larger errors are in principle expected for transition metal complexes, as it is the case for any dispersion corrected Hartree-Fock scheme.

The efficency of the approach allows the study of noncovalent interactions in systems with several hundreds of atoms. An input example is reported below.

```
!  HFLD aug-cc-pvdz aug-cc-pvdz/C verytightscf

*xyz 0 1
C(1)    0.18726407287156      0.08210467619149      0.19811955853244
H(1)    1.07120465088431     -0.00229078749404     -0.46002874025040
H(1)   -0.15524644515923      1.12171178448874      0.04316776563623
O(2)   -1.47509614629583     -1.29358571885374      2.29818864036820
H(2)   -0.87783948760158     -0.98540169212890      1.58987042714267
H(2)   -1.22399221548771     -2.20523304094991      2.47014489963764
*
```

In the corresponding output, after the DLPNO-CC iterations and the LED output, the following information is printed:

```
-----------------------------   -----------------
Inter-fragment dispersion        -0.001871763
-----------------------------   -----------------



-------------------------   --------------------
FINAL SINGLE POINT ENERGY      -114.932878050741
-------------------------   --------------------
```

The total HFLD energy of the adduct is thus -114.932878050741 a.u.. To compute interaction energies, we have to substract from this value the Hartree-Fock energies of the monomers in the geometry they have in the complex, i.e., -38.884413525377 and -76.040412827089 a.u. for CH_2 and H_2O, respectively. The total interaction energy is thus -0.00805 a.u. or -5.1 kcal/mol (the corresponding DLPNO-CCSD(T)/TightPNO/CBS value is -5.3 kcal/mol. [33]).

Note that, to obtain binding energies, the geometric preparation should be added to this value. This can be computed using a standard computational method, e.g, DFT or DLPNO-CCSD(T).

Some of the most important aspects of the method are summarized below:

- `Accuracy and Recommended Settings.` For noncovalent interactions, HFLD typically provides an accuracy comparable to that of the DLPNO-CCSD(T) method if default PNO settings are used. For the HFLD scheme, these are defined as TCutPNO = 3.3e-7 and TCutPairs 1e-5. If used in conjuction with a def2-TZVP(-f) basis set, these settings are typically denoted as "`HFLD*`" and are recommeded for standard applications on large systems.[24] For example, `HFLD*` settings were used in Ref.[25] to elucidate the complex pattern of interactions responsible for the stability of the DNA duplex. If great accuracy is required, it is recommened to use TightPNO settings in conjuction with TCutPNO 1e-8 and two-point basis set extrapolation (aug-cc-pVTZ/aug-cc-pVQZ) to approach the CBS limit. These settings are typically denoted as the "`gold`" HFLD settings.[24]
- `Reference determinant in the Open shell HFLD scheme.` In the open shell case, HFLD relies on a restricted reference determinant for the calculation of the LD energy. If the QRO determinant is used as reference, the reference interaction energy can in principle be computed at the UHF or QRO levels. This leads to two different schemes, namely the QRO/HFLD and UHF/HFLD. Alternatively, the restricted open-shell HF (ROHF) determinant can be used as reference in HFLD calculations, which leads to the ROHF/HFLD approach. The energy value reported as "FINAL SINGLE POINT ENERGY" in the output corresponds to the UHF/HFLD scheme by default, which is typically slightly more accurate. See Ref. [24] for details.
- `Efficency.` The calculation of the dispersion correction typically requires the same time as an HF calculation. This is true for small as well as for large systems.
- `Analysis of Intermolecular Interactions.` The HFLD method can be combined with the Local Energy Decomposition (LED) to study intermolecular interactions in great detail. The LED dispersion energy obtained with HFLD is often very close to that obtained from a full DLPNO-CCSD(T) calculation. Hence, HFLD can be used as a cost-effective alternative to DLPNO-CCSD(T) to study, among other things, the importance of London dispersion in molecular chemistry.
- `Additional considerations.` (i) One can specify "`NormalPNO`" or "`TightPNO`" settings in the simple input line. The corresponding DLPNO tresholds would be in this case fully consistent with those used in the DLPNO-CCSD(T) method. (ii) The dispersion energy in the HFLD approach slightly depends by the choice of the localization scheme used for occupied orbitals and PNOs. Default settings are recommended for all intent and purposes. However, it is important to note that the localization iterations for occupied and virtual orbitals must be fully converged in order to obtain consistent results. To achieve this goal, it might be necessary to increase "`LocMaxIter`" or "`LocMaxIterLed`" (see below). However, this is typically necessary only if very large basis sets (e.g. aug-cc-pV5Z) are used. (iii) Importantly, the method benefits from the use of tightly converged SCF solutions. For closed-shell systems, a useful diagnostic in this context is the "Singles energy" term that is printed in the LED part of the output. This term must be smaller than 1e-6 for closed shell species. If this is not the case, one should change the settings used for the SCF iterations. Note also that all the features of the LED scheme (e.g. automatic fragmentation) are also available for the HFLD method.

Note that, as HFLD relies on both the DLPNO-CCSD(T) and LED methods, the options of both schemes can be used in principle in conjuction with HFLD. Some examples are shown below:

```
! HFLD aug-cc-pVDZ aug-cc-pVDZ/C aug-cc-pVTZ/JK RIJK TightSCF

%mdci
  LED 1               # localization method for the PNOs. Choices:
                      # 1 = PipekMezey
                      # 2 = FosterBoys (default, recommeded for the HFLD method)
  PrintLevel 3        # Selects large output for LED and prints the
                      # detailed contribution
                      # of each DLPNO-CCSD strong pair
  LocMaxIterLed 600   # Maximum number of localization iterations for PNOs
  LocMaxIter    300   # Maximum number of localization iterations for
                      # occupied orbitals
  LocTolLed     1e-6  # Absolute threshold for the localization procedure for PNOs
```

(continues on next page)

(continued from previous page)

```
  DoLEDHF true        # Decomposes the reference energy in the LED part.
                      # By default, it is set to false in HFLD for efficency reasons.
  TCutPNO    3.33e-7  # cutoff for PNO occupation numbers.
  TCutPairs  1e-5     # cutoff for estimated pair correlation energies
                      # to be included in the CC treatment
end
```

6.15 ORCA MM Module

Since version 4.2 ORCA features its own independent MM implementation.

The minimum input necessary for a MM treatment looks as follows.

```
!MM
%mm
 ORCAFFFilename "UBQ.ORCAFF.prms"
end
```

In this section we discuss the basic keywords and options, i.e.

- the basic structure of the ORCA Forcefield File,
- how to generate the ORCA Forcefield File,
- how to manipulate the ORCA Forcefield File,
- how to speed up MM calculations,
- further MM options and keywords.

Further options important for QM/MM calculations will be discussed in section *ORCA Multiscale Implementation*.

6.15.1 ORCA Forcefield File

For the MM part of the QM/MM calculation force-field parameters are necessary. ORCA has its own parameter file format (ORCA forcefield file - ORCAFF.prms), which includes the atom specific parameters for nonbonded interactions:

- partial charge
- LJ coefficients

and parameters for bonded interactions:

- bonds
- angles
- Urey-Bradley terms
- dihedrals
- impropers
- CMAP terms (cross-terms for backbone, currently not used)

Individual parameters, like e.g. atomic charge, equilibrium bond length and force constant, . . . , can be conveniently modified directly within the ORCA Forcefield File.

How to generate the ORCA Forcefield File

The easiest way to generate a ORCAFF.prms file is currently to convert from psf (protein structure file) files. Psf files are specific to the CHARMM forcefield and its application via NAMD. Psf files for a specific protein system can easily be generated by the popular molecular visualization program VMD and its extension QwikMD, but also with other extensions in the VMD program (e.g. psfgen or fftk). The psf file contains information on the atom types and on the bonded interactions of all atoms. It does, however, not contain the parameters that belong to these interactions. These parameteres are stored in specific files, often ending with prm, but also par or str. The CHARMM parameter files come with VMD installation, can be directly downloaded, or can be generated with the VMD extension fftk (forcefield toolkit).

Once a ORCAFF.prms file is available, it can be manipulated, i.e. split up into several parts for individual molecules, new ORCAFF.prms files can be generated for non-standard molecules, and individual ORCAFF.prms files can be merged, as described in the following:

Conversion from psf or prmtop files to ORCAFF.prms: convff

The orca_mm module can convert psf and prm files (CHARMM), prmtop files (AMBER) or xml files (open force field from the openff toolkit, compatible to AMBER) to the ORCAFF file with the -convff flag. Input options are:

```
orca_mm -convff <optional:-verbose> <FFInput> <PSFFILE> <PRMFILE(S)>

Keywords:
          <FFInput> = -CHARMM
          <FFInput> = -AMBER
          <FFInput> = -OPENMM
```

For CHARMM topologies, when a psf file is available for a system with standard residues, prepared by e.g. QwikMD, psfgen or other vmd tools, the conversion needs the psf plus the prm files as input:

```
CHARMM example:
orca_mm -convff -CHARMM 1C1E.psf par_all36_prot.prm toppar_water_ions_namd.str
```

ORCA can also convert Amber topologies to the ORCAFF file. Here, only the prmtop file is required:

```
AMBER example:
orca_mm -convff -AMBER complex.prmtop
```

ORCA can also convert xml files from the openff toolkit (AMBER compatible) to the ORCAFF file. Here, only the xml file is required:

```
OPENFF example:
orca_mm -convff -OPENMM complex.xml
```

Divide a forcefield file: splitff

If a ORCAFF.prms file should be subdivided into several files, e.g. if the psf file stems from QWikMD with non-standard molecules included, e.g. a ligand. In that case first the parameters of the ligand are split from the remaining system, next the ligand needs to be protonated, then a simple ORCAFF.prms file is generated via orca_mm's makeff option, and finally the ligand's new ORCAFF.prms file is added to the main systems file via the above described mergeff option. Note that the file can only be split into files for nonbonded fragments.

Input options:

```
orca_mm -splitff <optional:-verbose> <ORCAFFFILE> <A1> <optional:A2> ...

Keywords:
```

(continues on next page)

(continued from previous page)

```
          <ORCAFFFILE>  = ORCA forcefield file.
          <A1>          = Atom number of first atom of fragment that should belong
                          to a new ORCA forcefield file
          <A2>          = Atom number of first atom of fragment that should belong
                          to a new ORCA forcefield file
          ...           = More split atoms possible
          Note that atoms start counting at 1.

Example:
orca_mm -splitff 1C1E_substrate_noH.ORCAFF.prms 7208
```

Merge forcefield files: mergeff

If several ORCAFF.prms files are available and should be merged for an ORCA calculation, e.g. for a standard plus a non-standard molecule.

Input options:

```
orca_mm -mergeff <optional:-verbose> <ORCAFFFILE1> <ORCAFFFILE2> ...

Keywords:

          <ORCAFFFILE1>   = First ORCA forcefield file
          <ORCAFFFILE2>   = Second ORCA forcefield file
          ...             = More ORCA forcefield files possible

Example:
orca_mm -mergeff 1C1E.ORCAFF.prms substrate_withH.ORCAFF.prms
```

Repeat forcefield files: repeatff

In case the same ORCAFF.prms file needs to be repeated multiple times, the repeatff functionality is available.

Input options:

```
orca_mm -repeatff <optional:-verbose> <ORCAFFFILE> <A>
          <ORCAFFFILE>      = ORCA forcefield file.
          <A>               = Factor (integer) defining how often this forcefield file should␣
↪be repeated.

Example:
orca_mm -mergeff methanol.ORCAFF.prms 580
```

This feature can be useful e.g. in the case of solvating a molecule, i.e. adding multiple copies of a solvent to a solute. First the solvent can be repeated N times, and subsequently the solute's prms file can be merged together with the solvent prms file.

Divide a forcefield file: splitpdb

When splitting a ORCAFF.prms file, also splitting of the pdb file is required. The file can be split into an arbitrary number of individual files.

This can be useful together with the splitff command.

Input options:

```
orca_mm -splitpdb <optional:-verbose> <PDBFILE> <A1> <optional:A2> ...

Keywords:

          <PDBFILE> = PDB file.
          <A1>      = Atom number of first atom of fragment that should belong
                      to a new ORCA forcefield file
          <A2>      = Atom number of first atom of fragment that should belong
                      to a new ORCA forcefield file
          ...       = More split atoms possible
          Note that atoms start counting at 1.

Example:
orca_mm -splitpdb 1C1E_substrate_noH.pdb 7208
```

Merge PDB files: mergepdb

If several PDB files are available and need to be merged for an ORCA calculation, e.g. a protein and a ligand or multiple ligands, or a ligand that was first removed from a complex, then modified, and finally should get back into the complex PDB file.

This can be useful together with the mergeff command.

Input options:

```
orca_mm -mergepdb <optional:-verbose> <1PDBFILE> <2PDBFILE> ...

Keywords:

          <1PDBFILE>   = First PDB file
          <2PDBFILE>   = Second PDB file
          ...          = More PDB files possible

Example:
orca_mm -mergepdb 1C1E.pdb substrate_withH.pdb
```

Create simple force field: makeff

The orca_mm tool can generate an approximate forcefield for a molecule, storing it in ORCAFF.prms format. Here, the LJ coefficients are based on UFF parameters and the partial charges are based on a simple PBE or XTB calculation. The resulting topology is certainly not as accurate as an original CHARMM topology, but can still be used for an approximate handling of the molecule. Herewith, the molecule can be part of the QM region (having at least the necessary LJ coefficients), or part of the MM region as a non-active spectator - being not too close to the region of interest. In the latter case it is important that the molecule is not active, since bonded parameters are not available. However, it can still be optimized as a rigid body, see sections *Geometry Optimizations using the L-BFGS optimizer* and the usage in QM/MM calculations in section *Optimization with the Cartesian L-BFGS Minimizer*, on MM level, optimizing its position and orientation with respect to the specific environment.

Input options:

```
orca_mm -makeff <optional:-verbose> <XYZ/PDBFILE> <optional:-C CHARGE>
                <optional:-M MULT> <optional:-nproc N> <optional:-CHARGE_OPTIONS>

Keywords:

          <CHARGE>              = charge of system
          <MULT>                = multiplicity of system
          <-nproc N>            = number of processors (Default 1)
          <-CHARGE_OPTIONS>     =   Structure            Charge calculation
               <-PBE>                input                    PBE
               <-PBEOpt>             PBE opt.                 PBE
               <-PBEOptH>            PBE H-opt.               PBE
               <-XTB>                input                    XTB
               <-XTBOpt>             XTB opt.                 XTB
               <-XTBOptH>            XTB H-opt.               XTB
               <-XTBOptPBE>          XTB opt.                 PBE
               <-noChargeCalc>       distribute net charge evenly

               PBE Opt and SP level: RI-PBE/def2-SVP CPCM(Water), CHELPG charges
               XTB Opt and SP level: GFN2-XTB GBSA(Water), Mulliken partial charges

Example:
orca_mm -makeff substrate_withH.xyz -M 2 -XTBOptPBE
```

Note that ORCA generates bonds based on simple distance rules, which enables ORCA to detect where to add link atoms between QM and MM atoms, see also section *QM-MM, QM-QM2 and QM2-MM Boundary*. But the user is advised to treat a molecule, for which the ORCAFF.prms file was generated with the makeff option, either fully in the QM, or fully in the MM region, unless the charge distribution has been properly taken care of (due to the need of integer charges in QM and MM system).

Get standard hydrogen bond lengths: getHDist

For the definition of the link atoms standard bond lengths between C, N and O and hydrogen are directly set by ORCA but can be modified by the user, see section *QM-MM, QM-QM2 and QM2-MM Boundary*. If other atom types are on the QM side of the QM-MM boundary, their distance to the link atom has to be defined. In this case a file can be provided to ORCA which defines the standard bond length to hydrogen for all possible atoms. Such a file can be generated via the following command:

Input options:

```
orca_mm -getHDist <optional:-verbose> <XYZ/PDBFILE>

Example:
orca_mm -getHDist 1C1E.xyz
```

This file can then be modified, the required values can be added, and the resulting file can be defined as input for the QMMM calculation.

Create ORCAFF.prms file for IONIC-CRYSTAL-QMMM

For IONIC-CRYSTAL-QMMM calculations, section *IONIC-CRYSTAL-QMMM*, an ORCAFF.prms file with initial charges and connectivities is required. If you are not using the orca_crystalprep tool for setting up such calculations, see section *orca_crystalprep*, you can directly prepare the ORCAFF.prms file with the command:

```
orca_mm -makeff <XYZ/PDBFILE> -CEL <ELEMENT1> <OXIDATION_STATE1>
                              -CEL <ELEMENT2> <OXIDATION_STATE2>

Keywords:
```

(continues on next page)

(continued from previous page)

```
         <ELEMENT1>             = element name
         <OXIDATION_STATE1>     = formal oxidation state of element 1
         ...                    = More elements possible

Example:
orca_mm -makeff na4cl4.xyz -CEL Na 1.0 -CEL Cl -1.0
```

Here, na4cl4.xyz is the supercell structure file (it can contain tens of thousands of atoms).

Note

- For supercells with more complex oxidation states', e.g. Co_3O_4, the ORCAFF.prms file can be generated conveniently via the orca_crystalprep tool, *orca_crystalprep*.

6.15.2 Speeding Up Nonbonded Interaction Calculation

For MM calculations of very large systems with hundreds of thousands of atoms, and for QMMM calculations with fast QM methods (e.g. XTB, AM1) and / or very small QM systems, the computation of the nonbonded interactions can become a bottleneck. Different schemes for speeding up the calculation of nonbonded interactions are available within the ORCA MM implementation. Two schemes truncate long-range interaction, another scheme can be used for calculations with active regions, i.e. calculations where only a part of the system is active or optimized. For more information on active regions see section *Active and Non-Active Atoms - Optimization, Frequency Calculation, Molecular Dynamics and Rigid MM Water*.

Force Switching for LJ Interaction

With force switching for the LJ interaction (described in reference [818]) a smooth switching function is used to truncate the LJ potential energy smoothly at the outer cutoff distance LJCutOffOuter. If switching is set to false, the LJ interaction is not truncated at LJCutOffOuter. The parameter LJCutOffInner specifies the distance at which the switching function starts taking an effect to bring the van der Waals potential to 0 smoothly at the LJCutOffOuter distance, ensuring that the force goes down to zero at LJCutOffOuter, without introducing discontinuities. Note that LJCutOffInner must always be smaller than LJCutOffOuter.

```
%mm
 SwitchForceLJ true       # Use the switch force scheme for the LJ interaction.
                          # Default: true.
 LJCutOffInner 10.        # Distance (in Ang). Default: 10 Ang.
 LJCutOffOuter 12.        # Distance (in Ang). Default: 12 Ang.
end
```

Force Shifting for Electrostatic Interaction

With force shifting for the electrostatic interaction (described in reference [818]) the electrostatic potential is shifted to zero at the cutoff distance CoulombCutOff. If shifting is set to false, the electrostatic interaction is not truncated at CoulombCutOff.

```
%mm
 ShiftForceCoulomb true   # Use the shift force scheme for the Coulomb interaction.
                          # Default: true.
 CoulombCutOff 12.        # Distance (in Ang). Default: 12 Ang.
end
```

Neglecting Nonbonded Interactions Within Non-Active Region

When using active regions (see section *Active and Non-Active Atoms - Optimization, Frequency Calculation, Molecular Dynamics and Rigid MM Water*) for optimizations and MD runs, the nonbonded interactions at the MM level can be neglected for those atom pairs, which are both non-active, without loss of accuracy for the results. Even relative energies between two structures are correct, if the atom positions of the non-active atoms are identical. For all other cases, i.e. if the positions of atoms in the non-active region differ, the full nonbonded interaction should be computed in the final single-point energy calculation. By default this option is switched off.

```
%mm
 Do_NB_For_Fixed_Fixed true # Compute MM-MM nonbonded interaction also for
                            # non-active atom pairs. Default true.
end
```

6.15.3 Rigid Water

As TIP3P water might have to be treated as rigid bodies due to its parametrization - please check out the specifics of the applied force field parametrization - we offer a keyword for the automated rigid treatment of all active MM waters. The following keyword applies bond and angle constraints to active MM waters in optimization as well as MD runs:

```
%mm
 Rigid_MM_Water false       # Default: false.
end
```

6.15.4 Available Keywords for the MM module

Here we list all keywords that are accessible from within the mm block and that are relevant to MM, but also QM/MM calculations. Some of the MM keywords can also be accessed via the qmmm block, see section *Additional Keywords*.

```
!MM # or QMMM, as the MM keywords will also affect the MM part of the QMMM calculation
%mm
# Schemes for the truncation of long-range
#  Coulomb and LJ interaction:
#  The Shift Force scheme for the Coulomb interaction shifts the Coulomb potential
#  such that it becomes zero at the cutoff distance.
 ShiftForceCoulomb true   # Use the shift force scheme for the Coulomb interaction.
                          # Default: true.
 CoulombCutOff 12.        # Distance (in Ang). Default: 12 Ang.
#  With the Switch Force scheme for the LJ interaction is unchanged up to
#  LJCutOffInner. Between LJCutOffInner and LJCutOffOuter a smooth switching function
#  is applied onto the LJ potential so that the force goes down to zero at
#  LJCutOffOuter, without introducing discontinuities.
 SwitchForceLJ true       # Use the switch force scheme for the LJ interaction.
                          # Default: true.
 LJCutOffInner 10.        # Distance (in Ang). Default: 10 Ang.
 LJCutOffOuter 12.        # Distance (in Ang). Default: 12 Ang.

 DielecConst 1.           # dielectric constant used in calc. of electrostatic
                          #  interaction. Default: 1.

 Coulomb14Scaling 1.      # Scaling factor for electrostatic interaction between
                          #  1,4-bonded atoms. Default: 1.

 PrintLevel 1            # Printing options: Can be 0 to 4, 0=nothing, 1=normal, ...
```

(continues on next page)

(continued from previous page)

```
# Keywords that can be accessed from the mm as well as the qmmm block.
#  For a description see qmmm block.

# Information about topology  and force field
 ORCAFFFilename "UBQ.ORCAFF.prms"# If available, e.g. from a previous run, or after
                              # modification, the ORCA Forcefield can be provided.

# Computing MM nonbonded interactions within non-active region.
 Do_NB_For_Fixed_Fixed true # Compute MM-MM nonbonded interaction also for
                            # non-active atom pairs. Default true.

# Optimization and MD of active MM waters
 RIGID_MM_WATER false    # Default: false.

# Extended active region
 ExtendActiveRegion distance    # rule to choose the atoms belonging to activeRegionExt.
                                #  no - do not use activeRegionExt atoms
                                #  cov_bonds - add only atoms bonded covalently to
                                #   active atoms
                                #  distance (default) - use a distance criterion (VDW
                                #   distance plus Dist_AtomsAroundOpt)
 Dist_AtomsAroundOpt 1.         # in Angstrom. Default 1 Ang.
 OptRegion_FixedAtoms { 2 9} end # Default: empty list.

# The following keywords will affect the behavior of MM (without QMMM) calculations,
#  but have to be provided via the qmmm block
 PrintOptRegion true        # Additionally print xyz and trj for opt region
 PrintOptRegionExt true     # Additionally print xyz and trj for extended opt region
 PrintQMRegion true         # Additionally print xyz and trj for QM region
 PrintPDB true              # Additionally print pdb file for entire system, is
                            # updated every iteration for optimization

end

*pdbfile 0 1 ubq.pdb  # structure input via pdb file, but also possible via xyz file
```

6.16 ORCA Multiscale Implementation

With ORCA 5.0 ORCA ‘s multiscale functionality has been extensively expanded. ORCA 5 features five different multiscale methods for

- proteins, DNA, large molecules, explicit solvation:
  - additive **QMMM** (*Additive QMMM*)
  - subtractive **QM1/QM2** methods (2-layered ONIOM) (*Subtractive QM/QM2 Method*)
  - **QM1/QM2/MM** methods (3-layered ONIOM) (*QM/QM2/MM Method*)
- **CRYSTAL-QMMM** for crystals:
  - **MOL-CRYSTAL-QMMM** for molecular crystals (*MOL-CRYSTAL-QMMM*)
  - **IONIC-CRYSTAL-QMMM** for semiconductors and insulators (*IONIC-CRYSTAL-QMMM*)

The multiscale features are optimally connected to all other modules and tools available in ORCA allowing the user to handle multiscale calculations from a QM-centric perspective in a simple and efficient way, with a focus on simplifying the process to prepare, set up and run multiscale calculations.

From the input side all methods share a common set of concepts and keywords, which will be outlined in the first part of this chapter. In the subsequent parts of this chapter, the different methods are described and further input options are discussed.

6.16.1 General Settings and Input Structure

Some of the keywords in this section are common to all five multiscale features, and some are not. If keywords are not available for one of the multiscale features, this will be mentioned.

Overview on Combining Multiscale Features with other ORCA Features

The multiscale features can be used together with all other possible ORCA methods:

Single Point Calculations
Use all kinds of available electronic structure methods as QM method.

Optimization
Use all kinds of geometry optimizations with all kinds of constraints, TS optimization, relaxed surface scans, and the ScanTS feature. Use the L-Opt and L-OptH features including the combination of all kinds of fragment optimizations (fix fragments, relax fragments, relax only specific elements in fragments, treat a fragment as a rigid body).

Transition States and Minimum Energy Paths
Use all kinds of Nudged-Elastic Band calculations (Fast-NEB-TS, NEB, NEB-CI, NEB-TS, including their ZOOM variants) and Intrinsic Reaction Coordinate calculations. (not implemented for MOL-CRYSTAL-QMMM and IONIC-CRYSTAL-QMMM)

Frequency Calculations
Use regular frequency calculations. If required, ORCA automatically switches on the Partial Hessian Vibrational Analysis (PHVA) calculation. (not tested for IONIC-CRYSTAL-QMMM)

Molecular Dynamics
Use the Molecular Dynamics (MD) module for Born-Oppenheimer MD (BOMD) with QM/MM in combination with all kinds of electronic structure methods. (not implemented for MOL-CRYSTAL-QMMM and IONIC-CRYSTAL-QMMM)

Property Calculation
All kinds of regular property calculations are available. For electrostatic embedding the electron density is automatically perturbed by the surrounding point charges.

Excited State Calculations
Use all kinds of excited state calculations (TD-DFT, EOM, single point calculations, optimizations, frequencies). (For the ONIOM calculations the low-level calculations are carried out in the ground state)

Overview on Basic Aspects of the Multiscale Feature

In the following, the basic concepts are introduced.

QM atoms
The user can define the QM region either directly, or via flags in a pdb file. See *QM Atoms*.

QM2 atoms
Only applicable for QM/QM2/MM. For the QM/QM2/MM method the low level QM region (QM2) is defined via the input or via flags in a pdb file. See *QM2 Atoms*. For QM/QM2 the low level region consists of all atoms but the QM atoms.

Active atoms
The user can choose an active region, e.g. for geometry optimizations the atoms that are optimized, for a frequency calculation the atoms that are allowed to vibrate for the PHVA, or for an MD run the atoms that are propagated. See *Active and Non-Active Atoms - Optimization, Frequency Calculation, Molecular Dynamics and Rigid MM Water* and Fig. 6.57.

Forcefield
ORCA has its own forcefield file format (stored in files called basename.ORCAFF.prms). For a convenient setup the orca_mm module offers the option to convert from other forcefield formats. Currently, the following formats can be converted to the ORCA forcefield file format:

CHARMM psf files
protein structure file from the CHARMM forcefield. The psf files can be easily prepared with the popular molecular visualizer VMD, together with its extensions (psfgen, QwikMD, fftk, which works together with ORCA).

AMBER prmtop files
topology files from the AMBER force field. Tutorials on how to generate AMBER prmtop files (for standard and non-standard molecules) can be found here.

Open Force Field
xml files from the openforcefield initiative. With the openff-toolkit xml topology files (compatible with the AMBER force field) can be easily generated for small and medium-sized non-standard molecules. For a tutorial see here.

Simple forcefield for small to medium-sized molecules
Alternatively, the orca_mm module can generate a simple approximate ORCAFF.prms file. For more options, see *ORCA Forcefield File*.

This concept has the following advantages:

Modification of forcefield parameters
Atom and bond specific parameters can be easily modified within the ORCA forcefield file, allowing the user maximum flexibility in modifying the forcefield, which might be particularly useful for chemical systems like transition metal complexes. See *ORCA Forcefield File*.

Standard and Non-Standard Ligands
Ligands can be easily and flexibly exchanged or added to the system, see *ORCA Forcefield File*.

Boundary Treatment
ORCA automatically detects QM-MM boundaries, i.e. bonds that have to be cut between QM and MM region. ORCA automatically generates the link atoms and keeps them at their relative position throughout the run, even allowing to optimize the bond along the boundary. See *QM-MM, QM-QM2 and QM2-MM Boundary*. Not applicable for CRYSTAL-QMMM.

Treatment of overpolarization
ORCA automatically adapts the charges at the QM-MM boundary. See *QM-MM, QM-QM2 and QM2-MM Boundary*. Not applicable for both CRYSTAL-QMMM.

Embedding types
The electrostatic and mechanical embedding schemes are available. See *Embedding Types*.

Detailed information on the different available input and runtime options and additionally available keywords (see *Additional Keywords*) are given below.

QM Atoms

QM atoms can be defined either directly

```
!QMMM
%qmmm
 QMAtoms {0 1 2 27 28} end
end
```

or via the occupancy column of a pdb file.

```
%qmmm                          # use either
 QMAtoms {0:4} end             # 1. list of atoms (counting starts from 0) or
 Use_QM_InfoFromPDB true       # 2. get the definition from the pdb file. Default false.
end                            # If (2) is set to true, (1) is ignored
*pdbfile 0 1 ubq.pdb
```

If `Use_QM_InfoFromPDB` is set to true, a pdb file should be used for the structural input. QM atoms are defined via 1 in the occupancy column, MM atoms via 0. QM2 atoms (for QM/QM2/MM calculations, see *QM2 Atoms*) can be defined via 2 in the occupancy column. The IONIC-CRYSTAL-QMMM method can have even further entries

in the PDB file, see *Different QM and MM regions Stored in the PDB file*. Note that the `Use_QM_InfoFromPDB` keyword needs to be written before the coordinate section.

```
ubq.pdb:
...
ATOM    327  N   ASP A  21      29.599  18.599   9.828  0.00  0.00      P1   N
ATOM    328  HN  ASP A  21      29.168  19.310   9.279  0.00  0.00      P1   H
ATOM    329  CA  ASP A  21      30.796  19.083  10.566  0.00  0.00      P1   C
ATOM    330  HA  ASP A  21      31.577  18.340  10.448  0.00  0.00      P1   H
ATOM    331  CB  ASP A  21      31.155  20.515  10.048  2.00  0.00      P1   C
ATOM    332  HB1 ASP A  21      30.220  21.082   9.865  2.00  0.00      P1   H
ATOM    333  HB2 ASP A  21      31.754  21.064  10.801  2.00  0.00      P1   H
ATOM    334  CG  ASP A  21      31.923  20.436   8.755  1.00  0.00      P1   C
ATOM    335  OD1 ASP A  21      32.493  19.374   8.456  1.00  0.00      P1   O
ATOM    336  OD2 ASP A  21      31.838  21.402   7.968  1.00  0.00      P1   O
ATOM    337  C   ASP A  21      30.491  19.162  12.040  0.00  0.00      P1   C
ATOM    338  O   ASP A  21      29.367  19.523  12.441  0.00  0.00      P1   O
...
```

Note that contrary to the hybrid36 standard of PDB files, ORCA writes non-standard pdb files as:

- atoms 1-99,999 in decimal numbers
- atoms 100,000 and beyond in hexadecimal numbers, with atom 100,000 corresponding to index 186a0.

This ensures a unique mapping of indices. If you want to select an atom with an idex in hexadecimal space (index larger than 100,000), convert the hexadecimal number into decimals when choosing this index in the ORCA input file. Note also, that in the pdb file, counting starts from 1, while in ORCA counting starts from zero.

Active and Non-Active Atoms - Optimization, Frequency Calculation, Molecular Dynamics and Rigid MM Water

The systems of multiscale calculations can become quite large with tens and hundreds of thousands of atoms. In multiscale calculations the region of interest is often only a particular part of the system, and it is common practice to restrict the optimization to a small part of the system, which we call the active part of the system. Usually this active part consists of hundreds of atoms, and is defined as the QM region plus a layer around the QM region. The same definition holds for frequency calculations, in particular since after optimization non-active atoms are not at stationary points, and a frequency calculation would lead to artifacts in such a scenario. MD calculations on systems with hundreds of thousands of atoms are not problematic, but there are applications where a separation in active and non-active parts can be important (e.g. a solute in a solvent droplet, with the outer shell of the solvent frozen).

Note

- If no active atoms are defined, the entire system is treated as active.
- The active region definitions also apply to MM calculations, but have to be provided via the qmmm block.

Input Format

Active atoms can be defined either directly or via the B-factor column of a pdb file.

```
%qmmm                            # use either
 ActiveAtoms {0:5 16 21:30} end # 1. list of atoms (counting starts from 0) or
 Use_Active_InfoFromPDB true    # 2. get the definition from the pdb file.
                                #  Default false.
end                             # If (2) is set to true, (1) is ignored
*pdbfile 0 1 ubq.pdb
```

If `Use_Active_InfoFromPDB` is set to true, a pdb file should be used for the structural input. Active atoms are defined via 1 in the B-factor column, non-active atoms via 0. Note that the `Use_Active_InfoFromPDB` keyword needs to be written before the coordinate section.

```
ubq.pdb:
...
ATOM    327  N   ASP A  21      29.599  18.599   9.828  0.00  0.00      P1   N
ATOM    328  HN  ASP A  21      29.168  19.310   9.279  0.00  0.00      P1   H
ATOM    329  CA  ASP A  21      30.796  19.083  10.566  0.00  1.00      P1   C
ATOM    330  HA  ASP A  21      31.577  18.340  10.448  0.00  1.00      P1   H
ATOM    331  CB  ASP A  21      31.155  20.515  10.048  1.00  1.00      P1   C
ATOM    332  HB1 ASP A  21      30.220  21.082   9.865  1.00  1.00      P1   H
ATOM    333  HB2 ASP A  21      31.754  21.064  10.801  1.00  1.00      P1   H
ATOM    334  CG  ASP A  21      31.923  20.436   8.755  1.00  1.00      P1   C
ATOM    335  OD1 ASP A  21      32.493  19.374   8.456  1.00  1.00      P1   O
ATOM    336  OD2 ASP A  21      31.838  21.402   7.968  1.00  1.00      P1   O
ATOM    337  C   ASP A  21      30.491  19.162  12.040  0.00  0.00      P1   C
ATOM    338  O   ASP A  21      29.367  19.523  12.441  0.00  0.00      P1   O
...
```

Note that in the above example also the QM atoms are defined along with the active atoms.

Optimization in redundant internal coordinates

In ORCA's QM/MM geometry optimization only the positions of the active atoms are optimized. The forces on these active atoms are nevertheless influenced by the interactions with the non-active surrounding atoms. In order to get a smooth optimization convergence for quasi-Newton optimization algorithms in internal coordinates, it is necessary that the Hessian values between the active atoms and the directly surrounding non-active atoms are available. For that reason the active atoms are extended by a shell of surrounding non-active atoms which are also included in the geometry optimization, but whose positions are constrained, see Fig. 6.57. This shell of atoms can be automatically chosen by ORCA. There are three options available:

Distance
: (Default) The parameter `Dist_AtomsAroundOpt` controls which non-active atoms are included in the extension shell, i.e. non-active atoms that have a distance of less than the sum of their VDW radii plus `Dist_AtomsAroundOpt` are included.

Covalent bonds
: All (non-active) atoms that are covalently bonded to active atoms are included.

No
: No non-active atoms are included.

The user can also provide the atoms for the extension shell manually. This will be discussed in section *Frequency Calculation*.

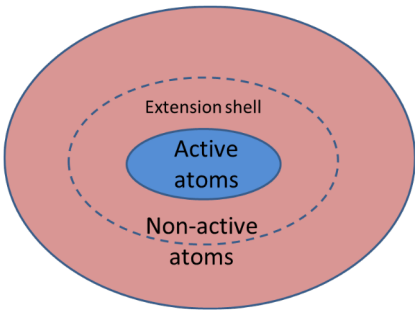


Fig. 6.57: Active and non-active atoms. Additionally shown is the extension shell, which consists of non-active atoms close in distance to the active atoms. The extension shell is used for optimization in internal coordinates and PHVA.

The set of active atoms is called the 'activeRegion', and the set of active atoms plus the surrounding non-active atoms is called 'activeRegionExt'. During geometry optimization the following trajectories are stored (which can be switched off):

basename_trj.xyz
Entire QM/MM system

basename.QMonly_trj.xyz
Only QM region

basename.activeRegion_trj.xyz
Only active atoms

basename.activeRegionExt_trj.xyz
Active atoms plus extension

The following files are stored containing the optimized structures - if requested:

basename.pdb
Optimized QM/MM system in pdb file format

basename.xyz
Optimized QM/MM system

basename.QMonly.xyz
Only QM region

basename.activeRegion.xyz
Only active atoms

basename.activeRegionExt.xyz
Active atoms plus extension

Optimization with the Cartesian L-BFGS Minimizer

For very large active regions the quasi-Newton optimization in internal coordinates can become costly and it can be advantageous to use the L-Opt or L-OptH feature, see section *Geometry Optimizations using the L-BFGS optimizer*. For the L-Opt(H) feature there exist two ways to define the active region:

- via the ActiveAtoms keyword (or the Use_Active_InfoFromPDB flag) or
- via fragment definition and the different keywords for fragment optimization. Available options are:

  FixFrags
  Freeze the coordinates of all atoms of the specified fragments.

  RelaxHFrags
  Relax the hydrogen atoms of the specified fragments. Default for all atoms if !L-OptH is defined.

RelaxFrags
Relax all atoms of the specified fragments. Default for all atoms if !L-Opt is defined.

RigidFrags
Treat each specified fragment as a rigid body, but relax the position and orientation of these rigid bodies.

Note

- The L-Opt(H) option together with the fragment optimization can be used in order to quickly preoptimize your system at MM level. E.g. you can optimize the hydrogen positions of the protein and water molecules, and at the same time relax non-standard molecules, for which no exact bonded parameters are available, as rigid bodies.

```
!MM L-OptH
%mm
 ORCAFFFilename "DNA_hexamer.ORCAFF.prms"
end
*pdbfile 0 1 protein_ligand.pdb
%geom
 Frags                # all other atoms belong to fragment 1 by default
  2 {22307:22396} end # cofactor
  3 {22397:22423} end # ligand
 end
 RigidFrags {2 3} end # treat cofactor and ligand as individual rigid bodies
end
```

Frequency Calculation

If all atoms are active, a regular frequency calculation is carried out when requesting !NumFreq. If there are also non-active atoms in the QM/MM system, the Partial Hessian Vibrational Analysis (PHVA, see section *Partial Hessian Vibrational Analysis*) is automatically selected. Here, the PHVA is carried out for the above defined activeRegionExt, where the extension shell atoms are automatically used as 'frozen' atoms. Note that the analytic Hessian is not available due to the missing analytic second derivatives for the MM calculation. Note that in a new calculation after an optimization it might happen that the new automatically generated extended active region is different compared to the previous region before optimization. This means that when using a previously computed Hessian (e.g. at the end of an optimization or a NEB-TS run) the Hessian does not fit to the new extended active region. ORCA tries to solve this problem by fetching the information on the extended region from the hess file. If that does not work (e.g. if you distort the geometry after the Hessian calculation) you should manually provide the list of atoms of the extended active region. This is done via the following keyword:

```
%qmmm
 OptRegion_FixedAtoms {27 1288:1290 4400} end  # manually define the
                                               #  activeRegionExt atoms.
end
```

Note that ORCA did print the necessary information in the output of the calculation in which the Hessian was computed:

```
Fixed atoms used in optimizer          ... 27 1288 1289 1290 4400
```

Nudged Elastic Band Calculations

NEB calculations (section *Nudged Elastic Band Method*) can be carried out in combination with the multiscale features in order to e.g. study enzyme catalysis. When automatically building the extension shell at the start of a Multiscale-NEB calculation, not only the coordinates of the main input structure ('reactant'), but also the atomic coordinates of the 'product' and, if available, of the 'transition state guess' are used to determine the union of the extension shell of the active region. For large systems it is advised to use the active region feature for the NEB calculation. Note that the atomic positions of the non-active atoms of reactant and product and, if available, transition state guess, should be identical.

Molecular Dynamics

If there are active and non-active atoms in the multiscale system, only the active atoms are allowed to propagate in the MD run. If all atoms are active, all atoms are propagated. For more information on the output and trajectory options, see section *Regions*.

Rigid MM Water

As TIP3P water might have to be treated as rigid bodies due to its parametrization - please check out the specifics of the applied force field parametrization - we offer a keyword for the automated rigid treatment of all active MM water molecules. The following keyword applies bond and angle constraints to active MM water molecules in optimizations as well as MD runs:

```
Rigid_MM_Water false       # Default: false.
```

Forcefield Input

For the MM part of the QM/MM calculation forcefield parameters are necessary. Internally, ORCA uses the ORCA forcefield. For a description on the format, how to obtain and manipulate the forcefield parameters, see section *ORCA Forcefield File*.

Note

- ORCAFF.prms files only need to be provided for QM/MM, QM/QM2/MM and IONIC-CRYSTAL-QMMM calculations.
- For QM/QM2 and MOL-CRYSTAL-QMMM calculations there is no need to provide a ORCAFF.prms file.
- The ORCAFF.prms file for the IONIC-CRYSTAL-QMMM calculation can be conveniently set up with the orca_crystalprep tool, see section *orca_crystalprep*.
- For IONIC-CRYSTAL-QMMM and MOL-CRYSTAL-QMMM calculations the self-consistently optimized MM point charges of the entire supercell are stored in an ORCAFF.prms file, see section *Charge Convergence between QM and MM region*. This ORCAFF.prms file can then be used in subsequent calculations with larger QM regions, different methods and basis sets, excited state calculations, etc.

The force field filename is provided via the keyword `ORCAFFFilename`:

```
%qmmm
 ORCAFFFilename "UBQ.ORCAFF.prms"
end
```

QM-MM, QM-QM2 and QM2-MM Boundary

This section is important for the QM/MM, QM/QM2 and QM/QM2/MM methods. For the latter method two boundary regions are present (between QM and QM2 as well as between QM2 and MM region), and both can go through covalent bonds. In the following we will only discuss the concept for the boundary between QM and MM, but the same holds true for the other two boundaries.

Link Atoms

ORCA automatically generates link atoms based on the information of the QM region and on the topology of the system (based on the ORCAFF.prms file). ORCA places link atoms on the bond between QM and MM atoms.

Important

- When defining the QM, QM2 and MM regions, make sure that you only cut through single bonds (not aromatic, double, triple bonds, etc.).

Bond Length Scaling factor

The distance between QM atom and link atom is determined via a scaling factor (in relation to the QM-MM bond length) that is computed as the ratio of the equilibrium bond length between QM and hydrogen atom (d0_QM-H) and the equilibrium bond length between QM and MM atom (d0_QM-MM).

Standard QM-H Bond Length

For the equilibrium bond lengths to hydrogen, ORCA uses tabulated standard values for the most common atoms involved in boundary regions (C, N, O), which can be modified via keywords as defined further below. ORCA stores these values on-the-fly in a simple file (basename.H_DIST.prms). If necessary, the user can modify these values atom-specific or add others to the file and provide this file as input to ORCA (see also paragraph *Get standard hydrogen bond lengths: getHDist*). For QM/QM2 and QM/QM2/MM methods the equilibrium bond lengths to hydrogen are explicitly calculated.

```
%qmmm
# standard equilibrium bond lengths with hydrogen can be modified
 Dist_C_HLA 1.09        # d0_C-H
 Dist_O_HLA 0.98        # d0_O-H
 Dist_N_HLA 0.99        # d0_N-H
# file can be provided which provides the used d0_X-H values specific to all atoms
 H_Dist_FileName "QM_H_dist.txt"
end
```

Bonded Interactions at the QM-MM Boundary

The MM bonded interactions within the QM region are neglected in the additive coupling scheme. However, at the boundary, it is advisable to use some specific bonded interactions which include QM atoms. By default ORCA neglects only those bonded interactions at the boundary, where only one MM atom is involved, i.e. all bonds of the type QM1-MM1, bends of the type QM2-QM1-MM1, and torsions of the type QM3-QM2-QM1-MM1. Other QM-MM mixed bonded interactions (with more than two MM atoms involved) are included. The user is allowed to include the described interactions, which is controlled via the following keywords:

```
%qmmm
 DeleteLADoubleCounting true     # Neglect bends (QM2-QM1-MM1) and torsions
                                 # (QM3-QM2-QM1-MM1). Default true.
```

(continues on next page)

(continued from previous page)

```
 DeleteLABondDoubleCounting true # Neglect bonds (QM1-MM1)
end
```

Charge Alteration

If QM and MM atoms are connected via a bond (defined in the topology file), the charge of the close-by MM atom (and its neighboring atoms) has to be modified in order to prevent overpolarization of the electron density of LA and QM region. This charge alteration is automatically taken care of by ORCA. ORCA provides the most popular schemes that can be used to prevent overpolarization:

CS
: Charge Shift - Shift the charge of the MM atom away to the MM2 atoms so that the overall charge is conserved

RCD
: Redistributed Charge and Dipole - Shift the charge of the MM atom so that the overall charge and dipole is conserved

Z0
: Keep charges as they are. This MM atom will probably lead to overpolarization

Z1
: Delete the charge on the MM1 atom (no overall charge conservation)

Z2
: Delete the charges on the MM1 atom and on its first (MM2) neighbors (no overall charge conservation)

Z3
: Delete the charges on the MM1 atom and on its first (MM2) and second (MM3) neighbors (no overall charge conservation)

Embedding Types

The following embeding schemes are available:

Electrostatic
: The electrostatic interaction between QM and MM system is computed at the QM level. Thus, the charge distribution of the MM atoms can polarize the electron density of the QM region. The LJ interaction between QM and MM system is computed at the MM level.

```
%qmmm
 Embedding Electrostatic  # Electrostatic (Default)
                          # Mechanical
end
```

In the scheme of electrostatic embedding, the evaluation of the electrostatic potential generated by the MM part can be accelerated by using the FMM algorithm (described in reference [317]). This will speed up the building of the Fock Matrix. The default recommended setup can be called using the FMM-QMMM keyword directly in the keywords line. However, more details about the algorithm parameters and all input options can be found in (see also paragraph *Fast Multipole Method*)

```
! FMM-QMMM
```

It is recommended to use that option whenever the MM part is composed of more than 10,000 atoms (or point charges for ECM methods).

6.16.2 Additive QMMM

The minimum input necessary for an additive QM/MM calculation looks as follows.

```
!QMMM
%qmmm
 QMAtoms {0 1 2 27 28} end
 ORCAFFFilename "UBQ.ORCAFF.prms"
end
```

6.16.3 ONIOM Methods

For the simulation of large systems with up to 10000 atoms, or for large QM regions in biomolecules, ORCA provides the QM/QM2 and QM/QM2/MM methods. The specifics of the two different methods are discussed further below. Here we are presenting the common concepts and keywords of both methods. For subtractive methods, we use a high level (QM) and a low level (QM2) of accuracy for different parts of the system. The advantages of this - in contrary to QM-MM methods - are as follows:

- QM2 methods are polarizable, the interaction with the high level region is more accurate.
- No MM parameters are needed for the atoms that are described at the QM2 level.

Available QM2 Methods

ORCA has several built in QM2 methods:

- Semiempirical methods (AM1, PM3)
- Tight-binding DFT (XTB0, XTB1, XTB (or XTB2))
- Composite methods (HF-3c, PBEh-3c, r2SCAN-3c)
- User-defined QM2 methods

The individual keywords for these methods are:

```
!QM/XTB          or   !QM/XTB/MM
!QM/XTB1         or   !QM/XTB1/MM
!QM/XTB0         or   !QM/XTB0/MM
!QM/HF-3C        or   !QM/HF-3C/MM
!QM/PBEH-3C      or   !QM/PBEH-3C/MM
!QM/R2SCAN-3C    or   !QM/R2SCAN-3C/MM
!QM/PM3          or   !QM/PM3/MM
!QM/AM1          or   !QM/AM1/MM
```

Users can define their own low-level methods in the following way

```
!QM/QM2          or   !QM/QM2/MM
%qmmm
  QM2CUSTOMMETHOD "B3LYP"
  QM2CUSTOMBASIS "def2-SVP def2/J"
end
```

Alternatively, a custom QM2 method / basis set file can be provided:

```
!QM/QM2  or  !QM/QM2/MM
%qmmm
  QM2CustomFile "myQM2Method.txt" # File should be available in working directory.
end
```

The custom QM2 method file can contain any desired input, as e.g. the file myQM2Method.txt:

```
!cc-pVDZ HF TightSCF NOSOSCF KDIIS
%basis
  NewAuxJKGTO Mg "AutoAux" end
end
```

Note

- Only add methods (including convergence settings) and basis sets for the QM2Custom options. Everything else (parallelization, memory, solvation, etc.) is taken care of by ORCA itself.

Electrostatic Interaction between high and low level

By default we are using electrostatic embedding, i.e. the high level system sees the atomic point charges of the low level (QM2) system. These point charges are derived from the full system low level (QM2) calculation. The following methods for determining these charges are available:

```
Charge_Method Hirshfeld # Hirshfeld (default)
                        # MBIS
                        # CHELPG
                        # Mulliken
                        # Loewdin, default for QM2 = AM1 or PM3
```

The QM2 point charges can be scaled with the following keyword.

```
%qmmm
  Scale_QM2Charges_MMAtom 1. # default is 1.
end
```

Boundary Region

The boundary between high and low level part of the system can contain covalent bonds. For the detection and realistic treatment of these covalent bonds, a topology of the large QM2 system is generated using the following keyword.

```
AutoFF_QM2_Method XTB   # XTB (default)
                        # XTB1
                        # XTB0
                        # GFNFF
                        # HF3C
                        # PBEH3C
                        # R2SCAN3C
                        # PM3
                        # AM1
```

Note

- By default ORCA uses the XTB method for the preparation of the QM2 topology. In order to use the default you need to make sure to have the otool_xtb binary in your ORCA PATH, see *Semi-empirical tight-binding methods: Grimme's GFN0-xTB, GFN-xTB and GFN2-xTB*.

Subtractive QM/QM2 Method

The QM/QM2 method is a very convenient way of running multiscale calculations without the need to prepare any parameters. This method is a subtractive QM-QM method, in which we treat a part of interest on a higher level of accuracy, and the remainder of the system on lower level of accuracy. The implementation follows similar works as e.g. described in reference [571].

The method can be used in a similar way as a regular QM calculation. Let us have a look at the proton transfer in propionic acid, which can be modeled as follows:

```
!QM/XTB BP86 def2-TZVP def2/J
!Fast-NEB-TS NumFreq
!pal8
%qmmm
  QMAtoms {0:3} end
end
%neb
 product "propionicAcid_prod.xyz"
 preopt true
end
*xyz 0 1
H       -0.738352472      0.000000000     -5.836214279
O       -0.738352472     -0.587240971     -5.061536853
O       -0.738352472      1.434717404     -4.069730302
C       -0.738352472      0.227304724     -3.975502162
C       -0.738352472     -0.566448428     -2.687358498
H        0.133951528     -1.231202352     -2.710760176
H       -1.610656472     -1.231202352     -2.710760176
C       -0.738352472      0.318369069     -1.443687014
H       -0.738352472     -0.294739868     -0.538164669
H        0.142397528      0.965221387     -1.423275731
H       -1.619102472      0.965221387     -1.423275731
*
```

with the product structure file (propionicAcid_prod.xyz):

```
11
C3H6O2
H       -0.738352472      1.628728096     -5.020130139
O       -0.738352472     -0.587240971     -5.061536853
O       -0.738352472      1.434717404     -4.069730302
C       -0.738352472      0.227304724     -3.975502162
C       -0.738352472     -0.566448428     -2.687358498
H        0.133951528     -1.231202352     -2.710760176
H       -1.610656472     -1.231202352     -2.710760176
C       -0.738352472      0.318369069     -1.443687014
H       -0.738352472     -0.294739868     -0.538164669
H        0.142397528      0.965221387     -1.423275731
H       -1.619102472      0.965221387     -1.423275731
```

As can be seen from the input, the only difference to a regular calculation is the necessity to define the high level region via the QMAtoms keyword.

System charges and multiplicities

The two subsystems can have different (integer) charges and multiplicities. Defining the correct charges and multiplicities is important. The charge and multiplicity defined via the coordinate section defines the charge and multiplicity of the high level region (QMAtoms). The user still needs to define the charge and multiplicity of the total system (corresponding to the sum of the charge of the high level and low level parts, and corresponding to the overall multiplicity).

```
%qmmm
  QMAtoms {0:3} end # high level region
  Charge_Total  0   # charge of the full system. Default 0.
  Mult_Total    1   # multiplicity of the full system. Default 1.
end
*xyz 0 1            # charge and mult. of the high level region, i.e. atoms 0 to 3
```

Available low level methods

The following QM2 (low level) methods are available:

```
!QM/XTB
!QM/XTB1
!QM/XTB0
!QM/HF-3C
!QM/PBEH-3C
!QM/R2SCAN-3C
!QM/PM3
!QM/AM1
!QM/QM2
```

For information on how to specify the custom QM/QM2 method please see *Available QM2 Methods*.

Solvation

Implicit Solvation effects can be included in QM/QM2 calculations. On the one hand, for QM/XTB calculations, one can adopt the analytical linearized Poisson-Boltzmann (ALPB) solvation model, the domain decomposition COSMO (ddCOSMO), or the extended conductor-like polarizable continuum model (CPCM-X), and on the other hand, if no XTB is requested, ORCA uses the C-PCM. The user just needs to add the following tags in the ORCA input file,

XTB for the QM2 region:

```
!QM/XTB ALPB(Water)

or

!QM/XTB DDCOSMO(Water)

or

!QM/XTB CPCMX(Water)
```

No XTB for the QM2 region:

```
!QM/HF-3c CPCM(Water)
```

If the ddCOSMO (XTB) or the C-PCM (non-XTB) are requested, there are two possible ONIOM/implicit-solvation methods:[874]

- **C-PCM/B:** The effect of the solvent is, in the first place, included in the calculation for the large QM2 system. Once this calculation finishes, the solvation charges located on the surface of the cavity for the large system are used as point charges for the subsequent low-level and high-level calculations for the small system.
- **C-PCM/C:** The effect of the solvent is only included in the calculation for the large QM2 system.

The user can choose one scheme or the other via the tag "`solv_scheme`" in the "`qmmm`" block:

```
%qmmm
solv_scheme CPCM_B   # CPCM_B (default)
                     # CPCM_C
end
```

If the ALPB model or the CPCM-X are requested (within QM/XTB methods), the solvation effect is just included in the calculation for the large QM2 system (as one does for the C-PCM/C scheme).

QM/QM2/MM Method

The QM/QM2/MM method uses a combination of the subtractive scheme for the QM-QM2 part, and the additive scheme for the (QM-QM2) - (MM) interaction. It can be used if very large QM regions are required for biomolecules and explicitly solvated systems. The system is divided into a high level (QM), low level (QM2), and MM region (MM).

QM2 Atoms

QM2 atoms need to be defined for QM/QM2/MM calculations. They can be defined either directly

```
%qmmm
 QMAtoms {0:4} end           # list of QM atoms (counting starts from 0) or
 QM2Atoms {5:22} end         # list of QM2 atoms
end                          # an atom should not occur in both lists
*pdbfile 0 1 ubq.pdb
```

or via the occupancy column of a pdb file (see *QM Atoms*).

System charges and multiplicities

The high and low level subsystems can have different (integer) charges and multiplicities. Defining the correct charges and multiplicities is important. The charge and multiplicity defined via the coordinate section defines the charge and multiplicity of the high level region (QMAtoms). The user still needs to define the charge and multiplicity of the medium system (corresponding to the sum of the charge of the high level and low level regions, and corresponding to the overall multiplicity of the combined high and low level region). The charge of the MM region is determined based on the MM parameters provided by the forcefield.

```
%qmmm
  QMAtoms {0:3} end # high level region
  Charge_Medium 0   # charge of the medium system. Default 0.
  Mult_Medium   1   # multiplicity of the medium system.  Default 0.
end
*xyz 0 1            # charge and mult. of the high level region, i.e. atoms 0 to 3
```

Available low level methods

The following QM2 (low level) methods are available:

```
!QM/XTB/MM
!QM/XTB1/MM
!QM/XTB0/MM
!QM/HF-3C/MM
!QM/PBEH-3C/MM
!QM/R2SCAN-3C/MM
!QM/PM3/MM
!QM/AM1/MM
!QM/QM2/MM
```

For information on how to specify the custom QM/QM2/MM method please see *Available QM2 Methods*.

Example Input

An example for a QM/QM2/MM calculation looks as follows:

```
!QM/HF-3c/MM Opt B3LYP def2-TZVP def2/J NumFreq
%qmmm
 ORCAFFFilename "peptideChain.ORCAFF.prms"
 QMAtoms {16:33 68:82} end
 QM2Atoms {0:12 83:104} end
 ActiveAtoms { 0:38 65:120} end
 Charge_Medium 0
end
*pdbfile -1 1 peptideChain.pdb
```

6.16.4 CRYSTAL-QMMM

For the simulation of extended systems ORCA provides the CRYSTAL-QMMM features:

MOL-CRYSTAL-QMMM
: for molecular crystals.

IONIC-CRYSTAL-QMMM
: for semiconductors and insulators.

In this section we present the concepts and keywords that are common to both methods. ORCA is a molecular code and cannot carry out periodic calculations, but ORCA has been used for modeling selected properties for ionic solids using embedded cluster models already in the past (see e.g reference [213]). With ORCA 5 we provide the CRYSTAL-QMMM method, which simplifies setting up and running these types of embedded cluster calculations. The user needs to provide the structure with a (large enough) supercell representative for the crystal, and provide a list of QM atoms. The QM region should be located in the central part of the supercell. The QM atoms are embedded in the remainder of the supercell, the surrounding MM atoms, which are represented by atom-centered point charges and influence the QM calculations. How these charges are obtained, is outlined in the next paragraph.

Charge Convergence between QM and MM region

The core idea of the CRYSTAL-QMMM method is to reach charge convergence between the QM and the MM atoms. For this purpose, the MM charges are updated self-consistently with the atomic (CHELPG) charges of the QM atoms. This idea follows the work of reference [213] for the IONIC-CRYSTAL-QMMM and of reference [109] for the MOL-CRYSTAL-QMMM.

```
!IONIC-CRYSTAL-QMMM or MOL-CRYSTAL-QMMM
%qmmm
  Conv_Charges true            # default true for both CRYSTAL-QMMM variants
  Conv_Charges_MaxNCycles 10   # default 30 for MOL-CRYSTAL-QMMM
                               # default 10 for IONIC-CRYSTAL-QMMM
  Conv_Charges_ConvThresh 0.01 # threshold (default 0.01) for maximum charge change
                              #for atom type between two subsequent charge convergence
                               # cycles
  Scale_FormalCharge_MMAtom  1. # default is 1. MM atomic charges used in QM part of
                               # the CRYSTAL-QMMM calculation are scaled by this factor
end
```

During optimizations, the charge convergence scheme can be used (i) only at the very beginning, (ii) every N geometry steps, or (iii) for each geometry step, using the following keyword:

```
%geom
  ReConvCharge 1 # default is 1. Redo charge convergence scheme every N steps.
end
```

After charge convergence, the converged charges are stored in the file basename.convCharges.ORCAFF.prms. It can be used for a later calculation (with the same or different electronic structure settings) with the following keyword combination:

```
!IONIC-CRYSTAL-QMMM    or    !MOL-CRYSTAL-QMMM
%qmmm
  Conv_Charges false
  ORCAFFFilename "firstjob.convCharges.ORCAFF.prms"
end
```

MM-MM Interaction

Unlike with the other multiscale methods (QMMM, QM/QM2, QM/QM2/MM) the MM region is only treated as a perturbation to the QM region. The (bonded and nonbonded) MM-MM interaction is not computed - since it would not affect the QM-MM interaction - and thus the active region (optimizations, frequencies, ...) in CRYSTAL-QMMM calculations should be restricted to atoms of the QM region.

Preparing CRYSTAL-QMMM Calculations

The input structures and input files for the CRYSTAL-QMMM calculations can be prepared with the orca_crystalprep module (see section *orca_crystalprep*).

MOL-CRYSTAL-QMMM

Conceptually we use an additive QM/MM calculation, in which the QM region interacts with the MM region only via nonbonded interactions. Lennard-Jones interaction between QM and MM region is computed using van der Waals parameters from the UFF force field. The charge convergence treatment between QM and MM region starts with zero charges on the MM atoms, and is iterated until convergence of the QM atomic (CHELPG) and MM point charges is achieved.

The MOL-CRYSTAL-QMMM method expects as structural input a supercell that consists of repeating, identical molecular subunits, i.e. the atom ordering of the molecules should always be the same, and one molecular subunit should follow the next one. The minimum input necessary for a MOL-CRYSTAL-QMMM run looks as follows.

```
%qmmm
  NUnitCellAtoms 16      # provide the number of atoms per molecular subunit
  QMAtoms {160:175} end  # QM atoms, should be in central part of the supercell.
                         # Must be at least one complete molecular subunit.
end
```

Note

- For molecular crystals it is assumed that the entire supercell consists of repeating units which each have zero charge. QM regions should consist of one or multiple of such charge-neutral units.

Extending the QM Region

ORCA can extend the QM region (which we call QM core region, defined by the original QMAtoms input) by one or multiple layers of molecular subunits using the `HFLayers` keyword. If the `HFLayers` keyword is used, the molecular subunits of the defined number of layers around the QM core region is added to the QM region. The layers of molecular subunits around the QM core region are detected via distance criteria.

```
%qmmm
  HFLayers 0                        # default 0
  HFLayerGTO "LANL2DZ"              # default. Use this basis set for the HFLayer atoms
  HFLayerECP "HayWadt"              # default. Use these ECPs for the HFLayer atoms.
  Conv_Charge_UseQMCoreOnly true    # only use the QM core region for the charge
                                    # convergence scheme
end
```

The HFLayer can be seen as a buffer region between the molecular subunit of interest (original QM atoms) and the surrounding subunits. The different possible reasons for HFLayers are as follows:

- For DLPNO calculations with HFLayers, the DLPNO multilevel feature is automatically switched on. The molecules of the HFlayer form an own fragment which is treated on HF level only, i.e. these molecules are not included in the Post-HF treatment.
- The HFLayers can also be used for non-DLPNO calculations. In such cases, the HFLayer molecules are in the same way included in the QM calculation as the other QM molecules. But their definition is a convenient way to choose a different basis set (and ECPs) for those molecules.
- It can be assumed that the QM charges computed for the QM core region are more realistic than the charges of the HFLayer atoms near the MM atomic charges. Thus, using only the QM atomic charges of the QM core region represent more realistic charges for the charge convergence scheme.

Note

- The term HFLayers might be misleading. Strictly speaking, only for MOL-CRYSTAL-QMMM calculations with DLPNO methods (e.g. DLPNO-CCSD(T), DLPNO-MP2, DLPNO-B2PLYP) the HF layer atoms are treated on HF level. For other methods (e.g. DFT) the HF layer atoms are treated with the same electronic structure method as the QM core region atoms.
- If necessary, the atoms of the HFLayer can be defined by the user. Make sure that complete molecular subunits are used here.

```
%qmmm
  HFLayerAtoms {0:15} end
end
```

Example Input

An example for a MOL-CRYSTAL-QMMM calculation looks as follows:

```
! MOL-CRYSTAL-QMMM
! PBE def2-SVP Opt NumFreq
%qmmm
 NUnitCellAtoms 13
 QMAtoms {221:233} end
 ActiveAtoms {221:233} end
end
*xyzfile 0 1 Ala_SC.xyz
```

IONIC-CRYSTAL-QMMM

Conceptually we use an embedded cluster calculation, in which the QM region interacts only with the MM atomic point charges of the surrounding. Unlike with MOL-CRYSTAL-QMMM, the Lennard-Jones interaction between QM and MM region is not computed. The charge convergence treatment between QM and MM region starts with the initial MM charges (the charges initially read from the ORCAFF.prms file) and is iterated until convergence of the QM atomic (CHELPG) and MM point charges is achieved.

Boundary Region

For ionic crystals a boundary region needs to be introduced between the QM region and the atomic point charges of the MM atoms in order to avoid spurious electron leakage from the clusters and to treat dangling bonds on the surface of the QM region. This boundary region consists of capped effective core potentials (cECPs). These cECPs are placed on the position of the MM atoms that are directly adjacent to the QM region. ORCA analyzes the connectivity of the atoms of the supercell and can thus detect the layers of atoms around the QM region, where the first layer consists of the atoms directly bonded to the QM atoms, the second layer consists of the atoms directly bonded to the atoms of the first layer and so on. The charges of these cECPs are determined in the same way as the MM atomic charges during the charge convergence scheme.

```
%qmmm
  ECPLayerECP "SDD"              # cECPs used for the boundary region
  ECPLayers 3                    # Number of cECP layers around the QM region.
                                 # Default is 3.
  Scale_FormalCharge_ECPAtom 1.  # default is 1. Charges on cECPs are scaled by
                                 # this factor
end
```

Extending the QM Region

ORCA can extend the QM region (which we call QM core region, defined by the original QMAtoms input) by one or multiple layers of atoms using the `HFLayers` keyword. If the `HFLayers` keyword is used, the atoms of the defined number of layers around the QM core region is added to the QM region. The layers of atoms around the QM core region are detected in the same way as defined above for the ECPLayers.

```
%qmmm
  HFLayers 0                        # default 0
  HFLayerGTO "LANL2DZ"              # default. Use this basis set for the HFLayer atoms
  HFLayerECP "HayWadt"              # default. Use these ECPs for the HFLayer atoms.
  Conv_Charge_UseQMCoreOnly true # only use the QM core region for the charge
                                    # convergence scheme
end
```

The HFLayer can be seen as a buffer region between the actual atoms of interest (original QM atoms) and the surrounding cECP and MM point charge environment.The different possible reasons for HFLayers are as follows:

- For DLPNO calculations with HFLayers, the DLPNO multilevel feature is automatically switched on. The atoms of the HFLayer form an own fragment which is treated at HF level only, i.e. these atoms are not included in the Post-HF treatment.
- It can be assumed that the QM charges computed for the QM core region are more realistic than the charges of the HFLayer atoms near the cECPs and MM atomic charges, in particular for highly charged systems. Thus, using only the QM atomic charges of the QM core region represent more realistic charges for the charge convergence scheme.

Note

- The term HFLayers might be misleading. Strictly speaking, only for IONIC-CRYSTAL-QMMM calculations with DLPNO methods (e.g. DLPNO-CCSD(T), DLPNO-MP2, DLPNO-B2PLYP) the HF layer atoms are treated on HF level. For other methods (e.g. DFT) the HF layer atoms are treated with the same electronic structure method as the QM core region atoms.
- If necessary, the atoms of the HFLayer can be defined by the user:

```
%qmmm
  HFLayerAtoms {29:35} end
end
```

- The charge of the HFLayer is automatically computed based on the formal charges of its atoms. If necessary, the HFLayer charge can be provided by the user.

```
%qmmm
  Charge_HFLayer 10 # by default it is computed automatically based on the formal
                    # charges from the ORCAFF.prms file
end
```

Net Charge of the Entire Supercell

For ionic crystals, the QM region (as well as the entire supercell) can consist of an unequal number of atoms of each atom type. As a result of that, the QM region may have non-zero net charge. Consequently, when assigning the atomic point charges to the MM atoms during the charge convergence scheme, the sum of the charge of the QM region, of the ECP layer, and of the atomic charges of the MM atoms, can deviate from the ideal charge of the supercell. In order to enforce the ideal charge of the supercell, the total charge of the entire system is enforced by equally shifting the charges of all MM (including boundary) atoms.

```
%qmmm
  Charge_Total  0         # default 0. Total charge of the supercell
```

(continues on next page)

(continued from previous page)

```
  EnforceTotalCharge true # enforce the total charge by shifting MM charges
end
```

The charge shifting procedure can be restricted to the MM atoms on the outer parts of the supercell by defining the number of OuterPCLayers. If `OuterPCLayers` is set to 1, only the atoms of the most outer layer of the supercell (recognized based on the connectivity of the atoms) are included in the charge shifting procedure. If `OuterPCLayers` is larger than 2, the atoms connected to the most outer layer are additionally included in the charge shifting procedure, etc.

```
%qmmm
  OuterPCLayers 0         # default 0, i.e. charge shifting for all MM atoms
end
```

Note

- The charge of the QM core region can be chosen to be assigned automatically. This overwrites the charge provided in the xyz or pdb section.

```
%qmmm
  AutoDetectQCCharge false # default is false
end
```

Example Input

A minimal example for an IONIC-CRYSTAL-QMMM calculation looks as follows:

```
! IONIC-CRYSTAL-QMMM
! NMR
! PBE pcSseg-2 def2/JK def2-svp/C
%qmmm
  QMAtoms {0:6} end
  ORCAFFFilename "nacl6.ORCAFF.prms"
  CHARGE_TOTAL    0
end
*xyzfile -5 1 nacl6.xyz
```

Different QM and MM regions Stored in the PDB file

The stored PDB file contains the following entries in its occupancy column.

0
: MM atoms mimicked by surrounding point charges.

1
: QM core region atoms

2
: HFLayer atoms

3
: cECPs

4
: OuterPCLayer atoms (subset of MM atoms) if defined, are the only atoms that are used in the charge shift procedure for enforcing the total charge.

6.16.5 Additional Keywords

Here we list additional keywords and options that are accessible from within the %qmmm block and that are relevant to QM/MM calculations. Some of these keywords can also be accessed via the %mm block, see section *Available Keywords for the MM module*.

```
!QMMM
%qmmm
# Charge alteration scheme preventing overpolarization.
 ChargeAlteration CS    # CS (Default)
                        # RCD
                        # Z0
                        # Z1
                        # Z2
                        # Z3
# Parameters for placing the shifted and redistributed charges for RCD and CS schemes.
 Scale_RCD 0.5          # Defines where on the bond between MM1 and MM2 the
                        # shifted charge is positioned. Default: midpoint.
 Scale_CS 0.06          # Defines where on the bond between MM2 and MM1/MM3 the
                        # shifted charge is positioned. Default: 0.06 x bond.

# The extended active region, activeRegionExt, contains the atoms surrounding the
# active atoms.
 ExtendActiveRegion distance  # rule to choose the atoms belonging to activeRegionExt.
                              #  no - do not use activeRegionExt atoms
                              #  cov_bonds - add only atoms bonded covalently to
                              #   active atoms
                              #  distance (default) - use a distance criterion (VDW
                              #   distance plus Dist_AtomsAroundOpt)
 Dist_AtomsAroundOpt 0.0      # in Angstrom (Default 0). Meaning only freeze atoms
                              # which touch the active atoms by their VDW spheres.
                              # only needed for ExtendActiveRegion distance
 OptRegion_FixedAtoms {27 1288:1290 4400} end  # manually define the activeRegionExt
                                               #   atoms. Default: empty list.

# Printing options. All are true by default.
 PrintLevel 1               # Can be 0 to 4, 0=nothing, 1=normal, ...
 PrintOptRegion true        # Additionally print xyz and trj for opt region
 PrintOptRegionExt false    # Additionally print xyz and trj for extended opt region
 PrintQMRegion true         # Additionally print xyz and trj for QM region
 PrintFullTRJ true          # Print trajectory of full system. Default true.
 PrintPDB true              # Additionally print pdb file for entire system, is
                            # updated every iteration for optimization

# Computing MM nonbonded interactions within non-active region.
 Do_NB_For_Fixed_Fixed true # Compute MM-MM nonbonded interaction also for
                             # non-active atom pairs. Default true.
# Treats all active water molecules that are treated on MM level as rigid bodies
#  in optimization and MD simulation, see section "Rigid Water".
 Rigid_MM_Water false       #  Default false.

end

*pdbfile 0 1 ubq.pdb  # structure input via pdb file, but also possible via xyz file
```

6.17 QM/MM via Interfaces to ORCA

ORCA is easy to interface as a QM engine in pretty much any QM/MM environment, as it will accept a set of point charges from an external file (see section *Inclusion of Point Charges*) and it will return, in a transparent format, all the required information for computing energies and gradients to the driving program. In our research group we have experience with four different QM/MM environments: ChemShell, Gromacs, pDynamo and NAMD. In the following each of the interfaces are described. Is beyond the scope of the manual to be exhaustive, and only minimal working examples are going to be presented. These will cover mainly the technical aspects with respect to the QM part of the QM/MM calculation. For the initial preparation of the system the user is referred to the documentation of the driving program.

6.17.1 ORCA and Gromacs

In cooperation with the developers of Gromacs software package we developed an interface to ORCA. The interface is available starting with Gromacs version 4.0.4 up to version 4.5.5.

As mentioned above, the initial setup has to be done with the Gromacs. In the QM/MM calculation Gromacs will call ORCA to get the energy and the gradients of the QM atoms. The interface can be used to perform all job types allowed by the Gromacs software package, e.g. optimizations, molecular dynamics, free energy. In addition, for geometry optimizations we have implemented a microiterative scheme that can be requested by setting the keyword `bOpt = yes` in the Gromacs .mdp file. This will cause ORCA to perform a QM geometry optimization every time it is called by Gromacs. During this optimization ORCA will also compute the Lennard-Jones interaction between the QM and MM atoms, and freeze the boundary atoms. This microiterative scheme can also be used to perform a transition state optimization. If you are looking for a transition state and have a good initial guess for the structure, you can carry out an optimization of the MM system and at the same time perform a transition state optimization of the QM system with ORCA via the microiterative scheme. Since it is expensive to calculate the Hessian for each microiterative microiterative step, the user can tell ORCA to use the (updated) Hessian matrix of the previous step via `read_temp_Hess` in the ORCA input.

In order to allow full flexibility to the user, the information for the QM run are provided in a separate plain text file, called `GromacsBasename.ORCAINFO`. When Gromacs writes the input for the ORCAcalculation, it will merge this file with the information on the coordinates, point charges, Lennard-Jones coefficients and type of calculation (EnGrad, Opt, TSOpt).

Below is a typical example of an input file created by Gromacs, where for each Gromacs optimization step, a full optimization of the QM-part will be performed by ORCA, instead of just doing the energy and gradient calculation.

```
# Optimization step in the Lennard-Jones and point charges field of the MM atoms
! QMMMOpt

# file containing the Lennard Jones coefficients for the Lennard-Jones interaction
%LJCoefficients "temp.LJ"
# file containing the point charges for the electrostatic interaction
%pointcharges "temp.pc"

%geom
    # calculate the exact Hessian before the first optimization step
    Calc_Hess true
    # in case of a TS optimization the updated Hessian of the previous
    # TS optimization run is read instead of calculating a new one
    read_temp_Hess true
 end
```

Note

- If you are using `bOpt` or `bTS` you have to take care of additional terms over the boundary. Either you can zero out the dihedral terms of the Q2-Q1-M1-M2 configurations, or you can fix the respective Q2 atoms.
- If you want to use the ORCA constraints, you have to also put them in the Gromacs part of the calculation.

- If there are no bonds between the QM and MM systems, the bOpt optimization of the QM system might have convergence problems. This is the case if the step size is too large and thus QM atoms come too close to MM atoms. The convergence problems can be circumvented by adding extra coordinates and Hessian diagonal values for Cartesian coordinates of selected QM atoms to the set of internal coordinates. This should be done for only few atoms that are in the boundary region.

```
%geom
    # add the Cartesian position of atoms 2 and 4 to the
    # set of internal coordinates with a diagonal Hessian value of 0.1
    Hess_Internal
    {C 2 D 0.1}
    {C 4 D 0.1}
    end
end
```

6.17.2 ORCA and pDynamo

The interface with the pDynamo library is briefly documented here. It uses the same plain text files to exchange information between the pDynamo library and ORCA. In order to have simiar functionality as presented above, we have extended the interface to generate more complex ORCA input files by accepting verbatim blocks of text. We have also implemented in pDynamo the capability of writing files containing Lennard-Jones parameters.

A simple example which calculates the QM/MM energy for a small designed peptide in which the side chain of one amino acid is treated QM is ilustrated below. For this example, a complete geometry optimization is going to be performed during the ORCA call.

```
import os
import pCore
import pBabel
import pMolecule
import pMoleculeScripts

# Read the initial coordinates from the .pdb file.
system = pMoleculeScripts.PDBFile_ToSystem(
        "1UAO.pdb", modelNumber=1, useComponentLibrary=True)

# Instantiate the required models.
mmmodel = pMolecule.MMModelOPLS("protein")
nbmodel = pMolecule.NBModelORCA()
qcmodel = pMolecule.QCModelORCA(
        command=os.getenv("ORCA_COMMAND"),
        deleteJobFiles=False, header="bp86 def2-svp qmmmopt/pdynamo",
        job="chignolin", run=True)

# Assign the models to the system.
system.DefineMMModel(mmmodel)
system.DefineQCModel(
        qcmodel, qcSelection=pCore.Selection([35, 36, 37, 34, 40, 41]))
system.DefineNBModel(nbmodel)
system.electronicState = pMolecule.ElectronicState(
        charge=-1, multiplicity=1)

# Print a summary and calculate the energy.
system.Summary()
system.Energy()
```

After the execution of the above Python program, a series of files are going to be created `chignolin.inp`, `chignolin.pc`, `chignolin.lj` and ORCA is going to be called. The generated ORCAinput file is listed below.

```
! bp86 def2-svp qmmmopt/pdynamo
% geom
    constraints
        {C 0 C}
        {C 1 C}
        end
    end

% pointcharges "chignolin.pc"
% ljcoefficients "chignolin.lj"
* xyz -1 1
H                  -1.0637532468        1.1350324675        2.4244220779
C                  -0.5230000000        0.6870000000        3.2490000000
C                   0.4180000000        1.7240000000        3.8660000000
O                  -0.0690000000        2.7620000000        4.2830000000
O                   1.6090000000        1.4630000000        3.9110000000
H                  -1.2240000000        0.3460000000        3.9970000000
H                   0.0550000000       -0.1510000000        2.8890000000
*
```

There are few points that have to be raised here. Because the keyword `qmmm/pdynamo` was specified in the `header` variable, the pDynamo library will automatically add the `constraint` block in the ORCA input, which will freeze the link atoms and the QM atoms to which they are bound. It will also generate the `chignolin.lj` file containing all the Lennard-Jones parameters. The important parts of this file, which is somewhat different than the one generated by Gromacs, are listed next.

```
# number of atoms  combination rule
138 0
#      x            y            z          sigma       epsilon   id
   -6.778000    -1.424000     4.200000     3.250000     0.711280   -1
   -6.878000    -0.708000     2.896000     3.500000     0.276144   -1
   -5.557000    -0.840000     2.138000     3.750000     0.439320   -1
...
    0.433000     0.826000     0.502000     2.960000     0.878640   -1
   -0.523000     0.687000     3.249000     3.500000     0.276144    1
    0.418000     1.724000     3.866000     3.750000     0.439320    2
   -0.069000     2.762000     4.283000     2.960000     0.878640    3
    1.609000     1.463000     3.911000     2.960000     0.878640    4
   -2.259000    -0.588000     1.846000     0.000000     0.000000   -1
   -1.795000     2.207000     2.427000     2.500000     0.125520   -1
   -1.224000     0.346000     3.997000     2.500000     0.125520    5
    0.055000    -0.151000     2.889000     2.500000     0.125520    6
   -0.311000     2.922000     0.557000     3.250000     0.711280   -1
...
   -1.387000    -2.946000     5.106000     2.500000     0.125520   -1
# number of special pairs
22
#        atom1        atom2    factor
          34           32     0.000000
          35           39     0.500000
          40           31     0.000000
          41           30     0.500000
          41           32     0.500000
          36           31     0.500000
          40           32     0.500000
          40           39     0.500000
          34           31     0.000000
          35           30     0.500000
          34           11     0.500000
          34           38     0.500000
          41           31     0.000000
          37           31     0.500000
```

(continues on next page)

(continued from previous page)

```
          34          33     0.500000
          34          39     0.000000
          40          30     0.500000
          41          39     0.500000
          34          30     0.000000
          35          31     0.000000
          34          42     0.500000
          35          32     0.500000
```

The second number on the first line refers to the type of combination rule used to calculate the Lennard-Jones interaction. It is 0 if a geometric average is used (OPLS force field), or 1 for the Lorentz-Berthelot rules (AMBER force field). The `id` on the last column is -1 for MM atoms and is equal to the atom number for the QM atoms. In this case the hydrogen link atom is atom 0. The last block of the file is composed of atom pairs and a special factor by which their Lennard-Jones interaction is scaled. In general this factor is equal to 1, but for atoms one or two bonds apart is zero, while for atoms three bonds apart depends on the type of force field, and in this case is 0.5.

After successful completion of the ORCA optimization run, the information will be relayed back the pDynamo library, which will report the total QM/MM energy of the system. At this point the type QM/MM of calculation is limited only by the capabilities of the pDynamo library, which are quite extensive.

6.17.3 ORCA and NAMD

Since version 2.12, NAMD is able to perform hybrid QM/MM calculations. A more detailed explanation of all available key words, setting up the calculation and information on tutorials and on the upcoming graphic interface to VMD are available on the NAMD website.

Similar to other calculations with NAMD, the QM/MM is using a pdb file to control the active regions. An example is shown below, where the sidechain of a histidine protonated at Nϵ is chosen to be the QM region. Either the occupancy column or the b-factor column of the file are used to indicate which atom are included in a QM area and which are treated by the forcefield. In the other column, atoms which are connecting the QM area and the MM part are indicated similarly. To clarify which column is used for which purpose, the keywords `qmColumn` and `qmBondColumn` have to be defined in the NAMD input.

```
...
ATOM   1737  CA  HSE P 117      14.762  47.946  31.597  1.00  0.00      PROT C
ATOM   1738  HA  HSE P 117      14.751  47.579  32.616  0.00  0.00      PROT H
ATOM   1739  CB  HSE P 117      14.129  49.300  31.501  1.00  1.00      PROT C
ATOM   1740  HB1 HSE P 117      14.407  49.738  30.518  0.00  1.00      PROT H
ATOM   1741  HB2 HSE P 117      13.024  49.194  31.509  0.00  1.00      PROT H
ATOM   1742  ND1 HSE P 117      13.899  51.381  32.779  0.00  1.00      PROT N
ATOM   1743  CG  HSE P 117      14.572  50.261  32.582  0.00  1.00      PROT C
ATOM   1744  CE1 HSE P 117      14.615  52.043  33.669  0.00  1.00      PROT C
ATOM   1745  HE1 HSE P 117      14.356  53.029  34.064  0.00  1.00      PROT H
ATOM   1746  NE2 HSE P 117      15.678  51.318  33.982  0.00  1.00      PROT N
ATOM   1747  HE2 HSE P 117      16.369  51.641  34.627  0.00  1.00      PROT H
ATOM   1748  CD2 HSE P 117      15.706  50.183  33.335  0.00  1.00      PROT C
ATOM   1749  HD2 HSE P 117      16.451  49.401  33.388  0.00  1.00      PROT H
ATOM   1750  C   HSE P 117      13.916  47.000  30.775  0.00  0.00      PROT C
ATOM   1751  O   HSE P 117      12.965  46.452  31.334  0.00  0.00      PROT O
...
```

NOTES:

- If one wants to include more than one QM region, integers bigger than 1 can be used to define the different regions.
- Charge groups cannot be split when selecting QM and MM region. The reason is that non-integer partial charges may occur if a charge group is split. Since the QM partial charges are updated in every QM iteration, this may lead to a change in the total charge of the system over the course of the MD simulation.

- The occupancy and b-factor columns are used for several declarations in NAMD. If two of these come together in one simulation, the keyword qmParamPDB is used to define which pdb file contains the information about QM atoms and bonds.
- To simplify the selection of QM atoms and writing the pdb file a set of scripts is planned to be included in future releases of NAMD.

To run the calculation, the keyword qmForces on must be set. To select ORCA qmSoftware "orca" must be specified and the path to the executables must be given to qmExecPath, as well as a directory where the calculation is carried out (qmBaseDir). To pass the method and specifications from NAMD to ORCA qmConfigLine is used. These lines will be copied to the beginning of the input file and can contain both simple input as well as block input. To ensure the calculation of the gradient, the engrad keyword should be used.

The geometry of the QM region including the selected links as well as the MM point charges are copied to the ORCA inputfile automatically. Multiplicity and charge can be defined using qmMult and qmCharge, although the latter can be determined automatically by NAMD using the MM parameters. It should be noted at this point that NAMD is capable to handle more than one QM region per QM/MM calculation. Therefore for each region, charge and multiplicity are expected. In the case of only one QM region, the input looks like the following:

```
qmMult          "1 1"
qmCharge        "1 0"
```

Currently, two charge modes are available: Mulliken and CHELPG. They have to be specified in the NAMD input using QMChargeMode and in the qmConfigLine, respectively. Different embedding schemes, point charge schemes and switching functions are available, which will be not further discussed here. Another useful tool worth mentioning is the possibility to call secondary executables before the first or after each QM software execution using QMPrepProc or QMSecProc, respectively. Both are called with the complete path and name to the QM input file, allowing e.g. storage of values during an QM/MM-MD.

It is strongly enphasized that at this points both programs are constantly developed further. For the latest information, either the ORCA forum or the NAMD website should be consulted.

6.18 Excited State Dynamics

ORCA can now also be used to compute dynamic properties involving excited states such as absorption spectra, fluorescence and phosphorescence rates and spectra, as well as resonant Raman spectra using the new ORCA_ESD module. We do this by analytically solving the Fermi's Golden Rule-like equation from Quantum Electrodynamics (see the section *More on the Excited State Dynamics module*), using a path integral approach to the dynamics, as described in our recent papers [198, 199]. The computation of these rates relies on the harmonic approximation for the nuclear normal modes. Provided this approximation holds, the results closely match experimental data.

The theory can do most of what ORCA_ASA can and more, such as including vibronic coupling in forbidden transitions (the so-called Herzberg-Teller effect, HT), considering Duschinsky rotations between modes of different states, solving the equations using different coordinate systems, etc. There are also seven new approaches to obtain the excited state geometry and Hessian without necessarily optimizing its geometry. Many keywords and options are available, but most of the defaults already give good results. Let's get into specific examples, starting with the absorption spectrum. Please refer to section *More on the Excited State Dynamics module* for a complete keyword list and details.

6.18.1 Absorption Spectrum

The ideal model, Adiabatic Hessian (AH)

To predict absorption or emission rates, including all vibronic transitions, ideally, one requires both the ground state (GS) and excited state (ES) geometries and Hessians. For instance, when predicting the absorption spectrum for benzene, which exhibits one band above 220 nm corresponding to a symmetry-forbidden excitation to the S1 state, the process is straightforward. Ground state information can be obtained from (Sec. *Geometry Optimizations, Surface Scans, Transition States, MECPs, Conical Intersections, IRC, NEB*):

```
!B3LYP DEF2-SVP OPT FREQ
* XYZFILE 0 1 BEN.xyz
```

and the S1 ES from (Sec. *Excited State Geometry Optimization*):

```
!B3LYP DEF2-SVP OPT FREQ
%TDDFT
  NROOTS    5
  IROOT     1
END
* XYZFILE 0 1 BEN_S1.xyz
```

Assuming DFT/TD-DFT here, but other methods can also be used (see *Tips, Tricks and Troubleshooting*). With both Hessians available, the ESD module can be accessed from:

```
!B3LYP DEF2-SVP TIGHTSCF ESD(ABS)
%TDDFT
  NROOTS      5
  IROOT       1
END
%ESD
  GSHESSIAN   "BEN.hess"
  ESHESSIAN   "BEN_S1.hess"
  DOHT        TRUE
END
* XYZFILE 0 1 BEN.xyz
```

Important

The geometry must match that in the GS Hessian when calling the ESD module. You can obtain it from the .xyz file after geometry optimization or directly copy it from the .hess file (remember to use BOHRS on the input to correct the units, if obtained from the .hess).

You must provide both names for the Hessians and set DOHT to TRUE here because the first transition of benzene is symmetry forbidden, with an oscillator strength of 2e-6. Therefore, all intensity arises from vibronic coupling (HT effect) [199]. In molecules with strongly allowed transitions, this parameter can typically remain FALSE (the default). Some calculation details are printed, including the computation of transition dipole derivatives for the HT component, and the spectrum is saved as BASENAME.spectrum.

```
Energy                 TotalSpectrum          IntensityFC            IntensityHT
10807.078728           2.545915e-02           2.067393e-07           2.545894e-02
10828.022679           2.550974e-02           2.071508e-07           2.550954e-02
10848.966630           2.556034e-02           2.075624e-07           2.556013e-02
...
```

The first column has the total spectrum, but the contributions from the Franck-Condon part and the Herzberg-Teller part are also discriminated. As you can see, the FC intensity is less than 1% of the HT intensity here, highlighting the importance of including the HT effect. It is important to note that, in theory, the absorbance intensity values correspond to the experimental ε (in L mol cm^{-1}), and they depend on the spectral lineshape. The TotalSpectrum

column can be plotted using any software to obtain the spectrum named Full AH spectrum (shown in blue), in Fig. 6.58 below.

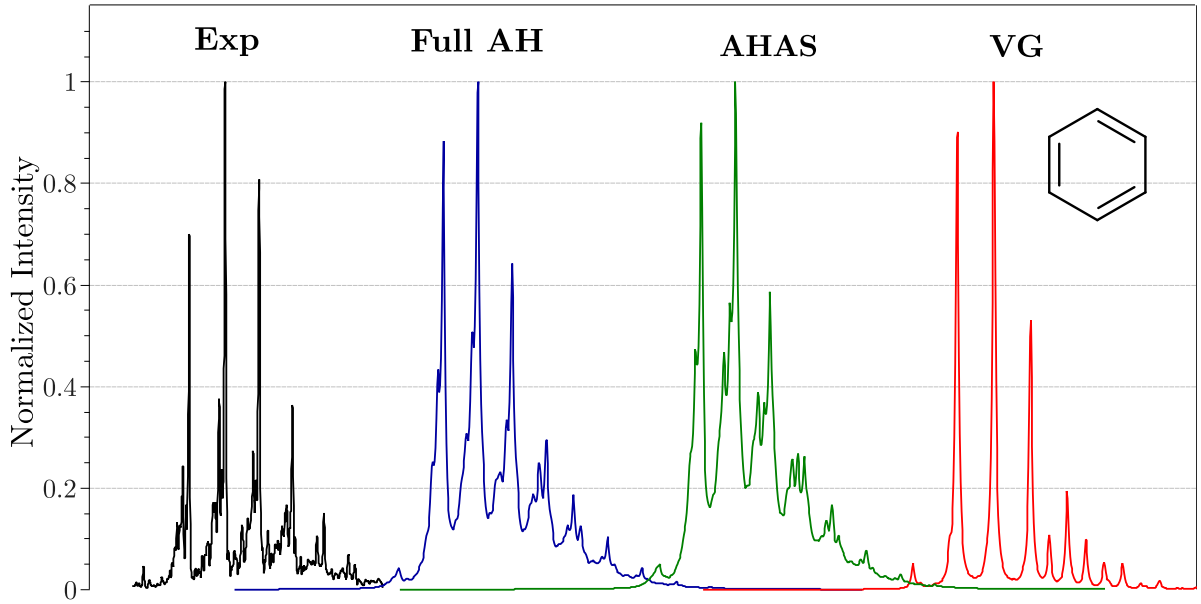


Fig. 6.58: Here is the experimental absorption spectrum for benzene (shown in black on the left), alongside predictions made using ORCA_ESD at various PES approximations.

The spectrum obtained is very close to the experimental results at 298K, even when simply using all the defaults, and it could be further improved by adjusting parameters such as lineshape, as discussed in detail in Sec. *General Aspects* and Sec. *More on the Excited State Dynamics module*.

Note

- The path integral approach in ORCA_ESD is much faster than the more traditional approach of calculating all vibronic transitions with non-negligible intensities, one by one [747]. This is especially true for large systems, where the number of bright vibronic transitions may potentially scale exponentially, but our approach's scaling remains polynomial (in fact near linear scaling in favorable cases [199]). The price to pay is that one can no longer read off the compositions of the vibronic states from the spectrum, in other words, one cannot assign the peaks without doing further calculations. However, one can know whether a given vibrational mode contributes to a given peak, by repeating the ESD calculation with a few modes removed, and see if the peak is still present. This can be conveniently done using the MODELIST or SINGLEMODE keywords. More information can be found in Sec. *More on the Excited State Dynamics module*.
- The Huang-Rhys factors are important tools for qualitative and quantitative analysis of the contributions of each vibrational mode to the vibrationally resolved spectrum (and also to the transition rate constants, as will be discussed later). They can be requested by setting PRINTLEVEL in the %ESD module to 3 or above.

Of course, it is not always possible to obtain the excited state (ES) geometry due to root flipping, or it might be too costly for larger systems. Therefore, some approximations to the ES Potential Energy Surface (PES) have been developed.

The simplest model, Vertical Gradient (VG)

The minimal approximation, known as Vertical Gradient (VG), assumes that the excited state (ES) Hessian equals the ground state (GS) Hessian and extrapolates the ES geometry from the ES gradient and that Hessian using some step (Quasi-Newton or Augmented Hessian, which is the default here). Additionally, in this scenario, the simplest Displaced Oscillator (DO) model is employed, ensuring fast computation [199]. To use this level of approximation, simply provide an input like:

```
!B3LYP DEF2-SVP TIGHTSCF ESD(ABS)
%TDDFT
  NROOTS      5
  IROOT       1
END
%ESD
  GSHESSIAN   "BEN.hess"
  DOHT        TRUE
  HESSFLAG    VG    #DEFAULT
END
* XYZFILE 0 1 BEN.xyz
```

OBS: If no GSHESSIAN is given, it will automatically look for an BASENAME.hess file.

Choosing one of the methods in ORCA to compute excited state information is essential. Here, we utilize TD(A)-DFT with IROOT 1 to compute properties for the first excited state. TD(A)-DFT is currently the sole method offering analytic gradients for excited states; selecting any other method will automatically enforce NUMGRAD.

Important

Please note that certain methods, such as STEOM-DLPNO-CCSD, require significant time to compute numerical gradients. In such cases, we recommend using DFT/TD-DFT Hessians and employing the higher-level method solely for single points.

If everything is set correctly, after the regular single point calculation, the ESD module in ORCA will initiate. It proceeds to obtain the excited state (ES) geometry, compute derivatives, and predict the spectrum. The resulting normalized spectrum can be observed in Fig. 6.58, depicted in red. Due to such simple model, the spectrum is also simplified. While this simplicity is less critical for larger molecules, it highlight the potential benefit of employing an intermediate model.

A better model, Adiabatic Hessian After a Step (AHAS)

A reasonable compromise between a full geometry optimization and a simple step with the same Hessian is to perform a step and then recalculate the ES Hessian at that geometry. This approach is referred to here as Adiabatic Hessian After Step (AHAS). In our tests, it can be invoked with the following input:

```
!B3LYP DEF2-SVP TIGHTSCF ESD(ABS)
%TDDFT
  NROOTS      5
  IROOT       1
END
%ESD
  GSHESSIAN   "BEN.hess"
  DOHT        TRUE
  HESSFLAG    AHAS
END
* XYZFILE 0 1 BEN.xyz
```

The spectrum obtained corresponds to the green line in Fig. 6.58. As shown, it closely resembles the spectrum obtained using AH, where a full geometry optimization was performed. Although not set as the default, this method comes highly recommended based on our experience [199]. Another advantage of this approach is that the

derivatives of the transition dipole are computed simultaneously over Cartesian displacements on the ES structure using the numerical Hessian. Subsequently, these modes are straightforwardly converted.

OBS: The transition dipoles used in our formulation are always those of the FINAL state geometry. For absorption, this corresponds to the ES, so in AHAS, the derivatives are computed over this geometry. For fluorescence, the default behavior is to recompute the derivatives over the GS geometry. Alternatively, you can choose to save time and convert directly from ES to GS by setting CONVDER TRUE (though this is an approximation). For more details, refer to Sec. *More on the Excited State Dynamics module*.

Other PES options

There are also a few other options that can be set using HESSFLAG. For example, you can calculate the vertical ES Hessian over the GS geometry and perform a step, known as the **Vertical Hessian (HESSFLAG VH)** method. This method has the advantage that the geometry step is expected to be better because it does not assume the initial ES Hessian is equal to the GS Hessian. However, it is likely to encounter negative frequencies on that VH, since you are not at the ES minimum. By default, ORCA will turn negative frequencies into positive ones, issuing a warning if any were lower than -300 cm^{-1}. You can also choose to completely remove them (and the corresponding frequencies from the GS) by setting IFREQFLAG to REMOVE or leave them as negative with IFREQFLAG set to LEAVE under %ESD. Just be aware that an odd number of negative frequencies might disrupt the calculation of the correlation function, so be sure to check.

If your excited state is localized and you prefer not to recalculate the entire Hessian, you can opt for a **Hybrid Hessian (HH)** approach. This involves recomputing the ES Hessian only for specific atoms listed in HYBRID_HESS under %FREQ (*Frequency calculations - numerical and analytical*). The HH method uses the GS Hessian as a base but adjusts it at the specified atoms. This computation can be performed either before or after the step, offering two variations: **Hybrid Hessian Before Step (HESSFLAG HHBS)** or **Hybrid Hessian After Step (HESSFLAG HHAS)**. When using either of these options, derivatives are recalculated across the modes as needed.

Another approach involves comparing the ES Hessian with the GS Hessian and selectively recomputing frequencies that differs. This method works by applying a displacement based on the GS Hessian and evaluating the resulting energy change. If the predicted mode matches the actual mode, the prediction should be accurate. However, if the difference exceeds a specified threshold, the gradient is computed, and the frequency for that mode is recalculated accordingly. The final ES Hessian is then derived from the **Updated Frequencies (UF)** and the original GS Hessian.

This approach offers the advantage of minimizing the computation of ES gradients typical in standard ES Hessians, thereby speeding up the process. By default, the system checks for frequency errors of approximately 20%. You can adjust this threshold using the UPDATEFREQERR flag; for instance, setting UPDATEFREQERR to 0.5 under %ESD allows for a larger error tolerance of 50%. Additionally, you can implement either the **Updated Frequencies Before Step (HESSFLAG UFBS)** or the **Updated Frequencies After Step (UFAS)** methods. Transition dipole derivatives are computed concurrently with the update process.

OBS: All these options apply to Fluorescence and resonant Raman as well.

Duschinsky rotations

The ES modes can sometimes be expressed as linear combinations of the GS modes (see Sec. *General Aspects of the Theory*), a phenomenon known in the literature as Duschinsky rotation [228]. In our formulation within ORCA_ESD, it is possible to account for this effect, which reflects a closer approximation to real-world scenarios, albeit at a higher computational cost.

You can enable this feature by setting USEJ TRUE; otherwise, the rotation matrix J defaults to unity. For instance, in the case of benzene, while the effect may not be pronounced, there is noticeable improvement in matching peak ratios with experimental data when incorporating rotations. Exploring this option may reveal more significant impacts in other cases.

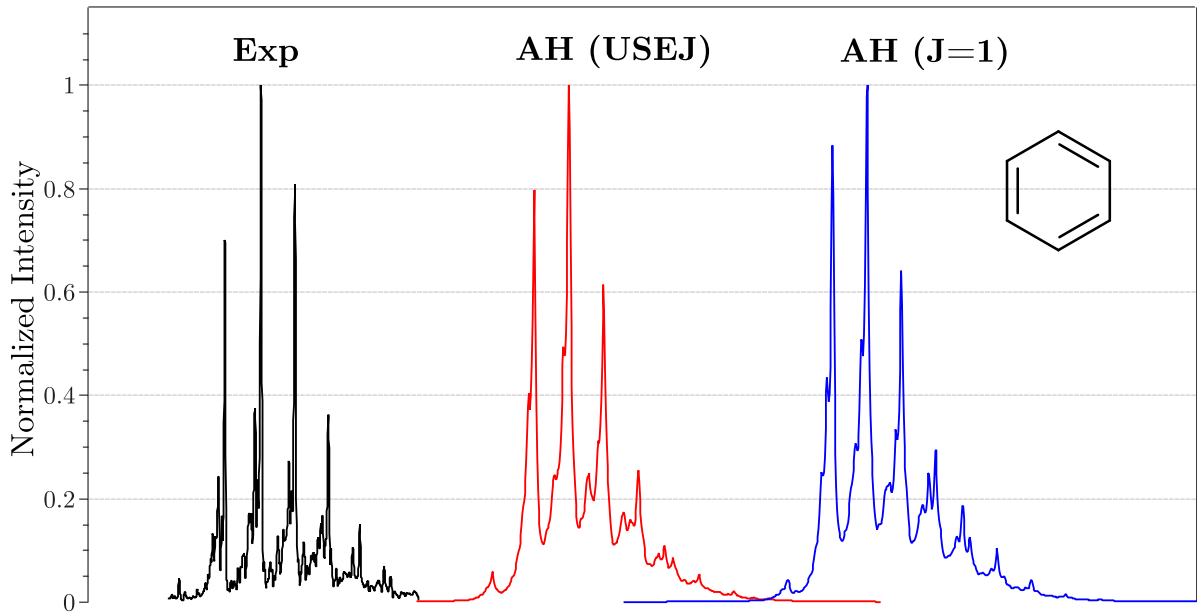


Fig. 6.59: Experimental absorption spectrum for benzene (black on the left) and the effect of Duschinsky rotation on the spectrum.

Temperature effects

In our model, the effects of the Boltzmann distribution due to temperature are exactly accounted for [199]. The default temperature is set to 298.15 K, but you can specify any other temperature by adjusting the TEMP parameter under %ESD. However, it is important to note that when approaching temperatures close to 0 K, numerical issues may arise. For instance, if you encounter difficulties modeling a spectrum at 5 K or wish to predict a jet-cooled spectrum, setting TEMP to 0 will activate a set of equations specifically tailored for T=0 K conditions.

As can be seen in Fig. 6.60, at 0 K there are no hot bands and fewer peaks, while at 600 K there are many more possible transitions due to the population distribution over the GS.

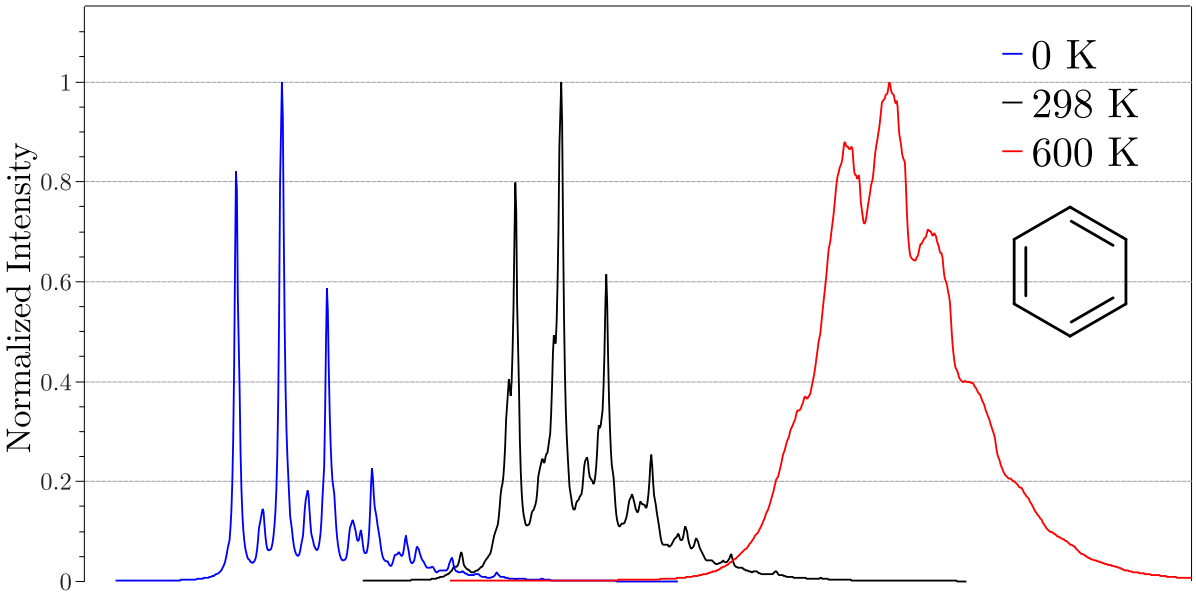


Fig. 6.60: Predicted absorption spectrum for benzene at different temperatures.

Multistate Spectrum

If you want to predict a spectrum that includes many different states, you should ignore the IROOT flag in all modules and instead use the STATES flag under %ESD. For example, to predict the absorption spectra of pyrene in the gas phase and consider the first twenty states, you would specify:

```
!B3LYP DEF2-TZVP(-F) TIGHTSCF ESD(ABS)
%TDDFT
  NROOTS      20
END
%ESD
  GSHESSIAN   "PYR.hess"
  ESHESSIAN   "PYR_S1.hess"
  DOHT        TRUE
  STATES      1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
  UNIT        NM
END
* XYZFILE 0 1 PYR.xyz
```

This input would result in the spectra shown in Fig. 6.61. In this case, each individual spectrum for every state will be saved as BASENAME.spectrum.root1, BASENAME.spectrum.root2, etc., while the combined spectrum, which is the sum of all individual spectra, will be saved as BASENAME.spectrum.

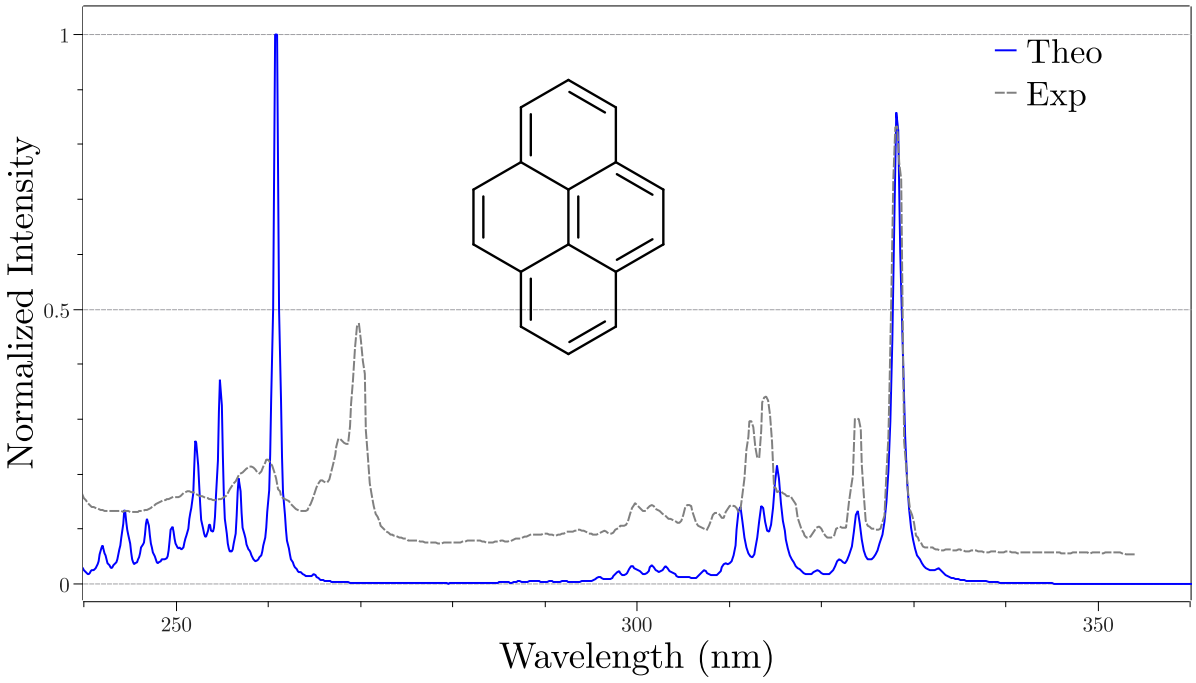


Fig. 6.61: Predicted absorption spectrum for pyrene in gas phase (solid blue) in comparison to the experiment (dashed grey) at 298 K.

OBS: The flag UNIT can be used to control the output unit of the X axis. Its values can be CM-1, NM or EV and it only affects the OUTPUT, the INPUT should always be in cm^{-1}

6.18.2 Fluorescence Rates and Spectrum

General Aspects

The prediction of fluorescence rates and spectra can be performed in a manner analogous to absorption, as described above, by using ESD(FLUOR) on the main input line. You can select any of the methods described earlier to obtain the Potential Energy Surface (PES) by setting the appropriate HESSFLAG. The primary distinction is that the transition dipoles must correspond to the geometry of the ground state (GS), while all other aspects remain largely unchanged.

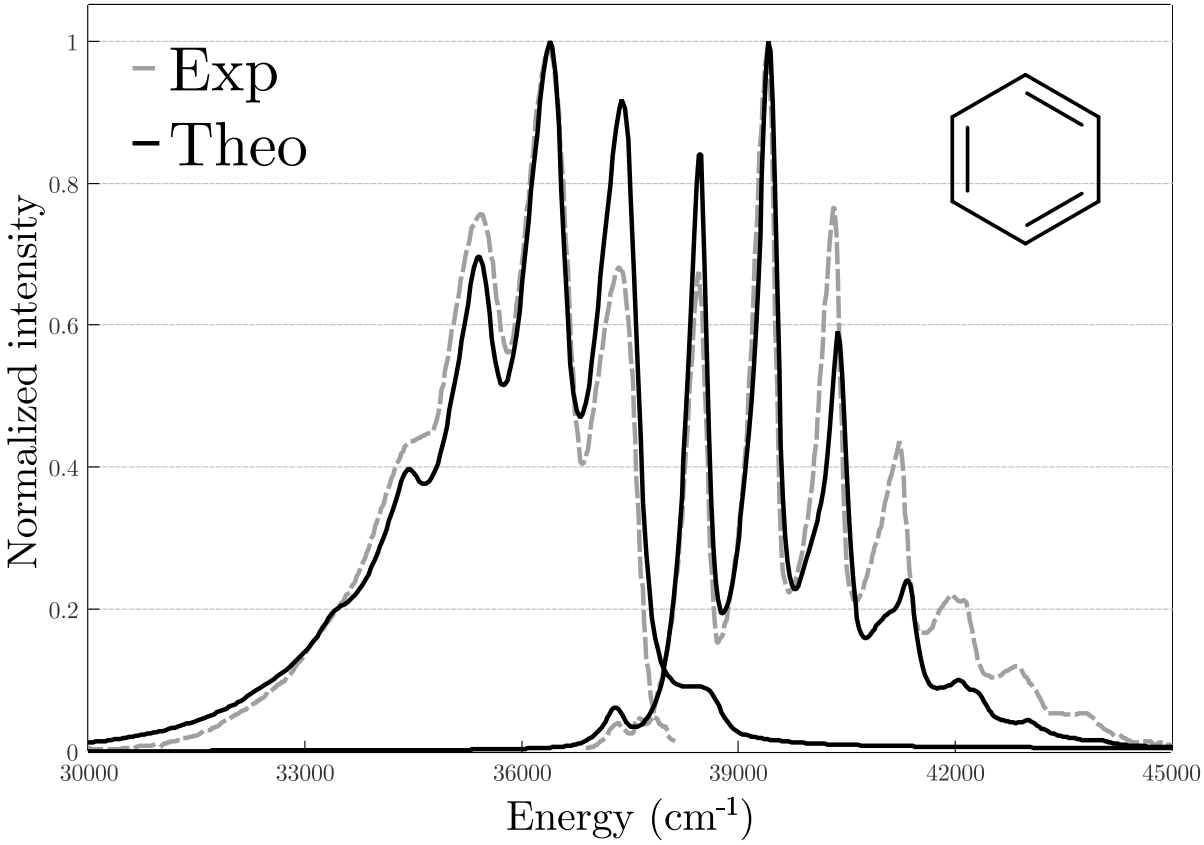


Fig. 6.62: Predicted absorption (right) and emission (left) spectrum for benzene in hexane at 298.15 K.

As depicted in Fig. 6.62, the fluorescence spectrum also closely matches the experimental data [199]. The difference observed in the absorption spectrum in Fig. 6.62, compared to previous spectra, arises because the experiment was conducted in a solvent environment. Therefore, we adjusted the linewidth to align with the experimental data.

OBS: It is common for the experimental lineshape to vary depending on the setup, and this can be adjusted using the LINEW flag (in cm^{-1}). There are four options for the lineshape function controlled by the LINES flag: DELTA (for a Dirac delta function), LORENTZ (default), GAUSS (for a Gaussian function), and VOIGT (a Voigt profile, which is a product of Gaussian and Lorentzian functions).

OBS2: The DELE and TDIP keywords can be used to input adiabatic (not vertical!) excitation energy and transition dipole moment computed at a higher level of theory. This enables calculating the computationally intensive Hessians (especially the excited state Hessian) at a low level of theory without compromising the accuracy. For more details, see Sec. *Mixing methods*.

If you need to control the lineshapes separately for Gaussian (GAUSS) and Lorentzian (LORENTZ), you can set LINEW for Lorentzian and INLINEW for Gaussian (where "I" stands for Inhomogeneous Line Width).

```
!B3LYP DEF2-SVP TIGHTSCF ESD(FLUOR)
%TDDFT
  NROOTS     5
  IROOT      1
END
%ESD
  GSHESSIAN  "BEN.hess"
  ESHESSIAN  "BEN_S1.hess"
  DOHT       TRUE
  LINES      VOIGT
  LINEW      75
  INLINEW    200
END
* XYZFILE 0 1 BEN.xyz
```

OBS: The LINEW and INLINEW are NOT the full width half maximum ($FWHM$) of these curves. However, they are related to them by: $FWHM_{lorentz} = 2 \times LINEW$; $FWHM_{gauss} = 2.355 \times INLINEW$.

For the VOIGT curve, it is a little more complicated but in terms of the other FWHMs, it can be approximated as: $FWHM_{voigt} = 0.5346 \times FWHM_{lorentz} + \sqrt{(0.2166 \times FWHM_{lorentz}^2 + FWHM_{gauss}^2)}$.

Rates and Examples

When you select ESD(FLUOR) on the main input, the fluorescence rate will be printed at the end of the output, with contributions from Franck-Condon (FC) and Herzberg-Teller (HT) mechanisms discriminated. If you use CPCM, the rate will be multiplied by the square of the refractive index, following Strickler and Berg [830].

If you calculate a rate without CPCM but still want to account for the solvent effect, remember to multiply the final rate by this factor. Below is an excerpt from the output of a calculation with CPCM (hexane):

> **Warning**
>
> Whenever using ESD with CIS/TD-DFT and solvation, CPCMEQ will be set to TRUE by default, since the excited state should be under equilibrium conditions! More info in *Including solvation effects via LR-CPCM theory*.

```
...
          ***Everything is set, now computing the correlation function***

Homogeneous linewidht is:                          50.00 cm-1
Number of points:                                  131072
Maximum time:                                       1592.65 fs
Spectral resolution:                               3.33 cm-1
Temperature used:                                  298.15 K
Z value:                                          5.099843e-42
Energy difference:                                  41049.37 cm-1
Reference transition dipole (x,y,z):                (0.00004 0.00000),
                                                    (0.00002 0.00000),
                                                    (-0.00058 0.00000)
Calculating correlation function:                  ...done
Last element of the correlation function:          0.000000,-0.000000
Computing the Fourier Transform:                  ...done

The calculated fluorescence rate constant is        1.688355e+06 s-1*
with 0.00% from FC and 100.00% from HT

*The rate is multiplied by the square of the refractive index

The fluorescence spectrum was saved in                BASENAME.spectrum
```

In one of our theory papers, we investigated the calculation of fluorescence rates for the set of molecules presented in Fig. 6.63. The results are summarized in Fig. 6.64 for some of the methods used to obtain the Potential Energy Surface (PES) mentioned.

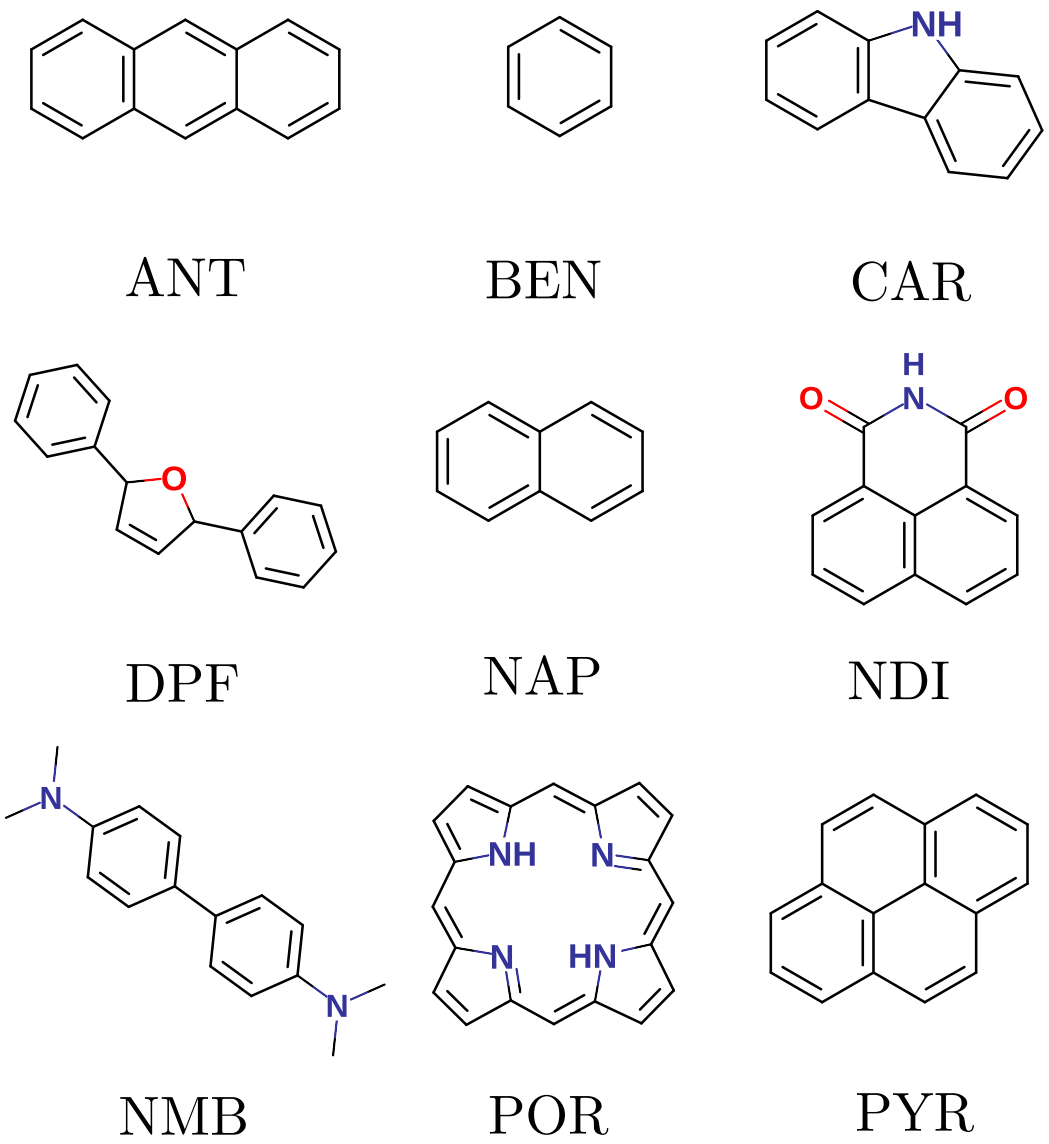


Fig. 6.63: The set of molecules studied, with rates on Fig. 6.64.

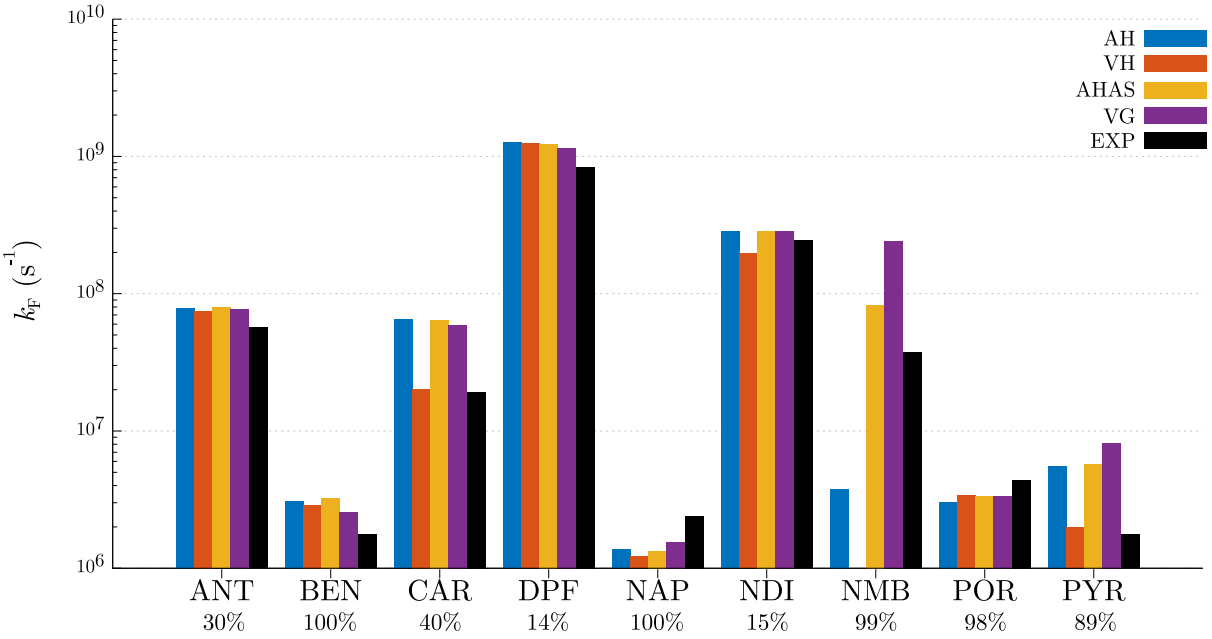


Fig. 6.64: Predicted emission rates for various molecules in hexane at 298.15 K. The numbers below the labels are the HT contribution to the rates.

6.18.3 Phosphorescence Rates and Spectrum

General Aspects

As with fluorescence, phosphorescence rates and spectra can be calculated if spin-orbit coupling is included in the excited state module (please refer to the relevant publication [198]). To enable this, ESD(PHOSP) must be selected in the main input, and both a GSHESSIAN and a TSHESSIAN must be provided. The triplet Hessian can be computed analytically from the spin-adapted triplets.

```
!B3LYP DEF2-TZVP(-F) CPCM(ETHANOL) OPT FREQ
%TDDFT
  NROOTS 5
  IROOTMULT TRIPLET
END
* XYZ 0 1
C         -0.82240       -0.05739        0.00515
C          0.42295        0.77803        0.02146
H         -0.85252       -0.69527        0.89195
H         -0.85090       -0.66429       -0.90325
H         -1.69889        0.59680        0.01431
C          1.74379        0.02561       -0.01818
C          2.98907        0.86121       -0.00686
H          3.01366        1.50199       -0.89176
H          3.86561        0.20724       -0.02398
H          3.02300        1.46514        0.90332
O          0.42398        2.00161        0.06749
O          1.74282       -1.19814       -0.05965
*
```

or, in this case, by computing the ground state triplet by simply setting the multiplicity to three:

```
!B3LYP DEF2-TZVP(-F) CPCM(ETHANOL) OPT FREQ
* XYZFILE 0 3 BIA.xyz
```

Alternatively, one can use methods like VG, AHAS, etc., to approximate the triplet geometry and Hessian. However, this approach requires preparing the Hessian in a separate ESD run (Sec. *Approximations to the excited state*

PES).

Additionally, you must input the adiabatic energy difference between the ground singlet and ground triplet states at their respective geometries (without any zero-point energy correction) using the DELE flag under %ESD. In this case, the spin-adapted triplet computed previously serves as our reference triplet state. An example input for the rate calculation using TDDFT is as follows:

```
!B3LYP DEF2-TZVP(-F) TIGHTSCF CPCM(ETHANOL) ESD(PHOSP) RI-SOMF(1X)
%TDDFT
  NROOTS  20
  DOSOC   TRUE
  TDA     FALSE
  IROOT   1
END
%ESD    GSHESSIAN       "BIA.hess"
        TSHESSIAN       "BIA_T1.hess"
        DOHT            TRUE
        DELE            17260
END
* XYZFILE 0 1 BIA.xyz

$NEW_JOB

!B3LYP DEF2-TZVP(-F) TIGHTSCF CPCM(ETHANOL) ESD(PHOSP) RI-SOMF(1X)
%TDDFT
  NROOTS  20
  DOSOC   TRUE
  TDA     FALSE
  IROOT   2
END
%ESD    GSHESSIAN       "BIA.hess"
        TSHESSIAN       "BIA_T1.hess"
        DOHT            TRUE
        DELE            17260
END
* XYZFILE 0 1 BIA.xyz

$NEW_JOB

!B3LYP DEF2-TZVP(-F) TIGHTSCF CPCM(ETHANOL) ESD(PHOSP) RI-SOMF(1X)
%TDDFT
  NROOTS  20
  DOSOC   TRUE
  TDA     FALSE
  IROOT   3
END
%ESD    GSHESSIAN       "BIA.hess"
        TSHESSIAN       "BIA_T1.hess"
        DOHT            TRUE
        DELE            17260
END
* XYZFILE 0 1 BIA.xyz
```

Phosphorescence rate calculation are always accompanied by the generation of the vibrationally resolved phosphorescence spectrum, which can be visualized in the same way as fluorescence spectra.

OBS.: When computing phosphorescence rates, each rate from individual spin sub-levels must be requested separately. You may use the $NEW_JOB option, just changing the `IROOT`, to write everything in a single input. After SOC, the three triplet states (T_1 with M_S = -1, 0 and +1) will split into IROOTs 1, 2 and 3, and all of them must be included when computing the final phosphorescence rate. In this case, it is reasonable to assume that the geometries and Hessians of these spin sub-levels are the same, and we will use the same .hess file for all three.

OBS2.: The ground state geometry should be used in the input file, similar to the case of fluorescence (vide supra).

OBS3.: Apart from DELE, one can also use the TDIP keyword to input a high-level transition dipole moment, similar to the fluorescence case. This enables e.g. the calculation of phosphorescence rates/spectra using e.g. NEVPT2 or DLPNO-STEOM-CCSD transition dipole moments, with (TD)DFT geometries and Hessians. Note however that the transition dipole moment in phosphorescence processes is complex, so 6 instead of 3 components are required.

Here, we are computing the rate and spectrum for biacetyl in ethanol at 298 K. The geometries and Hessians were obtained as previously described, with the ground triplet computed from a simple open-shell calculation. To compute the rate, the flag DOSOC must be set to TRUE under %TDDFT (Sec *Spin-orbit coupling*), or the respective module, and it is advisable to set a large number of roots to ensure a good mixing of states.

Please note that we have chosen the RI-SOMF(1X) option for the spin-orbit coupling integrals, but any of the available methods can be used (Sec. *The Spin-Orbit Coupling Operator*).

Calculation of rates

As you can see, the predicted spectra for biacetyl (Fig. 6.65) are quite close to the experimental results [198, 783]. The calculation of the phosphorescence rate is more complex because there are three triplet states that contribute. Therefore, the observed rate must be taken as an average of these three states.

$$k_{av}^{phosp} = \frac{k_1 + k_2 + k_3}{3}$$

To be even more strict and account for the Boltzmann population distribution at a given temperature T caused by the Zero Field Splitting (ZFS), one should use [592]:

$$k_{av}^{phosp} = \frac{k_1 + k_2 e^{-(\Delta E_{1,2}/k_B T)} + k_3 e^{-(\Delta E_{1,3}/k_B T)}}{1 + e^{-(\Delta E_{1,2}/k_B T)} + e^{-(\Delta E_{1,3}/k_B T)}} \tag{6.26}$$

where $\Delta E_{1,2}$ is the energy difference between the first and second states, and so on.

After completion of each calculation, the rates for the three triplets were 8.91 s^{-1}, 0.55 s^{-1}, and 284 s^{-1}. Using (6.26), the final calculated rate is about 98 s^{-1}, while the best experimental value is 102 s^{-1} (at 77K) [591], with about 40% deriving from the HT effect.

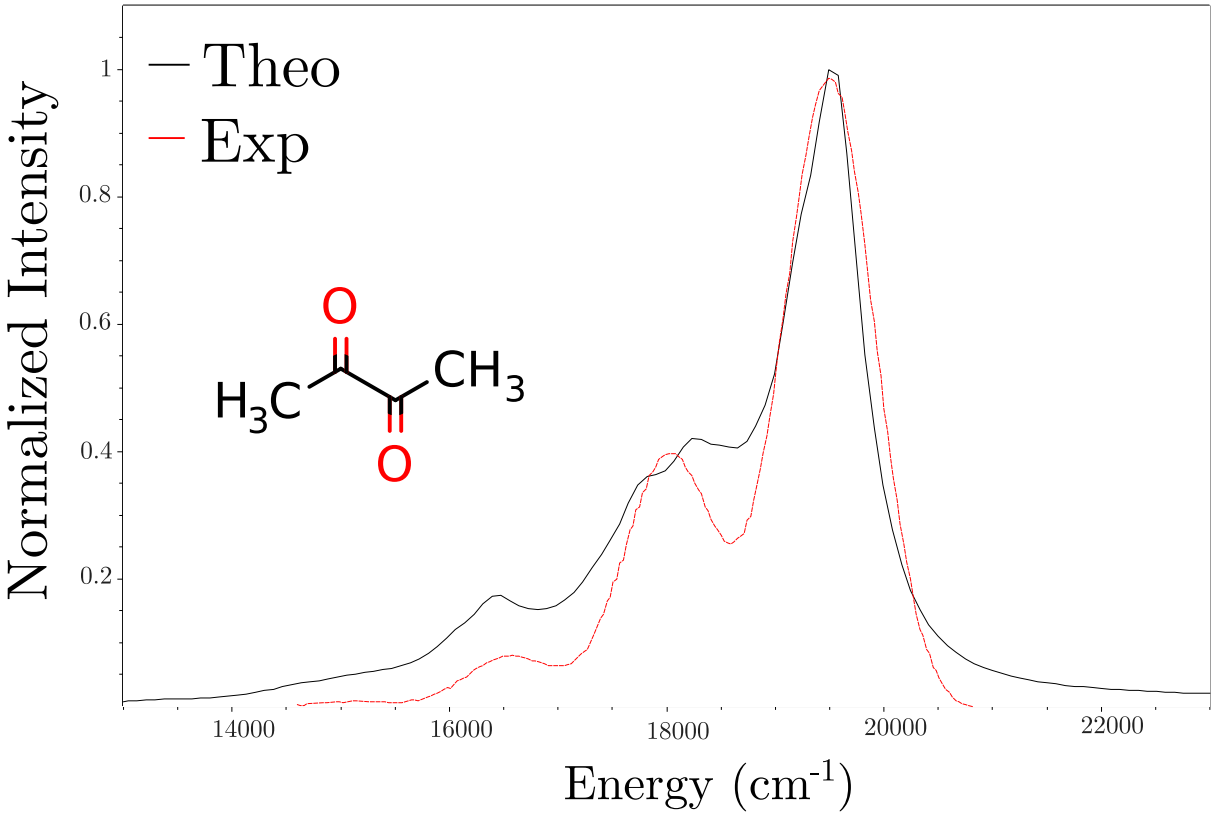


Fig. 6.65: The experimental (dashed red) and theoretical (solid black, displaced by about 2800 cm^{-1}) phosphorescence spectra for biacetyl, in ethanol at 298 K.

OBS: A subtlety arises when the final state is not a singlet state, for example in radical phosphorescence (doublet ground state) or singlet oxygen phosphorescence (triplet ground state). In this case the most rigorous treatment would be to sum over the final states but average over the initial states. For example, with quartet-to-doublet phosphorescence one gets

$$k_{av}^{phosp} = \frac{k_{1\to1} + k_{2\to1}e^{-(\Delta E_{1,2}/k_BT)} + k_{3\to1}e^{-(\Delta E_{1,3}/k_BT)} + k_{4\to1}e^{-(\Delta E_{1,4}/k_BT)}}{1 + e^{-(\Delta E_{1,2}/k_BT)} + e^{-(\Delta E_{1,3}/k_BT)} + e^{-(\Delta E_{1,4}/k_BT)}} + \frac{k_{1\to2} + k_{2\to2}e^{-(\Delta E_{1,2}/k_BT)} + k_{3\to2}e^{-(\Delta E_{1,3}/k_BT)} + k_{4\to2}e^{-(\Delta E_{1,4}/k_BT)}}{1 + e^{-(\Delta E_{1,2}/k_BT)} + e^{-(\Delta E_{1,3}/k_BT)} + e^{-(\Delta E_{1,4}/k_BT)}}$$

where $k_{2\to1}$ is the phosphorescence rate constant from the second sublevel of the initial quartet to the first sublevel of the final doublet, etc. Note that the Boltzmann factors of the final state do not enter the expression. Unfortunately, since U-TDDFT is spin contaminated and unsuitable for calculating SOC-corrected transition dipole moments, the transition dipoles in this case have to be calculated by more advanced methods, such as DFT/ROCIS or multireference methods. The transition dipole should then be given in the input file using the TDIP keyword.

6.18.4 Intersystem Crossing Rates (unpublished)

General Aspects

Yet another application of the path integral approach is to compute intersystem crossing rates, or non-radiative transition rates between states of different multiplicities. This can be calculated if one has two geometries, two Hessians, and the relevant spin-orbit coupling matrix elements.

The input is similar to those discussed above. Here ESD(ISC) should be used on the main input to indicate an InterSystem Crossing calculation, and the Hessians should be provided by ISCISHESSIAN and ISCFSHESSIAN for the initial and final states, respectively. Please note that the geometry used in the input file should correspond to that of the FINAL state, specified through the ISCFSHESSIAN flag. The relevant matrix elements can be computed

using any method available in ORCA and inputted as SOCME Re,Im under %ESD where Re and Im represent its real and imaginary parts (**in atomic units!**).

As a simple example, one could compute the excited singlet and ground triplet geometries and Hessians for anthracene using TD-DFT. Then, compute the spin-orbit coupling (SOC) matrix elements for a specific triplet spin-sublevel using the same method (see the details below), potentially employing methods like CASSCF, MRCI, STEOM-CCSD, or another theoretical level. Finally, obtain the intersystem crossing (ISC) rates using an input such as:

```
!ESD(ISC) NOITER
%ESD
  ISCISHESSIAN  "ANT_S1.hess"
  ISCFSHESSIAN  "ANT_T1.hess"
  DELE          11548
  SOCME         0.0, 2.33e-5
END
* XYZFILE 0 1 ANT_T1.xyz
```

The SOCMEs between a singlet state and a triplet state consist of three complex numbers, not just one, because the triplet state has three sublevels. If the user uses the SOCME of one of the sublevels as input to the ISC rate calculation, this gives the ISC rate associated with that particular sublevel. However, experimentally one usually does not distinguish the three sublevels of a triplet state, and experimentally ISC rates are reported as if the three sublevels of a triplet state are the same species. Therefore, the rate of singlet-to-triplet ISC is the *sum* of the ISC rates from the singlet state to the three triplet sublevels. Fortunately, in case the Herzberg-Teller effect (vide infra) can be neglected, it is not necessary to perform three rate calculations and add up the rates, since the rate is proportional to the squared modulus of the SOCME. Thus, one can run a *single* ESD calculation where the SOCME is the square root of the summed squared moduli of the three SOCME components.

As an illustration, consider the S_1 to T_1 ISC rate of acetophenone. First, we optimize the S_0 geometry, and (after manually tweaking the geometry to break mirror symmetry) use it as an initial guess for the geometry optimization and frequency calculations of the S_1 and T_1 states. Then, we calculate the S_1-T_1 SOCMEs at the T_1 geometry (note that, as usual, final state geometries should be used for ESD calculations; this may differ from some programs other than ORCA). These calculations are conveniently done using a compound script, although the individual steps can of course also be done using separate input files.

```
%pal nprocs 16 end

* xyz 0 1
  C      1.512698    7.783764   -0.013405
  C      2.900029    7.735359   -0.016012
  C      3.664170    8.950745    0.000408
  C      2.952045   10.199539    0.007630
  C      1.564497   10.214795    0.009890
  C      0.827311    9.015246    0.000726
  H      0.946857    6.847253   -0.027731
  H      3.394565    6.763402   -0.041976
  H      3.505090   11.140193    0.017889
  H      1.039592   11.175094    0.019479
  H     -0.265468    9.038211    0.000969
  C      5.059695    8.958412    0.000442
  O      5.733735   10.161120   -0.048672
  C      5.927824    7.713913    0.020029
  H      5.442615    6.938582    0.639989
  H      6.073177    7.326421   -1.014490
  H      6.923426    7.950459    0.447785
 *

%compound

new_step # Compound 1: S1 opt
! B3LYP def2-SV(P) opt tightopt freq
```

(continues on next page)

(continued from previous page)

```
%tddft
tda false
nroots 2
iroot 1
end

step_end

new_step # Compound 2: T1 opt
! B3LYP def2-SV(P) opt tightopt freq

* xyz 0 3
  C      1.512698    7.783764   -0.013405
  C      2.900029    7.735359   -0.016012
  C      3.664170    8.950745    0.000408
  C      2.952045   10.199539    0.007630
  C      1.564497   10.214795    0.009890
  C      0.827311    9.015246    0.000726
  H      0.946857    6.847253   -0.027731
  H      3.394565    6.763402   -0.041976
  H      3.505090   11.140193    0.017889
  H      1.039592   11.175094    0.019479
  H     -0.265468    9.038211    0.000969
  C      5.059695    8.958412    0.000442
  O      5.733735   10.161120   -0.048672
  C      5.927824    7.713913    0.020029
  H      5.442615    6.938582    0.639989
  H      6.073177    7.326421   -1.014490
  H      6.923426    7.950459    0.447785
 *

step_end

new_step # Compound 3: SOC at T1 geometry
! B3LYP def2-SV(P)

%tddft
tda false
nroots 3
triplets true
dosoc true
end

# Here it is assumed that the input file is named "a.inp"
* xyzfile 0 1 a_Compound_2.xyz

step_end
end
```

After verifying that neither of the Hessians have imaginary frequencies (which is very important!), we can read the computed SOCMEs:

```
     ---------------------------------------------------------------------------
                  CALCULATED SOCME BETWEEN TRIPLETS AND SINGLETS
     ---------------------------------------------------------------------------
          Root                          <T|HSO|S>  (Re, Im) cm-1
        T      S             Z                      X                      Y
     ---------------------------------------------------------------------------
        1      0     (   0.00 ,   -2.00 )    (   0.00 ,   29.24 )    (  -0.00 ,  -41.11 )
        1      1     (   0.00 ,   -0.59 )    (   0.00 ,    1.79 )    (  -0.00 ,   -2.57 )
        1      2     (   0.00 ,    0.52 )    (   0.00 ,   -3.19 )    (  -0.00 ,    3.21 )
```

(continues on next page)

(continued from previous page)

```
      1      3    (   0.00 ,   -2.39 )    (   0.00 ,   14.57 )    (  -0.00 ,  -19.53 )
      2      0    (   0.00 ,   -0.11 )    (   0.00 ,    2.78 )    (  -0.00 ,   -2.72 )
      2      1    (   0.00 ,    2.25 )    (   0.00 ,  -20.07 )    (  -0.00 ,   28.07 )
      2      2    (   0.00 ,    0.02 )    (   0.00 ,   -0.13 )    (  -0.00 ,    0.29 )
      2      3    (   0.00 ,   -0.17 )    (   0.00 ,    0.59 )    (  -0.00 ,   -1.27 )
      3      0    (   0.00 ,   -0.01 )    (   0.00 ,    0.05 )    (  -0.00 ,   -0.06 )
      3      1    (   0.00 ,   -0.01 )    (   0.00 ,    0.19 )    (  -0.00 ,   -1.36 )
      3      2    (   0.00 ,   -0.01 )    (   0.00 ,   -0.05 )    (  -0.00 ,   -0.14 )
      3      3    (   0.00 ,   -0.00 )    (   0.00 ,    0.14 )    (  -0.00 ,   -0.16 )
```

The "total" SOCME between S_1 and T_1 is then calculated, from the line that begins with "1 1", as

$$\sqrt{0.00^2 + (-0.59)^2 + 0.00^2 + 1.79^2 + 0.00^2 + 2.57^2} cm^{-1} = 3.19 cm^{-1} = 1.45 \times 10^{-5} au$$

One should therefore write `socme 1.45e-5` in the `%esd` block in the subsequent ISC rate calculation.

Importantly, the above approach is only applicable to singlet-to-triplet ISC, but not to triplet-to-singlet ISC (including, but not limited to, $T_1 \rightarrow S_0$ and $T_1 \rightarrow S_1$ processes). In the latter case, assuming that the triplet sublevels are degenerate and in rapid equilibrium, we obtain that the observed rate constant is the *average*, not the sum, of the rate constants of the three triplet sublevels, because each triplet sublevel has a Boltzmann weight of 1/3. Therefore, the "effective" SOCME that should be plugged into the ESD module to get the observed rate constant is (here the squared modulus $|SOCME_x|^2$ should be calculated as $Re(SOCME_x)^2 + Im(SOCME_x)^2$, etc.)

$$SOCME_{av} = \sqrt{\frac{|SOCME_x|^2 + |SOCME_y|^2 + |SOCME_z|^2}{3}} \tag{6.27}$$

i.e. a factor of $\sqrt{3}$ smaller than the singlet-to-triplet case. However, both of the two assumptions (degenerate triplet sublevels, and rapid equilibrium between sublevels) may fail under certain circumstances, which may contribute to the error of the predicted rate. Nevertheless, in many cases the present, simple approach still provides a rate with at least the correct order of magnitude.

OBS.: The adiabatic energy difference is NOT computed automatically for ESD(ISC), so you must provide it in the input. This is the energy of the initial state minus the energy of the final state, each evaluated at its respective geometry.

OBS2.: All the other options concerning changes of coordinate system, Duschinsky rotation, etc., are also available here.

OBS3.: For many molecules, the $S_1 \rightarrow T_1$ ISC process is not the dominant ISC pathway. This is because the excited state compositions of S_1 and T_1 are often similar, and therefore ISC transitions between them frequently do not satisfy the El-Sayed rule. Even if the only experimentally observed excited states are S_1 and T_1, it may still be that the initial ISC is dominated by $S_1 \rightarrow T_n (n > 1)$, followed by ultrafast $T_n \rightarrow T_1$ internal conversion.

OBS4.: Similarly, if the molecule starts at a high singlet state $S_n (n > 1)$, the dominant ISC pathway is not necessarily the direct ISC from S_n to one of the triplet states. Rather, it is possible (but not necessarily true) that S_n first decays to a lower singlet excited state before the ISC occurs.

OBS5.: If you calculate the DELE or SOCMEs at a higher level of theory and use it as an input for the ESD calculation, please make sure that you have chosen the same excited state (in terms of state composition, not energy ordering) in the Hessian and DELE/SOCME calculations. For example, suppose that you have obtained the geometry and Hessian of the T_2 state, but the T_2 state of the higher level of theory has a very different state composition than the T_2 state at the level of theory used in the Hessian calculation; rather, it is T_3 at the high level of theory that shares the same composition as the T_2 state at the lower level of theory. In this case, you should use the SOCME related to T_3 in the SOCME output file.

ISC, TD-DFT and the HT effect

In the anthracene example above, the result is an ISC rate (k_{ISC}) smaller than $1s^{-1}$, which is quite different from the experimental value of $10^8 s^{-1}$ at $77K$ [591]. The reason for this discrepancy, in this particular case, is because the ISC occurs primarily due to the Herzberg-Teller effect, which must also be included. To achieve this, one needs to compute the derivatives of the SOCMEs over the normal modes, currently feasible only using CIS/TD-DFT.

When using the %CIS/TDDFT option, you can control the SROOT and TROOT flags to select the singlet and triplet states for which SOCMEs are computed, and the TROOTSSL flag to specify the triplet spin-sublevel (1, 0, or -1).

In practice, to obtain an ISC rate (k_{ISC}) close enough to experimental values, one would need to consider all possible transitions between the initial singlet and all available final states. For anthracene, these are predicted to be the ground triplet (T_1) and the first excited triplet (T_2), consistent with experimental observations [402], while the next triplet (T_3) is energetically too high to be significant (Fig. 6.66 below). An example input used to calculate the k_{ISC} from S_1 to T_1 at $77K$ is:

```
!B3LYP DEF2-TZVP(-F) TIGHTSCF ESD(ISC)
%TDDFT  NROOTS  5
        SROOT   1
        TROOT   1
        TROOTSSL 0
        DOSOC   TRUE
END
%ESD    ISCISHESS       "ANT_S1.hess"
        ISCFSHESS       "ANT_T1.hess"
        USEJ            TRUE
        DOHT            TRUE
        TEMP            77
        DELE            11548
END
* XYZFILE 0 1 ANT_T1.xyz

$NEW_JOB

!B3LYP DEF2-TZVP(-F) TIGHTSCF ESD(ISC)
%TDDFT  NROOTS  5
        SROOT   1
        TROOT   1
        TROOTSSL -1
        DOSOC   TRUE
END
%ESD    ISCISHESS       "ANT_S1.hess"
        ISCFSHESS       "ANT_T1.hess"
        USEJ            TRUE
        DOHT            TRUE
        TEMP            77
        DELE            11548
END
* XYZFILE 0 1 ANT_T1.xyz

...
```

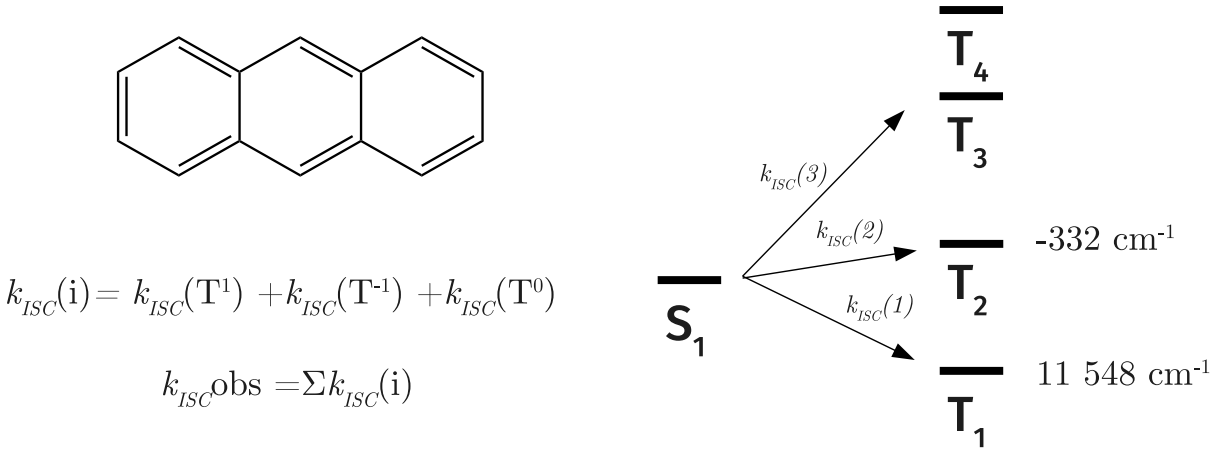


Fig. 6.66: Scheme for the calculation of the intersystem crossing in anthracene. The $k_{ISC}(i)$ between S_1 and each triplet is a sum of all transitions to the spin-sublevels and the actual observed k_{ISC}^{obs}, which consolidates these transitions. On the right, there is a diagram illustrating the distribution of excited states with $E(S_1) - E(T_n)$ depicted on the side. Since T_3 is energetically too high, intersystem crossing involving T_3 can be safely neglected.

Then, the derivatives of the SOCME are computed and the rates are printed at the end. By performing the same calculations for the T_2 states and summing up these values, a predicted $k_{ISC}^{obs} = 1.17 \times 10^8 s^{-1}$ can be obtained, much closer to the experimental value, which is associated with a large error anyway.

OBS.: In cases where the SOCME are relatively large, e.g., SOCME $> 5cm^{-1}$, the Herzberg-Teller effect might be negligible, and a simple Frank-Condon calculation should yield good results. This is typically applicable to molecules with heavy atoms, where vibronic coupling is less significant.

OBS2.: Always consider that there are actually THREE triplet spin-sublevels, and transitions from the singlet to all of them should be included.

OBS3.: ISC rates are extremely sensitive to energy differences. Exercise caution when calculating them. If a more accurate excited state method is available, it should be considered for prediction.

6.18.5 Internal Conversion Rates (unpublished)

The ESD module can also calculate internal conversion (IC) rates from an excited state to the ground state at the TDDFT and TDA levels. Apart from the ground state and excited state Hessians, the only additional quantity that needs to be calculated is the nonadiabatic coupling matrix elements (NACMEs).

The input file is simple:

```
# S1-S0 IC rate of azulene
! B3LYP D3 def2-SVP ESD(IC) CPCM(methanol)

%tddft
tda false # Full TDDFT is recommended over TDA
nroots 3
iroot 1 # Change to 2 for S2-S0 IC rate, etc.
nacme true # Calculate the NACME between the iroot-th root with the ground state
etf true # Use electron translation factor (recommended)
end

%esd
gshessian "azulene_S0.hess" # Ground state Hessian (B3LYP-D3/def2-SVP)
eshessian "azulene_S1.hess" # Excited state Hessian (TD-B3LYP-D3/def2-SVP)
usej true # Use Duschinsky rotation (recommended)
end
```

(continues on next page)

(continued from previous page)

```
# Excited state geometry (TD-B3LYP-D3/def2-SVP)
* xyzfile 0 1 azulene_S1.xyz
```

Here the S_0 geometry, as well as the S_0 and S_1 Hessian files, were obtained at the B3LYP-D3/def2-SVP level of theory. Note that the NACME calculation uses full TDDFT and also includes the electron-translation factor (ETF), which are the recommended practices in general. The "iroot 1" specifies that the initial state is S_1; the final state is always S_0 and this cannot be changed.

The computed IC rate constant is given near the end of the output file:

```
          ***Everything is set, now computing the correlation function***

Homogeneous linewidth is:     50.00 cm-1
Inhomogeneous linewidth is:     250.00 cm-1
Number of points:       32768
Maximum time:         157.86 fs
Temperature used:       298.15 K
Z value:          4.924300e-66
Cutoff for the correlation function:    1.00e-12
Adiabatic energy difference:      16699.89 cm-1
0-0 energy difference:        16382.09 cm-1
Calculating correlation function:   ...done
Last element of the correlation function: -0.000000,-0.000000

The calculated internal conversion rate constant is 3.823146e+08 s-1

Total run time: 0 hours 2 minutes 50 seconds

                      ****ORCA ESD FINISHED WITHOUT ERROR****
```

For more accurate results, one may add explicit solvation shells, since implicit solvation models only describe the electrostatic and dispersive effects of the solvent on the solute, but cannot provide the extra vibrational degrees of freedom that can help dissipating the excitation energy. Conversely, by using an implicit solvent one also misses the effect of the solvent viscosity on inhibiting the internal conversion, which is particularly important when one wants to compare against experiments conducted at low temperatures.

6.18.6 Resonant Raman Spectrum

General Aspects

Using a theoretical framework similar to what was published for Absorption and Fluorescence, we have also developed a method to compute resonant Raman spectra for molecules [197]. In this implementation, one can employ all the methods to obtain the excited state potential energy surfaces (PES) mentioned earlier using HESSFLAG, and include Duschinsky rotations and even consider the Herzberg-Teller effect on top of it. This calculation can be initiated by using ESD(RR) or ESD(RRAMAN) on the first input line. It is important to note that by default, we calculate the "Scattering Factor" or "Raman Activity," as described by D. A. Long [533] (see Sec. *General Aspects of the Theory* for more information).

When using this module, the laser energy can be controlled by the LASERE flag. If no laser energy is specified, the 0-0 energy difference is used by default. You can select multiple energies by using LASERE 10000, 15000, 20000, etc., and if multiple energies are specified, a series of files named BASENAME.spectrum.LASERE will be saved. Additionally, it is possible to specify several states of interest using the STATES flag, but not both simultaneously.

As an example, let's predict the resonant Raman spectrum of the phenoxyl radical. You need at least a ground state geometry and Hessian, and then you can initiate the ESD calculation using:

```
!PBE0 DEF2-SVP TIGHTSCF ESD(RR)
%TDDFT
  NROOTS     5
```

(continues on next page)

(continued from previous page)

```
  IROOT      3
END
%ESD
  GSHESSIAN  "PHE.hess"
  LASERE     28468
END
* XYZFILE 0 2 PHE.xyz
```

Important

The LASERE used in the input is NOT necessarily the same as the experimental one. It should be proportional to the theoretical transition energy. For example, if the experimental 0-0 ΔE is 30000 cm^{-1} and the laser energy used is 28000 cm^{-1}, then for a theoretical ΔE of 33000 cm^{-1}, you should use a laser energy of 31000 cm^{-1} to obtain the corresponding theoretical result. At the end of the ESD output, the theoretical 0-0 ΔE is printed for your information.

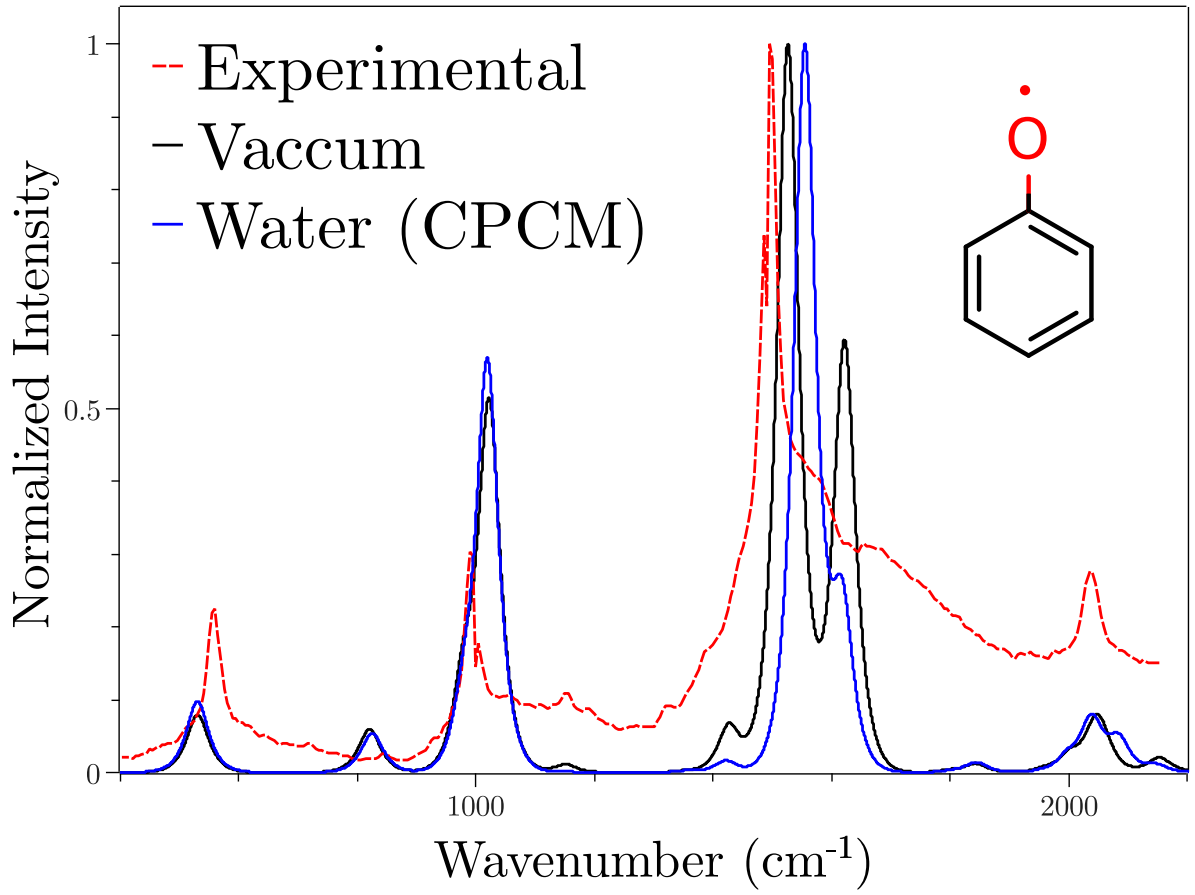


Fig. 6.67: The theoretical (solid black - vacuum and solid blue - water) and experimental (dashed red - water) resonant Raman spectrum for the phenoxyl radical.

OBS.: The actual Raman Intensity collected with any polarization at 90 degrees, the I($\pi/2; \|^s + \perp^s, \perp^i$) [533], can be obtained by setting RRINTES to TRUE under %ESD.

And the result is shown in Fig. 6.67. In this case, the default method VG was used. If one wants to include solvent effects, then CPCM(WATER) should be added. As can be seen, there is a noticeable difference in the main peak when calculated in water.

It is important to clarify some differences from the `ORCA_ASA` usage here. Using the ESD module, you do not

need to select which modes you will account for in the spectra; we include all of them. Additionally, we can only operate at 0 K, and the maximum "Raman Order" is 2. This means we account for all fundamental transitions, first overtones, and combination bands, without including hot bands. This level of approximation is generally sufficient for most applications.

If you are working with a very large system and want to reduce calculation time, you can request RORDER 1 under the %ESD options. This setting includes only the fundamental transitions, omitting higher-order bands. This approach may be relevant, especially when including both Duschinsky rotations and the Herzberg-Teller effect, which can significantly increase computation time.

The rRaman spectra are printed with the contributions from "Raman Order" 1 and 2 separated as follows:

```
Energy                    TotalSpectrum              Intensity01              Intensity02
0.000000              2.722264e-08              2.722264e-08              8.436299e-30
0.305176              2.824807e-08              2.824807e-08              9.043525e-30
0.610352              2.931074e-08              2.931074e-08              9.693968e-30
...
```

Isotopic Labeling

If you want to simulate the effect of isotopic labeling on the rRaman spectrum, there is no need to recalculate the Hessian again. Instead, you can directly modify the masses of the respective atoms in the Hessian files. This can be done by editing the $\$atoms$ section of the input file or directly in the Hessian file itself (see also Sec. *Isotope Shifts*). After making these adjustments, you can rerun ESD using the modified Hessian files, for example:

```
!PBE0 DEF2-SVP TIGHTSCF ESD(RR) CPCM(WATER)
%TDDFT
  NROOTS     5
  IROOT      3
END
%ESD
  GSHESSIAN  "PHE_WATER_ISO.hess"
  ESHESSIAN  "PHE_WATER_ISO.ES.hess"
END
* XYZFILE 0 2 PHE_WATER.xyz
```

As depicted in Fig. 6.68, the distinction between phenoxyl and its deuterated counterpart is evident. The peak around 1000 cm^{-1} corresponds to a C-H bond, which shifts to lower energy after deuteration. This difference of approximately 150 cm^{-1} aligns closely with experimental findings [851].

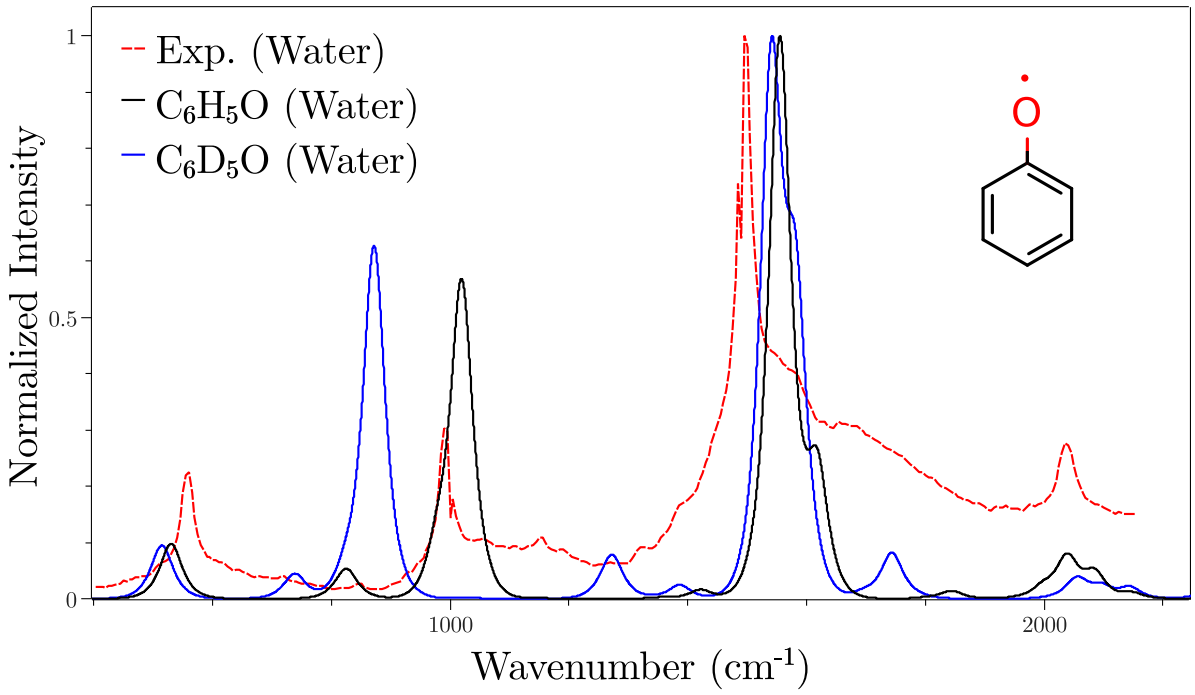


Fig. 6.68: The theoretical (solid black - C_6H_5O and solid blue - C_6D_5O) and experimental (dashed red) resonant Raman spectrum for the phenoxyl radical.

OBS: Whenever an ES Hessian is calculated using the HESSFLAG methods, it is saved in a file named BASENAME.ES.hess. If you want to repeat a calculation, you can simply use this file as an input without the need to recalculate everything.

RRaman and Linewidths

The keywords LINEW and INLINEW control the LINES function used in the calculation of the correlation function and are related to the lifetime of intermediate states and energy disordering. However, they do not determine the spectral linewidth itself, but rather the lineshape. The spectral linewidth is set independently using the RRSLINEW keyword, which defaults to 10 cm^{-1}.

It's important to note that LINEW and INLINEW significantly influence the final shape of the spectrum and should be chosen appropriately based on your specific needs. While the defaults are generally suitable, you may need to adjust them according to your requirements.

6.18.7 ESD and STEOM-CCSD or other higher level methods - the APPROXADEN option

If you plan to use the ESD module together with STEOM-CCSD, or other higher level methods such as EOM-CCSD, CASSCF/NEVPT2, some special advice must be given.

Since these methods currently do not have analytic gradients, numerical ones will be requested by default to compute the excited state geometries. This, of course, can take a significant amount of time, as they require approximately $3 \times N_{\text{atoms}}$ single-point calculations. We strongly recommend that, in these cases, you should use DFT/TD-DFT to obtain the ground/excited/triplet state geometry and Hessians, and only use the higher-level method for the final ESD step.

Also, we recommend using APPROXADEN under the %ESD options.

```
%ESD
    APPROXADEN TRUE
END
```

In this case, only one single point at the geometry of the ground state needs to be done, and the adiabatic energy difference will be automatically obtained from the ES Hessian information, without the need of a second single point at the extrapolated ES geometry, which could be unstable.

6.18.8 Circularly Polarized Spectroscopies

General Aspects

When circularly polarized (CP) light interacts with a chiral chemical structure (optically active), it differentially absorbs left and right CP lights ($I_{LCP} \neq I_{RCP}$) resulting in the electronic circular dichroism (ECD). Similarly, it can differentially emit left and right CP lights leading to CP luminescence (CPL), which includes CP fluorescence (CPF) and phosphorescence (CPP) spectra.

Vibration effects on ECD spectra

Vertically excited (VE) computed ECD spectra are known to often be unable to describe the experiment. This is for example the case in (R) methyl oxirane. The unshifted or shifted VE ECD BP86 computed spectra do not much the experiment in terms of shape and intensity. It has been shown that these spectra need to be computed by taking into account vibronic interactions[396].

Hence following structure optimization and frequencies calculations according to the input:

```
!BP86 DEF2-SVP TIGHTSCF OPT FREQ

* xyz 0 1
C   0.02461655377138      0.08670067686058     -5.20436273663217
C   -0.23485307714882     -0.31738971302751     -3.80610272711970
O   -0.15359212444282      1.06795113221760     -4.17749263689755
H   1.05055243293426     -0.00333310016875     -5.61325012342071
H   -0.78323750369168      0.02252140387747     -5.95969728166753
H   -1.26417590138099     -0.65347889194363     -3.56065466091728
C   0.84884023150886     -0.84537627906508     -2.89604274193698
H   0.68194744925825     -0.50794125098111     -1.85197471265376
H   0.85716491819315     -1.95559179229870     -2.89420656551675
H   1.84600617099845     -0.48435690147087     -3.21751813823758
*
```

we can compute ESD spectrum within in VE approximation and within the ESD modules according to the following input:

```
!BP86 DEF2-SVP TIGHTSCF PAL4

%TDDFT
NROOTS 10
END

%ESD
ESDFlag ECD
GSHESSIAN "c3h6o_opt_freq.hess "
PRINTLEVEL 2
DOHT TRUE
LINEW 500
SPECRANGE 40000, 70000
STATES 1,2,3,4,5,6,7,8,9,10
END

* xyzfile 0 1 c3h6o_opt_freq.xyz
```

The result is provided in Figure Fig. 6.69 where one can see that according to the expectations the computed spectrum agrees with the experiment only when FC and HT vibronic coupling schemes are taken into account

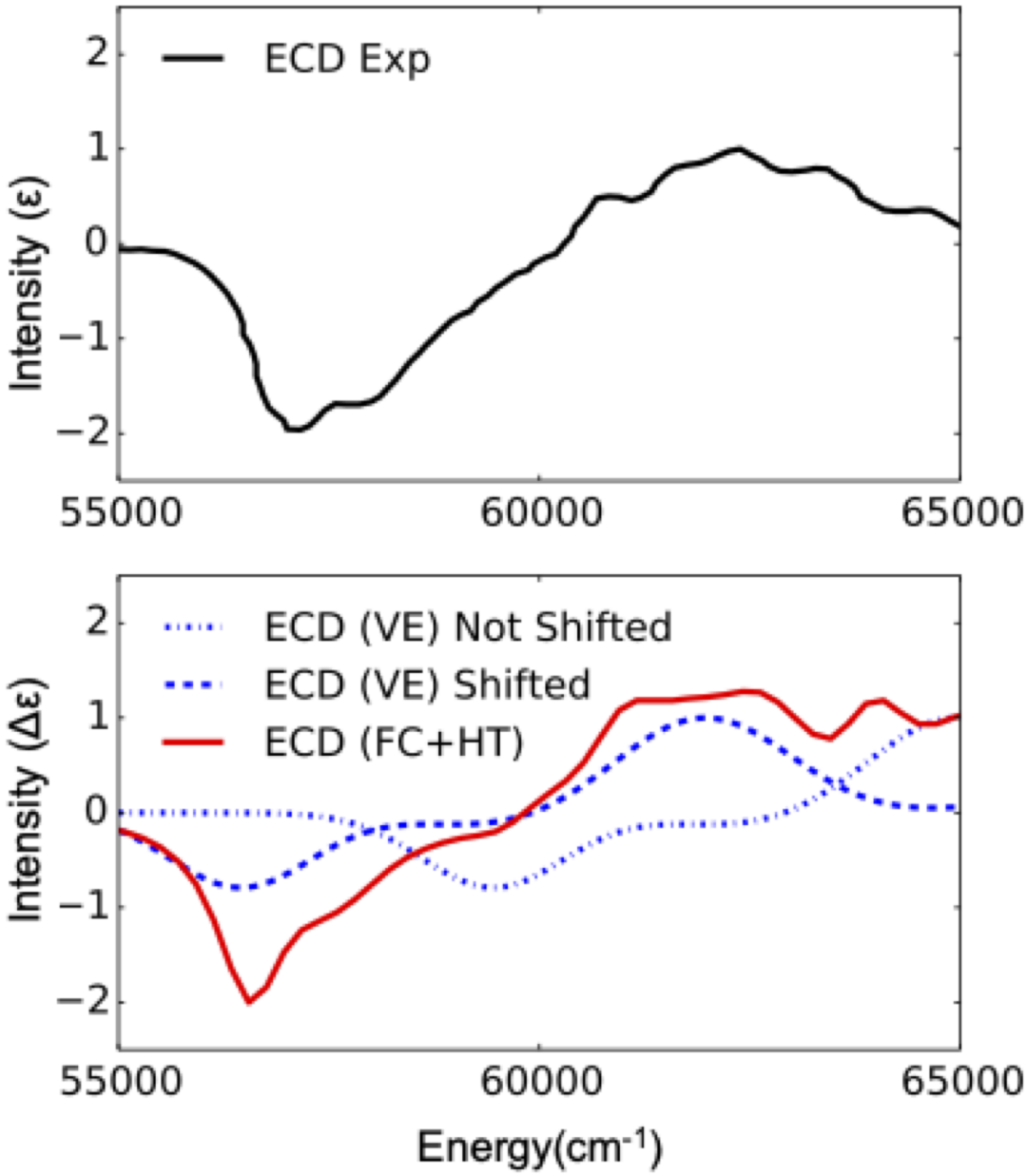


Fig. 6.69: Experimental (black) versus BP86/TDDFT (VE, blue and FC+HT red) ECD spectra for C3H6O molecule

Computation of CP-FLUOR vs CP-PHOS spectra. The case of C_3H_6O.

Following the strategy described for the computation of PL (Fluorescence (PF) of Phosphorescence (PP)) spectra in the case of C_3H_6O one can also access the respective CPL and CPP spectra.

For this one needs to compute the hessian of the 1st excited singlet (ES1) and triplet states respectively (ET0) according to the following inputs

```
!BP86 DEF2-SVP TIGHTSCF OPT FREQ

%TDDFT
NROOTS 10
IROOT 1
```

(continues on next page)

(continued from previous page)

```
IMULT 1
END

* xyzfile 0 1 c3h6o_opt_freq.xyz
```

and

```
!BP86 DEF2-SVP TIGHTSCF OPT FREQ

%TDDFT
NROOTS 10
IROOT 1
IMULT 3
SOCGRAD TRUE
TRIPLETS TRUE
END

* xyzfile 0 3 c3h6o_opt_freq.xyz
```

Then one can setup the respective PL and CPL inputs as:

PF:

```
!BP86 DEF2-SVP TIGHTSCF

%TDDFT
NROOTS 10
IROOT 1
IMULT 1
END

%ESD
ESDFlag FLUOR
GSHESSIAN "C3H6O_opt.hess"
PRINTLEVEL 2
DOHT TRUE
LINEW 500
SPECRANGE 40000, 70000
END

* xyzfile 0 1 c3h6o_opt_freq.xyz
```

CPF:

```
!BP86 DEF2-SVP TIGHTSCF

%TDDFT
NROOTS 10
IROOT 1
IMULT 1
END

%ESD
ESDFlag CPF
GSHESSIAN "C3H6O_opt.hess"
PRINTLEVEL 2
DOHT TRUE
LINEW 500
SPECRANGE 40000, 70000
END

* xyzfile 0 1 c3h6o_opt_freq.xyz
```

PP:

```
!BP86 DEF2-SVP TIGHTSCF

%TDDFT
NROOTS 10
IROOT 1
IMULT 3
DoSOC true
TRIPLETS TRUE
SOCGRAD TRUE
END

%ESD
ESDFlag PHOSP
GSHESSIAN "C3H6O_opt_freq.hess"
TSHESSIAN "C3H6O_et0.hess"
PRINTLEVEL 2
DELE 62313
DOHT TRUE
LINEW 500
TEMP  295
SPECRANGE 40000, 70000
END

* xyzfile 0 1 C3H6O_opt.xyz
```

CPP:

```
!BP86 DEF2-SVP TIGHTSCF

%TDDFT
NROOTS 10
IROOT 1
IMULT 3
DoSOC true
TRIPLETS TRUE
SOCGRAD TRUE
END

%ESD
ESDFlag CPP
GSHESSIAN "C3H6O_opt_freq.hess"
TSHESSIAN "C3H6O_et0.hess"
PRINTLEVEL 2
DELE 62313
DOHT TRUE
LINEW 500
TEMP  295
SPECRANGE 40000, 70000
END

* xyzfile 0 1 C3H6O_opt.xyz
```

The results are summarized in Figure Fig. 6.70

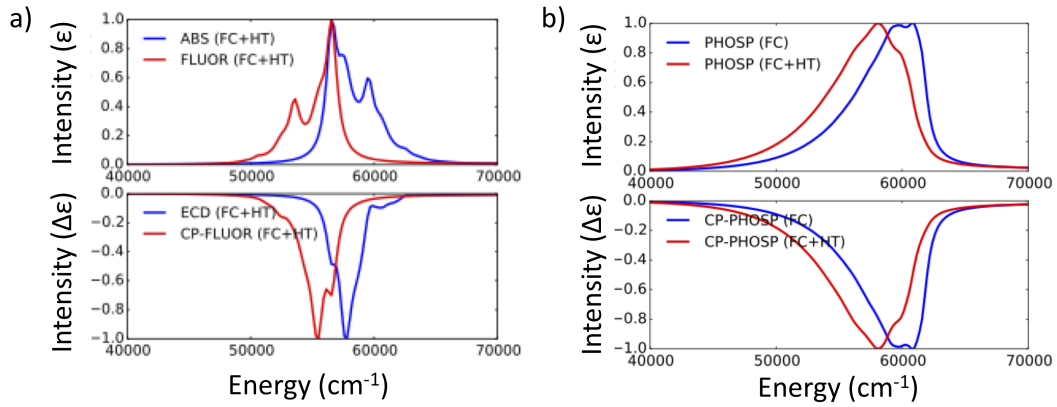


Fig. 6.70: a) Computed ABS and ECD (in blue) and Florescence and CPF (in red) under FC+HT vibronic coupling schemes b) Computed Phosphorescence and CPP under FC (in blue) and FC+HT (in red) vibronic coupling schemes

Use of ABS, ECD PL and CPL as a routine analysis computational tools

Having at hand the possibility to compute the above spectroscopic properties quartet. Consisting of Absorption, ECD, Luminescent/Emission and CPL spectroscopies creates an arsenal of useful analysis computational tools. Let us consider a practical example from the
the N- bridged triarylamine heterohelicenoid chiral family of molecules, which are known to be very good CPL emitters in the CPL community. [186] Namely the R-, L- isomers of oxygen-bridged diphenylnaphthylamine for which both ABS ECD, PL and CPL experimental spectra are available [634]

In a first step one needs to compute to calculate the ECD and CPL spectra, this implies that one needs to optimize the ground state (GS) geometry and at least the GS hessian of both isomers, (see examples in *Fluorescence Rates and Spectrum* and *Vibration effects on ECD spectra*). Lets suppose that we have generated the GSHessian file R_OptFreq.hess file for the R-isomer Then one can employ the ESD to calculate the Absorption, ECD, Fluorescence and CPL (CPluorescecne) as follows.

For Absorption or ECD spectra a representative input is given by:

```
! PBE0 def2-TZVP def2/J def2-TZVP VeryTightSCF PAL8
%MaxCore 5120

%TDDFT NROOTS 5
       End

%ESD   ESDFlag     ABS or ECD
       GSHessian  "R_diphenylnaphthylamine_OptFreq.hess"
       DoHT        True
       Lines       Gauss
       InLineW     500
       STATES      1,2,3,4,5
       End

*xyz 0 1
C           0.51103659781880     -1.54799809918165     -0.46957711710367
```

(continues on next page)

(continued from previous page)

```
C          -0.67682083480241     -2.19898547218227     -0.76456508506476
C           1.59698089595014     -2.23913762447750      0.04536520831045
C           1.49004594789848     -3.60107534512753      0.24696861333386
C          -0.75027591609003     -3.56343872947031     -0.57021086417552
C           0.33765559244610     -4.27346122451739     -0.10341675327345
N           0.21984508391204     -5.66346290617628      0.07289766826829
O           2.52716581975956     -4.28847037580032      0.82258958688483
C           2.54101738078760     -5.65282427633174      0.65190453446035
C           1.43184892145219     -6.36105379222857      0.25831057061542
C           3.77147940706722     -6.27965254373750      0.89001683729049
C           3.89719111077242     -7.62085864051752      0.70851807608572
C           1.58029288205151     -7.74060974139699     -0.06585842329724
C           2.81811727341129     -8.37881977445777      0.20527726022176
O          -1.91500963727701     -4.25152600059560     -0.80702339426568
C          -2.13250715287481     -5.31470429649343      0.04909464399864
C          -1.05242665026610     -6.06056594251354      0.52598562255970
C          -3.42444532048472     -5.61905191519688      0.41037068048411
C          -3.66399173496155     -6.68754268865431      1.26210310826770
C          -1.29901325564529     -7.09753540866930      1.41085143603056
C          -2.60176522746573     -7.41491597190339      1.76502960065907
H          -4.22973438581776     -5.01558835826860      0.01489276124005
H          -4.67873778203345     -6.93777656190575      1.53843511994436
H          -0.47290745868829     -7.66765085401275      1.80929822862278
H          -2.77926863803574     -8.23774715258380      2.44386581691750
H           0.58624233693590     -0.48095844424254     -0.62562692246504
H           2.51652973024519     -1.73662559987765      0.30924801794967
H          -1.54272285055189     -1.66670869597130     -1.13102495882678
H           4.59563142014220     -5.66684912938550      1.22771972849030
H           4.83828110533620     -8.11484947410606      0.91089239524435
C           0.56233992156145     -8.49575213625774     -0.69451663429121
C           2.96252715039542     -9.75415434575818     -0.08347675768224
C           0.74593486971180     -9.81550692161063     -0.98488391710801
C           1.95147468552282    -10.46257812809982     -0.65848843248314
H          -0.36538542100730     -8.01665224889233     -0.96504181763392
H          -0.04242675451776    -10.36879326589460     -1.47770458618347
H           2.07785937833070    -11.51320955617077     -0.88255640316097
H           3.90687057621082    -10.23141194413179      0.14665882143507
*
```

For Fluorescence or CP Fluorescence spectra a representative input is given by:

```
! PBE0 def2-TZVP def2/J def2-TZVP VeryTightSCF PAL8
%MaxCore 5120

%TDDFT NROOTS 5
       End

%ESD   ESDFlag     FLUOR or CPF
       GSHessian  "R_diphenylnaphthylamine_OptFreq.hess"
       DoHT        True
       Lines       Gauss
       InLineW     500
       STATES      1,2,3,4,5
       End

*xyz 0 1
C           0.51103659781880     -1.54799809918165     -0.46957711710367
C          -0.67682083480241     -2.19898547218227     -0.76456508506476
C           1.59698089595014     -2.23913762447750      0.04536520831045
C           1.49004594789848     -3.60107534512753      0.24696861333386
C          -0.75027591609003     -3.56343872947031     -0.57021086417552
```

(continues on next page)

(continued from previous page)

```
C           0.33765559244610     -4.27346122451739     -0.10341675327345
N           0.21984508391204     -5.66346290617628      0.07289766826829
O           2.52716581975956     -4.28847037580032      0.82258958688483
C           2.54101738078760     -5.65282427633174      0.65190453446035
C           1.43184892145219     -6.36105379222857      0.25831057061542
C           3.77147940706722     -6.27965254373750      0.89001683729049
C           3.89719111077242     -7.62085864051752      0.70851807608572
C           1.58029288205151     -7.74060974139699     -0.06585842329724
C           2.81811727341129     -8.37881977445777      0.20527726022176
O          -1.91500963727701     -4.25152600059560     -0.80702339426568
C          -2.13250715287481     -5.31470429649343      0.04909464399864
C          -1.05242665026610     -6.06056594251354      0.52598562255970
C          -3.42444532048472     -5.61905191519688      0.41037068048411
C          -3.66399173496155     -6.68754268865431      1.26210310826770
C          -1.29901325564529     -7.09753540866930      1.41085143603056
C          -2.60176522746573     -7.41491597190339      1.76502960065907
H          -4.22973438581776     -5.01558835826860      0.01489276124005
H          -4.67873778203345     -6.93777656190575      1.53843511994436
H          -0.47290745868829     -7.66765085401275      1.80929822862278
H          -2.77926863803574     -8.23774715258380      2.44386581691750
H           0.58624233693590     -0.48095844424254     -0.62562692246504
H           2.51652973024519     -1.73662559987765      0.30924801794967
H          -1.54272285055189     -1.66670869597130     -1.13102495882678
H           4.59563142014220     -5.66684912938550      1.22771972849030
H           4.83828110533620     -8.11484947410606      0.91089239524435
C           0.56233992156145     -8.49575213625774     -0.69451663429121
C           2.96252715039542     -9.75415434575818     -0.08347675768224
C           0.74593486971180     -9.81550692161063     -0.98488391710801
C           1.95147468552282    -10.46257812809982     -0.65848843248314
H          -0.36538542100730     -8.01665224889233     -0.96504181763392
H          -0.04242675451776    -10.36879326589460     -1.47770458618347
H           2.07785937833070    -11.51320955617077     -0.88255640316097
H           3.90687057621082    -10.23141194413179      0.14665882143507
*
```

The results are summarized in Figure Fig. 6.71

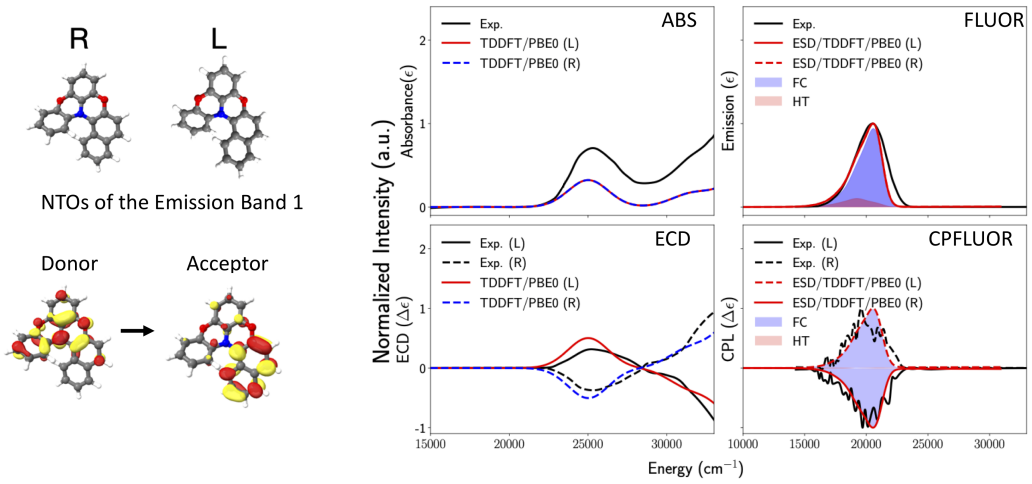


Fig. 6.71: Black Experimental vs Calculated ABS, ECD, Fluorescence and CPF spectra for R- (in blue) and L- (in red) isomers of diphenylnaphthylamine under FC+HT vibronic coupling schemes of the $\pi \rightarrow \pi^*$ transition located at 25000 cm^{-1}.

6.18.9 Magnetic Circular Dichroism (unpublished)

General Aspects

By applying a quasi-degenerate perturbative theory, similar to the inclusion of spin-orbit coupling effects in the phosphorescence calculations, the effect of an external magnetic field may be included in the representation of the quantum states [268]. As a result, the differential absorption of left and right circularly polarized light may be computed to obtain the vibrationally-corrected magnetic circularly dichroism spectrum. The input for the calculation is similar to the absorption case described above; nevertheless, ESD(MCD) should be used. Additionally, the intensity of the external magnetic field "B" (in Gauss) should be included, and a Lebedev grid for a semi-numerical molecular orientational average should be selected.

The method is only available with an electronic structure generated by TDDFT. The calculation supports the inclusion of Herzberg-Teller effects by setting DOHT TRUE, the ground state Hessian needs to be provided similarly to the absorption case, while the excited state Hessian can be provided or computed under a no external magnetic field approximation. An input example is:

```
!B3LYP DEF2-TZVP TIGHTSCF ESD(MCD)

%TDDFT NROOTS 40
       TDA FALSE
END

%ESD    GSHESSIAN "pbq.hess"
        Hessflag AHAS
        DOHT    TRUE
        STATES 1
        B 50000.0
END

* xyzfile 0 1 pbq.xyz
```

Similarly, to the ESB(ABS) calculation the MCD spectrum is saved in a BASENAME.MCD file as:

```
Energy          TotalSpectrum           IntensityFC             IntensityHT
4817.11         -6.324671e-05           -4.528264e-06           -5.871844e-05
5026.55         -6.717718e-05           -4.809014e-06           -6.236816e-05
5235.99         -7.126386e-05           -5.100843e-06           -6.616302e-05
5445.43         -7.551756e-05           -5.404513e-06           -7.011304e-05
5654.87         -7.994996e-05           -5.720850e-06           -7.422911e-05
5864.31         -8.457379e-05           -6.050750e-06           -7.852304e-05
...
```

6.18.10 Tips, Tricks and Troubleshooting

- Currently, the ESD module works optimally with TD-DFT (Sec. *Excited States Calculations*), but also with ROCIS (Sec. *Excited States with RPA, CIS, CIS(D), ROCIS and TD-DFT*), EOM/STEOM (Sec. *Excited States with EOM-CCSD* and Sec. *Excited States with STEOM-CCSD*) and CASSCF/NEVPT2 (Sec. *Complete Active Space Self-Consistent Field Method* and Sec *N-Electron Valence State Perturbation Theory (NEVPT2)*). Of course you can use any two Hessian files and input a custom DELE and TDIP obtained from any method (see Sec. *More on the Excited State Dynamics module*), if your interested only in the FC part.
- If you request for the HT effect, calculating absorption or emission, you might encounter phase changes during the displacements during the numerical derivatives of the transition dipole moment. There is a phase correction for TD-DFT and CASSCF, but not for the other methods. Please be aware that phase changes might lead to errors.
- Please check the K*K value if you have trouble. When it is too large (in general larger than 7), a warning message is printed and it means that the geometries might be too displaced and the harmonic approximation might fail. You can try removing some modes using TCUTFREQ or use a different method for the ES PES.
- If using DFT, the choice of functional can make a big difference on the excited state geometry, even if it is small on the ground state. Hybrid functionals are much better choices than pure ones.
- In CASSCF/NEVPT2, the IROOT flag has a different meaning from all other modules. In this case, the ground state is the IROOT 1, the first excited state is IROOT 2 and so on. If your are using a state-averaged calculation with more than one multiplicity, you need also to set an IMULT to define the right block, IMULT 1 being the first block, IMULT 2 the second and etc.
- If using NEVPT2 the IROOT should be related to the respective CASSCF root, don't consider the energy ordering after the perturbation.
- After choosing any of the HESSFLAG options, a BASENAME.ES.hess file is saved with the geometry and Hessian for the ES. If derivatives with respect to the GS are calculated, a BASENAME.GS.hess is also saved. Use those to avoid recalculating everything over and over. If you just want to get an ES PES, you can set WRITEHESS TRUE under
- Although in principle more complete, the AH is not NECESSARILY better, for we rely on the harmonic approximation and large displacements between geometries might lead to errors. In some cases the VG, AHAS and so one might be better options.
- If you use these .hess files with derivatives over normal modes in one coordinate system, DO NOT MIX IT with a different set of coordinates later! They will not be converted.
- Sometimes, low frequencies have displacements that are just too large, or the experimental modes are too anharmonic and you might want to remove them. It is possible to do that setting the TCUTFREQ flag (in cm^{-1}), and all frequencies below the given threshold will be removed.
- If you want to change the parameters related to the frequency calculations, you can do that under %FREQ (Sec. *Vibrational Frequencies*). The numerical gradient settings are under %NUMGRAD (Sec. *Numerical Gradients*).
- When computing rates, the use of any LINES besides DELTA is an approximation. It is recommended to compute the rate at much smaller lineshape (such as 10 cm^{-1}) to get a better value, even if the spectrum needs a larger lineshape than that.

- When in doubt, try setting a higher PRINTLEVEL. some extra printing might help with your particular problem.

6.19 The ORCA DOCKER: An Automated Docking Algorithm

The most important aspects of chemistry/physics do not occur with single molecules, but when they interact with each other. Now, given any two molecules, how to put them together in the best interacting "pose"? That is what we try to answer when using the ORCA DOCKER. Docking here refers to the process of taking two systems and putting them together in their best possible interaction.

6.19.1 Example 1: A Simple Water Dimer

Let us start with a very simple example. Given two water molecules, how to find the optimal dimer? With the DOCKER that is simple and can be done with:

```
!XTB
%DOCKER GUEST "water.xyz" END
* xyz 0 1
O  -2.13487    2.63905   -0.01809
H  -1.16698    2.61938    0.02397
H  -2.41372    2.24598    0.82256
*
```

where the file `water.xyz` is a `.xyz` file which contains the same water structure, optionally with charge and multiplicity (in that order) on the comment line (the second line by default):

```
3
0 1
O  -2.13487    2.63905   -0.01809
H  -1.16698    2.61938    0.02397
H  -2.41372    2.24598    0.82256
```

The molecule given on the regular ORCA input will be the `HOST`, and the `GUEST` is always given through an external file.

The output will start with:

```
                            ***************
                            * ORCA Docker *
                            ***************

Reading guests from file             water.xyz
Number of structures read from file  1
Charge and multiplicity of guest     from file
Docking approach                     independent
Docking level                        normal
Optimizing host                      .... -5.070544 Eh
Optimizing guest                     .... -5.070544 Eh
```

where it writes the name of the file with the `GUEST` structure, the number of structures read, some extra info and will optimize both host and guest (in this case they are the same), here by default using GFN2-XTB.

> **Note**
>
> If no multiplicity or charge are given, the `GUEST` is assumed to be neutral and closed-shell.

Note

The DOCKER right now is **only** working with the GFN-XTB and GFN-FF methods and the ALPB solvation model. It will be expanded later to others.

That is followed by some extra info that is explained in more details on its own detailed section (see *More details on the ORCA DOCKER*):

```
Starting Docker
---------------
Guest structure                      .... structure number 1
Guest charge and multiplicity        .... (0 , 1)
Final charge and multiplicity        .... (0 , 1)
PES used during evolution            .... GFN2-XTB
Setting random seed                  .... done
Creating spatial grid
   Grid Max Dimension                5.50 Angs
   Angular Grid Step                 32.73 degrees
   Cartesian Grid Step               0.50 Angs
   Points per Dimension              11 points
Initializing workers
   Population Density                0.50 worker/Ang^2
   Population Size                   57
Evolving structures
   Minimization Algorithm            mutant particle swarm
   Min, Max Iterations               (3 , 10)
```

That is followed by the docking itself, which will stop after a few iterations:

```
   Iter    Emin          avDE         stdDE           Time
           (Eh)       (kcal/mol)    (kcal/mol)       (min)
   ------------------------------------------------------

       1 -10.147462     2.756033     1.821981         0.03
       2 -10.147462     2.121389     1.610208         0.03
       3 -10.148583     2.313606     1.365227         0.03
       4 -10.148583     1.846998     1.188680         0.02
       5 -10.148583     1.587332     1.168207         0.02
No new minimum found after 3 (MinIter) steps.
```

The idea here is to collect as many local minima as possible, that is, collect a first guess for all possible modes of interaction between the different structures. We do this by allowing both structures to partially optimize, but it is important to say we will not look for multiple conformers of the host and guest here.

With all solutions collected, we will take a fraction of them and do a final full optimization:

```
Running final optimization
   Maximum number of structures      7
   Minimum energy difference         0.10 kcal/mol
   Maximum RMSD                      0.25 Angs
   Optimization strategy             regular
   Coordinate system                 redundant 2022
   Fixed host                        false

   Struc    Eopt       Interaction Energy       Time
            (Eh)               (kcal/mol)      (min)
   -------------------------------------------------

    1      -10.149006          -4.968378       0.01
    2      -10.149005          -4.967965       0.01
    3      -10.149007          -4.968825       0.01
```

(continues on next page)

(continued from previous page)

```
    4      -10.149007          -4.968641       0.01
    5      -10.149007          -4.968743       0.01
    6      -10.149006          -4.968116       0.01
    7      -10.149007          -4.968678       0.01
```

And as you can see, we also automatically print the `Interaction Energy`, which is simple an energy difference between the final complex, host and guest. The final best structure with lowest interaction energy is then saved on the `Basename.docker.xyz` file. If needed, all other structures are saved on the `Basename.docker.struc1.allopt.xyz`, as written on the output:

```
All optimized structures saved to  :      Basename.docker.struc1.allopt.xyz

-------------------------------------
LOWEST INTERACTION ENERGY:  -4.968825 kcal/mol (structure 3)
-------------------------------------

(...)

The lowest energy structure was 1, with energy -10.149007.
Docked structures saved to    Basename.docker.xyz
```

Note

The name `Basename.docker.struc1.allopt.xyz` refers to `struc1` because that is the first docked guest. Later that can be done with multiple guest and that is only a way to organize the outputs.

We are all set, the output can be visualized and it is, as expected:

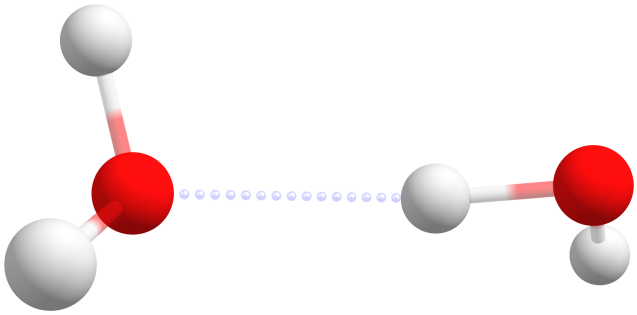


Fig. 6.72: The final water dimer found using the GFN2-XTB PES.

6.19.2 Example 2: A Uracil Dimer

Now for a slightly more complex example, a uracil dimer:

```
! XTB PAL16
%DOCKER GUEST "uracil.xyz" END
*xyz 0 1
N    -0.2707028    0.7632994    1.0276159
H    -0.5957915    1.3097757    1.8163465
C    -0.3386212    1.3810817   -0.2276640
O    -0.7270425    2.5346295   -0.3329857
```

(continues on next page)

(continued from previous page)

```
N     0.3638189   -1.2896563    0.1949192
H     0.0796815    0.9143946   -2.3190044
C     0.3781329   -0.7736192   -1.0714063
H     0.6499130   -1.4675080   -1.8526542
C     0.0669084    0.5154897   -1.3194961
H     0.4818502   -2.2779688    0.3498201
C    -0.0016589   -0.5616540    1.3117092
O    -0.0864879   -1.0482643    2.4227999
*
```

where the `uracil.xyz` is a simple repetition of the structure, as with the water before.

In this case the output is more diverse, and in fact many different poses appear as candidates for the final optimization:

```
   Struc    Eopt       Interaction Energy       Time
            (Eh)               (kcal/mol)      (min)
   ------------------------------------------------

    1      -49.248577         -11.723457       0.08
    2      -49.250442         -12.893758       0.08
    3      -49.245624          -9.870339       0.03
    4      -49.252991         -14.493130       0.06
    5      -49.248470         -11.656256       0.05
    6      -49.259335         -18.474228       0.05
    7      -49.259269         -18.432902       0.08
    8      -49.254913         -15.699019       0.03
    9      -49.254927         -15.708244       0.03
   10      -49.241672          -7.390198       0.02
   11      -49.246534         -10.441269       0.03
```

and structure number 6 is found to be the one with lowest interaction energy:

```
-------------------------------------
LOWEST INTERACTION ENERGY: -18.474228 kcal/mol (structure 6)
-------------------------------------
```

Here is a scheme with the structures found and their relative energies:

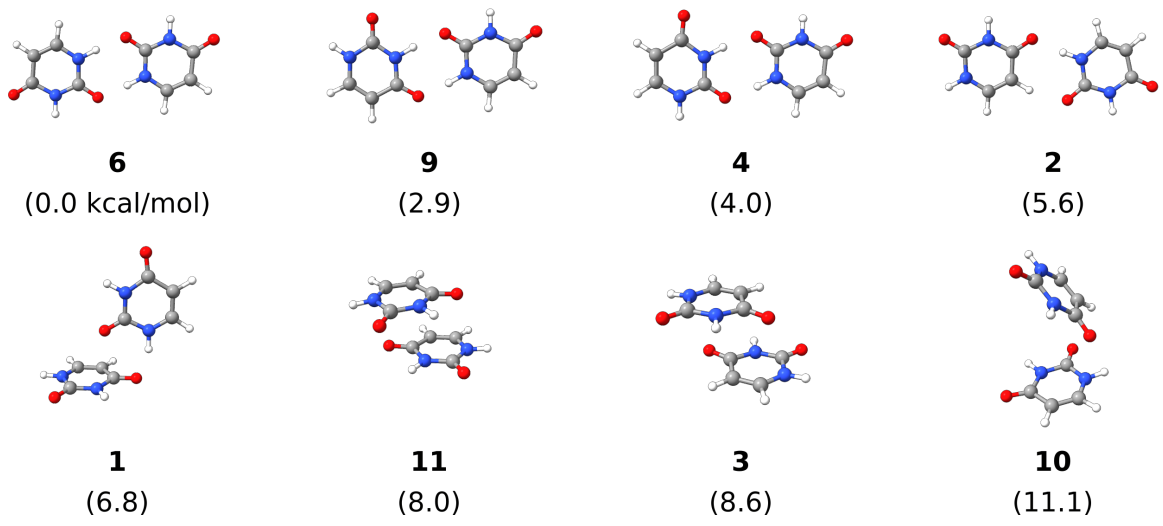


Fig. 6.73: Uracil dimer structures generated by DOCKER (duplicates removed) with relative energies in kcal/mol.

Note

There might be duplicated results after the final optimization, these are currently **not** automatically removed. Here they were manually removed.

Important

The PAL16 flag means XTB will run in parallel, but the ORCA DOCKER could be parallelized in a much more efficient way by paralleizing over the workers. That will be done for the next versions and it will be significantly more efficient.

6.19.3 Example 3: Adding Multiple Copies of a Guest

Suppose you want to add multiple copies of the same guest, for example three water molecules on top of the uracil one after the other. That can be simply done with:

```
! XTB PAL16
%DOCKER
    GUEST        "water.xyz"
    NREPEATGUEST 3
END
*xyz 0 1
N    -0.2707028    0.7632994    1.0276159
H    -0.5957915    1.3097757    1.8163465
C    -0.3386212    1.3810817   -0.2276640
O    -0.7270425    2.5346295   -0.3329857
N     0.3638189   -1.2896563    0.1949192
H     0.0796815    0.9143946   -2.3190044
C     0.3781329   -0.7736192   -1.0714063
H     0.6499130   -1.4675080   -1.8526542
C     0.0669084    0.5154897   -1.3194961
H     0.4818502   -2.2779688    0.3498201
C    -0.0016589   -0.5616540    1.3117092
O    -0.0864879   -1.0482643    2.4227999
*
```

and the guests on `water.xyz` will be added on top of the previous best complex three times. Now, there will be files with names `Basename.docker.struc1.allopt.xyz`, `Basename.docker.struc2.allopt.xyz` and `Basename.docker.struc3.allopt.xyz`, one for each step. Still, the same final `Basename.docker.xyz` file and now a `Basename.docker.build.xyz` is also printed, with the best result after each iteration.

That's how the results look like, from the `Basename.docker.xyz`:

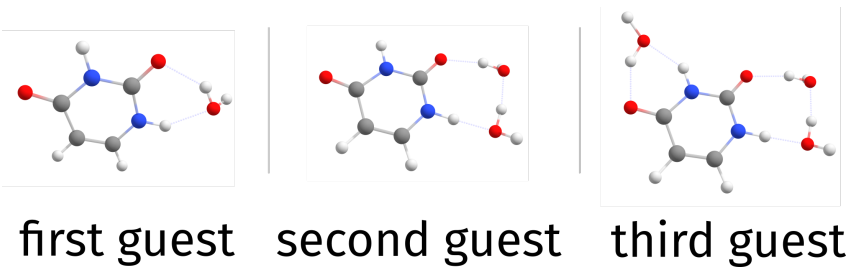


Fig. 6.74: Cumulative docking of three guests

Note

By default the HOST is always optimized. It can be changed with %DOCKER FIXHOST TRUE END.

6.19.4 Example 4: Find the Best Guest

Another common case would be: given a list of guests - or conformers of the same guest (see *GOAT: global geometry optimization and ensemble generator*) - one might want to know what is the "best guest", that is the one with the lowest interaction energy.

In order to do that, simply pass a multixyz file and the DOCKER will try to dock all structures from that file, one by one:

```
! XTB
%DOCKER GUEST "uracil_water.xyz" END
*xyz 0 1
N    -0.2707028    0.7632994    1.0276159
H    -0.5957915    1.3097757    1.8163465
C    -0.3386212    1.3810817   -0.2276640
O    -0.7270425    2.5346295   -0.3329857
N     0.3638189   -1.2896563    0.1949192
H     0.0796815    0.9143946   -2.3190044
C     0.3781329   -0.7736192   -1.0714063
H     0.6499130   -1.4675080   -1.8526542
C     0.0669084    0.5154897   -1.3194961
H     0.4818502   -2.2779688    0.3498201
C    -0.0016589   -0.5616540    1.3117092
O    -0.0864879   -1.0482643    2.4227999
*
```

Here the file `uracil_water.xyz` looks like:

```
3
0 1
O  -2.13487    2.63905   -0.01809
H  -1.16698    2.61938    0.02397
H  -2.41372    2.24598    0.82256
12
0 1
N    -0.2707028    0.7632994    1.0276159
H    -0.5957915    1.3097757    1.8163465
C    -0.3386212    1.3810817   -0.2276640
O    -0.7270425    2.5346295   -0.3329857
N     0.3638189   -1.2896563    0.1949192
H     0.0796815    0.9143946   -2.3190044
C     0.3781329   -0.7736192   -1.0714063
H     0.6499130   -1.4675080   -1.8526542
C     0.0669084    0.5154897   -1.3194961
H     0.4818502   -2.2779688    0.3498201
C    -0.0016589   -0.5616540    1.3117092
O    -0.0864879   -1.0482643    2.4227999
```

with a water followed by an uracil molecule. First, the water will be added, then the uracil, but both separately. The initial output is a bit different:

```
                            ***************
                            * ORCA Docker *
                            ***************

Reading guests from file            uracil_water.xyz
```

(continues on next page)

(continued from previous page)

```
Number of structures read from file  2
Charge and multiplicity of guests    from file
Docking approach                     independent
Docking level                        normal
```

with now two structures being read from file, and the `Docking approach` is labeled as `independent`, meaning each structure will be docked independently of each other.

After everything, the output is:

```
------------------------------------
LOWEST INTERACTION ENERGY: -18.482854 kcal/mol (structure 6)
------------------------------------

Total time for docking:       4.84 minutes

The lowest energy structure was 2, with energy -49.259349.
Docked structures saved to    Basename.docker.xyz
```

and one can see that the lowest interaction energy was that of structure 2 (the uracil), meaning it interacts strongly with the `HOST` than the water molecule given. Now the file `Basename.docker.xyz` will contain all final structures, ordered by interaction energy.

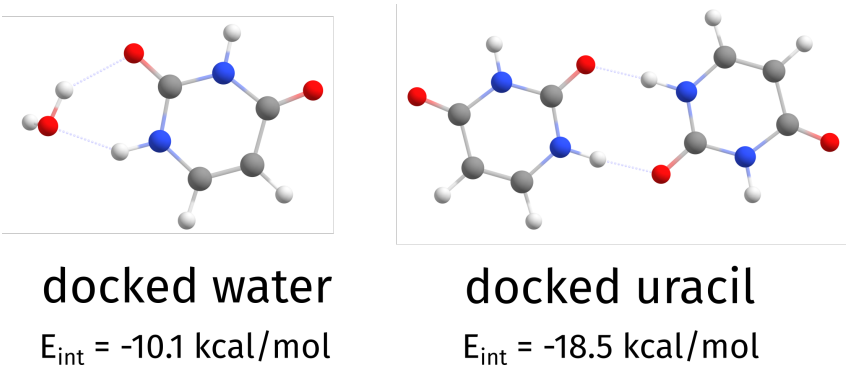


Fig. 6.75: Independent docking of water and uracil on top of an uracil molecule

Note

By default, the docking approach uses a fixed random seed and should always give the same result on the same machine. To make it always completely random add `%DOCKER RANDOMSEED TRUE END`.

Note

In order to use the faster GFN-FF instead of GFN2-XTB, use `!DOCK(GFNFF)`. For a quicker (and less accurate) docking, use `!QUICKDOCK`.

Note

To try multiple conformers of the `GUEST`, the ensemble file printed by GOAT `Basename.finalensemble.xyz` can be directly given here and the whole ensemble will be tested against a give `HOST`.

A detailed description of the other options can be found on *More details on the ORCA DOCKER*

6.19.5 Reduced Keyword List

```
!QUICKDOCK       # simple keyord to set DOCKLEVEL QUICK
!NORMALDOCK      # simple keyord to set DOCKLEVEL NORMAL
!COMPLETEDOCK    # simple keyord to set DOCKLEVEL COMPLETE

!DOCK(GFN-FF)    # simple keyord to set EVPES GFNFF
!DOCK(GFN0-XTB)  # simple keyord to set EVPES GFN0XTB
!DOCK(GFN1-XTB)  # simple keyord to set EVPES GFN1XTB
!DOCK(GFN2-XTB)  # simple keyord to set EVPES GFN2XTB

%DOCKER

   #
   # general options
   #

   GUEST           "filename.xyz" # an .xyz file (can be multistructure), from where
                                  # the guest(s) will be read. can contain different
                                  # charges and multiplicities for each guests on the
                                  # comment line. will only be read if exactly two
                                  # integer numbers are given, otherwise ignored.

   DOCKLEVEL       SCREENING  # defines a general strategy for docking.
                          # will alter things like that population density
                   NORMAL     # and final number of optimized structrures.
                   COMPLETE   # default is NORMAL.

   NREPEATGUEST    1          # number of times to repeat the content of the "GUEST" file
   CUMULATIVE      TRUE       # add the contents of the "GUEST" file one on
                              # top of each other?
                              # default is FALSE, meaning each will be done independently.

   FIXHOST         TRUE       # freeze coordinatef for the HOST during all steps?
                              # (default FALSE)

   #
   # evolution step
   #

   EVPES           GFNFF      # which PES to use **only** during the evolution step.
                   GFN0XTB    # can be different from the final optimization.
                   GFN1XTB
                   GFN2XTB
   #
   # final optimization
   #

   NOPT            10         # a fixed number of structures to be optimized
   NOOPT           FALSE      # do not optimize any structure at all? (default FALSE)
```

6.20 Compound Methods

Compound Methods is a form of sophisticated scripting language that can be used directly in the input of ORCA. Using '*Compound*' the user can combine various parts of a normal ORCA calculation to evaluate custom functions of his own. In order to explain its usage, we will use an example. For a more detailed description of this module the user is referred to section *Compound Methods*.

6.20.1 example

As a typical example we will use the constrained optimization describing the "umbrella effect" of NH_3. The script will perform a series of calculations and in the end it will print the potential of the movement plus it will identify the minima and the maximum. The corresponding compound script is the one shown below.

```
# ----------------------------------------------
# Umbrella coordinate mapping for NH3
# Author: Frank Neese
# ----------------------------------------------
variable JobName = "NH3-umbrella";
variable amin    = 50.0;
variable amax    = 130.0;
variable nsteps  = 21;
Variable energies[21];

Variable angle;
Variable JobStep;
Variable JobStep_m;
variable step;

Variable method = "BP86";
Variable basis  = "def2-SVP def2/J";

step  = 1.0*(amax-amin)/(nsteps-1);

# Loop over the number of steps
# -----------------------------
for iang from 0 to nsteps-1 do
  angle    = amin + iang*step;
  JobStep  = iang+1;
  JobStep_m= JobStep-1;
  if (iang>0) then
    Read_Geom(JobStep_m);
    New_step
      ! &{method} &{basis} TightSCF Opt
      %base "&{JobName}.step&{JobStep}"
      %geom constraints
        {A 1 0 2 &{angle} C}
        {A 1 0 3 &{angle} C}
        {A 1 0 4 &{angle} C}
        end
      end

    Step_End
  else
    New_step
      ! &{method} &{basis} TightSCF Opt
      %base "&{JobName}.step&{JobStep}"
      %geom constraints
        {A 1 0 2 &{angle} C}
        {A 1 0 3 &{angle} C}
        {A 1 0 4 &{angle} C}
```

(continues on next page)

(continued from previous page)

```
        end
      end

      * int 0 1
      N 0 0 0 0.0 0.0 0.0
      DA 1 0 0 2.0 0.0 0.0
      H 1 2 0 1.06 &{angle} 0.0
      H 1 2 3 1.06 &{angle} 120.0
      H 1 2 3 1.06 &{angle} 240.0
      *
    Step_End
  endif
   Read energies[iang] = SCF_ENERGY[jobStep];
   print(" index: %3d Angle %6.2lf Energy: %16.12lf Eh\n", iang, angle, energies[iang]);
EndFor

# Print a summary at the end of the calculation
# ---------------------------------------------
print("////////////////////////////////////////////////////////\n");
print("// POTENTIAL ENERGY RESULT\n");
print("////////////////////////////////////////////////////////\n");
variable minimum,maximum;
variable Em,E0,Ep;
variable i0,im,ip;
for iang from 0 to nsteps-1 do
  angle   = amin + 1.0*iang*step;
  JobStep = iang+1;
  minimum = 0;
  maximum = 0;
  i0      = iang;
  im      = iang-1;
  ip      = iang+1;
  E0      = energies[i0];
  Em      = E0;
  Ep      = E0;
  if (iang>0 and iang<nsteps-1) then
    Em = energies[im];
    Ep = energies[ip];
  endif
  if (E0<Em and E0<Ep) then minimum=1; endif
  if (E0>Em and E0>Ep) then maximum=1; endif
  if (minimum = 1 ) then
    print(" %3d  %6.2lf %16.12lf (-)\n",JobStep,angle, E0 );
  endif
  if (maximum = 1 ) then
    print(" %3d  %6.2lf %16.12lf (+)\n",JobStep,angle, E0 );
  endif
  if (minimum=0 and maximum=0) then
    print(" %3d  %6.2lf %16.12lf    \n",JobStep,angle, E0 );
  endif
endfor
print("////////////////////////////////////////////////////////\n");

End # end of compound block
```

Let's start with how somebody can execute this input. In order to run it, the easiest way is to save it in a normal text file, using the name "umbrella.cmp" and then use the following ORCA input file:

```
%Compound "umbrella.cmp"
```

nothing more is needed. ORCA will read the compound file and act appropriately.

A few notes about this ORCA input. First, there is no simple input line, (starting with *"!"*). A simple input is not required when one uses the *Compound* feature, but In case the user adds a simple input, all the information from the simple input will be passed to the actual compound jobs.

In addition, if one does not want to create a separate compound text file, it is perfecly possible in ORCA to use the compound feature as any other ORCA block. This means that after the *%Compound* directive, instead of giving the filename one can append the contents of the Compound file.

As we will see, inside the compound script file each compound job can contain all information of a normal ORCA input file. There are two very important exceptions here: The number of processors and the *MaxCore*. These information should be set in the initial ORCA input file and not in the actual compound files.

The Compound block has the same structure like all ORCA blocks. It starts with a *"%"* and ends with *"End"*, if the input is not read from a file. In case the compound directives are in a file, as in the example above, then simply the filename inside brackets is needed and no final *END*.

6.20.2 Defining variables

As we pointed out already, it is possible to either give all the information for the calculations and the manipulation of the data inside the Compound block or create a normal text file with all the details and let ORCA read it. The latter option has the advantage that one can use the same file with more than one geometries. In the previous example we refer ORCA to an external file. The file *"umbrella.cmp"*, that contains all necessary information.

Let's try to analyse now the Compound *"umbrella.cmp"* file.

```
# ----------------------------------------------
# Umbrella coordinate mapping for NH3
# Author: Frank Neese
# ----------------------------------------------
variable JobName = "NH3-umbrella";
variable amin    = 50.0;
variable amax    = 130.0;
variable nsteps  = 21;
Variable energies[21];

Variable angle;
Variable JobStep;
Variable JobStep_m;
variable step;

Variable method = "BP86";
Variable basis  = "def2-SVP def2/J";

step  = 1.0*(amax-amin)/(nsteps-1);
```

The first part contains some general comments and variable definitions. For the comments we use the same syntax as in the normal ORCA input, through the *"#"* symbol. Plase not that more than one *"#"* symbols in the same line cause an error.

After the initial comments we see some declarations and definitions. There are many different ways to declare variables described in detail in section *Variables - General*.

All variable declarations begin with the directive *Variable* which is a sign for the program to expect the declaration of one or more new variables. Then there are many options, including defining more than one variable, assigning also a value to the variable or using a list of values. Nevertheless all declarations **MUST** finish with the *;* symbol. This symbol is a message to the program that this is the end of the current command. The need of the *;* symbol in the end of each command is a general requirement in *Compound* and there are only very few exceptions to it.

6.20.3 Running calculations

```
# Loop over the number of steps
# ----------------------------
for iang from 0 to nsteps-1 do
  angle    = amin + iang*step;
  JobStep  = iang+1;
  JobStep_m= JobStep-1;
  if (iang>0) then
    Read_Geom(JobStep_m);
    New_step
      ! &{method} &{basis} TightSCF Opt
      %base "&{JobName}.step&{JobStep}"
      %geom
        constraints
          {A 1 0 2 &{angle} C}
          {A 1 0 3 &{angle} C}
          {A 1 0 4 &{angle} C}
        end
      end
    Step_End
  else
    New_step
      ! &{method} &{basis} TightSCF Opt
      %base "&{JobName}.step&{JobStep}"
      %geom
        constraints
          {A 1 0 2 &{angle} C}
          {A 1 0 3 &{angle} C}
          {A 1 0 4 &{angle} C}
        end
      end
      * int 0 1
        N 0 0 0 0.0 0.0 0.0
        DA 1 0 0 2.0 0.0 0.0
        H 1 2 0 1.06 &{angle} 0.0
        H 1 2 3 1.06 &{angle} 120.0
        H 1 2 3 1.06 &{angle} 240.0
      *
    Step_End
  endif
  Read energies[iang] = SCF_ENERGY[jobStep];
  print(" index: %3d Angle %6.2lf Energy: %16.12lf Eh\textbackslash{}n", iang, angle,␣
↪energies[iang]);
EndFor
```

Then we have the most information dense part. We start with the definition of a *for* loop. The syntax in compound for *for* loops is:

For *variable* **From** *startValue* **To** *endValue* **Do**
directives
EndFor

As we can see in the example above, the *startValue* and *endValue* can be constant numbers or previously defined variables, or even functions of these variables. Keep in mind that they have to be integers. The signal that the loop has reached it's end is the *EndFor* directive. For more details with regard to the *for* loops please refer to section *For*.

Then we proceed to assign some variables.

```
angle     = amin + iang*step;
JobStep   = iang+1;
JobStep_m = JobStep-1;
```

The syntax of the variable assignement is like in every programming language with a variable, followed with the = symbol and then the value or an equation. Please keep in mind, that the assignement **must** always finish with the *;* symbol.

The next step is another significant part of every programming language, namely the *if* block. The syntax of the *if* block is the following:

if (*expression to evaluate*) **Then**
directives
else if (*expression to evaluate*) **Then**
directives
else
directives
EndIf

The *else if* and *else* parts of the block are optional but the final *EndIf* must always signal the end of the *if* block. For more details concerning the usage of the *if* block please refer to section *If* of the manual.

Next we have a command which is specific for compound and not a part of a normal programming language. This is the *ReadGeom* command. It's syntax is:

Read_Geom(*integer value*);

Before explaining this command we will proceed with the next one in the compound script and return for this one.

The next command is the basis of all compound scripts. This is the *New_Step* Command. This command signals compound that a normal ORCA calculation follows. It's syntax is:

New_Command *Normal ORCA input* **Step_End**

Some comments about the *New_Step* command. Firstly, inside the *New_Step - Step_End* commands one can add all possilbe commands that a normal ORCA input accepts. We should remember here that the commands that define the number of processors and the *MaxCore* command will be ignored.

A second point to keep in mind is the idea of the *step*. Every *New_Step - Step_End* structure corresponds to a step, starting counting from 1 (The first ORCA calculation). This helps us define the property file that this calculation will create, so that we can use it to retrieve information from it.

A singificant feature in the *New_Step - Step_End* block. is the usage of the structure **&{*variable*}** . This structure allows the user to use variables that are defined outside the *New_Step - Step_End* block inside it, making the ORCA input more generic. For example, in the script given above, we build the *basename* of the calculations

```
%base "&{JobName}.step&{JobStep}"
```

using the defined variables *JobName* and *JobStep*. For more details regarding the usage of the **&{}** structure please refer to section *&* while for the *New_Step - Step_End* structure please refer to the section *New_Step*.

Finally, a few comments about the geometries of the calculation. There are 3 ways to provide a geometry to a *New_Step - Step_End* calculation. The first one is the traditional ORCA input way, where we can give the coordinates or the name of a file with coordinates, like we do in all ORCA inputs. In *Compound* though, if we do not pass any information concerning the geometry of the calculation, then *Compound* will automatically try to read the geometry of the previous calculation. This is the second (implicit) way to give a geometry to a compound Step. Then there is a third way and this is the one we used in the example above. This is the **Read_Geom** command. The syntat of this command is:
Read_Geom (*Step number*);
We can use this command when we want to pass a specific geometry to a calculation that is not explicitly given inside the *New_Step - Step_End* structure and it is also not the one from the previous step. Then we just pass the number of the step of the calculation we are interesting in just before we run our new calculation. For more details regarding the *Read_Geom* command please refer to section *Read_Geom*.

6.20.4 Data manipulation

One of the most powerfull features of *Compound* is it's direct access to properties of the calculation. In order to use these properties we defined the *Read* command. In the previous example we use it to read the SCF energy of the calculation:

```
Read energies[iang] = SCF\_ENERGY[jobStep];
```

The syntax of the command is:

Read *variable name* **=** *property*

where *variable name* is the name of a variable that is already defined, *property* is the property from the known properties found in table `TablePredefined` and *step* is the step of the calculation we are interested in. For more details in the *Read* command please refer to section *Read*.

We note here that *Compoud* has the ability to also read property files from older calculations. We can achieve this using the command *Get_From_Prop_File* that works exactly like the *Read* command but in older compound files. For more details concerning the *Get_From_Prop_File* please refer to section sec:compound.commands.get_from_prop_file.detailed.

```
# Print a summary at the end of the calculation
  # -----------------------------------------------
  print("////////////////////////////////////////////////////////\\n");
  print("// POTENTIAL ENERGY RESULT\\n");
  print("////////////////////////////////////////////////////////\\n");
  variable minimum,maximum;
  variable Em,E0,Ep;
  variable i0,im,ip;
  for iang from 0 to nsteps-1 do
    angle   = amin + 1.0*iang*step;
    JobStep = iang+1;
    minimum = 0;
    maximum = 0;
    i0      = iang;
    im      = iang-1;
    ip      = iang+1;
    E0      = energies[i0];
    Em      = E0;
    Ep      = E0;
    if (iang>0 and iang<nsteps-1) then
      Em = energies[im];
      Ep = energies[ip];
    endif
    if (E0<Em and E0<Ep) then minimum=1; endif
    if (E0>Em and E0>Ep) then maximum=1; endif
    if (minimum = 1 ) then
      print(" %3d  %6.2lf %16.12lf (-)\textbackslash{}n",JobStep,angle, E0 );
    endif
    if (maximum = 1 ) then
      print(" %3d  %6.2lf %16.12lf (+)\textbackslash{}n",JobStep,angle, E0 );
    endif
    if (minimum=0 and maximum=0) then
      print(" %3d  %6.2lf %16.12lf    \textbackslash{}n",JobStep,angle, E0 );
    endif
  endfor
  print("////////////////////////////////////////////////////////\\n");
```

Once all data are available we can use them in equations like in any programming language.

The syntax of the print statement is:

print(*format string, [variables]***);**

For example in the previous script we use it like:

```
print(" %3d  %6.2lf %16.12lf  \n",JobStep,angle, E0 );
```

where *%3d, %6.2lf* and *%16.2lf* are format identifiers and *JobStep, angle* and *E0* are previously defined variables. The syntax follows closely the widely accepted syntax of the *printf* command in the programming language C. For more details regarding the *print* statememnt please refer to section: *Print*.

Similar to the *print* command are the *write2file* and *write2string* commands that are used to write instead of the output file, either to a file we choose or to produce a new string.

Finally it is really importnat not to forget that every compound file should finish with a final **End**.

Once we run the previous example we get the following output:

```
////////////////////////////////////////////////////
// POTENTIAL ENERGY RESULT
////////////////////////////////////////////////////
1   50.00 -56.486626696200
2   54.00 -56.498074637200
3   58.00 -56.505200120800
4   62.00 -56.508823168800
5   66.00 -56.509732863600 (-)
6   70.00 -56.508724734300
7   74.00 -56.506590613800
8   78.00 -56.504070086000
9   82.00 -56.501791816800
10   86.00 -56.500229017900
11   90.00 -56.499674856600 (+)
12   94.00 -56.500229018100
13   98.00 -56.501791817200
14  102.00 -56.504070082800
15  106.00 -56.506590613300
16  110.00 -56.508724733100
17  114.00 -56.509732863700 (-)
18  118.00 -56.508823172900
19  122.00 -56.505200132200
20  126.00 -56.498074642900
21  130.00 -56.486626729200
////////////////////////////////////////////////////
```

with the step, the angle for the corresponding step, the energy of the constrained optimized energy plus the symbols for the two minima and the maximum in the potential.

CHAPTER

SEVEN

DETAILED DOCUMENTATION

7.1 The SHARK Integral Package and Task Driver

7.1.1 Preface

Starting with ORCA 5.0 very large changes have taken place in the way that the program handles integrals and integral related tasks like building Fock matrices. SHARK is a powerful and efficient infrastructure that greatly facilitates the handling of these tasks. This allows developers to write highly streamlined code with optimal performance and a high degree of reliability. Compared to the way ORCA handled integrals before ORCA 5.0, tens of thousands of lines of codes, often duplicated or nearly duplicated from closely related parts of the program could be eliminated. From the perspective of the user, the visible changes to the input and output of the program compared to ORCA 4.2.1 and earlier are relatively limited. However, under the hood, the changes are vast and massive and will ensure that ORCA's infrastructure is modern and very well suited for the future of scientific computing.

The benefits of SHARK for the users of ORCA are:

1. Improved code efficiency that is consistent through all program tasks. In particular, complicated two-electron integrals, for example in the context of GIAOs, two-electron spin-orbit coupling and two-electron spin-spin coupling integrals are handled with vastly improved efficiency. Also, integral digestion has been vastly improved with very large benefits for calculations that build many Fock matrices at a time, for example in CIS/TD-DFT, analytic Hessians or response property calculations.
2. Improved code reliability, since all integrals now run through a well debugged, common interface
3. Shorter development times. The new infrastructure is so user friendly to programmers that writing new code that makes use of SHARK is much faster than in the past.
4. SHARK handles basis sets much better than the old infrastructure. Whether the basis sets used follow a segmented contraction, general contraction or partial general contraction is immaterial since the algorithms have been optimized carefully for each kind of basis throughout.

7.1.2 The SHARK integral algorithm

One cornerstone of SHARK is a new integral algorithm that allows for highly efficient evaluation of molecular integrals. The algorithm is based on the beautiful McMurchie-Davidson algorithm which leads to the following equation for a given two-electron integral:

$$(\mu_A \nu_B|\kappa_C \tau_D) = C \sum_{tuv} E_t^{\mu\nu;x} E_u^{\mu\nu;y} E_v^{\mu\nu;z} \sum_{t'u'v'} E_{t'}^{\kappa\tau;x} E_{u'}^{\kappa\tau;y} E_{v'}^{\kappa\tau;z} (-1)^{t'+u'+v'} R_{t+t',u+u',v+v'}$$

Here

$$C = 8\pi^{5/2} = 139.947346620998902770103$$

and the primitive Cartesian Gaussian basis functions $\{\mu_A\}$ where A is the atomic center, where basis function μ is centered at position $\mathbf{R}_A$. In order to catch a glimpse of what the McMurchie-Davidson algorithm is about, consider

two unnormalized, primitive Gaussians centered at atoms A and B, respectively:

$$G_A = x_A^i y_A^j z_A^k \exp(-\alpha R_A^2)$$
$$G_B = x_B^{i'} y_B^{j'} z_B^{k'} \exp(-\beta R_B^2)$$

By means of the Gaussian product theorem, the two exponentials are straightforwardly rewritten as:

$$\exp(-\alpha R_A^2)\exp(-\beta R_B^2) = K_{AB}\exp\Big(-(\alpha+\beta)r_P^2\Big)$$

With

$$K_{AB} = \exp\left(-\tfrac{\alpha\beta}{\alpha+\beta}|\mathbf{R}_A - \mathbf{R}_B|^2\right)$$

$r_P^2 = |\mathbf{r} - \mathbf{R}_P|^2$ is the electronic position relative to the point

$$\mathbf{R}_P = \tfrac{\alpha}{\alpha+\beta}\mathbf{R}_A + \tfrac{\beta}{\alpha+\beta}\mathbf{R}_B$$

at which the new Gaussian is centered. The ingenious invention of McMurchie and Davidson was to realize that the complicated polynomial that arises from multiplying the two primitive Cartesian Gaussians can be nicely written in terms of Hermite polynomials $\{\Lambda\}$. In one dimension:

$$x_A^i x_B^{i'} = \sum_{t=0}^{i+i'} E_t$$

And hence:

$$G_A G_B = K_{AB}\sum_{t=0}^{i+i'} E_t^{AB}\sum_{u=0}^{j+j'} E_u^{AB}\sum_{v=0}^{k+k'} E_v^{AB}\Lambda_{tuv}^{AB}$$

With

$$\Lambda_{tuv}^{AB} = \left(\frac{\partial}{\partial X_P}\right)^t\left(\frac{\partial}{\partial Y_P}\right)^u\left(\frac{\partial}{\partial Z_P}\right)^v \exp\Big(-(\alpha+\beta)R_P^2\Big)$$

This means that the original four center integral is reduced to a sum of two-center integrals over Hermite Gaussian functions. These integrals are denoted as

$$R_{t+t',u+u',v+v'} = \int\int \Lambda_{tuv}^{AB}(\mathbf{r}_1;\mathbf{R}_P)\Lambda_{t'u'v'}^{CD}(\mathbf{r}_2;\mathbf{R}_Q) r_{12}^{-1} d\mathbf{r}_1 d\mathbf{r}_2$$

With these definitions one understands the McMurchie Davidson algorithm as consisting of three steps:

1. Transformation of the Bra function product into the Hermite Gaussian Basis
2. Transformation of the Ket function product into the Hermite Gaussian Basis
3. Calculation of the Hermite Gaussian electron repulsion integral

SHARK is the realization that these three steps can be efficiently executed by a triple matrix product:

$$(\mu_A\nu_B|\kappa_C\tau_D) = \left(\mathbf{E}^{\text{bra}}\mathbf{R}\mathbf{E}^{\text{ket}}\right)_{\mu\nu,\kappa\tau}$$

Here $\mathbf{E}^{\text{bra}}$ and $\mathbf{E}^{\text{ket}}$ collect the E coefficients for all members of the shell product on the bra and ket side ($E^{\text{bra}}_{\mu\nu,tuv}$ and $E^{\text{ket}}_{\kappa\tau,tuv}$), respectively, and $\mathbf{R}$ collects the integrals over Hermite Gaussian functions ($R_{tuv,t'u'v'}$).

There are many benefits to this formulation:

1. The integral is factorized allowing steps to be performed independent of each other. For example, the E matrices can be calculated at the beginning of the calculation and reused whenever needed. Their storage is unproblematic
2. Matrix multiplications lead to extremely efficient formation of the target integrals and drive the hardware at peak performance

3. Steps like contraction of primitive integrals and transformation from the Cartesian to the spherical Harmonics basis can be folded into the definition of the E matrices thus leading to extremely efficient code with next to no overhead creates by short loops.
4. Programming integrals becomes very easy and efficient. Other types of integrals as well as derivative integrals are readily approached in the same way. Also, two- and three-index repulsion integrals, as needed for the RI approximation are also readily formulated in this way.
5. One-electron integrals are equally readily done with this approach.

There is a very large number of technicalities that we will not describe in this manual which is only intended to provide the gist of the algorithm.

7.1.3 SHARK and libint

Up to ORCA 4.2.1, ORCA has almost entirely relied on the libint2 integral library which is known to be very efficient and powerful. Starting from ORCA 5.0, both SHARK and libint are used for integral evaluations and libint is fully integrated into the SHARK programming environment. Integrals that are only available in one of the packages are done with this package (e. g. GIAO, SOC and Spin-Spin integrals in SHARK; F12 or second derivative integrals in libint). For the integrals available in both packages, the program makes a judicious choice about the most efficient route. The reason for this hybrid approach is the following:

The SHARK integral algorithm is at its best for higher angular momentum functions ($l > 2$; d-functions) which is where the efficiency of the matrix multiplications leads to very large computational benefits. Integrals over, say, four f- or g-functions perform much faster (up to a factor of five) than with traditional integral algorithms. However, for low angular momenta, there is overhead created by the matrix multiplications and also by the fact that the McMurchie Davidson algorithm is known to not be the most FLOP count efficient algorithm. To some extent, this is take care of by using highly streamlined routines for low angular momenta that perform extremely well. However, there are penalties for intermediate angular momenta, where the efficiency of the matrix multiplications has not set in and the integrals are too complicated for hand coding. These integrals perform best with libint and consequently, the program will, by default, select libint to perform such integral batches.

7.1.4 Basis set types

One significant aspect of molecular integral evaluation is the type of contraction that is present in a Gaussian basis set. The most general type of basis set is met in the "general contraction" scheme. Here all primitive Gaussian basis functions of a given angular momentum are collected in a vector $\{\phi\}$. In general, all primitives will contribute to all basis functions $\{\varphi\}$ of this same angular momentum. Hence, we can write:

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{N_l} \end{pmatrix} = \begin{pmatrix} d_{11} & d_{11} & \cdots & d_{1M_l} \\ d_{21} & d_{21} & \cdots & d_{2M_l} \\ \vdots & \vdots & \ddots & \vdots \\ d_{N_l1} & d_{N_l2} & \cdots & d_{N_lM_l} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{M_l} \end{pmatrix}$$

Where N_l and M_l are the number of actual basis functions and primitives respectively. Typically, the number of primitives is much larger than the number of basis functions. The matrix $\mathbf{d}$ collects the contraction coefficients for each angular momentum. Typical basis sets that follow this contraction pattern are atomic natural orbital (ANO) basis sets. They are typically based on large primitive sets of Gaussians. Such basis sets put very demands on the integral package since there are many integrals over primitive Gaussian basis functions that need to be generated. If the integral package does not take advantage of the general contraction, then this integral evaluation will be highly redundant since identical integrals will be calculated N_l times (and hence, integrals over four generally contracted shells will be redundantly generated N_l^4 times). SHARK takes full advantage of general contraction for all one- and two-electron integrals that it can generate. Here, the unique advantages of the integral factorization come to full benefit since all integral quadruples of a given atom quadruple/angular momentum quadruple can be efficiently generated by just two large matrix multiplications.

The opposite of general contraction is met with segmented contraction. Here each basis function involves a number

of primitives:

$$\varphi_\mu = \sum_k d_{k\mu}\phi_k$$

Quite typically, none of the ϕ_k that occur in the contraction of one basis functions occurs in any other basis function. Typical basis sets of this form are the "def2" basis sets of the Karlsruhe group. They are readily handled by most integral packages and both SHARK and libint are efficient in this case.

The third class of basis sets is met, when general contraction is combined with segmented contraction. Basis sets of this type are, for example, the correlation consistent (cc) basis sets. We call such basis sets "partially generally contracted". In such basis sets, part of the basis functions are generally contracted (for example, the s- and p-functions in main group elements), while other basis functions (e. g. polarization functions, diffuse functions, core correlation functions) are not generally contracted. It is difficult to take full advantage of such basis sets given their complicated structure. In ORCA 5, special code has been provided that transforms the basis set into an intermediate basis set that does not contain any redundancies and hence drives SHARK or libint at peak performance.

In assessing the efficiency vs the accuracy of different integral algorithms, it is clear that segmented basis sets lead to the highest possible efficiency if they are well constructed. For such basis set the pre-screening that is an essential step of any integral direct algorithm performs best. The highest possible accuracy (per basis functions) is met with generally contracted basis sets. However, here the pre-screening becomes rather inefficient since it can only be performed at the level of atom/angular momentum combinations rather than individual shell quadruples. Thus, as soon as a given atom/angular momentum combination leads to any non-negligible integral, all integrals for this combination need to be calculated. This created a sizeable overhead. Consequently, SCF calculations can never be as efficient as with segmented basis sets. If this is immaterial, for example, because a subsequent coupled cluster or MRCI calculation is dominating the calculation time, general contraction is very worthwhile to be explored. For partial general contraction, our algorithm performs very nearly as efficiently as for segmented contraction in SCF calculations. However, since the intermediate basis set is larger than the original orbital basis, certain limited performance penalties can arise in some job types.

7.1.5 Task drivers

In traditional algorithms, quantum chemical programs frequently contain many instances of nested loops over basis function shells, the integral package is called and the integrals are "digested" for a given task. While these steps are inevitable, programming them repeatedly is laborious and error prone. In addition, improvements, say in the handling of contractions or symmetry, need to implemented in many different places. In the SHARK infrastructure all of this is unnecessary since it is programmed in an object-oriented fashion, where the programmer does not need to take care of any detail. Hence, developers only need to write short code sections that distribute the generated integrals into whatever data structure they need, while the SHARK interface takes care of all technical aspects and triggers the sophisticated and efficient machinery that underlies it.

Given this situation, the future of ORCA will involve SHARK taking care of nearly of the compute intensive, laborious tasks, while ORCA will organize and trigger all of these tasks. ORCA and SHARK communicate via a lean and well-defined interface to exchange the necessary data. In this way, a modern, efficient, easy to use and readily maintainable development environment is created.

7.1.6 SHARK User Interface

While SHARK is a large and complicated machinery, we have deliberately kept the interface as straightforward and simple as possible. There are only a few flags that can be set that are explained below:

In the simple input line there is:

```
! UseShark
! NoUseShark
```

This turns SHARK on (default) or off. Note that the option to turn SHARK off, will be unique to ORCA 5.0. Future versions of ORCA will always make use of SHARK and the legacy code will disappear from the program for good.

```
%shark
 UseGeneralContraction false  # turns general contraction algorithm on or
                              # off. There normally is no need to set this
                              # flag since the program will find the
                              # contraction case automatically
 Printlevel 1                 # Amount of output generated. Choose 0 to
                              # suppress output and 2 for more output.
                              # Everything else is debug level printing and
                              # will fill your harddrive very quickly with
                              # unusable information
 PartialGCFlag -1             # Let the program decide whether to use PGC
               0              # do not use it
               1              # Enforce PGC (even for ANO bases)
 FockFlag SHARK_libint_hybrid # default: best of both worlds
          force_shark         # Force Shark where possible
          force_libint        # Force libint where possible
 RIJFlag  RIJ_Auto            # default: program decides the best way
          Split_rij           # new SHARK Split-RI-J algorithm
          Split_rij_2003      # Highly efficient re-implementation of the
                              # Original 2003 algorithm. Mostly used!
          rij_regular         # Use traditional 3 center integrals
                              # (not recommended)
end
```

7.2 More on Coordinate Input

We will now enter the detailed discussion of the features of ORCA. Note that some examples are still written in the "old syntax" but there is no need for the user to adopt that old syntax. The new syntax works as well.

7.2.1 Fragment Specification

The atoms in the molecule can be assigned to certain *fragments*. This helps to organize the output in the population analysis section, is used for the fragment optimization feature, for the local energy decomposition and for multi-level calculations. There are two options to assign atoms to fragments. The first option is to assign a given atom to a given fragment by putting a `(n)` directly after the atomic symbol. Fragment enumeration starts with fragment 1!

```
%coords
 CTyp   xyz  # the type of coordinates xyz or internal
 Charge -2   # the total charge of the molecule
 Mult   2    # the multiplicity = 2S+1
 coords
    Cu(1)  0    0    0
    Cl(2)  2.25 0    0
    Cl(2) -2.25 0    0
    Cl(2)  0    2.25 0
    Cl(2)  0   -2.25 0
 end
end
```

In this example the fragment feature is used to divide the molecule into a "metal" and a "ligand" fragment and consequently the program will print the metal and ligand characters contained in each MO in the population analysis section.

Alternatively you can assign atoms to fragments in the geom block:

```
*xyz -2 2
    Cu  0    0    0
```

(continues on next page)

(continued from previous page)

```
    Cl  2.25 0    0
    Cl -2.25 0    0
    Cl  0    2.25 0
    Cl  0   -2.25 0
*
%geom
 Fragments
  1 {0} end    # atom 0 for fragment 1
  2 {1:4} end  # atoms 1 to 4 for fragment 2
 end
end
```

Note

- With the second option (geom-fragments) the %geom block has to be written after the coordinate section.
- geom-fragments also works with coordinates that are defined via an external file.
- For the geom-fragments option the atoms are assigned to fragment 1 if no assignment is given.

7.2.2 Defining Geometry Parameters and Scanning Potential Energy Surfaces

ORCA lets you define the coordinates of all atoms as functions of user defined geometry parameters. By giving not only a value but a range of values (or a list of values) to this parameters potential energy surfaces can be scanned. In this case the variable `RunTyp` is automatically changed to `Scan`. The format for the parameter specification is straightforward:

```
%coords
 CTyp   internal
 Charge 0
 Mult   1
 pardef
  rCH  = 1.09;    # a C-H distance
  ACOH = 120.0;   # a C-O-H angle
  rCO  = 1.35, 1.10, 26; # a C-O distance that will be scanned
 end
 coords
    C  0  0  0   0     0         0
    O  1  0  0  {rCO}  0         0
    H  1  2  0  {rCH}  {ACOH}    0
    H  1  2  3  {rCH}  {ACOH}  180
 end
end
```

In the example above the geometry of formaldehyde is defined in internal coordinates (the geometry functions work exactly the same way with Cartesian coordinates). Each geometric parameter can be assigned as a function of by enclosing an expression within function braces, "`{} `". For example, a function may look like `*cos(Theta)*rML+R`. Note that all trigonometric functions expect their arguments to be in degrees and not radians. The geometry parameters are expected to be defined such that the lengths come out in Ångströms and the angles in degrees. *After* evaluating the functions, the coordinates will be converted to atomic units. In the example above, the variable `rCO` was defined as a "Scan parameter". Its value will be changed in 26 steps from 1.3 Å down to 1.1 Å and at each point a single point calculation will be done. At the end of the run the program will summarize the total energy at each point. This information can then be copied into the spreadsheet of a graphics program and the potential energy surface can be plotted. Up to three parameters can be scan parameters. In this way grids or cubes of energy (or property) values as a function of geometry can be constructed.

If you want to define a parameter at a series of values rather than evenly spaced intervals, the following syntax is to be used:

```
%coords
 CTyp   internal
 Charge 0
 Mult   1
 pardef
  rCH = 1.09;   # a C-H distance
  ACOH= 120.0;  # a C-O-H angle
  rCO [1.3 1.25 1.22 1.20 1.18 1.15 1.10];  # a C-O distance that will be scanned
 end
 coords
    C  0  0  0   0     0         0
    O  1  0  0  {rCO}  0         0
    H  1  2  0  {rCH}  {ACOH}    0
    H  1  2  3  {rCH}  {ACOH}  180
 end
end
```

In this example the C-O distance is changed in seven non-equidistant steps. This can be used in order to provide more points close to a minimum or maximum and fewer points at less interesting parts of the surface.

A special feature has also been implemented into ORCA - the parameters themselves can be made functions of the other parameters as in the following (nonsense) example:

```
%coords
 CTyp   internal
 Charge 0
 Mult   1
 pardef
  rCOHalf= 0.6;
  rCO    = { 2.0*rCOHalf };
 end
 coords
    C  0  0  0   0     0         0
    O  1  0  0  {rCO}  0         0
    O  1  0  0  {rCO}  180       0
 end
end
```

In this example the parameter `rCO` is computed from the parameter `rCOHalf`. In general the geometry is computed (assuming a `Scan` calculation) by: (a) incrementing the value of the parameter to be scanned (b) evaluating the functions that assign values to parameters, and (c) evaluating functions that assign values to geometrical variables.

Although it is not mandatory, it is good practice to *first* define the static or scan-parameters and then define the parameters that are functions of these parameters.

Finally, ORCA has some special features that may help to reduce the computational effort for surface scans:

```
%method
 SwitchToSOSCF true      # switches the converger to SOSCF
                         # after the first point. SOSCF may
                         # converge better than DIIS if the
                         # starting orbitals are good.
                         # default = false
 ReducePrint true        # reduce printout after the first point
                         # default=true

 # The initial guess can be changed after the first point.
 # The default is MORead. The MOs of the previous point will,
 # in many cases, be a very good guess for the next point.
 # However, in some cases you may want to be more conservative
 # and use a general guess.
```

(continues on next page)

(continued from previous page)

```
ScanGuess   OneElec    # the one-electron matrix
            Hueckel    # the extended Hueckel guess
            PAtom      # the PAtom guess
            PModel     # the PModel guess
            MORead     # MOs of the previous point
end
```

Note

- You can scan along normal modes of a Hessian using the `NMScan` feature as described in section *Normal Mode Scan Calculations Between Different Structures*.
- The surface scan options are also supported in conjunction with TD-DFT/CIS or MR-CI calculations (see section *Potential Energy Surface Scans*).

7.2.3 Mixing internal and Cartesian coordinates

In some cases it may be practical to define some atomic positions in Cartesian and some in internal coordinates. This can be achieved by specifying all coordinates in the `*int` block: using "0 0 0" as reference atoms indicates Cartesian coordinates. Note that for the first atom the flags are "1 1 1", as "0 0 0" would be the normal values for internal coordinates. Consider, for example, the relaxed surface scan from section *Relaxed Surface Scans*, where the methyl group is given first in an arbitrary Cartesian reference frame and then the water molecule is specified in internal coordinates:

```
! UKS B3LYP SV(P) TightSCF Opt SlowConv
%geom scan B 4 0 = 2.0, 1.0, 15 end end
* int 0 2
# First atom - reference atoms 1,1,1 mean Cartesian coordinates
  C      1   1   1    -0.865590    1.240463   -2.026957

# Next atoms - reference atoms 0,0,0 mean Cartesian coordinates
  H      0   0   0    -1.141534    2.296757   -1.931942
  H      0   0   0    -1.135059    0.703085   -2.943344
  H      0   0   0    -0.607842    0.670110   -1.127819

# Actual internal coordinates
  H      1   2   3     1.999962   100.445    96.050
  O      5   1   2     0.984205   164.404    27.073
  H      6   5   1     0.972562   103.807    10.843
*
```

Internal and Cartesian coordinates can thus be mixed in any order but it is recommended that the first 3 atoms are specified in Cartesian coordinates in order to define a unique reference frame.

7.2.4 Inclusion of Point Charges

In some situations it is desirable to add point charges to the system. In ORCA there are two mechanisms to add point-charges. If you only want to add a few point charges you can "mask" them as atoms as in the following (nonsense) input:

```
# A water dimer
! BP86 def2-SVP

* xyz 0 1
O           1.4190    0.0000    0.0597
H           1.6119    0.0000   -0.8763
```

(continues on next page)

(continued from previous page)

```
H           0.4450    0.0000    0.0898
Q  -0.834  -1.3130    0.0000   -0.0310
Q   0.417  -1.8700    0.7570    0.1651
Q   0.417  -1.8700   -0.7570    0.1651
*
```

Here the "Q"'s define the atoms as point charges. The next four numbers are the magnitude of the point charge and its position. The program will then treat the point charges as atoms with no basis functions and nuclear charges equal to the "Q" values.

If you have thousands of point charges to treat, as in a QM/MM calculation, it is more convenient, and actually necessary, to read the point charges from an external file as in the following example:

```
# A water dimer
! BP86 def2-SVP

% pointcharges "pointcharges.pc"

* xyz 0 1
O    1.4190    0.0000    0.0597
H    1.6119    0.0000   -0.8763
H    0.4450    0.0000    0.0898
*
```

The program will now read the file "`pointcharges.pc`" that contains the point-charge information and then call the module `orca_pc` which adds the point charge contribution to the one-electron matrix and the nuclear repulsion. The file "`pointcharges.pc`" is a simple ASCII file in the following format:

```
3
 -0.834  -1.3130    0.0000   -0.0310
  0.417  -1.8700    0.7570    0.1651
  0.417  -1.8700   -0.7570    0.1651
```

The first line gives the number of point charges. Each consecutive line gives the magnitude of the point charge (in atomic units) and its position (in Ångström units!). However, it should be noted that ORCA treats point charges from an external file differently than "Q" atoms. When using an external point charge file, the interaction between the point charges is not included in the nuclear energy. This behavior originates from QM/MM, where the interactions among the point charges is done by the MM program. These programs typically use an external point charge file when generating the ORCA input. To add the interaction of the point charges to the nuclear energy, the `DoEQ` keyword is used either in the simple input or the `%method` block as shown below.

```
# A non QM/MM pointcharge calculation
! DoEQ

%pointcharges "pointcharges.pc"

%method
   DoEQ true
end
```

7.3 Details on the numerical integration grids

As in all other popular grid schemes, our grids are constructed from assembling a set of atomic grids into a molecular one, using Becke's approach. Each individual atomic grid is build based on optimized parameters for that atom, and are composed of an angular and a radial part, that are defined separately.

The whole scheme was updated from ORCA 5.0, but we tried to keep things as close as possible to the previous one. First, the overall construction of these grids needs to be explained.

7.3.1 The angular grid scheme

Instead of using a single angular grid throughout the whole atom, most schemes apply a so-called grid pruning in order to reduce the number of grid points outside of the most important regions, as we do in ORCA. In the current scheme, we split the atomic grids into five regions, using Lebedev grids with the following number of points on each of those:

Table 7.1: Different angular grid schemes used in ORCA. The numbers indicate the Lebedev grids used.

AngularGrid	Region 1	Region 2	Region 3	Region 4	Region 5
1	14	26	50	50	26
2	14	26	50	110	50
3	26	50	110	194	110
4	26	110	194	302	194
5	26	194	302	434	302
6	50	302	434	590	434
7	110	434	590	770	590

The ideal cutoffs between those regions were subjected to optimization, and are defined for all atoms. Whenever we refer to a AngularGrid flag in ORCA, one of these schemes is chosen.

7.3.2 The radial grid scheme

The number of radial points (n_r) for a given atom is simply defined using the equation first defined by Krack and Köster:

$$n_r = (15 \times \varepsilon - 40) + b \times ROW$$

where ε is called the IntAcc of that grid in ORCA, b is any number and ROW refers to the row of the periodic table for that atom. In its original formulation, b was set to 5, but here it as now optimized and varies slightly depending on the AngularGrid schemes shown above.

One important thing to note is that each increase of IntAcc by 1, adds 15 radial points to the atomic grids, as in the previous grid scheme. These IntAcc values were optimized for each angular grid:

Table 7.2: Optmimized IntAcc parameters for the exchange-correlation and COSX grids.

AngularGrid	XC	COSX
1	4.004	3.816
2	4.004	4.020
3	4.159	4.338
4	4.388	4.871
5	4.629	4.871
6	4.959	4.871
7	4.959	4.871

After defining the number of radial points n_r, the actual radial grid is then defined from a Gauss-Chebyschev quadrature using the so-called M3 mapping from Treutler and Ahlrich:

$$r = \frac{\xi}{ln2} ln \frac{2}{1-x}$$

where $-1 \leq x \leq 1$, and the center of the grid ($x = 0$) coincides with the value of ξ. These ξ parameters were also optimized for each atom type.

7.3.3 The DEFGRIDs

With all that in mind, we can now present how the DEFGRIDs are built in terms of their AngularGrid scheme and IntAccs, which define the angular and radial parts of the atomic grids. More details can be found in Ref. [383].

Table 7.3: Angular grid schemes used in different part of ORCA. The XC and COSX grids are separated by a slash, and multiple COSX grid schemes are separated by a comma.

Grid Name	SCF	TD-DFT	CP-SCF	MP2 grad	MP2 2ndder
DEFGRID1	3 / 1, 1, 2	1 / 1	1 / 1	3	1
DEFGRID2	4 / 1, 2, 3	1 / 1	1 / 1	4	4
DEFGRID3	6 / 2, 3, 4	3 / 2	3 / 2	5	4

OBS.: The IntAccs for TD-DFT and the CP-SCF are 3.467 for the XC and 3.067 for the COSX instead of the default. These numbers can be smaller here and we exploit this to increase the overall speed.

From the Table 7.3 one can see, for instance, that the default SCF XC grid now is defined from AngularGrid 4 (with no extra final integration in the end). The default COSX uses a 1,2,3 grid scheme, with the COSX third grid being used to update the energy after the SCF converges and for the gradients.

7.3.4 Other details and options

The new adaptive pruning. The current pruning scheme uses lower grids close to the nucleus, and far away from the bonding region. However, if the basis set has polarized functions close to the nuclei, or diffuse Gaussians, this might not be sufficient.

To improve the grids for these problems, we now use by default an adaptive pruning scheme, that detects core-polarization, diffuse functions and steep basis set orbitals by analyzing the expectation value of the position operator, $\langle \hat{r} \rangle$, and fixes the grid accordingly. This can increase the grids in these cases by 10-20%, but gives significantly better results. To use the non-adaptive scheme, just set %METHOD GRIDPRUNING OLDPRUNING END. For a completely unpruned grid, set GRIDPRUNING to UNPRUNED.

A simpler Gauss-Legendre angular grid. By setting AngularGrid to 0, instead of using the Lebedev grids, a Gauss-Legendre angular grid will be built, as suggested by Treutler and Ahlrich [850]. The number of θ points is defined as $0.4n_r$ by default and the number of ϕ points is chosen as to avoid crowding close to the poles.

These grids are in general less efficient than the Lebedev's, but are useful if one needs to construct extremely large grids for specific applications.

The SpecialGrid Option. Sometimes, you might want to increase the integration accuracy for some atoms that need special care, while it is not necessary to enlarge the grid generally. ORCA provides you with a basic mechanism to increase the radial integration accuracy for a few atoms while maintaining the chosen grid for all others.

```
%METHOD
   # a maximum of 64 assignments can be made
   #  in = 0 : no changes are made
   #  in > 0 : the grid will be changed for all atoms with
   #           atomic number=in to IntAcc=an
   #  in < 0 : only the specific n'th atom will have its
   #           IntAcc value changed to an
```

(continues on next page)

(continued from previous page)

```
   SpecialGridAtoms  i1, i2, i3,...,in;
   SpecialGridIntAcc a1,a2,a3,...,an;
END
```

OBS.: Starting from ORCA 5.0 it is not necessary to use this option anymore unless you have very specific reasons. The basis set is considered during the grid construction and that is automatically extended if needed.

7.3.5 SCF grid keyword list

A complete description of the SCF grid options is given below. There are keywords specific to the XC integration, COSX integration and a general part that applies to all:

```
%METHOD
# XC grids
AngularGrid    1    #Lebedev50
               2    #Lebedev110
               3    #Lebedev194
               4    #Lebedev302
               5    #Lebedev434
               6    #Lebedev590
               7    #Lebedev770
               0    #SimpleGrid
IntAcc         5.0  # determines no. of radial points

# COSX grids
AngularGridX   1,1,2,4,5    # the first three are used in the SCF
                            # the 4th in the MP2 gradient and
                            # the 5th for MP2 second derivatives
IntAccX        3.56,-1,4.5  # if a -1 is given, the default IntAcc is used.

# General
NThetaMax      0.4          # only for AngularGrid=0, multiplier for nr
GridPruning    Unpruned     # no Pruning
               OldPruning   # the old pruning
               Adaptive     # default (and recommended)
HGridReduced   true         # Reduce grids for H and He by one
                            # unit (default and recommended)
BFCut          1e-10        # basis fcn. cut. Is adjusted according to
                            # convergence tolerances
WeightCut      1e-14        # grid weight cut. default: 1e-14
END
```

7.3.6 Changing TD-DFT, CP-SCF and Hessian grids

TD-DFT. The grids used in CIS or TD-DFT can be changed in their respective block:

```
%TDDFT  # or %CIS, they are equivalent
  # XC grids
  IntAccXC 3.467
  GridXC   1

  #COSX grids
  IntAccX  3.076
  GridX    1
END
```

CP-SCF. The CP-SCF grids are changed in the %METHOD block:

```
%METHOD
  # XC grids
  Z_IntAccXC  3.467
  Z_GridXC    1

  #COSX grids
  Z_IntAccX   3.076
  Z_GridX     1
  Z_GridX_RHS 2
END
```

OBS.: The Z_Grid_RHS is only used in MP2 and the number here has a different meaning. It refers to which of the COSX grids used in the SCF will be chosen, rather than an AngularGrid scheme. The default is to use the second COSX grid.

Hessian. The XC grids used to compute the DFT terms in the Hessian are automatically chosen to be one unit higher than the SCF grids. Because of the second derivative terms, we found that it is better to have a slightly higher XC grid here. The COSX grid can be changed freely:

```
%FREQ
  HessGridX = 2,2,2,2
END
```

These four numbers refer to the possible usages of COSX in the Hessian, as explained in Sec. *Frequency calculations - numerical and analytical*

Non-local functionals (VV10 and alike). The default non-local grid is defined by AngularGrid 2, and is not recommended to be changed. In any case, these can be altered by using:

```
%METHOD
  # non-local grids
  VdwAngularGrid    2        # same scheme as the SCF ones
  VdWIntAcc         5.0      # determines no. of radial points
  VdwGridPruning    Adaptive # default
  VdwDistTCut       10       # cutoff distance between grid points, in angstroem
END
```

7.3.7 When should I change from the default grids?

In general, the errors from the default grids are rather small and reasonable for most applications. After benchmarking against the GMTNK55 test set with the default !DEFGRID2, we found an error of about 0.01 ± 0.03 kcal/mol from DFT (compared to a reference grid), and 0.05 ± 0.10 kcal/mol for the COSX (compared to the analytical integration). We also benchmarked geometries, excitation energies and frequencies, and all errors are systematically low.

However, there might still be cases where an improved grid is needed:

- If you need to be confident that your energy error is below 0.2 kcal/mol;
- When dealing with anions with large negative charges (< -3);
- For very subtle intermolecular interactions;
- When dealing with weird electronic structures;
- With large conjugated systems - graphene-like structures and large polyaromatics.

When needed, the !DEFGRID3 is very large and conservative - it was built to cover almost all these cases. In contrast, !DEFGRID1 will yield grids of the size close to previous ORCA versions defaults, but still with increased accuracy.

7.4 Choice of Computational Model

7.4.1 Features Common to All Calculations

The computational model is specified in the block %method. The following choices exist:

```
%method
  Method  HFGTO  (or HF)   # Hartree-Fock with GTOs
          DFGTO  (or DFT)  # Density Functional Theory with GTOs
          MP2              # Second-order Møller-Plesset Perturbation Theory
          CNDO             # Complete neglect of differential overlap
          INDO             # Intermediate neglect of differential overlap
          NDDO             # Neglect of diatomic differential overlap
end
```

In the case of Hartree-Fock calculations [839], nothing else is required in this block. Density functional calculations [451, 646] need slightly more attention, as will be seen in the next section.

The type of calculation to be performed can be chosen by the RunTyp flag as follows:

```
%method
  RunTyp  Energy (or SinglePoint)     # Single point calculation (default)
          Scan   (or Trajectory)      # Geometric scan
          Gradient                    # Single point energy and gradient
          Opt    (or GeometryOpt)     # Geometry optimization
          MD                          # Molecular dynamics calculation
end
```

7.4.2 Density Functional Calculations

Choice of Functional

Basic Choice of Density Functional. If you are doing a DFT calculation [451, 646], the following choices for local, gradient-corrected, and hybrid density functionals are available. See also the simple input keywords in section *Density Functional Methods* for a more complete list of density functionals (including double-hybrids, which cannot be called from the %method block for technical reasons).

```
%method

  Functional
      #***************************************
      # Local functionals
      #***************************************
        HFS        # Hartree-Fock Slater (Slater exchange only)
        LSD        # Local spin density (VWN-5A form)
        VWN5       # Local spin density (VWN-5)
        VWN3       # Local spin density (VWN-3)
        PWLDA      # Local spin density (PW-LDA)
      #***************************************
      # ``Pure'' GGA functionals
      #***************************************
        BNULL      # Becke '88 exchange, no correlation
        BVWN       # Becke '88 exchange, VWN-5 correlation
        BP         # Becke '88 exchange, Perdew '86 correlation
        PW91       # Perdew-Wang (PW) GGA-II '91 functional
        mPWPW      # Modified PW exchange, PW correlation
        mPWLYP     # Modified PW exchange, Lee-Yang-Parr (LYP) correlation
        BLYP       # Becke '88 exchange, LYP correlation
        GP         # Gill '96 exchange, Perdew '86 correlation
```

(continues on next page)

(continued from previous page)

```
        GLYP       # Gill '96 exchange, LYP correlation
        PBE        # Perdew-Burke-Ernzerhof (PBE) functional
        revPBE     # Revised PBE (exchange scaling)
        RPBE       # Revised PBE (modified exchange functional)
        PWP        # PW '91 exchange, Perdew '86 correlation
        OLYP       # Hoe/Cohen/Handy's optimized exchange, LYP correlation
        OPBE       # Hoe/Cohen/Handy's optimized exchange, PBE correlation
        XLYP       # Xu/Goddard exchange, LYP correlation
        B97D       # Grimme's GGA including D2 dispersion correction
        PW86PBE    # PW '86 exchange, PBE correlation (as used for vdw-DF and related)
        RPW86PBE   # Revised PW '86 exchange, PBE correlation
      #***************************************
      # Meta-GGA functionals
      #***************************************
        M06L       # Truhlar's semi-local functional
        TPSS       # TPSS functional
        revTPSS    # Revised TPSS functional
        SCANfunc   # Perdew's SCAN functional
        RSCAN      # Regularized SCAN functional
        R2SCAN     # Regularized and restored SCAN functional
      #***************************************
      # GGA Hybrid functionals
      #***************************************
        B1LYP      # 1-parameter hybrid of BLYP (25% HF exchange)
        B1P        # Similar with Perdew '86 correlation
        G1LYP      # 1-parameter analog with Gill '96 exchange
        G1P        # Similar with Perdew '86 correlation
        B3LYP      # 3-parameter Hybrid of BLYP (20% HF exchange)
        B3P        # Similar with Perdew '86 correlation
        G3LYP      # 3-parameter analog with Gill '96 exchange
        G3P        # Similar with Perdew '86 correlation
        PBE0       # 1-parameter version of PBE (25% HF exchange)
        PWP1       # 1-parameter version of PWP (analog of PBE0)
        mPW1PW     # 1-parameter version of mPWPW (analog of PBE0)
        mPW1LYP    # 2-parameter version of mPWLYP (analog of PBE0)
        PW91_0     # 1-parameter version of PW91 (analog of PBE0)
        O3LYP      # 3-parameter version of OLYP
        X3LYP      # 3-parameter version of XLYP
        B97        # Becke's original hybrid functional
        BHANDHLYP  # Becke's half-and-half hybrid functional (50% HF exchange)
      #***************************************
      # Meta-GGA Hybrid functionals
      #***************************************
        TPSSh      # TPSS hybrid with 10% HF exchange
        TPSS0      # TPSS hybrid with 25% HF exchange
        PW6B95     # Truhlar's 6-parameter hybrid functional
        M06        # Truhlar's 2006 low-HF hybrid (27% HF exchange)
        M062X      # Truhlar's 2006 high-HF hybrid (54% HF exchange)
        r2SCANh    # r2SCAN global hybrid with 10% HF exchange
        r2SCAN0    # r2SCAN global hybrid with 25% HF exchange
        r2SCAN50   # r2SCAN global hybrid with 50% HF exchange
      #***************************************
      # Range-Separated Hybrid functionals
      #***************************************
        wB97       # Head-Gordon's fully variable DF
        wB97X      # Head-Gordon's DF with minimal HF exchange
        CAMB3LYP   # Handy's fit
        LC_BLYP    # Hirao's original application
        LC_PBE     # Hirao's PBE-based range-separated hybrid
        wr2SCAN    # r2SCAN range-separated hybrid
end
```

Note that Functional is a *compound keyword*. It chooses specific values for the variables Exchange, Correlation, and ACM described below. If given as a simple input keyword, in some cases, it will also activate a dispersion correction. You can explicitly give these variables instead or in addition to Functional. However, make sure that you specify these variables *after* you have assigned a value to Functional or the values of Exchange, Correlation and ACM will be reset to the values chosen by Functional.

Empirical Parameters in Density Functionals. Some functionals incorporate empirical parameters that can be changed to improve agreement with experiment. Currently, there are a few parameters that can be changed (other than the parameters used in the hybrid functionals, which are described later).

The first of these parameters is α of Slater's X_α method. Theoretically, it has a value of 2/3 and this is used in the HFS and LSD functionals. However, the exchange contribution is underestimated by about 10% by this approximation (quite significant!) and a value around 0.70-0.75 is recommended for molecules.

The second parameter is the β for Becke's gradient-corrected exchange functional. Becke determined the value 0.0042 by fitting the exchange energies for rare gas atoms. There is some evidence that with smaller basis sets, a slightly smaller value such as 0.0039 gives improved results for molecules.

The next parameter is the value κ, which occurs in the PBE exchange functional. It has been given the value 0.804 by Perdew et al. in order to satisfy the Lieb-Oxford bound. Since then, other workers have argued that a larger value for this parameter (around 1.2) gives better energetics, which is explored in the revPBE functional. Note, however, that while revPBE gives slightly better energetics, it also gives slightly poorer geometries.

The last two parameters are also related to PBE. Within the PBE correlation functional, there is β_C (not to be confused with the β exchange parameter in Becke's exchange functional), whose original value is $\beta_C = 0.066725$. Modified variants exist with different β_C values, e.g., the PBEsol functional and the PBEh-3c compound method. Furthermore, the μ parameter in the PBE exchange functional may be modified. In the original formulation, it is related to β_C via $\mu = \beta_C \frac{\pi^2}{3}$, but has been modified in later variants as well.

```
%method
  XAlpha   0.75       # Slater's alpha parameter         (default 2/3)
  XBeta    0.0039     # Becke's beta parameter           (default 0.0042)
  XKappa   0.804      # PBE(exchange) kappa parameter    (default 0.804)
  XMuePBE  0.21952    # PBE(exchange) mue parameter      (default 0.21952)
  CBetaPBE 0.066725   # PBE(correlation) beta parameter  (default 0.066725)
end
```

Specifying Exchange and Correlation approximations individually. The following keywords are available for specifying the exchange and correlation approximations individually. In addition, scaling parameters can be defined to construct user-defined hybrid or "extended" hybrid functionals:

```
%method

  Exchange
      X_NOX      # No exchange
      X_SLATER   # Slater's local exchange
      X_B88      # Becke '88 gradient exchange
      X_wB88     # Short-range Becke '88 exchange for range-separated functionals
      X_G96      # Gill '96 gradient exchange
      X_PW91     # Perdew-Wang (PW) '91 gradient exchange
      X_mPW      # Adamo-Barone modification of PW exchange
      X_PBE      # Perdew-Burke-Ernzerhof (PBE) exchange
      X_RPBE     # Revised PBE exchange
      X_OPTX     # Hoe/Cohen/Handy's optimized exchange
      X_X        # Xu/Goddard exchange
      X_TPSS     # TPSS meta-GGA exchange
      X_B97D     # Grimme's modified exchange for the B97-D GGA
      X_B97BECKE # Becke's original exchange for the B97 hybrid
      X_SCAN     # Perdew's constrained exchange for the SCAN meta-GGA
      X_RSCAN    # Perdew's constrained exchange for the rSCAN meta-GGA
      X_R2SCAN   # Perdew's constrained exchange for the r2SCAN meta-GGA

  Correlation
```

(continues on next page)

(continued from previous page)

```
      C_NOC      # No correlation
      C_VWN5     # Local Vosko-Wilk-Nusair correlation (parameter set "V")
      C_VWN3     # Local Vosko-Wilk-Nusair correlation (parameter set "III")
      C_PWLDA    # Local PW correlation
      C_P86      # Perdew '86 correlation
      C_PW91     # PW '91 correlation
      C_PBE      # PBE correlation
      C_LYP      # LYP correlation
      C_TPSS     # TPSS meta-GGA correlation
      C_B97D     # Grimme's modified correlation for the B97-D GGA
      C_B97BECKE # Becke's original correlation for the B97 hybrid
      C_SCAN     # Perdew's constrained correlation for the SCAN meta-GGA
      C_RSCAN    # Perdew's constrained correlation for the rSCAN meta-GGA
      C_R2SCAN   # Perdew's constrained correlation for the r2SCAN meta-GGA

  # Parameters for hybrid functionals
  ACM   ACM-A, ACM-B, ACM-C
                 # ACM-A: fraction of HF exchange in hybrid DFT
                 # ACM-B: scaling of GGA exchange part of DFT
                 # ACM-C: scaling of GGA correlation part of DFT

  # Parameters for "extended" hybrid functionals
  ScalLDAC 1.0   # scaling of the LDA correlation part
  ScalMP2C 0.0   # fraction of MP2 correlation mixed into the density functional

end
```

Hybrid Density Functionals. The hybrid DFs [87, 88] are invoked by choosing a nonzero value for the variable `ACM`. (ACM stands for "adiabatic connection model"). Specifically, these functionals have the following form:

$$E_{\mathrm{XC}} = aE_{\mathrm{HF}}^{\mathrm{X}} + (1-a)\,E_{\mathrm{LSD}}^{\mathrm{X}} + bE_{\mathrm{GGA}}^{\mathrm{X}} + E_{\mathrm{LSD}}^{\mathrm{C}} + cE_{\mathrm{GGA}}^{\mathrm{C}} \tag{7.1}$$

Here, E_{XC} is the total exchange/correlation energy, $E_{\mathrm{HF}}^{\mathrm{X}}$ is the Hartree-Fock exchange, $E_{\mathrm{LSD}}^{\mathrm{X}}$ is the local (Slater) exchange, $E_{\mathrm{GGA}}^{\mathrm{X}}$ is the gradient correction to the exchange, $E_{\mathrm{LSD}}^{\mathrm{C}}$ is the local, spin-density based part of the correlation energy, and $E_{\mathrm{GGA}}^{\mathrm{C}}$ is the gradient correction to the correlation energy.

This brings us to a slightly awkward subject: several hybrid functionals with the same name give different values in different programs. The reason for this is that they choose slightly different default values for the parameters a, b, and c and/or differ in the way they treat the local part of the correlation energy.

Different parametrizations exist. The most popular example of this is due to Vosko, Wilk, and Nusair (VWN, [873]). VWN in their classic paper give two sets of parameters — one in the main body (parametrization of RPA results; known as VWN-III) and one in their table 5 of correlation energies (parametrization of the Ceperley/Alder Monte Carlo results; known as VWN-V). Some programs use VWN-III, while others use VWN-V. Additionally, a slightly better fit to the uniform electron gas has been produced by Perdew and Wang [667]. The results from this fit are very similar to those using `VWN5`, whereas `VWN3` yields quite different values.

In ORCA, almost all functionals choose `PWLDA` as the underlying LDA functional. A special situation arises if LYP is the correlation functional [502] since LYP is *not* a correction to the correlation, but rather includes the full correlation. It is therefore used in the B3LYP functional as:

$$E_{\mathrm{B3LYP}}^{\mathrm{C}} = E_{\mathrm{LSD}}^{\mathrm{C}} + c\left(E_{\mathrm{LYP}}^{\mathrm{C}} - E_{\mathrm{LSD}}^{\mathrm{C}}\right) \tag{7.2}$$

In ORCA, `VWN5` is chosen for the local correlation part of B3LYP. This choice is consistent with the TurboMole program [12, 13, 14], but not with the Gaussian program [278]. However, the user has full control over this setting. The underlying local part of any correlation functional can be set in the input with the variable `LDAOpt`:

```
%method
  LDAOpt  C_PWLDA
          C_VWN5
          C_VWN3
end
```

Specifying C_VWN3 for LDAOpt together with Functional=B3LYP should give results very close to the B3LYP functional as implemented in the Gaussian series of programs[1]. Due to the popularity of the B3LYP functional, the following aliases are defined in order to facilitate comparisons with other major electronic structure packages:

```
%method
  Functional B3LYP     # consistent with TurboMole
             B3LYP_TM  #    := Functional B3LYP
                       #       LDAOpt     C_VWN5
             B3LYP_G   # consistent with Gaussian
                       #    := Functional B3LYP
                       #       LDAOpt     C_VWN3
end
```

One-Parameter Hybrid Density Functionals. The underlying LDA-dependence of the three-parameter (i.e. ACM) hybrids causes slightly different results from programs which use a different choice of LDA. It has been argued from a theoretical viewpoint that the optimal mixing of HF exchange is 25% [248]. It has been further shown that use of this fixed ratio and *not* scaling the GGA correlation or exchange contributions gives results that are as good as the original three-parameter hybrids [9]. The one-parameter hybrid PBE0 has been advertised as a hybrid functional of overall well-balanced accuracy [8].

As such, the one-parameter hybrids are based more in theory and remove some arbitrariness from the hybrid procedures. The slightly higher HF-exchange (0.25 in favor of 0.20 used in the original three-parameter hybrids) is also reasonable for many systems. The one-parameter hybrids have the simple form:

$$E_{\mathrm{XC}} = E_{\mathrm{DFT}}^{\mathrm{X}} + a' \left(E_{\mathrm{HF}}^{\mathrm{X}} - E_{\mathrm{DFT}}^{\mathrm{X}} \right) + E_{\mathrm{DFT}}^{\mathrm{C}} \tag{7.3}$$

with $a' = \frac{1}{4}$. This is the same as the three-parameter hybrids (equation (7.1)) with $a = a'$, $b = 1 - a'$, and $c = 1$ and is how it is implemented.

Extended "double-hybrid" functionals. In addition to mixing the HF-exchange into a given DF, Grimme has proposed to mix in a fraction of the MP2 correlation energy as calculated with hybrid DFT orbitals [320]. Such functionals may be referred to as "extended" or "double" hybrid functionals. Grimme's expression is:

$$E_{\mathrm{XC}} = a E_{\mathrm{X}}^{\mathrm{HF}} + (1-a)\, E_{\mathrm{X}}^{\mathrm{DFT}} + (1-c)\, E_{\mathrm{C}}^{\mathrm{DFT}} + c E_{\mathrm{C}}^{\mathrm{MP2}} \tag{7.4}$$

Such functionals can be defined by the user in ORCA as follows:

```
%method
  ScalHFX  = a
  ScalDFX  = 1-a
  ScalGGAC = 1-c
  ScalLDAC = 1-c
  ScalMP2C = c
end
```

Grimme recommends the B88 exchange functional, the LYP correlation functional and the parameters $a = 0.53$ and $c = 0.27$. This gives the B2PLYP functional which appears to be a fair bit better than B3LYP based on Grimme's detailed evaluation study.

Presently, this methodology covers single point calculations, analytic gradients (and thus also geometry optimizations, relaxed scans, and transition state searches), and frequencies and other second-derivative properties (without the frozen core approximation in the MP2 part). Note that `%mp2 density relaxed end` should be specified in order to get the response density that is consistent with first-order properties from analytic first-derivatives. By default, this response density is not calculated since its construction adds significant overhead to the calculation. Therefore, you have to specifically request it to look at the consistent density. The considerably less costly unrelaxed (expectation value-like) density may instead requested by `%mp2 density unrelaxed end`. However, this is not recommended since the changes to the relaxed density are considerable in our experience and the unrelaxed density has a much weaker theoretical status than its relaxed counterpart.

[1] There is some evidence that the version used in the Gaussian program gives slightly better results in molecular applications than the TurboMole variant, but the differences are very small [387].

Range-separated hybrid functionals. ORCA supports functionals based on the error function splitting of the two-electron operator used for exchange as first realized by Hirao and coworkers [410]:

$$r_{12}^{-1} = \underbrace{\text{erfc}(\mu \cdot r_{12}) \cdot r_{12}^{-1}}_{\text{SR}} + \underbrace{\text{erf}(\mu \cdot r_{12}) \cdot r_{12}^{-1}}_{\text{LR}} \tag{7.5}$$

where $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)dt$ and $\text{erfc}(x) = 1 - \text{erf}(x)$. Note that the splitting is only applied to the exchange contribution; all other contributions (one-electron parts of the Hamiltonian, the electron-electron Coulomb interaction and the approximation for the DFT correlation) are not affected. Later, Handy and coworkers generalized the ansatz to [899]:

$$r_{12}^{-1} = \underbrace{\frac{1 - [\alpha + \beta \cdot \text{erf}(\mu \cdot r_{12})]}{r_{12}}}_{\text{SR}} + \underbrace{\frac{\alpha + \beta \cdot \text{erf}(\mu \cdot r_{12})}{r_{12}}}_{\text{LR}} \tag{7.6}$$

This form of the splitting used in ORCA is shown visually (according to Handy and coworkers) in Fig. 7.1.

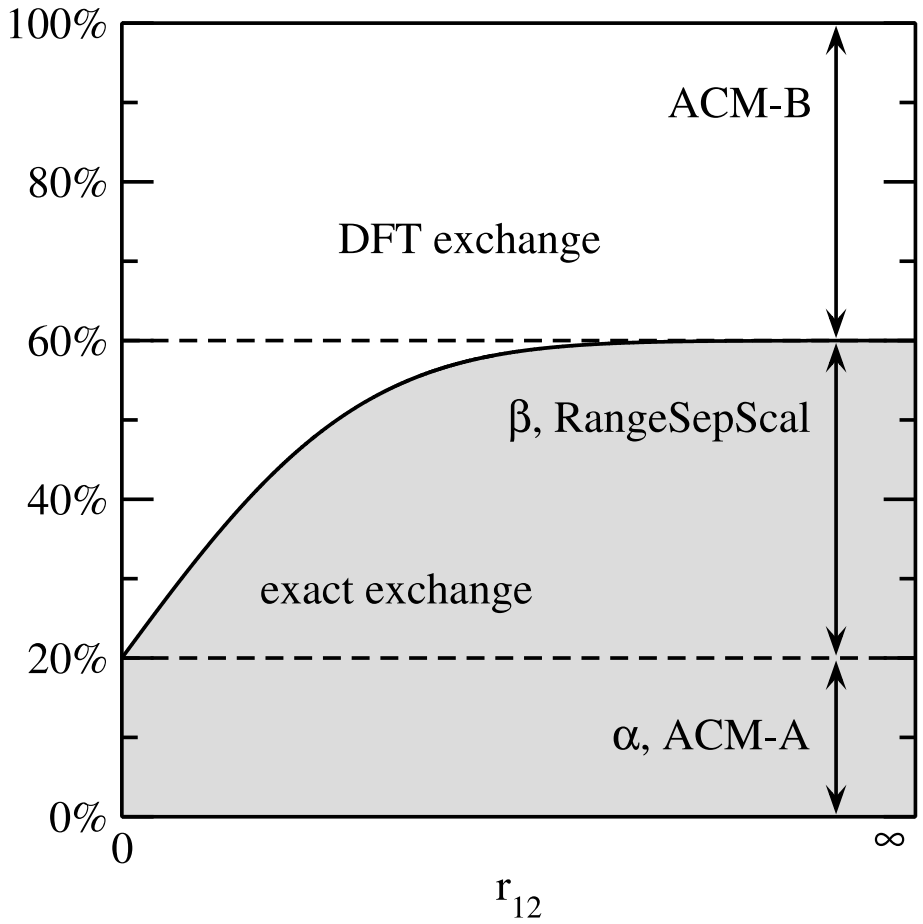


Fig. 7.1: Graphical description of the Range-Separation ansatz. The gray area corresponds to Hartree-Fock exchange. α and β follow Handy's terminology [899].

The splitting has been used to define the ωB97 family of functionals, wherein the short-range part (SR) is described by DFT exchange and the long-range part by exact exchange, i.e. Hartree-Fock exchange. The same is true for CAM-B3LYP, LC-BLYP, and LC-PBE. It is possible to use a fixed amount of Hartree-Fock exchange (EXX) and/or a fixed amount of DFT exchange in this ansatz. Here are some examples of range-separated hybrid functionals available in ORCA and their parameters:

Functional	Keyword	fixed EXX	variable part	μ/bohr^{-1}	fixed DFT-X	Reference
ωB97	WB97	—	100%	0.40	—	[151]
ωB97X	WB97X	15.7706%	84.2294%	0.30	—	[151]
ωB97X-D3	WB97X-D3	19.5728%	80.4272%	0.25	—	[525]
ωB97X-V	WB97X-V	16.7%	83.3%	0.30	—	[554]
ωB97X-D3BJ	WB97X-D3BJ	16.7%	83.3%	0.30	—	[602]
CAM-B3LYP	CAM-B3LYP	19%	46%	0.33	35%	[899]
LC-BLYP	LC-BLYP	—	100%	0.33	—	[845]
LC-PBE	LC-PBE	—	100%	0.47	—	[410]

The currently available speed-up options are RIJONX and RIJCOSX. Integral-direct single-point calculations, calculations involving the first nuclear gradient (i.e. geometry optimizations), frequency calculations, TDDFT, TDDFT nuclear gradient, and EPR/NMR calculations are the supported job types with range-separated hybrid functionals thus far.

In principle, it is possible to change the amount of fixed Hartree-Fock exchange (ACM-A), the amount of variable exchange (RangeSepScal), and μ, though this is not recommended. The amount of fixed DFT Exchange (ACM-B) can only be changed for CAM-B3LYP and LC-BLYP. In other words, ACM-B is ignored by the ωB97 approaches as no corresponding μ-independent exchange functional has been defined.

```
! RKS CAM-B3LYP DEF2-SVP

# default parameters for CAM-B3LYP:
%method
  RangeSepEXX  True    # must be set
  RangeSepMu   0.33    # should not be set to 0.0 or below
  RangeSepScal 0.46    # should sum to 1 with ACM-A and ACM-B
  ACM 0.19, 0.35, 0.81 # ACM-A, ACM-B, ACM-C(same as B3LYP)
end

* xyz 0 1
H 0.0 0.0 0.0
H 0.0 0.0 0.7
*
```

Note

For information on the ACM formalism, see preceding section called "Specifying Exchange and Correlation approximations individually". While it is technically possible to choose an exchange functional that has no μ-dependence, this makes no sense conceptually.

LibXC Functionals

Since ORCA 4.2, it is possible to use the functionals provided by LibXC[2] within the ORCA framework. The LibXC version used by ORCA is printed at the beginning of the output. The LibXC reference should be cited when LibXC is used as part of your calculations. For reference, see [506].

The complete list of functionals available via the LibXC interface can always be inspected by typing at the command line

```
orca -libxcfunctionals
```

The list of functionals has the following form:

[2] https://libxc.gitlab.io

```
Functionals available via LibXC:
  No.:     ID / Keyword                    - Name
    0:     18 / lda_c_1d_csc               - Casula, Sorella & Senatore
    1:     26 / lda_c_1d_loos              - P-F Loos correlation LDA
    2:     15 / lda_c_2d_amgb              - AMGB (for 2D systems)
    3:     16 / lda_c_2d_prm               - PRM (for 2D systems)
    4:    552 / lda_c_br78                 - Brual & Rothstein 78

   ...
```

The list is grouped by "level" of functional (LDA, GGA, meta-GGA, etc.) and then by part of the energy it models (correlation, exchange, exchange-correlation). Correlation functionals carry a `'_c_'` in their names, exchange functionals an `'_x_'`, and combined exchange-correlation functionals an `'_xc_'`. Additional information for a specific functional can be requested using

```
orca -libxcinfo [functional name]
```

where the functional name is the keyword in the above list.

Specification of LibXC functionals in the ORCA input follows the standard format:

```
%method
  method     dft
  functional hyb_gga_xc_b3lyp
end
```

or in the case of separate exchange and correlation specifications:

```
%method
  method      dft
  exchange    mgga_x_m06_l
  correlation mgga_c_m06_l
end
```

CAM-type range-separated functionals are supported through the LibXC interface since ORCA 5.0. So are functionals that include non-local (VV10) correlation (see *DFT Calculations with the Non-Local, Density Dependent Dispersion Correction (VV10): DFT-NL*). The VV10 part is computed internally by ORCA. Other non-local functionals, such as BEEF-vdW, are not supported. Meta-GGA functionals that depend on the kinetic energy density τ are supported, but not those that depend on the Laplacian of the density $\nabla^2\rho$.

Double-hybrid functionals can be constructed with LibXC components as described in section *DFT Calculations with Second Order Perturbative Correction (Double-Hybrid Functionals)*, but only with separate exchange and correlation components. So `exchange=gga_x_pbe` and `correlation=gga_c_pbe` can be used, but `functional=hyb_gga_xc_pbeh` **cannot** be used in a double-hybrid formulation. Beware that the exchange and correlation contributions calculated by LibXC are simply scaled by `ScalDFX` and `ScalGGAC`, respectively, and no care is taken to separately scale LDA components or alter other internal parameters!

Screening Thresholds

When the density is smaller than a certain threshold, LibXC skips the evaluation of the functional and instead sets all the output quantities to zero in order to avoid under- and/or overflows. The default thresholds for different functionals are set by the LibXC developers, but may sometimes be too low. We have not performed extensive testing, but allow the user to set the threshold. Similarly, there are thresholds for minimum values of ζ and τ in order to avoid division by zero. The default values are functional-independent and can be changed using the following keywords.

```
%method
  LibXCDensityThreshold 1e-12    # seems to be reasonable
  LibXCZetaThreshold    2e-16    # default value in LibXC
```

(continues on next page)

(continued from previous page)

```
  LibXCTauThreshold     1e-20    # default value in LibXC
end
```

Modifying LibXC Functional Parameters

Starting with ORCA 5.1 it is possible to modify the "external parameters" of a functional that the LibXC interface provides. The available parameters and their default values can be seen in the output or with the `orca -libxcinfo` command above. Please exercise caution when using this interface and if using a published functional reparametrization, make sure you can reproduce results from the original publication. The syntax requires one of the `ExtParamX`, `ExtParamC`, or `ExtParamXC` keywords (depending on which functional is modified), followed by the parameter name in quotation marks, and finally the new value. For example, here is an input for the PWPB95 double-hybrid functional, where the simple input keyword is used to set most parameters (such as the MP2 scaling factors) and only the exchange and correlation parameters of the mPW91 and B95 LibXC functionals, respectively, are modified.

```
! Opt PWPB95 def2-SVP def2-SVP/C

%method
  Exchange    gga_x_mpw91
  Correlation mgga_c_bc95
  ExtParamX   "_bt"     0.00444
  ExtParamX   "_alpha" 19.823391
  ExtParamX   "_expo"   3.7868
  ExtParamC   "_css"    0.03241
  ExtParamC   "_copp"   0.00250
end

*xyz 0 1
  H 0 0 0
  F 0 0 0.9
*
```

Note that the variable definitions in LibXC may be different from the ones used internally in ORCA or in the original publication, due to various constant factors, etc. When in doubt, please consult the LibXC documentation or source code repository — we simply provide access to the interface.

Simple Input of LibXC Functionals

Some LibXC functionals are accessible via the simple input keyword `!LibXC(Keyword)`, e.g. `!LibXC(TPSS)`. The keywords match those in section *Density Functional Methods* for functionals that are also implemented natively in ORCA. Using this syntax, parameters for the DFT-D dispersion corrections are also set automatically (if implemented). Table Table 7.4 lists the available functionals and their LibXC names.

Table 7.4: LibXC functionals available via the simple input `!LibXC(Keyword)`

Keyword	LibXC name(s)	Notes
GGA functions		
B97-D	gga_xc_b97_d	Uses D2
B97-D3	gga_xc_b97_d	Uses D3(0)
B97-D4	gga_xc_b97_d	Uses D4
BLYP	gga_x_b88 + gga_c_lyp	
BP86	gga_x_b88 + gga_c_p86	
KT2	gga_xc_kt2	
KT3	gga_xc_kt3	
PBE	gga_x_pbe + gga_c_pbe	

continues on next page

Table 7.4 – continued from previous page

Keyword	LibXC name(s)	Notes
Meta-GGA functionals		
B97M-V	mgga_xc_b97m_v	Uses VV10
B97M-D3BJ	mgga_xc_b97m_v	Uses D3(BJ)
B97M-D4	mgga_xc_b97m_v	Uses D4
M06L	mgga_x_m06_l + mgga_c_m06_l	
MN15L	mgga_x_mn15_l + mgga_c_mn15_l	
r2SCAN	mgga_x_r2scan + mgga_c_r2scan	
rSCAN	mgga_x_rscan + mgga_c_rscan	
SCAN	mgga_x_scan + mgga_c_scan	
TPSS	mgga_x_tpss + mgga_c_tpss	
Hybrid GGA functionals		
B3LYP	hyb_gga_xc_b3lyp5	
B3LYP/G	hyb_gga_xc_b3lyp	
B97	hyb_gga_xc_b97	
PBE0	hyb_gga_xc_pbeh	
wB97	hyb_gga_xc_wb97	
wB97X	hyb_gga_xc_wb97x	
wB97X-D3	hyb_gga_xc_wb97x_d3	Uses D3(0)
wB97X-V	hyb_gga_xc_wb97x_v	Uses VV10
wB97X-D3BJ	hyb_gga_xc_wb97x_v	Uses D3(BJ)
wB97X-D4	hyb_gga_xc_wb97x_v	Uses D4
Hybrid meta-GGA functionals		
M05	hyb_mgga_x_m05 + mgga_c_m05	
M052X	hyb_mgga_x_m05_2x + mgga_c_m05_2x	
M06	hyb_mgga_x_m06 + mgga_c_m06	
M062X	hyb_mgga_x_m06_2x + mgga_c_m06_2x	
MN15	hyb_mgga_x_mn15 + mgga_c_mn15	
PW6B95	hyb_mgga_xc_pw6b95	
SCAN0	hyb_mgga_x_scan0 + mgga_c_scan	
TPSS0	hyb_mgga_xc_tpss0	
TPSSH	hyb_mgga_xc_tpssh	
wB97M-V	hyb_mgga_xc_wb97m_v	Uses VV10
wB97M-D3BJ	hyb_mgga_xc_wb97m_v	Uses D3(BJ)
wB97M-D4	hyb_mgga_xc_wb97m_v	Uses D4
Double-hybrid functionals		
DSD-PBEB95-D3	gga_x_pbe + mgga_c_bc95	Custom mixing, uses D3(BJ)
DSD-PBEB95-D4	gga_x_pbe + mgga_c_bc95	Custom mixing, uses D4
PWPB95	gga_x_mpw91 + mgga_c_bc95	Custom mixing and `ExtParam`'s

Using the RI-J Approximation to the Coulomb Part

Note

This is the default for non-hybrid DFT! Can be turned off by using `!NORI`.

A very useful approximation that greatly speeds up DFT calculations unless the molecule gets very large is the so called "RI-approximation" [65, 225, 245, 246, 440, 861, 887]. RI stands for "Resolution of the identity". In short, charge distributions arising from *products of basis functions* are approximated by a linear combination of *auxiliary basis functions*.

$$\phi_i(\vec{r})\,\phi_j(\vec{r}) \approx \sum_k c_k^{ij}\,\eta_k(\mathbf{r}) \tag{7.7}$$

There are a variety of different possibilities to determine the expansion coefficients c_k^{ij}. A while ago, Almlöf and coworkers [859] have shown that for the approximation of electron repulsion integrals, the best choice is to minimize the *residual repulsion*[3].

Define:

$$R_{ij} \equiv \phi_i(\vec{r})\,\phi_j(\vec{r}) - \sum_k c_k^{ij}\eta_k(\vec{r}) \tag{7.8}$$

and

$$T_{ij} = \int\int R_{ij}(\vec{r})\,\frac{1}{|\vec{r}-\vec{r}'|}R_{ij}(\vec{r})\,d^3r d^3r' \tag{7.9}$$

Determining the coefficients that minimize T_{ij} leads to

$$\mathbf{c}^{ij} = \mathbf{V}^{-1}\mathbf{t}^{ij} \tag{7.10}$$

where:

$$t_k^{ij} = \left\langle \phi_i\phi_j \left| r_{12}^{-1} \right| \eta_k \right\rangle \tag{7.11}$$

$$V_{ij} = \left\langle \eta_i \left| r_{12}^{-1} \right| \eta_j \right\rangle \tag{7.12}$$

Thus, an ordinary two-electron integral becomes

$$\begin{aligned}\left\langle \phi_i\phi_j \left| r_{12}^{-1} \right| \phi_k\phi_l \right\rangle &\approx \sum_{p,q} c_p^{ij} c_q^{kl} V_{pq} \\ &= \sum_{p,q} V_{pq} \sum_r \left(\mathbf{V}^{-1}\right)_{pr} t_r^{ij} \sum_s \left(\mathbf{V}^{-1}\right)_{qs} t_s^{kl} \\ &= \sum_{r,s} \left(\mathbf{V}^{-1}\right)_{rs} t_r^{ij} t_s^{kl}\end{aligned} \tag{7.13}$$

and the total Coulomb energy becomes

$$\begin{aligned}E_J &= \sum_{i,j}\sum_{k,l} P_{ij}P_{kl}\left\langle \phi_i\phi_j \left| r_{12}^{-1} \right| \phi_k\phi_l \right\rangle \\ &\approx \sum_{i,j}\sum_{k,l} P_{ij}P_{kl} \sum_{r,s} \left(\mathbf{V}^{-1}\right)_{rs} t_r^{ij} t_s^{kl} \\ &= \sum_{r,s} \left(\mathbf{V}^{-1}\right)_{rs} \underbrace{\sum_{i,j} P_{ij} t_r^{ij}}_{\mathbf{x}_r} \underbrace{\sum_{k,l} P_{kl} t_s^{kl}}_{\mathbf{x}_s}\end{aligned} \tag{7.14}$$

where $\mathbf{P}$ is the total density matrix.

In a similar way, the Coulomb contribution to the Kohn-Sham matrix is calculated. There are substantial advantages from this approximation: the quantities to be stored are the matrix $\mathbf{V}^{-1}$ — which depends only on two indices — and the three-index auxiliary integrals t_r^{ij}. This leads to a tremendous reduction of storage requirements compared to a four-index list of repulsion integrals. In addition, the two- and three-index electron repulsion integrals are easier to compute than the four-index integrals, leading to further reductions in processing time. Furthermore, the Coulomb energy and the Kohn-Sham matrix contributions can be quickly assembled by simple vector/matrix operations, leading to large time savings. This arises because each auxiliary basis function $\eta_k(\vec{r})$ appears in the expansion of many charge distributions $\phi_i(\vec{r})\,\phi_j(\vec{r})$. Unfortunately, a similar strategy is less easily applied (or, at least, with less benefit) to the Hartree-Fock exchange term.

If the auxiliary basis set $\{\eta\}$ is large enough, the approximation is also highly accurate. Since any DFT procedure already has a certain, sometimes sizable, error from the noise in the numerical integration of the XC part, it might be argued that a similarly large error in the Coulomb part is perfectly acceptable without affecting the overall accuracy

[3] The basic theory behind the RI method has been known for a long time and since at least the late sixties, methods similar to the RI approximation have been used — mainly in the context of "approximate *ab initio* methods" such as LEDO, PDDO, and MADO, but also in density functional theory in the mid and late seventies by Baerends, Dunlap, and others [65, 225, 861, 887].

of the calculation much. Furthermore, the errors introduced by the RI method are usually much smaller than the errors in the calculation due to basis set incompleteness. It is therefore recommended to use the RI procedure for pure DFs. However, one should probably not directly mix absolute total energies obtained from RI and non-RI calculations as the error in the total energy accumulates and will rise with increased molecular size, while the errors in the relative energies will tend to cancel.

There are several choices for auxiliary basis sets described in the next section, which depend on the choice of the primary GTO basis set used to expand the molecular orbitals[4].

In ORCA, the RI approximation is toggled by the input

```
%method
  RI  on   # do use the RI-J approximation
      off  # do not use the RI-J approximation
end
```

Note

If you use RI, you *must* specify an auxiliary basis set (in the `%basis` section or using the appropriate simple keyword) or use the `!AutoAux` simple keyword.

The Split-RI-J Coulomb Approximation

There is an improved version of the RI-algorithm that has been implemented since ORCA 2.2.09. This Split-RI-J algorithm yields the same Coulomb energy as the standard RI-algorithm, but is significantly faster if the basis set contains many high angular momentum functions (d-, f-, g-functions). For small basis sets, there is virtually no difference between the two algorithms, except that Split-RI-J uses more memory than standard RI. However, calculations with ca. 2000 basis functions only need about an extra 13 MB for Split-RI-J, which is a trivial requirement on present-day hardware.

The Split-RI-J algorithm is invoked with the `!Split-RI-J` simple keyword. Split-RI-J is presently only available for SCF and gradient calculations.

Note

- The Split-RI-J algorithm is the default if RI is turned on via `!RI`. If you do not want to use Split-RI-J, please also use the keyword `!NoSplit-RI-J`

Using the RI Approximation for Hartree-Fock and Hybrid DFT (RIJONX)

The RI approximation can be used, although with less benefit, for hybrid DFT and Hartree-Fock (RHF and UHF) calculations. In this case, a different algorithm[5] is used that allows a fair approximation to the Hartree-Fock exchange matrix (RI-JK). It can be difficult to make this approximation highly accurate. It is, however, usefully fast compared to direct SCF if the molecule is "dense" enough. There are special auxiliary basis sets for this purpose (see section *Basis Sets*).

```
%method
  RI  on
end
```

(continues on next page)

[4] It probably should be noted that a slightly awkward step in the procedure is the inversion of the auxiliary integral matrix $\mathbf{V}$, which can easily become very large. Matrix inversion is an $O(N^3)$ process such that for large molecules, this step takes some real time. However, in ORCA, this is only done *once* during the calculation, whereas other programs that constrain the fit to also exactly reproduce the number of electrons perform a similar process *each iteration*. Starting from ORCA 2.2.09, the Cholesky decomposition is used in favor of matrix inversion, removing any bottleneck concerning the solution of the linear equation system.

[5] This algorithm was described by Kendall and Früchtl [440].

(continued from previous page)

```
%basis
  Aux "def2/JK"
end
```

Note

There has been little experimentation with this feature. It is provided on an experimental basis here.

Alternatively, the RI method can be used for the Coulomb term and the standard treatment for the exchange term. This method is called RIJONX since the exchange term should tend towards linear scaling for large molecules. This feature can be used for Hartree-Fock and hybrid DFT calculations by using:

```
%method
  RI      on      # do use the RI approximation
  RIFlags 1       # ...but treat exchange exactly
end

# Equivalently, use the following simple keyword:
! RIJONX

# Remember to assign an auxiliary basis set!
```

The requirements for the auxiliary basis are the same as for the normal RI-J method.

Using the RI Approximation for Hartree-Fock and Hybrid DFT (RIJCOSX)

Note

This is the default for hybrid DFT! Can be turned off by using `!NOCOSX`.

The aim of this approximation is to efficiently compute the elements of exchange-type matrices[6]:

$$K_{\mu\nu} = \sum_{\kappa\tau} P_{\kappa\tau}(\mu\kappa|\nu\tau) \tag{7.15}$$

where $\mathbf{P}$ is some kind of density-type matrix (not necessarily symmetric) and the two-electron integrals are defined over the basis set $\{\varphi\}$ by:

$$(\mu\kappa|\nu\tau) = \int \mu(\mathbf{r}_1)\kappa(\mathbf{r}_1)\nu(\mathbf{r}_2)\tau(\mathbf{r}_2)r_{12}^{-1}d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.16}$$

The approximation pursued here can be written as follows:

$$K_{\mu\nu} \approx \sum_g X_{\mu g} \sum_\tau A_{\upsilon\tau}(\mathbf{r}_g) \sum_\kappa P_{\kappa\tau} X_{\kappa g} \tag{7.17}$$

Here, the index g refers to grid points $\mathbf{r}_g$ and:

$$X_{\kappa g} = w_g^{1/2}\kappa(\mathbf{r}_g) \tag{7.18}$$

$$A_{\upsilon\tau}(\mathbf{r}_g) = \int \frac{\nu(\mathbf{r})\tau(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_g|}d\mathbf{r} \tag{7.19}$$

where w_g denotes the grid weights. Thus, the first integration is carried out numerically and the second one analytically. Note that this destroys the Hermitian character of the two-electron integrals.

[6] The theory of this approach together with all evaluations and implementation details is described in [383, 623]. References to earlier work can also be found there.

Equation (7.17) is perhaps best evaluated in three steps:

$$F_{\tau g} = (\mathbf{PX})_{\tau g} \tag{7.20}$$

$$G_{\nu g} = \sum_{\tau} A_{\nu\tau}(\mathbf{r}_g) F_{\tau g} \tag{7.21}$$

$$K_{\mu\nu} = (\mathbf{XG}^{+})_{\mu\nu} \tag{7.22}$$

As such, the equations are very similar to the pseudo-spectral method extensively developed and discussed by Friesner and co-workers since the mid 1980s and commercially available in the Jaguar quantum chemistry package. The main difference at this point is that instead of $X_{\kappa g}$ there appears a least-square fitting operator $Q_{\kappa g}$ in Friesner's formulation. Note that an analogue of the fitting procedure has also been implemented in ORCA and — in contrast to Friesner's pseudo-spectral method — does not need specially optimized grids. The basic idea is to remove the grid errors within the basis set by "fitting" the numerical overlap to the analytical one. Due to its nature, overlap fitting is supposed to work better with larger basis sets.

The RIJCOSX gradient

Given the exchange matrix, the exchange energy is given by (a sum over spin cases is left out here for simplicity):

$$E_{\mathrm{X}} = \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} K_{\mu\nu}(\mathbf{P}) \tag{7.23}$$

Previous to ORCA6, the gradient of the COSX contribution to the energy was taken as an approximation:

$$\frac{\partial E_X}{\partial \lambda} \approx 2 \sum_{g} \sum_{\mu\nu} \frac{\partial F_{\mu g}}{\partial \lambda} G_{\nu g} \tag{7.24}$$

with

$$\frac{\partial F_{\mu g}}{\partial \lambda} = w_g^{1/2} \sum_{\kappa} P_{\kappa\mu} \frac{\partial X_{\mu g}}{\partial \lambda} \tag{7.25}$$

as published in the original implementation paper [623].

Starting from ORCA6, this was updated to the full derivative of the energy component, including the derivative of all terms: grid weight derivatives, derivative of the SFitting matrices and all derivatives of F and G[7]. The gradient is thus now more accurate and less noisy. In case one wants to revert to the previous approximated version (not recommended), just set:

```
%method
  cosxgradtype grad_n
  useqgradfit false
end
```

Working with the COSX Numerical Grids

For expert users, the grid parameters for the exchange grids can be even more finely controlled:

```
%method
  IntAccX  Acc1, Acc2, Acc3
  GridX    Ang1, Ang2, Ang3
end
```

There are three grids involved: the smallest grid (Acc1, Ang1) that is used for the initial SCF iterations, the medium grid (Acc2, Ang2) that is used until the end of the SCF and the largest grid (Acc3, Ang3) that is used for the final energy and the gradient evaluations. `UseFinalGridX` can be used to turn this last grid on or off, though changing this is not generally recommended. More details about the grid constructions can be found in *Details on the numerical integration grids*.

[7] The theory is not yet published, but will be soon.

Some SFitting Parameters

To modify the overlap fitting parameters, the following input can be specified:

```
%method
  UseSFitting false   # Same as NoSFitting in the simple input
                      # (Default is true)
  UseQGradFit false   # Uses the SCF fitting matrix for gradient calculations
                      # (Default is true)
end
```

Note that overlap fitting also works for HF and MP2 gradients without specifying any additional keyword. The `UseQGradFit` parameter uses the same fitting matrix for the gradients as for the energy calculation and is the default behavior since ORCA6.

Use of Partial Contraction Scheme

Since ORCA 5.0, generally-contracted basis sets can be handled efficiently by using an intermediate partially contracted (pc) atomic-orbital basis for the exchange-matrix computation without affecting the results [383]. Depending on the basis set and element type, computational speedups by many orders of magnitude are possible. For testing or benchmark purposes, the K matrix computation can be done in the original basis by using the flag

```
%method
  COSX_PartialContraction false   # No intermediate basis for generally contracted shells
                                  # (Default is true)
end
```

Restoring Full Symmetry

The semi-numerical integration scheme in the default COSX algorithm breaks the permutational symmetry of the two-electron integrals. We have observed that this flaw is often the cause of convergence problems for iterative algorithms, in particular for multi-reference theories [382]. An input option is available since ORCA 6.0 to preserve the full eight-fold permutational symmetry of the two-electron integrals:

```
%method
  COSX_IntSymmetry Full      # Fully symmetrized integrals
                   Standard  # Original COSX algorithm
end
```

The full-symmetrization algorithm often improves the numerical stability and is enabled by default for TRAH-CASSCF and CASSCF linear-response calculations. However, the full-symmetrization algorithm may come with additional costs that depend on the number of $\mathbf{F}$ intermediates. The number of $\mathbf{F}$ intermediates depends on the symmetry of the density matrix (symmetric (S), anti-symmetric (A), and non-symmetric (N)) and whether overlap fitting is employed, as summarized in the table below.

Table 7.5: Number of COSX $\mathbf{F}$ intermediates per density matrix

S fitting	P symmetry (S, A, N)	Number of $\mathbf{F}$ intermediates
No	S / A	1
No	N	2
Yes	S / A	2
Yes	N	4

Note that for symmetric (S) and anti-symmetric (A) densities, we symmetrize and anti-symmetrize, respectively, exchange matrices at the end, which reduces the number of $\mathbf{F}$ intermediates by a factor of 2. The actual computational costs usually do not increase linearly with the number of $\mathbf{F}$ intermediates since we compute the costly

analytic integrals ($A_{\nu\tau}(\mathbf{r}_g)$) only once and then contract them with the additional $\mathbf{F}$ intermediates. From our experience the overhead of the full-symmetrization algorithm is roughly between 1.2 and 1.5 times that of the original algorithm.

COSX Grid and Convergence Issues

Symptoms of convergence issues: Erratic convergence behavior, often starting from the first SCF step or possibly setting in at a later stage, which produce crazy energy values with "megahartree" jumps. If overlap fitting is on, the following error message may also be encountered: "Error in Cholesky inversion of numerical overlap".

Convergence issues may arise when the chosen grid has difficulties in representing the basis set. This is the "grid equivalent" of a linear dependence problem, as discussed in *Linear Dependence*. It should also be mentioned that the grid-related problem discussed here often goes hand in hand with basis set linear dependence, although not necessarily. The most straightforward way of dealing with this is to increase the size of the integration grid. This, however, is not always desirable or possible.

One way to avoid the Cholesky inversion issue is to turn overlap fitting off (`!NoSFitting`). However, this means that the numerical problems are still present, but are ignored. Due to the fact that overlap fitting operates with the numerical overlap and its inverse, it is more sensitive to linear dependence issues, so turning off the fitting procedure may lead to convergence. This may be a pragmatic — but by no means clean — solution, since it relies on the assumption that the numerical errors are small.

On the other hand, overlap fitting also gives a similar tool to deal with linear dependence issues as the one discussed for basis sets. The eigenvalues of the numerical overlap can be similarly inspected and small values screened out. There is unfortunately no universal way to determine this screening parameter, but see *Linear Dependence* for typical values.

The parameters influencing the method used for inversion and obtaining the fitting matrix are:

```
%method
  SFitInvertType Cholesky    # Cholesky inversion (default)
                 Cholesky_Q  # Cholesky + full Q matrix
                 Diag        # Inversion via diagonalization
                 Diag_Q      # Diag + full Q matrix
  SInvThresh     1e-8        # Inversion threshold for Diag and Diag_Q (default 1e-8)
end
```

By default, the inversion procedure proceeds through Cholesky decomposition as it is the fastest option. Ideally, the overlap matrix is non-singular if the basis set is not linearly dependent. For singular matrices, the Cholesky procedure will fail. It should be noted at this point that the numerical overlap can become linearly dependent even if the overlap of basis functions is not, and so a separate parameter will be needed to take care of grid-related issues. To achieve this, a diagonalization procedure (`Diag`) can be used instead of Cholesky with the corresponding parameter to screen out eigenvectors belonging to eigenvalues below a threshold (`SInvThresh`). For both Cholesky and diagonalization procedures, a "full Q" approach is also available (`Cholesky_Q` and `Diag_Q`), which corresponds to the use of a more accurate (untruncated) fitting matrix.

Treatment of Dispersion Interactions with DFT-D3

Introduction

DFT-D3 is an atom-pairwise (atom-triplewise) dispersion correction which can be added to the KS-DFT energies and gradient [324]:

$$E_{\text{DFT-D3}} = E_{\text{KS-DFT}} + E_{\text{disp}} \tag{7.26}$$

E_{disp} is then the sum of the two- and three-body contributions to the dispersion energy, $E_{\text{disp}} = E^{(2)} + E^{(3)}$. Most important is the two-body term, which is given at long range by:

$$E_{\text{disp}} = -\sum_{A<B}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_{AB}^n} \tag{7.27}$$

where C_n^{AB} denotes the averaged (isotropic) n^{th}-order dispersion coefficient for atom pair *AB* and r_{AB} is their internuclear distance. s_n is a functional-dependent scaling factor (see below). In the general case, an adequate damping function must be employed.

Damping Functions

In order to avoid near-singularities for small r_{AB}, the dispersion contribution needs to be damped at short distances. One possible way is to use rational damping as proposed by Becke and Johnson [90, 424, 425]:

$$E^{(2)} = -\sum_{A<B}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_{AB}^n + f(R_0^{AB})^n} \tag{7.28}$$

with [425]

$$R_0^{AB} = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}} \tag{7.29}$$

and

$$f(R_0^{AB}) = a_1 R_0^{AB} + a_2. \tag{7.30}$$

Damping the dispersion contribution to zero for short ranges (as in Ref. [324]) is also possible:

$$E^{(2)} = -\sum_{A<B}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_{AB}^n} f_{d,n}(r_{AB}) \tag{7.31}$$

with

$$f_{d,n} = \frac{1}{1 + 6(\frac{r_{AB}}{s_{r,n} R_0^{AB}})^{-\alpha_n}} \tag{7.32}$$

Note that the R_0^{AB} used with this damping are from Ref. [324]. For more information on the supported damping functions, see Ref [326]. The recommended variant is that with Becke-Johnson damping and is invoked by the keyword `!D3BJ` or simply `!D3`. The dispersion correction with zero damping is invoked by the keyword `!D3ZERO`.

Three-body term

It is possible to calculate the three-body dispersion contributions with DFT-D3, according to

$$E^{(3)} = -\sum_{A<B<C} \frac{C_9^{ABC}(3\cos\theta_a \cos\theta_b \cos\theta_c + 1)}{(r_{AB} r_{BC} r_{CA})^3} f_{d,(3)}(\overline{r}_{ABC}) \tag{7.33}$$

where θ_a, θ_b and θ_c are the internal angles of the triangle formed by r_{AB}, r_{BC} and r_{CA}. The C_9 coefficient is approximated by

$$C_9^{ABC} \approx -\sqrt{C_6^{AB} C_6^{AC} C_6^{BC}} \tag{7.34}$$

The three-body contribution has a small effect on medium-sized molecules and is damped according to equation (7.33). The damping function $f_{d,(3)}(\overline{r}_{ABC})$ is similar to the one shown in equation (7.32) with $\overline{r}_{ABC}$ being the geometric mean of r_{AB}, r_{BC} and r_{CA}. It can be used with both variants of the $E^{(2)}$ term, although the three-body term itself will always be calculated using the zero damping scheme. Adding the three-body correction has proven to give quite accurate results for the thermochemistry of supramolecular systems[322]. The three-body term is invoked by the keyword `!ABC`, along with either the `!D3ZERO` keyword for zero damping of $E^{(2)}$ or `!D3BJ` for Becke-Johnson damping of $E^{(2)}$.

Options

Note that correcting Hartree-Fock (HF) is only parametrized with BJ-damping. For a constantly updated list of supported functionals, check the website of the Grimme group [329]. If there is a functional on this website that is parametrized, but the parameters are not yet implemented into ORCA, you can specify the parameters manually as shown below (using the respective parameters from [329]). In the same fashion, one could also use one's own parameters, but this is not recommended.

Important: GGA and hybrid functionals should only be used with $s_6 = 1.0$ to ensure asymptotically correct behavior. Double-hybrid functionals already account for parts of the dispersion interaction and should therefore be used with $s_6 < 1.0$. In the `%method` block, it is possible to change the s_6, a_1, s_8, and a_2 parameters for the variant with Becke-Johnson damping:

```
! d3bj b2plyp

%method
  D3S6 0.64
  D3A1 0.3065
  D3S8 0.9147
  D3A2 5.0570
end
```

The variant with zero damping offers the parameters s_6, rs_6, s_8, and α_6.

```
! d3zero blyp

%method
  D3S6     1.0
  D3RS6    1.094
  D3S8     1.682
  D3alpha6 14
end
```

If a geometry optimization is performed (`!OPT`), then the program automatically calls the DFT-D3 gradient. There are also special functional parameters, which were optimized for triple-zeta basis sets. This option is only available with zero damping and can be invoked by the keywords `!D3ZERO D3TZ`. Preliminary results in the SI of Ref. [324] indicate that results are only slightly worse than with quadruple-zeta basis sets using the default parameters. This option should be carefully tested for future use in large computations.

Example input files

Following are some example input files. In this first example, a computation using the DFT-D3 dispersion correction with BJ-damping is run. The use of `!D3BJ` is identical to `!D3` as the BJ-damping is on by default.

```
! pbe svp d3bj

* xyz 0 1
  C     0.000000     0.000000     0.000000
  O     0.000000     0.000000     1.400000
  O     0.000000     0.000000    -1.400000
*
```

The output for the dispersion correction in the ORCA output will look like this:

```
-----------------------------------------------------------------------------
                         DFT DISPERSION CORRECTION

                             DFTD3 V3.1  Rev 1
                         USING Becke-Johnson damping
-----------------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
The PBE functional is recognized
Active option DFTDOPT                   ...         4

molecular C6(AA) [au] = 156.562480


            DFT-D V3
 parameters
 s6 scaling factor         :     1.0000
 a1 scaling factor         :     0.4289
 s8 scaling factor         :     0.7875
 a2 scaling factor         :     4.4407
 ad hoc parameters k1-k3   :    16.0000     1.3333    -4.0000

 Edisp/kcal,au: -0.563071585638  -0.000897311593
 E6   /kcal   :  -0.390909076
 E8   /kcal   :  -0.172162510
 % E8         :  30.575598941

-------------------------   ----------------
Dispersion correction           -0.000897312
-------------------------   ----------------


-------------------------   --------------------
FINAL SINGLE POINT ENERGY      -188.136908447288
-------------------------   --------------------
```

E_{disp} is given as the "`Dispersion correction`" and is automatically included in the final single point energy. As discussed above, the parameters s_6, a_1, s_8, and a_2 may be manually defined by:

```
! pbe svp d3bj

%method
  D3S6 1.0
  D3A1 0.4289
  D3S8 0.7875
  D3A2 4.4407
end

*xyz 0 1
  C     0.000000     0.000000     0.000000
  O     0.000000     0.000000     1.400000
  O     0.000000     0.000000    -1.400000
*
```

This results in the same output as above, but with additional messages that user inputs were found for the parameters:

```
A user input s6-coefficient scaling factor has been recognized
A user input a1-coefficient scaling factor has been recognized
A user input s8-coefficient scaling factor has been recognized
A user input a2-coefficient scaling factor has been recognized
```

The calculation of the same system using zero damping is run by the input:

```
! pbe svp d3zero

*xyz 0 1
  C     0.000000     0.000000     0.000000
  O     0.000000     0.000000     1.400000
  O     0.000000     0.000000    -1.400000
```

(continues on next page)

(continued from previous page)

```
*
```

As stated above, the parameters s_6, rs_6, s_8 and α_6 can be defined by the user. The next example shows this along with the call for the three-body correction (`!ABC`):

```
! pbe svp d3zero abc

%method
  D3S6 1.0000
  D3RS6 1.2170
  D3S8 0.7220
  D3ALPHA6 14.0
end

*xyz 0 1
  C     0.000000     0.000000     0.000000
  O     0.000000     0.000000     1.400000
  O     0.000000     0.000000    -1.400000
*
```

DFT Calculations with the Non-Local, Density Dependent Dispersion Correction (VV10): DFT-NL

Accounting for the missing van der Waals (vdW, dispersion) forces in standard Kohn-Sham Density Functional Theory (DFT) has become essential in many studies of chemical and physical electronic structure problems. Common approaches use atom pair-wise additive schemes such as the popular DFT-D3 [324, 326] method, which is also available in ORCA by invoking the keyword `!D3` (see section *Treatment of Dispersion Interactions with DFT-D3*), but these are not self-consistent, and thus do not correct the MOs or any computed property besides the energies.

A different route is followed by the van der Waals Density Functional (vdW-DF) as pioneered by Langreth and Lundquist [503]. These methods use only the electron density to include such dispersion/correlation effects. The vdW correlation functional VV10 of Vydrov and Van Voorhis [875] currently seems to be the most promising candidate for a general and accurate electronic structure method.

We use the term DFT-NL for any density functional in combination with the non-local part of the VV10 functional with an optimized parameter b, which will be defined below. The performance of these methods has been evaluated in Ref. [400] using the GMTKN30 [307, 308, 309] database and the S66 set [6]. The performance of weak hydrogen bonds were evaluated in Ref. [401].

DFT-NL and DFT-D3 perform very similarly, but NL is to be preferred for metallic systems or when the basic electronic structure changes significantly (e.g. oxidations or ionizations). As a recent example, Janes and Iron showed that for functionals such as wB97X-V, including VV10 correlation results in very high quality reaction barriers when metals are involved [411].

The total exchange-correlation (XC) energy of VV10-type functionals is defined in eq. (7.35). It is composed of standard exchange (X) and correlation (C) parts and the non-local (NL) term, which covers (mainly) long-range dispersive energy:

$$E_{\mathrm{XC}}^{\mathrm{DFT\text{-}NL}} = E_{\mathrm{X}}^{\mathrm{(hybrid)GGA}} + E_{\mathrm{C}}^{\mathrm{GGA}} + E_{\mathrm{C\text{-}NL}}^{\mathrm{VV10}} \tag{7.35}$$

The NL term is given by the following double integral:

$$E_{\mathrm{C\text{-}NL}}^{\mathrm{VV10}} = \int dr \rho(r) \left[\beta + \frac{1}{2} \int dr' \rho(r') \varphi(r, r') \right] \tag{7.36}$$

where ρ is the total electron density, and the definitions of the kernel $\varphi(r, r')$ and β are as follows (in a.u.):

$$\varphi\left(r, r'\right) = -\frac{3}{2gg'\left(g + g'\right)} \tag{7.37}$$

$$\beta = \frac{1}{32} \left[\frac{3}{b^2} \right]^{3/4} \tag{7.38}$$

with

$$g(r) = \omega_0(r) R^2 + \kappa(r)$$
$$R = |r - r'|$$
$$\omega_0(r) = \sqrt{C \left| \frac{\nabla\rho(r)}{\rho(r)} \right|^4 + \frac{4\pi}{3}\rho(r)}$$
$$\kappa(r) = b\frac{3\pi}{2}\left[\frac{\rho(r)}{9\pi}\right]^{1/6}$$

In the original definition, the short-range attenuation parameter b appearing in κ and β was fitted to the S22 set [429] of non-covalent interactions ($b = 5.9$ for the rPW86PBE GGA). The other parameter $C = 0.0093$, appearing in ω_0, determines the long-range behavior, and was set to its original value. Other DFT-NL functionals are constructed analogously. For a detailed discussion of the derivation of the formulas and their physical meaning and basis, see the references given above.

The defined energy of the non-local DFT-NL exchange-correlation functional can be computed in two ways: as a post-SCF addition based on a converged SCF density or in a self-consistent treatment. We take B3LYP as an example.

In our implementation of the non-self-consistent B3LYP-NL functional, a self-consistent B3LYP computation is performed as the first step. In the second step, the optimized electron density from the B3LYP computation is taken as input for the energy calculation of the non-local part. This procedure is invoked by the combination of the keywords `!B3LYP NL`. The use of the keywords `!B3LYP SCNL` would request a self-consistent treatment in which orbitals and density are optimized in the presence of the full B3LYP + VV10 exchange-correlation potential. According to many test calculations, an SCNL treatment is rarely necessary for normal energy evaluations.

The computation of the double integral given in eq. (7.36) requires using an integration grid, just like for normal exchange-correlation functionals. This grid is built similarly to the regular grids available in the ORCA, see Sec. *Details on the numerical integration grids*.

In the following, we compute the energy of the argon dimer at a distance of 3.76 Å with the def2-TZVP basis set and using the B3LYP hybrid functional as an example. Here, we choose the non-self-consistent variant of the DFT-NL dispersion correction.

```
! B3LYP NL
! def2-TZVP def2/JK RIJK

*xyz 0 1
  Ar  0.0 0.0 0.0
  Ar  0.0 0.0 3.76
*
```

The DFT-NL output for this example is shown below:

```
-------------------------------------------------------------------------------
                    post-SCF DFT-NL dispersion correction
-------------------------------------------------------------------------------

SCF Energy:    -1054.960547980
NL    Energy:      0.209449625
SC+NL Energy:  -1054.751098355
NL done in  :              0.1 sec
-------------------------------------------------------------------------------

----------------
TOTAL SCF ENERGY
----------------

(...)

DFT components:
```

(continues on next page)

(continued from previous page)

```
N(Alpha)           :       17.999996327923 electrons
N(Beta)            :       17.999996327923 electrons
N(Total)           :       35.999992655845 electrons
E(X)               :      -47.880990295917 Eh
E(C)               :       -1.761924720763 Eh
NL Energy, E(C,NL) :        0.209449624689 Eh
E(XC)              :      -49.433465391991 Eh
```

Here, we find the B3LYP total energy ("`SCF Energy`"), the non-local contribution ("`NL Energy`"), and their sum ("`SC+NL Energy`"), which is the final total energy. In the "`DFT components`" section, the non-local contribution is listed separately ("`NL Energy, E(C,NL)`") in order to be consistent with the `!SCNL` output.

In the current version of ORCA, there are several functionals with pre-fitted b parameters (the parameter C is not changed), available for use by the keyword `!DF NL` or `!DF SCNL`, where DF stands for one of the following density functional keywords:

DF keyword
(meta-)GGA
BLYP
PBE
REVPBE
RPBE
RPW86PBE
SCANfunc
TPSS
hybrid
B3LYP
B3LYP/G
B3P86
B3PW91
mPW1PW
PBE0
PW1PW
PW6B95
REVPBE0
REVPBE38
TPSSh
TPSS0
R2SCANh (see [128])
R2SCAN0 (see [128])
R2SCAN50 (see [128])
range-separated hybrid
WR2SCAN50 (see [890])
double-hybrid
B2PLYP (see [51])
DSD-BLYP (see [905]). Same b used for DSD-BLYP/2013
DSD-PBEP86 (see [905]). Same b used for DSD-PBEP86/2013
PWPB95 (see [905])
PR2SCAN50 (see [890])
KPR2SCAN50 (see [890])
PR2SCAN69 (see [890])
range-separated double-hybrid
WPR2SCAN50 (see [890])

Additionally, one can include the non-local term in Hartree-Fock (HF) with a parameter of $b = 3.9$ using the keywords "`!HF NL`".

The B97-V Family

Head-Gordon's ωB97X-V functional [554] is a reparametrized version of the range-separated ωB97X, which makes use of the non-local VV10 kernel to capture London-dispersion effects ($b = 6.0$ and $C = 0.01$; note that C is, unlike for the other functionals, changed for ωB97X-V). The keyword `!wB97X-V` evaluates the VV10 kernel in a post-SCF way by default (i.e. only the semi-local exchange-correlation part is converged self-consistently and the resulting density is then used to assess the VV10-type energy contribution). A recent study showed that this saves computer time and does not have a significant effect on the resulting relative energies [602]. The keyword "`!NL`" does not have to be specified in this case as the VV10 kernel is invoked automatically. If a user wishes to carry out fully self-consistent calculations with ωB97X-V, the "`!SCNL`" keyword must be specified.

The range-separated meta-GGA hybrid ωB97M-V [556] and the meta-GGA B97M-V [555] are also available (via keywords `!wB97M-V` and `!B97M-V`, respectively). In the spirit of ωB97X-V, the VV10 (NL) correction is called automatically in the post-SCF way by default.

Changing the b, C and Scaling Parameters

All density functionals that are available in ORCA, but for which no b parameter is available, can be used with the DFT-NL method by providing a value for the parameter b as shown here:

```
%method
  NLb 5.0
end
```

The other parameter C, appearing in ω_0, may also be changed, as shown in the following example.

```
! B3LYP NL
! def2-TZVP def2/JK RIJK nososcf nopop

%method
  NLC    0.0083
end

*xyz 0 1
  Ar  0.0 0.0 0.0
  Ar  0.0 0.0 3.76
*
```

Of course, both parameters may be changed by using the `NLb` and `NLC` keywords in the `%method` block at the same time.

Note

In order to improve the scaling for larger systems, a distance cutoff was also introduced, controlled by the `vdWdistTCUT` flag in the `%method` block. The default value is 10 Å, so two grid points more than 10 Å away from each other do not correlate via the VV10 potential. This is already very conservative and has practically zero effect on the final energy, but can be altered if needed.

For methods where the long-range correlation is already partly covered, e.g. in double-hybrids, the NL energy can be scaled down using the `NLScal` parameter as shown below. By default, the scaling factor is 1.0 if not otherwise specified by the employed functional.

```
%method
  NLScal 0.5
end
```

Self-consistent Computations with the DFT-NL Dispersion Correction

Self-consistent calculations with the DFT-NL dispersion correction are possible by using the keyword `!SCNL` in combination with one of the available density functionals. Note that due to technical reasons, self-consistent calculations are not possible with the Hartree-Fock method.

For example, we can use the B3LYP hybrid functional with the self-consistent DFT-NL variant by the following input:

```
! B3LYP SCNL
! def2-TZVP def2/JK RIJK

*xyz 0 1
  Ar  0.0 0.0 0.0
  Ar  0.0 0.0 3.76
*
```

The output is the same as a normal SCF calculation, but after convergence, a line with `NL Energy, E(C,NL)` is printed under "`DFT components`", as it was for the post-SCF DFT-NL variant.

- Analytic gradients are already available, thus geometry optimizations with numerical frequencies can be computed using these functionals.
- TD-DFT calculations are not yet available.
- Any calculation that requires second derivatives of the NL functional is not yet possible. These are needed for real type perturbations in the CP-SCF solutions, e.g. for analytic Hessians, dipole polarizabilities, and double-hybrid gradients.
- Strictly imaginary perturbations such as NMR shielding and EPR g-tensors (both also with GIAOs), and hyperfine couplings are available.

DFT and HF Calculations with the Geometrical Counterpoise Correction: gCP

The central idea of the gCP correction [476] is to add a semi-empirical correction ΔE_{gCP} to the energies of molecular systems that removes artificial overbinding effects from BSSE (see section *Counterpoise Correction*). gCP uses atomic corrections and therefore also has the ability to correct for intramolecular BSSE. The parametrization is constructed such that it approximates the Boys and Bernadi counterpoise (CP) correction ΔE_{CP} in the intermolecular case. That is,

$$\Delta E_{CP} \approx \Delta E_{\text{gCP}} \tag{7.39}$$

where e.g. for a complexation reaction $A + B \rightarrow C$, our correction is given by

$$\Delta E_{\text{gCP}} = E_{\text{gCP}}(C) - E_{\text{gCP}}(A) - E_{\text{gCP}}(B) \tag{7.40}$$

In practice, E_{gCP} can be simply added to the HF/DFT energy

$$E_{\text{total}} = E_{\text{HF/DFT}} + E_{\text{gCP}} \tag{7.41}$$

which is done automatically in ORCA (the `FINAL SINGLE POINT ENERGY` includes the gCP correction).

The central equation of the gCP correction over all atoms N reads:

$$E_{\text{gCP}} = \sigma \cdot \sum_{a}^{N} \sum_{b \neq a}^{N} e_a^{\text{miss}} \cdot f_{\text{dec}}(R_{ab}) \tag{7.42}$$

where the energy e_a^{miss} is a measure of the incompleteness of the chosen target basis set (which is typically small) and $f_{\text{dec}}(R_{ab})$ is a decay function that depends on the interatomic distance R_{ab}. The scaling factor σ is one of 4 parameters needed for every `method/basis set` combination. More details on this can be found in the original gCP paper [476].

Analytical gradients with gCP are available for geometry optimization. Due to its semi-empirical nature, the correction is calculated within seconds, even for very large systems.

The gCP correction can be invoked by using the `!GCP(level)` keyword, where `level` is a compound of the method (`HF` or `DFT`) and the `basis set` keyword. See Table 7.7 for the available basis sets and the corresponding keywords. For example, gCP can be invoked in a B3LYP calculation with the def2-SV(P) basis set using the input:

```
! B3LYP def2-SV(P) GCP(DFT/SV(P))
*xyzfile 0 1 example.xyz
```

Table 7.7: Overview of parametrized basis sets. The `level` keyword in `!GCP(level)` is a compound of `HF` or `DFT` and the `basis set` keyword. Valid inputs are, for example: `!GCP(HF/MINIS)`, `!GCP(DFT/LANL)`, `!GCP(HF/TZ)`, `!GCP(DFT/631GD)`, ...

parametrized basis set	HF	DFT	`basis set`
MINIS	yes	yes	MINIS
SV	yes	yes	SV
6-31G(d)	yes	yes	631GD
6-31G(d) + LANL2DZ (Sc-Zn)	no	yes	LANL
def2-SV(P)	no	yes	SV(P)
def2-SV(P/h,c)	no	yes	SVX or SV(P/H,C)
def2-SVP	yes	yes	SVP
def2-TZVP	yes	yes	TZ

At all print levels, warnings from the gCP program are printed. Using the default print level, the only additional output is the final gCP correction before `FINAL SINGLE POINT ENERGY`. Using `!LargePrint` or `%output Print[P_gCP] 2 end` states the gCP `level`, the 4 parameters mentioned above, and the computed correction. A larger print level can be specified to get more details (more information on this below).

Several warnings and notices may be issued. Elements of the 5th and higher periods are treated as their 4th period analogs — e.g. Sn is treated with the same parameters as Ge. If this is the case, a note is printed. Another note is printed if there is a mismatch between the basis used for the SCF calculation and that of the requested gCP calculation. For example, the following input with tetramethyltin

```
! HF def2-SVP GCP(HF/MINIS)
*xyzfile 0 1 tetramethyltin.xyz
```

should use the parameters of Ge in place of Sn and there should be a mismatch between the basis set in ORCA (def2-SVP) and gCP (MINIS). Sure enough, the output is as follows:

```
 ** NOTE ** -> element sn will be treated as ge
NOTIFICATION: Different basis set in ORCA and otool_gcp:
ORCA: 142  gCP: 32
If you are NOT using ECPs, check your basis set inputs!
------------------   -----------------
gCP correction             0.073031339
------------------   -----------------
```

A mismatch between the basis sets used is allowed since a minor mismatch may only result in a small error. One should still be careful with such results; use your own judgment! This also allows gCP in calculations that use an unparametrized basis set. However, in this case, the number of basis functions and exponents should be very similar!

It should be noted that some elements are not parametrized, depending on the gCP `level` used. If only a few atoms in a large molecule are treated inaccurately or not at all, the error is expected to be small. To check all parameters used and the individual atomic contributions, specify the print level `%output Print[P_gCP] 3 end`. For example, the above tetramethyltin input with this print level has the following output:

```
--------------------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
                g C P  - geometrical counterpoise correction

-----------------------------------------------------------------------------
Method:  hf/minis

 ** NOTE ** -> element sn will be treated as ge
Parameters:   sigma        eta         alpha        beta
              0.1290      1.1526      1.1549       1.1763
Nbf:   32
NAtoms:   17

gCP element parameters:
  elem   emiss   nbas    elem   emiss   nbas    elem   emiss   nbas
    h    0.04240    1      he   0.02832    1      li   0.25266    2
    be   0.19720    2      b    0.22424    5      c    0.27995    5
    n    0.35791    5      o    0.47901    5      f    0.63852    5
    ne   0.83235    5      na   1.23292    6      mg   1.34339    6
    al   1.44828    9      si   1.61336    9      p    1.76814    9
    s    1.99201    9      cl   2.23311    9      ar   2.49323    9
    k    3.02964   10      ca   3.38998   10      sc   0.00000    0
    ti   0.00000    0      v    0.00000    0      cr   0.00000    0
    mn   0.00000    0      fe   0.00000    0      co   0.00000    0
    ni   0.00000    0      cu   0.00000    0      zn   0.00000    0
    ga   0.00000    0      ge   0.00000    0      as   0.00000    0
    se   0.00000    0      br   0.00000    0      kr   0.00000    0

      #     ON  sites    Nvirt           Emiss    BSSE [kcal/mol]
      1      6     15      2.0          0.2799         10.0090
      2     32     16    -16.0          0.0000          0.0000
      3      6     15      2.0          0.2799         10.0084
      4      6     15      2.0          0.2799         10.0095
      5      6     15      2.0          0.2799         10.0083
      6      1     15      0.5          0.0424          0.4827
      7      1     15      0.5          0.0424          0.4828
      8      1     15      0.5          0.0424          0.4828
      9      1     15      0.5          0.0424          0.4827
     10      1     15      0.5          0.0424          0.4827
     11      1     15      0.5          0.0424          0.4827
     12      1     15      0.5          0.0424          0.4828
     13      1     15      0.5          0.0424          0.4827
     14      1     15      0.5          0.0424          0.4828
     15      1     15      0.5          0.0424          0.4827
     16      1     15      0.5          0.0424          0.4827
     17      1     15      0.5          0.0424          0.4827
Egcp:       0.0730313386 a.u.
NOTIFICATION: Different basis set in ORCA and otool_gcp:
ORCA: 142  gCP: 32
If you are NOT using ECPs, check your basis set inputs!
------------------   -----------------
gCP correction             0.073031339
------------------   -----------------
```

From this, it can be seen that the Sn atom (atom 2 in the list of atomic contributions) gives no contribution because its `Emiss` is zero (unparametrized for the given gCP `level`). This is confirmed by looking at the `gCP element parameters` section, which lists the `emiss` of Sn as zero for `!GCP(HF/MINIS)`. Rerunning this example with `!GCP(HF/SVP)` now gives a contribution for Sn, as seen by the following output. Note that this calculation also has a much smaller basis set mismatch, and so should be the more accurate gCP correction of the two.

```
-----------------------------------------------------------------------------

                g C P  - geometrical counterpoise correction
```

(continues on next page)

(continued from previous page)

```
-------------------------------------------------------------------------------
Method:  hf/svp

 ** NOTE ** -> element sn will be treated as ge
Parameters:   sigma        eta         alpha        beta
              0.2054      1.3157      0.8136       1.2572
Nbf:   148
NAtoms:   17

gCP element parameters:
  elem   emiss   nbas    elem   emiss   nbas    elem   emiss   nbas
    h    0.00811    5      he   0.00805    5      li   0.11358    9
    be   0.02837    9      b    0.04937   14      c    0.05538   14
    n    0.07278   14      o    0.10031   14      f    0.13327   14
    ne   0.17360   14      na   0.18114   15      mg   0.12556   18
    al   0.16719   18      si   0.14984   18      p    0.14540   18
    s    0.16431   18      cl   0.18299   18      ar   0.20567   18
    k    0.20096   24      ca   0.29966   24      sc   0.32599   31
    ti   0.30549   31      v    0.29172   31      cr   0.29380   31
    mn   0.29179   31      fe   0.29673   31      co   0.30460   31
    ni   0.24204   31      cu   0.35419   31      zn   0.35072   31
    ga   0.35002   32      ge   0.34578   32      as   0.34953   32
    se   0.36731   32      br   0.38201   32      kr   0.39971   32

      #     ON  sites    Nvirt           Emiss    BSSE [kcal/mol]
      1      6     16     11.0          0.0554          2.5093
      2     32     16     16.0          0.3458          6.8274
      3      6     16     11.0          0.0554          2.5092
      4      6     16     11.0          0.0554          2.5094
      5      6     16     11.0          0.0554          2.5092
      6      1     16      4.5          0.0081          0.1703
      7      1     16      4.5          0.0081          0.1703
      8      1     16      4.5          0.0081          0.1703
      9      1     16      4.5          0.0081          0.1703
     10      1     16      4.5          0.0081          0.1703
     11      1     16      4.5          0.0081          0.1703
     12      1     16      4.5          0.0081          0.1703
     13      1     16      4.5          0.0081          0.1703
     14      1     16      4.5          0.0081          0.1703
     15      1     16      4.5          0.0081          0.1703
     16      1     16      4.5          0.0081          0.1703
     17      1     16      4.5          0.0081          0.1703
Egcp:       0.0301322002 a.u.
NOTIFICATION: Different basis set in ORCA and otool_gcp:
ORCA: 142  gCP: 148
If you are NOT using ECPs, check your basis set inputs!
------------------   -----------------
gCP correction             0.030132200
------------------   -----------------
```

The gCP input can also be defined in the %method block of the input:

```
%method
  DoGCP       true/false    # turn gCP on/off
  GCPMETHOD   "method"      # define method string for otool_gcp, e.g. "dft/svp"
  GCP.D3MINIS true/false    # use special DFT-D3 refit for HF/MINIS (default=true)
end
```

General advice:

- Small basis sets show not only a large BSSE, but general shortcomings. These effects are not always clearly

distinguishable.

- If computationally affordable, large basis sets (triple-ζ and higher) are always preferable for a given system.
- gCP only makes sense for somewhat large molecules
- gCP should always be applied together with a dispersion correction for DFT: gCP-D3 is well tested, but gCP-NL also works well (see sections *Treatment of Dispersion Interactions with DFT-D3* and *DFT Calculations with the Non-Local, Density Dependent Dispersion Correction (VV10): DFT-NL*)

Note

- `!GCP(HF/MINIS)` automatically sets the refitted D3 parameter, as proposed in the original gCP paper.
- The gCP method is implemented via an external tool called **otool_gcp**, which is based on the original Fortran program used in the publication. Thus, the otool_gcp binary can also be called directly via the command line (`otool_gcp -h` gives an overview of the options).
- It is also possible to read an external parameter file (`$HOME/.gcppar`) using the `!GCP(FILE)` keyword. For further information, please look at the manual for the gcp program as provided by Prof. S. Grimme[8].
- During the calculation, some temporary output files are written by ORCA: `BASENAME.gcp.in.tmp` and `BASENAME.gcp.out.tmp` will contain the input for otool_gcp and its output, respectively.
- It has been demonstrated that the popular combination of B3LYP with 6-31G(d) can be strongly improved using DFT-D3 and gCP [475]. For convenience, the keyword `!B3LYP-gCP-D3/6-31G*` has been defined. This is equivalent to `!B3LYP 6-31G* D3BJ GCP(DFT/631GD)`.

HF-3c: Hartree-Fock with three corrections

HF-3c is a fast Hartree-Fock based method developed for computation of structures, vibrational frequencies and non-covalent interaction energies in huge molecular systems [833]. The starting point for evaluating the electronic energy is a standard HF calculation with a small Gaussian AO basis set. The used so-called MINIX basis set consists of different sets of basis functions for different groups of atoms as shown in table Table 7.8.

Table 7.8: Composition of the MINIX basis set.

element	basis
H-He, B-Ne	MINIS
Li-Be	MINIS+1(p)
Na-Mg	MINIS+1(p)
Al-Ar	MINIS+1(d)
K-Zn	SV
Ga-Kr	SVP
Rb-Rn	def2-SVP with Stuttgart-Dresden ECPs

Three terms are added to correct the HF energy $E_{\text{tot}}^{\text{HF/MINIX}}$ in order to include London dispersion interactions, to account for the BSSE and to correct for basis set deficiencies, i.e. overestimated bond lengths. The corrected total energy is therefore calculated as

$$E_{\text{tot}}^{\text{HF-3c}} = E_{\text{tot}}^{\text{HF/MINIX}} + E_{\text{disp}}^{\text{D3(BJ)}} + E_{\text{BSSE}}^{\text{gCP}} + E_{\text{SRB}}. \tag{7.43}$$

The first correction term $E_{\text{disp}}^{\text{D3(BJ)}}$ is the atom-pair wise London dispersion energy from the D3 dispersion correction scheme[324] applying Becke-Johnson (BJ) damping [90, 424, 425] (see section *Treatment of Dispersion Interactions with DFT-D3*). The second term $E_{\text{BSSE}}^{\text{gCP}}$ denotes the geometrical counterpoise (gCP) correction [476] to treat the BSSE (see section *DFT and HF Calculations with the Geometrical Counterpoise Correction: gCP*). For the HF-3c method, the three usual D3 parameters s_8, a_1 and a_2 were re-fitted using reference interaction energies of the complexes of the S66 test set [6]. This results in $s_8 = 0.8777$, $a_1 = 0.4171$ and $a_2 = 2.9149$. The parameter

[8] http://www.thch.uni-bonn.de/tc/

s_6 was set to unity as usual to enforce the correct asymptotic limit and the gCP correction was already applied in this fitting step.

The last term E_{SRB} is a short-ranged correction to deal with basis set deficiencies which occur when using small or minimal basis sets. It corrects for systematically overestimated covalent bond lengths for electronegative elements and is calculated as a sum over all atom pairs:

$$E_{\text{SRB}} = -s \sum_{A}^{\text{Atoms}} \sum_{B \neq A}^{\text{Atoms}} (Z_A Z_B)^{3/2} \exp(-\gamma (R_{AB}^{0,\text{D3}})^{3/4} R_{AB})$$

Here, $R_{AB}^{0,\text{D3}}$ are the default cut-off radii as determined *ab initio* for the D3 scheme [324] and Z_A, Z_B are the nuclear charges. This correction is applied for all elements up to argon. The empirical fitting parameters $s = 0.03$ and $\gamma = 0.7$ were determined to produce vanishing HF-3c total atomic forces for B3LYP-D3(BJ)/def2-TZVPP equilibrium structures of 107 small organic molecules. More details can be found in the original publication [833].

The HF-3c method can only be invoked with a simple keyword:

```
! HF-3c
```

`! HF-3c` is a compound keyword and equals `! HF MINIX D3BJ GCP(HF/MINIX) PATOM`, hence no basis set etc. has to be specified. The `PATOM` guess is selected since the grid construction for the default guess can take more time than an actual SCF step. The guess can only be overwritten manually in the %method section.

The default mode for the integral handling is set to `Conventional`. The storing of the two-electron integrals on disk or in memory if possible leads to large computational savings. In case one want to use the `Direct` mode, this has to be specified in the %scf input section:

```
%scf
  SCFmode Direct
end
```

The output gives the used parameters and the correction itself for D3 and gCP separately. As the SRB correction is also calculated with the otool_gcp, the results are given in the gCP output section. The `Total correction to HF/MINIX` is the sum of all three corrections (D3, gCP and SRB) and `FINAL SINGLE POINT ENERGY` is the total HF-3c energy as given in equation (7.43).

```
-------------------------------------------------------------------------------
                         DFT DISPERSION CORRECTION

                              DFTD3 V2.1  Rev 6
                         USING Becke-Johnson damping
-------------------------------------------------------------------------------
The default Hartree-Fock  is recognized
Active option DFTDOPT                   ...         4

molecular C6(AA) [au] = 1689.256597


            DFT-D V3
 parameters
 using HF/MINIX parameters
 s6 scaling factor         :     1.0000
 a1 scaling factor         :     0.4171
 s8 scaling factor         :     0.8777
 a2 scaling factor         :     2.9149
 ad hoc parameters k1-k3   :    16.0000     1.3333    -4.0000

 Edisp/kcal,au: -32.163184627631  -0.051255291794
 E6   /kcal   : -18.007221978
 E8   /kcal   : -14.155962649
 % E8         :  44.012938437
```

(continues on next page)

(continued from previous page)

```
-------------------------   ----------------
Dispersion correction           -0.051255292
-------------------------   ----------------


-------------------------------------------------------------------------------

                 g C P  - geometrical counterpoise correction

-------------------------------------------------------------------------------
Method:  hf/minix

Parameters:   sigma        eta         alpha        beta
              0.1290      1.1526      1.1549       1.1763
Egcp:       0.0723150636 a.u.
Ebas:      -0.0636976872 a.u.
------------------   -----------------
gCP+bas correction         0.008617376
------------------   -----------------
----------------------------   ----------------
Total correction to HF/MINIX       -0.042637915
----------------------------   ----------------


-------------------------   --------------------
FINAL SINGLE POINT ENERGY      -163.002895262171
-------------------------   --------------------
```

For the elements up to Xe, the default initial guess is a Hückel guess. Beyond Xe, the guess mode is changed to `HCORE` since no Hückel parameters for the respective ECP bases are available and other models are not implemented at the moment. For calculations with only lighter elements and therefore no ECPs, the ECP printouts in the output file can be ignored.

PBEh-3c: A PBE hybrid density functional with small AO basis set and two corrections

PBEh-3c is a highly efficient electronic structure approach performing particularly well in the optimization of geometries and for interaction energies of non-covalent complexes.[325] Here, a global hybrid variant of the Perdew-Burke-Ernzerhof (PBE) functional with a relatively large amount of non-local Fock-exchange (42%) is employed with a valence-double-zeta Gaussian AO basis set (def2-mSVP). Basis set superposition errors (BSSE) and London dispersion effects are accounted for by the gCP and D3 schemes, respectively (see above). The basis set is constructed such that:

Table 7.9: Composition of the def2-mSVP basis set.

element	basis
H	def2-SV(P) (α scaled by 1.2)
He	def2-SVP(-p)
Li-Be,Na-Kr	def2-SV(P)
B,Ne	Ahlrichs' DZ + Polarization from def2-SVP
C-F	Ahlrichs' DZ + Polarization from 6-31G*
Rb-Rn	def2-SVP with Stuttgart-Dresden ECPs

For inter- and intramolecular BSSE the gCP expression from Eq. (7.42) is used but with a damping function (similar to the zero-damping in Eq. (7.32)). This damping improves the thermochemistry of the method significantly compared with the non-damped version. London dispersion effects are accounted for by the DFT-D3 (BJ-damping) scheme including the three-body term. Compared to the related HF-3c approach, the PBEh-3c is somewhat more costly, however, it yields much better geometries. These are roughly of MP2-quality (or even better for non-covalent structures) but may be computed at much lower cost. Due to the moderate amount of non-local Fock exchange, the

method is less prone to self-interaction errors (as in GGAs) but still applicable in cases when Hartree-Fock fails (strongly correlated systems).

The PBEh-3c method may be invoked with the simple keyword:

```
! PBEh-3c
```

Identical to HF-3c, the default initial guess for all elements up to Xe is a Hückel guess. Beyond Xe, the guess mode is changed to HCORE. For calculations with only lighter elements and therefore no ECPs, the ECP printouts in the output file can be ignored.

Recently, a new composite 'low-cost' method for accurate thermochemistry, structures, and noncovalent interactions specifically also for transition metal chemistry and other stronger correlated systems was implemented. As it is based on the B97 GGA including D3 with three-body contribution, a short range bond length correction, and a modified, stripped-down triple-ζ basis, def2-mTZVP, the computational cost of this method termed B97-3c are between that of HF-3c and PBEh-3c (for large systems roughly two times more expensive than HF-3c). It is invoked with a simple keyword analogously to the latter methods. Some more detailed information can be found in Ref. [121] .

r^2SCAN-3c: A robust "Swiss army knife" composite electronic-structure method

The r^2SCAN-3c composite method[333] is available as robust "Swiss army knife" electronic structure method for thermochemistry, geometries and non-covalent interactions and has shown in preliminary tests consistent performance for both open and closed shell transition metal complexes. It is based on the r^2SCAN[282] meta-GGA combined with the D4 dispersion correction[244] and the geometrical counter poise-correction[476]. The modified triple-ζ basis set, def2-mTZVPP, is larger and more consistent for the light main-group elements and almost as computationally efficient as the def2-mTZVP basis set of B97-3c. The computational cost of r^2SCAN-3c is slightly larger than B97-3c. It is invoked with the simple keyword

```
! r2SCAN-3c
```

ωB97X-3c: A composite range-separated hybrid DFT method with a molecule-optimized polarized valence double-ζ basis set

The ωB97X-3c composite method[540] is based on the ωB97X-V functional and combines a tailored and molecule-optimized polarized valence double-ζ (vDZP) basis set and a specifically adapted D4 dispersion correction. The vDZP basis set employs large-core ECPs and shows only very small basis set superposition and incompleteness errors compared to conventional double-ζ basis sets. In thorough tests on standard benchmarks sets, the ωB97X-3c method was shown to be on par with well-performing hybrid DFT methods in a quadruple-ζ basis set at a fraction of their computational cost. ωB97X-3c is consistently available for all elements up to Rn (Z = 1–86).

It is invoked with the simple keyword:

```
! wB97X-3c
```

The vDZP basis set alone is utilized as follows (note that the corresponding large-core ECPs are called automatically):

```
! vDZP
```

7.4.3 Semiempirical Methods

The present version of ORCA has inherited the capability of doing semiempirical calculations from the earlier versions. A number of methods based on the "neglect of differential overlap" [691, 778] are currently implemented for energies and analytic gradients (for geometry optimization).

```
%method
  Method  CNDO
          INDO
          NDDO
  # for Method=CNDO
  Version CNDO_1
          CNDO_2
          CNDO_S
  # for Method=INDO
  Version INDO_1
          INDO_2
          INDO_S
          ZINDO_1
          ZINDO_2
          ZINDO_S
  # for Method=NDDO
  Version ZNDDO_1
          ZNDDO_2
          MNDO
          AM1
          PM3
end
```

The methods MNDO [206, 207, 847], AM1 [208] and PM3 [821] are available for main group elements only and arise from the work of the Dewar group. They have been optimized to reproduce molecular structure and energetics. The older CNDO/1,2 and INDO/1,2 were developed by the Pople group [61, 175, 176, 177, 692, 695, 696, 748, 749] and were designed to roughly mimic minimal basis *ab initio* calculations. The methods of the Zerner group (ZINDO/1,2 and ZINDO/S) are closely related to the older methods but have been well parameterized for transition metals too [36, 37, 38, 63, 183, 466, 719, 907, 908, 909, 911]. ZINDO/1 (and less so ZINDO/2) are suitable for geometry optimization. ZINDO/S gives good results for electronically excited states at moderate configuration interaction levels and is also successful for the calculation of electron and spin distributions in large transition metal complexes [37, 183, 466, 907, 908, 909]. The ZNDDO/1,2 methods have been implemented into ORCA as straightforward extensions of the corresponding INDO methods without changing any parameter. However, the methods benefit from the somewhat more accurate representation of the Coulomb interaction within the NDDO approximation [431, 633]. The preliminary experience with these methods is that they are better than the corresponding INDO methods for calculation of transition metal complex structures but on the whole have also similar deficiencies.

The analytic gradients are available for all of these methods and can be used to produce reasonable molecular structures at low computational cost or to get preliminary insight in the behavior of the system under investigation[9].

There is also a mechanism for simplified input. Instead of giving values for `Method` and `Version` separately you can also assign the value that would normally belong to `Method` to `Version`. The program will recognize that and assign the correct values to both `Method` and `Version`.

```
%method
  # shortcut to Method=NDDO; and Version=AM1;
  Method  AM1
end
```

- If you want you can also combine semiempirical methods with MP2 (energies only). For example use `Method =AM1;` and `DoMP2 =true;` It is questionable if this makes the results of semiempirical calculations any better but at least it is possible in ORCA.

[9] However, do not try to use ZINDO/S (or CNDO/S) for structure optimizations - it does not make sense and will lead to disastrous results because there is no accurate representation of nuclear repulsion in these methods.

You can change the built-in semiempirical parameters in a straightforward fashion. For example:

```
! ZINDO/S TightSCF DIIS NoMOPrint

%cis    NRoots 20
        MaxDim 3 # Davidson expansion space = MaxDim * NRoots
        end

%ndoparas  P[6,25] 20
           P[6,26] 20
           end
```

The %ndoparas block is there in order to let you input your favorite personal parameters. The "molecular" parameters are set using "INTFA" ("interaction factors");

```
%ndoparas  INTFA[PP_PI] 0.585
  # The interaction factors exist for
  #  ss_sigma
  #  sp_sigma
  #  sd_sigma
  #  pp_sigma
  #  pd_sigma
  #  dd_sigma
  #  pp_pi
  #  pd_pi
  #  dd_pi
  #  dd_delta
  # the parameter entering the Coulomb integrals
  # in INDO/S
  FGAMMA 1.2
  end
```

All atomic parameters are collected in an array "P". The first index is the atomic number of the element whose parameters you want to change. The second index identifies which parameter. The list of parameters follows below. Most of them will only be interesting for expert users. The most commonly modified parameters are the Beta's (number 25 through 28). Note that most programs require a negative number here. In ORCA the resonance integrals are defined in a way that makes the Beta's positive.

```
# core integrals (in eV)
US      0
UP      1
UD      2
UF      3
# Basis set parameters (double-zeta for generality)
NSH     4  # number of shells for the element
NZS     5  # number of Slater type orbitals for the s shell
ZS1     6  # first exponent
ZS2     7  # second exponent
CS1     8  # first contraction coefficient
CS2     9  # second contraction coefficient
NZP    10  # number of Slater type orbitals for the p shell
ZP1    11  # ...
ZP2    12
CP1    13
CP2    14
NZD    15  # number of Slater type orbitals for the d shell
ZD1    16  # ...
ZD2    17
CD1    18
CD2    19
NZF    20  # number of Slater type orbitals for the f shell
ZF1    21  # ...
```

(continues on next page)

(continued from previous page)

```
ZF2    22
CF1    23
CF2    24
# Resonance integral parameters (in eV)
BS     25  # s shell beta
BP     26  # p shell beta
BD     27  # d shell beta
BF     28  # f shell beta
# Number of electrons in the g.s.
NEL    29  # total number of electrons (integer)
NS     30  # fractional occupation number of the s shell
NP     31  # fractional occupation number of the p shell
ND     32  # fractional occupation number of the d shell
NF     33  # fractional occupation number of the f shell
# The one center repulsion (gamma) integrals (in eV)
GSS    34
GSP    35
GSD    36
GSF    37
GPP    38
GPD    39
GPF    40
GDD    41
GDF    42
GFF    43
# The Slater Condon parameters (in eV)
F2PP   44
F2PD   45
F2DD   46
F4DD   47
G1SP   48
G1PD   49
G2SD   50
G3PD   51
R1SPPD 52
R2SDPP 53
R2SDDD 54
# The nuclear repulsion parameters for Dewar type models
NR1    55
NR2    56
NR3    57
NR4    58
NR5    59
NR6    60
NR7    61
NR8    62
NR9    63
NR10   64
NR11   65
NR12   66
NR13   67
# The nuclear attraction/repulsion parameter for MNDO/d
RHO    68
# Spin orbit coupling parameters
SOCP   69 # SOC for the p shell
SOCD   70 # SOC for the d shell
SOCF   71 # SOC for the f shell
```

Semi-empirical tight-binding methods: Grimme's GFN0-xTB, GFN-xTB and GFN2-xTB

ORCA is interfaced to the XTB tool by Grimme and coworkers, allowing the user to request all kinds of calculations using the popular GFN0-xTB, GFN-xTB and GFN2-xTB Hamiltonians.[70, 332, 698] From the technical side, the user has to provide the executable provided by the Grimme group. The xtb program package can be obtained free of charge from https://github.com/grimme-lab/xtb/releases and detailed information on the usage of the `xtb` standalone program and its utilities can be found at https://xtb-docs.readthedocs.io/en/latest/contents.html. Only the file `bin/xtb` is used by ORCA . The user should copy this file into the directory where the other ORCA executables are located, and rename it as `otool_xtb`.

Please use the 6.7.1 version (or any later version) of xtb; older versions are not fully compatible with ORCA anymore or are missing features, for example it may not be possible to invoke the solvation model! Additionally, Windows users should copy `libiomp5md.dll` from the XTB directory to the ORCA directory.

XTB is invoked by the following keywords:

```
! XTB0  # for GFN0-xTB. Synonym: GFN0-XTB
! XTB1  # for GFN1-xTB. Synonym: GFN-XTB
! XTB2  # for GFN2-xTB. Synonym: GFN2-XTB
! XTBFF # for GFN-FF. Synonym: GFN-FF
```

The following methods can be used in conjunction with XTB:

- Single Point Energy
- Energy and Gradient
- Optimization, using all kinds of constraints, relaxed surface scans, etc.
- Nudged-Elastic Band calculations
- Numerical Frequency Calculations
- Intrinsic Reaction Coordinate
- Molecular Dynamics Calculations
- QM/MM calculations

Note

- XTB0 is a non-self-consistent tight-binding method, and as such, its accuracy is generally inferior to XTB1 and XTB2 (and sometimes even XTBFF), despite that it is a few times faster than XTB1 and XTB2. From our experience, we only recommend XTB0 when both XTB1 and XTB2 exhibit qualitative failures for the system of interest.
- Please note that XTB0, XTB1 and XTB2 can also be used for the initial path generation or for the calculation of an initial TS structure on XTB level, both as input for the subsequent NEB calculation on a higher level of theory. For more details, please consult section *Nudged Elastic Band Method*.

Solvation

Three implicit solvation models can be requested in XTB calculations: (1) the analytical linearized Poisson-Boltzmann (ALPB) solvation model, (2) the domain decomposition COSMO (ddCOSMO)[136], and (3) the extended conductor-like polarizable continuum model (CPCM-X).[814] These three models can be requested via the following tags in the simple input

```
! ALPB(solvent) # use ALPB

or

! DDCOSMO(solvent) # use ddCOSMO
```

(continues on next page)

(continued from previous page)

```
or

! CPCMX(solvent) # use CPCMX
```

where `solvent` is any of the solvents in Table 7.29.

Keywords for XTB

A list of additional keywords for XTB is detailed here:

```
! XTB
! ALPB(water)       # use ALPB for implicit solvation, solvent water,
! DDCOSMO(water)    # use ddCOSMO for implicit solvation, solvent water,
! CPCMX(water)      # use CPCMX for implicit solvation, solvent water,
                    # they can also be defined in the xtb block
%xtb
  XTBINPUTSTRING  "argument 1" # string passed on to XTB call
  XTBINPUTSTRING2 "keyword 2"  # string passed on to XTB call
  ETemp           0 # electronic temperature, value for --etemp
  DOALPB      false # use implicit solvation, ALPB
  ALPBSOLVENT    "" # ALPB solvent, no default, string for --alpb
  DODDCOSMO   false # use implicit solvation, ddCOSMO
  DDCOSMOSOLVENT "" # ddCOSMO solvent, no default, string for --cosmo
  DOCPCMX     false # use implicit solvation, CPCMX
  CPCMXSOLVENT ""   # CPCMX solvent, no default, string for --cpcmx
  EPSILON       3.5 # Dielectric constant (only for ddCOSMO)
  ACCURACY        1 # accuracy, value for --acc, default is ORCA's accuracy x 1.e6
  MAXCORE      1024 # memory in MB reserved for XTB calculation,
                    # default is ORCA's maxcore
  NPROCS          1 # number of processors for running XTB",
                    # default is ORCA's PAL command
end
```

Note

- If jobs are run over several nodes, the number of cores used by the XTB tool might be lower than requested via the pal keyword.

7.5 Choice of Basis Set

A fair number of reasonable basis sets is hardwired in the program as will be described in the next section. In addition, whole basis sets can be read from a file, basis sets can be assigned for all atoms of a given type or, at the highest resolution, basis sets can be assigned to individual atoms which is convenient if different parts of the molecule are to be treated at different levels of accuracy. Most hard wired basis sets were obtained from the EMSL library [251] and the input format in ORCA is closely related to the "GAMESS-US" format.

Note

As of ORCA version 4.0, the basis set handling has been significantly modified!
Please check your basis sets very carefully!

7.5.1 Built-in Basis Sets

The basis set is specified in the block `%BASIS`. Note that there are three distinguished slots for auxiliary basis sets (AuxJ, AuxC and AuxJK) to be used with RI approximation. Which auxiliary basis slot is used in the actual program depends on the context. The AuxJ and AuxJK slots are used in the context of Fock matrix construction, whereas the AuxC slot is used for all other integral generation steps e.g. in post-Hartree Fock methods. Assigning the auxiliary basis with the simple input, takes care of the individual slots. However, in specific cases they must be set explicitly in the block input. For example, a "/JK" basis may be assigned to AuxJ in this way.

Note

As of ORCA 4.0, the basis set name has to be put in quotation marks, and the basis set name identifiers are the same as in the simple input!

```
%basis

  Basis "Def2-TZVP"         # The orbital expansion basis set
  AuxJ  "Def2/J"            # RI-J auxiliary basis set
  AuxJK "Def2/JK"           # RI-JK auxiliary basis set
  AuxC  "Def2-TZVP/C"       # Auxiliary basis set for correlated
                            # calcualtions, e.g. RI-MP2
  CABS  "cc-pVDZ-F12-OptRI" # complementary auxiliary basis set
                            # for F12 calculations

  DecontractBas   false  # if chosen "true" the program will
                         # decontract the orbital basis set
  DecontractAuxJ  false  # if "true" - decontract the AuxJ basis set
  DecontractAuxJK false  # if "true" - decontract the AuxJK basis set
  DecontractAuxC  false  # if "true" - decontract the AuxC basis set
  DecontractCABS  true   # if "false" - do not decontract the CABS
  Decontract      false  # if "true" - decontract all basis sets
end
```

Warning

- ORCA uses pure d and f functions (5D and 7F instead of Cartesian 6D and 10F) for all basis sets. This needs to be taken into account when results are compared with other programs, especially for Pople-style basis sets that were optimized with Cartesian (6D) functions.
- If you use `Decontract`: if your basis set arises from general contraction it will contain duplicate primitives in several contractions and these will be removed such that only unique primitives remain and there is no problem with redundancy.

A complete list of predefined basis sets and their availability is given in Table 7.10.

Table 7.10: Basis sets availability

Keyword	Availability
Orbital basis sets (`Basis`)	
STO-3G	H–I
MINI	H–Ca
MINIS	H–Ca
MINIX[1]	H–Rn
MIDI	H–Na, Al–K
3-21G	H–Cs
3-21GSP	H–Ar

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
4-22GSP	H–Ar
6-31G	H–Zn
6-31G*	H–Zn
m6-31G	Sc–Cu
m6-31G*	Sc–Cu
6-31G**	H–Zn
6-31G(d)	H–Zn
6-31G(d,p)	H–Zn
6-31G(2d)	H–Zn
6-31G(2d,p)	H–Zn
6-31G(2d,2p)	H–Zn
6-31G(2df)	H–Zn
6-31G(2df,2p)	H–Zn
6-31G(2df,2pd)	H–Zn
6-31+G*	H–Zn
6-31+G**	H–Zn
6-31+G(d)	H–Zn
6-31+G(d,p)	H–Zn
6-31+G(2d)	H–Zn
6-31+G(2d,p)	H–Zn
6-31+G(2d,2p)	H–Zn
6-31+G(2df)	H–Zn
6-31+G(2df,2p)	H–Zn
6-31+G(2df,2pd)	H–Zn
6-31++G**	H–Zn
6-31++G(d,p)	H–Zn
6-31++G(2d,p)	H–Zn
6-31++G(2d,2p)	H–Zn
6-31++G(2df,2p)	H–Zn
6-31++G(2df,2pd)	H–Zn
6-311G	H–Br
6-311G*	H–Br
6-311G**	H–Br
6-311G(d)	H–Br
6-311G(d,p)	H–Br
6-311G(2d)	H–Br
6-311G(2d,p)	H–Br
6-311G(2d,2p)	H–Br
6-311G(2df)	H–Br
6-311G(2df,2p)	H–Br
6-311G(2df,2pd)	H–Br
6-311G(3df)	H–Br
6-311G(3df,3pd)	H–Br
6-311+G*	H–Br
6-311+G**	H–Br
6-311+G(d)	H–Br
6-311+G(d,p)	H–Br
6-311+G(2d)	H–Br
6-311+G(2d,p)	H–Br
6-311+G(2d,2p)	H–Br
6-311+G(2df)	H–Br
6-311+G(2df,2p)	H–Br
6-311+G(2df,2pd)	H–Br
6-311+G(3df)	H–Br
6-311+G(3df,2p)	H–Br

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
6-311+G(3df,3pd)	H–Br
6-311++G**	H–Br
6-311++G(d,p)	H–Br
6-311++G(2d,p)	H–Br
6-311++G(2d,2p)	H–Br
6-311++G(2df,2p)	H–Br
6-311++G(2df,2pd)	H–Br
6-311++G(3df,3pd)	H–Br
SV	H–Kr
SV(P)	H–Kr
SVP	H–Kr
TZV	H–Kr
TZV(P)	H–Kr
TZVP	H–Kr
TZVPP	H–Kr
QZVP	H–Kr
QZVPP	H–Kr
DKH-SV(P)	H–Kr
DKH-SVP	H–Kr
DKH-TZV(P)	H–Kr
DKH-TZVP	H–Kr
DKH-TZVPP	H–Kr
DKH-QZVP	H–Kr
DKH-QZVPP	H–Kr
ZORA-SV(P)	H–Kr
ZORA-SVP	H–Kr
ZORA-TZV(P)	H–Kr
ZORA-TZVP	H–Kr
ZORA-TZVPP	H–Kr
ZORA-QZVP	H–Kr
ZORA-QZVPP	H–Kr
def2-mSVP[2]	H–Rn
def2-mTZVP[Page 502, 2]	H–Rn
def2-mTZVPP[Page 502, 2]	H–Lr
def2-SV(P)[Page 502, 2]	H–Rn
def2-SVP[Page 502, 2]	H–Rn
def2-TZVP(-f)[Page 502, 2]	H–Rn
def2-TZVP[Page 502, 2]	H–Rn
def2-TZVPP[Page 502, 2]	H–Rn
def2-QZVP[Page 502, 2]	H–Rn
def2-QZVPP[Page 502, 2]	H–Rn
def2-SVPD[Page 502, 2]	H–Rn
def2-TZVPD[Page 502, 2]	H–Rn
def2-TZVPPD[Page 502, 2]	H–Rn
def2-QZVPD[Page 502, 2]	H–Rn
def2-QZVPPD[Page 502, 2]	H–Rn
dhf-SV(P)[3]	H–Kr, Rb–Rn
dhf-SVP[Page 502, 3]	H–Kr, Rb–Rn
dhf-TZVP[Page 502, 3]	H–Kr, Rb–Rn
dhf-TZVPP[Page 502, 3]	H–Kr, Rb–Rn
dhf-QZVP[Page 502, 3]	H–Kr, Rb–Rn
dhf-QZVPP[Page 502, 3]	H–Kr, Rb–Rn
dhf-SV(P)-2c[Page 502, 3]	H–Kr, Rb–Rn
dhf-SVP-2c[Page 502, 3]	H–Kr, Rb–Rn
dhf-TZVP-2c[Page 502, 3]	H–Kr, Rb–Rn

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
dhf-TZVPP-2c[Page 502, 3]	H–Kr, Rb–Rn
dhf-QZVP-2c[Page 502, 3]	H–Kr, Rb–Rn
dhf-QZVPP-2c[Page 502, 3]	H–Kr, Rb–Rn
DKH-def2-SV(P)	H–Kr
DKH-def2-SVP	H–Kr
DKH-def2-TZVP(-f)	H–Kr
DKH-def2-TZVP	H–Kr
DKH-def2-TZVPP	H–Kr
DKH-def2-QZVPP	H–Kr
ZORA-def2-SV(P)	H–Kr
ZORA-def2-SVP	H–Kr
ZORA-def2-TZVP(-f)	H–Kr
ZORA-def2-TZVP	H–Kr
ZORA-def2-TZVPP	H–Kr
ZORA-def2-QZVPP	H–Kr
ma-def2-mSVP[Page 502, 2]	H–Rn
ma-def2-SV(P)[Page 502, 2]	H–Rn
ma-def2-SVP[Page 502, 2]	H–Rn
ma-def2-TZVP(-f)[Page 502, 2]	H–Rn
ma-def2-TZVP[Page 502, 2]	H–Rn
ma-def2-TZVPP[Page 502, 2]	H–Rn
ma-def2-QZVP[Page 502, 2]	H–Rn
ma-def2-QZVPP[Page 502, 2]	H–Rn
ma-DKH-def2-SV(P)	H–Kr
ma-DKH-def2-SVP	H–Kr
ma-DKH-def2-TZVP(-f)	H–Kr
ma-DKH-def2-TZVP	H–Kr
ma-DKH-def2-TZVPP	H–Kr
ma-DKH-def2-QZVPP	H–Kr
ma-ZORA-def2-SV(P)	H–Kr
ma-ZORA-def2-SVP	H–Kr
ma-ZORA-def2-TZVP(-f)	H–Kr
ma-ZORA-def2-TZVP	H–Kr
ma-ZORA-def2-TZVPP	H–Kr
ma-ZORA-def2-QZVPP	H–Kr
old-SV	H–I
old-SV(P)	H–I
old-SVP	H–I
old-TZV	H–I
old-TZV(P)	H–I
old-TZVP	H–I
old-TZVPP	H–I
old-DKH-SV(P)	H–I
old-DKH-SVP	H–I
old-DKH-TZV(P)	H–I
old-DKH-TZVP	H–I
old-DKH-TZVPP	H–I
old-ZORA-SV(P)	H–I
old-ZORA-SVP	H–I
old-ZORA-TZV(P)	H–I
old-ZORA-TZVP	H–I
old-ZORA-TZVPP	H–I
ANO-SZ	H–Ar, Sc–Zn
ANO-pVDZ	H–Ar, Sc–Zn
ANO-pVTZ	H–Ar, Sc–Zn

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
ANO-pVQZ	H–Ar, Sc–Zn
ANO-pV5Z	H–Ar, Sc–Zn
ANO-pV6Z	H–Ar, Sc–Zn
aug-ANO-pVDZ	H–Ar, Sc–Zn
aug-ANO-pVTZ	H–Ar, Sc–Zn
aug-ANO-pVQZ	H–Ar, Sc–Zn
aug-ANO-pV5Z	H–Ar, Sc–Zn
saug-ANO-pVDZ	H–Ar, Sc–Zn
saug-ANO-pVTZ	H–Ar, Sc–Zn
saug-ANO-pVQZ	H–Ar, Sc–Zn
saug-ANO-pV5Z	H–Ar, Sc–Zn
ANO-RCC-DZP	H–Cm
ANO-RCC-TZP	H–Cm
ANO-RCC-QZP	H–Cm
ANO-RCC-Full	H–Cm
pc-0	H–Ca, Ga–Kr
pc-1	H–Kr
pc-2	H–Kr
pc-3	H–Kr
pc-4	H–Kr
aug-pc-0	H–Ca, Ga–Kr
aug-pc-1	H–Kr
aug-pc-2	H–Kr
aug-pc-3	H–Kr
aug-pc-4	H–Kr
pcJ-0	H–He, B–Ne, Al–Ar
pcJ-1	H–He, B–Ne, Al–Ar
pcJ-2	H–He, B–Ne, Al–Ar
pcJ-3	H–He, B–Ne, Al–Ar
pcJ-4	H–He, B–Ne, Al–Ar
aug-pcJ-0	H–He, B–Ne, Al–Ar
aug-pcJ-1	H–He, B–Ne, Al–Ar
aug-pcJ-2	H–He, B–Ne, Al–Ar
aug-pcJ-3	H–He, B–Ne, Al–Ar
aug-pcJ-4	H–He, B–Ne, Al–Ar
pcseg-0	H–Kr
pcseg-1	H–Kr
pcseg-2	H–Kr
pcseg-3	H–Kr
pcseg-4	H–Kr
aug-pcseg-0	H–Kr
aug-pcseg-1	H–Kr
aug-pcseg-2	H–Kr
aug-pcseg-3	H–Kr
aug-pcseg-4	H–Kr
pcSseg-0	H–Kr
pcSseg-1	H–Kr
pcSseg-2	H–Kr
pcSseg-3	H–Kr
pcSseg-4	H–Kr
aug-pcSseg-0	H–Kr
aug-pcSseg-1	H–Kr
aug-pcSseg-2	H–Kr
aug-pcSseg-3	H–Kr
aug-pcSseg-4	H–Kr

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
W1-mtsmall	H–Ar
W1-DZ	H–Ar
W1-TZ	H–Ar
W1-QZ	H–Ar
W1-Opt	H–Ar
Sapporo-DZP-2012	H–Xe
Sapporo-TZP-2012	H–Xe
Sapporo-QZP-2012	H–Xe
Sapporo-DKH3-DZP-2012	K–Rn
Sapporo-DKH3-TZP-2012	K–Rn
Sapporo-DKH3-QZP-2012	K–Rn
LANL08[4]	Na–La, Hf–Bi
LANL08(f)[Page 502, 4]	Sc–Cu, Y–Ag, La, Hf–Au
LANL2DZ[Page 502, 4]	H, Li–La, Hf–Bi, U–Pu
LANL2TZ[Page 502, 4]	Sc–Zn, Y–Cd, La, Hf–Hg
LANL2TZ(f)[Page 502, 4]	Sc–Cu, Y–Ag, La, Hf–Au
vDZP[5]	H–Rn
def-TZVP[Page 502, 1]	Fr–Lr
ma-def-TZVP[Page 502, 1]	Fr–Lr
HGBS-5	H–Og
HGBS-7	H–Og
HGBS-9	H–Og
HGBSP1-5	H–Og
HGBSP1-7	H–Og
HGBSP1-9	H–Og
HGBSP2-5	H–Og
HGBSP2-7	H–Og
HGBSP2-9	H–Og
HGBSP3-5	H–Og
HGBSP3-7	H–Og
HGBSP3-9	H–Og
AHGBS-5	H–Og
AHGBS-7	H–Og
AHGBS-9	H–Og
AHGBSP1-5	H–Og
AHGBSP1-7	H–Og
AHGBSP1-9	H–Og
AHGBSP2-5	H–Og
AHGBSP2-7	H–Og
AHGBSP2-9	H–Og
AHGBSP3-5	H–Og
AHGBSP3-7	H–Og
AHGBSP3-9	H–Og
cc-pVDZ	H–Ar, Ca–Kr
cc-pVTZ	H–Ar, Ca–Kr, Y, Ag, Au
cc-pVQZ	H–Ar, Ca–Kr
cc-pV5Z	H–Ar, Ca–Kr
cc-pV6Z	H–He, Be–Ne, Al–Ar
aug-cc-pVDZ	H–Ar, Sc–Kr
aug-cc-pVTZ	H–Ar, Sc–Kr, Ag, Au
aug-cc-pVQZ	H–Ar, Sc–Kr
aug-cc-pV5Z	H–Ar, Sc–Kr
aug-cc-pV6Z	H–He, B–Ne, Al–Ar
cc-pVD(+d)Z	Na–Ar
cc-pVT(+d)Z	Na–Ar

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
cc-pVQ(+d)Z	Na–Ar
cc-pV5(+d)Z	Na–Ar
apr-cc-pV(Q+d)Z	H–Ar
may-cc-pV(T+d)Z	H–Ar
may-cc-pV(Q+d)Z	H–Ar
jun-cc-pV(D+d)Z	H–Ar
jun-cc-pV(T+d)Z	H–Ar
jun-cc-pV(Q+d)Z	H–Ar
jul-cc-pV(D+d)Z	H–Ar
jul-cc-pV(T+d)Z	H–Ar
jul-cc-pV(Q+d)Z	H–Ar
maug-cc-pV(D+d)Z	H–Ar
maug-cc-pV(T+d)Z	H–Ar
maug-cc-pV(Q+d)Z	H–Ar
aug-cc-pVD(+d)Z	Al–Ar
aug-cc-pVT(+d)Z	Al–Ar
aug-cc-pVQ(+d)Z	Al–Ar
aug-cc-pV5(+d)Z	Al–Ar
aug-cc-pV6(+d)Z	Al–Ar
aug-cc-pVTZ-J	H, B–F, Al–Cl, Sc–Zn, Se
cc-pCVDZ[6]	H–Ar, Ca, Ga–Kr
cc-pCVTZ[Page 502, 6]	H–Ar, Ca, Ga–Kr
cc-pCVQZ[Page 502, 6]	H–Ar, Ca, Ga–Kr
cc-pCV5Z[Page 502, 6]	H–Ar, Ca, Ga–Kr
cc-pCV6Z[Page 502, 6]	H–He, B–Ne, Al–Ar
aug-cc-pCVDZ[Page 502, 6]	H–Ar, Ga–Kr
aug-cc-pCVTZ[Page 502, 6]	H–Ar, Ga–Kr
aug-cc-pCVQZ[Page 502, 6]	H–Ar, Ga–Kr
aug-cc-pCV5Z[Page 502, 6]	H–Ar, Ga–Kr
aug-cc-pCV6Z[Page 502, 6]	H–He, B–Ne, Al–Ar
cc-pwCVDZ[Page 502, 6]	H–Ar, Ca, Ga–Kr
cc-pwCVTZ[Page 502, 6]	H–Ar, Ca–Kr, Ag, Au
cc-pwCVQZ[Page 502, 6]	H–Ar, Ca–Kr
cc-pwCV5Z[Page 502, 6]	H–Ar, Ca–Kr
aug-cc-pwCVDZ[Page 502, 6]	H–Ar, Ga–Kr
aug-cc-pwCVTZ[Page 502, 6]	H–Ar, Sc–Kr, Ag, Au
aug-cc-pwCVQZ[Page 502, 6]	H–Ar, Sc–Kr
aug-cc-pwCV5Z[Page 502, 6]	H–Ar, Sc–Kr
cc-pVDZ-PP[7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra, U
cc-pVTZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra, U
cc-pVQZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra, U
cc-pV5Z-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pVDZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pVTZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pVQZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pV5Z-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
cc-pCVDZ-PP[Page 502, 7]	Ca, Sr, Ba, Ra
cc-pCVTZ-PP[Page 502, 7]	Ca, Sr, Ba, Ra
cc-pCVQZ-PP[Page 502, 7]	Ca, Sr, Ba, Ra
cc-pCV5Z-PP[Page 502, 7]	Ca, Sr, Ba, Ra
aug-cc-pCVDZ-PP[Page 502, 7]	Ca, Sr, Ba, Ra
aug-cc-pCVTZ-PP[Page 502, 7]	Ca, Sr, Ba, Ra
aug-cc-pCVQZ-PP[Page 502, 7]	Ca, Sr, Ba, Ra
aug-cc-pCV5Z-PP[Page 502, 7]	Ca, Sr, Ba, Ra
cc-pwCVDZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra, U

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
cc-pwCVTZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra, U
cc-pwCVQZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra, U
cc-pwCV5Z-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pwCVDZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pwCVTZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pwCVQZ-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
aug-cc-pwCV5Z-PP[Page 502, 7]	Ca, Cu–Kr, Sr–Xe, Ba, Hf–Rn, Ra
cc-pVDZ-DK	H–Ar, Sc–Kr
cc-pVTZ-DK	H–Ar, Sc–Kr, Y–Xe, Hf–Rn
cc-pVQZ-DK	H–Ar, Sc–Kr, In–Xe, Tl–Rn
cc-pV5Z-DK	H–Ar, Sc–Kr
cc-pVDZ-DK3	U
cc-pVTZ-DK3	U
cc-pVQZ-DK3	U
aug-cc-pVDZ-DK	H–Ar, Sc–Kr
aug-cc-pVTZ-DK	H–Ar, Sc–Kr, Y–Xe, Hf–Rn
aug-cc-pVQZ-DK	H–Ar, Sc–Kr, In–Xe, Tl–Rn
aug-cc-pV5Z-DK	H–Ar, Sc–Kr
cc-pwCVDZ-DK[Page 502, 6]	H–Be, Na–Mg, Ca–Zn
cc-pwCVTZ-DK[Page 502, 6]	H–Be, Na–Mg, Ca–Zn, Y–Xe, Hf–Rn
cc-pwCVQZ-DK[Page 502, 6]	H–Be, Na–Mg, Ca–Zn, In–Xe, Tl–Rn
cc-pwCV5Z-DK[Page 502, 6]	H–Be, Na–Mg, Ca–Zn
cc-pwCVDZ-DK3	U
cc-pwCVTZ-DK3	U
cc-pwCVQZ-DK3	U
aug-cc-pwCVDZ-DK[Page 502, 6]	H–Be, Na–Mg, Sc–Zn
aug-cc-pwCVTZ-DK[Page 502, 6]	H–Be, Na–Mg, Sc–Zn, Y–Xe, Hf–Rn
aug-cc-pwCVQZ-DK[Page 502, 6]	H–Be, Na–Mg, Sc–Zn, In–Xe, Tl–Rn
aug-cc-pwCV5Z-DK[Page 502, 6]	H–Be, Na–Mg, Sc–Zn
cc-pVDZ-F12	H–Ar
cc-pVTZ-F12	H–Ar
cc-pVQZ-F12	H–Ar
cc-pVDZ-PP-F12[Page 502, 7]	Ga–Kr, In–Xe, Tl–Rn
cc-pVTZ-PP-F12[Page 502, 7]	Ga–Kr, In–Xe, Tl–Rn
cc-pVQZ-PP-F12[Page 502, 7]	Ga–Kr, In–Xe, Tl–Rn
cc-pCVDZ-F12	Li–Ar
cc-pCVTZ-F12	Li–Ar
cc-pCVQZ-F12	Li–Ar
haV(T+d)Z	H–Ar
haV(Q+d)Z	H–Ar
haV(5+d)Z	H–Ar
Partridge-1	H, Li–Sr
Partridge-2	H, Li–Kr
Partridge-3	H, Li–Zn
Partridge-4	Sc–Zn
x2c-SV(P)all	H–Rn
x2c-SVPall	H–Rn
x2c-TZVPall	H–Rn
x2c-TZVPPall	H–Rn
x2c-QZVPall	H–Rn
x2c-QZVPPall	H–Rn
x2c-SV(P)all-2c	H–Rn
x2c-SVPall-2c	H–Rn
x2c-TZVPall-2c	H–Rn
x2c-TZVPPall-2c	H–Rn

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
x2c-QZVPall-2c	H–Rn
x2c-QZVPPall-2c	H–Rn
x2c-SV(P)all-s	H–Rn
x2c-SVPall-s	H–Rn
x2c-TZVPall-s	H–Rn
x2c-TZVPPall-s	H–Rn
x2c-QZVPall-s	H–Rn
x2c-QZVPPall-s	H–Rn
x2c-QZVPall-2c-s	H–Rn
x2c-QZVPPall-2c-s	H–Rn
SARC-DKH-SVP	Hf–Hg
SARC-DKH-TZVP	Rb–Rn, Ac–Lr
SARC-DKH-TZVPP	Rb–Rn, Ac–Lr
SARC-ZORA-SVP	Hf–Hg
SARC-ZORA-TZVP	Rb–Rn, Ac–Lr
SARC-ZORA-TZVPP	Rb–Rn, Ac–Lr
SARC2-DKH-QZV	La–Lu
SARC2-DKH-QZVP	La–Lu
SARC2-ZORA-QZV	La–Lu
SARC2-ZORA-QZVP	La–Lu
D95	H, Li, B–Ne, Al–Cl
D95p	H, Li, B–Ne, Al–Cl
EPR-II	H, B–F
EPR-III	H, B–F
IGLO-II	H, B–F, Al–Cl
IGLO-III	H, B–F, Al–Cl
UGBS	H–Th, Pu–Am, Cf–Lr
CP	Sc–Zn
CP(PPP)	Sc–Zn
Wachters+f	Sc–Cu
Coulomb-fitting auxiliary basis sets (`AuxJ``)	
def2/J	H–Rn
def2-mTZVP/J	H–Rn
def2-mTZVPP/J	H–Rn
x2c/J	H–Rn
SARC/J	H–Rn, Ac–Lr
Coulomb and exchange-fitting auxiliary basis sets (`AuxJK`)	
def2/JK	H–Ba, Hf–Rn
def2/JKsmall	H–Ra, Th–Lr
cc-pVTZ/JK	H, B–F, Al–Cl, Ga–Br
cc-pVQZ/JK	H, B–F, Al–Cl, Ga–Br
cc-pV5Z/JK	H, B–F, Al–Cl, Ga–Br
aug-cc-pVTZ/JK	H, B–F, Al–Cl, Ga–Br
aug-cc-pVQZ/JK	H, B–F, Al–Cl, Ga–Br
aug-cc-pV5Z/JK	H, B–F, Al–Cl, Ga–Br
SARC2-DKH-QZV/JK	La–Lu
SARC2-DKH-QZVP/JK	La–Lu
SARC2-ZORA-QZV/JK	La–Lu
SARC2-ZORA-QZVP/JK	La–Lu
Auxiliary basis sets for correlated methods (`AuxC`)	
def2-SVP/C	H–Rn
def2-TZVP/C	H–Rn
def2-TZVPP/C	H–Rn
def2-QZVPP/C	H–Rn
def2-SVPD/C	H–La, Hf–Rn

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
def2-TZVPD/C	H–La, Hf–Rn
def2-TZVPPD/C	H–La, Hf–Rn
def2-QZVPPD/C	H–La, Hf–Rn
cc-pVDZ/C	H–Ar, Ga–Kr
cc-pVTZ/C	H–Ar, Sc–Kr
cc-pVQZ/C	H–Ar, Sc–Kr
cc-pV5Z/C	H–Ar, Ga–Kr
cc-pV6Z/C	H–He, B–Ne, Al–Ar
aug-cc-pVDZ/C	H–He, Be–Ne, Mg–Ar, Ga–Kr
aug-cc-pVTZ/C	H–He, Be–Ne, Mg–Ar, Sc–Kr
aug-cc-pVQZ/C	H–He, Be–Ne, Mg–Ar, Sc–Kr
aug-cc-pV5Z/C	H–Ne, Al–Ar, Ga–Kr
aug-cc-pV6Z/C	H–He, B–Ne, Al–Ar
cc-pwCVDZ/C[Page 502, 6]	H–He, B–Ne, Al–Ar, Ga–Kr
cc-pwCVTZ/C[Page 502, 6]	H–He, B–Ne, Al–Ar, Sc–Kr
cc-pwCVQZ/C[Page 502, 6]	H–He, B–Ne, Al–Ar, Ga–Kr
cc-pwCV5Z/C[Page 502, 6]	H–Ne, Al–Ar
aug-cc-pwCVDZ/C[Page 502, 6]	H–He, B–Ne, Al–Ar, Ga–Kr
aug-cc-pwCVTZ/C[Page 502, 6]	H–He, B–Ne, Al–Ar, Sc–Kr
aug-cc-pwCVQZ/C[Page 502, 6]	H–He, B–Ne, Al–Ar, Ga–Kr
aug-cc-pwCV5Z/C[Page 502, 6]	H–Ne, Al–Ar
cc-pVDZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
cc-pVTZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
cc-pVQZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
aug-cc-pVDZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
aug-cc-pVTZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
aug-cc-pVQZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
cc-pwCVDZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
cc-pwCVTZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
cc-pwCVQZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
aug-cc-pwCVDZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
aug-cc-pwCVTZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
aug-cc-pwCVQZ-PP/C	Cu–Kr, Y–Xe, Hf–Rn
cc-pVDZ-F12-MP2Fit	H–Ar
cc-pVTZ-F12-MP2Fit	H–Ar
cc-pVQZ-F12-MP2Fit	H–Ar
cc-pVDZ-PP-F12-MP2Fit	Ga–Kr, In–Xe, Tl–Rn
cc-pVTZ-PP-F12-MP2Fit	Ga–Kr, In–Xe, Tl–Rn
cc-pVQZ-PP-F12-MP2Fit	Ga–Kr, In–Xe, Tl–Rn
cc-pCVDZ-F12-MP2Fit	Li–Ar
cc-pCVTZ-F12-MP2Fit	Li–Ar
cc-pCVQZ-F12-MP2Fit	Li–Ar
Complementary auxiliary basis sets for F12 calculations (CABS)	
cc-pVDZ-F12-CABS	H, B–Ne, Al–Ar
cc-pVTZ-F12-CABS	H, B–Ne, Al–Ar
cc-pVQZ-F12-CABS	H, B–Ne, Al–Ar
cc-pVDZ-F12-OptRI	H–Ar
cc-pVTZ-F12-OptRI	H–Ar
cc-pVQZ-F12-OptRI	H–Ar
cc-pVDZ-PP-F12-OptRI	Ga–Kr, In–Xe, Tl–Rn
cc-pVTZ-PP-F12-OptRI	Ga–Kr, In–Xe, Tl–Rn
cc-pVQZ-PP-F12-OptRI	Ga–Kr, In–Xe, Tl–Rn
aug-cc-pVDZ-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg
aug-cc-pVTZ-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg
aug-cc-pVQZ-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg

continues on next page

Table 7.10 – continued from previous page

Keyword	Availability
aug-cc-pV5Z-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg
cc-pCVDZ-F12-OptRI	Li–Ar
cc-pCVTZ-F12-OptRI	Li–Ar
cc-pCVQZ-F12-OptRI	Li–Ar
aug-cc-pwCVDZ-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg
aug-cc-pwCVTZ-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg
aug-cc-pwCVQZ-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg
aug-cc-pwCV5Z-PP-OptRI	Cu–Zn, Ag–Cd, Au–Hg

Note

Check these pointers for more information on the basis sets: (indicated in ecah element of the table as well)

- [1]Used with the Def-ECP pseudopotentials (Rb–Lr).
- [2]Used with the Def2-ECP pseudopotentials (Rb–Rn).
- [3]Used with the dhf-ECP or dhf-ECP-2c pseudopotentials (Rb–Rn). For elements H–Kr equivalent to the respective def2-XVP basis set.
- [4]Used with the HayWadt pseudopotentials (Na–La, Hf–Bi, U–Pu).
- [5]Valence double-ζ with large-core pseudopotentials. For the respective ECP types per element, see Ref. [15] and table 6.5.
- [6]The respective basis sets without core correlation functions, i.e. (aug-)cc-pVXZ(-DK)(/C), are used for H and He.
- [7]Used with the SK-MCDHF-RSC pseudopotentials (Ca, Cu–Kr, Sr–Xe, Ba, Hf–Ra, U).

A note on RI and auxiliary basis sets: one thing that is certainly feasible and reasonable if you do not want to depend on the RI approximation is to converge a RI-J calculation and then take the resulting orbitals as initial guess for a calculation with exact Coulomb term. This should converge within a few cycles and the total execution time should still be lower than just converging the calculation directly with exact Coulomb treatment.

7.5.2 Automatic generation of auxiliary basis sets

If no auxiliary basis set is available for your chosen orbital basis set, one can be generated automatically by ORCA using the keyword `AutoAux`. This is specified as any other fitting basis set: as a value to the `AuxJ/AuxJK/AuxC` variables in the `%basis` block or as a separate keyword in the simple input line (in which case all three Aux slots are populated with identical fitting basis sets). `AutoAux` can also be assigned to individual elements or atoms - see sections *Assigning or Adding Basis Functions to an Element* and *Assigning or Adding Basis Functions to Individual Atoms*. The generated basis sets can be used for Coulomb, exchange and correlation fitting and are as accurate as the optimized auxiliary basis sets at the cost of being up to twice as large. The exact generation procedure is described elsewhere [827] but notably **it has been significantly altered since ORCA 3.1 and will not produce the same results!** For compatibility, the old version is still accessible via the setting `OldAutoAux true` in the `%basis` block. Some additional settings for `AutoAux` are given below with their default values.

[1] Used with the Def-ECP pseudopotentials (Rb–Lr).
[2] Used with the Def2-ECP pseudopotentials (Rb–Rn).
[3] Used with the dhf-ECP or dhf-ECP-2c pseudopotentials (Rb–Rn).
[4] Used with the HayWadt pseudopotentials (Na–La, Hf–Bi, U–Pu).
[5] Valence double-ζ with large-core pseudopotentials. For the respective ECP types per element, see Ref. [540] and Table 4.3.
[6] The respective basis sets without core correlation functions, i.e. (aug-)cc-pVXZ(-DK)(/C), are used for H and He.
[7] Used with the SK-MCDHF-RSC pseudopotentials (Ca, Cu–Kr, Sr–Xe, Ba, Hf–Ra, U).

```
%basis
  AuxJ  "AutoAux"      # Use AutoAux to generate the AuxJ fitting basis set
  AuxJK "AutoAux"      # Use AutoAux to generate the AuxJK fitting basis set
  AuxC  "AutoAux"      # Use AutoAux to generate the AuxC fitting basis set
  AutoAuxSize    0-3   # 0 use minimal effective rather than minimal
                       # primitive exponent (suitable for ANO basis sets)
                       # 1 (default) increases the maximal exponent
                       # for the shells with low angular momenta.
                       # 2 increases the maximal exponent for all shells
                       # 3 directly uses the primitives and produces
                       # the largest fitting basis
  AutoAuxLmax  false   # If true increase the maximal angular momentum of
                       # the fitting basis set to the highest value
                       # permitted by ORCA and by the orbital basis set.
  AutoAuxLLimit   -1   # If >0, do not exceed the given angular momentum.
  OldAutoAux   false   # If true selects the ORCA 3.1 generation procedure

  # Only change the defaults below if you know what you are doing

  AutoAuxF[0]   20.0   # The factor to increase the maximal s-exponent
  AutoAuxF[1]    7.0   # Same for the p-shell
  AutoAuxF[2]    4.0   # Same for the d-shell
  AutoAuxF[3]    4.0   # Same for the f-shell
  AutoAuxF[4]    3.5   # Same for the g-shell
  AutoAuxF[5]    2.5   # Same for the h-shell
  AutoAuxF[6]    2.0   # Same for the i-shell
  AutoAuxF[7]    2.0   # Same for the j-shell
  AutoAuxB[0]    1.8   # Even-tempered expansion factor for the s-shell
  AutoAuxB[1]    2.0   # Same for the p-shell
  AutoAuxB[2]    2.2   # Same for the d-shell
  AutoAuxB[3]    2.2   # Same for the f-shell
  AutoAuxB[4]    2.2   # Same for the g-shell
  AutoAuxB[5]    2.3   # Same for the h-shell
  AutoAuxB[6]    3.0   # Same for the i-shell
  AutoAuxB[7]    3.0   # Same for the j-shell
  AutoAuxTightB true   # Only use B[l] for shells with high l and B[0] for the rest
end
```

Note that if the orbital basis set contains diffuse functions, as is the case for the aug-cc-pVXZ sets, the AutoAux fitting basis may contain (near-)linear dependencies. In this case, the Cholesky decomposition of the Coulomb metric will fail and the program will likely crash. One may print the offending auxiliary basis using `!PrintBasis` and manually remove the most diffuse s- and/or p-functions, which will usually resolve the problem. An alternative, automatic solution is implemented in ORCA 5.0 – see section *Removal of Redundant Basis Functions*.

7.5.3 Assigning or Adding Basis Functions to an Element

In order to assign a new basis set to a given element, use:

```
%basis
  NewGTO 8    # New basis for oxygen.
# NewGTO O    # This works as well.
   S   3
    1  910.10034975         0.03280967
    2  137.19711335         0.23422391
    3   30.85279077         0.81490980
   S   2
    1     1.72885887         0.27389659
    2     0.39954770         0.79112437
   P   1
    1     8.35065975         1.00000000
```

(continues on next page)

(continued from previous page)

```
  end
end
```

Note that for simplicity and consistency the input format is the same as that used in the basis set files. In this format the first line carries first the angular momentum of the shell to be added (s, p, d, f, g, h, i, j) and the number of primitives. Then for each primitive one line follows which has (a) the index of the primitive (1, 2, 3, ...) (b) the exponent of the primitive and (c) the contraction coefficient (unnormalized). Note that ORCA always uses spherical harmonic Gaussian functions. L-shells (not to be confused with angular momentum equal to 9) can only be dealt with as separate s- and p-shells. There also is the possibility to include a `SCALE X` statement after the number of primitives in the first line to indicate that the basis function exponents should be scaled.

In order to add basis functions to the basis of a given element (for example because you do not like the standard polarization functions) use `AddGTO` instead of `NewGTO`. In `NewGTO` or `AddGTO` you can also use the nicknames of internally stored basis sets. An example is:

```
%basis
  NewGTO 8    # new basis for oxygen
   "6-31G"
    D  1
    1  0.4  1.0
  end
end
```

In this example the 6-31G basis is assigned to oxygen and in addition a polarization function with exponent 0.4 is added to the oxygen basis.

Note that the `NewGTO` keyword does not change the ECP for the given element - you must use `NewECP` or `DelECP` (see section *Advanced Specification of Effective Core Potentials*).

A similar mechanism was established for the auxiliary basis sets in RI calculations:

```
%basis
  NewAuxJGTO 8    # new auxiliary basis for oxygen
    s  1
    1  350  1.0
    ... etc
  end
  AddAuxJGTO 8   # add a shell to the auxiliary basis for
                 # oxygen
    D  1
    1  0.8  1.0
  end
end
```

New basis functions can be specifically assigned to any auxiliary basis sets. The keywords `NewAuxCGTO`, `AddAuxCGTO`, `NewAuxJKGTO`, `AddAuxJKGTO`, `NewCABSGTO`, `AddCABSGTO` are used in the same way. The keywords `NewAuxGTO` and `AddAuxGTO` are the same as `NewAuxJGTO` and `AddAuxJGTO`, that is, they only influence the Coulomb auxiliary basis (/J basis)!

7.5.4 Assigning or Adding Basis Functions to Individual Atoms

Sometimes you may want to not treat all carbon atoms with the same basis set but to assign a specific basis set to a specific atom in the molecules. This is also possible in ORCA and takes place in the coordinate section (`%coords`, `*xyz`, etc.). The format is the same as described above. An example may help to make things clear:

```
*int 0 1
  C(1)  0 0 0 0.00   0.0   0.00
   AddGTO
     D  1
```

(continues on next page)

(continued from previous page)

```
      1  1.0  1.0
   end
  O(2)  1 0 0 1.13   0.0   0.00
   NewGTO
     "6-311G"
     D  1
      1  1.2  1.0
   end
*
```

In this example an extra d-shell with exponent 1.0 is added to the first carbon atom and the basis for the oxygen atom is changed to 6-311G with an extra d-function of exponent 1.2 added.

Analogously, AUX basis functions can be assigned or added to individual atoms using the keywords NewAuxJGTO, AddAuxJGTO, NewAuxCGTO, AddAuxCGTO, NewAuxJKGTO, AddAuxJKGTO, NewCABSGTO, AddCABSGTO.

A note on the use of AutoAux: if you change the basis set on a given atom and want to generate a fitting basis, you have to specify it again in the COORDS section, even if AutoAux is already present in the simple input line or in the %basis block. For example:

```
! Def2-SVP Def2/JK
%basis
  NewAuxJKGTO H
    "AutoAux"
  end
end
*xyz 0 1
  O  0.00 0.00 0.00
  H -0.25 0.93 0.00
  H  0.96 0.00 0.00
   AddGTO
     P 1
      1 1.6 1.0
     D 1
      1 1.0 1.0
   end
   NewAuxJKGTO
     "AutoAux"
   end
*
```

Here the oxygen atom is assigned the Def2-SVP basis and the Def2/JK fitting basis, the first hydrogen atom is assigned the Def2-SVP basis and an automatically generated fitting basis and the second hydrogen atom is assigned the Def2-SVP basis with two additional polarization functions and a larger automatically generated fitting basis that accounts for these functions.

Tip

When assigning custom basis sets it is always a good idea to print the basis set information (%output print[p_basis] 2 end or simply !PrintBasis) and check that everything is correct.

7.5.5 Assigning Basis Sets and ECPs to Fragments

In multi-level or QM/QM calculations it may be convenient to assign different basis sets to different fragments. This can be done with the keywords `FragBasis`, `FragAuxJ`, `FragAuxJK`, `FragAuxC`, `FragCABS`, and `FragECP` in the `%basis` block, followed by the number of the fragment (numbering starts at 1!) and a standard basis set or ECP from the ORCA library (see Tables Table 7.10 and Table 4.3). Note that unlike the `NewGTO` keyword, `FragBasis` also changes the ECP, if applicable. Fragment basis sets will overload the global or element-specific (*Assigning or Adding Basis Functions to an Element*) choice but can be overloaded for individual atoms (*Assigning or Adding Basis Functions to Individual Atoms*). If AutoAux is requested for a fragment, it will be generated for the actual orbital basis set chosen for each atom, even if it is changed in the coordinates section. However, if AutoAux was requested for an element or in the simple input, the auxiliary basis will be generated before the fragment basis is assigned (which is not desired), therefore AutoAux must be requested again for the fragment. An example is given below:

```
! PrintBasis BP86 NoIter
! def2-SVP def2/J
%basis
  FragBasis 1 "def2-TZVP"
  FragBasis 2 "cc-pVTZ-PP"
  FragAuxJ  2 "AutoAux"
  FragECP   3 "SK-MCDHF-RSC"
  FragAuxJ  3 "def2/JK"
end
*xyz 0 1
  H(1) 0 0 0
  I(1) 0 0 1.6
  H(2) 0 5 0   NewGTO "cc-pVTZ" end
  I(2) 0 5 1.6
  H(3) 5 0 0
  I(3) 5 0 1.6
*
# Final basis sets:
# Atom Basis      ECP          AuxJ
# 0H   def2-TZVP  def2-ECP     def2/J
# 1I   def2-TZVP  def2-ECP     def2/J
# 2H   cc-pVTZ    -            AutoAux(cc-pVTZ)
# 3I   cc-pVTZ-PP SK-MCDHF-RSC AutoAux(cc-pVTZ-PP)
# 4H   def2-SVP   -            def2/JK
# 5I   def2-SVP   SK-MCDHF-RSC def2/JK
```

It is also possible to read fragment-specific basis sets from a file. The syntax is analogous, using the keywords `ReadFragBasis`, `ReadFragAuxJ`, `ReadFragAuxJK`, `ReadFragAuxC`, `ReadFragCABS`, and `ReadFragECP`. In this case, the input string is expected to be an existing basis set file in GAMESS-US format (see section *Reading Orbital and Auxiliary Basis Sets from a File*). All other details above (e.g., regarding ECPs and AutoAux) also apply here.

7.5.6 Reading Orbital and Auxiliary Basis Sets from a File

By using the variables `GTOName`, `GTOAuxJName`, `GTOAuxJKName`, `GTOAuxCName`, and `GTOCABSName` (`GTOAuxName` is a synonym for `GTOAuxJName`) a basis set can be read from an ASCII file. In this way you can construct or modify your favorite standard basis set and load it easily into the program.

```
%basis
  # read an externally specified orbital basis
  GTOName      = "MyBasis.bas"
  # read an externally specified Coulomb-fitting basis for RI calculations
  GTOAuxJName  = "MyAuxJBasis.bas"
  # read an externally specified Coulomb- and exchange-fitting basis
  GTOAuxJKName = "MyAuxJKBasis.bas"
  # read an externally specified correlation-fitting basis
  GTOAuxCName  = "MyAuxCBasis.bas"
```

(continues on next page)

(continued from previous page)

```
  # read an externally specified complementary auxiliary basis set
  GTOCABSName  = "MyCABSBasis.bas"
end
```

A word of caution: in C/C++ the backslashes in directory assignments must be given twice to be correctly understood! The format is that used for "GAMESS-US" in the EMSL library [251]. To give an example of what this format looks like here is a part of the 3-21GSP basis of Buenker and coworkers [587, 588]:

```
!                         lines in the beginning with '!' or '#' are comments
! BASIS="3-21GSP"
!Elements                             References
!--------                             ----------
! H - Ne: A.V. Mitin, G. Hirsch, R. J. Buenker, Chem. Phys. Lett. 259, 151 (1996)
! Na - Ar: A.V. Mitin, G. Hirsch, R. J. Buenker, J. Comp. Chem. 18, 1200 (1997).
!
$DATA        ! Optional
HYDROGEN     ! (3s) -> [2s]      Element symbols are also recognized
 S   2
  1           4.50036231         0.15631167
  2           0.68128924         0.90466909
 S   1
  1           0.15137639         1.00000000
CARBON       ! (6s,3p) -> [3s,2p]
 S   3
  1         499.24042249         0.03330322
  2          75.25419194         0.23617745
  3          16.86538669         0.81336259
 L   2       ! L shells are a s and a p shell with identical exponents
  1           0.89739483         0.24008573         0.46214684
  2           0.21746772         0.81603757         0.66529098
 L   1
  1           4.52660451         1.00000000         1.00000000
$END         ! Optional
```

The file format for the auxiliary basis sets is exactly the same. Basis sets can be also exported in GAMESS-US format by the `orca_exportbasis` utility (section *orca_exportbasis*). Note that in order to read basis sets printed by ORCA (using `!PrintBasis`), the `NewGTO` and `end` keywords must be removed.

7.5.7 Advanced Specification of Effective Core Potentials

Library ECPs and Basis Sets

Besides the simple input line (section *Effective Core Potentials*), assignment of ECPs can be done within the `%basis` block using the `ECP` and `NewECP` keywords as in the following example:

```
%basis
  ECP       "def2-ECP"     # All elements (for which the ECP is defined)
  NewECP Pt "def2-SD" end  # Different ECP for Pt
end
```

A variant of the `NewECP` keyword can be used for individual atoms inside the geometry definition:

```
* xyz ...
  ...
  S   0.0  0.0  0.0  NewECP "SDD" end
  ...
*
```

Note that these keywords only affect the ECP and not the valence basis set!

In case the basis set for an element/atom has been changed using the NewGTO keyword (see sections *Assigning or Adding Basis Functions to an Element* and *Assigning or Adding Basis Functions to Individual Atoms* above) it may be necessary to remove the ECP from that element/atom. This can be done with the DelECP keyword in the %basis block or coordinates input, respectively:

```
! LANL2DZ                   # Uses HayWadt ECPs by default, starting from Na
%basis
  NewGTO S "Def2-TZVP" end  # All-electron up to Kr
  DelECP S                  # Remove HayWadt ECP
end
* xyz ...
...
Cu   0.0  0.0  0.0
  DelECP                    # Remove HayWadt ECP
  NewGTO "Def2-QZVPP" end   # All-electron up to Kr
...
*
```

To remove all ECPs loaded by default (e.g. in case no global basis set is chosen) you can use the !NoECP simple keyword.

Manual Input of ECP Parameters

To manually specify ECP parameters, the NewECP keyword is followed by the element for which an ECP is to be entered, the number of core electrons to be replaced (N_core) and the maximum angular momentum (lmax). The ECP specification is finished by giving the definitions of the individual shells that constitute the angular dependent potentials U_l.

```
%basis
    NewECP <element>
      N_core <number of core electrons>
      lmax <max. angular momentum>
      [shells]
    end
end
```

For each ECP shell, first the angular momentum l has to be given, followed by the number of primitives. The primitives themselves are then specified by giving a running index and the respective tuple of exponent a_{kl}, expansion coefficient d_{kl} and radial power n_{kl}.

```
# ECP shell
      l <number of primitives>
      1    a1l    d1l    n1l
      2    a2l    d2l    n3l
      ...
```

As an example, consider the SD(10,MDF) for Vanadium. The name indicates a Stuttgart–Dresden type ECP that replaces 10 core electrons and is derived from a relativistic calculation for the neutral atom. It consists of 4 shells with angular momentum s, p, d, and f. Note that the f shell has an expansion coefficient of 0.0 and thus will not contribute at all to this effective core potential. This is typical for all SD potentials (but may be different for program packages like TURBOMOLE that do not support arbitrary angular momentum with respect to the ECP and therefore use recontractions of the original parameter sets).

```
%basis
  # ECP SD(10,MDF) for V
  # M. Dolg, U. Wedig, H. Stoll, H. Preuss,
  # J. Chem. Phys. 86, 866 (1987).
  NewECP V
    N_core 10
    lmax f
```

(continues on next page)

(continued from previous page)

```
    s 2
      1      14.4900000000    178.4479710000  2
      2       6.5240000000     19.8313750000  2
    p 2
      1      14.3000000000    109.5297630000  2
      2       6.0210000000     12.5703100000  2
    d 2
      1      17.4800000000    -19.2196570000  2
      2       5.7090000000     -0.6427750000  2
    f 1
      1       1.0000000000      0.0000000000  2
  end
end
```

ECPs and ghost atoms

When ghost atoms are defined on the input (see section *Special definitions*), ECPs are **not** added to these atoms by default. If that is somehow needed, please add `GhostECP true` under the `%basis` block.

```
%basis
 GhostECP true
 # AllowGhostECP true # synonym
end
```

7.5.8 Embedding Potentials

Computations on cluster models sometimes require the presence of embedding potentials in order to account for otherwise neglected repulsive terms at the border [304]. In order to simplify these kind of calculations with ORCA the ECP embedding can be accomplished quite easily:

```
*xyz ...
# atom>  charge x    y    z    optional ECP declaration
Zr>      4.0    0.0  0.0  0.0  NewECP "SDD" end
...
*
```

The declaration of such a coreless ECP center takes place in the coordinates section by appending a bracket ">" to the element symbol. Note that embedding ECPs are treated as point charges in ORCA, so the charge has to be given next. The coordinates of the coreless ECP center have to be specified as usual and may be followed by an optional ECP assignment. In general, calculations that employ an ECP embedding procedure should be single point calculations. However if the need arises to perform a geometry optimization, make sure to set up explicit Cartesian constraints for the coreless ECP centers.

7.5.9 Linear Dependence

The previous sections describe the assessment of a desired molecular basis set from appropriately parametrized functions at various locations within the molecule (normally centered on atoms). The parametrization of these functions is such that the chance for redundancy is minimal. Since however, one is limited to work with finite numerical precision, and furthermore these parameters also depend on the molecular geometry, redundancies cannot be completely eliminated in advance. Redundancy means that the subspace spanned by the given basis functions at given values of parameters (including geometry), can be identically spanned by a smaller number of *linear independent* basis functions. Linear dependent (redundant) function sets however may cause numerical instabilities. Linear dependence is normally identified by searching for zero eigenvalues of the overlap matrix. Note that the inverse of the overlap (or related matrices) are used for orthogonalization purposes, and it follows that if near zero eigenvalues are not treated properly, the inverse becomes ill-defined, and the SCF procedure numerically unstable.

From the previous discussion, it is evident that the crucial parameter for curing linear dependence is the threshold below which an overlap eigenvalue is considered zero. This parameter may be changed using the following keyword

```
%scf
  sthresh 1e-6 # default 1e-7
end
```

Although there is no strict limit for the value of the above parameter, it should reasonably be somewhere between 1e-5 and 1e-8 (the default is 1e-7). One may get away with 1e-9 or perhaps even lower without convergence problem, but there is a risk that the result is contaminated with noise caused by the near zero vectors. In difficult cases, an 1e-6 threshold was often found to work smoothly, and above that one risks throwing away more and more functions, which also influence comparability of results with other calculations. To monitor the behavior of the small eigenvalues, one should look for the following block in the output

```
Diagonalization of the overlap matrix:
Smallest eigenvalue                        ... -1.340e-17
Time for diagonalization                   ...    0.313 sec
Threshold for overlap eigenvalues          ... 1.000e-07
Number of eigenvalues below threshold      ... 1
Smallest eigenvalue above threshold        ... 6.013e-07
Time for construction of square roots      ...    0.073 sec
Total time needed                          ...    0.387 sec
```

Here, the smallest eigenvalue is printed, along with the currently used overlap threshold, and the number of functions below this (which will be dropped). It is a recommended consistency check to look for an equal number of zero entries among orbital energies once the SCF procedure converged. Note that for functions belonging to zero eigenvalues *no level shifts* are applied!

In case that redundant vectors were removed from the basis, `! MORead NoIter` should only be used in conjunction with the same `SThresh` as in the original calculation, otherwise the results will be inconsistent. `! MORead` may still be used together with a change in `SThresh`, but a few SCF iterations will be required.

Automatic Adjustments for Near Linear-Dependent Cases

Starting from ORCA6, there is now a keyword called DiffSThresh, which controls an automatic tightening of the integral cutoff parameters Thresh and TCut in case small eigenvalues of the overlap matrix are found. We found this to be important in some calculations using diffuse basis, and these parameters are set to a minimum value of Thresh=1e-12 and TCut=1e-13 in case the "Smallest eigenvalue" shown above gets below that number. If the cutoffs are already tighter than that, for instance when using `!VeryTightSCF`, than nothing will happen.

We found empirically that these are safe numbers to mitigate noise and increase the robustness of the SCF procedure, thus they are enforced by default. The default is 1e-6 and this can be turned off by setting `%SCF DiffSThresh -1 END` on the input in case you don't want this automatic adjustment to happen.

Removal of Redundant Basis Functions

While the approach described above is usually successful in removing linear dependencies from the orbital basis set, the auxiliary basis used in RI is not orthogonalized the same way. Instead, the RI linear equation system is solved using a Cholesky decomposition (CD) of the auxiliary basis Coulomb metric. If the auxiliary basis is redundant, the CD fails and the program usually aborts. One simple solution implemented in ORCA is to perform a pivoted Cholesky decomposition (PCD) of the metric, terminating at a given threshold. Then, the shells contributing to the nullspace are removed from the basis at the beginning of the calculation. This can be requested for any of the basis sets using either the overlap or the Coulomb metric. It is most appropriate for the AuxJ/AuxJK/AuxC basis using the Coulomb metric. The truncated basis can be examined using the `!PrintBasis` keyword. Often, functions may be removed for some atoms of a given element, but kept for others. As long as the threshold is low enough, i.e. only truly redundant functions are removed, this should not affect the molecular symmetry of the results.

```
%basis
  PCDTrimBas   Overlap # Trim the orbital basis in the overlap metric
  PCDTrimAuxJ  Coulomb # Trim the AuxJ basis in the Coulomb metric
  PCDTrimAuxJK Coulomb # Trim the AuxJK basis in the Coulomb metric
  PCDTrimAuxC  Coulomb # Trim the AuxC basis in the Coulomb metric
  PCDThresh    -1      # Threshold for the PCD: chosen automatically if <0
end
```

7.6 Choice of Initial Guess and Restart of SCF Calculations

The initial guess is an important issue in each SCF calculation. If this guess is reasonable, the convergence of the procedure will be much better. ORCA makes some effort to provide a good initial guess and gives the user enough flexibility to tailor the initial guess to his or her needs.

The initial guess is also controlled via the `%scf` block and the variables `Guess`, `MOInp` and `GuessMode`.

```
%scf
  Guess     HCore       # One electron matrix
            Hueckel     # Extended H\"uckel guess
            PAtom       # Polarized atomic densities
            PModel      # Model potential
            MORead      # Restart from an earlier calc.
  MOInp     "Name.gbw"  # orbitals used for MORead
  GuessMode FMatrix     # FMatrix projection
            CMatrix     # Corresponding orbital projection
  AutoStart true        # try to use the orbitals from the existing
                        # GBW file of the same name (if possible)
                        # (default)
            false       # don't use orbitals from existing GBW file
end
```

7.6.1 AutoStart feature

Older versions of ORCA always created a new GBW file at the beginning of the run no matter whether a file of the same name existed or perhaps contained orbitals. Now, in the case of single-point calculations the program automatically checks if a `.gbw` file of the same name exists. If yes, the program checks if it contains orbitals and all other necessary information for a restart. If yes, the variable `Guess` is set to `MORead`. The existing `.gbw` file is renamed to `BaseName.ges` and `MOInp` is set to this filename. If the `AutoStart` feature is not desired, set `AutoStart false` in the `%scf` block or give the keyword `!NoAutoStart` in the simple input line. Note that `AutoStart` is ignored for geometry optimizations: in this case, using previously converged orbitals contained in a `.gbw` file (of a different name) can be achieved via `MORead` and `MOInp`.

7.6.2 One Electron Matrix Guess

The simplest guess is to diagonalize the one electron matrix to obtain starting orbitals. This guess is very simple but usually also a disaster because it produces orbitals that are far too compact.

7.6.3 Basis Set Projection

The remaining guesses (may) need the projection of initial guess orbitals onto the actual basis set. In ORCA there are two ways this can be done. `GuessMode FMatrix` and `GuessMode CMatrix`. The results from the two methods are usually rather similar. In certain cases `GuessMode CMatrix` may be preferable. `GuessMode FMatrix` is simpler and faster. In short the `FMatrix` projection defines an effective one electron operator:

$$\hat{f} = \sum_p \varepsilon_p a_p^\dagger a_p \tag{7.44}$$

where the sum is over all orbitals of the initial guess orbital set, $a_p^\dagger$ is the creation operator for an electron in guess MO p, a_p is the corresponding annihilation operator and ε_i is the orbital energy. This effective one electron operator is diagonalized in the actual basis and the eigenvectors are the initial guess orbitals in the target basis. For most wavefunctions this produces a fairly reasonable guess.

`CMatrix` is more involved. It uses the theory of corresponding orbitals to fit each MO subspace (occupied, partially occupied or spin-up and spin-down occupied) separately [34, 444]. After fitting the occupied orbitals, the virtual starting orbitals are chosen in the orthogonal complement of the occupied orbitals. In some cases, especially when restarting ROHF calculations, this may be an advantage. Otherwise, it is not expected that `CMatrix` will be grossly superior to `FMatrix` for most cases.

7.6.4 PModel Guess

The PModel guess (chosen by `Guess PModel` in the `%scf` block or simply a keyword line with `!PModel`) is one that is usually considerably successful. It consists of building and diagonalizing a Kohn–Sham matrix with an electron density which consists of the superposition of spherical neutral atoms densities which are predetermined for both relativistic and nonrelativistic methods. This guess is valid for both Hartree–Fock and DFT methods, but not for semiempirical models. However, due to the complexity of the guess it will also take a little computer time (usually less than one SCF iteration). The model densities are available for most atoms of the periodic table and consequently the PModel guess is usually the method of choice (particularly for molecules containing heavy elements) unless you have more accurate starting orbitals available.

7.6.5 Hückel and PAtom Guesses

The extended Hückel guess proceeds by performing a minimal basis extended Hückel calculation and projecting the MOs from this calculation onto the actual basis set using one of the two methods described above. The minimal basis is the STO-3G basis set. The Hückel guess may not be very good because the STO-3G basis set is so poor. There is also accumulating evidence that the superposition of atomic densities produces a fairly good initial guess. The critique of the atomic density method is that the actual shape of the molecule is not taken into account and it is more difficult to reliably define singly occupied orbitals for ROHF calculations or a reasonable spin density for UHF calculations. Therefore ORCA chooses a different way in the `PAtom` guess (which is the default guess): the Hückel calculation is simply carried out *for all electrons* in a minimal basis of atomic SCF orbitals. These were determined once and for all and are stored inside the program. This means that the densities around the atoms are very close to the atomic ones, all orbitals on one center are exactly orthogonal, the initial electron distribution already reflects the molecular shape and there are well defined singly occupied orbitals for ROHF calculations.

7.6.6 Restarting SCF Calculations

To restart SCF calculations, it can be very helpful and time-saving to read in the orbital information of a previous calculation. To do this, specify:

```
! moread
%moinp "name.gbw"
```

This is done by default for single-point calculations if the `.gbw` file of the same name exists.

The program stores the current orbitals in every SCF cycle. Should a job crash, it can be restarted from the orbitals that were present at this time by just re-running the calculation to use the present `.gbw` file. In addition, an effort has

been made to make .gbw files from different releases compatible with each other. If your input .gbw file is from an older release, use ! rescue moread noiter with % moinp "name.gbw" to produce an up-to-date .gbw. When the rescue keyword is invoked, only the orbital coefficients are read from the .gbw file, and everything else from the input file. Thus, make sure that the geometry and the basis set of the old .gbw file and the new input match.

Within the same ORCA version, neither the geometry nor the basis set stored in name.gbw need to match the present geometry or basis set. The program merely checks if the molecules found in the current calculation and name.gbw are consistent with each other and then performs one of the possible orbital projections. If the two basis sets are identical the program by default only reorthogonalizes and renormalizes the input orbitals. However, this can be overruled by explicitly specifying GuessMode in the % scf block as CMatrix or FMatrix.

If redundant components were removed from the basis (see *Linear Dependence*), then ! moread noiter must not be used to read SCF orbitals from a previous calculation, as it is going to lead to wrong results. In that case, ! rescue moread may be used (without noiter) if doing the entire calculation in one go is not possible.

For pre 2.5-03 versions of ORCA the input .gbw file from the earlier calculation must have a different name than the new calculation, because in the very beginning of a calculation, a new .gbw file is written. If the names are the same, the .gbw file from the earlier calculation is overwritten and all information is lost. Therefore, if you want to restart a calculation with an input file of the same name as the previous calculation, you have to rename the .gbw file first. This is good practice anyway to avoid surprises, particularly for expensive calculations.

There is an additional aspect of restarting SCF calculations — if you have chosen SCFMode = Conventional the program stores a large number of integrals that might have been time consuming to calculate on disk. Normally the program deletes these integrals at the end of the calculation. However, if you want to do a closely related calculation that requires the same integrals (i.e. the *geometry*, the *basis set* and the *threshold* Thresh are the same) it is faster to use the integrals generated previously. This is done by using KeepInts = true in the % scf block of the first calculation and then use ReadInts = true in the % scf block of the second calculation. If the second calculation has a different name than the first calculation you have to use IntName = "FirstName" to tell the program the name of the integral files. Note that the file containing the integrals does not have an extension — it is simply the name of the previous input file with .inp stripped off.

```
%scf
  KeepInts true         # Keep integrals on disk
  ReadInts true         # Read integrals from disk
  IntName  "MyInts"     # Name of the integral files without extension
end
```

Note that, in general, restarting calculations with old integral files requires the awareness and responsibility of the user. If properly used, this feature can save considerable amounts of time.

7.6.7 Changing the Order of Initial Guess MOs and Breaking the Initial Guess Symmetry

Occasionally you will want to change the order of initial guess MOs — be it because the initial guess yielded an erroneous occupation pattern or because you want to converge to a different electronic state using the orbitals of a previous calculation. Reordering of MOs and other tasks (like breaking the symmetry of the electronic wavefunction) are conveniently handled with the Rotate feature in ORCA. Rotate is a subblock of the SCF block that allows you to linearly transform pairs of MOs.

```
%scf
  Rotate
   { MO1, MO2, Angle }
   { MO1, MO2, Angle, Operator1, Operator2 }
   { MO1, MO2} # Shortcut to swap MO1 and MO2. Angle=90 degrees.
  end
end
```

Here, MO1 and MO2 are the indices of the two MOs of interest. Recall that ORCA starts counting MOs with index 0, i.e. the MO with index 1 is the *second* MO. Angle is the rotation angle in degrees. A rotation angle of 90°

corresponds to flipping two MOs, an angle of 45° leads to a 50:50 mixture of two MOs, and a 180° rotation leads to a change of phase. `Operator1` and `Operator2` are the orbitals sets for the rotation. For UHF calculations spin-up orbitals belong to operator 0 and spin-down orbitals to operator 1. RHF and ROHF calculations only have a single orbital set.

Among others, the `Rotate` feature can be used to produce broken-symmetry solutions, for example in transition metal dimers. In order to do that, first perform a high-spin calculation, then find the pairs of MOs that are symmetric and antisymmetric combinations of each other. Take these MOs as the initial guess and use rotations of 45° for each pair to localize the starting MOs. If you are lucky and the broken symmetry solution exists, you have a good chance of finding it this way. See section *Broken-Symmetry Wavefunctions and Exchange Couplings* for more details on the broken-symmetry approach.

7.6.8 Automatically Breaking of the Initial Guess Symmetry

Another simple way to break the initial guess symmetry for more trivial cases, is to simply use the keyword `! GUESSMIX`. This will automatically mix 50% of the alpha LUMO into the alpha HOMO. That is equivalent to a 45 degree rotation as done above and only for these orbitals. It might be useful when one wants an open-shell singlet and needs the alpha and beta orbitals to start differently.

The specific angle of rotation can be controlled with:

```
%scf
  guessmix 75 # angle in degrees, default is 45
end
```

7.6.9 Calculating only the energy of an input density

In case you want to give the result of a previous SCF and recalculate the energy, or maybe some other property (like the MP2 energy) using that density without changing the orbitals, you can use the flags `!CALCGUESSENERGY NOITER`.

The SCF program will read the orbitals, compute the density and one Fock matrix necessary to get the energy and move on with no orbital updates. This can be used to combine DFT orbitals with DLPNO-CCSD(T) for example. Be careful with the results you get from this because these orbitals are not variational anymore!

7.7 SCF Convergence

SCF convergence is a pressing problem in any electronic structure package because the total execution times increases linearly with the number of iterations. Thus, it remains true that the best way to enhance the performance of an SCF program is to make it converge better. In some cases, especially for open-shell transition metal complexes, convergence may be very difficult. ORCA makes a dedicated effort to achieve reasonable SCF convergence for these cases without compromising efficiency.

Another issue is whether the solution found by ORCA is stable, i.e. a minimum on the surface of orbital rotations. Especially for open-shell singlets it can be hard to achieve a broken-symmetry solution. The SCF stability analysis (section *SCF Stability Analysis*) may be able to help in such situations. Please also note that if `! TRAH` is used the solution must be a true local minimum though not necessarily a global.

7.7.1 Convergence Tolerances

Before discussing how to converge a SCF calculation it should be defined what is meant by "converged". ORCA has a variety of options to control the target precision of the energy and the wavefunction that can be selected in the `% scf` block, or with a simple input line keyword that merges the criterion label with "SCF", e.g. `! StrongSCF` or `! VeryTightSCF`:

```
%scf
  Convergence             # The default convergence is between medium and strong
               Sloppy     # very weak convergence
               Loose      # still weak convergence
               Medium     # intermediate accuracy
               Strong     # stronger
               Tight      # still stronger
               VeryTight  # even stronger
               Extreme    # close to numerical zero of the computer
                          # in double precision arithmetic
 end
```

Like other keys, `Convergence` is a compound key that assigns default values to a variety of other variables given in the box below. In table *Threshold choices for compound convergence keys* we present the chosen values for each compound key. If the corresponding simple inputs are given (`StrongSCF`, `VeryTightSCF`, ...etc), then in addition the values of table *Additional threshold choices set by the simple input keys (strongSCF, ... etc.)* are also set. The default convergence criteria are reasonable and should be sufficient for most purposes. For a cursory look at populations weaker convergence may be sufficient, whereas other cases may require stronger than default convergence. Note that `Convergence` does not only affect the target convergence tolerances but also the integral accuracy as discussed in the section about direct SCF and alike. **This is very important: if the error in the integrals is larger than the convergence criterion, a direct SCF calculation cannot possibly converge.**

The convergence criteria are always printed in the output. Given below is a list of the convergence criteria for `! TightSCF`, which is often used for calculations on transition metal complexes.

```
%scf
  TolE        1e-8  # energy change between two cycles
  TolRMSP     5e-9  # RMS density change
  TolMaxP     1e-7  # maximum density change
  TolErr      5e-7  # DIIS error convergence
  TolG        1e-5  # orbital gradient convergence
  TolX        1e-5  # orbital rotation angle convergence
  ConvCheckMode  2  # = 0: check all convergence criteria
                    # = 1: stop if one of criterion is met, this is sloppy!
                    # = 2: check change in total energy and in one-electron energy
                    #      Converged if delta(Etot)<TolE and delta(E1)<1e3*TolE
  ConvForced        # = 0: convergence not mandatory for next calculation step
                    # = 1: break, if you did not meet the convergence criteria
end
```

Table 7.11: Threshold choices for compound convergence keys

Sloppy	
TolE	3e-5
TolMAXP	1e-4
TolRMSP	1e-5
TolErr	1e-4
Thresh	1e-9
TCut	1e-10
DFTGrid.BFCut	1e-10

continues on next page

Table 7.11 – continued from previous page

TolG	3e-4
TolX	3e-4
Z_Tol	5e-3
Loose	
TolE	1e-5
TolMAXP	1e-3
TolRMSP	1e-4
TolErr	5e-4
Thresh	1e-9
TCut	1e-10
DFTGrid.BFCut	1e-10
TolG	1e-4
TolX	1e-4
Z_Tol	3e-3
Medium	
SCFConvMode	_CONVCHECK_ENERGY
TolE	1e-6
TolMAXP	1e-5
TolRMSP	1e-6
TolErr	1e-5
Thresh	1e-10
TCut	1e-11
DFTGrid.BFCut	1e-10
TolG	5e-5
TolX	5e-5
Z_Tol	1e-3
Strong	
SCFConvMode	_CONVCHECK_ENERGY
TolE	3e-7
TolMAXP	3e-6
TolRMSP	1e-7
TolErr	3e-6
Thresh	1e-10
TCut	3e-11
DFTGrid.BFCut	3e-11
TolG	2e-5
TolX	2e-5
Z_Tol	7e-4
Tight	
SCFConvMode	_CONVCHECK_ENERGY
TolE	1e-8
TolMAXP	1e-7
TolRMSP	5e-9
TolErr	5e-7
Thresh	2.5e-11
TCut	2.5e-12
DFTGrid.BFCut	1e-11
TolG	1e-5
TolX	1e-5
Z_Tol	1e-4
VeryTight	

continues on next page

Table 7.11 – continued from previous page

SCFConvMode	_CONVCHECK_ENERGY
TolE	1e-9
TolMAXP	1e-8
TolRMSP	1e-9
TolErr	1e-8
Thresh	1e-12
TCut	1e-14
DFTGrid.BFCut	1e-12
TolG	2e-6
TolX	2e-6
Z_Tol	3e-5
Extreme	
SCFConvMode	_CONVCHECK_ALL
TolE	1e-14
TolMAXP	1e-14
TolRMSP	1e-14
TolErr	1e-14
Thresh	3e-16
TCut	3e-16
TolG	1e-09
TolX	1e-09
Z_Tol	3e-06
DFTGrid.BFCut	3e-16

Table 7.12: Additional threshold choices set by the simple input keys (strongSCF, . . . etc.)

SloppySCF	
DCAS.TolG	5.0e-3
DCAS.TolE	1.0e-6
DMDCI.STol	1.0e-4
DMRCI.ETol	1.0e-5
DMRCI.RTol	1.0e-5
DCIS.ETol	1.0e-5
DCIS.RTol	1.0e-5
LooseSCF	
IN.DCAS.TolG	5.0e-3
DCAS.TolE	1.0e-6
DMDCI.STol	1.0e-4
DMRCI.ETol	1.0e-5
DMRCI.RTol	1.0e-5
DCIS.ETol	1.0e-5
DCIS.RTol	1.0e-5
NORMALSCF	
DCAS.TolG	1.0e-3
DCAS.TolE	1.0e-7
DMDCI.STol	2.5e-5

continues on next page

Table 7.12 – continued from previous page

DMRCI.ETol	1.0e-6
DMRCI.RTol	1.0e-6
DCIS.ETol	1.0e-6
DCIS.RTol	1.0e-6
STRONGSCF	
DCAS.TolG	5.00e-4
DCAS.TolE	6.66e-8
DMDCI.STol	7.50e-6
DMRCI.ETol	6.66e-7
DMRCI.RTol	6.66e-7
DCIS.ETol	6.66e-7
DCIS.RTol	6.66e-7
TIGHTSCF	
DCAS.TolG	2.5e-4
DCAS.TolE	2.5e-8
DMDCI.STol	1.0e-5
DMRCI.ETol	2.5e-7
DMRCI.RTol	2.5e-7
DCIS.ETol	2.5e-7
DCIS.RTol	2.5e-7
VERYTIGHTSCF	
DCAS.TolG	1.0e-5
DCAS.TolE	1.0e-8
DMDCI.STol	1.0e-6
DMRCI.ETol	1.0e-7
DMRCI.RTol	1.0e-7
DCIS.ETol	1.0e-7
DCIS.RTol	1.0e-7
EXTREMESCF	
DCAS.TolG	1.0e-9
DCAS.TolE	1.0e-12
DMDCI.STol	1.0e-9
DMDCI.TCutInt	0.0
DMRCI.ETol	1.0e-12
DMRCI.RTol	1.0e-12
DCIS.ETol	1.0e-12
DCIS.RTol	1.0e-12

If `ConvCheckMode=0`, *all* convergence criteria have to be satisfied for the program to accept the calculation as converged, which is a quite rigorous criterion. In this mode, the program also has mechanisms to decide that a calculation is converged even if one convergence criterion is not fulfilled but the others are overachieved. `ConvCheckMode=1` means that one criterion is enough. This is quite dangerous, so ensure that none of the criteria are too weak, otherwise the result will be unreliable. The default `ConvCheckMode=2` is a check of medium rigor — the program checks for the change in total energy and for the change in the one-electron energy. If the ratio of total energy and one-electron energy is constant, the self-consistent field does not fluctuate anymore and the calculation can be considered converged. If you have small eigenvalues of the overlap matrix, the density may not be converged to the number of significant figures requested by `TolMaxP` and `TolRMSP`.

`ConvForced` is a flag to prevent time consuming calculations on non-converged wave functions. It will default to `ConvForced=1` for Post-HF methods, Excited States runs and Broken Symmetry calculations. You can overwrite this default behavior by setting `ConvForced=0`.

Irrespective of the ConvForced value that has been chosen, properties or numerical calculations (NumGrad, NumFreq) will not be performed on non-converged wavefunctions!

7.7.2 Dynamic and Static Damping

Damping is the oldest and simplest convergence aid. It was already invented by Douglas Hartree when he did his famous atomic calculations. Damping consists of mixing the old density with the new density as:

$$P_{\text{new, damped}} = (1 - \alpha)\, P_{\text{new}} + \alpha P_{\text{old}} \tag{7.45}$$

where α is the damping factor, which must have a value of less than 1. Thus the permissible range (not checked by the program) is $0 \ldots 0.999999$. For α values larger than 1, the calculation cannot proceed since no new density is admixed. Damping is important in the early stages of a calculation where P_{old} and P_{new} are very different from each other and the energy is strongly fluctuating. Many schemes have been suggested that vary the damping factor dynamically to give strong damping in the beginning and no damping in the end of an SCF. The scheme implemented in ORCA is that by Hehenberger and Zerner [910] and is invoked with CNVZerner=true. Static damping is invoked with CNVDamp=true. These convergers are mutually exclusive. They can be used in the beginning of a calculation when it is not within the convergence radius of DIIS or SOSCF. Damping works reasonably well, but most other convergers in ORCA are more powerful.

If damping used in conjunction with DIIS or SOSCF, the value of DampErr is important: once the DIIS error falls below DampErr, the damping is turned off. In case SOSCF is used, DampErr refers to the orbital gradient value at which the damping is turned off. The default value is 0.1 Eh. In difficult cases, however, it is a good idea to choose DampErr much smaller, e.g. 0.001. This is — to some extent — chosen automatically together with the keyword ! SlowConv.

```
%scf
  # control of the Damping procedure
  CNVDamp   true    # default: true
  CNVZerner false   # default: false
  DampFac   0.98    # default: 0.7
  DampErr   0.05    # default: 0.1
  DampMin   0.1     # default: 0.0
  DampMax   0.99    # default: 0.98
  # more convenient:
  Damp fac 0.98 ErrOff 0.05 Min 0.1 Max 0.99 end
end
```

7.7.3 Level Shifting

Level shifting is a frequently used technique. The basic idea is to shift the energies of the virtual orbitals such that after diagonalization the occupied and virtual orbitals mix less strongly and the calculation converges more smoothly towards the desired state. Also, level shifting should prevent flipping of electronic states in near-degenerate cases. In a special context it has been shown by Saunders and Hillier [335, 750] to be equivalent to damping.

Similar to DampErr described in the previous section, ShiftErr refers to the DIIS error at which the level shifting is turned off.

```
%scf
  # control of the level shift procedure
  CNVShift   true  # default: true
  LShift     0.1   # default: 0.25, energy unit is Eh.
  ShiftErr   0.1   # default: 0.0
  # more convenient:
  Shift Shift 0.1 ErrOff 0.1 end
end
```

7.7.4 Direct Inversion in Iterative Subspace (DIIS)

The direct inversion in iterative subspace (DIIS) is a technique that was invented by Pulay [701, 702]. It has become the *de facto* standard in most modern electronic structure programs, because DIIS is robust, efficient and easy to implement. Basically DIIS uses a criterion to judge how far a given trial density is from self-consistency. The commutator of the Fock and density matrices [**F**,**P**] is a convenient measure for this error. With this information, an extrapolated Fock matrix from the present and previous Fock matrices is constructed, which should be much closer to self-consistency. In practice this is usually true, and better than linear convergence has been observed with DIIS. In some rare (open-shell) cases however, DIIS convergence is slow or absent after some initial progress. As self-consistency is approached, the set of linear equations to be solved for DIIS approaches linear dependency and it is useful to bias DIIS in favor of the SCF cycle that had the lowest energy using the factor `DIISBfac`. This is achieved by multiplying all diagonal elements of the DIIS matrix with this factor unless it is the Fock matrix/density which leads to the lowest energy. The default value for `DIISBfac` is 1.05.

The value of `DIISMaxEq` is the maximum number of old Fock matrices to remember. Values of 5-7 have been recommended, while other users store 10-15 Fock matrices. Should the standard DIIS not achieve convergence, some experimentation with this parameter can be worthwhile. In cases where DIIS causes problems in the beginning of the SCF, it may have to be invoked at a later stage. The start of the DIIS procedure is controlled by `DIISStart`. It has a default value of 0.2 Eh, which usually starts DIIS after 0-3 cycles. A different way of controlling the DIIS start is adjusting the value `DIISMaxIt`, which sets the maximum number of cycles after which DIIS will be started irrespective of the error value.

```
%scf
  # control of the DIIS procedure
  CNVDIIS     true  # default: true
  DIISStart   0.1   # default: 0.2
  DIISMaxIt   5     # default: 12
  DIISMaxEq   7     # default: 5
  DIISBFac    1.2   # default: 1.05
  DIISMaxC    15.0  # default: 10.0
  # more convenient:
  DIIS Start 0.1  MaxIt 5  MaxEq 7  BFac 1.2  MaxC 15.0  end
end
```

Note that for troublesome or lacking SCF convergence the `TRAH` algorithm should be used (see Sec. *Trust-Region Augmented Hessian (TRAH) SCF*). If not turned off explicitly, `TRAH` is switched on automatically whenever convergence problems are present by means of the `AutoTRAH` feature (see Sec. *Trust-Region Augmented Hessian (TRAH) SCF*).

7.7.5 An alternative DIIS algorithm: KDIIS

An alternative algorithm that makes use of the DIIS concept is called KDIIS (Kolmar's DIIS[452, 453]) in ORCA. The KDIIS algorithm is designed to bring the orbital gradient of any energy expression to zero using a combination of DIIS extrapolation and first order perturbation theory. Thus, the method is diagonalization-free. In our hands it is superior to the standard DIIS algorithm in many cases, but not always. The algorithm is invoked with the keyword `! KDIIS` and is available for RHF, UHF and CASSCF.

7.7.6 Approximate Second Order SCF (SOSCF)

SOSCF is an approximately quadratically convergent variant of the SCF procedure [264, 607]. The theory is relatively involved and will not be described here. In short – SOSCF computes an initial guess to the inverse orbital Hessian and then uses the BFGS formula in a recursive way to update orbital rotation angles. As information from a few iterations accumulates, the guess to the inverse orbital Hessian becomes better and better and the calculation reaches a regime where it converges superlinearly. As implemented, the procedure converges as well or slightly better than DIIS and takes a somewhat less time. However, it is also *a lot* less robust, so that DIIS is the method of choice for many problems (see also the description of the full second-order trust-region augmented Hessian (`TRAH`) procedure in the next section). On the other hand, SOSCF is useful when DIIS gets stuck at some error around $\sim$ 0.001 or 0.0001. Such cases were the primary motive for the implementation of SOSCF into ORCA.

The drawback of SOSCF is the following: in the beginning of the SCF, the orbital gradient (the derivative of the total energy with respect to rotations that describe the mixing of occupied and virtual MOs) is large, so that one is far from the quadratic regime. In such cases, the procedure is not successful and may even wildly diverge. Therefore it is recommended to only invoke the SOSCF procedure in the very end of the SCF where DIIS may lead to "trailing" convergence. SOSCF is controlled by the variables `SOSCFStart` and `SOSCFMaxIt`. `SOSCFStart` is a threshold for the orbital gradient. When the orbital gradient, or equivalently the DIIS Error, fall below `SOSCFStart`, the SOSCF procedure is initiated. `SOSCFMaxIt` is the latest iteration to start the SOSCF even if the orbital gradient is still above `SOSCFStart`.

```
%scf
  # control of the SOSCF procedure
  CNVSOSCF    true  # default: false
  SOSCFStart  0.1   # default: 0.01
  SOSCFMaxIt  5     # default: 1000
  # more convenient:
  SOSCF Start 0.1  MaxIt 5  end
end
```

For many calculations on transition metal complexes, it is a good idea to be conservative in the startup criterion for SOSCF, it may diverge otherwise. A choice of 0.01 or lower is recommended.

7.7.7 Trust-Region Augmented Hessian (TRAH) SCF

The trust-region augmented Hessian (`! TRAH`) approach[64, 349, 381, 734] should be used if the standard SCF solver[264, 607, 702, 703] DIIS+SOSCF fails or is expected to fail. TRAH-SCF should converge for any system. Convergence to the electronic ground state is also guaranteed because information of the electronic Hessian is exploited.

The TRAH approach constructs a quadratic model of the SCF energy as a function of the orbital rotation parameters $\mathbf{x}$,

$$E(\mathbf{x}) = E_0 + \mathbf{g}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x}$$

and minimizes $E(\mathbf{x})$ w.r.t $\mathbf{x}$ with the constraint that orbital rotations should lie within a trust region h

$$||\mathbf{x}|| \leq h\ .$$

Such a constraint minimization leads to the level-shifted Newton equations

$$(\mathbf{H} - \mu\mathbf{I})\,\mathbf{x} = -\mathbf{g},$$

which have the two unknowns $(\mu, \mathbf{x})$. Instead of solving the level-shifted Newtons equations, the eigenvalues and eigenvectors of the scaled augmented Hessian are solved,

$$\begin{pmatrix} 0 & \alpha\mathbf{g}^T \\ \alpha\mathbf{g} & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \tilde{\mathbf{x}} \end{pmatrix} = \mu \begin{pmatrix} 1 \\ \tilde{\mathbf{x}} \end{pmatrix},$$

The TRAH eigenvalue equations are solved iteratively with the Davidson algorithm until the residual norm for $\tilde{\mathbf{x}}$ is below a scaled (by `TolFacMicro`) norm of the electronic gradient. The scaling parameter α is adjusted in every Davidson iteration (micro iteration) such that

$$||\mathbf{x}|| \approx ||\frac{1}{\alpha}\tilde{\mathbf{x}}|| \leq h$$

by using a bisection search within $[\alpha_0, \alpha_1]$ and ensures that the orbital rotation (update) vectors are within the trust region. Once $\mathbf{x}$ is found, the orbitals are updated and a new macro iteration starts with the SCF energy and electronic gradient computation $\mathbf{g}$. TRAH terminates if the gradient norm $||\mathbf{g}||$ is below a user-given threshold `TolG`.

The most time consuming steps of the algorithm are the computation of the electronic gradient $\mathbf{g}$ and the linear transformations of the electronic Hessian $\mathbf{H}$ with some trial vectors during the Davidson micro iterations. Both

intermediates can be computed efficiently in the atomic-orbital basis using AO-Fock matrices as done for TD-DFT or CP-SCF. Hence, TRAH-SCF can be also used for very large molecules. However, in contrast to a standard DIIS approach, difference density matrices cannot be used which makes the shell pair screening based on Schwarz estimates and density matrices less effective for TRAH. To accelerate the various sigma vector computations, we choose for those steps in the micro iterations as for CP-SCF smaller grids for evaluation of XC functionals and semi-numerical exchange COSX, which is controlled via

```
%method
  Z_GridXC   1      // Lebedev Grid
  Z_IntAccXC 3.467  // eps parameter of radial grid
  Z_GridX    1      // Lebedev Grid
  Z_IntAccX  3.067  // eps parameter of radial grid
end
```

TRAH-SCF is currently implemented for restricted closed-shell (RHF and RKS) and unrestricted open-shell determinants (UHF and UKS) and can be accelerated with RIJ, RIJONX, RIJK, or RIJCOSX. Solvation effects can also be accounted for with the C-PCM model. The implementation is also parallelized with MPI. A restricted open-shell implementation (ROHF and ROKS) is not yet available.

The default preconditioner (diag) for the Davidson algorithm uses approximate matrix element of the diagonal Hessian. We have also added an improved preconditioner (full) that uses the exact orbital Hessian for a subset of the most important occupied-virtual MO pairs (with smallest orbital-energy difference). This number PreconMaxRed (default 250) cannot be too large.

```
%scf
  trah
    Precond         Diag   # DIAG, FULL, or NONE
    PreconMaxRed    250    # maximum dimension of FULL Hessian for preconditioning
  end
end
```

Otherwise, additional computational bottlenecks would be introduced when transforming the two-electron integrals in the MO basis or when diagonalizing this reduced-space Hessian. Note that for the integral transformation the RI approximation is used and an auxiliary /C basis must be provided. Please also note that, so far, we have only implemented an XC Hessian contribution for LDA functionals. From our experience, the "full" precondition is very advantageous for RHF and UHF calculations of small molecules but does not provide any advantage for other (TRAH-)SCF calculations.

In cases for which the conventional SCF procedures (DIIS/KDIIS/SOSCF) struggle, we invoke TRAH-SCF automatically (AutoTRAH). For this purpose, we perform a linear interpolation of the norm of the electronic gradient on the $\log_{10}$ scale after a minimum number of SCF iterations AutoTRAHIter (default 20). The number of iterations for interpolations is controlled by AutoTRAHNInter (default 10). If the slope s of the interpolated gradient norm 10^{-s} becomes smaller than AutoTRAHTol (default 1.125), conventional SCF is shut down and TRAH-SCF start from the current set of MOs. Those parameters were optimized for a benchmark set with the purpose to minimize the calculation times even for SCF calculations that are hard to converge. There is not really a need to modify the AutoTRAH parameters except for turning TRAH off entirely.

```
! NoTRAH
```

The accuracy of the SCF calculation is controlled by via the simple keywords NORMALSCF, LOOSESCF, etc. The accuracy threshold that checks the gradient norm for TRAH-SCF calculations is is read from the SCF input block. Note that checking the Frobensius norm of $||\mathbf{g}||$ is the **only convergence check** in TRAH.

```
%scf
  TolG 1.e-6  # gradient norm threshold (converging macro iterations)
end
```

Below a complete list of input parameters is given for TRAH. Please note that all parameters influence the convergence and should not be changed carelessly!

```
%scf
 # AutoTRAH parameter
 AutoTRAH       true
 AutoTRAHTol    1.125
 AutoTRAHIter   20
 AutoTRAHNInter 10

 # TRAH parameter
   trah
     MaxRed          16     # maximum number of Davidson micro iterations
     NStart          2      # number of start vectors for Davidson (at least 2)
     TolFacMicro     0.1    # Scaling factor for Davidson convergence
                            #  threshold =  TolFacMicro * || G ||
     MinTolMicro     1.e-2  # minimum accuracy of micro iterations
     QuadRegionStart 1.e-3  # start Newton-Raphson if || G || < QuadRegionStart
     tradius         0.4    # initial trust radius
     AlphaMin        0.1    # lower bound of gradient scaling parameter
     AlphaMax        1000.  # upper bound of gradient scaling parameter
     Randomize       true   # add white noise to Davidson start vectors
     PseudoRand      false  # use pseudo random numbers for comparibility
     MaxNoise        1.e-2  # maximum random number magnitude
     OrbUpdate       Taylor # orbital update algorithm (TAYLOR or CAYLEY)
     InactiveMOs     Canonical # CANONICAL or NOTSET
     Precond         Diag   # DIAG, FULL, or NONE
     PreconMaxRed    250    # maximum dimension of FULL Hessian for preconditioning
   end
 end
```

OBS.: The maximum number of macro iterations is defined by MAXITER under the %SCF block.

7.7.8 Finite Temperature HF/KS-DFT

A finite temperature can be used to apply a Fermi-like occupation number smearing over the orbitals of the system, which may sometimes help to get convergence of the SCF equations in near hopeless cases. Through the smearing, the electrons are distributed according to Fermi statistics among the available orbitals. The "chemical potential" is found through the condition that the total number of electrons remains correct. Gradients can be computed in the presence of occupation number smearing.

```
%scf SmearTemp 5000  # ``temperature'' in Kelvin
 end
```

Note

- Finite temperature SCF (fractional occupation numbers or FOD analysis, see sections *Fractional Occupation Numbers* and *Fractional Occupation Number Weighted Electron Density (FOD)*, respectively) *cannot* be used together with the `CNVRico` or `SOSCF` methods.

Fractional Occupation Numbers

Only a very basic implementation of fractional occupation numbers is presently provided. It is meant to deal with orbitally degenerate states in the UHF/UKS method. Mainly it was implemented to avoid symmetry breaking in DFT calculations on orbitally degenerate molecules and atoms. The program checks the orbital energies of the initial guess orbitals, finds degenerate sets and averages the occupation numbers among them. Currently the criterion for degenerate orbitals is 10^{-3} Eh. The fractional occupation number option is invoked by:

```
%scf
  FracOcc true
end
```

Clearly, the power of fractional occupation numbers goes far beyond what is presently implemented in the program and future releases will likely make more use of them. The program prints a warning whenever it uses fractional occupation numbers. The fractionally occupied orbitals should be checked to ensure they are actually the intended ones.

Note

- Using `GuessMode = CMatrix` will cause problems because there are no orbital energies for the initial guess orbitals. The program will then average over *all* orbitals — which makes no sense at all.

Fractional Occupation Number Weighted Electron Density (FOD)

Many approximate QC methods do not yield reliable results for systems with significant static electron correlation (SEC) but it is often difficult to predict if the system in question suffers from SEC or not. Existing scalar SEC diagnostics (e.g., the T_1 diagnostic) do not provide any information where the SEC is located in the molecule. Furthermore, often quite expensive calculations have to be performed first (e.g., CCSD) in order to judge the reliability of the results based on a single number. Molecular systems with strong SEC (e.g. covalent bond-breaking, biradicals, open-shell transition metal complexes) are usually characterized by small energy gaps between frontier orbitals, and hence, the appearance of many equally important determinants in their electronic wavefunction. This finding is used in the FOD analysis[327] which is based on finite temperature KS-DFT where the fractional occupation numbers are determined from the Fermi distribution ("Fermi smearing")

$$f_i = \frac{1}{e^{(\varepsilon_i - E_F)/kT_{el}} + 1}$$

The central quantity of the FOD analysis is the fractional occupation number weighted electron density (ρ^{FOD}), a real-space function of the position vector r:

$$\rho^{FOD}(r) = \sum_i^N (\delta_1 - \delta_2 f_i)|\varphi_i(r)|^2$$

(δ_1 and δ_2 are unity if the level is lower than E_F while they are 0 and -1, respectively, for levels higher than E_F). The f_i represent the fractional occupation numbers ($0 \leq f_i \leq 1$; sum over all electronic single-particle levels obtained by solving self-consistently the KS-SCF equations minimizing the *free*-electronic energy).

$\rho^{FOD}(r)$ can be plotted using a pre-defined contour surface value (see *FOD plots*). FOD plots only show the contribution of the 'hot' (strongly correlated) electrons and can thus be used to choose a suitable QC method for the system in question based on some rules of thumb (see *FOD plots*). Mulliken reduced orbital charges based on $\rho^{FOD}(r)$ (see *Mulliken Population Analysis*) offer a fast alternative to get the information of the FOD plot.

The integration of ρ^{FOD} over all space yields as additional information a single size-extensive number termed N_{FOD} which correlates well with other scalar SEC diagnostics and can be used to globally quantify SEC effects in the molecule.

ρ^{FOD} (and N_{FOD}) strongly depend on the orbital energy gap which itself depends almost linearly on the amount of the non-local Fock exchange admixture a_x. The following (empirical) function of the optimal electronic temperature T_{el} on a_x

$$T_{el} = 20000\,\mathrm{K} \times a_x + 5000\,\mathrm{K}$$

is used to ensure that similar results of the FOD analysis are obtained with various functionals. For example, the SmearTemp has to be 5000 K for TPSS ($a_x = 0$), 9000 K for B3LYP ($a_x = 20\%$), 10000 K for PBE0 ($a_x = 25\%$), and 15800 K for M06-2x ($a_x = 54\%$). The result of the FOD analysis is not strongly dependent on the employed basis set (see supplementary information of the original publication[327]). TPSS/def2-TZVP/TightSCF was chosen as the default since it is fast and robust. The FOD analysis is a very efficient and practicable tool to get information about the amount and localization of SEC in the system of question. It is called by a simple keyword:

```
# ground state of p-benzyne
! FOD

* xyz 0 1
C 0.0000000  1.2077612  0.7161013
C 0.0000000  0.0000000  1.3596219
C 0.0000000 -1.2077612  0.7161013
C 0.0000000 -1.2077612 -0.7161013
C 0.0000000  0.0000000 -1.3596219
C 0.0000000  1.2077612 -0.7161013
H 0.0000000  2.1606260  1.2276695
H 0.0000000 -2.1606260  1.2276695
H 0.0000000 -2.1606260 -1.2276695
H 0.0000000  2.1606260 -1.2276695
*
```

The respective output reads:

```
-------------------------------------------------------------------------------------------
                                    ORCA LEAN-SCF
                               memory conserving SCF solver
-------------------------------------------------------------------------------------------


----------------------------------------D-I-I-S--------------------------------------------
Iteration    Energy (Eh)           Delta-E    RMSDP     MaxDP     DIISErr   Damp  Time(sec)
-------------------------------------------------------------------------------------------
               ***  Starting incremental Fock matrix formation  ***
    1    -230.8982516003082139     0.00e+00  5.01e-03  1.01e-01  1.12e-01  0.700   1.7
Warning: op=0 Small HOMO/LUMO gap (  -0.021) - skipping pre-diagonalization
         Will do a full diagonalization
    2    -230.9463607195993120    -4.81e-02  1.15e-03  2.59e-02  4.06e-02  0.700   1.6
                              ***Turning on AO-DIIS***
  ... etc.
   12    -231.0033984839932089    -5.02e-09  3.33e-07  7.37e-06  7.95e-06  0.000   1.1
                 **** Energy Check signals convergence ****

FOD:
 Fermi smearing:E(HOMO(Eh)) = -0.201252 MUE = -0.179318 gap=   1.119 eV

 N_FOD =  0.920364
```

The default functional and basis set are TPSS and def2-TZVP respectively. If the FOD analysis should be done employing a different functional, one has to explicitly specify the functional and basis set in the simple keyword line and adjust the SmearTemp accordingly.

```
# ozone molecule
! B3LYP def2-TZVP TightSCF

%scf
 SmearTemp 9000
end

* xyz 0 1
  O   0.00000000017911      0.00000000000000      0.43702029765776
  O  -1.09512651993192      0.00000000000000     -0.21851064888325
```

(continues on next page)

(continued from previous page)

```
  O    1.09512651975281      0.00000000000000     -0.21851064877451
*
```

The FOD analysis may also be useful for finding a suitable active space for e.g. CASSCF calculations.

Note

- The FOD analysis will be always printed (including Mulliken reduced orbital charges based on ρ^{FOD}) if `SmearTemp` > 0 K ρ^{FOD} is stored on disk in the file `Basename.scfp_fod` which is included in the general `Basename.densities` container).
- Since the $\hat{S}^2$ expectation value is not defined for fractional occupation numbers, its printout is omitted.

7.8 Choice of Wavefunction and Integral Handling

7.8.1 Choice of Wavefunction Type

The basic variable that controls the type of wavefunction to be computed is the variable `HFTyp` in the `%scf` block. If nothing is specified for `HFTyp`, the program will check the multiplicity given in the input: for closed-shell molecules with multiplicity 1, RHF/RKS is assumed; for open shell molecules with multiplicity larger than 1, UHF/UKS is invoked. `RHF` will lead to a spin restricted closed-shell type computation [839]. For DFT calculations, `RKS`, `UKS` and `ROKS` can be used as synonyms for `RHF`, `UHF` and `ROHF`. The restricted open-shell DFT method (`ROKS`) is only operative for high-spin states that have n unpaired electrons and $S = n/2$. `UKS` wavefunctions will not be spin-purified.

```
%scf
  HFTyp  RHF     # closed-shell (RKS for DFT)
         UHF     # unrestricted open-shell (UKS for DFT)
         ROHF    # restricted open-shell (ROKS for DFT)
         CASSCF  # complete active space SCF
end
```

In certain cases you may want to run open-shell molecules with RHF/RKS to get a "half-electron" type wavefunction [205]. The total energy is not corrected! Sometimes these half-electron computations lead to acceptable convergence, and the resulting orbitals may be used as input for ROHF, UHF or MRCI calculations. Especially for transition metal complexes the orbitals are quite different from ROHF or UHF orbitals, so that it is not recommended to over-interpret the wavefunctions from such calculations. The calculation is set up in the following way:

```
%method AllowRHF true end
# or simply: ! AllowRHF
```

7.8.2 ROHF Options

For ROHF calculations[102, 112, 125, 141, 243, 452, 453, 574, 597] the program will try to figure out what type of open-shell situation is present on the basis of the initial guess orbitals and their energies. Most "simple" cases are well recognized, but sometimes a little help from the user is needed.

The simplest ROHF case is the `HIGHSPIN` case, where all unpaired electrons in the open-shell are coupled parallel to each other, resulting in the highest multiplicity possible. The user can request this case as follow:

```
%scf
  HFTyp       ROHF
  ROHF_case   HIGHSPIN
```

(continues on next page)

(continued from previous page)

```
  ROHF_NEl[1] 4          # number of electrons in the open-shell
end
```

The ROHF code also has a very powerful feature that goes back to insights of Mike Zerner [817, 906]. It can average over either *all* states of a given configuration (CAHF) or *all states of a given spin* for a given configuration (SAHF). Especially the SAHF feature gives you easy access to most degenerate high symmetry situations and the orbitals resulting from such calculations will be very convenient as input for CI calculations.

```
%scf
  HFTyp         ROHF
  ROHF_case     CAHF    # configuration averaged HF
                SAHF    # spin averaged HF
  ROHF_NumOp    3       # number of operators (3, 2 or 1)
  ROHF_NOrb[1]  2,1     # number of orbitals in each open-shell
  ROHF_NEl[1]   1,1     # number of electrons in each open-shell
end
```

The hypothetical example below could represent an excited state of an octahedral d^3 transition metal complex. In this case there are five open-shell orbitals. The first three open-shell orbitals contain two electrons and the last two one electron. The input for a SAHF calculation is identical, just replace CAHF with SAHF.

```
%scf
  HFTyp         ROHF
  ROHF_case     CAHF   # configuration averaged HF
  ROHF_NumOp    3      # 3 operators in this case: closed, open1, open2
  ROHF_NOrb[1]  3,2    # 3 orbitals in first open shell, 2 in the second
  ROHF_NEl[1]   2,1    # 2 electrons in first open shell, 1 in the second
end
```

Another feature of the ROHF code is the ability to converge the SCF to a given Configuration State Function (CSF-ROHF) [512]. In this way one can approach results from MCSCF calculations. This can be requested in two ways.

The user can give a specific coupling situation.

```
%scf
  HFTyp      ROHF
  ROHF_CASE  USER_CSF          # User defined CSF
  ROHF_REF   {1 1 -1 -1} end   # CSF to be converged to
end
```

Or the user can give how many orbitals per shell. Where each open-shell will couple with antiparallel spin with the previous one.

```
%scf
  HFTyp       ROHF
  ROHF_CASE   AF_CSF          # User defined CSF
  ROHF_AFORBS 2,2             # Coupling Situation
end
```

As an example, one can think of a Fe(III) dimer, where each center is locally high spin, but they couple antiferromagnetically to each other. In order to get the ROHF solution for this system, first one need a set of guess orbitals. The guess orbitals can be obtained either from the QROs of an UHF calculation, or a high spin ROHF calculation, or even a SAHF or CAHF. Independently of the method used, the orbitals need to be localized and ordered in a way that the 5 3d orbitals of each iron are grouped together in sequence. From this, one can run a CSF-ROHF calculation for the antiferromagnetic CSF as shown bellow:

```
%scf
  HFTyp      ROHF
  ROHF_CASE  USER_CSF                        # User defined CSF
```

(continues on next page)

(continued from previous page)

```
  ROHF_REF   {1 1 1 1 1 -1 -1 -1 -1 -1} end   # 2 open shells coupling antiparallel
end

or

%scf
  HFTyp       ROHF
  ROHF_CASE   AF_CSF                          # User defined CSF
  ROHF_AFORBS 5,5                             # 2 open shells coupling antiparallel
end
```

The CSF-ROHF procedure can recognize doubly occupied and virtual orbitals in the definition of the CSF when the USER_CSF case is invoked. When detected, these orbitals will be rotated out of the open-shells defined in the ROHF method and the calculation will run normally:

```
%scf
  HFTyp      ROHF
  ROHF_CASE  USER_CSF                   # User defined CSF
  ROHF_REF   {1 1 2 1  -1 -1 0 -1} end  # the DOMO will be rotated to the closed-shell and
                                        # the VMO will be rotated to the virtual space.
end
```

The user can also directly input the ROHF variables by means of the ROHFOPT Case User keyword. For example for the high spin case with three electrons in three orbitals gives two operators with vector coupling coefficients $a = 1$ and $b = 2$ (Zerner convention).

```
%scf
  HFTyp      ROHF
  ROHFOP  Case    User  # manual input of ROHF variables
    Nop     2           # number of operators
    Norb[1] 3           # number of open-shell orbitals
    Nel[1]  3           # number of open-shell electrons
    A[1,1]  1           # Coulomb vector in the open shell
    B[1,1]  2           # Exchange vector in the open shell
  end
end
```

One awkward feature of the ROHF theory is that the Fock operator is somewhat arbitrarily defined. Different choices lead to the same wavefunction, but have different convergence properties that may vary from system to system. ORCA thus lets the user choose the desired variant. Playing around with these choices may turn a divergent or slowly converging ROHF calculation into a successful calculation!

The ROHF_Restrict feature is another feature that may be useful. If you suspect that the ROHF calculation does not converge because an open-shell and a closed-shell orbital are flipping back and forth, you can try to avoid this behavior by choosing ROHF_Restrict true. Of course there is no guarantee that it will work, and no guarantee that the system stays in the desired state. However, it decreases the chances of large, uncontrolled steps.

```
%scf
  ROHF_Mode 0   # construct F according to Pulay (default)
            1   # construct F as in the Gamess program
            2   # construct F according to Kollmar

  ROHF_Restrict false # restrict orbital interchanges and off-diagonal elements
                      # (default=false)

  # a complete list of ROHF variables
  ROHFOP
    Case      User   # manual input of ROHF variables
    Nop       2      # number of operators
    Norb[1]   3      # number of open-shell orbitals
    Nel[1]    3      # number of open-shell electrons
```

(continues on next page)

(continued from previous page)

```
    A[1,1]    1      # Coulomb vector in the open shell
    B[1,1]    2      # Exchange vector in the open shell
    Mode      2      # use the Kollmar operator
    Restrict  false  # do not restrict
  end
end
```

7.8.3 UHF Natural Orbitals

The program can produce the UHF natural orbitals (UNOs). With these, the open-shell wavefunction can be pictured conveniently. The syntax is simple:

```
%scf
  UHFNO true
end
# or simply: ! UNO
```

There are various printing options for UNOs described in the output section *Population Analyses and Control of Output*. The UNOs can also be plotted as described in the plots section *Orbital and Density Plots*. In general the program stores a file `BaseName.uno`, where `BaseName` is by default the name of you input file with `.inp` stripped off. Accordingly, the gbw file is named `BaseName.gbw`. The `.uno` file is a normal gbw file that contains the geometry, basis set and the UNO orbitals. It could be used, for example, to start a ROHF calculation.

7.8.4 Integral Handling (Conventional and Direct)

As the number of nonzero integrals grows very rapidly and reaches easily hundreds of millions even with medium sized basis sets in medium sized molecules, storage of all integrals is not generally feasible. This desperate situation prevented SCF calculations on larger molecules for quite some time, so that Almlöf [19, 20, 21] made the insightful suggestion to repeat the integral calculation, which was already the dominant step, in *every SCF cycle* to solve the storage problem. Naively, one would think that this raises the effort for the calculation to $n_{\text{iter}} t_{\text{integrals}}$ (where n_{iter} is the number of iterations and $t_{\text{integrals}}$ is the time needed to generate the nonzero integrals). However, this is not the case because only the change in the Fock matrix is required from one iteration to the next, but not the Fock matrix itself. As the calculations starts to converge, more and more integrals can be skipped. The integral calculation time will still dominate the calculation quite strongly, so that ways to reduce this burden are clearly called for. As integrals are calculated in *batches*[1] the cost of evaluating the given batch of shells p, q, r, s may be estimated as:

$$\text{cost} \approx n_p n_q n_r n_s \left(2l_p + 1\right) \left(2l_q + 1\right) \left(2l_r + 1\right) \left(2l_s + 1\right) \tag{7.46}$$

Here, n_p is the number of primitives involved in shell p, and l_p is the angular momentum for this shell. *Large* integrals are also good candidates for storage, because small changes in the density that multiply large integrals are likely to give a nonzero contribution to the changes in the Fock matrix.

ORCA thus features two possibilities for integral handling, which are controlled by the variable `SCFMode`. In the mode `Conventional`, all integrals above a given threshold are stored on disk (in a compressed format that saves much disk space). In the mode `Direct`, all two-electron integrals are recomputed in each iteration.

Two further variables are of importance: In the `Conventional` mode the program may write enormous amounts of data to disk. To ensure this stays within bounds, the program first performs a so-called "statistics run" that gives a pessimistic estimate of how large the integral files will be. Oftentimes, the program will overestimate the amount of disk space required by a factor of two or more. The maximum amount of disk space that is allowed for the integral files is given by `MaxDisk` (in Megabytes).

On the other hand, if the integral files in `Conventional` run are small enough to fit into the central memory, it is faster to do this since it avoids I/O bottlenecks. The maximum amount of memory allocated for integrals in this

[1] A *batch* is a set of integrals that arises from all components of the shells involved in the integral. For example a $\langle pp|pp\rangle$ batch gives rise to $3 \times 3 \times 3 \times 3 = 81$ integrals due to all possible combinations of p_x, p_y and p_z functions in the four shells. Computations based on batches lead to great computational advantages because the 81 integrals involved in the $\langle pp|pp\rangle$ batch share many common intermediate quantities.

way is specified by MaxIntMem (in Megabytes). If the integral files are larger than MaxIntMem, no integrals will be read into memory.

```
%scf
  MaxIter   100    # Max. no. of SCF iterations
  SCFMode Direct   # default, other choice: Conventional
  Thresh  1e-8     # Threshold for neglecting integrals / Fock matrix contributions
                   # Depends on the chosen convergence tolerance (in Eh).
  TCut    1e-10    # Threshold for neglecting primitive batches. If the prefactor
                   # in the integral is smaller than TCut, the contribution of the
                   # primitive batch to the total batch is neglected.
  UseCheapInts false   # default: false
  DirectResetFreq 20   # default: 15
  MaxDisk  2500        # Max. amount of disk for 2 el. ints. (MB)
  MaxIntMem 400        # Max. amount of RAM for 2 el. ints. (MB)
end
```

The flag UseCheapInts has the following meaning: In a Direct SCF calculation, the oscillations in the total energy and density are initially quite large. High accuracy in the integrals is therefore not crucial. If UseCheapInts is switched on, the program loosens the threshold for the integrals and thus saves a lot of computational time. After having obtained a reasonable initial convergence, the thresholds are tightened to the target accuracy. One pitfall with this method is that the number of cycles required to reach convergence may be larger relative to a calculation with full integral accuracy throughout.[2] When restarting calculations that are close to convergence, it is recommended to switch UseCheapInts off. UseCheapInts has no meaning in a conventional SCF.

The value of DirectResetFreq sets the number of incremental Fock matrix builds after which the program should perform a full Fock matrix build in a Direct SCF calculation. To prevent numerical instabilities that arise from accumulated errors in the recursively build Fock matrix, the value should not be too large, since this will adversely affect the SCF convergence. If the value is too small, the program will update more frequently, but the calculation will take considerably longer, since a full Fock matrix build is more expensive than a recursive one.

The thresholds TCut and Thresh also deserve a closer explanation. Thresh is a threshold that determines when to neglect two-electron integrals. If a given integral is smaller than Thresh Eh, it will not be stored or used in Fock matrix construction. Additionally, contributions to the Fock matrix that are smaller than Thresh Eh will not be calculated in a Direct SCF.

Clearly, it would be wasteful to calculate an integral, then find out it is good for nothing and thus discard it. A useful feature would be an efficient way to estimate the size of the integral *before it is even calculated*, or even have an estimate that is a *rigorous upper bound* on the value of the integral. Häser and Ahlrichs [346] were the first to recognize that such an upper bound is actually rather easy to calculate. They showed that:

$$|\langle ij\,|kl\rangle| \leqslant \sqrt{\langle ij\,|ij\rangle}\sqrt{\langle kl\,|kl\rangle} \tag{7.47}$$

where:

$$\langle ij\,|kl\rangle = \int\int \phi_i\left(\vec{r}_1\right)\phi_j\left(\vec{r}_1\right) r_{12}^{-1}\phi_k\left(\vec{r}_2\right)\phi_l\left(\vec{r}_2\right) d\vec{r}_1 d\vec{r}_2 \tag{7.48}$$

Thus, in order to compute an upper bound for the integral only the right hand side of this equation must be known. This involves only two index quantities, namely the matrix of two center exchange integrals $\langle ij\,|ij\rangle$. These integrals are easy and quick to calculate and they are all $\geqslant 0$ so that there is no trouble with the square root. Thus, one has a powerful device to avoid computation of small integrals. In an actual calculation, the Schwartz prescreening is not used on the level of individual basis functions but on the level of shell batches because integrals are always calculated in batches. To realize this, the largest exchange integral of a given exchange integral block is looked for and its square root is stored in the so called *pre-screening* matrix $\mathbf{K}$ (that is stored on disk in ORCA). In a Direct SCF this matrix is not recalculated in every cycle, but simply read from disk whenever it is needed. The matrix of exchange integrals on the level of individual basis function is used in Conventional calculations to estimate the disk requirements (the "statistics" run).

Once it has been determined that a given integral batch survives it may be calculated as:

$$\langle ij\,|kl\rangle = \sum_p d_{pi}\sum_q d_{qj}\sum_r d_{kr}\sum_s d_{sl}\,\langle i_p j_q\,|k_r l_s\rangle \tag{7.49}$$

[2] This might be an undesirable feature of the current implementation.

where the sums p, q, r, s run over the primitive Gaussians in each basis function i, j, k, l and the d's are the contraction coefficients. There are more powerful algorithms than this one and they are also used in ORCA. However, if many terms in the sum can be skipped and the total angular momentum is low, it is still worthwhile to compute contracted integrals in this straightforward way. In equation (7.49), each primitive integral batch $\langle i_p j_q \,|k_r l_s\,\rangle$ contains a prefactor $I_{Gaussians}$ that depends on the position of the four Gaussians and their orbital exponents. Since a contracted Gaussian usually has orbital exponents over a rather wide range, it is clear that many of these primitive integral batches will contribute negligibly to the final integral values. In order to reduce the overhead, the parameter `TCut` is introduced. If the common prefactor I_{pqrs} is smaller than `TCut`, the primitive integral batch is skipped. However, I_{pqrs} is *not* a rigorous upper bound to the true value of the primitive integral. Thus, one has to be more conservative with `TCut` than with `Thresh`. In practice it appears that choosing `TCut=0.01*Thresh` provides sufficient accuracy, but the user is encouraged to determine the influence of `TCut` if it is suspected that the accuracy reached in the integrals is not sufficient.

Hint

- If the direct SCF calculation is close to convergence but fails to finally converge, this maybe related to a numerical problem with the Fock matrix update procedure – the accumulated numerical noise from the update procedure prevents sharp convergence. In this case, set `Thresh` and `TCut` lower and/or let the calculation more frequently reset the Fock matrix (`DirectResetFreq`).

Note

- For a `Direct` calculation, there is no way to have `Thresh` larger than `TolE`. If the errors in the Fock matrix are larger than the requested convergence of the energy, the change in energy can never reach `TolE`. The program checks for that.
- The actual disk space used for *all* temporary files may easily be larger than `MaxDisk`. `MaxDisk` only pertains to the two-electron integral files. Other disk requirements are not currently checked by the program and appear to be uncritical.

7.9 DeltaSCF: Converging to Arbitrary Single-Reference Wavefunctions

The regular SCF procedure is supposed to bring the wavefunction to a stationary point, and most times that means a minimum. However, sometimes one might *want* to converge to some kind of "excited state", that is, a higher order saddle-point on the SCF surface.

Using the conventional SCF technology, it is usually not enough to start from an excited-state guess, but one needs to make some extra effort to keep the convergence towards that state. The stack of such methods to keep convergence to a given non-trivial state is called in the literature the `DeltaSCF` approach.

The general idea was first introduced by the group of Peter Gill [299] as the Maximum Overlap Method (MOM), and in ORCA we also feature the more recent PMOM from the group of Hrant Hratchian [184]. It has also been referred to as as "orbital optimized DFT for electronic excited states" [351].

To be very clear, let's show one example in a picture:

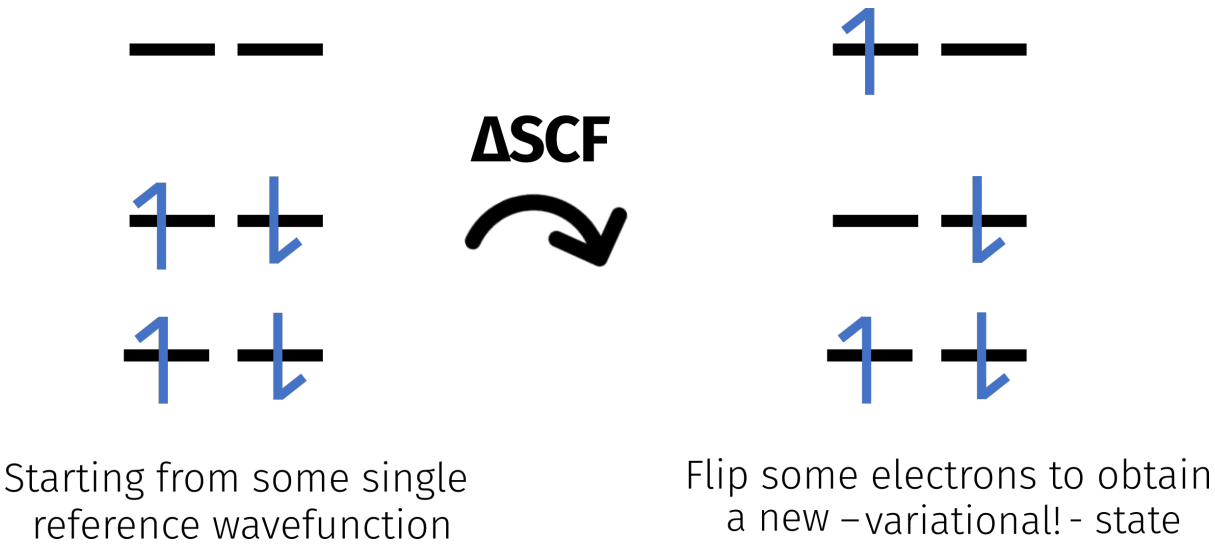


Fig. 7.2: A simple scheme of a HOMO-LUMO state using DeltaSCF

Using ORCA's `DeltaSCF`, we can **choose** to converge the SCF to a HOMO/LUMO excited state. Now that excited state was obtained by fully relaxing the orbitals and can include any contributions such as the VV10 correlation or CPCM for solvation. We will also allow for gradients, geometry optimization, Hessian, EPR, NMR, or anything else that ORCA can do for a "normal" ground state calculation.

Important

The states obtained here are still represented by single-determinant wavefunctions, and in some cases might not even have physically correct meaning. Be careful and conscious of what you are doing! These are relatively reasonable wavefunctions for cases when:

1. the excited state can be simply described by a particle-hole interaction. That is for example **NOT** the case of a benzene molecule or most pi-pi* excited states.
2. the occupied and virtual orbitals are orthogonal or separated in space such that the relevant exchange integral is zero. Eg. some n-pi* states, orthogonal or long-distance charge separated states, etc;
3. for open-shell and doubly-excited states which can be represented by a single-determinant wavefunction.

7.9.1 First Example: HOMO-LUMO Excited State of Formaldehyde

Let's begin by trying to converge and optimize to the first excited state of formaldehyde, starting from its regular planar structure:

```
!PBE DEF2-TZVP OPT FREQ DELTASCF UHF
%SCF ALPHACONF 0,1 END
* XYZ 0 1
C      0.000000    0.000000   -0.602985
O      0.000000    0.000000    0.605394
H      0.000000    0.934673   -1.182175
H      0.000000   -0.934673   -1.182175
*
```

Besides the regular keywords like method, basis set, `OPT` and `FREQ`, one needs to specify `DELTASCF` on the main input, and in this case, `UHF`, since the alpha and beta orbitals will be different (for doubly-excited states `RHF` is sufficient).

It is also necessary to add the `ALPHACONF` or `BETACONF` under the `%SCF` block. That is a minimal representation of the configuration you want to converge to. In this case, `0,1` means a HOMO/LUMO transition, where the HOMO

has occupation zero and LUMO occupation one. For a HOMO-1/LUMO transition, it would be `ALPHACONF 0,1,1`. For a HOMO/LUMO+1 it would be `ALPHACONF 0,0,1` and so on. Just picture how the frontier orbitals should look like. Any orbital below the first zero is assumed to be occupied and any orbital above the last occupied is assumed to be empty.

After the regular startup, the `DeltaSCF`-specific print shows:

```
------------------------------
DELTA-SCF INITIAL CONFIGURATION
------------------------------

Alpha: 1.00 1.00 1.00 0.00 1.00 0.00 0.00 0.00
Beta : 1.00 1.00 1.00 1.00 0.00 0.00 0.00 0.00

Hessian update                              ... L-SR1
Aufbau metric                               ... MOM
Keep initial reference                      ... true
```

Here you can follow the initial configuration and some other important things:

1. `Hessian update` refers to which method will be used for the SOSCF Hessian update. ORCA's default is `L-BFGS`, which forces the electronic Hessian to be positive definite and will always push the system down to a minimum. As we want to go to a saddle point the `L-SR1` is set by default for `DeltaSCF`. More details on Hessian updates and their consequences at this reference from the group of Hannes Jónsson [510].
2. `Aufbau metric` is the way one measures the "overlap" between the actual and reference wavefunctions (more details on [184]). It is `MOM` by default, but can be also set to `%SCF PMOM TRUE END` to use `PMOM`.
3. `Keep initial reference TRUE` means we will always try to keep the initial reference state defined after the guess phase. That is sometimes called `IMOM` in the literature [71]. If `%SCF KEEPINITIALREF FALSE END` is set, it is always the last SCF iteration that is taken as reference.

Important

Here are starting the orbitals from the `PMODEL` guess because it is trivial. In general we recommend always starting the SCF by reading the orbitals of a previously converged ground-state SCF! Please check *Restarting SCF Calculations* for more info on that.

This calculation trivially converges in 12 steps:

```
---------------------------------------D-I-I-S-----------------------------------------
Iteration    Energy (Eh)           Delta-E    RMSDP     MaxDP     DIISErr   Damp  Time(sec)
---------------------------------------------------------------------------------------
               ***  Starting incremental Fock matrix formation  ***
MOM changed the orbital occupation numbers
    1    -114.2533654537761123     0.00e+00  2.06e-03  6.64e-02  7.14e-02  0.700   0.1
MOM changed the orbital occupation numbers
    2    -114.2675698961360382    -1.42e-02  4.71e-04  1.46e-02  3.08e-02  0.700   0.1
                              ***Turning on AO-DIIS***
MOM changed the orbital occupation numbers
    3    -114.2754325617175226    -7.86e-03  1.98e-04  3.98e-03  2.21e-02  0.700   0.1
MOM changed the orbital occupation numbers
    4    -114.2806616906060100    -5.23e-03  4.98e-04  6.40e-03  1.60e-02  0.000   0.1
MOM changed the orbital occupation numbers
                             *** Initializing SOSCF ***
---------------------------------------S-O-S-C-F-----------------------------------------
Iteration    Energy (Eh)           Delta-E    RMSDP     MaxDP     MaxGrad    Time(sec)
---------------------------------------------------------------------------------------
    5    -114.2932665018211225    -1.26e-02  7.19e-05  1.55e-03  2.33e-03     0.1
               *** Restarting incremental Fock matrix formation ***
                          *** Restarting Hessian update ***
    6    -114.2932913947584979    -2.49e-05  5.82e-05  1.33e-03  1.11e-03     0.1
```

(continues on next page)

(continued from previous page)

```
   7    -114.2932671856877818     2.42e-05  2.54e-05  5.00e-04  2.58e-03     0.1
   8    -114.2932914986330530    -2.43e-05  1.76e-05  3.20e-04  1.02e-03     0.1
   9    -114.2932961472105404    -4.65e-06  2.13e-06  6.94e-05  6.42e-05     0.1
  10    -114.2932961779292640    -3.07e-08  2.99e-05  1.10e-03  8.27e-06     0.1
  11    -114.2932921911087050     3.99e-06  3.05e-05  1.12e-03  5.44e-04     0.1
  12    -114.2932961794015654    -3.99e-06  2.45e-08  5.31e-07  2.39e-07     0.1
                *** Gradient check signals convergence ***
```

Note

The statement `MOM changed the orbital occupation numbers` is normal, it is just printing what it is doing. Nothing to worry about.

and one can see from the spin contamination, that this is indeed an open-shell singlet:

```
----------------------
UHF SPIN CONTAMINATION
----------------------

Warning: in a DFT calculation there is little theoretical justification to
         calculate <S**2> as in Hartree-Fock theory. We will do it anyways
         but you should keep in mind that the values have only limited relevance

Expectation value of <S**2>     :     1.006910
Ideal value S*(S+1) for S=0.0   :     0.000000
Deviation                       :     1.006910
```

The gradient is then computed, and the geometry is optimized until convergence. Finally the frequencies show this is actually not a minimum, but a saddle point on the geometry space!

```
-----------------------
VIBRATIONAL FREQUENCIES
-----------------------

Scaling factor for frequencies =  1.000000000  (already applied!)

     0:       0.00 cm**-1
     1:       0.00 cm**-1
     2:       0.00 cm**-1
     3:       0.00 cm**-1
     4:       0.00 cm**-1
     5:       0.00 cm**-1
     6:    -591.31 cm**-1 ***imaginary mode***
     7:     799.56 cm**-1
     8:    1228.37 cm**-1
     9:    1271.29 cm**-1
    10:    2971.77 cm**-1
    11:    3067.96 cm**-1
```

The reason is: the HOMO/LUMO transition on formaldehyde populated the π^* LUMO, thus breaking the double bound and making the carbon atom pyramidal. If one starts from a slightly distorted structure, it then converges to the actual geometry minimum. Starting from a pyramidal structure:

```
!wB97M-D4 DEF2-TZVP DELTASCF UHF OPT FREQ
%SCF ALPHACONF 0,1 END
* XYZ 0 1
C          0.00000        0.00000       -0.60298
O         -0.74131       -0.10909        0.45746
H          0.00000        0.93467       -1.18217
```

(continues on next page)

(continued from previous page)

```
H          0.00000       -0.93467       -1.18217
*
```

now converges to a minimum, as shown by the absence of negative frequencies:

```
-----------------------
VIBRATIONAL FREQUENCIES
-----------------------

Scaling factor for frequencies =  1.000000000  (already applied!)

     0:       0.00 cm**-1
     1:       0.00 cm**-1
     2:       0.00 cm**-1
     3:       0.00 cm**-1
     4:       0.00 cm**-1
     5:       0.00 cm**-1
     6:     713.25 cm**-1
     7:     826.75 cm**-1
     8:    1206.57 cm**-1
     9:    1292.70 cm**-1
    10:    2826.06 cm**-1
    11:    2893.16 cm**-1
```

The final geometry is surprisingly accurate for the gas phase formaldehyde too! We will also show the results here using the range-corrected, meta-GGA hybrid `wB97M-D4` and the double-hybrid DFT `!B2PLYP DEF2-QZVPP AUTOAUX`, just to show that MP2 also works.

Table 7.13: Geometry of gas phase formaldehyde versus the DeltaSCF results, in Angstroem and degrees.

parameter	exp.	PBE	wB97M-D4	B2PLYP
r(C-O)	1.323	1.300	1.310	1.311
r(C-H)	1.098	1.113	1.093	1.090
∠HCH	118.8	115.0	117.0	116.8
∠OOP	34.0	41.1	36.5	37.0

- exp. taken from [299].

Important

We are using the default SCF algorithm here, `AO-DIIS + SOSCF` because this is relatively simple. In general and for more complicated cases we suggest using directly the second order method, to avoid escaping back to the ground state with `!NODIIS`.

Important

Do **NOT** combine DeltaSCF wavefunctions with CCSD, or any such method with single excitations. It requires a speciallized version of CC which we don't have yet.

Important

When running the same calculation above with `wB97M-D4`, there will **not** be a virtual orbital between the alpha HOMO-1 and the HOMO (so no negative HOMO-LUMO gap). There is nothing wrong here, it just optimized the orbitals to the excited state such that this is now a minimum on the SCF surface.

The energy is still higher than the non-DeltaSCF solution and if you plot the orbitals you will see that the alpha HOMO is now a π orbital instead of an n.

7.9.2 Core-ionized States

Another big advantage of the `DeltaSCF` is the possibility to converge to core-excited and/or core-ionized states. We have a simple keyword to kick out electrons from any orbital, even the deep core ones:

```
%SCF IONIZEALPHA 2 END
```

and the electron from orbital number two will be removed. `IONIZEBETA` works for beta orbitals. One can start from an anion UHF -electronic structure obtained by adding one extra electron and remove a core electron like this to obtain core-excited states too. Geometry optimization, EPR, and even TD-DFT calculations are all valid for these states.

As an example, the input below will ionize the 1s electron from a water molecule, which corresponds to MO 0 here:

```
!PBE0 DEF2-TZVP DELTASCF NODIIS UHF
%SCF IONIZEALPHA 0 END
* xyz 0 1
  O      2.127880   -0.361920    0.104770
  H      3.117210   -0.387460    0.070360
  H      1.838520   -0.926280   -0.655730
*
```

Note

If the orbital is not localized over a single atom one might need to localized them first!

and one can see from the results that is exactly where it converged to.

```
----------------
ORBITAL ENERGIES
----------------
                  SPIN UP ORBITALS
  NO   OCC          E(Eh)            E(eV)
   0   0.0000     -20.108086      -547.1688
   1   1.0000      -1.567415       -42.6515
   2   1.0000      -1.057257       -28.7694
   3   1.0000      -0.939315       -25.5601
   4   1.0000      -0.882871       -24.0241
   5   0.0000      -0.304254        -8.2792
   6   0.0000      -0.239285        -6.5113
   7   0.0000       0.044344         1.2066
       (...)
```

Note

ORCA automatically adjusts charge and multiplicity here. **The input should contain those from the reference system!**

Here are some examples of binding energies for 1s electrons. The atom from where it was removed is highlighted in bold:

Table 7.14: Binding energies from 1s electrons found by DeltaSCF using wB97M-V/DEF2-TZVPP, in eV.

1s ionization	exp.	wB97M-V
H2**O**	539.82	541.17
CO2	297.69	299.10
NH3	405.56	406.82
CH3C**N**	405.64	406.92

- exp. taken from [143]

7.9.3 Diabatic Couplings

`DeltaSCF` is also a quite accurate method to obtained diabatic couplings, which can later be used in Markus theory to compute electron transfer rates. These can be computed by calculating the energy difference between electron transfer states and using the Generalized Mulliken-Hush Approach (GMH). For more details please check for example this paper from the group of Blumberger [477].

There is not enough space to go through the details here, but one can get these diabatic coupling from essentially one regular SCF for the ground state + a `DeltaSCF` for the excited state. For symmetric systems, this is trivial:

$$2|H_{ab}| = \Delta E_{12}$$

where states a and b are diabatic states ΔE_{12} is the energy difference between adiabatic states 1 and 2 (which are obtained via SCF solution). Here is an example of the diabatic couplings obtained for a benzene dimer, obtained by starting from the ground state cation and exciting the beta electron with:

```
%SCF BETACONF 0,1 END
```

Table 7.15: Diabatic couplings found by DeltaSCF using wB97M-V/DEF2-TZVPP, in meV.

Distance in Ang	MRCI+Q	wB97M-V	TD-DFT
3.5	435.2	473.1	593
4.0	214.3	236.7	374
4.5	104.0	115.9	267.5
5.0	51.70	56.7	218.5

- MRCI+Q taken from [477]

Note

There is more to come with respect to `DeltaSCF`. We are collaborating further with Prof. Hannes Jónsson's group - stay tuned.

7.9.4 Full keyword list

Here we present a complete list of options to be given under `%SCF` related to `DeltaSCF`:

```
%SCF

  #
  # general options
  #

  DOMOM          TRUE  # do the reordering of the orbitals at all? (default TRUE)
  KEEPINITIALREF TRUE  # always keep initial reference: IMOM? (default TRUE)

  PMOM           FALSE # use the PMOM metric instead of the regular MOM? (default FALSE)
```

(continues on next page)

(continued from previous page)

```
  #
  # occupation number
  #

  ALPHACONF      0,1    # define the occupation of the frontier orbitals.
                        # for RHF and doubly occupied states it could be 0,2.
  BETACONF       0,0,1  # same for the beta orbitals.

  IONIZEALPHA    23     # remove electron from alfa MO number 23?
  IONIZEBETA     12     # remove electron from alfa MO number 12?

  #
  # SOSCF Hessian update
  #

  SOSCFHESSUP    LSR1     # symmetryc rank-1 update (default and recommended).
                 LBFGS    # L-BFGS update.
                 LPOWELL  # regular L-Powell update.
                 LBOFILL  # Bofill update, a combination of SR1 and Powell.
```

7.10 CP-SCF Options

The coupled perturbed self-consistent field (CP-SCF) equations have to be solved in many cases, such as when second derivative properties (e.g. vibrational frequencies, polarizability, NMR shielding, indirect spin-spin coupling, hyperfine coupling, g-tensor) or the MP2 relaxed density (in this case they are referred to as Z-vector equations) are calculated. They are a set of linear equations generally expressed as

$$\mathbf{A}\mathbf{U}^x = \mathbf{B}^x,$$

where $\mathbf{U}^x$ is the vector of solutions for perturbation x, the right-hand side (RHS) matrix $\mathbf{B}^x$ is perturbation-specific and the left-hand side (LHS) matrix $\mathbf{A}$ is perturbation-independent and contains, among other terms, the two-electron repulsion integrals $(ij|ab)$ and $(ia|jb)$. The equations are solved iteratively and the LHS is reassembled at every step, while the RHS does not change. The generation and transformation of the two-electron integrals are therefore the most time-consuming parts of the CP-SCF solution.

The ORCA module which solves these equations accepts several options given below with their default values:

```
%method
  Z_Solver     Pople # (default) Use the Pople algorithm to solve the equations
               DIIS  # Use the DIIS algorithm
               CG    # Use the conjugate gradient algorithm
  Z_Tol        1e-3  # Convergence tolerance for the residual norm.
                     # Default is 1e-5 for VeryTightOpt
                     # and varies from 3e-3 to 3e-6 from LooseSCF to ExtremeSCF
  Z_MaxIter    128   # Maximum number of iterations
  Z_MaxDIIS    12    # Maximum number of DIIS vectors
  Z_Shift      0.3   # Level shift for DIIS
  Z_GridXC     1     # XC angular grid used for the LHS
  Z_IntAccXC   3.467 # XC radial grid accuracy used for the LHS
  Z_GridX      1     # COSX angular grid used for the LHS
  Z_IntAccX    3.067 # COSX radial grid accuracy used for the LHS
  Z_GridX_RHS  2     # COSX grid used for the RHS of MP2 Z-vector eqs (see below)
  Z_COSX_Alg   0     # (default) choose the best COSX algorithm automatically
               1     # better prescreening, more efficient for few densities
               2     # uses more memory, more efficient for many densities
end
```

Since ORCA 6, the same settings are used for all electric response property calculations as well as for CIS/TD-DFT gradients and relaxed densities. For convenience, the keywords in the `%elprop` input block are still available but

they modify the same internal variables as those in %method. For magnetic response properties, the solver and convergence tolerance are set separately in %eprnmr, because the convergence behavior of the magnetic response CP-SCF equations is sometimes different.

```
%elprop
  Solver     # Alias, see: %method Z_Solver
  Tol        # Alias, see: %method Z_Tol
  MaxIter    # Alias, see: %method Z_MaxIter
  MaxDIIS    # Alias, see: %method Z_MaxDIIS
  LevelShift # Alias, see: %method Z_Shift
end

%eprnmr
  Solver     # Solver for magnetic response, see options at: %method Z_Solver
  Tol        # Convergence tolerance for magnetic response
  MaxIter    # Alias, see: %method Z_MaxIter
  MaxDIIS    # Alias, see: %method Z_MaxDIIS
  LevelShift # Alias, see: %method Z_Shift
end
```

The keywords Z_GridX and Z_IntAccX are applicable if the RIJCOSX approximation is chosen for the treatment of two-electron integrals. They determine the angular and radial COSX integration grids, as discussed in section *Changing TD-DFT, CP-SCF and Hessian grids*. Analogously, the keywords Z_Grid and Z_IntAcc determine the integration grid for DFT XC functionals.

Integrals on the RHS are evaluated differently for different perturbations - refer to sections *Using the RI Approximation for Hartree-Fock and Hybrid DFT (RIJCOSX)*, *EPR and NMR properties*, *RIJCOSX-RI-MP2 Gradients*, and *MP2 and RI-MP2 Second Derivatives* and *RI-MP2 and Double-Hybrid DFT Response Properties* for SCF-level gradients, EPR/NMR calculations with GIAOs, MP2 gradients, and MP2 second derivatives, respectively. For MP2 Z-vector equations, the RIJCOSX Fock-response terms in the RHS are evaluated with the COSX grid specified by Z_GridX_RHS. Note that it is used differently to Z_GridX: instead, it selects one of the three grids used in the SCF (see Sections *Using the RI Approximation for Hartree-Fock and Hybrid DFT (RIJCOSX)*, *COSX Grid and Convergence Issues*, and *Details on the numerical integration grids* for details) and it is not recommended to change the default value of 2.

If the RIJONX or RIJK approximation is used in the SCF, the same is also employed in the CP-SCF. Note, however, that the RI-K approximation is not efficient for these terms.

7.11 SCF Stability Analysis

The SCF stability analysis evaluates the electronic Hessian (with respect to orbital rotations) at the point indicated by the SCF solution to determine the lowest eigenvalues of the Hessian. If one or more negative eigenvalues are found, the SCF solution corresponds to a saddle point and not a true local minimum in the space considered in the analysis. A typical case are stretched bonds of diatomics, where the symmetry of the initial guess leads to a restricted solution instead of the often preferred unrestricted one. Several spaces are theoretically possible[779]. However, ORCA limits itself to the analysis RHF/RKS in the space of UHF/UKS or UHF/UKS in the space of UHF/UKS. As such, it is on the available for the SCF parts of DFT and HF.[80] We mention passing, that a stability analysis is also available for the CASSCF type wave function and is described elswhere in more detail (Section *Detecting CASSCF Instabilities*). In the following, HF is used to indicate both HF and KS. Consider the following input (unless indicated otherwise, default values are shown):

```
! BHLYP def2-SVP NORI

%scf
 guess hcore # for illustrative purposes only
 HFTyp UHF # default based on spin multiplicity
 STABPerform true # default false
 STABRestartUHFifUnstable true # restart the UHF-SCF if unstable
 STABNRoots 3                  # number of eigenpairs sought
```

(continues on next page)

(continued from previous page)

```
 STABMaxDim 3                  # Davidson expansion space = MaxDim * NRoots
 STABMaxIter 100               # maximum number of Davidson iterations
 STABNGuess 4096               # size of initial guess matrix: 4096 x 4096
 STABDTol 0.0001               # convergence criterion from iteration to iteration
 STABRTol 0.0001               # convergence criterion max residual norm
 STABlambda +0.5               # mixing parameter
 STABORBWIN -1, -1, -1, -1,  -1, -1, -1, -1 # defines the donor / acceptor spaces
                               # 4 parameters for RHF
                               # 8 paramters for UHF (4 alpha, 4 beta)
                               # orbital window, -1 refers to automatic determination
 STABEWIN -5.0, 5.0            # lower and upper cutoff in Eh for automatic freezing
 #-------------------------------------------------------------------------------
 # alternative specification using a sub-block:
 stab
  NRoots 3
  MaxDim 3 # etc.
 end
end

* xyz 0 1
h 0.0 0.0 0.0
h 0.0 0.0 1.4
*
```

The determination of the electronic Hessian is structurally comparable to the TDHF/CIS/TDDFT procedure. Thus, many options are very similar and the user is encouraged to read the section on TDDFT (Section *Excited States via RPA, CIS, TD-DFT and SF-TDA*) to clarify some of the options given here. Since one is usually only interested in the qualitative determination "stable or not?", three roots should be sufficient to find the lowest eigenvalue. By the same philosophy, `StabMaxDim`, `StabMaxIter`, `StabNGuess` and the convergence criteria were chosen. The parameter `StabLambda` refers to the λ of equation 37 of reference [779], which determines the mixing of the original SCF solution and the new orbitals to yield a new guess. Choosing this value is not trivial, since positive and negative values can lead to different new solutions (at least in principle). The convergence of the ensuing SCF depends on it, as well, since all SCF procedures require a sufficiently good guess to converge in a decent number of iterations (or even at all).

The orbital window and the energy window can be specified. Note that the `StabEWIN` will be overridden by the appropriate `StabORBWIN` values. The automatic determination is also influenced by the `%method FrozenCore` settings. Tests have shown that significant curtailing of the actual orbital window can drastically influence the results to the point of qualitative failure.

Current limitations on the method are:

- Only single-point-like calculations are supported. For geometry optimizations etc., one must use the guess MORead feature *Choice of Initial Guess and Restart of SCF Calculations* to employ the guess obtained here. Likewise, one must extract a geometry and run a separate calculation if one is interested in the SCF stability.
- As for TDDFT, NORI, RIJONX, and RIJCOSX are supported. RI-JK is not supported.
- Other, more advanced features like finite-temperature calculations and relativistic calculations (beside ECPs) are not possible at this time.

Overall, the user is cautioned against using the stability analysis blindly without critically evaluating the result in terms of energy difference and by investigating the orbitals (by the printout or by plotting). Its usefulness cannot be denied, but it is certainly not black-box.

An SCF stability analysis with default settings can be requested via STABILITY, SCFSTABILITY, SCFSTAB or STAB on the simple input line.

7.12 Frozen Core Options

The frozen core (FC) approximation is usually applied in correlated calculation and consists in neglecting correlation effects for electrons in the low-lying core orbitals. The FC approximation and the number of core electrons per element can be adjusted in the %method block. The default number of core electrons per element is listed in Table 7.16.

Table 7.16: Default values for number of frozen core electrons.

H																	He
0																	0
Li	Be											B	C	N	O	F	Ne
0	0											2	2	2	2	2	2
Na	Mg											Al	Si	P	S	Cl	Ar
2	2											10	10	10	10	10	10
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
10	10	10	10	10	10	10	10	10	10	10	10	18	18	18	18	18	18
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
18	18	28	28	28	28	28	28	28	28	28	28	36	36	36	36	36	36
Cs	Ba	Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
36	36	46	46	46	46	46	46	46	46	46	46	68	68	68	68	68	68
Fr	Ra	Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn						
68	68	68	100	100	100	100	100	100	100	100	100						
Lanthanid			La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	
			36	36	36	36	36	36	36	36	36	36	36	36	36	36	
Actinides			Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	
			68	68	68	68	68	68	68	68	68	68	68	68	68	68	

For systems containing heavy elements, core electrons might have higher orbital energies compared to the orbital energies of valence MOs of some lighter elements. In that case, core electrons might be included in the correlation calculation, which ultimately leads to large errors in correlation energy. In order to prevent this, the MO ordering is checked: Do all lower energy MOs in the core region have core electron character, i.e. are they strongly localized on the individual elements? For post-(CAS)SCF calulations, this check is always performed both after the SCF calculation, and after the initial guess (because the SCF may be skipped with `!NoIter`). For other calculations, the check is off by default but may be switched on with the `CheckFrozenCore` keyword in the `%method` block. If core orbitals are found in the valence region, while more delocalized orbitals are found in the core region, the corresponding MO pairs are swapped. This behavior can be disabled using the `CorrectFrozenCore` keyword.

```
%method
  FrozenCore FC_ELECTRONS #Freeze all core electrons
             FC_EWIN      #Freeze selected core electrons via an energy window
```

(continues on next page)

(continued from previous page)

```
                          #e.g. for MP2: %mp2 EWin EMin,EMax
             FC_NONE      #No frozencore approximation
             -n           #Freeze a total of n electrons

  NewNCore Bi 68 end      #Set the number of core electrons for Bi to 68
  CheckFrozenCore   true  #Check whether frozen core orbitals are ordered correctly
                          #Default: true only for post-(CAS)SCF calculations
  CorrectFrozenCore true  #Whether to rotate valence orbitals out of the core region
end
```

Note

- The FrozenCore options are applied to all post Hartree-Fock methods.
- If including all electrons is desired, the `!NoFrozenCore` keyword can be simply inserted. For MP2: Frozen virtual orbitals are not allowed in gradient runs or geometry optimization!
- If ECPs are used, the number for NewNCore has to include the electrons represented by the ECPs as well. E.g. if an element is supposed to have 60 electrons in the ECP and additional 8 electrons should be frozen in the correlation calculation, NewNCore should be 68.
- In ORCA we use rather conservative frozencore settings, i. e. a large number of electrons are included in the correlation treatment. Therefore, we recommend to use properly optimized correlating basis functions in all cases, such as the cc-pwCVXZ basis sets.
- For DLPNO calculations the virtual space for core-core and core-valence correlation is adjusted by default, which is described in detail in section *Including (semi)core orbitals in the correlation treatment*.
- In general, `NewNCore` only has an effect in calculations with `FC_ELECTRONS`. In calculations using the DLPNO approximation (except DLPNO-NEVPT2), `NewNCore` has also an effect in the other cases, as is described in section *Including (semi)core orbitals in the correlation treatment*.
- Double-hybrid density functional (section *DFT Calculations with Second Order Perturbative Correction (Double-Hybrid Functionals)*) calculations by default use the FrozenCore option for the perturbative part, as is the case for MP2.

7.13 The Second Order Many Body Pertubation Theory Module (MP2)

Throughout this section, indices $i, j, k, \ldots$ refer to occupied orbitals in the reference determinant, $a, b, c, \ldots$ to virtual orbitals and $p, q, r, \ldots$ to general orbitals from either set while $\mu, \nu, \kappa, \tau, \ldots$ refer to basis functions.

7.13.1 Standard MP2

The standard (or full accuracy) MP2 module has two different branches. One branch is used for energy calculations, the other for gradient calculations.

For standard MP2 energies, the program performs two half-transformations and the half-transformed integrals are stored on disk in compressed form. This appears to be the most efficient approach that can also be used for medium sized molecules.The module should parallelize acceptably well as long as I/O is not limiting.

For standard MP2 gradients, the program performs four quarter transformations that are ordered by occupied orbitals. Here, the program massively benefits from large core memory (`%maxcore`) since this minimizes the number of batches that are to be done. I/O demands are minimal in this approach.

In "memory mode" (Q1Opt>0) basically the program treats batches of occupied orbitals at the same time. Thus, there must be at least enough memory to treat a single occupied MO at each pass. Otherwise the MP2 module will

fail. Thus, potentially, MP2 calculations on large molecules take significant memory and may be most efficiently done through the RI approximation.

Alternatively, in the "disk based mode" (Q1Opt=-1) the program performs a half transformation of the exchange integrals and stores the transformed integrals on disk. A bin-sort then leads to the AO operator $K^{ij}(\mu,\nu) = (i\mu|j\nu)$ in (11|22) integral notation. These integrals are then used to make the final K^{ij}(a,b) (a,b=virtual MOs) and the EMP2 pair energy contributions. In many cases, and in particular for larger molecules, this algorithm is much more efficient than the memory based algorithm. It depends, however, much more heavily on the I/O system of the computer that you use. It is important, that the program uses the flags `CFLOAT, UCFLOAT, CDOUBLE` or `UCDOUBLE` in order to store the unsorted and sorted AO exchange integrals. Which flag is used will influence the performance of the program and to some extent the accuracy of the result (float based single precision results are usually very slightly less accurate; μEh-range deviations from the double precision result[1]). Finally, gradients are presently only available for the memory based algorithm since in this case a much larger set of integrals is required.

The `! MP2` command does the following: (a) it changes the `Method` to `HFGTO` and (b) it sets the flag `DoMP2` to `true`. The program will then first carry out a Hartree-Fock SCF calculation and then estimate the correlation energy by MP2 theory. RHF, UHF and high-spin ROHF reference wavefunctions are permissible and the type of MP2 calculation to be carried out (for high-spin ROHF the gradients are not available) is automatically chosen based on the value of `HFTyp`. If the SCF is carried out conventionally, the MP2 calculation will also be done in a conventional scheme unless the user forces the calculation to be direct. For `SCFMode` =`Direct` the MP2 energy evaluation will be fully in the integral direct mode.

The following variables can be adjusted in the block for conventional MP2 calculations:

```
%mp2
  EMin       -1.5       # orbital energy cutoff that defines the
                        # frozen core in Eh
  EMax        1.0e3     # orbital energy cutoff that defines the
                        # neglected virtual orbitals in Eh
  EWin        EMin,EMax # the same, but accessed as array
                        # (respects settings in %method block!)
  MaxCore     256       # maximum amount of memory (in MB) to be
                        # used for integral buffering
  ForceDirect false     # Force the calculation to be integral
                        # direct
  RI          false     # use the RI approximation
  F12         false     # apply F12 correction
  Q1Opt                 # For non-RI calculations a flag how to perform
                        # the first quarter transformation
                        #  1 - use double precision buffers
                        #      (default for gradient runs)
                        #  2 - use single precision buffers. This reduces
                        #      the memory usage in the bottleneck step by
                        #      a factor of two. If several passes are re-
                        #      quired, the number of passes is reduced by
                        #      a factor of two.
                        # -1 - Use a disk based algorithm. This respects
                        #      the flags UCFLOAT,CFLOAT,UCDOUBLE and
                        #      CDOUBLE. (but BE CAREFUL with FLOAT)
                        #      (default for energy runs)
  PrintLevel  2         # How much output to produce. PrintLevel 3 produces
                        # also pair correlation energies and other info.
  DoSCS       false     # use spin-component scaling
  Ps          1.2       # scaling factor for ab pairs
  Pt          0.333     # scaling factor for aa and bb pairs
  Density     none      # no density construction
              unrelaxed # only "unrelaxed densities"
              relaxed   # full relaxed densities
  NatOrbs     false     # calculate natural orbitals
```

[1] However, sometimes, and in particular when transition metals and core orbitals are involved we have met unpleasantly large errors. So – be careful and double check when using floats!

7.13.2 RI-MP2

The RI-MP2 module is of a straightforward nature. The program first transforms the three-index integrals $(ia|\tilde{P})$, where "i" is a occupied, "a" is a virtual MO and "$\tilde{P}$" is an auxiliary basis function that is orthogonalized against the Coulomb metric. These integrals are stored on disk, which is not critical, even if the basis has several thousand functions. The integral transformation is parallelized and has no specifically large core memory requirements.

In the next step, the integrals are read ordered with respect to the occupied labels and the exchange operators $K^{ij}(a,b) = (ia|jb) = \sum_{\tilde{P}}^{\text{NAux}} (ia|\tilde{P})(\tilde{P}|jb)$ are formed in the rate limiting O(N^5) step. This step is done with high efficiency by a large matrix multiplication and parallelizes well. From the exchange operators, the MP2 amplitudes and the MP2 energy is formed. The program mildly benefits from large core memory (%maxcore) as this minimizes the number of batches and hence reads through the integral list.

The RI-MP2 gradient is also available. Here, all necessary intermediates are made on the fly.

In the RI approximation, one introduces an auxiliary fitting basis $\eta_P(\mathbf{r})$ and then approximates the two-electron integrals in the Coulomb metric as:

$$(pq|rs) \approx \sum_{PQ} (pq|P)\, V_{PQ}^{-1}\, (Q|rs) \tag{7.50}$$

where $V_{PQ} = (P|Q)$ is a two-index electron-electron repulsion integral. As first discussed by Weigend and Häser, the closed-shell case RI-MP2 gradient takes the form:

$$E^x_{\text{RI-MP2}} = 2\sum_{\mu\nu P} (\mu\nu|P)^{(x)} \sum_i c_{\mu i}\Gamma^P_{i\nu} + \sum_{RS} V^x_{RS} \left(\mathbf{V}^{-1/2}\gamma\mathbf{V}^{-1/2}\right)_{RS} + \langle \mathbf{D}\mathbf{F}^x\rangle \tag{7.51}$$

The **F**-matrix derivative terms are precisely handled as in the non-RI case and need not be discussed any further. Γ^P_{ia} is a three-index two-particle "density":

$$\Gamma^P_{ia} = \sum_{jbQ} (1+\delta_{ij})\, \tilde{t}^{ij}_{ab} V^{-1/2}_{PQ}\, (Q|jb) \tag{7.52}$$

Which is partially transformed to the AO basis by:

$$\Gamma^P_{i\nu} = \sum_a c_{\nu a}\Gamma^P_{ia} \tag{7.53}$$

The two-index analogue is given by:

$$\gamma_{PQ} = \sum_{iaR} \Gamma^Q_{ia}\, (ia|R)\, V^{-1/2}_{RP} \tag{7.54}$$

The RI contribution to the Lagrangian is particularly convenient to calculate:

$$L^{RI}_{ai} = \sum_\mu c_{\mu a} \left[2\sum_{PQ\nu} \Gamma^P_{i\nu}\, (\mu\nu|Q)\, V^{-1/2}_{PQ} \right] \tag{7.55}$$

In a similar way, the remaining contributions to the energy weighted density matrix can be obtained efficiently. Note, however, that the response operator and solution of the CP-SCF equations still proceed via traditional four-index integrals since the SCF operator was built in this way. Thus, while the derivatives of the three-index integrals are readily and efficiently calculated, one still has the separable contribution to the gradient, which requires the derivatives of the four-index integrals.

The RI-MP2 energy and gradient calculations can be drastically accelerated by employing the RIJCOSX or the RIJDX approximation.

7.13.3 "Double-Hybrid" Density Functional Theory

A slightly more general form is met in the double-hybrid DFT gradient. The theory is briefly described below.

The energy expression for perturbatively and gradient corrected hybrid functionals as proposed by Grimme is:

$$E = V_{NN} + \left\langle \mathbf{Ph}^{+} \right\rangle + \frac{1}{2}\int\int \frac{\rho\left(\mathbf{r}_1\right)\rho\left(\mathbf{r}_2\right)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 - \frac{1}{2} a_x \sum_{\mu\nu\kappa\tau\sigma} P^{\sigma}_{\mu\kappa} P^{\sigma}_{\nu\tau} \left(\mu\nu|\kappa\tau\right) + c_{\mathrm{DF}} E_{\mathrm{XC}}\left[\rho_\alpha, \rho_\beta\right] + c_{\mathrm{PT}} E_{\mathrm{PT}}$$

$$E = E_{\mathrm{SCF}} + c_{\mathrm{PT}} E_{\mathrm{PT}} \tag{7.56}$$

Here V_{NN} is the nuclear repulsion energy and $h_{\mu\nu}$ is a matrix element of the usual one-electron operator which contains the kinetic energy and electron-nuclear attraction terms ($\langle \mathbf{ab} \rangle$ denotes the trace of the matrix product $\mathbf{ab}$). As usual, the molecular spin-orbitals are expanded in atom centered basis functions ($\sigma = \alpha, \beta$):

$$\psi^{\sigma}_{p}\left(\mathbf{r}\right) = \sum_{\mu} c^{\sigma}_{\mu p} \varphi_{\mu}\left(\mathbf{r}\right) \tag{7.57}$$

with MO coefficients $c^{\sigma}_{\mu p}$. The total density is given by (real orbitals are assumed throughout):

$$\rho\left(\mathbf{r}\right) = \sum_{i\sigma} \left|\psi^{\sigma}_{i}\left(\mathbf{r}\right)\right|^2 = \sum_{\mu\nu\sigma} P^{\sigma}_{\mu\nu} \varphi_{\mu}\left(\mathbf{r}\right) \varphi_{\nu}\left(\mathbf{r}\right) = \rho^{\alpha}\left(\mathbf{r}\right) + \rho^{\beta}\left(\mathbf{r}\right) \tag{7.58}$$

Where $\mathbf{P} = \mathbf{P}^{\alpha} + \mathbf{P}^{\beta}$ and $P^{\sigma}_{\mu\nu} = \sum_{i_\sigma} c^{\sigma}_{\mu i} c^{\sigma}_{\nu i}$.

The second term of (7.56) represents the Coulombic self-repulsion. The third term represents the contribution of the Hartree-Fock exchange with the two-electron integrals being defined as:

$$\left(\mu\nu|\kappa\tau\right) = \int\int \phi_{\mu}\left(\mathbf{r}_1\right) \phi_{\nu}\left(\mathbf{r}_1\right) r_{12}^{-1} \phi_{\kappa}\left(\mathbf{r}_2\right) \phi_{\tau}\left(\mathbf{r}_2\right) d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.59}$$

The mixing parameter a_x controls the fraction of Hartree-Fock exchange and is of a semi-empirical nature. The exchange correlation contribution may be written as:

$$E_{\mathrm{XC}}\left[\rho_\alpha, \rho_\beta\right] = \left(1 - a_x\right) E^{\mathrm{GGA}}_{\mathrm{X}}\left[\rho_\alpha, \rho_\beta\right] + b E^{\mathrm{GGA}}_{\mathrm{C}}\left[\rho_\alpha, \rho_\beta\right] \tag{7.60}$$

Here, $E^{\mathrm{GGA}}_{\mathrm{X}}\left[\rho_\alpha, \rho_\beta\right]$ is the exchange part of the XC- functional in question and $E^{\mathrm{GGA}}_{\mathrm{C}}\left[\rho_\alpha, \rho_\beta\right]$ is the correlation part. The parameter b controls the mixing of DFT correlation into the total energy and the parameter c_{DF} is a global scaling factor that allows one to proceed from Hartree-Fock theory ($a_{\mathrm{X}} = 1$, $c_{\mathrm{DF}} = 0$, $c_{\mathrm{PT}} = 0$) to MP2 theory ($a_{\mathrm{X}} = 1$, $c_{\mathrm{DF}} = 0$, $c_{\mathrm{PT}} = 1$) to pure DFT ($a_{\mathrm{X}} = 1$, $c_{\mathrm{DF}} = 0$, $c_{\mathrm{PT}} = 1$) to hybrid DFT ($0 < a_{\mathrm{X}} < 1$, $c_{\mathrm{DF}} = 1$, $c_{\mathrm{PT}} = 0$) and finally to the general perturbatively corrected methods discussed in this work ($0 < a_{\mathrm{X}} < 1$, $c_{\mathrm{DF}} = 1$, $0 < c_{\mathrm{PT}} < 1$). As discussed in detail by Grimme, the B2- PLYP functional uses the Lee-Yang-Parr (LYP) functional as correlation part, the Becke 1988 (B88) functional as GGA exchange part and the optimum choice of the semi-empirical parameters was determined to be $a_{\mathrm{X}} = 0.53$, $c_{\mathrm{PT}} = 0.27$, $c_{\mathrm{DF}} = 1$, $b = 1 - c_{\mathrm{PT}}$. For convenience, we will suppress the explicit reference to the parameters a_{X} and b in the XC part and rewrite the gradient corrected XC energy as:

$$E_{\mathrm{XC}}\left[\rho^{\alpha}, \rho^{\beta}\right] = \int f\left(\rho^{\alpha}, \rho^{\beta}, \gamma^{\alpha\alpha}, \gamma^{\beta\beta}, \gamma^{\alpha\beta}\right) d\mathbf{r} \tag{7.61}$$

with the gradient invariants $\gamma^{\sigma\sigma'} = \vec{\nabla}\rho^{\sigma}\vec{\nabla}\rho^{\sigma'}$. The final term in eq (48) represents the scaled second order perturbation energy:

$$E^{\mathrm{PT2}} = \frac{1}{2}\sum_{i_\alpha < j_\alpha} \left\langle \mathbf{t}^{i_\alpha j_\alpha} \bar{\mathbf{K}}^{i_\alpha j_\alpha +} \right\rangle + \frac{1}{2}\sum_{i_\beta < j_\beta} \left\langle \mathbf{t}^{i_\beta j_\beta} \bar{\mathbf{K}}^{i_\beta j_\beta +} \right\rangle + \sum_{i_\alpha, j_\beta} \left\langle \mathbf{t}^{i_\alpha j_\beta} \bar{\mathbf{K}}^{i_\alpha j_\beta +} \right\rangle \tag{7.62}$$

The PT2 amplitudes have been collected in matrices $\mathbf{t}^{i_\sigma j_{\sigma'}}$ with elements:

$$t^{i_\sigma j_{\sigma'}}_{a_\sigma b_{\sigma'}} = \bar{K}^{i_\sigma j_{\sigma'}}_{a_\sigma b_{\sigma'}} \left(\varepsilon^{\sigma}_{i} + \varepsilon^{\sigma'}_{j} - \varepsilon^{\sigma}_{a} - \varepsilon^{\sigma'}_{b}\right)^{-1} \tag{7.63}$$

Where the orbitals were assumed to be canonical with orbital energies ε_p^σ. The exchange operator matrices are $K_{a_\sigma b_{\sigma'}}^{i_\sigma j_{\sigma'}} = (i_\sigma a_\sigma | j_{\sigma'} b_{\sigma'})$ and the anti-symmetrized exchange integrals are defined as $\bar{K}_{a_\sigma b_{\sigma'}}^{i_\sigma j_{\sigma'}} = (i_\sigma a_\sigma | j_{\sigma'} b_{\sigma'}) - \delta_{\sigma\sigma'} (i_\sigma b_\sigma | _\sigma a_\sigma)$.

The orbitals satisfy the SCF equations with the matrix element of the SCF operator given by:

$$F_{\mu\nu}^\sigma = h_{\mu\nu} + \sum_{\kappa\tau} P_{\kappa\tau} (\mu\nu|\kappa\tau) - a_{\mathrm{X}} P_{\kappa\tau}^\sigma (\mu\kappa|\nu\tau) + c_{\mathrm{DF}} (\mu|V_{\mathrm{XC}}^\sigma|\nu) \tag{7.64}$$

The matrix elements of the XC–potential for a gradient corrected functional are: [839]

$$(\mu|V_{\mathrm{XC}}^\alpha|\nu) = \int \left\{ \frac{\delta f}{\delta \rho_\alpha (\mathbf{r})} (\varphi_\mu \varphi_\nu) + 2 \frac{\delta f}{\delta \gamma_{\alpha\alpha}} \vec{\nabla} \rho_\alpha \vec{\nabla} (\varphi_\mu \varphi_\nu) + \frac{\delta f}{\delta \gamma_{\alpha\beta}} \vec{\nabla} \rho_\beta \vec{\nabla} (\varphi_\mu \varphi_\nu) \right\} d\mathbf{r} \tag{7.65}$$

The energy in equation (7.56) depends on the MO-coefficients, the PT2-amplitudes and through V_{NN}, V_{eN} (in h) and the basis functions also explicitly on the molecular geometry. Unfortunately, the energy is only stationary with respect to the PT2 amplitudes since they can be considered as having been optimized through the minimization of the Hylleraas functional:

$$E_{\mathrm{PT2}} = \min_{\mathbf{t}} \left\{ \frac{1}{2} \sum_{i_\alpha < j_\alpha} \left\langle \mathbf{t}^{i_\alpha j_\alpha} \bar{\mathbf{K}}^{i_\alpha j_\alpha +} \right\rangle + \frac{1}{2} \sum_{i_\beta < j_\beta} \left\langle \mathbf{t}^{i_\beta j_\beta} \bar{\mathbf{K}}^{i_\beta j_\beta +} \right\rangle + \sum_{i_\alpha j_\beta} \left\langle \mathbf{t}^{i_\alpha j_\beta} \bar{\mathbf{K}}^{i_\alpha j_\beta +} \right\rangle + \left\langle \mathbf{D}'^\alpha \mathbf{F}^{\alpha +} \right\rangle + \left\langle \mathbf{D}'^\beta \mathbf{F}^{\beta +} \right\rangle \right\} \tag{7.66}$$

The unrelaxed PT2 difference density is defined as:

$$D_{ij}^{'\alpha} = -\frac{1}{2} \sum_{k_\alpha} \left\langle \mathbf{t}^{i_\alpha k_\alpha} \mathbf{t}^{k_\alpha j_\alpha} \right\rangle - \sum_{k_\beta} \left\langle \mathbf{t}^{i_\alpha k_\beta} \mathbf{t}^{k_\beta j_\alpha} \right\rangle \tag{7.67}$$

$$D_{ab}^{'\alpha} = \sum_{i_\alpha < j_\alpha} \mathbf{t}^{i_\alpha j_\alpha} \mathbf{t}^{i_\alpha j_\alpha +} + \sum_{i_\beta j_\alpha} \mathbf{t}^{i_\beta j_\alpha +} \mathbf{t}^{i_\beta j_\alpha} \tag{7.68}$$

With analogous expressions for the spin-down unrelaxed difference densities. Minimization of this functional with respect to the amplitudes yields the second order perturbation energy. The derivative of the SCF part of equation (7.56) with respect to a parameter "x" is straightforward and well known. It yields:

$$\begin{aligned} E_{\mathrm{SCF}}^x = {} & V_{NN}^x + \langle \mathbf{P}\mathbf{h}^x \rangle + \left\langle \mathbf{W}^{SCF} \mathbf{S}^{(x)} \right\rangle + \textstyle\sum_{\mu\nu\kappa\tau} \Gamma_{\mu\nu\kappa\tau} (\mu\nu|\kappa\tau)^{(x)} \\ & + \underbrace{\sum_{\sigma}}_{(\sigma' \neq \sigma)} \int \left\{ \tfrac{\delta f}{\delta \rho_\sigma (\mathbf{r})} \rho_\sigma^{(x)} + 2 \tfrac{\delta f}{\delta \gamma_{\sigma\sigma}} \vec{\nabla} \rho_\sigma \vec{\nabla} \rho_\sigma^{(x)} + \tfrac{\delta f}{\delta \gamma_{\sigma\sigma'}} \vec{\nabla} \rho_{\sigma'} \vec{\nabla} \rho_\sigma^{(x)} \right\} d\mathbf{r} \end{aligned} \tag{7.69}$$

Superscript "x" refers to the derivative with respect to some perturbation "x" while a superscript in parentheses indicates that only the derivative of the basis functions with respect to "x" is to be taken. For example:

$$\begin{aligned} \rho_\sigma^{(x)} &= \textstyle\sum_{\mu\nu} P_{\mu\nu}^\sigma \left\{ \frac{\partial \varphi_\mu}{\partial x} \varphi_\nu + \varphi_\mu \frac{\partial \varphi_\nu}{\partial x} \right\} \\ h_{\mu\nu}^x &= \left(\tfrac{\partial \varphi_\mu}{\partial x} | \hat{h} | \varphi_\nu \right) + \left(\varphi_\mu | \hat{h} | \tfrac{\partial \varphi_\nu}{\partial x} \right) + \left(\varphi_\mu | \tfrac{\partial \hat{h}}{\partial x} | \varphi_\nu \right) \end{aligned} \tag{7.70}$$

In equation (7.69), $\mathbf{S}$ is the overlap matrix and $\mathbf{W}$ $^{\mathrm{SCF}}$ the energy weighted density:

$$W_{\mu\nu}^{\mathrm{SCF}} = W_{\mu\nu}^{\alpha;\mathrm{SCF}} + W_{\mu\nu}^{\beta;\mathrm{SCF}} = -\sum_{i\sigma} c_{\mu i}^\sigma c_{\nu i}^\sigma \varepsilon_i^\sigma \tag{7.71}$$

At this point, the effective two-particle density matrix is fully separable and reads:

$$\Gamma_{\mu\nu\kappa\tau} = \frac{1}{2} P_{\mu\nu} P_{\kappa\tau} - \frac{1}{2} a_x P_{\mu\kappa}^\alpha P_{\nu\tau}^\alpha - \frac{1}{2} a_x P_{\mu\kappa}^\beta P_{\nu\tau}^\beta \tag{7.72}$$

The derivative of the PT2 part is considerably more complex, since E_{PT2} is not stationary with respect to changes in the molecular orbitals. This necessitates the solution of the coupled-perturbed SCF (CP-SCF) equations. We follow the standard practice and expand the perturbed orbitals in terms of the unperturbed ones as:

$$\psi_p^{\sigma;x} (\mathbf{r}) = \sum_q U_{qp}^{\sigma;x} \psi_q^\sigma (\mathbf{r}) \tag{7.73}$$

The occupied-occupied and virtual-virtual blocks of $\mathbf{U}$ are fixed, as usual, through the derivative of the orthonormality constraints:

$$U_{ij}^{\sigma;x} = -\frac{1}{2} S_{ij}^{\sigma(x)} \tag{7.74}$$

$$U_{ab}^{\sigma;x} = -\frac{1}{2} S_{ab}^{\sigma(x)} \tag{7.75}$$

$$U_{ia}^{\sigma;x} = -S_{ia}^{\sigma(x)} - U_{ai}^{\sigma;x} \tag{7.76}$$

Where $S_{pq}^{\sigma(x)} = \sum_{\mu\nu} c_{\mu p}^{\sigma} c_{\nu q}^{\sigma} S_{\mu\nu}^{(x)}$. The remaining virtual-occupied block of $\mathbf{U}^{\,x}$ must be determined through the solution of the CP-SCF equations. However, as shown by Handy and Schaefer, this step is unnecessary and only a single set of CP-SCF equations (Z-vector equations) needs to be solved. To this end, one defines the Lagrangian:

$$\begin{aligned} L_{ai}^{\alpha} = R^{\sigma}\left(\mathbf{D}'\right)_{ai} + 2\sum_{j_\alpha b_\alpha c_\alpha} \left(a_\alpha c_\alpha | j_\alpha b_\alpha\right) t_{c_\alpha b_\alpha}^{i_\alpha j_\alpha} - 2\sum_{j_\alpha k_\alpha b_\alpha} \left(k_\alpha i_\alpha | j_\alpha b_\alpha\right) t_{a_\alpha b_\alpha}^{k_\alpha j_\alpha} \\ + 2\sum_{j_\beta b_\beta c_\alpha} \left(a_\alpha c_\alpha | j_\beta b_\beta\right) t_{b_\beta c_\alpha}^{j_\beta i_\alpha} - 2\sum_{j_\beta k_\alpha b_\beta} \left(k_\alpha i_\alpha | j_\beta b_\beta\right) t_{b_\beta a_\alpha}^{j_\beta k_\alpha} \end{aligned} \tag{7.77}$$

An analogous equation holds for L_{ai}^{β}. The matrix elements of the response operator $R^{\alpha}\left(\mathbf{D}'\right)$ are best evaluated in the AO basis and then transformed into the MO basis. The AO basis matrix elements are given by:

$$\begin{aligned} R^{\alpha}\left(\mathbf{D}'\right)_{\mu\nu} = & \sum_{\kappa\tau} 2D'_{\kappa\tau}\left(\mu\nu|\kappa\tau\right) - D'^{\alpha}_{\kappa\tau}\left[\left(\mu\kappa|\nu\tau\right) + \left(\nu k|\mu\tau\right)\right] \\ & + \sum_{\zeta} \int \left[\frac{\delta^2 f}{\delta\rho_\alpha \delta\zeta} \zeta\left(\mathbf{D}'\right)\left(\phi_\mu\phi_\nu\right) + \left(2\frac{\delta^2 f}{\delta\gamma_{\alpha\alpha}\delta\zeta}\vec{\nabla}\rho_{\mathbf{P}}^{\alpha} + \frac{\delta^2 f}{\delta\gamma_{\alpha\beta}\delta\zeta}\vec{\nabla}\rho_{\mathbf{P}}^{\beta} \right) \zeta\left(\mathbf{D}'\right)\vec{\nabla}\left(\phi_\mu\phi_\nu\right) \right. \\ & \left. + \left(2\frac{\delta f}{\delta\gamma_{\alpha\alpha}}\vec{\nabla}\rho_{\mathbf{D}'}^{\alpha} + \frac{\delta f}{\delta\gamma_{\alpha\beta}}\vec{\nabla}\rho_{\mathbf{D}'}^{\beta} \right)\vec{\nabla}\left(\phi_\mu\phi_\nu\right) \right] d\mathbf{r} \end{aligned} \tag{7.78}$$

where

$$\zeta\left(\mathbf{D}'\right) = \rho_{\mathbf{D}'}^{\alpha}, \rho_{\mathbf{D}'}^{\beta}, \gamma_{\alpha\alpha}\left(\mathbf{D}'\right), \gamma_{\beta\beta}\left(\mathbf{D}'\right), \gamma_{\alpha\beta}\left(\mathbf{D}'\right) \tag{7.79}$$

The ζ-gradient-parameters are evaluated as a mixture of PT2 difference densities and SCF densities. For example:

$$\gamma_{\alpha\alpha}(\mathrm{D}') = 2\vec{\nabla}\rho_{\mathrm{D}'}^{\alpha}\vec{\nabla}\rho_{\mathrm{P}'}^{\alpha} \tag{7.80}$$

With

$$\rho_{\mathbf{D}'}^{\alpha}\left(\mathbf{r}\right) = \sum_{\mu\nu} D'^{\alpha}_{\mu\nu}\phi_\mu\left(\mathbf{r}\right)\phi_\nu\left(\mathbf{r}\right) \tag{7.81}$$

$$\rho_{\mathbf{P}}^{\alpha}\left(\mathbf{r}\right) = \sum_{\mu\nu} P_{\mu\nu}^{\alpha}\phi_\mu\left(\mathbf{r}\right)\phi_\nu\left(\mathbf{r}\right) \tag{7.82}$$

Having defined the Lagrangian, the following CP-SCF equations need to be solved for the elements of the "Z-vector":

$$\left(\varepsilon_a^{\sigma} - \varepsilon_i^{\sigma}\right) Z_{ai}^{\sigma} + R^{\sigma}\left(\mathbf{Z}\right)_{ai} = -L_{ai}^{\sigma} \tag{7.83}$$

The solution defines the occupied-virtual block of the relaxed difference density, which is given by:

$$\mathbf{D}^{\sigma} = \mathbf{D}'^{\sigma} + \mathbf{Z}^{\sigma} \tag{7.84}$$

For convenience, $\mathbf{D}^{\sigma}$ is symmetrized since it will only be contracted with symmetric matrices afterwards. After having solved the Z-vector equations, all parts of the energy weighted difference density matrix can be readily calculated:

$$W_{ij}^{\alpha;\mathrm{PT2}} = -\frac{1}{2} D_{ij}^{\alpha}\left(\varepsilon_i^{\alpha} + \varepsilon_j^{\alpha}\right) - \frac{1}{2} R\left(\mathbf{D}\right)_{ij} - \sum_{k_\alpha a_\alpha b_\alpha} \left(i_\alpha a_\alpha | k_\alpha b_\alpha\right) t_{a_\alpha b_\alpha}^{j_\alpha k_\alpha} - \sum_{k_\beta a_\alpha b_\beta} \left(i_\alpha a_\alpha | k_\beta b_\beta\right) t_{b_\beta a_\alpha}^{k_\beta j_\alpha} \tag{7.85}$$

$$W_{ab}^{\alpha;\mathrm{PT2}} = -\frac{1}{2} D_{ab}^{\alpha}\left(\varepsilon_a^{\alpha} + \varepsilon_b^{\alpha}\right) - \sum_{i_\alpha j_\alpha c_\alpha} \left(i_\alpha a_\alpha | j_\alpha c_\alpha\right) t_{b_\alpha c_\alpha}^{i_\alpha j_\alpha} - \sum_{i_\alpha j_\beta c_\beta} \left(i_\alpha a_\alpha | j_\beta c_\beta\right) t_{c_\beta b_\alpha}^{j_\beta i_\alpha} \tag{7.86}$$

$$W_{ai}^{\alpha;\text{PT2}} = -2 \sum_{j_\alpha k_\alpha b_\alpha} (k_\alpha i_\alpha | j_\alpha b_\alpha)\, t_{a_\alpha b_\alpha}^{k_\alpha j_\alpha} - 2 \sum_{j_\beta k_\alpha b_\beta} (k_\alpha i_\alpha | j_\beta b_\beta)\, t_{b_\beta a_\alpha}^{j_\beta k_\alpha} \tag{7.87}$$

$$W_{ia}^{\alpha;\text{PT2}} = -\varepsilon_i^\alpha Z_{ai}^\alpha \tag{7.88}$$

Once more, analogous equations hold for the spin-down case. With the relaxed difference density and energy weighted density matrices in hand, one can finally proceed to evaluate the gradient of the PT2 part as ($\mathbf{W}^{\text{PT2}} = \mathbf{W}^{\alpha;\text{PT2}} + \mathbf{W}^{\beta;\text{PT2}}$):

$$\begin{aligned} E_{\text{PT2}}^x = &\langle \mathbf{D}\mathbf{h}^x \rangle + \left\langle \mathbf{W}^{\text{PT2}} \mathbf{S}^{(x)} \right\rangle + \textstyle\sum_{\mu\nu\kappa\tau} \Gamma_{\mu\nu\kappa\tau}^{\text{PT2}} \left(\mu\nu|\kappa\tau\right)^{(x)} \\ &+ \textstyle\sum_{\substack{\sigma \\ (\sigma \neq \sigma')}} \int \left\{ \frac{\delta f}{\delta \rho_\sigma(\mathbf{r})} \rho_\sigma^{(x)} + 2 \frac{\delta f}{\delta \gamma_{\sigma\sigma}} \overset{r}{\nabla} \rho_\sigma \overset{r}{\nabla} \rho_\sigma^{(x)} + \frac{\delta f}{\delta \gamma_{\sigma\sigma}} \overset{r}{\nabla} \rho_{\sigma'} \overset{r}{\nabla} \rho_\sigma^{(x)} \right\} d\mathbf{r} \end{aligned} \tag{7.89}$$

The final derivative of eq. (7.56) is of course the sum $E_{SCF}^x + c_{PT} E_{\text{PT2}}^x$. Both derivatives should be evaluated simultaneously in the interest of computational efficiency.

Note that the exchange-correlation contributions to the gradient take a somewhat more involved form than might have been anticipated. In fact, from looking at the SCF XC-gradient (eq. (7.69)) it could have been speculated that the PT2 part of the gradient is of the same form but with $\rho_{\mathbf{P}}^{\sigma(x)}$ being replaced by $\hat{H}$, the relaxed PT2 difference density. This is, however, not the case. The underlying reason for the added complexity apparent in equation (7.89) is that the XC contributions to the PT2 gradient arise from the contraction of the relaxed PT2 difference density with the derivative of the SCF operator. Since the SCF operator already contains the first derivative of the XC potential and the PT2 energy is not stationary with respect to changes in the SCF density, a response type term arises which requires the evaluation of the second functional derivative of the XC-functional. Finally, as is well known from MP2 gradient theory, the effective two- particle density matrix contains a separable and a non-separable part:

$$\Gamma_{\mu\nu\kappa\tau}^{\text{PT2}} = D_{\mu\nu} P_{\kappa\tau} - D_{\mu\kappa}^\alpha P_{\nu\tau}^\alpha - D_{\mu\kappa}^\beta P_{\nu\tau}^\beta + \Gamma_{\mu\nu\kappa\tau}^{\text{NS}} \tag{7.90}$$

$$\Gamma_{\mu\nu\kappa\tau}^{\text{NS}} = \sum_{i_\alpha j_\alpha a_\alpha b_\alpha} c_{\mu i}^\alpha c_{\nu a}^\alpha c_{\kappa j}^\alpha c_{\tau b}^\alpha t_{a_\alpha b_\alpha}^{i_\alpha j_\alpha} + \sum_{i_\beta j_\beta a_\beta b_\beta} c_{\mu i}^\beta c_{\nu a}^\beta c_{\kappa j}^\beta c_{\tau b}^\beta t_{a_\beta b_\beta}^{i_\beta j_\beta} + 2 \sum_{i_\alpha j_\beta a_\alpha b_\beta} c_{\mu i}^\alpha c_{\nu a}^\alpha c_{\kappa j}^\beta c_{\tau b}^\beta t_{a_\alpha b_\beta}^{i_\alpha j_\beta} \tag{7.91}$$

Thus, the non-separable part is merely the back-transformation of the amplitudes from the MO to the AO basis. It is, however, important to symmetrize the two-particle density matrix in order to be able to exploit the full permutational symmetry of the AO derivative integrals.

7.13.4 Orbital Optimized MP2

The MP2 energy can be regarded as being stationary with respect to the MP2 amplitudes, since they can be considered as having been optimized through the minimization of the Hylleraas functional:

$$E_{\text{MP2}} = \min_{\mathbf{t}} \left\{ 2 \left\langle \Psi_1 | \hat{H} | \Psi_0 \right\rangle + \left\langle \Psi_1 | \hat{H}_0 - E_0 | \Psi_1 \right\rangle \right\} \tag{7.92}$$

$\hat{H}$ is the 0^{th} order Hamiltonian as proposed by Møller and Plesset, Ψ_0 is the reference determinant, Ψ_1 is the first-order wave function and $E_0 = E_{\text{HF}} = \langle \Psi_{\text{HF}} | \hat{H} | \Psi_{\text{HF}} \rangle$ is the reference energy. The quantities $\mathbf{t}$ collectively denote the MP2 amplitudes.

The fundamental idea of the OO-MP2 method is to not only minimize the MP2 energy with respect to the MP2 amplitudes, but to minimize the total energy additionally with respect to changes in the orbitals. Since the MP2 energy is not variational with respect to the MO coefficients, no orbital relaxation due to the correlation field is taken into account. If the reference determinant is poor, the low-order perturbative correction then becomes unreliable. This may be alleviated to a large extent by choosing better orbitals in the reference determinant. Numerical evidence for the correctness of this assumption will be presented below.

In order to allow for orbital relaxation, the Hylleraas functional can be regarded as a functional of the wavefunction amplitudes $\mathbf{t}$ and the orbital rotation parameters $\mathbf{R}$ that will be defined below. Through a suitable parameterization it becomes unnecessary to ensure orbital orthonormality through Lagrange multipliers. The functional that we minimize reads:

$$L\{\mathbf{t}, \mathbf{R}\} = E_0[\mathbf{R}] + 2 \langle \Psi_1 | \hat{H} | \Psi_0 \rangle + \langle \Psi_1 | \hat{H}_0 - E_0 | \Psi_1 \rangle \tag{7.93}$$

Ψ_0 is the reference determinant. However, it does no longer correspond to the Hartree-Fock (HF) determinant. Hence, the reference energy $E_0[\mathbf{R}] = \langle\Psi_0[\mathbf{R}]|\hat{H}|\Psi_0[\mathbf{R}]\rangle$ also changes during the variational process and is no longer stationary with respect to the HF MO coefficients. Obviously, $E_0[\mathbf{R}] \geqslant E_{\text{HF}}$ since the HF determinant is, by construction, the single determinant with the lowest expectation value of the full Hamiltonian.

The reference energy is given as:

$$E_0[\mathbf{R}] = \sum_i \langle i|h|i\rangle + \frac{1}{2}\sum_{ij}\langle ij||ij\rangle \tag{7.94}$$

The first-order wave function excluding single excitations is:

$$|\Psi_1\rangle = \frac{1}{4}\sum_{ijab} t_{ab}^{ij}|\Psi_{ij}^{ab}\rangle \tag{7.95}$$

A conceptually important point is that Brillouin's theorem [419] is no longer obeyed since the Fock matrix will contain off-diagonal blocks. Under these circumstances the first-order wavefunction would contain contributions from single excitations. Since the orbital optimization brings in all important effects of the singles we prefer to leave them out of the treatment. Any attempt to the contrary will destroy the convergence properties. We have nevertheless contemplated to include the single excitations perturbatively:

$$E_{\text{Singles}}^{(2)} = -\sum_{ia}\frac{|F_{ia}|^2}{\varepsilon_a - \varepsilon_i} \tag{7.96}$$

The perturbative nature of this correction would destroy the stationary nature of the total energy and is hence not desirable. Furthermore, results with inclusion of single excitation contributions represent no improvement to the results reported below. They will therefore not be documented below and henceforth be omitted from the OO-MP2 method by default.

The explicit form of the orbital-optimized MP2 Hylleraas functional employing the RI approximation (OO-RI-MP2) becomes:

$$L_\infty[\mathbf{t},\mathbf{R}] = \sum_i \langle i|\hat{h}|i\rangle + \frac{1}{2}\sum_{ij}\langle ij||ij\rangle + \sum_{iaP}(ia|P)\Gamma_{ia}^{'P} + \sum_{ij}D_{ij}F_{ij} + \sum_{ab}D_{ab}F_{ab} \tag{7.97}$$

with:

$$\Gamma_{ia}^{'P} = \sum_Q V_{PQ}^{-1}\sum_{jb}(Q|jb)t_{ab}^{ij} \tag{7.98}$$

$$(ia|P) = \int\int \psi_i(\mathbf{r}_1)\psi_a(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\eta_P(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.99}$$

$$(P|Q) = \int\int \eta_p(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\eta_Q(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.100}$$

Here, $\{\psi\}$ is the set of orthonormal molecular orbitals and $\{\eta\}$ denotes the auxiliary basis set. F_{pq} denotes a Fock matrix element:

$$F_{pq} = \langle p|\hat{h}|q\rangle + \sum_k \langle pk||qk\rangle \tag{7.101}$$

and it is insisted that the orbitals diagonalize the occupied and virtual subspaces, respectively:

$$\begin{aligned} F_{ij} &= \delta_{ij}F_{ii} = \delta_{ij}\varepsilon_i \\ F_{ab} &= \delta_{ab}F_{aa} = \delta_{ab}\varepsilon_a \end{aligned} \tag{7.102}$$

The MP2 like density blocks are,

$$\begin{aligned} D_{ij} &= -\tfrac{1}{2}\textstyle\sum_{kab} t_{ab}^{ik}t_{ab}^{jk} \\ D_{ab} &= \tfrac{1}{2}\textstyle\sum_{ijc} t_{ac}^{ij}t_{bc}^{ij} \end{aligned} \tag{7.103}$$

where the MP2 amplitudes in the case of a block diagonal Fock matrix are obtained through the condition $\frac{\partial L_\infty}{\partial t_{ab}^{ij}} = 0$:

$$t_{ab}^{ij} = -\frac{\langle ij||ab\rangle}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{7.104}$$

The orbital changes are parameterized by an anti-Hermitian matrix $\mathbf{R}$ and an exponential Ansatz,

$$\begin{aligned} \mathbf{c}^{\text{new}} &= \mathbf{c}^{\text{old}} \exp(\mathbf{R}) \\ \mathbf{R} &= \begin{pmatrix} 0 & \mathbf{R}_{ia} \\ -\mathbf{R}_{ia} & 0 \end{pmatrix} \end{aligned} \tag{7.105}$$

The orbitals changes to second order are,

$$\begin{aligned} \exp(\mathbf{R})\,|i\rangle &= |i\rangle + \textstyle\sum_a \mathbf{R}_{ai}\,|a\rangle - \frac{1}{2}\sum_{jb}\mathbf{R}_{bi}\mathbf{R}_{bj}\,|j\rangle + \ldots \\ \exp(\mathbf{R})\,|a\rangle &= |a\rangle - \textstyle\sum_i \mathbf{R}_{ai}\,|i\rangle - \frac{1}{2}\sum_{jb}\mathbf{R}_{aj}\mathbf{R}_{bj}\,|b\rangle + \ldots \end{aligned} \tag{7.106}$$

Through this Ansatz it is ensured that the orbitals remain orthonormal and no Lagrangian multipliers need to be introduced. The first-order expansion of the Fock operator due to the orbital rotations are:

$$F_{pq}\,[R] = F_{pq}\,[0] + R_{pq}^{(1)} + \sum_r R_{rp}F_{rq}\,[0] + R_{rq}F_{pr}\,[0] \tag{7.107}$$

$$R_{pq}^{(1)} = \sum_{kc} R_{ck}\left\{\langle pc||qk\rangle + \langle pk||qc\rangle\right\} \tag{7.108}$$

The first-order energy change becomes $\left(h_{pq} \equiv \langle p|\,\hat{h}\,|q\rangle\,, g_{pqrs} \equiv \langle pq||rs\rangle\right)$:

$$\begin{aligned} L_\infty\,[\mathbf{t},\mathbf{R}] = &\textstyle\sum_{ic} R_{ci}\,(h_{ci} + h_{ic}) + \frac{1}{2}\sum_{ijc} R_{ci}\,(g_{cjij} + g_{ijcj}) + R_{cj}\,(g_{icij} + g_{ijic}) \\ &+2\textstyle\sum_{iacP} R_{ci}(ac|P)\Gamma_{ia}^{'P} - 2\sum_{ikaP} R_{ak}(ik|P)\Gamma_{ia}^{'P} \\ &-\textstyle\sum_{ij} D_{ij}\left(R_{ij}^{(1)} + \sum_c\left(R_{ci}F_{cj} + R_{cj}F_{ic}\right)\right) \\ &+\textstyle\sum_{ab} D_{ab}\left(R_{ab}^{(1)} - \sum_k\left(R_{ak}F_{kb} + R_{bk}F_{ak}\right)\right) \end{aligned} \tag{7.109}$$

The condition for the energy functional to be stationary with respect to the orbital rotations $\left(\frac{\partial L_\infty[\mathbf{t},\mathbf{R}]}{\partial R_{ai}} = 0\right)$, yields the expression for the orbital gradient and hence the expression for the OO-RI-MP2 Lagrangian.

$$\begin{aligned} \frac{\partial L_\infty[\mathbf{t},\mathbf{R}]}{\partial R_{ai}} \equiv g_{ai} = 2F_{ai} + 2\sum_j D_{ij}F_{aj} - 2\sum_b D_{ab}F_{ib} + R^{(1)}(\mathbf{D})_{ai} \\ +2\sum_{cP}(ac|P)\Gamma_{ia}^{\prime P} - 2\sum_{kP}(ik|P)\Gamma_{ia}^{\prime P} \end{aligned} \tag{7.110}$$

The goal of the orbital optimization process is to bring this gradient to zero. There are obviously many ways to achieve this. In our experience, the following simple procedure is essentially satisfactory. We first build a matrix $\mathbf{B}$ in the current MO basis with the following structure:

$$\begin{aligned} \mathbf{B}_{ij} &= \delta_{ij}\mathbf{F}_{ii} \\ \mathbf{B}_{ab} &= \delta_{ab}(\mathbf{F}_{aa} + \boldsymbol{\Delta}) \\ \mathbf{B}_{ai} &= \mathbf{B}_{ia} = \mathbf{g}_{ai} \end{aligned} \tag{7.111}$$

where Δ is a level shift parameter. The occupied/occupied and virtual/virtual blocks of this matrix are arbitrary but their definition has a bearing on the convergence properties of the method. The orbital energies of the block diagonalized Fock matrix appear to be a logical choice. If the gradient is zero, the **B**-matrix is diagonal. Hence one obtains an improved set of orbitals by diagonalizing **B**.

In order to accelerate convergence a standard DIIS scheme is used. [451, 575] However, in order to carry out the DIIS extrapolation of the **B**-matrix it is essential that a common basis is used that does not change from iteration to iteration. Since the **B**-matrix itself is defined in the molecular orbitals of the current iteration we choose as a common set of orthonormal orbitals the MOs of the HF calculation. The extrapolation is carried out in this basis and the extrapolated **B**-matrix is transformed back to the current set of MOs prior to diagonalization. Obviously, the same strategy can be used for orbital optimization in any method for which an orbital gradient is available.

For well behaved cases this simple scheme converges in 5-10 iterations. Transition metals and more complicated molecules may require up to 20 iterations and level shifting in order to achieve convergence.

Upon convergence the sum of the matrix $\mathbf{D}$ and the density of the reference determinant $P_{\mu\nu} = \sum_i c_{\mu i}c_{\nu i}$ form the true one-particle density matrix of the OO-MP2 approach that can be used for property or gradient calculations.

7.13.5 Regularized MP2 and RI-MP2

Regularized MP2 is a variant of second-order Moller-Plesset theory (MP2) introduced by J. Shee, M. Loipersberger, A. Rettig, J. Lee, and M. Head-Gordon [790] that aims to improve its accuracy for systems with π-driven dispersion interactions and dative bonds in transition metal complexes. The approach achieves this by introducing a single-parameter, energy-gap dependent regularization that dampens overestimated pairwise additive contributions, thus renormalizing first-order amplitudes to empirically mimic higher-order correlations.

For this, the standard MP2 energy and thus the standard algorithms are modified. For the σ ($p = 1$) and σ^2 ($p = 2$) regularization, the energy is modified according to

$$E_{\sigma^p\text{-MP2}} = -\frac{1}{4}\sum_{ijab}\frac{|\langle ij||ab\rangle|^2}{\Delta_{ij}^{ab}}(1 - e^{-\sigma(\Delta_{ij}^{ab})^p})$$

which corresponds to regularizing the first-order amplitudes. For the κ regularization, the MP2 energy is modified according to

$$E_{\kappa\text{-MP2}} = -\frac{1}{4}\sum_{ijab}\frac{|\langle ij||ab\rangle|^2}{\Delta_{ij}^{ab}}\left(1 - e^{-\kappa(\Delta_{ij}^{ab})}\right)^2$$

which corresponds to regularizing the first-order amplitudes and the exchange integrals.

Regularized MP2 is available for standard MP2 in "memory mode" (Q1Opt>0) and RI-MP2 (RIJDX, RIJCOSX, RIJK).

The usage of regularized MP2 is controlled by the `DoRegMP2` keyword, the type of regularization can be specified by setting the `RegMP2Type` parameter to 0 for κ, 1 for σ, or 2 for σ^2. The value of the regularizers can be specified by `RegMP2Kappa` and `RegMP2Sigma` respectively.

```
%mp2
    DoRegMP2     true   # required
    RegMP2Type   0      # kappa regularizer
                 1      # sigma regularizer
                 2      # sigma-sqaured regularizer
    RegMP2Kappa  1.1    # kappa value
    RegMP2Sigma  0.5    # sigma value
end
```

It is important to note that only single point energies are available and tested for regularized MP2. Density, Gradient, and Hessian calculations are not yet supported.

7.13.6 RIJCOSX-RI-MP2 Gradients

Additional grids are introduced for the RIJCOSX-MP2 gradient. They have sensible default settings and therefore do not usually require any intervention from the user. However, a number of expert options are available, as described below.

The COSX terms in the Z-vector equations are calculated on a grid, controlled by the keywords `Z_GridX` and `Z_IntAccX`, as discussed in sections *Changing TD-DFT, CP-SCF and Hessian grids* and *CP-SCF Options*. For example, the `DefGrid3` CP-SCF COSX grid can be requested as:

```
%method
  Z_GridX   2     # Lebedev 110-point grid
  Z_IntAccX 3.067 # radial integration accuracy
end
```

The grid used for evaluation of the response operator on the right-hand side of the Z-vector equations (see for example eqs (7.77) and (7.78)) can be independently selected using the keyword `Z_GridX_RHS`. Note that starting with ORCA 5, the usage is different to `Z_GridX` - the choice is between one of the three grids used during the RIJCOSX SCF procedure: a small grid for the initial iterations, a medium grid for the final iterations (default in ORCA 5), and a large grid to evaluate the energy more accurately after the iterations have converged.

```
%method
  Z_GridX_RHS 1  # small SCF grid
              2  # medium SCF grid (default)
              3  # large SCF grid
end
```

Yet another grid is used to evaluate basis functions derivatives. Appropriate parameters are chosen through ! DefGridn (in addition to the three SCF grids), but one can override this by setting the angular (GridX) and radial (IntAccX) grids explicitly through:

```
%mp2 GridX    4       # default 4: angular Lebedev grid 302
     IntAccX  4.871   # radial grid
end
```

7.13.7 MP2 and RI-MP2 Second Derivatives

Analytical second-order properties with the MP2, RI-MP2 and double-hybrid DFT methods are available in ORCA for calculations without frozen core orbitals. The most expensive term in the second derivative calculations is the four-external contribution which can be evaluated either via an AO direct (default) or a semi-numerical Chain-of-Spheres approach. In case that the latter approach is chosen, appropriate grid parameters are defined through the ! DefGridn settings. However, a more fine-grained specification is available to expert users as follows:

```
%mp2 KCOpt _AOBLAS          # (default) AO direct with BLAS routines
           _COSX            # semi-numerical evaluation using the COSX method
     KC_GridX    2          # default 2: angular Lebedev grid 110
     KC_IntAccX  4.020      # radial grid
end
```

Alternatively, all the grid settings can be defined in the %method block, as discussed in section *SCF grid keyword list*. The first three entries define the three SCF grids, the fourth entry the MP2 grid for basis function derivatives (refer to section *RIJCOSX-RI-MP2 Gradients*) and the fifth entry the grid for the four-external contribution.

```
%method
   IntAccX     Acc1,  Acc2,  Acc3,  Acc4,   Acc5
   GridX       Ang1,  Ang2,  Ang3,  Ang4,   Ang5
end
```

7.13.8 RI-MP2 and Double-Hybrid DFT Response Properties

Starting from ORCA 5, both the electric (for the dipole polarizability) and the magnetic (for NMR shielding and the EPR g-tensor) field response as well as the nucleus-orbit response (hyperfine couplings A_{orb} term) for RI-MP2 (and double-hybrid functionals) is handled by a different implementation of the RI-MP2 second derivatives than that used for geometric Hessian calculations (*MP2 and RI-MP2 Second Derivatives*). This code is more efficient, uses the RI approximation throughout (including the four-external contribution) and supports frozen core orbitals. The implementation is described in detail in refs [828, 849]. Consider the following input for a GIAO-RI-MP2 NMR shielding calculation:

```
! RIJK RI-MP2 def2-SVP def2/JK def2-SVP/C TightSCF NMR NoFrozenCore
%mp2
  Density         relaxed # required
  UsePertCanOrbs  true    # Whether to use perturbed canonical orbitals for
                          # the internal block of the perturbed Fock matrix
  PertCan_EThresh 1e-6    # Energy threshold for special treatment of
                          # degenerate orbital pairs
  PertCan_UThresh 10      # Coefficient threshold for special treatment of
                          # strongly interacting orbital pairs
  FCut            1e-5    # Threshold for internal perturbed Fock elements
```

(continues on next page)

(continued from previous page)

```
  RespStoreT      true    # Whether to precalculate and store all necessary
                          # unperturbed amplitudes on disk
  RespDijConv     false   # Whether to store intermediates required for the
                          # internal block of the response density on disk
end
* int 0 1
O 0 0 0 0      0      0
H 1 0 0 1.1056 0      0
H 1 2 0 1.1056 109.62 0
*
```

By default perturbed canonical orbitals are used for the occupied block, i.e., the internal orbital rotation coefficients are chosen as

$$U_{ij}^{\mathbf{B}} = \frac{F_{ij}^{(\mathbf{B})} - S_{ij}^{(\mathbf{B})}\varepsilon_j}{\varepsilon_j - \varepsilon_i}$$

which results in $F_{ij}^{\mathbf{B}} = 0$, thereby eliminating its contribution to the perturbed amplitudes:

$$T_{ab}^{ij,\mathbf{B}} \leftarrow -\sum_k \left[T_{ab}^{ik} F_{kj}^{\mathbf{B}} + T_{ab}^{kj} F_{ki}^{\mathbf{B}}\right] \tag{7.112}$$

If $|\epsilon_j - \epsilon_i| <$ `PertCan_EThresh` or $|U_{ij}^B| >$ `PertCan_UThresh`, then $U_{ij}^{\mathbf{B}}$ is chosen using the standard formula

$$U_{ij}^{\mathbf{B}} = -\frac{1}{2} S_{ij}^{(\mathbf{B})}$$

And the relevant contributions to eq (7.112) are added, unless $\left|F_{ij}^{\mathbf{B}}\right| <$ `FCut`. The required amplitudes $\mathbf{T}^{ik}$ and $\mathbf{T}^{kj}$ (all amplitudes, in case `UsePertCanOrbs = false`) are stored on disk if `RespStoreT = true` or recalculated as needed otherwise. The latter option is significantly slower and not recommended unless disk space is an issue. Similarly, in the case of insufficient RAM, the option `RespDijConv = true` tells ORCA to store all amplitudes in the batch (required to calculate $D_{ij}^{\mathbf{B}}$) on disk, rather than keep them in memory. The 3-index 2-particle densities, needed for the right-hand side of the Z-vector equations, are always stored on disk.

Note also that in this implementation the RIJCOSX Fock-response terms are calculated with one of the SCF grids, chosen with `Z_GridX_RHS` (see section *CP-SCF Options*).

7.13.9 Local MP2

In analogy to the domain-based local pair natural orbital coupled-cluster methods, there is also a local linear scaling version of MP2 (DLPNO-MP2) implemented in ORCA. Its default thresholds are chosen to reproduce about 99.9% of the total RI-MP2 correlation energy, resulting in an accuracy of a fraction of 1 kcal/mol for energy differences. The theory has been described in the literature.[653, 684]

Further information of local correlation methods in ORCA can be found in section *Local correlation*. The local MP2 method becomes truly beneficial for very large molecules and can be used to compute energies of systems containing several hundred atoms. Fig. 7.3 shows the scaling behavior for linear alkane chains. Note that this represents an idealized situation. For three-dimensional molecules the crossover with canonical RI-MP2 is going to occur at a later point.

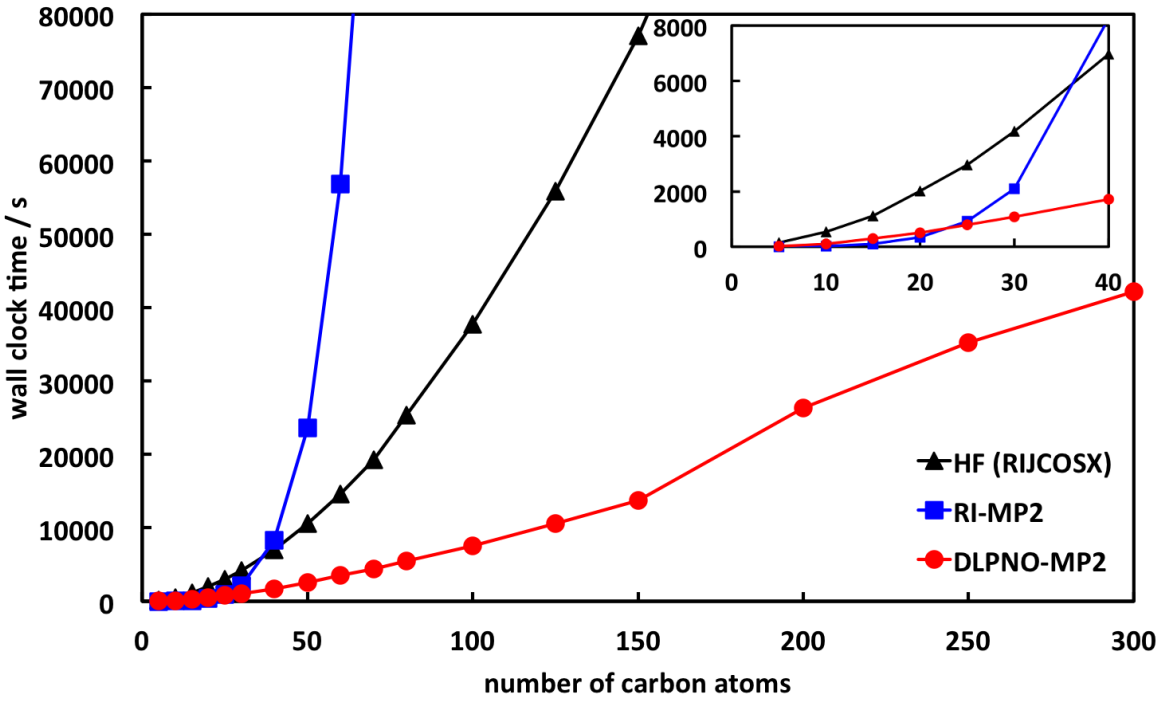


Fig. 7.3: Scaling of the DLPNO-MP2 method with default thresholds for linear alkane chains in def2-TZVP basis. Shown are also the times for the corresponding Hartree-Fock calculations with RIJCOSX and for RI-MP2.

In the following, the most important design principles of the RHF-DLPNO-MP2 are pointed out.

- Unlike in the 2013 version of the DLPNO methodology, domains are selected by means of the differential overlap $\sqrt{\int i^2(\mathbf{r})\tilde{\mu}'^2(\mathbf{r})d\mathbf{r}}$ between localized MOs i and projected atomic orbitals (PAOs) $\tilde{\mu}'$ which are normalized to unity. The default value for the corresponding cutoff is $T_{\text{CutDO}} = 10^{-2}$.
- MP2 amplitudes for each pair of localized orbitals ij are expressed in a basis of pair natural orbitals (PNOs) $\tilde{a}_{ij}$. PNOs are obtained from diagonalization of an approximate, "semicanonical" MP2 pair density $\mathbf{D}^{ij}$. Only PNOs with an occupation number $> T_{\text{CutPNO}}$ are retained, with a default value of $T_{\text{CutPNO}} = 10^{-8}$ for DLPNO-MP2. The pair density is given by:

$$\mathbf{D}^{ij} = \mathbf{T}^{ij\dagger}\tilde{\mathbf{T}}^{ij} + \mathbf{T}^{ij}\tilde{\mathbf{T}}^{ij\dagger} \quad \text{where} \quad \begin{aligned} T^{ij}_{\tilde{\mu}\tilde{\nu}} &= -\frac{(i\tilde{\mu}|j\tilde{\nu})}{\varepsilon_{\tilde{\mu}} + \varepsilon_{\tilde{\nu}} - F_{ii} - F_{jj}} \\ \tilde{\mathbf{T}}^{ij} &= (1+\delta_{ij})^{-1}\left(4\mathbf{T}^{ij} - 2\mathbf{T}^{ij\dagger}\right) \end{aligned}$$

- Since the occupied block of the Fock matrix is not diagonal in the basis of localized orbitals, the MP2 amplitudes $\mathbf{T}^{ij}$ are obtained by solving the following set of residual equations iteratively (where subscripts of PNOs have been dropped):

$$R^{ij}_{\tilde{a}\tilde{b}} = \left(i\tilde{a}\middle|j\tilde{b}\right) + \left(\varepsilon_{\tilde{a}} + \varepsilon_{\tilde{b}} - F_{ii} - F_{jj}\right) T^{ij}_{\tilde{a}\tilde{b}} - \sum_{k\neq i}\sum_{\tilde{c}\tilde{d}} F_{ik} S^{ij,kj}_{\tilde{a}\tilde{c}} T^{kj}_{\tilde{c}\tilde{d}} S^{kj,ij}_{\tilde{d}\tilde{b}} - \sum_{k\neq j}\sum_{\tilde{c}\tilde{d}} F_{kj} S^{ij,ik}_{\tilde{a}\tilde{c}} T^{ik}_{\tilde{c}\tilde{d}} S^{ik,ij}_{\tilde{d}\tilde{b}} = 0$$

- The largest part of the error relative to canonical RI-MP2 is controlled by the domain (`TCutDO`) and PNO (`TCutPNO`) thresholds, which should be adequate for most applications. If increased accuracy is needed (e.g. for studying weak interactions), tighter truncation criteria can be invoked by means of the `! TightPNO` keyword. Conversely, a less accurate but faster calculation can be performed with the `! LoosePNO` keyword. For more details refer to table Table 7.17.
- Fitting domains are determined by means of orbital Mulliken populations with a threshold $T_{\text{CutMKN}} = 10^{-3}$. This threshold results in an error contribution that is typically about an order of magnitude smaller than the overall total energy deviation from RI-MP2.

- Prior to performing the local MP2 calculation, pairs of localized molecular orbitals ij are prescreened using an MP2 energy estimate with a dipole approximation, and the differential overlap integral between orbitals i and j. This procedure has been chosen conservatively and leads to minimal errors.
- Residual evaluation can be accelerated significantly by neglecting terms with associated Fock matrix elements F_{ik} and F_{kj} below $F_{\text{Cut}} = 10^{-5}$. Errors resulting from this approximation are typically below 1 μE_h and thus negligible.
- Sparsity of the MO and PAO coefficient matrices in atomic orbital basis is exploited to accelerate integral transformations for large systems. Truncation of the coefficients is controlled by a parameter T_{CutC}. Neglect of these coefficients has to be performed very carefully in order to avoid uncontrollable errors. The threshold has been chosen so as to make the errors essentially vanish.
- By default, core orbitals are frozen in the MP2 module. However, if core orbitals are subject to an MP2 treatment, it is necessary to use a tighter PNO cutoff for pairs involving at least one core orbital. For this purpose core orbitals and valence orbitals are localized separately. The cutoff for pairs involving core orbitals is given by $T_{\text{CutPNO}} \times T_{\text{ScalePNO\_Core}}$, where $T_{\text{ScalePNO\_Core}} = 10^{-2}$ by default. For more details refer to section *Including (semi)core orbitals in the correlation treatment*.

The UHF-DLPNO-MP2 implementation differs somewhat from the RHF case, particularly regarding construction of PNOs, as described below.

- A separate set of PAOs $\tilde{\mu}'_\alpha$ and $\tilde{\mu}'_\beta$ is obtained for each spin case.
- For $\alpha\beta$ pairs, separate pair domains of PAOs need to be determined for each spin case. For example, the α pair domain $[i_\alpha j_\beta]_\alpha$ is the union of the domains $[i_\alpha]_\alpha$ and $[j_\beta]_\alpha$. The latter domain $[j_\beta]_\alpha$ is determined by evaluating the spatial differential overlap between j_β and α-spin PAOs $\tilde{\mu}'_\alpha$.
- One set of PNOs is needed for each same-spin pair. Opposite-spin pairs require a set of α-PNOs and a set of β-PNOs. In total this results in four types of PNO sets.
- Semicanonical amplitudes are obtained as follows, where i, j are spin orbitals and $\tilde{\mu}\tilde{\nu}$ are nonredundant spin PAOs.

$$T^{ij}_{\tilde{\mu}\tilde{\nu}} = -\frac{\langle ij||\tilde{\mu}\tilde{\nu}\rangle}{\varepsilon_{\tilde{\mu}} + \varepsilon_{\tilde{\nu}} - F_{ii} - F_{jj}}$$

  In the same-spin case $\langle i_\alpha j_\alpha||\tilde{\mu}_\alpha\tilde{\nu}_\alpha\rangle = \langle ij|\tilde{\mu}\tilde{\nu}\rangle - \langle ij|\tilde{\nu}\tilde{\mu}\rangle$, while in the opposite-spin case $\langle i_\alpha j_\beta||\tilde{\mu}_\alpha\tilde{\nu}_\beta\rangle = \langle ij|\tilde{\mu}\tilde{\nu}\rangle$.
- For opposite-spin pairs, α-PNOs and β-PNOs are obtained from diagonalisation of $\mathbf{T}^{ij}\mathbf{T}^{ij\dagger}$ and $\mathbf{T}^{ij\dagger}\mathbf{T}^{ij}$, respectively. For same-spin pairs the pair density is symmetric and only one set of PNOs is needed. PNOs are discarded whenever the absolute value of their natural occupation number is below the threshold T_{CutPNO}.
- The following residual equations need to be solved for the cases $R^{i_\alpha j_\alpha}_{\tilde{a}_\alpha\tilde{b}_\alpha}$, $R^{i_\beta j_\beta}_{\tilde{a}_\beta\tilde{b}_\beta}$ and $R^{i_\alpha j_\beta}_{\tilde{a}_\alpha\tilde{b}_\beta}$:

$$\begin{aligned} R^{i_\sigma j_\tau}_{\tilde{a}_\sigma\tilde{b}_\tau} = &\left\langle ij\middle|\middle|\tilde{a}\tilde{b}\right\rangle + \left(\varepsilon_{\tilde{a}} + \varepsilon_{\tilde{b}} - F_{ii} - F_{jj}\right) T^{ij}_{\tilde{a}\tilde{b}} \\ &- \sum_{k_\sigma \neq i_\sigma} \sum_{\tilde{c}_\sigma \tilde{d}_\tau} F_{ik} S^{ij,kj}_{\tilde{a}\tilde{c}} T^{kj}_{\tilde{c}\tilde{d}} S^{kj,ij}_{\tilde{d}\tilde{b}} - \sum_{k_\tau \neq j_\tau} \sum_{\tilde{c}_\sigma \tilde{d}_\tau} F_{kj} S^{ij,ik}_{\tilde{a}\tilde{c}} T^{ik}_{\tilde{c}\tilde{d}} S^{ik,ij}_{\tilde{d}\tilde{b}} = 0 \end{aligned}$$

- Most approximations are consistent between the RHF and UHF schemes, with the exception of the PNO truncation. This means that results would match for closed-shell molecules with $T_{\text{CutPNO}} = 0$ (provided both Hartree-Fock solutions are identical), but this is not true whenever the PNO space is truncated. Therefore, UHF-DLPNO-MP2 energies should only be compared to other UHF-DLPNO-MP2 energies, even for closed-shell species.
- We found that it is necessary to use tighter PNO thresholds for UHF-DLPNO-MP2. With NormalPNO settings the default value is $T_{\text{CutPNO}} = 10^{-9}$. For an overview of accuracy settings refer to table Table 7.17. As in the RHF implementation, the PNO cutoff for pairs involving core orbitals is scaled with $T_{\text{ScalePNO\_Core}}$.

Table 7.17: Accuracy settings for DLPNO-MP2.

Setting	T_{CutDO}	T_{CutPNO} (RHF)	T_{CutPNO} (UHF)
LoosePNO	2×10^{-2}	10^{-7}	10^{-8}
NormalPNO	1×10^{-2}	10^{-8}	10^{-9}
TightPNO	5×10^{-3}	10^{-9}	10^{-10}

Options specific to DLPNO-MP2 are listed below.

```
%mp2 DLPNO            false  # Do DLPNO-MP2 (also requires RI true)
     TolE             1e-6   # Energy convergence threshold. Default: TolE of SCF
     TolR             5e-6   # Residual convergence threshold. Default: 5 * TolE
     MaxPNOIter       100    # Maximum number of residual iterations
     MaxLocIter       128    # Maximum number of iterations for orbital localization
     LocMet           AHFB   # Localization method
                             # options: FB, PM, IAOIBO, IAOBOYS, NEWBOYS, AHFB
     LocTol           1e-6   # Localization convergence tolerance
                             # Default: 0.1 * TolG from SCF
     DIISStart_PNO    0      # First iteration to invoke DIIS extrapolation
     MaxDIIS_PNO      7      # length of DIIS vector
     Damp1_PNO        0.5    # Damping before DIIS is started
     Damp2_PNO        0.0    # Damping with DIIS
     MP2Shift_PNO     0.2    # level shift in amplitude update (Eh)
# Truncation parameters:
     TCutPNO          1e-8   # PNO occupation number cutoff (RHF)
                      1e-9   # PNO occupation number cutoff (UHF)
     TScalePNO_Core   1e-2   # Core PNO scaling factor
     TCutDO           1e-2   # Differential overlap cutoff for domain selection
     TCutMKN          1e-3   # Mulliken population cutoff for fitting domain selection
     FCut             1e-5   # Occupied Fock matrix element cutoff
     TCutPre          1e-6   # Energy threshold for dipole prescreening
     TCutDOij         1e-5   # Maximum differential overlap between screened MOs
     TCutDOPre        3e-2   # Cutoff to select PAOs for domains in prescreening
     TCutC            1e-3   # Cutoff for PAO coefficient truncation
     ScaleTCutC_MO    1.0    # Cutoff for MO truncation: TCutC * ScaleTCutC_MO
     PAOOverlapThresh 1e-8   # Threshold for constructing non-redundant PAOs
end
```

Local MP2 Gradient

The analytical gradient has been implemented for the RHF variant of the DLPNO-MP2 method.[685, 686] It is a complete derivative of all components in the DLPNO-MP2 energy, and the results are therefore expected to coincide with numerical derivatives of DLPNO-MP2 (minus the noise). General remarks:

- No gradient is presently implemented for the UHF-DLPNO-MP2 variant.
- Spin-component scaled MP2 is supported by the gradient.
- Double-hybrid density functionals are supported by the gradient.
- Only Foster-Boys localization is presently supported. The default converger is `AHFB` with a convergence tolerance that is automatically bound by a constant factor to the SCF orbital gradient tolerance. Using a different converger is possible, but discouraged, as the orbital localization needs to be sufficiently tightly converged.
- When calculating properties without the full nuclear gradient, the relaxed MP2 density should be requested.

A number of points regarding geometry optimizations (not all of them specific to DLPNO-MP2) are worth keeping in mind:

- As of 2018, we expect that the DLPNO-MP2 gradient can most beneficially be used for geometry optimizations of systems containing around 70-150 atoms. It may be faster than RI-MP2 even for systems containing 50-60 atoms or less, but the timing difference is probably not going to be very large. Of course, structures containing 200 atoms and above can (and have been) optimized, but this may take long if many geometry

steps are required. On the other hand, single point gradient or density calculations can be performed for systems containing many hundred atoms.

- DLPNO-MP2 is a substantially more expensive method for geometry optimizations than GGA or hybrid DFT functionals. Therefore, it is generally a good idea to start a geometry optimization with a structure that is already optimized at dispersion-corrected DFT level.
- RIJCOSX can be used to accelerate exchange evaluation substantially. However, great care needs to be exercised with the grid settings. Insufficiently large grids may lead, for example, to non-planar distortions of planar molecules. The updated default grids in ORCA 5 (`DefGrid2-3`) should be sufficiently accurate to optimize neutral main group compounds. We therefore recommend these grids for general use with some careful checking in more complicated cases. Even with these grids, the calculation is a lot faster than "regular" Hartree-Fock with basis sets of triple zeta quality (or larger).

  Using RIJONX is also possible.
- Sufficiently large grids should be used for the exchange-correlation functional of double hybrids. The SCF calculation takes only a fraction of the time that is needed for DLPNO-MP2, and sacrificing quality because of an insufficiently accurate grid is a waste of computer time.
- Optimization of large structures is often a challenge for the geometry optimizer. It may help to change the trust radius settings, to modify the settings of the `AddExtraBonds` feature, or to change other settings of the geometry optimizer. Sometimes it may be beneficial to check the geometry optimizer settings with a less demanding electronic structure method.
- Finally, problems with a geometry optimization may in some cases indeed be caused by the DLPNO approximations. Using `LoosePNO` for accurate calculations is not recommended anyway, and difficulties with `NormalPNO` settings are possibly rectified by switching to `TightPNO`.

During the development process, a number of difficulties were encountered related to the orbital localization Z-vector equations. Great care was taken to work around these problems and to make the procedures as robust as possible, but a number of settings can be changed. For more information on these aspects, we recommend consulting the full paper on the DLPNO-MP2 gradient [686].

- Several different solvers are implemented for the orbital localization Z-vector equations. The default is an iterative conjugate gradient solver. As an alternative, the DIIS-accelerated Jacobi solver can be used, but it tends to be inferior to the conjugate gradient solver. Moreover, a direct solver is available as a fail-safe alternative for smaller systems. As the dimension of the linear equation system is $n(n-1)/2$ for n occupied orbitals, the memory requirement and FLOP count increase as $O(n^4)$ and $O(n^6)$, respectively, and using the direct solver becomes unfeasible for large systems.
- Settings for the CPSCF solver are specified the same way as for canonical MP2.
- Under specific circumstances, the orbital Hessian of the orbital localization function may have zero or near-zero eigenvalues, which can lead to singular localization Z-vector equations. In particular, it is typically a consequence of continuously degenerate localized orbitals, which may (but do not need to) appear in some molecular symmetries.[760] A typical symptom are natural occupation number above 2 and below 0 for systems that would be expected to have MP2 density eigenvalues between 2 and 0 without the DLPNO approximations.
- In order to work around the aforementioned problem, a procedure has been implemented to eliminate singular or near-singular eigenvectors of the localization function orbital Hessian. Vectors with an eigenvalue smaller than `ZLoc_EThresh` (or `ZLoc_EThresh_core` for the core orbitals) are subject to the modified procedure. If the program eliminates any eigenvectors, it might sometimes be a good idea to check if calculated properties are reasonable (or at least to check the natural occupation numbers). Eigenvectors of the Hessian are calculated by Davidson diagonalization by default, but direct diagonalization can be chosen for smaller systems, instead.
- Diagonalization of the localization orbital Hessian can be switched off altogether by setting `ZLoc_EThresh` to 0.
- If the "Asymmetric localization equation residual norm" exceeds the localization Z-vector equation tolerance (`ZLoc_Tol`), there are typically two plausible reasons: (1) the localized orbitals are not sufficiently tightly converged (too large `LocTol`) or unconverged, or (2) the orbital localization Hessian has got small eigenvalues that were not eliminated.

This is an overview over the options related to the gradient:

```
# Settings specific to the localization equation z-solver
%mp2 ZLoc_Solver        CG      # Use conjugate gradient solver (default)
                        DIR     # Use direct solver
                        JAC     # Use DIIS-accelerated Jacobi solver
     ZLoc_Tol           1.0e-3  # Residual convergence tolerance for the
                                # localization Z-solver
                                # Default: same value as Z_Tol for CPSCF
     ZLoc_MaxIter       1024    # Maximum localization Z-solver iterations
     ZLoc_MaxDIIS       10      # Number of DIIS vectors for the Jacobi solver
     ZLoc_Shift         0.2     # Shift for the Jacobi solver
# Options for eliminating (near-)singular eigenvectors of the
# orbital Hessian of the localization function.
     ZLoc_EThresh       3.0e-4  # Eigenvectors with an eigenvalue below
                                # this threshold are eliminated.
     ZLoc_EThresh_core  3.0e-4  # Same as ZLoc_EThresh, but for the core orbitals.
                                # Default: identical value as ZLoc_EThresh.
# Options for determining eigenvectors of the localization orbital Hessian.
     ZLoc_UseDavidson   True    # Use Davidson diagonalization.
                                # If false, use direct diagonalization.
     ZLoc_DVDRoots      32      # Number of Davidson roots to be determined.
     ZLoc_DVDNIter      256     # Number of Davidson iterations.
     ZLoc_DVDTolE       3.0e-10 # Eigenvalue tolerance for the Davidson solver.
                                # Default: 1e-6 * ZLoc_EThresh
     ZLoc_DVDTolR       1.0e-7  # Residual tolerance for the Davidson solver.
                                # Default: 0.1 * (ZLoc_Tol)^2
     ZLoc_DVDMaxDim     10      # During Davidson diagonalization, the space of trial
                                # vectors is expanded up to MaxDim * DVDRoots.
# Choice of the PNO processing algorithm.
     DLPNOGrad_Opt      AUTO      # Chooses automatically between RAM and DISK
                                  # (default and recommended)
                        RAM       # Enforce memory-based one-pass algorithm
                        DISK      # Enforce disk-based two-pass algorithm
                        BUFFERED  # Buffered algorithm. Usage is discouraged.
                                  # Experimental, unpredictable I/O performance.
end
```

Local MP2 Response Properties

Analytical second derivatives with respect to electric and magnetic fields are implemented for closed-shell DLPNO-MP2 (as well as double-hybrid DFT).[826] Thus, analytic dipole polarizability and NMR shielding tensors (with our without GIAOs) are available. All considerations and options discussed in sections *Local MP2* and *Local MP2 Gradient* apply here as well, while additional remarks specific to second derivatives are given below.

- DLPNO-MP2 response property calculations are expected to be faster than the RI-MP2 equivalents for systems larger than about 70 atoms or 300 correlated electrons.
- Using the `NormalPNO` default thresholds, relative errors in the calculated properties, due to the local approximations, are smaller than 0.5%, or 5–10 times smaller than the inherent inaccuracy of MP2 vs a more accurate method like CCSD(T).
- DLPNO-MP2 second derivatives are much more sensitive to near-linear dependencies and other numerical issues than the energy or first-order properties. We have made efforts to choose reasonable and robust defaults, however we encourage the user to be critical of the results and to proceed with caution, especially if diffuse basis sets or numerical integration (DFT, COSX) are used. In the latter case, `DefGrid3` is recommended.
- In particular, the near-redundancy of PAO domains introduces numerical instabilities in the algorithm. Hence, these should be truncated at `PAOOverlapThresh=1e-5`, which is higher than the default for the energy and gradient. Therefore, the energy and gradient in jobs, which include response property calculations, may deviate from jobs, which do not. The difference is much smaller than the accuracy of the method

(vs RI-MP2) but it is still advisable to use the same value of `PAOOverlapThresh` in all calculations, when calculating, e.g., relative energies.

- For the same reason, if diffuse basis sets are used, it is advised to set `SThresh=1e-6` in the `%scf` block.
- Another instability arises due to small differences between the occupation number of kept and discarded PNOs and may result in very large errors. The smallest difference is printed during the DLPNO-MP2 relaxed density calculation:

  Smallest occupation number difference between PNOs and complementary PNOs. Absolute: 3.10e-10 Relative: 3.28e-02

  We found that a relative difference under 10^{-3}, which is not uncommon, may be cause for concern. To regularize the unstable equations, a level shift is applied, which is equal to T_{CutPNO} multiplied by $T_{\text{ScalePNO\_LShift}}$. A reasonable value of `TScalePNO_LShift=0.1` is set by default for response property calculations but not for gradient (or energy) calculations, as these were not found to suffer from this issue, so the same considerations as above apply.
- The option `DLPNOGrad_Opt=BUFFERED` is not implemented for response properties.

A summary of the additional options used for DLPNO-MP2 response properties is given below:

```
%mp2
     PAOOverlapThresh 1e-5   # Threshold for constructing non-redundant PAOs
                             # Default is 1e-8 for energy/gradient calculations!
     TScalePNO_LShift 0.1    # Level shift for PNO constraint equations:
                             # TScalePNO_LShift * TCutPNO
                             # Default is 0 for energy/gradient calculations!
end
```

An example input for a DSD-PBEP86 calculation of the NMR shielding and dipole polarizability tensors employing DLPNO-MP2 is given below. Note that the def2-TZVP basis set is not necessarily ideal for either shielding or polarizability.

```
! DLPNO-DSD-PBEP86/2013 D3BJ def2-TZVP def2-TZVP/C TightSCF NoFrozenCore
! RIJCOSX def2/J DefGrid3      # Use RIJCOSX with tighter grid settings
! NMR                         # Request NMR shielding
%elprop Polar 1 end           # Request polarizability
%mp2                          # These settings are default for response properties
   Density Relaxed
   PAOOverlapThresh 1e-5
   TScalePNO_LShift 0.1
end
%eprnmr
  GIAO_2el GIAO_2el_RIJCOSX   # Also use RIJCOSX for GIAO integrals
end                           # (This is the default for !RIJCOSX)
*xyzfile 0 1 geometry.xyz
```

Numerical DLPNO-MP2 derivatives

The various truncations in local correlation methods introduce small discontinuities in the potential energy surface. For example, a small displacement may change the sizes of correlation domains, leading to a slightly larger or smaller error from the domain approximation. The default DLPNO-MP2 truncation thresholds are conservative enough, so that these discontinuities should not cause problems in geometry optimizations using analytic gradients.[686] However, if one wishes to calculate (semi-)numerical derivatives of the DLPNO-MP2 energy, gradient, or properties using finite differences, large errors can occur. Therefore, in these cases it is advisable to keep the pair lists and correlation domains fixed upon displacement. Currently, this can be achieved using the following procedure: first, the calculation at the reference geometry is carried out with the additional setting `StoreDLPNOData=true`:

```
! DLPNO-MP2 def2-SVP def2-SVP/C VeryTightSCF
%base "calc0"
```

(continues on next page)

(continued from previous page)

```
%mp2
  StoreDLPNOData true
end
*xyzfile 0 1 geom0.xyz
```

This will produce the additional files *calc0.MapDLPNO00.tmp*, *calc0.MapDLPNOPre0.tmp*, *calc0.IJLIST.0.tmp*, *calc0.IJLISTSCR.0.tmp*, and *calc0.IJNPNO.0.tmp*, which are needed in the working directory for the next step together with the localized orbitals in *calc0.loc*. The calculation at the displaced geometry is then requested as:

```
! DLPNO-MP2 def2-SVP def2-SVP/C VeryTightSCF
%mp2
  RefBaseName "calc0"
end
*xyzfile 0 1 geom1.xyz
```

The program will use the orbitals from *calc0.loc* as a starting guess for the localization and map the reference orbitals to the new ones based on maximal overlap. The lists of correlated and screened out pairs are read from the files *calc0.IJLIST.0.tmp* and *calc0.IJLISTSCR.0.tmp*, while the domain information (MO-PAO, MO-Aux, etc.) is read from *calc0.MapDLPNO00.tmp* and *calc0.MapDLPNOPre0.tmp*. The number of PNOs for a each pair (stored in *calc0.IJNPNO.0.tmp*) is also kept consistent with the reference calculation: the ones with the higest occupation numbers are kept, disregarding T_{CutPNO}.

This procedure should improve the accuracy and numerical stability for manually calculated geometric derivatives of various DLPNO-MP2 properties (including those that require analytic first or second derivatives at the displaced geometries). For semi-numerical Hessian calculations (`NumFreq`), it is sufficient to add `StoreDLPNOData=true` as shown below and ORCA will handle the rest. For the sake of numerical stability, it is also recommended to increase `PAOOverlapThresh` and add a PNO level shift for the reasons described in section *Local MP2 Response Properties*.

```
! DLPNO-MP2 def2-SVP def2-SVP/C VeryTightSCF NumFreq
%mp2
  StoreDLPNOData true
  PAOOverlapThresh 1e-5
  TScalePNO_LShift 0.1
end
# geometry definition
```

Note that in case the orbital localization Hessian is (near-)singular, the mapping of orbitals from reference to displaced geometries will likely fail. No solution is presently implemented for this problem.

Multi-Level DLPNO-MP2 calculations

With the DLPNO-MP2 method it is possible to treat the interactions among different fragments of a system with varying accuracy, or exclude some interactions from the electron correlation treatment entirely. A more detailed discussion in the DLPNO-CCSD(T) context is given in section *Multi-Level Calculations* and in ref. [811]. Here we just present the technical capabilities of the MP2 module and the required input. Currently, multilayer calculations are only available for closed-shell DLPNO-MP2. Multilayer gradients and response properties are also possible. Fragments must be defined – see *Fragment Specification*.

```
! DLPNO-MP2 NoFrozenCore TightPNO
%mp2
  LoosePNOFragInter  {1 *} # * can be used as a wildcard for either or both indices
  NormalPNOFragInter {1 1} {1 2} # multiple fragment pairs can be listed like this
  TightPNOFragInter  {2 3}
  HFFragInter        {3 1} {4 2}
  CustomFragInter
    FragPairs        {4 4} {3 4} # pair definition is required
    HFOnly           false # flag to skip MP2 for these pairs - same as HFFragInter
    FrozenCore       false # flag to skip core correlation - requires !NoFrozenCore
```

(continues on next page)

(continued from previous page)

```
    TCutPNO          1e-8  # custom value for these pairs
    TCutDO           1e-2  # custom value for these pairs
    TCutDOij         1e-5  # custom value for these pairs
    TCutPre          1e-6  # custom value for these pairs
  end
end
# geometry and fragment definition
```

Note that a given pair or fragments can only belong to a single layer and definitions later in the input overwrite previous ones. This means that if the above input is used in a 4-fragment calculation, the 1-4 interfragment interactions will be treated with `LoosePNO` thresholds, the interactions within fragment 1 and with fragment 2 – with `NormalPNO` thresholds, 2-3 pairs – with `TightPNO`, 1-3 and 2-4 pairs will be left at the HF level, 3-4 and 4-4 pairs will be treated with $T_{\text{CutPNO}} = 10^{-8}$ and $T_{\text{CutDO}} = 10^{-2}$ (i.e. the `NormalPNO` defaults), and 2-2 and 3-3 pairs will be left at the global (`TightPNO`) settings.

7.14 The Single Reference Correlation Module

ORCA features a variety of single-reference correlation methods for single point energies (restricted to a RHF or RKS determinant in the closed-shell case and a UHF or UKS determinant in the open-shell case; quasi-restricted orbitals (QROs)[611] are also supported in the open-shell case). They are all fairly expensive but maybe be used in order to obtain accurate results in the case that the reference determinant is a good starting point for the expansion of the many-body wavefunction. The module is called `orca_mdci` for "matrix driven configuration interaction". This is a rather technical term to emphasize that if one wants to implement these methods (CCSD, QCISD etc.) efficiently, one needs to write them in terms of matrix operations which pretty much every computer can drive at peak performance. Let us first briefly describe the theoretical background of the methods that we have implemented in ORCA.

7.14.1 Theory

We start from the full CI hierarchy in which the wavefunction is expanded as:

$$|\Psi\rangle = |0\rangle + |S\rangle + |D\rangle + |T\rangle + |Q\rangle + ... \tag{7.113}$$

where $|0\rangle$ is a single-determinant reference and S, D, T, Q, ... denote the single, double, triple quadruple and higher excitations relative to this determinant at the spin-orbital level. As usual, labels i, j, k, l refer to occupied orbitals in $|0\rangle$, a, b, c, d to unoccupied MOs and p, q, r, s to general MOs. The action of the second quantized excitation operators $a_i^a = a_a^\dagger a_i$ on $|0\rangle$ lead to excited determinants $|\Phi_i^a\rangle$ that enter $|\Psi\rangle$ with coefficients C_a^i. The variational equations are:

$$\langle \Phi_i^a | H - E_0 | 0 + S + D \rangle = E_C C_a^i - \langle \Phi_i^a | H - E_0 | T \rangle \tag{7.114}$$

$$\left\langle \Phi_{ij}^{ab} | H - E_0 | 0 + S + D \right\rangle = E_C C_{ab}^{ij} - \left\langle \Phi_{ij}^{ab} | H - E_0 | T + Q \right\rangle \tag{7.115}$$

Further equations coupling triples with singles through pentuples etc.

The total energy is the sum of the reference energy $E_0 = \langle 0 | H | 0 \rangle$ and the correlation energy

$$E_C = \langle 0 | H | S + D \rangle \tag{7.116}$$

which requires the exact singles- and doubles amplitudes to be known. In order to truncate the series to singles- and doubles one may either neglect the terms containing the higher excitations on the right hand side (leading to CISD) or approximate their effect thereby losing the variational character of the CI method (CCSD, QCISD and CEPA methods). Defining the one- and two-body excitation operators as $\hat{C}_1 = \sum_{ia} C_a^i a_i^a$, $\hat{C}_2 = \frac{1}{4} \sum_{ijab} C_{ab}^{ij} a_{ij}^{ab}$ one can proceed to approximate the triples and quadruples by the disconnected terms:

$$|T\rangle = \hat{C}_1 \hat{C}_2 |0\rangle \tag{7.117}$$

$$|Q\rangle = \frac{1}{2}\hat{C}_2^2 |0\rangle \tag{7.118}$$

As is well known, the CCSD equations contain many more disconnected contributions arising from the various powers of the $\hat{C}_1$ operator (if one would stick to CC logics one would usually label the cluster amplitudes with t_a^i, t_{ab}^{ij},... and the n-body cluster operators with $\hat{T}_n$; we take a CI point of view here). In order to obtain the CEPA type equations from ((7.114)-(7.118)), it is most transparent to relabel the singles and doubles excitations with a compound label P for the internal indices (i) or (ij) and x for (a) or (ab). Then, the approximations are as follows:

$$\frac{1}{2}\left\langle \Phi_P^x \middle| (H-E_0)\,\hat{C}_2^2 \middle| 0 \right\rangle = \frac{1}{2}\sum_{QRyz} C_y^Q C_z^R \left\langle \Phi_P^x \middle| H-E_0 \middle| \Phi_{QR}^{yz} \right\rangle \tag{7.119}$$

$$\approx C_x^P \sum_{Qy} C_y^Q \left\langle \Phi_P^x \middle| H \middle| \Phi_{PQ}^{xy} \right\rangle \tag{7.120}$$

$$= C_x^P \sum_{Qy} C_y^Q \left\langle 0 \middle| H \middle| \Phi_Q^y \right\rangle - C_x^P \sum_{Qy\cup Px} C_y^Q \left\langle 0 \middle| H \middle| \Phi_Q^y \right\rangle \tag{7.121}$$

$$\approx C_x^P \left(E_\mathrm{C} - \sum_{Q\cup P} \varepsilon_Q \right) \tag{7.122}$$

Here the second line contains the approximation that only the terms in which either *Qy* or *Rz* are equal to *Px* are kept (this destroys the unitary invariance) and the fourth line contains the approximation that only "exclusion principle violating" (EPV) terms of internal labels are considered. The notation $Qy\cup Px$ means "*Qy* joint with *Px*" (containing common orbital indices) and ε_Q is the pair correlation energy. The EPV terms must be subtracted from the correlation energy since they arise from double excitations that are impossible due to the fact that an excitation out of an occupied or into an empty orbital of the reference determinant has already been performed. Inserting eq. (7.122) into eq. (7.115) $C_x^P E_C$ cancels and effectively is replaced by the "partial correlation energy" $\sum_{Q\cup P}\varepsilon_Q$.

The resulting equations thus have the appearance of a diagonally shifted ("dressed") CISD equation $\langle \Phi_P^x | H - E_0 + \Delta | 0 + S + D\rangle = 0$. If the second approximation mentioned above is avoided Malrieu's (SC)^2-CISD arises. [38, 63, 646, 719, 911] Otherwise, one obtains CEPA/3 with the shift:

$$-\Delta_{ab}^{ij} = \sum_k (\varepsilon_{ik} + \varepsilon_{jk}) - \varepsilon_{ij} \tag{7.123}$$

CEPA/2 is obtained by $-\Delta_{ab}^{ij} = \varepsilon_{ij}$ and CEPA/1 is the average of the CEPA/2 and CEPA/3. As mentioned by Ahlrichs, [11] no consensus appears to exist in the literature for the appropriate shift on the single excitations. If one proceeds straightforwardly in the same way as above, one obtains:

$$\left\langle \Phi_i^a \middle| (H-E_0)\,\hat{C}_1\hat{C}_2 \middle| 0 \right\rangle \approx C_a^i \left(E_\mathrm{C} - 2\sum_k \varepsilon_{ik} \right) \tag{7.124}$$

as the appropriate effect of the disconnected triples on the singles. It has been assumed here that only the singles $|\Phi_i^a\rangle$ in $\hat{C}_1$ contribute to the shift. If $|0\rangle$ is a HF determinant, the effect of the disconnected triples in the doubles projection vanishes under the same CEPA approximations owing to Brillouin's theorem. Averaged CEPA models are derived by assuming that all pair correlation energies are equal (except $\varepsilon_{ii} = 0$). As previously discussed by Gdanitz [291], the averaging of CEPA/1 yields $\frac{2}{n}E_\mathrm{C}$ and CEPA/3 $E_\mathrm{C}\frac{4n-6}{n(n-1)}$ where n is the number of correlated electrons. These happen to be the shifts used for the averaged coupled-pair functional (ACPF [292]) and averaged quadratic coupled-cluster (AQCC [840]) methods respectively. However, averaging the singles shift of eq. (7.124) gives $\frac{4}{n}E_\mathrm{C}$. The latter is also the leading term in the expansion of the AQCC shift for large n. In view of the instability of ACPF in certain situations, Gdanitz has proposed to use the AQCC shift for the singles and the original ACPF shift for the doubles and called his new method ACPF/2 [292]. Based on what has been argued above, we feel that it would be most consistent with the ACPF approach to simply use $\frac{4}{n}E_\mathrm{C}$ as the appropriate singles shift. We refer to this as NACPF.

It is readily demonstrated that the averaged models may be obtained by a variation of the modified correlation energy functional:

$$E_\mathrm{C} = \frac{\langle 0 + S + D | H - E_0 | 0 + S + D\rangle}{1 + g_s \langle S|S\rangle + g_D \langle D|D\rangle} \tag{7.125}$$

with g_S and g_D being the statistical factors $\frac{4}{n}$, $\frac{2}{n}$, $\frac{4n-6}{n(n-1)}$, as appropriate for the given method. Thus, unlike the CEPA models, the averaged models fulfill a stationarity principle and are unitarily invariant. However, if one thinks about localized internal MOs, it appears evident that the approximation of equal pair energies must be one of rather limited validity and that a more detailed treatment of the electron pairs is warranted. Maintaining a stationarity principle while providing a treatment of the pairs that closely resembles that of the CEPA methods was achieved by Ahlrichs and co-workers in an ingenious way with the development of the CPF method [16]. In this method, the correlation energy functional is written as:

$$E_C = 2\sum_{Px} \frac{\langle \Phi_p^x | H | 0 \rangle}{N_P} + \sum_{PQxy} \frac{\left\langle \Phi_p^x | H - E_0 | \Phi_Q^y \right\rangle}{\sqrt{N_P N_Q}} \tag{7.126}$$

with

$$N_P = 1 + \sum_Q T_{PQ} \sum_y (C_y^Q)^2 \tag{7.127}$$

The topological matrix for pairs P =(ij) and $Q = (kl)$ is chosen as: [440]

$$T_{PQ} = \frac{\delta_{ik} + \delta_{il}}{2n_i} + \frac{\delta_{jk} + \delta_{jl}}{2n_j} \tag{7.128}$$

with n_i being the number of electrons in orbital i in the reference determinant. The singles out of orbital i are formally equated with $P = (ii)$. At the spin-orbital level, $n_i = 1$, for closed shells $n_i = 2$. Using the same topological matrix in $\Delta_P = \sum_Q T_{PQ}\varepsilon_Q$ one recovers the CEPA/1 shifts for the doubles in eq. (7.124). It is straightforward to obtain the CPF equivalents of the other CEPA models by adjusting the T_{PQ} matrix appropriately. In our program, we have done so and we refer below to these methods as CPF/1, CPF/2 and CPF/3 in analogy to the CEPA models (CPF/1 ≡CPF). In fact, as discussed by Ahlrichs and co-workers, variation of the CPF-functional leads to equations that very closely resemble the CEPA equation and can be readily implemented along the same lines as a simple modification of a CISD program. Ahlrichs *et al.* argued that the energies of CEPA/1 and CPF/1 should be very close. We have independently confirmed that in the majority of cases, the total energies predicted by the two methods differ by less than 0.1 mEh.

An alternative to the CPF approach which is also based on variational optimization of an energy functional is the VCEPA method [455]. The equations resulting from application of the variational principle to the VCEPA functional are even closer to the CEPA equations than for CPF so that the resulting energies are practically indistinguishable from the corresponding CEPA values. The VCEPA variants are referred to as VCEPA/1, VCEPA/2, and VCEPA/3 in analogy to CEPA and CPF. A strictly size extensive energy functional (SEOI) which is invariant with respect to unitary transformations within the occupied and virtual orbital subspaces is also available [456] (an open-shell version is not implemented yet).

Again, a somewhat critical point concerns the single excitations. They do not account for a large fraction of the correlation energy. However, large coefficients of the single excitations lead to instability and deterioration of the results. Secondly, linear response properties are highly dependent on the effective energies of the singles and their balanced treatment is therefore important. Since the CEPA and CPF methods amount to shifting down the diagonal energies of the singles and doubles, instabilities are expected if the effective energy of an excitation approaches the reference energy of even falls below it. In the CPF method this would show up as denominators N_P that are too small. The argument that the CPF denominators are too small has led Chong and Langhoff to the proposal of the MCPF method which uses a slightly more elaborate averaging than $(N_P N_Q)^{1/2}$ [173].[1] However, their modification was solely based on numerical arguments rather than physical or mathematical reasoning. In the light of Eq. (7.124) and the performance of the NACPF, it appears to us that for the singles one should use twice the T_{PQ} proposed by Ahlrichs and co-workers. The topological matrix T_{PQ} is modified in the following way for the (very slightly) modified method to which we refer to as NCPF/1:

$$T_{ij,kl} = \frac{\delta_{ik} + \delta_{il}}{2n_i} + \frac{\delta_{jk} + \delta_{jl}}{2n_j} \tag{7.129}$$

$$T_{ij,k} = 0 \tag{7.130}$$

[1] This method – although it has been rather extensively used in the past – is not implemented in ORCA. We recommend to use our NCPF/1 instead.

$$T_{i,kl} = 2\frac{\delta_{ik} + \delta_{il}}{n_i} \tag{7.131}$$

$$T_{i,k} = 0 \tag{7.132}$$

(note that $T_{PQ} \neq T_{QP}$ for this choice). Thus, the effect of the singles on the doubles is set to zero based on the analysis of the CEPA approximations and the effect of the singles on the singles is also set to zero. This is a sensible choice since the product of two single excitations is a double excitation which is already included in the SD space and thus none of them can belong to the outer space. It is straightforward to adapt this reasoning about the single excitations to the CEPA versions as well as to NCPF/2 and NCPF/3.

The aforementioned ambiguities arising from the use of single excitations in coupled-pair methods can be avoided by using correlation-adapted orbitals instead of Hartree-Fock orbitals thus eliminating the single excitations. There are two alternatives: (a) Brueckner orbitals and (b) optimized orbitals obtained from the variational optimization of the electronic energy with respect to the orbitals. Both approaches have already been used for the coupled-cluster doubles (CCD) method [360, 776] and later been extended to coupled-pair methods [454]. In the case of CCD, orbital optimization requires the solution of so-called Λ (or Z vector) equations [744]. There is, however, a cheaper alternative approximating the Z vector by a simple analytical formula [457].

Furthermore, the parametrized coupled-cluster (pCCSD) method of Huntington and Nooijen [405], which combines the accuracy of coupled-pair type methods for (usually superior to CCSD, at least for energies and energy differences) with the higher stability of the coupled-cluster methods, is an attractive alternative. Comprehensive numerical tests [404] indicate that particularly pCCSD(-1,1,1) (or pCCSD/1a) and pCCSD (-1.5,1,1) (or pCCSD/2a) have great potential for accurate computational thermochemistry. These methods can be employed by adding the "simple" keywords `pCCSD/1a` or `pCCSD/2a` to the first line of input. As mentioned in section *Local Coupled Pair and Coupled-Cluster Calculations*, the LPNO variants of the pCCSD methods are also available for RHF and UHF references via usage of the simple keywords `LPNO-pCCSD/1a` and `LPNO-pCCSD/2a`.

7.14.2 Closed-Shell Equations

Proceeding from spin-orbitals to the spatial orbitals of a closed-shell determinant leads to the actual working equations of this work. Saebo, Meyer and Pulay have exploited the generator state formalism to arrive at a set of highly efficient equations for the CISD problem [704]. A similar set of matrix formulated equations for the CCSD and QCISD cases has been discussed by Werner and co-workers [355] and the MOLPRO implementation is widely recognized to be particularly efficient. Equivalent explicit equations for the CISD and CCSD methods were published by Scuseria *et al.* [777][2] The doubles equations for the residual "vector" σ are ($i \leqslant j$, all a, b):

$$\begin{aligned}
\sigma_{ab}^{ij} = K_{ab}^{ij} &+ K\left(\mathbf{C}^{ij}\right)_{ab} + \left\{\mathbf{F}^V\mathbf{C}^{ij} + \mathbf{C}^{ij}\mathbf{F}^V\right\}_{ab} - \sum_k \left\{F_{jk}C_{ab}^{ik} + F_{ik}C_{ab}^{kj}\right\} + \sum_{kl} K_{kl}^{ij}C_{ab}^{kl} \\
&+ \sum_k \left\{\left(2\mathbf{C}^{ik} - \mathbf{C}^{ik+}\right)\left(\mathbf{K}^{kj} - \tfrac{1}{2}\mathbf{J}^{kj}\right) + \left(\mathbf{K}^{ik} - \tfrac{1}{2}\mathbf{J}^{ik}\right)\left(2\mathbf{C}^{kj} - \mathbf{C}^{kj+}\right)\right\}_{ab} \\
&- \sum_k \left\{\tfrac{1}{2}\mathbf{C}^{ik+}\mathbf{J}^{jk+} + \tfrac{1}{2}\mathbf{J}^{ik}\mathbf{C}^{kj+} + \mathbf{J}^{jk}\mathbf{C}^{ik} + \mathbf{C}^{kj}\mathbf{J}^{ik+}\right\}_{ab} \\
&+ C_a^i F_b^j + C_b^j F_a^i - \sum_k \left\{K_{ka}^{ji}C_b^k + K_{kb}^{ij}C_a^k\right\} + \left\{\mathbf{K}^{ia}\mathbf{C}^j + \mathbf{K}^{ja}\mathbf{C}^i\right\}_b \\
&- \Delta^{ij}C_{ab}^{ij}
\end{aligned} \tag{7.133}$$

The singles equations are:

$$\begin{aligned}
\sigma_a^i = F_a^i &+ \left\{\mathbf{F}^V\mathbf{C}^i\right\}_a - \sum_j F_{ij}C_a^j - \sum_{jkb}\left(2K_{jb}^{ik} - J_{jb}^{ik}\right)C_{ba}^{kj} \\
&+ \sum_j \left\{\left(2\mathbf{K}^{ij} - \mathbf{J}^{ij}\right)\mathbf{C}^j + \mathbf{F}^j\left(2\mathbf{C}^{ij+} - \mathbf{C}^{ij}\right) + \left\langle\left(2\mathbf{K}^{ia} - \mathbf{K}^{ia+}\right)\mathbf{C}^{ij+}\right\rangle\right\}_a \\
&- \Delta^i C_a^i
\end{aligned} \tag{7.134}$$

The following definitions apply:

$$K\left(\mathbf{C}^{ij}\right)_{ab} = \sum_{cd}(ac|bd)\,C_{cd}^{ij} \tag{7.135}$$

$$K_{rs}^{pq} = (pr|qs) \tag{7.136}$$

[2] Our coupled-cluster implementation is largely based on this nice paper. The equations there have been extensively verified to be correct.

$$J_{rs}^{pq} = (pq|rs) \tag{7.137}$$

$$\langle \mathbf{AB} \rangle = \sum_{pq} A_{pq} B_{qp} \tag{7.138}$$

The two-electron integrals are written in (11|12) notation and $\mathbf{F}$ is the closed-shell Fock operator with $\mathbf{F}^V$ being its virtual sub-block. We do not assume the validity of Brillouin's theorem. The amplitudes C_a^i, C_{ab}^{ij} have been collected in vectors $\mathbf{C}^i$ and matrices $\mathbf{C}^{ij}$ wherever appropriate. The shifts Δ^i and Δ^{ij} are dependent on the method used and are defined in Table Table 7.18 for each method implemented in ORCA.

Table 7.18: Summary of the diagonal shifts used in various singles- and doubles methods discussed in this chapter. The quantities ε_i and ε_{ij} are the correlation energy increments brought about by the single- and the double excitations respectively. The partial denominators for the CPF type methods N_i and N_{ij} are specified in eq. (7.127).

Method	Doubles Shift	Singles Shift
CISD	E_{C}	$E_{\mathbf{C}}$
CEPA/0	0	0
CEPA/1	$\frac{1}{2}(\varepsilon_i + \varepsilon_j) + \frac{1}{4}\sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$\frac{1}{2}\varepsilon_{ii} + \frac{1}{2}\sum_k \varepsilon_{ik}$
CEPA/2	$\delta_{ij}\varepsilon_i + \varepsilon_{ij}$	$\varepsilon_i + \varepsilon_{ii}$
CEPA/3	$(\varepsilon_i + \varepsilon_j) - \delta_{ij}\varepsilon_i - \varepsilon_{ij} + \frac{1}{2}\sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$\varepsilon_i + \sum_k \varepsilon_{ik}$
NCEPA/1	$\frac{1}{4}\sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$\varepsilon_{ii} + \sum_k \varepsilon_{ik}$
NCEPA/2	ε_{ij}	$2\varepsilon_{ii}$
NCEPA/3	$-\varepsilon_{ij} + \frac{1}{2}\sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$2\sum_k \varepsilon_{ik}$
CPF/1	$N_{ij}\left\{\frac{1}{2}(\frac{\varepsilon_i}{N_i} + \frac{\varepsilon_j}{N_j}) + \frac{1}{4}\sum_k (\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}})\right\}$	$N_i\left\{\frac{1}{2}\frac{\varepsilon_{ii}}{N_{ii}} + \frac{1}{2}\sum_k \frac{\varepsilon_{ik}}{N_{ik}}\right\}$
CPF/2	$N_{ij}\left\{\delta_{ij}\frac{\varepsilon_i}{N_i} + \frac{\varepsilon_{ij}}{N_{ij}}\right\}$	$N_i\left\{\frac{\varepsilon_i}{N_i} + \frac{\varepsilon_{ii}}{N_{ii}}\right\}$
CPF/3	$N_{ij}\left\{\frac{\varepsilon_i}{N_i}(1-\delta_{ij}) + \frac{\varepsilon_j}{N_j} - \frac{\varepsilon_{ij}}{N_{ij}} + \frac{1}{2}\sum_k (\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}})\right\}$	$N_i\left\{\frac{\varepsilon_i}{N_i} + \sum_k \frac{\varepsilon_{ik}}{N_{ik}}\right\}$
NCPF/1	$\frac{1}{4}N_{ij}\sum_k (\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}})$	$N_i\left\{\frac{\varepsilon_{ii}}{N_{ii}} + \sum_k \frac{\varepsilon_{ik}}{N_{ik}}\right\}$
NCPF/2	$N_{ij}\frac{\varepsilon_{ij}}{N_{ij}}$	$2N_i\frac{\varepsilon_{ii}}{N_{ii}}$
NCPF/3	$N_{ij}\left\{-\frac{\varepsilon_{ij}}{N_{ij}} + \frac{1}{2}\sum_k (\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}})\right\}$	$2N_i\sum_k \frac{\varepsilon_{ik}}{N_{ik}}$
ACPF	$\frac{2}{n}E_{\mathrm{C}}$	$\frac{2}{n}E_{\mathrm{C}}$
ACPF/2	$\frac{2}{n}E_{\mathrm{C}}$	$\left[1 - \frac{(n-3)(n-2)}{n(n-1)}\right]E_{\mathrm{C}}$
NACPF	$\frac{2}{n}E_{\mathrm{C}}$	$\frac{4}{n}E_{\mathrm{C}}$
AQCC	$\left[1 - \frac{(n-3)(n-2)}{n(n-1)}\right]E_{\mathrm{C}}$	$\left[1 - \frac{(n-3)(n-2)}{n(n-1)}\right]E_{\mathrm{C}}$

The QCISD method requires some slight modifications. We found it most convenient to think about the effect of the nonlinear terms as a "dressing" of the integrals occurring in equations (7.133) and (7.134). This attitude is close to the recent arguments of Heully and Malrieu and may even open interesting new routes towards the calculation of excited states and the incorporation of connected triple excitations.[391] The dressed integrals are given by:

$$\bar{F}_{ik} = F_{ik} + \sum_l \left\langle \mathbf{C}^{il}\left(2\mathbf{K}^{kl} - \mathbf{K}^{kl+}\right)\right\rangle \tag{7.139}$$

$$\bar{F}_{ab} = F_{ab} - \sum_{kl} \left\{\mathbf{C}^{kl}\left(2\mathbf{K}^{kl} - \mathbf{K}^{kl+}\right)\right\}_{ab} \tag{7.140}$$

$$\bar{F}_{kc} = F_{kc} + \sum_l \left(2\mathbf{K}^{kl} - \mathbf{K}^{kl+}\right)\mathbf{C}^l \tag{7.141}$$

$$\bar{K}_{kl}^{ij} = K_{kl}^{ij} + \left\langle \mathbf{K}^{kl}\mathbf{T}^{kl+}\right\rangle \tag{7.142}$$

$$\bar{K}_{ab}^{ij} = K_{ab}^{ij} + \sum_k \left\{\mathbf{C}^{ik}\left(\mathbf{K}^{kj} - \frac{1}{2}\mathbf{K}^{jk}\right) + \mathbf{C}^{ki}\mathbf{K}^{kj}\right\}_{ab} \tag{7.143}$$

$$\bar{J}_{ab}^{ij} = J_{ab}^{ij} + \sum_{k} \left\{ \mathbf{C}^{ki} \mathbf{K}^{jk} \right\}_{ab} \tag{7.144}$$

The CCSD method can be written in a similar way but requires 15 additional terms that we do not document here. They may be taken conveniently from our paper about the LPNO-CCSD method [616].

A somewhat subtle point concerns the definition of the shifts in making the transition from spin-orbitals to spatial orbitals. For example, the CEPA/2 shift becomes in the generator state formalism:

$$-\left\langle \tilde{\Phi}_{ij}^{ab} | \Delta^{ij} | \Psi \right\rangle = C_{ab}^{ij} \left(\frac{1}{3} \varepsilon_{ij}^{\alpha\alpha} + \frac{2}{3} \varepsilon_{ij}^{\alpha\beta} \right) + C_{ba}^{ij} \left(-\frac{1}{3} \varepsilon_{ij}^{\alpha\alpha} + \frac{1}{3} \varepsilon_{ij}^{\alpha\beta} \right) \tag{7.145}$$

($\tilde{\Phi}_{ij}^{ab}$ is a contravariant configuration state function, see Pulay *et al.* [744]. The parallel and antiparallel spin pair energies are given by:

$$\varepsilon_{ij}^{\alpha\alpha} = \frac{1}{2} \sum_{ab} \left[K_{ab}^{ij} - K_{ba}^{ij} \right] \left(C_{ab}^{ij} - C_{ba}^{ij} \right) \tag{7.146}$$

$$\varepsilon_{ij}^{\alpha\beta} = \frac{1}{2} \sum_{ab} K_{ab}^{ij} C_{ab}^{ij} \tag{7.147}$$

This formulation would maintain the exact equivalence of an orbital and a spin-orbital based code. Only in the (unrealistic) case that the parallel and antiparallel pair correlation energies are equal the CEPA/2 shift of Table 7.18 arise. However, we have not found it possible to maintain the same equivalence for the CPF method since the electron pairs defined by the generator state formalism are a combination of parallel and antiparallel spin pairs. In order to maintain the maximum degree of internal consistency we have therefore decided to follow the proposal of Ahlrichs and co-workers and use the topological matrix T_{PQ} in equation (7.128) and the equivalents thereof in the CEPA and CPF methods that we have programmed.

7.14.3 Open-Shell Equations

We have used a non-redundant set of three spin cases ($\alpha\alpha$, $\beta\beta$, $\alpha\beta$) for which the doubles amplitudes are optimized separately. The equations in the spin-unrestricted formalism are straightforwardly obtained from the corresponding spin orbital equations by integrating out the spin. For implementing the unrestricted QCISD and CCSD method, we applied the same strategy (dressed integrals) as in the spin-restricted case. The resulting equations are quite cumbersome and will not be shown here explicitly [361].

Note that the definitions of the spin-unrestricted CEPA shifts differ from those of the spin-restricted formalism described above (see Kollmar et al. [361]). Therefore, except for CEPA/1 and VCEPA/1 (and of course CEPA/0), for which the spin-adaptation of the shift can be done in a consistent way, CEPA calculations of closed-shell molecules yield slightly different energies for the spin-restricted and spin-unrestricted versions. Since variant 1 is also the most accurate among the various CEPA variants [617], we recommend to use variant 1 for coupled-pair type calculations. For the variants 2 and 3, reaction energies of reactions involving closed-shell and open-shell molecules simultaneously should be calculated using the spin-unrestricted versions only.

A subtle point for open-shell correlation methods is the choice of the reference determinant [618]. Single reference correlation methods only yield reliable results if the reference determinant already provides a good description of the systems electronic structure. However, an UHF reference wavefunction suffers from spin-contamination which can spoil the results and lead to convergence problems. This can be avoided if quasi-restricted orbitals (QROs) are used [391, 611] since the corresponding zeroth-order wavefunction is an eigenfunction of the $\hat{S}^2$ operator and thus, no severe spin-contamination will appear. The coupled-pair and coupled-cluster equations will be still solved in a spin-unrestricted formalism but the energy will be slightly higher compared to the results obtained with a spin-polarized UHF reference determinant. Furthermore, especially for more difficult systems like e.g. transition metal complexes, it is often advantageous to use Kohn-Sham (KS) orbitals instead of HF orbitals.

7.14.4 Local correlation

As described in previous sections of the manual, ORCA features the extremely powerful LPNO and DLPNO methods. "LPNO" stands for "local pair natural orbital" approximation and DLPNO for "domain based LPNO". These methods are designed to provide results as close as possible to the canonical coupled-cluster results while gaining orders of magnitude of efficiency through a series of well-controlled approximations. In fact, typically about 99.8% to 99.9% of the canonical correlation energy is recovered in such calculations. Even higher accuracy is achievable but will ultimately come at much higher computational cost. The default cut-offs are set such that the vast majority of chemically meaningful energy differences are reproduced to better than 1 kcal/mol relative to the canonical results with the same basis set. Of the LPNO and DLPNO methods, the LPNO is the older one and will eventually be discarded from ORCA. It has some higher order scaling steps (up to N^5) while DLPNO is linear scaling and of similar accuracy. Amongst the DLPNO methods, the first generation implementation of the DLPNO methods (DLPNO2013) is only near-linear scaling, while the DLPNO implementation since ORCA 4.0 is linear scaling.

It is important to understand that the LPNO and DLPNO implementations are intimately tied to the RI approximation. Hence, in these calculations one *must* specify a fitting basis set. The same rules as for RI-MP2 apply: the auxiliary basis sets optimized for MP2 are just fine for PNO calculations. You can specify several aux bases for the same job and the program will sort it out correctly.

The theory of the LPNO methods has been thoroughly described in the literature in a number of original research papers.[616, 622, 720, 722]

Hence, it is sufficient to only point to a few significant design principles and features of these methods:

1. The correlation energy of any molecule can be written as a sum over the correlation energies of pairs of electrons, labelled by the internal indices (ij) of pairs of orbitals that are occupied in the reference determinant. If the orbitals (i) and (j) are localized, the pair correlation energy ϵ_{ij} falls off very quickly with distance, quite typically by about an order of magnitude per chemical bond that is separating orbitals (i) and (j). Hence, by using a suitable cut-off for a reasonable pair correlation energy estimate many electron pairs can be removed from the high-level treatment and only a linear scaling number of electron pairs will make a significant contribution to the correlation energy.

2. The natural estimate for the pair correlation energy comes from second order many body perturbation theory (MP2). However, canonical MP2 is scaling with the fifth power of the molecular size and hence, is not really attractive from a theoretical nor computational point of view. Owing to the small pre-factor RI-MP2 goes a long way to provide reasonably cheap estimates for pair correlation energies. However, if one uses localized internal orbitals, then the MP2 energy expression must be cast in form of linear equations. On the other hand, if one uses canonical virtual orbitals together with localized internal orbitals and neglects the coupling terms coming from purely internal Fock matrix elements F(i,k) and F(k,j) then one ends up with a fair approximation that is termed "semi-canonical MP2" in ORCA. It serves as a guess in the older LPNO method. For DLPNO this method is also not attractive.

3. In DLPNO, the guess is more sophisticated. Here the virtual space is spanned by projected atomic orbitals (PAOs) that are assigned to domains of atoms that are associated with each local internal orbital (i) and with the union of two such domains (i) and (j) for the electron pair (ij). If one applies the semi-local approximation, one obtains an excellent approximation to the semi-canonical MP2 energy. This is called the "semi-local" approximation and it scales linearly with respect to computational effort. If one iterates these equations to self-consistency to eliminate the coupling terms F(i,k) and F(k,j) then one obtains the full local MP2 method (LMP2 or L-MP2). By making the domains large enough the results approach the canonical MP2 energy to arbitrary accuracy while still being linear scaling with respect to computational resources. This method is the default for the DLPNO method.

4. Basically, the high-spin open-shell version of the DLPNO uses the same strategy as the closed-shell variant to efficiently generate the open-shell PNOs in a consistent manner to the closed-shell formalism. Through the development of the UHF-LPNO-CCSD method, we have realized that use of the unrestricted MP2 (UMP2) approach to define the open-shell PNOs introduces a few complexities; (1) the PNOs for β spin orbitals cannot be defined for α-α electron pairs and vice versa, (2) the diagonal PNOs for singly occupied orbitals cannot be properly defined, and (3) the PNO space does not become identical to that in the closed-shell LPNO framework in the closed-shell limit. However, to program all the unrestricted CCSD terms in the LPNO basis, those PNOs are certainly necessary. Therefore, in the UHF-LPNO-CCSD implementation, several terms, which, in many cases, give rather minor contributions in the correlation energy are omitted. Due to these facts, the UHF-LPNO-CCSD does not give identical results to that of RHF-LPNO-CCSD in

the closed-shell limit. In addition, screening of the weak pairs on the basis of the semi-canonical UMP2 pair-energy results in somewhat unbalanced treatment of the closed- and open-shell states in some cases, leading to rather larger errors in the reaction energies. To overcome those issues, in the high-spin open-shell DLPNO-CCSD method, the PNOs are generated in the framework of semi-canonical NEVPT2 which smoothly converges to the RHF-MP2 counterpart in the closed-shell limit. The open-shell DLPNO-CCSD, which is made available from ORCA 4.0, includes a full set of the open-shell CCSD equations and is designed as a natural extension to the RHF-DLPNO-CCSD.

5. Screening of the electron pairs according to a truncation parameter (in ORCA it is called T_{CutPairs}) is not sufficient to obtain a highly performing local electronic structure method. The original work of Pulay suggested to limit excitations out of the internal orbitals (i) and (j) to the domain associated with the pair (ij). While this works well and has been implemented to perfection by Werner, Schütz and co-workers over the years,[754, 755, 774, 775] the pre-factor of such calculations is high, since the domains have to be chosen large in order to recover 99.9% or more of the canonical correlation energy.

6. The ORCA developers have therefore turned to an approach that has been used with a high degree of success in the early 1970s by Meyer, Kutzelnigg, Staemmler and their co-workers, namely the method of "pair natural orbitals" (PNOs).[10, 582, 583, 886]
   As shown by Loewdin in his seminal paper from 1955, natural orbitals (the eigenfunctions of the one-particle density matrix) provide the fastest possible convergence of the correlated wavefunction with respect to the number of one-particle functions included in the virtual space. It has been amply established that approximate natural orbitals are almost as succesful as the true natural orbitals (which would require the knowledge of the exact wavefunction) in this respect. While the success of approximate correlation treatments of many particle systems that use approximate natural orbitals of the whole systems are somewhat limited, this is not the case for pair natural orbitals. The latter have first been suggested as a basis for correlation calculations by England and co-workers and, at the time, were given the name "pseudonatural orbitals", a term that was used by Meyer throughout his pioneering work.

7. The PNOs are approximate natural orbitals of a given electron pair. In order to generate them one requires a one particle pair density matrix D_{ij} the eigenfunctions of which are the PNOs themselves while the corresponding eigenvalues are the PNO occupation numbers. While there are many creative possibilities that can lead to slightly different PNO sets, a quite useful and natural approximation is to generate such a density from the MP2 amplitudes as an expectation value (the "unrelaxed" MP2 density. One then uses a second threshold (in ORCAT_{CutPNO}) that controls the PNOs to be included in the calculation. PNOs with an occupation number $< T_{\mathrm{CutPNO}}$ are neglected.

8. PNOs of a given electron pair form an orthonormal set. However, PNOs belonging to different electron pairs are not orthonormal and hence they overlap. This non-orthogonality leads to surprisingly few complications because the PNOs stay orthogonal to all occupied orbitals. In practice, the equations for PNO-based correlation calculations are hardly more complex than the canonical equations.

9. The nice feature of these pair densities is that they become small when the pair interaction becomes small. Hence weak pairs are correlated by very few PNOs. Therefore, the PNO expansion of the wavefunction is extremely compact and there only is a linear scaling number of significant excitation amplitudes that need to be considered.

10. A great feature of the PNOs is that they are "self-adapting" to the correlation situation that they are supposed to describe. Hence, they are as delocalized as required by the physics and there is no ad-hoc assumption about their location in space. However, it is clear that the PNOs are located in the same region of space as the internal orbitals that they correlate because otherwise they would not contribute to the correlation energy.

11. In the iterative local MP2, a set of PNOs is calculated for the MP2 calculation from the semicanonical amplitudes first using the cutoff `TCutPNO` $\times$`LMP2ScaleTCutPNO`. The size of the resulting PNO space should be comparable with DLPNO-MP2. After the iterations have converged, a (smaller) set of PNOs for the CCSD double excitation amplitudes is generated from the iterated local MP2 amplitudes in the (larger) PNO basis. The PNOs for the single excitation amplitudes are always calculated using the semicanonical MP2 amplitudes, as they require a much more conservative PNO truncation threshold.

12. Capitalizing on this feature one can define generous domains and expand the PNOs in terms of the PAOs and auxiliary fit functions (for the RI approximations) that are contained in these domains. In ORCA this is controlled by the third significant truncation parameter T_{CutMKN}. This is the basis of the DLPNO method. In the older LPNO method, the PNOs are expanded in terms of the canonical virtual orbitals and T_{CutMKN}

is only used for a local fit to the PNOs. In the linear-scaling DLPNO implementation the domains expanding the PNOs in terms of the PAOs are controlled via T_{CutDO}.

13. PNO expansions have the amazing advantage that the PNOs converge to a well-defined set as the basis set is approaching completeness. Hence, the increase in the number of PNOs per electron pair is very small upon enlargement of the orbital basis. In other words, correlation calculations with large basis sets are not that much more expensive than correlation calculations with small basis sets. Thus, the advantage of PNO over canonical calculations increases with the size of the basis set. This is a great feature as large and flexible basis sets are a requirement for meaningful correlation calculations.

14. In summary, DLPNO and LPNO calculations are controlled by only three cut-off parameters with well-defined meanings: a) T_{CutPairs}, the cut-off for the electron pairs to be included in the coupled-pair or coupled-cluster iterations, b) T_{CutPNO} which controls how many PNOs are retained for a given electron pair and c) T_{CutMKN} that controls how large the domains are that the PNOs expand over. For the linear-scaling DLPNO calculations the domain sizes are controlled via T_{CutDO}.

15. It is clear that owing to the truncations a certain amount of error is introduced in the results. However, having the LMP2 results available, one can compensate for the errors coming from T_{CutPairs} and T_{CutPNO}. This is done in ORCA and the correction is added to the final correlation energy, thus bringing it very close (to mEh accuracy or better) to the canonical result. T_{CutMKN} is best dealt with by making it conservative (at 1e$-$3 to 1e$-$4 the domains are about 20–30 atoms large, which is sufficient for an accurate treatment).

16. Note that the LPNO and DLPNO methods do not introduce any real space cut-offs. We refrain from doing so and insist in our method development by making all truncations based on wavefunction or energy parameters. We feel that this is the most unbiased approach and it involves no element of "chemical intuition" or "prejudice".

17. In the DLPNO method a highly efficient screening mechanism is operative. In this method one first obtains a (quadratically scaling) multipole estimate for the pair correlation energy that is extremely fast to compute (a few seconds, even for entire proteins). Only if this estimate is large enough, a given electron pair is even considered for a LMP2 treatment. Quite typically pairs with energy contributions of 1e$-$6 Eh and smaller are very well described by the dipole-dipole estimate. Specifically, we drop pairs with estimated energies $< 0.01 \times T_{\mathrm{CutPairs}}$ and add their multipole energy sum to the final correlation energy. These corrections tend to be extremely small, even for large molecules and are insignificant for the energy. However, importantly, the multipole estimate ensures linear scaling in the MP2 treatment. The pairs that then do not survive the pair-prescreening are called "weak pairs" in the ORCA or DLPNO sense. They still play a role in the calculation of the triple excitation correction.

18. The calculation of triple excitation contributions is more involved and one does not have a perturbative estimate available since the (T) contribution is perturbative itself. While the (T) contribution is much smaller than the CCSD correlation energy, the errors introduced by the various local and PNO approximations can be significant. We found that one has to include triples with at least one pair being a "weak" LMP2 pair (with its LMP2 amplitudes) in order to arrive at sufficiently accurate results.

Given these explanations the various cut-off parameters that can be controlled in LPNO and DLPNO calculations should be understandable and are listed below. We emphasize again that only the three thresholds T_{CutPairs}, T_{CutPNO} and T_{CutMKN} should be touched by the user, unless very specific questions are addressed. The recommended way to control the accuracy of calculations is to specify "TightPNO", "NormalPNO" or "LoosePNO" keywords, rather than to change numeric values of cutoffs. Individual thresholds should normally not be changed, as the defaults are sensible and lead to good cost/performance ratios.

```
%mdci TCutPairs        1e-4   # cut-off for the pair truncation
      TCutPNO       3.33e-7   # cut-off for the PNO truncation
      TCutDO           1e-2   # cut-off for the DLPNO domain construction
      TCutMKN          1e-3   # cut-off for the local fit
                              # for DLPNO2013: also domain construction
                              # remaining options, tied to the three main cut-offs,
                                EXPERTS ONLY!
      Localize         true   # flag for using localized orbitals
      LocMet           AHFB   # Localization method.
                              # Options: PM, FB, IAOIBO, IAOBOYS, NEWBOYS, AHFB
      LocTol           1e-6   # Absolute threshold for the localization procedure
```

(continues on next page)

(continued from previous page)

```
                              # Automatically adjusted by default.
     LocTolRel        1e-8    # Relative threshold for the localization procedure
     LocMaxIter        128    # Maximum number of localization iterations
     LocRandom           1    # default, take random seed for any localization
                              # For internal orbitals: choose best of 32 localizations
                              # Switched off for AHFB
                         0    # take constant seed for any localization (for testing)
     LocNAttempts       np    # number of localization attempts
                              # default: number of processes, minimum 8, if
                              #  randomize true
                              # 1, if randomize false
                              # any number larger or equal np, if randomize true

     PNONorm       MP2Norm    # default, old IEPANorm can also be used
                                (near identical results)
     NrMP2Pairs_Trip     1    # number of MP2 pairs to be included in the triples
                                calculation
     PAOOverlapThresh 1e-8    # generation of non-redundant PAOs from redundant ones
     UseFullLMP2Guess false   # Use iterative full LMP2 (for DLPNO)
     SinglesFockUsePNOs true  # compute the Singles Fock matrix (SFM) in PNOs.
                              # DLPNO2013: default for SinglesFockUsePNOs is false,
                              #    by default RIJCOSX is used for the SFM, except when
                              #    RIJK is given. In that case the RIJK-SFM is used.
     LMP2MaxIter         50   # max no of iterations in the MP2 equations
     LMP2TolE          1e-7   # LMP2 energy convergence tolerance
     LMP2TolR          5e-7   # LMP2 residual convergence tolerance
     LMP2ScaleTCutPNO         # PNO cutoff for LMP2 is: TCutPNO*LMP2ScaleTCutPNO
                              # Default: TCutPNO(DLPNO-MP2)/TCutPNO(DLPNO-CCSD) with
                              # respective TCutPNOs specific to Loose/Normal/TightPNO
     LMP2FCut          1e-5   # LMP2 neglect cut-off for
                                off-diagonal Fock matrix elements
     TCutTNO              0   # Cut-off for triples natural orbitals (0=automatic)
     TCutPNOSingles      -1   # -1= use 0.03*TCutPNO
     TCutPreScr          -1   # -1= use 0.01*TCutPairs for
                                multipole estimate based screening
     TCutDeloc          0.1   # delocalization threshold for specification
                                of extended domains.
                                Necessary because PAOs are not strictly localized
     TCutOSV           1e-6   # orbital specific virtuals used for pre-screening.
                                No critical
     TScaleMP2Pairs    0.1
     TScaleMKNStrong    10
     TScaleMKNWeak     100
```

For larger systems and tighter thresholds the disk I/O of a DLPNO calculation may become challenging. In this case, it might be usefull to keep some integrals in memory, if enough RAM is available. With the following flag

```
%mdci
     StorageType Shared
end
```

ORCA will try to store certain much-used integrals in shared memory. If the amount of memory is not sufficient, ORCA will fall back to on-disk storage. **NOTE:** This flag will only work if all processes work on the same node.

IMPORTANT NOTE REGARDING ORBITAL LOCALIZATION

- **Localized orbitals for DLPNO are obtained via the Foster-Boys method with an augmented Hessian converger (`AHFB`) by default.**
- The localization tolerance (`LocTol`) is coupled to the SCF gradient tolerance (`TolG`) with a constant factor by default. Selecting specific SCF convergence settings (such as `TightSCF`) therefore also ensures obtaining a set of appropriately well converged localized orbitals. This can be overridden by setting a different value

for LocTol.

- An important feature of the augmented Hessian converger is that it systematically approaches a local maximum of the localization function (even though convergence to the global maximum cannot be guaranteed). As opposed to that, the conventional localization method (FB) may stop, for example, at a saddle point. In bad cases, this can lead to deviations of several kJ/mol in the DLPNO energy. Likewise, it can contribute towards lack of reproducibility of results.
- No similar procedure has been implemented for the other localization methods (such as Pipek-Mezey) yet. The same problems as with the FB converger can occur in these cases.
- No randomization is used for the AHFB converger.

The old default orbital localization settings of ORCA 4.0 can be reproduced with the following options:

```
%MDCI
  LocMet    FB
  LocTol    1.0e-6
  LocRandom 1
End
```

Regarding the methods that employ randomization (FB, PM, IAOIBO, IAOBOYS) only, the following notes apply:

- Generally, better DLPNO results are obtained when several runs of localization are undertaken using different initial guesses. The different initial guesses are obtained using randomization (LocRandom).
- However, randomization of the initial guess can lead to differently localized MOs in different calculations. This can yield non-identical correlation energies, varying in the sub-kJ/mol range, for different runs on the same machine.
- In order to yield identical correlation energy results, randomization can be switched off (LocRandom 0). However, switching off randomization only leads to identical results on the same machine, but can still lead to slightly different results (in the sub-kJ/mol range) on different machines.
- Reproducibility of the correlation energy is expected to increase further if LocNAttempts is set to higher values.

The input below shows how to perform a DLPNO calculation with settings that exactly reproduce the canonical RI-MP2 result. They are not recommended for production use, but merely to show that if the local approximations are pushed, then the result coincides with the canonical one. If one would set T_{CutPNO} to zero this would give canonical RI-CCSD. However, this is an absurdly inefficient calculation and hence not done.

```
#
! def2-SV(P) def2/JK def2-SVP/C RI-JK DLPNO-CCSD VeryTightSCF RI-MP2

# obtain a result that only contains errors from the PNO approximation
# but no others
%mdci TCutPairs        0
      TCutMKN          0
      UseFullLMP2Guess true
      LMP2FCut         1e-9
      LMP2MaxIter      25
      LMP2TolE         1e-10
      LMP2TolR         1e-11
      PAOOVerlapThresh 1e-9
      end

! Bohrs
* xyz 0 1
   C  -1.505246952209632    1.048213673267046   -3.005665895986369
   C   1.289678561934891    0.246429688933291   -3.259735682020124
   C   2.834670835163566    1.157307360133605   -0.990383454919828
   C   1.924119415395082   -0.128330938291771    1.465070676514038
   C  -0.931529472233802   -0.722841293992075    1.397639867298547
   C  -2.347670084056626    1.213332291655600   -0.217984867773136
```

(continues on next page)

(continued from previous page)

```
   H   2.084955694093313    0.973408301535989   -5.037750251258102
   H   1.426532559234904   -1.831017720289521   -3.371063003813707
   H  -1.795307501459984    2.891278294563413   -3.927855043896308
   H  -2.709613973668925   -0.308515546176734   -4.026627646697411
   H  -4.404246093821399    0.941639912907262   -0.071175054238094
   H  -1.962867323232915    3.122079490952855    0.528101313545138
   H  -1.245096579039474   -2.621186110634707    0.594784162223769
   H  -1.699155144887690   -0.782162821007662    3.328959985756973
   H   2.347109421287126    1.104305785540561    3.087624818244846
   H   2.990679065503112   -1.888017241218143    1.775287572161196
   H   4.862301668284708    0.796425411350593   -1.279131939569907
   H   2.634027658640572    3.226752635113244   -0.827936424652650
*
```

Including (semi)core orbitals in the correlation treatment

In some chemical applications some or all of the chemical (semi)core electrons must be included in the correlation treatment. In this case, it is necessary to tighten the TCutPNO thresholds for electron pairs in which chemical (semi)core electrons are involved. This is now the default in DLPNO calculations.

For instance, one can decide to switch off the frozen-core approximation and include all the electrons in the correlation treatment. In this case, the program will use tighter thresholds by default for all electron pairs and Singles that involve chemical core electrons. Note that, in this case, the use of properly optimized basis functions for correlating the inner electrons is highly recommened.

```
! DLPNO-CCSD(T) def2-SVP def2-SVP/C NoFrozenCore
%mdci TSCALEPNO_CORE 0.01 # scaling factor for TCutPNO for electron pairs and
                          # Singles involving chemical core electrons
      end
* xyz 0 1
   Ti   0.0001595288      0.0000775041      0.0000000000
   F    1.7595996122      0.0000634675     -0.0000000011
   F   -0.5865076471      1.6586935196      0.0000000018
   F   -0.5866248292     -0.8294172469     -1.4362516915
   F   -0.5866248311     -0.8294172443      1.4362516907
*
```

Another option is to choose the involved chemical core electrons by using an energy window. In this way all electron pairs and Singles that involve chemical core electrons, which are in the defined energy window, are affected by TScalePNO_CORE.

```
! DLPNO-CCSD(T) def2-SVP def2-SVP/C
%method
  FrozenCore  FC_EWIN
end
%mdci
  EWIN -40, 10000
end
* xyz 0 1
   Ti   0.0001595288      0.0000775041      0.0000000000
   F    1.7595996122      0.0000634675     -0.0000000011
   F   -0.5865076471      1.6586935196      0.0000000018
   F   -0.5866248292     -0.8294172469     -1.4362516915
   F   -0.5866248311     -0.8294172443      1.4362516907
*
```

A summary with the number of electrons affected by TScalePNO_Core for the examples just discussed is shown in Table Table 7.19.

Table 7.19: Number of chemical core electrons included in the DLPNO calculation and affected by TScalePNO_Core for the TiF_4 examples

Keyword	Chemical Core Electrons		Valence Electrons
	Frozen	Included[a]	
FrozenCore (default)	18	0	40
NoFrozenCore	0	18	40
EWIN -40, 10000	16	2	40

[a] using TScalePNO_Core.

By default, ORCA provides a chemical meaningful definition for the number of electrons which belong to the chemical core of each element. As already discussed, these default values define which pairs are affected by TScalePNO_Core. However, the user can modify the number of chemical core electrons for a specific element via the NewNCore keyword.

```
! DLPNO-CCSD(T) def2-SVP def2-SVP/C NoFrozenCore
%method
   NewNCore Ti 8 end
end
* xyz 0 1
   Ti   0.0001595288      0.0000775041      0.0000000000
   F    1.7595996122      0.0000634675     -0.0000000011
   F   -0.5865076471      1.6586935196      0.0000000018
   F   -0.5866248292     -0.8294172469     -1.4362516915
   F   -0.5866248311     -0.8294172443      1.4362516907
*
```

In the previous example, the number of chemical core electrons for Ti has been fixed to 8.

Starting from ORCA 6.0, in DLPNO calculations, tightened TCutPNO thresholds are used by default for "semicore" electron pairs involving the 3s and 3p orbitals of first-row transition metals.[32] This improves not only the accuracy of the results noticeably but also the efficiency of the computations as the system size grows (ref.). To reproduce results obtained with earlier versions, the number of semicore orbitals on first-row transition metals needs to be set to zero (the old default) instead of eight (the current default), as done below for a Zn atom with the NewNSemiCore keyword in the method block:

```
%method NewNSemiCore Zn 0 end end
```

NOTE

- Of course, if electrons are replaced by ECPs, they are not included in the correlation treatment.
- If ECPs are used, the number for NewNCore has to include the electrons represented by the ECPs as well. E.g. if an element is supposed to have 60 electrons in the ECP and additional 8 electrons should be frozen in the correlation calculation, NewNCore should be 68.
- The different sets of orbitals (chemical core electrons included in the correlation treatment and valence electrons) are localized separately in order to avoid the mixing of core and valence orbitals.

Multi-Level Calculations

In many applications events are investigated that are located in a relatively small region of the system of interest. In these cases, combined quantum-mechanics/molecular-mechanics (QM/MM) approaches have been proved to be extremely useful, especially in the modeling of enzymatic reactions. The basic idea of any QM/MM method is to treat a small region of the system at the QM level and to use an MM treatment for the remaining part of the system. Alternatively, QM/QM methods, where different parts of a system are studied at various quantum mechanical levels, are also available. Quantum mechanical methods are more computationally demanding than the molecular mechanics treatment, and this limits the applicability of all-QM approaches. Nevertheless, QM/QM methods retain some strong advantages over QM/MM schemes. For instance, force field parameters for the molecular mechanics part of the calculation are not necessary, and thus there are no restrictions on the type of chemical systems that can be treated. Moreover, problems usually caused by boundaries between QM and MM parts do not occur. Finally, the accuracy of an all-QM calculation is expected to be higher compared to the accuracy of QM/MM approaches, that is limited by the MM treatment.

In ORCA, the different methods that can be used in a QM/QM calculations are:

- DLPNO-CCSD(T) with TightPNO thresholds
- DLPNO-CCSD(T) with NormalPNO thresholds
- DLPNO-CCSD(T) with LoosePNO thresholds
- DLPNO-CCSD
- DLPNO-MP2
- HF

The user can define an arbitrary number of fragments in the input, the level of theory to be used for each fragment and for the interaction between different fragments. Localized molecular orbitals are then assigned to a given fragment on the basis of their Mulliken populations.

The following example shows the calculation of a benzene dimer, for which the individual monomers are calculated on MP2 level, and the interaction between the two monomers is calculated on TightPNO DLPNO-CCSD(T) level. More realistic use cases are discussed in ref. [811].

```
! DLPNO-CCSD(T) Def2-SVP Def2-SVP/C
%mdci
  UseFullLmp2Guess True
  TightPNOFragInter  {1 2}
  MP2FragInter  {1 1} {2 2}
end

*xyz 0 1
C(1)       1.393      0.000      0.0
H(1)       2.474      0.000      0.0
C(1)       0.695      1.206      0.0
H(1)       1.238      2.143      0.0
C(1)      -0.695      1.206      0.0
H(1)      -1.238      2.143      0.0
C(1)      -1.393      0.000      0.0
H(1)      -2.474      0.000      0.0
C(1)      -0.695     -1.206      0.0
H(1)      -1.238     -2.143      0.0
C(1)       0.695     -1.206      0.0
H(1)       1.238     -2.143      0.0
C(2)       2.333      1.33       3.5
H(2)       3.414      1.33       3.5
C(2)       1.635      2.536      3.5
H(2)       2.178      3.473      3.5
C(2)       0.245      2.536      3.5
H(2)      -0.298      3.473      3.5
C(2)      -0.453      1.33       3.5
```

(continues on next page)

(continued from previous page)

```
H(2)      -1.534      1.33       3.5
C(2)       0.245      0.124      3.5
H(2)      -0.298     -0.813      3.5
C(2)       1.635      0.124      3.5
H(2)       2.178     -0.813      3.5
*
```

- For the calculation of the interaction energy, the energy of the individual benzene monomer should be calculated on the accuracy level of the monomer in the dimer calculation, i.e. using MP2 with full LMP2 guess for the above example.

All possible settings for the multi-level calculation are listed below.

```
# The one-keyword line defines the default method for the multi-level calculation.
# Options here are DLPNO-CCSD(T) or DLPNO-CCSD with the addition of the
# LoosePNO, NormalPNO and TightPNO keyword
!DLPNO-CCSD(T)

# The below given keywords define the changes with respect to the
# above given default method. The user should take care that each intra- or
# interfragment combination is defined only once (unlike in the example given below)

%mdci
  LoosePNOFragInter {1 1} {2 2}       # use LoosePNO settings for the intrafragment
                                      # pair energies of fragments 1 and 2
  NormalPNOFragInter {1 2}            # use NormalPNO settings for the interfragment
                                      # pair energies between fragment 1 and 2
  TightPNOFragInter {1 3}             # use TightPNO settings for the interfragment
                                      # pair energies between fragment 1 and 3
  NormalPNOTightPairFragInter {1 2}   # use NormalPNO settings but with TCutPairs
                                      # 1.e-5 for the interfragment pair energies
                                      # between fragment 1 and 2
  LoosePNOTightPairFragInter {1 3}    # use LoosePNO settings but with TCutPairs 1.e-5
                                      # for the interfragment pair energies between
                                      # fragment 1 and 3
  NoTriplesFragments  1, 2            # if all MOs of a triple are located on fragment
                                      # 1 and / or 2, the triple is neglected
  MP2FragInter  {1 1} {2 2}           # compute the intrafragment pair energies of
                                      # fragments 1 and 2 on MP2 level
  HFFragInter  {1 1} {2 2}            # compute the intrafragment energies on HF level
  UseFullLmp2Guess false              # default = false,
                                      # if MP2FragInter is used: default = true

# The fragments need to be defined in the xyz section.
*xyz 0 1
C(1)       1.393      0.000      0.0
H(1)       2.474      0.000      0.0
...
```

Multi-Level Calculations for IP and EA-EOM-DLPNO-CCSD

The multi-layer method can be used to include the environmental effect in IP-and EA-EOM-DLPNO-CCSD method. A typical input file for the multi-layer IP-EOM-CCSD method will look like

```
! IP-EOM-DLPNO-CCSD TightSCF NORMALPNO def2-SVP def2-SVP/C def2/J RIJCOSX

%mdci
nroots 4
DTOl 1e-7
NORMALPNOFragInter { 1 1}
```

(continues on next page)

(continued from previous page)

```
LOOSEPNOFragInter { 1 1}
HFFRAGMENTINTERACTION { 2 2}
end

*xyz 0 1
C(1)       2.782064    -1.456235     0.000000
C(1)       1.478695    -0.729491     0.000000
C(1)       0.274461    -1.343436     0.000000
N(1)      -0.914790    -0.659079     0.000000
C(1)      -0.988897     0.709718     0.000000
N(1)       0.241239     1.324758     0.000000
H(1)       0.224165     2.335424     0.000000
C(1)       1.507368     0.726274     0.000000
O(1)       2.518005     1.411594     0.000000
O(1)      -2.043648     1.337449     0.000000
H(1)      -1.808257    -1.143873     0.000000
H(1)       0.182931    -2.420317     0.000000
H(1)       2.626386    -2.532891     0.000000
H(1)       3.370859    -1.183830    -0.874506
H(1)       3.370859    -1.183830     0.874506
O(2)      -3.661424    -0.883408     0.000000
H(2)      -3.462053     0.068032     0.000000
H(2)      -4.615649    -0.964529     0.000000
*
```

Here the example is a mono-hydrated thymine molecule, where the thymine is treated at the main fragment and water is treated at the environment. It will result in the following output

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT=  1:  0.322826 au     8.785 eV   70852.1 cm**-1
  Amplitude    Excitation
  -0.689911    37 -> x
Percentage singles character=     96.89

IROOT=  2:  0.364959 au     9.931 eV   80099.2 cm**-1
  Amplitude    Excitation
  -0.689956    35 -> x
Percentage singles character=     95.49

IROOT=  3:  0.378175 au    10.291 eV   82999.9 cm**-1
  Amplitude    Excitation
  -0.691827    36 -> x
Percentage singles character=     93.93

IROOT=  4:  0.403845 au    10.989 eV   88633.7 cm**-1
  Amplitude    Excitation
  -0.690254    34 -> x
Percentage singles character=     96.55
```

The result of a full a IP-EOM-DLPNO-CCSD calculation with NORMALPNO setting would have looked like

```
IROOT=  1:  0.322576 au     8.778 eV   70797.3 cm**-1
  Amplitude    Excitation
   0.689734    37 -> x
Percentage singles character=     96.88

IROOT=  2:  0.364691 au     9.924 eV   80040.3 cm**-1
  Amplitude    Excitation
```

(continues on next page)

(continued from previous page)

```
  -0.689947    35 -> x
Percentage singles character=     95.50

IROOT=  3:  0.377966 au    10.285 eV   82954.0 cm**-1
  Amplitude    Excitation
  -0.691801    36 -> x
Percentage singles character=     93.94

IROOT=  4:  0.402497 au    10.953 eV   88337.9 cm**-1
  Amplitude    Excitation
  -0.690138    34 -> x
Percentage singles character=     96.50
```

The results in multi-layer IP-EOM-DLPNO-CCSD method has been found to be in excellent agreement with standard variant. The $MP2FragInter$ treatment is not available for the EOM method. To get a reasonable accuracy one need to treat the fragment from where the ionization is happening (thymine in the above example) at the highest possible level. The interaction between the main fragment (thymine) and environment (water) should be treated at the intermediate level accuracy. The environment can safely be treated with $HFFragInter$ for almost all the cases. One can decide the size of the main fragment by looking at HF occupied orbitals as the Koopmans' approximation is a very good zeroth order guess for the IP values. The electron attached states are much less localized as compared to the ionization problem. Consequently, the multi-layer EA-EOM-DLPNO-CCSD requires much more tighter thresholds than the IP variant. An typical input file for multi-layer EA-EOM-DLPNO-CCSD will look as follows.

```
! EA-EOM-DLPNO-CCSD NORMALPNO  ma-def2-SVP RIJCOSX aug-cc-pVDZ/C def2/J

%mdci
  NRoots             4
  FollowCIS          true
  TCutPNOSingles     1e-12
  MaxIter            2000
  DTol               1e-7
  NDAV               10
  NormalPNOFragInter { 1 1 }
  LoosePNOFragInter  { 1 2 } { 2 2 }
end

* xyz 0 1
N(1)   -1.114   -0.934   -3.554
C(1)   -0.343   -0.202   -4.483
H(1)   -0.668   -0.311   -5.520
C(1)   0.635   1.107   -2.633
O(1)   1.241   2.018   -2.013
N(1)   0.022   0.050   -1.776
H(1)   -0.069   0.339   -0.800
C(1)   -0.975   -0.782   -2.233
O(1)   -1.697   -1.422   -1.333
C(1)   1.087   1.852   -4.986
H(1)   1.673   2.594   -4.418
H(1)   1.771   1.340   -5.697
H(1)   0.348   2.403   -5.609
H(1)   -1.823   -1.635   -3.888
N(2)   -4.904   -4.773   -4.958
H(2)   -4.634   -3.987   -5.572
C(2)   -4.875   -4.201   -3.599
C(2)   -3.704   -3.226   -3.310
H(2)   -5.818   -3.646   -3.423
H(2)   -4.859   -5.012   -2.850
O(2)   -3.026   -2.826   -4.301
H(2)   -4.078   -5.386   -5.029
```

(continues on next page)

(continued from previous page)

```
O(2)   -3.559   -2.899   -2.077
H(2)   -2.494   -2.057   -1.754
C(2)   0.440   0.898   -4.018
end
```

It is a thymine-glycine complex where the thymine is treated as the main fragment and glycine as the environment. One needs to use a more tighter value of $TCutPNOSingles$ for EA as in the case of standard EA-EOM-DLPNO-CCSD. The $TCutPNOSingles$ for the respective fragments automatically gets adjusted based on their respective $TCutPNO$ values. The output will be

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT=  1: -0.039822 au    -1.084 eV   -8739.9 cm**-1
  Amplitude    Excitation
   0.689012     x ->  53
Percentage singles character=     93.33

IROOT=  2:  0.025156 au     0.685 eV    5521.0 cm**-1
  Amplitude    Excitation
  -0.614385     x ->  54
   0.139410     x ->  55
   0.297903     x ->  59
Percentage singles character=     96.86

IROOT=  3:  0.044569 au     1.213 eV    9781.7 cm**-1
  Amplitude    Excitation
   0.116867     x ->  54
   0.684240     x ->  55
Percentage singles character=     98.16

IROOT=  4:  0.053677 au     1.461 eV   11780.7 cm**-1
  Amplitude    Excitation
   0.695211     x ->  56
Percentage singles character=     98.12
```

The results are in excellent agreement with the standard EA-EOM-DLPNO-CCSD method.

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT=  1: -0.038862 au    -1.057 eV   -8529.3 cm**-1
  Amplitude    Excitation
  -0.689412     x ->  53
Percentage singles character=     93.47

IROOT=  2:  0.025448 au     0.692 eV    5585.2 cm**-1
  Amplitude    Excitation
  -0.654101     x ->  54
   0.131404     x ->  55
   0.207630     x ->  59
Percentage singles character=     97.47

IROOT=  3:  0.044651 au     1.215 eV    9799.8 cm**-1
  Amplitude    Excitation
   0.112183     x ->  54
   0.684562     x ->  55
Percentage singles character=     98.17
```

(continues on next page)

(continued from previous page)

```
IROOT=  4:  0.053780 au      1.463 eV   11803.3 cm**-1
  Amplitude    Excitation
   0.695635     x ->  56
Percentage singles character=     98.16
```

To get the reasonable accuracy in multi-layer EA-EOM-CCSD one need to treat the environment and inter-fragment interaction atleast at $LoosePNOFragInter$ level.

7.14.5 The singles Fock term

In most MDCI calculations, there is an intermediate, which resembles closely to the SCF Fock matrix, and similar methods are available to efficiently calculate it. In the followings, a short discussion will be given of the so-called singles Fock term, which in the closed shell case has the form

$$G(\mathbf{t}_1)_{pq} = \sum_{jb} t_b^j (2(pq|jb) - (pj|qb)) = \sum_{\mu\nu} c_\mu^p c_\nu^q G(\mathbf{t}_1)_{\mu\nu},$$

The singles Fock matrix can be obtained via transformation from its counterpart ($G(\mathbf{t}_1)_{\mu\nu}$) in the atomic orbital (AO) basis

$$G(\mathbf{t}_1)_{\mu\nu} = \sum_{jb} t_b^j (2(\mu\nu|jb) - (\mu j|\nu b)) = \sum_{\kappa\tau} P(\mathbf{t}_1)_{\kappa\tau} (2(\mu\nu|\kappa\tau) - (\mu\kappa|\nu\tau)), \tag{7.148}$$

where

$$P(\mathbf{t}_1)_{\kappa\tau} = \sum_{jb} t_b^j c_\kappa^j c_\tau^b$$

is the analogue of the SCF density matrix for the singles Fock case. For the singles Coulomb ($J(\mathbf{t}_1)_{\mu\nu}$) case, the density may be symmetrized ($\tilde{P}(\mathbf{t}_1)_{\kappa\tau} = P(\mathbf{t}_1)_{\kappa\tau} + P(\mathbf{t}_1)_{\tau\kappa}$), and one may use the resolution of identity approximation

$$J(\mathbf{t}_1)_{\mu\nu} = \sum_{\kappa\tau} \tilde{P}(\mathbf{t}_1)_{\kappa\tau} (\mu\nu|\kappa\tau) \approx \sum_{AB} \sum_{\kappa\tau} \tilde{P}(\mathbf{t}_1)_{\kappa\tau} (\mu\nu|r_{12}^{-1}|A) V_{AB}^{-1} (B|r_{12}^{-1}|\kappa\tau),$$

where A, B are elements of the RI/DF auxiliary fitting basis. Note that the factor of 2 in ((7.148)) is taken care of by symmetrization. Since we are using a symmetric density, we may use the same efficient routine to evaluate the singles Coulomb term as in the SCF case, see *Using the RI-J Approximation to the Coulomb Part* and *The Split-RI-J Coulomb Approximation*.

For the exchange case ($K(\mathbf{t}_1)_{\mu\nu}$), one possibility is to use the COSX approximation (see *Using the RI Approximation for Hartree-Fock and Hybrid DFT (RIJCOSX)*)

$$K(\mathbf{t}_1)_{\mu\nu} = \sum_{\kappa\tau} P(\mathbf{t}_1)_{\kappa\tau} (\mu\kappa|\nu\tau) \approx \sum_g Q_{\mu g} \sum_\tau A_{\nu\tau}(\mathbf{r}_g) \sum_\kappa P(\mathbf{t}_1)_{\kappa\tau} X_{\kappa g},$$

The COSX routine is able to deal with asymmetric densities as well, and thus, it can be used here similar to the SCF case.

The other possibility is to use RI for exchange (RIK),

$$K(\mathbf{t}_1)_{\mu\nu} = \sum_{j\kappa\tau} c_\kappa^j C(\mathbf{t}_1)_\tau^j (\mu\kappa|\nu\tau) \approx \sum_{jAB} (\mu j|r_{12}^{-1}|A) V_{AB}^{-1} (B|r_{12}^{-1}|\nu\tilde{j}),$$

where

$$C(\mathbf{t}_1)_\tau^j = \sum_b t_b^j c_\tau^b,$$

and the $\tilde{j}$ is an "orbital" transformed using $C(\mathbf{t}_1)$.

Using these approximations, there are two approximations for the total singles Fock term, RIJCOSX called by the simple keyword RCSinglesFock and RIJK called by RIJKSinglesFock, see *Basics*. For canonical and LPNO methods, by default the program chooses the same approximation used in the SCF calculation. DLPNO2013 uses RIJCOSX by default, while in DLPNO, the singles Fock term is also evaluated using PNOs via SinglesFockUsePNOs, see *Local correlation*. This behavior can also be changed by keywords in the method block.

```
%method RIJKSinglesFock 1   # 0 false, 1 true
        RCSinglesFock   0   # 0 false, 1 true
end
```

7.14.6 Use of the MDCI Module

The MDCI module is fairly easy to use. The flags for the "simple" input lines have been described in section *Keyword Lines*. The detailed listing of options is found below:

```
%mdci citype CISD     # CI singles+doubles
             QCISD    # quadratic CI (singles+doubles)
             CCSD     # coupled-cluster singles+doubles
             CEPA_1   # coupled-electron pair approximation ''1''
             CEPA_2   #
             CEPA_3   #
             NCEPA_1  # our slightly modified versions of CEPA
             NCEPA_2  # and CPF
             NCEPA_3  #
             NCPF_1   #
             NCPF_2   #
             NCPF_3   #
             ACPF     # averaged coupled-pair functional
             ACPF_2   # Gdanitz modification of it
             NACPF    # our modification of it
             AQCC     # Szalay + Bartlett
             SEOI     # a strictly size extensive energy functional
                      # maintaining unitary invariance (not yet
                      # available for UHF)
        ewin -3,1e3   # orbital energy window to determine which
                      # MOs are included in the treatment
                   # (respects settings in %method block)
           Singles true  # include single excitations in the
                         # treatment (default true)
           Triples 0     # (T) correction in CCSD(T)/QCISD(T)
                         # default is no triples
                   1     # Algorithm 1 (lots of memory, fast)
                   2     # Algorithm 2 (less memory, about 2x slower,
                         # not yet available for UHF)
           Brueckner true   # use Brueckner orbitals
                            # (default false)
           Denmat none       # no evaluation of density matrices
                  linearized # density matrix obtained by retaining
                             # only CEPA_0-like terms, i.e., those
                             # linear in the excitation amplitudes
                  unrelaxed  # unrelaxed density matrices, i.e.,
                             # density matrices without orbital
                             # relaxation
                  orbopt     # perform orbital optimization yielding
                             # fully relaxed density matrices (if
                             # citype chosen as CCSD or QCISD this option
                             # implies evaluation of the Z vector).
                             # (default: linearized)
           ZSimple true     # simplified evaluation of the Z vector
                            # in case of orbital optimized CCD
```

(continues on next page)

(continued from previous page)

```
                  # (citype chosen as CCSD or QCISD and
                  # Denmat as orbopt) by using an
                  # analytical formula
         false    # explicit solution of Z vector
                  # equations
                  # in case of orbital optimized CCD
                  # (default: false)
UseQROs           # use of quasi-restricted orbitals
                  # (default false)
Localize 0    # use localized MOs. Presently very little
              # use is made of locality. It may help
              # for interpretations. Localization is
              # incompatible with the (T) correction
          PM  # Use Pipek-Mezey localized MOs
          FB  # use Foster-Boys localized MOs
NatOrbIters 0 # Perform natural orbital iterations.
              # default is none. Not possible for CCSD
              # and QCISD
pCCSDAB       # the three parameters for parametrized
pCCSDCD       # coupled-cluster (default is 1.0 which
pCCSDEF       # corresponds to normal CCSD
# this defines how the rate limiting step is handled
# MO and AOX need lots of disk and I/O but if they
# can be done they are fast
KCOpt KC_MO    # Perform full 4-index transformation
      KC_AOBLAS# AO direct with BLAS (preferred)
               # (not yet available for UHF, switch to KC_AOX)
      KC_AO    # AO direct handling of 3,4 externals
               # (not yet available for UHF, switch to KC_AOX)
      KC_RI    # RI approximation of 3,4 externals
               # (not yet available for UHF)
      KC_RI2   # Alternative RI (not recommended)
               # (not yet available for UHF)
      KC_AOX   # Do it from stored AO exchange integrals
PrintLevel 2   # Control the amount of output. For 3 and
               # higher things like pair correlation
               # energies are printed.
MaxIter 35     # Max. number of iterations
# How the integral transformation is done.
# Note that it is fine to do AOX or AO or AOBLAS
# together with trafo_ri
TrafoType trafo_jk # Partial trafo to J+K operators
          trafo_ri   # RI transformation of all
                     # integrals up to 2-externals
                     # (3-ext for (T))and rest on the
                     # fly
          trafo_full # Full four index transformation.
                     # Automatically chosen for
                     # KCOpt=KC_MO
MaxCore 350 # Memory in MB - used for integral
               # trafos and batching and for storage of
               # integrals and amplitudes
               # don't be too generous
STol  1e-5     # Max. element of the residual vector
              # for convergence check
LShift 0.3     # Level shift to be used in update of
               # coefficients
MaxDIIS  7     # Max number of DIIS vectors to be stored
# this lets you control how much and what is residing
# in central memory. May speed up things. Note that
# MaxCore is not respected here
```

(continues on next page)

(continued from previous page)

```
        InCore   0   # nothing in core
                 1   #  + sigma-vector and amplitudes (default)
                 2   #  + Jij(a,b) Kij(a,b) operators
                 3   #  + DIIS vectors
                 4   #  + 3-exernal integral Kia(b,c)
                 5   #  + 4-external integrals Kab(c,d)
                     # this is identical to ALL
                 # the default is AUTO which means that incore
                 # is chosen based on MaxCore
        end
```

7.15 The Complete Active Space Self-Consistent Field (CASSCF) Module

7.15.1 General Description

The complete active space self-consistent field (CASSCF) method is a special form of a multiconfigurational SCF method and can be thought of as an extension of the Hartree-Fock method. It is a very powerful method to study static correlation effects and a solid basis for MR-CI and MR-PT treatments. It can be applied to the ground state and excited states or averages thereof. The implementation in ORCA is fairly general and reasonably efficient. However, CASSCF calculations are fairly complex and ultimately require a lot of insight from the user in order to be successful. In addition to detailed description here, the manual explores some typical examples in section *CASSCF Natural Orbitals as Input for Coupled-Cluster Calculations*. Furthermore, the manual is supplemented with a tutorial for CASSCF that covers many practical tips on the calculation design and usage of the program.

The wavefunction. The wavefunction of a given CASSCF state is written as

$$\left|\Psi_I^S\right\rangle = \sum_k C_{kI} \left|\Phi_k^S\right\rangle. \tag{7.149}$$

Here, $\left|\Psi_I^S\right\rangle$ is the CASSCF N-electron wavefunction for state I with total spin S. The set of $\left|\Phi_k^S\right\rangle$ is a set of configuration state functions (for example linear combination of Slater determinants) each adapted to a total spin S. The expansion coefficients C_{kI} represent the first set of variational parameters. Each CSF is constructed from a common set of orthonormal molecular orbitals $\psi_i(\mathbf{r})$ which are in turn expanded in basis functions $\psi_i(\mathbf{r}) = \sum_\mu c_{\mu i}\phi_\mu(\mathbf{r})$. The MO coefficients $c_{\mu i}$ form the second set of variational parameters.

The energy. The energy of the CASSCF wavefunction is given by the Rayleigh quotient

$$E(\mathbf{c},\mathbf{C}) = \frac{\left\langle \Psi_I^S \middle| \hat{H}_{\mathrm{BO}} \middle| \Psi_I^S \right\rangle}{\left\langle \Psi_I^S \middle| \Psi_I^S \right\rangle}, \tag{7.150}$$

and represents an upper bound to the true total energy. However, CASSCF calculations are *not* designed to provide values for total energy which are close to the exact energy. The purpose of a CASSCF calculation is to provide a qualitatively correct wavefunction, which forms a good starting point for a treatment of dynamic electron correlation.

The CASSCF method is fully variational in the sense that the energy is made stationary with respect to variations in both sets of MO and CI coefficients. At convergence, the gradient of the energy with respect to the MO and CI coefficients vanishes

$$\frac{\partial E(\mathbf{c},\mathbf{C})}{\partial c_{\mu i}} = 0, \tag{7.151}$$

$$\frac{\partial E(\mathbf{c},\mathbf{C})}{\partial C_{kI}} = 0. \tag{7.152}$$

Orbital spaces. In CASSCF calculations, the MO space is divided into three user defined subspaces:

- The "inactive orbitals" are the orbitals which are doubly occupied in all configuration state functions (labels i, j, k, l).
- The "active orbitals" are the orbitals with variable occupation numbers in the various CSFs (labels t, u, v, w).
- The "external orbitals" (labels a, b, c, d)

Note that in older publications, the inactive and active orbitals are distinguished and referred to as "internal" orbitals.

The wavefunction and energy is invariant with respect to unitary transformations within the three subspaces. The special feature of a CASSCF wavefunction is that a fixed number of electrons is assigned to each subspace. The internal subspace is of course completely filled but the CSFs in the active space constitute a full-CI of n-electrons in m-orbitals. The CSF list is constructed such, however, that a wavefunction of well defined total spin (and potential space) symmetry results. Such a wavefunction is referred to as a **CASSCF(n,m)** wavefunction. The CSF list grows extremely quickly with the number of active orbitals and the number of active electrons (basically factorially). Depending on the system, the limit of feasibility is roughly around ~14 active orbitals or about one million CSFs in the active space. Larger active spaces are tractable with approximate CI solver such as the Iterative-Configuration-Expansion CI (ICE-CI) described in *Approximate Full CI Calculations in Subspace: ICE-CI* or the Density Matrix Renormalization Group (DMRG) discussed in *Density Matrix Renormalization Group*.[1]

Since the orbitals within the subspaces are only defined up to a unitary transformation, the program needs to make some canonicalization choice.

In ORCA, the final orbitals by default are:

1. natural orbitals in the active space,
2. orbitals which diagonalize the CASSCF Fock matrix in the internal space and
3. orbitals which diagonalize the CASSCF Fock matrix in the external space.

State averaging. In many circumstances, it is desirable to optimize the orbitals not for a single state but for the average of several states. In order to see what is done, the energy for state I is re-written as:

$$E_I\left(\mathbf{c},\mathbf{C}\right) = \sum_{pq} \Gamma_q^{p(I)} h_{pq} + \sum_{pqrs} \Gamma_{qs}^{pr(I)} \left(pq\,|rs\right) \tag{7.153}$$

Here, $\Gamma_q^{p(I)}$ and $\Gamma_{qs}^{pr(I)}$ are the one-and two-particle reduced electron density matrices for this state (labels p, q, r, s span the internal and active subspaces):

$$\Gamma_q^{p(I)} = \left\langle \Psi_I^S \left| E_q^p \right| \Psi_I^S \right\rangle \tag{7.154}$$

$$\Gamma_{qs}^{pr(I)} = \frac{1}{2} \left\langle \Psi_I^S \left| E_q^p E_s^r - \delta_{qr} E_s^p \right| \Psi_I^S \right\rangle \tag{7.155}$$

The average energy is simply obtained from averaging the density matrices using arbitrary weights w_I that are user defined but are constrained to sum to unity.

$$\Gamma_q^{p(av)} = \sum_I w_I \Gamma_q^{p(I)} \tag{7.156}$$

$$\Gamma_{qs}^{pr(av)} = \sum_I w_I \Gamma_{qs}^{pr(I)} \tag{7.157}$$

$$\sum_I w_I = 1 \tag{7.158}$$

Optimization of CASSCF wavefunctions. In general, except for trivial cases, CASSCF wavefunctions are considerably more difficult to optimize than RHF (or UHF) wavefunctions. The underlying reason is that variations in **c** and **C** maybe strongly coupled and the energy functional may have many local minima in (**c**,**C**) space. Consequently, the choice of starting orbitals is of really high importance and the choice which orbitals and electrons are included in the active space has decisive influence on the success of a CASSCF study. In general, after transformation to natural orbitals, one can classify the active space orbitals by their occupation numbers which vary between 0.0 and 2.0. In general, convergence problems are almost guaranteed if orbitals with occupation numbers close

[1] For approximate full CI approaches, CASSCF is neither invariant to active-active rotations nor exactly size-consistent.

to zero or close to 2.0 are included in the active space. Occupation numbers between 0.02 and 1.98 are typically very reasonable and should not lead to large convergence problems. The reason for the occurrence of convergence problems is that the energy is only very weakly dependent on rotations between internal and active orbitals if the active orbital is almost doubly occupied and similarly for the rotations between external and weakly occupied active orbitals. However, in some cases (for example in the study of potential energy surfaces) it may not be avoidable to include weakly or almost inactive orbitals in the active space and in these cases the use of the most powerful convergence aids is necessary (*vide infra*). As in the case of single-determinant wavefunctions (RHF, UHF, RKS, UKS) there are first and second order converging methods available. The first order CASSCF methods require the transformed integrals $(tu|vx)$ with x belonging to any subspace. This is a very small subspace of the total transformed integral list and is readily held in central storage even for larger calculations. On the other hand, second order CASSCF methods require the integrals $(pq|xy)$ and $(px|qy)$ ($p, q =$ internal, active; $x, y =$ any orbital). This is a fairly large set of integrals and their generation is laborious in terms of CPU time and disk storage. Second order CASSCF calculations are therefore more limited in the size of the molecules which can be well treated. It would be possible to basically avoid the integral transformation also in the case of second-order CASSCF calculations and proceed to fully direct calculations. Such calculations may become quite time consuming since there may be a large number of Fock matrix builds necessary.

The augmented Hessian method (Newton-Raphson) solves the eigenvalue problem:

$$\begin{pmatrix} 0 & \mathbf{g} \\ \mathbf{g} & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{t} \end{pmatrix} = \varepsilon \begin{pmatrix} 1 \\ \mathbf{t} \end{pmatrix} \tag{7.159}$$

Here, $\mathbf{g}$ is the orbital gradient (derivative of the total energy with respect to a non-redundant rotation between two orbitals) and $\mathbf{H}$ is the orbital Hessian (second derivative of the energy with respect to two non-redundant orbital rotations). The vector $\mathbf{t}$ (in intermediate normalization obtained from the CI like vector) summarizes the rotation angles. The angles are used to define the antisymmetric matrix ($X_{pq} = -X_{qp}$ is thus the rotation angle between orbitals p and q):

$$\mathbf{X} = \begin{pmatrix} \mathbf{0} & \mathbf{t} \\ -\mathbf{t} & \mathbf{0} \end{pmatrix}, \tag{7.160}$$

which is used to parametrize the unitary matrix $\mathbf{U} = \exp(\mathbf{X})$ which is used to update the orbitals according to:

$$\mathbf{c}^{\text{new}} = \mathbf{c}^{\text{old}}\mathbf{U} \tag{7.161}$$

(where $\mathbf{c}$ is an MO coefficient matrix).

Starting orbitals. You cannot be careful enough with your starting orbitals. What type of initial guess works best depends on the system. Quite often it is not the magnitude of the initial gradient, but the similarity between initial and final active orbitals. The CASSCF tutorial discusses a number of guess options in more detail. Generally speaking, canonical orbitals HF orbitals from a RHF calculation are not good choice, as the identification and selection of the active space orbitals is often difficult. Usually DFT orbitals (quasi-restricted or RKS) perform better in this respect. Alternatively, if CASSCF orbitals from a previous run or a close-by geometry are available this is a good choice. For coordination chemistry complexes, the guess generated with `orca_mergefrag` (see the CASSCF tutorial), is probably the best choice - especially for heave metals. Natural orbitals from a simple correlation calculation like MP2 or a calculation with the MRCI module are usually a good choice and easily generated. For example:

```
#
# First job provides reasonable natural orbitals
#
! RI-MP2 SVP def2-SVP/C

%mp2 natorbs true
     density unrelaxed # or relaxed (more expensive)
     end

* int 0 1
    C  0 0 0 0.00   0.0    0.00
    O  1 0 0 1.20   0.0    0.00
    H  1 2 0 1.10 120.0    0.00
    H  1 2 3 1.10 120.0  180.00
*
```

Now examine the occupation numbers of the natural orbitals (you will find that in the output of the MP2 part of the calculation):

```
Natural Orbital Occupation Numbers:
N[  0] =   2.00000000
N[  1] =   2.00000000
N[  2] =   1.98676733
N[  3] =   1.97726840
N[  4] =   1.97500109
N[  5] =   1.96759239
N[  6] =   1.96423113
N[  7] =   1.93719340
N[  8] =   0.05427454
N[  9] =   0.02555886
N[ 10] =   0.02530580
N[ 11] =   0.01358500
N[ 12] =   0.01096092
N[ 13] =   0.01028129
N[ 14] =   0.00702048
N[ 15] =   0.00627820
```

A rule of thumb is that orbitals with occupation numbers between 1.98 and 0.02 should be in the active space. Thus, in the present case we speculate that a 10 electrons in 8 orbitals active space would be appropriate for the CASSCF of the ground state. Let's try:

```
#
# Run a CASSCF calculation for the ground state of H2CO
#
! SVP def2-SVP/C SmallPrint
! moread
%moinp "Test-CASSCF-MP2-H2CO.mp2nat"

%casscf nel        10
        norb        8
        mult        1
        end

* int 0 1
    C  0 0 0 0.00   0.0    0.00
    O  1 0 0 1.20   0.0    0.00
    H  1 2 0 1.10 120.0    0.00
    H  1 2 3 1.10 120.0  180.00
*
```

If we run that calculation, it converges and produces the following:

```
MACRO-ITERATION  10:
        --- Inactive Energy E0 = -82.97337099 Eh
        E(CAS)=  -113.889438276 Eh DE=    -0.000000807
        --- Energy gap subspaces: Ext-Act = -0.431   Act-Int = -0.240
        N(occ)=  1.99763 1.99696 1.98360 1.97923 1.94253 0.05958 0.02153 0.01894
        ||g|| =      0.000361782 Max(G)=     0.000189613 Rot=9,2
        ---- THE CAS-SCF GRADIENT HAS CONVERGED ----
                  --- FINALIZING ORBITALS ---
        ---- DOING ONE FINAL ITERATION FOR PRINTING ----
        --- Forming Natural Orbitals
        --- Canonicalize Internal Space
        --- Canonicalize External Space
```

From which we see that we had *two orbitals too many in the active space* with occupation numbers very close to two. The presence of barely correlated orbitals (occupation close to 0.0 or 2.0) can cause convergence problems. Their inclusion in the active space does not significantly change the energy and it might better to omit these orbitals from the start.

In the present case, we re-run the CASSCF with 6 active electrons in six orbitals. The result is:

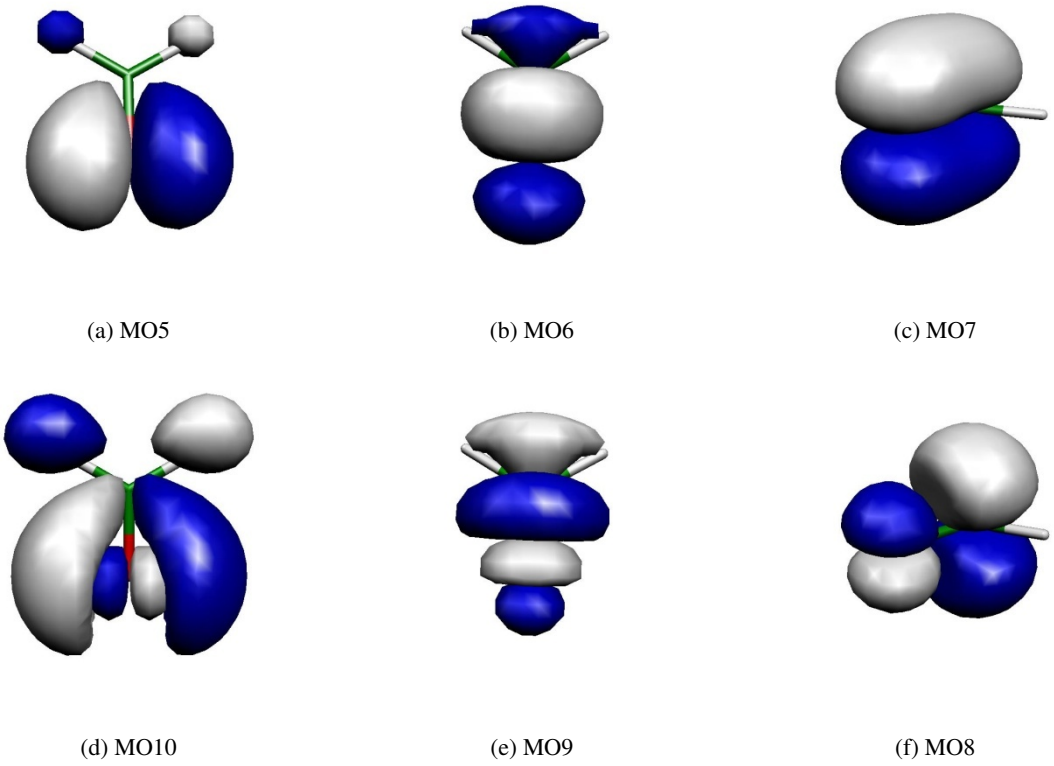
```
MACRO-ITERATION   2:
        --- Inactive Energy E0 = -101.16144179 Eh
        E(CAS)=  -113.882700257 Eh DE=    -0.012049926
        --- Energy gap subspaces: Ext-Act = -0.411   Act-Int = -0.142
        N(occ)=  1.98172 1.97921 1.94092 0.05983 0.02089 0.01743
        ||g|| =      0.052811635 Max(G)=     0.025065586 Rot=19,7
        --- Orbital Update [SuperCI(PT)]
        --- Canonicalize Internal Space
        --- Canonicalize External Space
        --- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.160674186 Max(X)(5,4) =     -0.128053569
        --- SFit(Active Orbitals)

        MACRO-ITERATION   3:
        --- Inactive Energy E0 = -100.78371592 Eh
        E(CAS)=  -113.885011169 Eh DE=    -0.002310912
        --- Energy gap subspaces: Ext-Act = -0.434   Act-Int = -0.199
        N(occ)=  1.98150 1.97909 1.94143 0.05924 0.02108 0.01766
        ||g|| =      0.017438409 Max(G)=     0.009231446 Rot=10,4
        --- Orbital Update [SuperCI(PT)]
        --- Canonicalize Internal Space
        --- Canonicalize External Space
        --- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.050337699 Max(X)(6,2) =     -0.033671129
        --- SFit(Active Orbitals)

        MACRO-ITERATION   4:
        --- Inactive Energy E0 = -100.72313195 Eh
        E(CAS)=  -113.885258854 Eh DE=    -0.000247685
        --- Energy gap subspaces: Ext-Act = -0.438   Act-Int = -0.219
        N(occ)=  1.98141 1.97918 1.94178 0.05886 0.02102 0.01776
        ||g|| =      0.009726271 Max(G)=     0.004281706 Rot=9,2
        --- Orbital Update [SuperCI(PT)]
        --- Canonicalize Internal Space
        --- Canonicalize External Space
        --- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.031123960 Max(X)(22,9) =      0.015789781
        --- SFit(Active Orbitals)

        MACRO-ITERATION   5:
        --- Inactive Energy E0 = -100.65264536 Eh
        E(CAS)=  -113.885424851 Eh DE=    -0.000165997
        --- Energy gap subspaces: Ext-Act = -0.440   Act-Int = -0.238
        N(occ)=  1.98140 1.97918 1.94202 0.05857 0.02105 0.01776
        ||g|| =      0.006606671 Max(G)=     0.003548636 Rot=9,2
        --- Orbital Update [SuperCI(PT)]
        --- Canonicalize Internal Space
        --- Canonicalize External Space
        --- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.019988497 Max(X)(6,2) =     -0.014410848
        --- SFit(Active Orbitals)

        MACRO-ITERATION   6:
        --- Inactive Energy E0 = -100.56070274 Eh
        E(CAS)=  -113.885549550 Eh DE=    -0.000124699
        --- Energy gap subspaces: Ext-Act = -0.440   Act-Int = -0.268
        N(occ)=  1.98138 1.97925 1.94206 0.05849 0.02104 0.01778
        ||g|| =      0.004483296 Max(G)=     0.002939015 Rot=9,2
        --- Orbital Update [SuperCI(PT)]
        --- Canonicalize Internal Space
        --- Canonicalize External Space
        --- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.011383690 Max(X)(5,4) =      0.005997355
        --- SFit(Active Orbitals)

        MACRO-ITERATION   7:
```

(continues on next page)

(continued from previous page)

```
        --- Inactive Energy E0 = -100.52522560 Eh
        E(CAS)=  -113.885583124 Eh DE=    -0.000033574
        --- Energy gap subspaces: Ext-Act = -0.437   Act-Int = -0.283
        N(occ)=  1.98132 1.97929 1.94192 0.05861 0.02103 0.01783
        ||g|| =      0.002275031 Max(G)=    -0.001215398 Rot=10,4
        --- Orbital Update [SuperCI(PT)]
        --- Canonicalize Internal Space
        --- Canonicalize External Space
        --- SX_PT (Skipped TA=0 IT=0): ||X|| =      0.002106033 Max(X)(19,10) =     -0.
↪001056121
        --- SFit(Active Orbitals)

        MACRO-ITERATION   8:
        --- Inactive Energy E0 = -100.52457962 Eh
        E(CAS)=  -113.885584276 Eh DE=    -0.000001152
        --- Energy gap subspaces: Ext-Act = -0.438   Act-Int = -0.283
        N(occ)=  1.98134 1.97931 1.94184 0.05868 0.02101 0.01781
        ||g|| =      0.000752012 Max(G)=    -0.000357510 Rot=13,4
        ---- THE CAS-SCF GRADIENT HAS CONVERGED ----
        --- FINALIZING ORBITALS ---
        ---- DOING ONE FINAL ITERATION FOR PRINTING ----
        --- Forming Natural Orbitals
        --- Canonicalize Internal Space
        --- Canonicalize External Space

        MACRO-ITERATION   9:
        --- Inactive Energy E0 = -100.52457962 Eh
        --- All densities will be recomputed
        E(CAS)=  -113.885584276 Eh DE=    -0.000000000
        --- Energy gap subspaces: Ext-Act = -0.858   Act-Int = -0.283
        N(occ)=  1.98172 1.97932 1.94207 0.05845 0.02100 0.01743
        ||g|| =      0.000752012 Max(G)=    -0.000327367 Rot=12,4
        --------------
        CASSCF RESULTS
        --------------

        Final CASSCF energy       : -113.885584276 Eh   -3098.9843 eV
```

The calculation converges very quickly and the occupation numbers show you that all of these orbitals are actually needed in the active space. The omission of the two orbitals from the active space came at an increase of the energy by ∼4 mEh which seems to be tolerable. Let's look what we have in the active space in figure Fig. 7.4.



(a) MO5 (b) MO6 (c) MO7

(d) MO10 (e) MO9 (f) MO8

Fig. 7.4: Orbitals of the active space for the CASSCF(6,6) calculation of H_2CO.

Thus, we can see that we got a fairly nice result: our calculation has correlated the in-plane oxygen lone pair, the C-O σ and the C-O π bond. For each strongly occupied bonding orbital, there is an accompanying weakly occupied antibonding orbital in the active space that is characterized by one more node. In particular, the correlating lone pair and the C-O σ^* orbital would have been hard to find with any other procedure than the one chosen based on natural orbitals. We have now done it blindly and looked at the orbitals only after the CASSCF — a better approach is normally to look at the starting orbitals *before* you enter a potentially expensive CASSCF calculation. If you have bonding/antibonding pairs in the active space plus perhaps the singly-occupied MOs of the system you probably have chosen a reasonable active space.

We can play the game now somewhat more seriously and optimize the geometry of the molecule using a reasonable basis set:

```
! def2-TZVP def2-TZVP/C SmallPrint Opt
! moread
%moinp "Test-CASSCF-MP2-H2CO.mp2nat"

%casscf nel         6
        norb        6
        end

* int 0 1
    C  0 0 0 0.00   0.0    0.00
    O  1 0 0 1.20   0.0    0.00
    H  1 2 0 1.10 120.0    0.00
    H  1 2 3 1.10 120.0  180.00
*
```

and get:

```
-------------------------------------------------------------------------
    Redundant Internal Coordinates

    --- Optimized Parameters ---
    (Angstroem and degrees)

    Definition                    OldVal   dE/dq     Step     FinalVal
    ----------------------------------------------------------------------------
    1. B(O   1,C   0)               1.2101  0.000259 -0.0002    1.2100
    2. B(H   2,C   0)               1.0942 -0.000029  0.0001    1.0943
    3. B(H   3,C   0)               1.0942 -0.000029  0.0001    1.0943
    4. A(O   1,C   0,H   3)         122.07  0.000023   -0.00    122.07
    5. A(H   2,C   0,H   3)         115.85 -0.000046    0.01    115.86
    6. A(O   1,C   0,H   2)         122.07  0.000023   -0.00    122.07
    7. I(O   1,H   3,H   2,C   0)     -0.00 -0.000000    0.00     -0.00
    ----------------------------------------------------------------------------
```

Let us compare to MP2 geometries (this job was actually run first):

```
! RI-MP2 def2-TZVP def2-TZVP/C Opt

%mp2 natorbs true
     end

* int 0 1
    C  0 0 0 0.00   0.0    0.00
    O  1 0 0 1.20   0.0    0.00
    H  1 2 0 1.10 120.0    0.00
    H  1 2 3 1.10 120.0  180.00
*
```

```
-------------------------------------------------------------------------
                        Redundant Internal Coordinates

                         --- Optimized Parameters ---
                           (Angstroem and degrees)

        Definition                    OldVal   dE/dq     Step     FinalVal
    ----------------------------------------------------------------------------
     1. B(O   1,C   0)               1.2127  0.000374 -0.0002    1.2125
     2. B(H   2,C   0)               1.0991 -0.000031  0.0001    1.0992
     3. B(H   3,C   0)               1.0991 -0.000031  0.0001    1.0992
     4. A(O   1,C   0,H   3)         121.77  0.000023   -0.00    121.77
     5. A(H   2,C   0,H   3)         116.45 -0.000046    0.01    116.46
     6. A(O   1,C   0,H   2)         121.77  0.000023   -0.00    121.77
     7. I(O   1,H   3,H   2,C   0)     -0.00 -0.000000    0.00     -0.00
    ----------------------------------------------------------------------------
```

The results are actually extremely similar (better than 1 pm agreement). Compare to RHF:

```
-------------------------------------------------------------------------
                        Redundant Internal Coordinates

                         --- Optimized Parameters ---
                           (Angstroem and degrees)

        Definition                    OldVal   dE/dq     Step     FinalVal
    ----------------------------------------------------------------------------
     1. B(O   1,C   0)               1.1784 -0.000164  0.0001    1.1785
     2. B(H   2,C   0)               1.0921  0.000010 -0.0000    1.0921
     3. B(H   3,C   0)               1.0921  0.000010 -0.0000    1.0921
     4. A(O   1,C   0,H   3)         121.93 -0.000003   -0.00    121.93
```

(continues on next page)

(continued from previous page)

```
    5. A(H   2,C   0,H   3)          116.13  0.000005    0.00    116.13
    6. A(O   1,C   0,H   2)          121.93 -0.000003   -0.00    121.93
    7. I(O   1,H   3,H   2,C   0)      0.00  0.000000   -0.00     -0.00
   -----------------------------------------------------------------------------
```

Thus, one can observe that the correlation brought in by CASSCF or MP2 has an important effect on the C=O distance (∼4 pm), while the rest of the geometry is not much affected.

More on the technical use of the CASSCF program.

The most elementary input information which is always required for CASSCF calculations is the specification of the number of active electrons and orbitals.

```
%casscf nel  4  # number of active space electrons
        norb 6  # number of active orbitals
        end
```

The CASSCF program in ORCA can average states of several multiplicities. The multiplicities are given as a list. For each multiplicity the number of roots should be specified:

```
%casscf mult   1,3 # here: multiplicities singlet and triplet

        nroots 4,2 # four singlets, two triplets
        end
```

If the symmetry handling in ORCA is enabled (`! UseSym`) each multiplicity block must have an irreducible representation assigned. Numbers corresponding to the "irrep" within a given symmetry are printed in the output of ORCA.

```
%casscf mult   1,3 # here: multiplicities singlet and triplet
        irrep  0,1 # here: irrep for each mult. block (mandatory!)
        nroots 4,2 # four singlets, two triplets
```

Several roots and multiplicities usually imply a state average CASSCF (SA-CASSCF) calculation. Note that the program by default chooses equal weights for the multiplicity blocks. Roots within a given block have equal weight. Users can define a custom weighting scheme for the multiplicity blocks and roots:

```
%casscf mult 1,3   # here: multiplicities singlet and triplet
        nroots 4,2  # four singlets, two triplets
        bweight 2,1 # singlets and triplets weighted 2:1
        weights[0] = 0.5,0.2,0.2,0.2 # singlet weights
        weights[1] = 0.7,0.3         # triplet weights
        end
```

The program automatically normalizes these weights such that the sum over all weights is unity. If convergence on an excited state is desired then the `weights[0]` array may look like `0.0,0.0,1.0` (this would optimize the orbitals for the third excited state. If several states cross during the orbital optimization this will ultimately cause convergence problems.

We note passing that the converged orbitals of the state averaged procedure are a compromise for the set of states. ORCA by default only prints the SA-CASSCF gradient norm. State-specific gradients are summarized at the end of the calculation with the keyword `PrintGState`.

```
%casscf
  ...
  printgstate true # optional printing of the state-specific orbital gradients
end
```

Orbital optimization methods. In the following we discuss the available options for orbital optimization. A number of convergence problems can be resolved changing the guess orbitals. **The following keywords are optional and should only be used facing severe convergence difficulties.** Aside from the `SuperCI_PT` (default),[459] several orbital optimization methods (list below) are implemented.

```
# Keywords to be used as Orbstep/Switchstep
SuperCI_PT  # perturbative SuperCI (first order)
SuperCI     # SuperCI (first order)
DIIS        # DIIS (first order)
KDIIS       # KDIIS (first order)
SOSCF       # approx. Newton-Raphson (first order)
NR          # augmented Hessian Newton-Raphson
            # unfolded two-step procedure
            # - still not true second order
```

The different convergers have different strengths. First order method are cheap but typically require more iterations compared to second order methods. When the gradient is far off from convergence the program uses the converger defined as `orbstep` while close to convergence the `switchstep` is used. The actual criteria for switchstep are defined with the keywords `SwitchConv` and `SwitchIter`.

```
%casscf
  OrbStep    SuperCI # or any other from the list above
  SwitchStep DIIS    # or any other from the list above

  SwitchConv 0.03 # gradient at which to switch
  SwitchIter 15   # iteration at which the switch takes place
                  # irrespective of the gradient

  MaxIter     75  # Maximum number of macro-iterations
end
```

Picking a convergence strategy, the program has to balance speed and robustness. The default strategy uses the `SuperCI_PT` as converger for `orbstep` and `switchstep`.[459] This approach determines the elements X_{pq} of the anti-Hermitean matrix used in the orbital update according to

$$\mathbf{C}^{new} = \mathbf{C}^{old} e^{\mathbf{X}}$$

from first order perturbation theory using the Dyall-Hamiltonian [238] in zeroth order and a first-order perturbed wave function given as $\Psi^{(1)} = \sum_{pq} \Psi_p^q X_{qp}$ where the Ψ_p^q represent singly excited functions obtained from the CASSCF wave function by excitation from orbital ψ_p to orbital ψ_q. The `SuperCI_PT` is robust with respect to orbitals that are exactly doubly occupied or empty. Rotations with orbital close to this critical occupations can further be eliminated with the keyword `DThresh` (default=1e-6). However, the method is quiet aggressive in the orbital optimization. In some cases, such as basis set projection or `PATOM` guess (intrinsic basis set projection), the program might pick a step-size that is too big. Then restricting the step-size via the keyword `MaxRot` (default=0.2) might be useful. The keywords `DThresh` and `MaxRot` described below are specific to `SuperCI_PT`. For many users, `MaxRot` is less palpable than level shifting. Therefore, the present version allows level shifts as well. **In contrast to other convergers, level shifts are not needed and highly discouraged.** With the exception of `GradScaling` (*vide infra*), other damping techniques described further below do not apply to the `SuperCI_PT`.

```
MaxRot 0.05  # cap stepsize for SuperCI_PT
DThresh 1e-6 # thresh for critical occupation
```

In case of convergence problems with the default settings, it is recommended to try the combination of `orbstep SuperCI` and `switchstep DIIS`, which in conjuction with a large level shift (2 Eh), which may be immediately successful. The proposed scheme typically requires more iterations. Moreoever, in contrast to the SuperCI(PT), the `SuperCI`, `DIIS` and `KDIIS` should not be used when the active orbitals have an occupation of exactly 2.0 or 0.0! The `DIIS` may sometimes converge slowly or "trail" towards the end such that real convergence is never reached. The `KDIIS` [452, 453] — based on perturbation theory — is an approximation to the regular `DIIS` procedure avoiding redundant rotations. Both DIIS schemes avoid linear dependencies in the expansion space.

```
MaxDIIS    15    # max. no of DIIS vectors to keep
DIISThresh 1e-7  # overlap criteria for linear dependency
```

The combination of `SuperCI` and `DIIS` (switchstep) is particularly suited to protect the active space composition. Adjusting the level shift will do the job. **Here, level shift is the single most important lever to control convergence**.

```
# default = dynamic level-shifting based on Ext-Act, Int-Act
ShiftUp      2.0   # static up-shift the virtual orbitals
ShiftDn      2.0   # static down-shift the internal orbitals
MinShift     0.6   # minimum separation subspaces
```

Level-shift is particularly important if the active, inactive and virtual orbitals overlap in their orbital energies. The energy separation of the subspaces is printed in the output. Ideally, the entries `Ext-Act` and `Act-Int` should be positive and larger than 0.2 Eh. This will help the program to preserve your active space composition throughout the iterations. If no shift is specified in the input, ORCA will choose a level-shift to guarantee an energy separation between the subspaces (`MinShift`).

```
E(CAS)=  -230.590325053 Eh DE=    -0.000798832
        --- Energy gap subspaces: Ext-Act = -0.244   Act-Int = -0.002
        --- current l-shift: Up(Ext-Act) = 0.54   Dn(Act-Int) = 0.30
```

In difficult cases the use of the Newton-Raphson method (`NR`) is recommended even if each individual iteration is considerably more expensive. It is strong towards the end but it would be a waste to start orbital optimization with the expensive `NR` method since its radius of quadratic convergence is quite small. The computationally cheaper alternative is the `SOSCF` procedure belonging to the family of quasi-Newton updates.

Keep in mind that the Newton-Raphson is designed for optimization on a **convex surface** (Hessian is semidefinite). If the NR is switched on too early, there is a good chance that this condition is not fulfilled. In this case the program will complain about negative eigenvalues or diagonal elements of the Hessian as can be seen in the snippet below. The next optimization step is large and unpredictable. It is a wildcard that can get you closer to convergence or immediate divergence of the CASSCF procedure.

```
||g|| =      0.771376945 Max(G)=     0.216712933 Rot=140,53
        --- Orbital Update [        NR]
        Warning:  NEGATIVE diagonal element D(81,53)=    -4.733590
        Warning:  NEGATIVE diagonal element D(82,53)=    -4.737955
        ...
```

For larger system, the augmented Hessian equations are solved iteratively (NR iterations). The augmented Hessian is considered solved when the residual norm, $< r|r >$, is small enough. Aside from the overall CASSCF convergence, negative eigenvalues affect these NR iterations.

```
--- Orbital Update [        NR]
        AugHess Tolerance (auto): Tol= 1.00e-07
        AUGHESS-ITER   0: E=    -0.174480747 <r|r>=      0.558679452
        AUGHESS-ITER   1: E=    -0.308672359 <r|r>=      0.468254671
        AUGHESS-ITER   2: E=    -0.434272813 <r|r>=      0.286305469
        AUGHESS-ITER   3: E=    -0.439149451 <r|r>=      0.286514628
        AUGHESS-ITER   4: E=    -0.605787445 <r|r>=      0.191691955
        AUGHESS-ITER   5: E=    -0.607766529 <r|r>=      0.310450670
        AUGHESS-ITER   6: E=    -0.611674930 <r|r>=      0.141402593
        AUGHESS-ITER   7: E=    -0.623145299 <r|r>=      0.394505306
        AUGHESS-ITER   8: E=    -0.658410333 <r|r>=      0.166915094
        AUGHESS-ITER   9: E=    -0.790571374 <r|r>=      4.722929453
        AUGHESS-ITER  10: E=    -0.790590554 <r|r>=      4.716012014
        AugHess: No convergence in the Davidson procedure
        ...
```

There are a number of refined NR settings that influence the convergence behavior on a non-convex energy surface. We mention the keywords for completeness and dis-encourage from changing the default settings. If overall convergence cannot be changed due to negative eigenvalues, it is recommended to delay the NR switchstep (`switchconv`, `switchiter`). This will require some trial and error, since the curvature of the surface is *a priori* not know.

```
%casscf
  ...
  aughess
    Solver   0       # Davidson (default)
```

(continues on next page)

(continued from previous page)

```
    1       # Pople (pure NR steps)
    2       # DIIS
    MaxIter 35       # max. no. of CI iters.
    MaxDim  35       # Davidson expansion space
    MaxDIIS  12      # max. number of DIIS vectors
    UseSubMatrixGuess true # diag a submatrix of the Hessian
    # as an initial guess
    NGuessMat 512   # size of initial guess matrix (part of
    # the Hessian exactly diagonalized)
    ExactDiagSwitch 512 # up to this dimension the Hessian
    # is exactly diagonalized (small problems)
    PrintLevel 1     # amount of output during AH iterations
    Tol    1e-6      # convergence tolerance
    Compress true    # use compressed storage
    DiagShift 0.0    # shift of the diagonal elements of the
    # Hessian
    UseDiagPrec true # use the diagonal in updating
    SecShift 1e-4     # shift the higher roots in the Davidson
    # secular equations
    UpdateShift 0.5  # shift of the denominator in the
    # update of the AH coefficients
  end
end
```

In general, convergence is strongly influenced by numerical noise, especially in the final iterations. One source of numerical noise is the incremental Fock build. Thus, it can help to enforce more frequent full Fock matrix formation.

```
ResetFreq 1 # reset frequency for direct SCF
```

If the orbital change in the active space is small, the active Fock matrix in ORCA is approximated using the density matrix from the previous cycle saving a second Fock matrix build. However, this approximation might also be a source of numerical instability. The threshold "SwitchDens" can be set to zero to enable the exact build. The program default starts with a rather large value (1e-2) and will reduce this parameter automatically when necessary.

```
switchdens  0.0001 # ~gtol * 0.1
```

In all of the implemented orbital optimization schemes the step-size correlates with the gradient-norm. A constant damping factor can be set with the keyword `GradScaling`. Note, damping and level shifting techniques are not recommended for the default converger (`SuperCI_PT`).

```
GradScaling  0.5   # constant damping in all steps
```

There are situations when the active space has been chosen carefully, but the initial gradient is still far off. To keep the "good" active space, we can suppress all rotation but the inactive-external ones until the gradient-norm is small enough to continue safely. The threshold can be set with `FreezeIE` keyword. Once the components of the gradient in the inactive-external direction have a weight of less than `FreezeIE`, all constraints are lifted. ORCA by default freezes active rotations if the total gradient norm is larger than 1.0 and the active rotations have a weight of less than 5%. The feature can be turned off setting the threshold to zero.

Similarly, rotations of the almost doubly occupied orbitals with the inactive orbitals can be damped using the threshold `FreezeActive`. Rotations of this type are damped as long as all their weight is smaller than `FreezeActive`. In contrast to the `ShiftDn`, it damps just the "troublesome" parts of internal-active rotations. This option applies to all of the orbital optimization schemes but the `SuperCI_PT`.

```
FreezeIE 0.4       # keep active space until int-ext rotation have
# a contribution of less than 40% to the ||g||
FreezeActive 0.03  # keep almost doubly occupied orbitals as long as
# their contribution is less than 3%  to the ||g||
```

If the calculation starts from a converged Hartree-Fock orbitals, the core orbitals should not change dramatically

by the CASSCF optimization. Often trailing convergence is associated to rotations with low lying orbitals. Their contribution to the total energy is fairly small. With the keyword `FreezeGrad` these rotations can be omitted from the orbital optimization procedure.

```
FreezeGrad   0.2  # omit  hitting a gradient norm ||g|| <0.2
```

The affected orbitals are printed at the startup of CASSCF.

```
FreezeGrad                          ... enabled if ||g|| is below 0.02
       Note Convergence can be signaled if the reduced gradient reaches GTol

       Last  frozen  orbital             ... 9
       First deleted orbital             ... 320
       Once rotations with core and deleted orbitals are stabilized they will be damped.
```

By default rotations with frozencore (or deleted virtuals) are not omitted. If the option FreezeGrad is active, the ratio with respect to the total gradient is printed.

```
||g|| =      0.001240414 Max(G)=    -0.000431747 Rot=319,1
       --- Option=FreezeGrad: ||g|| =      0.001081707
       = 87.21%
       Omitting frozencore elements
```

Using the RI Approximation.

Aside from the Fock matrices, integrals appearing in the orbital gradient and Hessian require substantial computation time. A good way to speed up the calculations at the expense of "only" obtaining approximate results is to introduce the RI approximation. `TrafStep RI` approximated the aforementioned integrals. Here are sufficiently large auxiliary basis must be provided - ideally a /JK or /C. Further acceleration can be achieved approximating the Fock matrix construction with `!RIJCOSX` or `!RIJK` as described in section *RI, RIJCOSX and RIJK approximations for CASSCF*. More details can also be found in the CASSCF tutorial. Note that with ORCA 4.1, there are three destinct auxiliary basis slots, that need to be set if the auxiliary basis is defined via the `%basis` block.

```
TrafoStep  RI    # RI used in transformation
                        # Note: Needs an auxiliary basis for
                        #       AuxC slot.
                  Exact # exact transformation (default)
```

Monitoring the active space

During the iterations, the **active orbitals are chosen to maximize the overlap** with active orbitals from the previous iterations. Maximizing the overlap does not make any restrictions on the nature of the orbitals. Thus initially localized orbitals will stay localized and ordered, which is sometimes a desired feature e.g. in the density matrix renormalization group approach (DMRG). This feature is set with the keyword `ActConstraints` and is enabled by default (after the first 3 macroiterations). For some orbital optimization procedures, such as the SuperCI, natural orbitals are more advantageous. Therefore, the `ActConstraints` can be turned off in favor of natural orbital construction (see below). If the keyword is not set by the user, ORCA picks the best choice for the given orbital optimization step.

```
ActConstraints  0 # no checks and no changes
                1 # maximize overlap of active orbitals and check sanity. (default for DIIS)
                2 # make natural orbitals in every iteration (default SuperCI)
                3 # make canonical orbitals in every iteration
                4 # localize orbitals
```

In addition to maximizing the overlap, **"`ActConstraints 1`" checks if the composition of the active space has changed** i.e. an orbital has been rotated out of the active space. In this case, ORCA aborts and stores the last valid set of orbitals. Below is an example error message.

```
--- Orbital Update [      DIIS]
       --- Failed to constrain active orbitals due to rotations:
```

(continues on next page)

(continued from previous page)

```
        Rot( 37, 35) with OVL=0.960986
        Rot( 38, 34) with OVL=0.842114
        Rot( 43,104) with OVL=0.031938
```

In the snippet above, the active space ranges from 37-43. The program reports that orbitals 37,38 and 43 have changed their character. The overlap of orbital 43 (active) with the previous set of active orbitals is just 3% and the program aborts. There are a number of reasons, why this happens in the calculation. If this error occurs constantly with the same orbitals, it is worthwhile to inspect the rotating partner orbitals (visualize). It might be sign that the active space is not balanced and should be extended. In many instances changing the level-shift or lowering switchconv is sufficient to protect the active space. In some cases, turning off the sanity check (`"ActConstraints 0"`) and re-rotating orbitals will bring CASSCF closer to convergence. Some problems can be avoided by a better design of the calculation. The CASSCF tutorial elaborates on the subject in more detail.

There are situations such as parameter scans, where "actconstraints" is counter-productive and should be disabled. In other words, we want to allow changes in the active space composition. As an example, consider the rotation of ethylene around its double-bond represented by a CAS(2,2). Although the active space consists of the bonding and anti-bonding orbitals π-orbitals, their composition in terms of atomic orbitals changes from the eclipsed to the staggered conformation. Depending on the actual input settings (orbstep and number of scan points), this might trigger an abort.

Final orbitals options.

Once the calculation has converged, ORCA will do a final macro-iteration, where the orbital are "finalized". For complete active spaces (CAS), these transformations do not alter the final energy and wavefunction. Note, that solutions from approximate CAS-CI solvers such as the ICE approach or the DMRG ansatz are affected by the final orbital transformation. The magnitude depends on the truncation level (e.g. `TGen`, `TVar` and `MaxM`) of the approximated wavefunction. The default final orbitals are canonical in the internal and external space with respect to state-averaged Fock operator. The active orbitals are chosen as natural orbitals. Other orbital choices are equally valid and can be selected for the individual subspaces.

```
#internal space
IntOrbs CanonOrbs # canonical
        LocOrbs   # localized
        unchanged # no changes

        # partner orbitals for the active space based
        # on various concepts
        PMOS      # based on orthogonalization tails.
        OSZ       # based on oszillator orbital
        DOI       # based on differential overlap

#external space
ExtOrbs CanonOrbs # canonical
        LocOrbs   # localized
        unchanged # no changes

        # partner orbitals for the active space based
        # on various concepts
        PMOS      # based on orthogonalization tails.
        OSZ       # based on oszillator orbital
        DOI       # based on differential overlap
        DoubleShell # based on the shell and angular momentum
                    # of the highest active orbitals, the first few
                    # virtual orbitals correspond to the doubled-shell.
                    # All other virt. orbitals are canonicalized.
                    # For 3d-metal complexes, these are the 4d orbitals!
                    # For 4d-metal complexes, these are the 5d orbitals!
                    # And so on...
#active space
ActOrbs NatOrbs   # natural
        CanonOrbs # canonical
```

(continues on next page)

(continued from previous page)

```
       LocOrbs   # localized
       unchanged # no changes
       dOrbs     # purify metal d-orbital and call the AILFT
       fOrbs     # purify metal f-orbital and call the AILFT
       SDO       # Single Determinant Orbitals: this is only possible if the
                 # active space has a single hole or a single electron.
                 # SDOs are then the unique choice of orbitals that simultaneously
                 # turns each CASSCF root into a single determinant.
```

SDOs are specific for the active orbital space.[491] The set of options (`PMOS, OSZ, DOI, DoubelShell`) are specific for the inactive and external space. They aim to assist the extension of the current active space. All four options, re-design the first `NOrb` (number of active orbitals) next to the active space, while the remaining orbitals of the same subspace are canonical. The re-designed orbital are based on different concepts.

- `PMOS` generates the bonding / anti-bonding partner orbitals for the chosen active space. It is based on the orthogonalization tail of the active orbitals.
- `OSZ` generates a single orbital for each active orbital, that maximizes the dipole-dipole interaction.
- `DOI` follows the same principle as OSZ, but the differential overlap is maximized instead.
- `DoubleShell` is specific to the external space. The highest active MO or `DoubleShellMO` is analyzed. A set of orbitals with the same angular momentum but larger radial distribution is generated.

Optionally, the four options above can be supplemented with a reference MO using the keyword `RefMO/DoubleShellMO`. The presence of `RefMO/DoubleShellMO` changes the default behavior. In case of `PMOS`, `OSZ` and `DOI`, all orbitals of the given subspace are chosen to maximize a single objective function with respect to the reference MO (must be active). This contrasts the default settings, where for each active MO an objective function is maximized and a single "best" orbital is picked. In other words, in the default setting, each active orbital has a corresponding "best" orbital in the selected subspace neighboring the active space.

```
RefMO 17 # MO with number 17 (default =-1)
       DoubleShellMO 17 # MO with number 17 (default=-1)
```

The aforementioned options are aids and the resulting orbitals should be inspected prior extension of the active space. In particular the `PMOS` option is useful in the context of transition metal complexes to find suitable Ligand based orbitals. There are more options (`dorbs, forbs, DoubleShell`), that are specifically designed for coordination chemistry. A more detailed description is found in the CASSCF tutorial that supplements the manual.

If the active space consists of a single set of metal d-orbitals, natural orbitals may be a mixture of the d-orbitals. The active orbitals are remixed to obtain "pure" d-orbitals (ligand field orbitals) if the `actorbs` is set to `dorbs`. The same holds for f-orbitals and the option `forbs`. Furthermore, the keyword `dorbs` automatically triggers the *ab initio* ligand field analysis (AILFT).[57, 490]The approach has been reported in a number of applications.[53, 56, 154, 169, 170, 428, 798, 836] Note that the actual representation depends on the chosen axis frame. It is thus recommended to align the molecule properly. For more details on the AILFT approach, we refer to the AILFT section (*1- and 2-shell Abinitio Ligand Field Theory*), the original paper and the CASSCF tutorial, where examples are shown. For a few applications, a printing of the complete wavefunction is useful and can be requested.

```
PrintWF 0     # (default) prints only the CFGs
       csf   # Printing of the wavefunction in the basis of CSFs
       det   # Printing of the wavefunction in the basis of Determinants
```

The CI-step default setting is CSF based and is done in the present program by generating a partial "formula tape" which is read in each CI iteration. The tape may become quite large beyond several hundred thousand CSFs which limits the applicability of the CASSCF module. The accelerated CI (ACCCI) has the same limitations, but uses a slightly different algorithm that handles multi-root calculations much more efficiently. For now, properties (spin-orbit coupling, g-Tensor...) as well as NEVPT2 corrections are not available with ACCCI. Nevertheless, it is the recommended option to converge a CASSCF calculation with multiple roots. The resulting .gbw file may be used in a successive run to obtain properties or NEVPT2 corrections.

Larger active spaces are tractable with the DMRG approach or the iterative configuration expansion (ICE) developed in our own group.[171, 172] DMRG and ICE return approximate full CI results. The maximum size of the

active space depends on the system and the required accuracy. Active spaces of 10–20 orbitals should be feasible with both approaches. The CASSCF tutorial covers examples with ACCCI and ICE as CI solvers.

```
%casscf
  CIStep CSFCI # CSF based CI (default)
  ACCCI  # CSF based CI solver with faster algorithm for multi-root calculations
  ICE    # CSF based approximate CI -> ICE/CIPSI algorithm
  DMRGCI # density matrix renormalization group approach instead of the CI
end
```

In the ICE approach, the computation of the coupling coefficients is time-consuming and by default repeated in every macro-iteration. To avoid the reconstruction, it is recommended to once generate a coupling coefficient library (`cclib`) and to use it for all of your ICE calculations. The details of the methodology and the `cclib` are described in the ICE section *Approximate Full CI Calculations in Subspace: ICE-CI*.

Detailed settings for the conventional CI solvers (CSFCI, ACCCI, ICE) can be controlled in a sub-block. **Not all of the options and properties are available for CISteps apart from the default!** NEVPT2, transition densities and spin-dependent properties such as spin-orbit coupling are not yet available for `ACCCI` and `ICE`.

```
%casscf ci
          MaxIter 64   # max. no. of CI iters.
          MaxDim  10    # Davidson expansion space = MaxDim * NRoots
          NGuessMat 512  # Initial guess matrix: 512x512
          PrintLevel 3 # amount of output during CI iterations
          ETol   1e-10    # default 0.1*ETol in CASSCF
          RTol   1e-10    # default 0.1*ETol in CASSCF
          TGen   1e-4     # ICE generator thresh
          TVar   1e-11    # ICE selection thresh, default = TGen*1e-7
          end
```

The `CI-step` DMRGCI interfaces to the **BLOCK** program developed in the group of G. K.-L. Chan [156, 157, 296, 788]. A detailed description of the **BLOCK** program, its input parameters, general information and examples on the density matrix renormalization group (DMRG) approach, are available in the section *Density Matrix Renormalization Group* of the manual.

The implementation of DMRG in **BLOCK** is fully spin-adapted. However, spin-densities and related properties are not available in the current version of the **BLOCK** code. To start a DMRG calculation add the keyword "`CIStep DMRGCI`" into a regular CASSCF input. ORCA will set default parameters and generate and input for the **BLOCK** program. In general, DMRG is not invariant to rotation in the active space. The program by **default** will run an automatic ordering procedure (`Fiedler`). More and refined options can be set in the `dmrg` sub-block of CASSCF — see section *Density Matrix Renormalization Group* for a complete list of keywords.

```
%casscf
  nel  8
  norb 6
  mult 3
  CIStep DMRGCI

  # Detailed settings
  dmrg
    # more/refined options
    ...
  end
end
```

It is highly recommended to start the calculation with split-localized orbitals. Any set of starting orbitals (gbw file) can be localized using the `orca_loc` program. Typing `orca_loc` in the shell will return a small help-file with details on how to setup an input for the localization. Examples for DMRG including the localization are in the corresponding section of the manual *Density Matrix Renormalization Group*. The utility program `orca_loc` is documented in section *orca_loc*. Split-localization refers to an independent localization of the internal and virtual part of the desired active orbitals.

NOTE:

- Let us stress again: it is strongly recommended to first LOOK at your orbitals and make sure that the ones that will enter the active space are really the ones that you want to be in the active space! Many problems can be solved by thinking about the desired physical contents of the reference space before starting a CASSCF. A poor choice of orbitals results in poor convergence or poor accuracy of the results! Choosing the active orbitals always requires chemical and physical insight into the molecules that you are studying!
- Please try the program with default settings before playing with the more advanced options. If you encounter convergence problems, have a look into your output, read the warning and see how the gradient and energy evolves. Increasing `MaxIter` will not help in many cases.
- Be careful with keywords such as `!tightscf`, `!verytightscf` and so on. These keywords set higher integral thresholds, which is a good idea, but also tighten the CASSCF convergence thresholds. If you do not need a tighter energy convergence, reset the criteria in the casscf block using `ETol`. For many applications an energy convergence beyond 10^{-7} is unnecessary.

7.15.2 CASSCF Densities

The one-particle electron and spin density can be stored on disk using the keyword `!KeepDens`. ORCA stores all densities in a container (**.densities** file on disk), which can be used in conjunction with `orca_plot` to plot the charge and spin densities. Please check Section *orca_plot* for more details on the procedure. The state-specific densities will have a name postfix that reflects the root, multiplicity and potentially irreducible representation of the state. Densities arising from a calculation with the spin-orbit coupling, will have an additional flag in the density container marking their origin (e.g. "cas_qdsoc" or "nev_qdsoc").

7.15.3 CASSCF Properties

The CASSCF program is able to calculate UV transition, CD spectra, SOC, SSC, Zeeman splittings, EPR g-matrices and A-matrices (the latter implemented in the same way as in the DCD-CAS(2) method[491]), magnetization, magnetic susceptibility and MCD spectra. Note that the results for the Fermi contact contribution to A will not be reliable if the spin density is dominated by spin polarization, which is a dynamic correlation effect. The properties are exercised in more detail in the CASSCF tutorial. The techniques used to calculate SOC, and Zeeman splittings are identical to those implemented into the MRCI program. Input and keywords mimic the ones in the MRCI module described in section *Properties Calculation Using the SOC Submodule*. As an example, the input file to calculate g-values and HFC constants A of CO^+ is listed below:

```
!TZVPP  Bohrs TightSCF #TightSCF for more accurate integrals
%casscf nel  9
        norb 8
        nroots 9
        mult   2
        switchstep NR
        etol 1e-7 #reset energy convergence
        rel
            dosoc true   #spin-orbit coupling (and ZFS)
            gtensor true
            amatrix true
        end
        end
* xyz 1 2
 C 0 0  0.0
 O 0 0  2.3504
*
```

In addition to pseudo-spin 1/2 A-tensors for individual Kramers doublets, the CASSCF module also features the calculation of "intrinsic" HFC A-tensors for the whole lowest nonrelativistic spin multiplet in the effective Hamiltonian approach.[489]

In contrast to the MRCI module, the CASSCF module also supports the calculation of susceptibility tensors at non-zero magnetic fields. The corresponding keywords are

```
...
%casscf
  ...
  rel
    dosoc true
    dosusceptibility true
    susctensor_nfields 2                    # number of user-defined magnetic fields
    susctensor_magfields[0] = 35000,0,0     # 1st user-defined magnetic field
    susctensor_magfields[1] = 70000,0,0     # 2nd user-defined magnetic field
  end
end
```

This example input calculates the susceptibility tensor at the two (vector-valued!) magnetic fields (35000,0,0) and (70000,0,0) (in Gauss). Note that for practical reasons it is necessary to specify the number of user-defined magnetic fields using the keyword `susctensor_nfields`.

Until ORCA 4.0 it was possible to access spin-spin couplings only via running CAS-CI type calculations in MRCI. Converged CASSCF orbitals can be read setting the following flags

```
!MOREAD NOITER ALLOWRHF TZVPP  TightSCF Bohrs
%moinp "convergedCASSCF.gbw"

%mrci
  ...
  TPre 0.0
  citype mrci

  newblock 2 *
    excitations none
    refs CAS(9,8) end
  end

  soc
    DoSSC true # spin-spin coupling
    DoSOC true # spin-orbit coupling
    ...
  end
end
* xyz 1 2
 C 0 0  0.0
 O 0 0  2.3504
*
```

Starting with ORCA 4.1, spin-spin couplings are also directly accessible in the CASSCF module via the keyword `DoSSC true` in the `rel` subblock. Note that the calculation of SSC requires the definition of an auxiliary basis set (`AuxC` auxiliary basis set slot), since it is only implemented in conjunction with RI integrals. A common way to introduce dynamical correlation for the property computation, is to replace the energies entering the quasi-degenerate perturbation theory. If the NEVPT2 energy correction is computed in CASSCF, there will be additional printings where CASSCF energies are replaced by the more accurate NEVPT2 values. Alternatively, these diagonal energies can be taken from the input file similarly how it is described for the MRCI module. A more detailed documentation is presented in the MRCI property section.

Note

- The program does NOT print the SOC matrix by default! To obtain SOCMEs at the CASSCF/NEVPT2/... levels, please set the `PrintLevel` in the `rel` block to at least 2.

7.15.4 1- and 2-shell Abinitio Ligand Field Theory

Starting from ORCA 5.0, ORCA features a 1- and 2-shell AILFT module. AILFT was originally developed for 1-shell d- and f- LFT problems [57, 490]. In ORCA 5.0 an extenion to 2-shell AILFT provides access to all common 1- and 2-shell AILFT problems namely:

1. Valence LFT problems, involving the d-, f-, sp-, ds- and df-shells
2. Core LFT problems involving the sd-, pd-, sf- and pf-shells become readily accesible.

Requesting an CAS-AILFT calcultion withing the CASSCF module is provided in two ways:

1. Through the ActOrbs xOrbs keywords (e.g. xOrbs: dOrbs, fOrbs spOrbs, psOrbs, sdOrbs, dsOrbs, sfOrbs, fsOrbs, pdOrbs, dpOrbs, pfOrbs, fpOrbs, dfOrbs, fdOrbs)
2. Through the LFTCase keyword where particular LFT problems can be requested according to the above 1- and 2-shell combinations (e.g. LFTCase 3d, LFTCase 4f, LFTCase 1s3d, LFTCase 2p3d ...)

Note: that the LFTCase keyword overwrites the ActOrbs keyword and as it will be discussed below provides a particular utility that simplifies the 2-shell AILFT input.

A simple input for the Ni^{2+} d^8 ion is provided below:

```
!NEVPT2 def2-SVP def2-SVP/C
%casscf
 nel 8
 norb 5
 ActOrbs dOrbs
 mult 3,1
 nroots 10,15
 rel
  dosoc true
 end
end

*xyz 2 3
Ni      0.0000000000      0.0000000000      0.0000000000
*
```

The programm after the CASSCF convergence will undergo few important steps and sanity checks which involve

1. an Orbital purification step
2. a Phase correction of the 1 and 2-electron integrals

It is then important from the user's perspective to monitor that these steps have been succesfully performed. The relevant parts of the output are provided below:

```
---- THE CAS-SCF GRADIENT HAS CONVERGED ----
--- FINALIZING ORBITALS ---
---- DOING ONE FINAL ITERATION FOR PRINTING ----
--- d-orbitals (depends on the molecular axis frame)
L-Center:  0 Ni [0.000, 0.000, 0.000]
--- The active space contains 5 d Orbitals : OK
Setting 9 active MO to AO dz2     (11)
Setting 10 active MO to AO dxz     (12)
Setting 11 active MO to AO dyz     (13)
Setting 12 active MO to AO dx2y2   (14)
Setting 13 active MO to AO dxy     (15)
--- Canonicalize Internal Space
--- Canonicalize External Space

...
```

(continues on next page)

(continued from previous page)

```
============================
AB INITIO LIGAND FIELD THEORY
d8 configuration
2 CI blocks
MOs 9 to 13
============================

Metal/Atom center is atom 0
orbital phases =   1.0   1.0   1.0   1.0   1.0
Metal/Atom d-orbital parts of active orbitals
Shell 7
9        10        11        12        13
dz2     :  0.848522 -0.000000  0.000000 -0.000000 -0.000000
dxz     : -0.000000  0.848522  0.000000 -0.000000  0.000000
dyz     :  0.000000  0.000000  0.848522 -0.000000 -0.000000
dx2y2   : -0.000000 -0.000000 -0.000000  0.848522  0.000000
dxy     : -0.000000  0.000000 -0.000000  0.000000  0.848522
Shell 8
9        10        11        12        13
dz2     :  0.300072  0.000000 -0.000000  0.000000  0.000000
dxz     :  0.000000  0.300072 -0.000000  0.000000 -0.000000
dyz     : -0.000000 -0.000000  0.300072  0.000000  0.000000
dx2y2   :  0.000000  0.000000  0.000000  0.300072 -0.000000
dxy     :  0.000000 -0.000000  0.000000 -0.000000  0.300072

Adjusting phases of one-electron integrals               ... done
Adjusting phases of two-electron integrals               ... done
```

In a subsequent step the program will

1. compute the AI Hamiltonian
2. construct the parameterized LFT Hamiltonian
3. and perform the fit

The relevant output can be seen below:

```
Calculating ab initio Hamiltonian matrices               ...
------------------------------------------------------------
NRoots (NEVPT2) for this block = 10
NEVPT2 correction for this block is calculated
Full NEVPT2 Hamiltonian constructed
Full NEVPT2 Hamiltonian diagonalized
------------------------------------------------------------
------------------------------------------------------------
NRoots (NEVPT2) for this block = 15
NEVPT2 correction for this block is calculated
Full NEVPT2 Hamiltonian constructed
Full NEVPT2 Hamiltonian diagonalized
------------------------------------------------------------
Calculating fit matrices                                 ... done
Fitting                                                  ... done
```

In following the fitted 1-electron energies and SCP parameters also Racah parameters for 1-shells will be printed at the CASSCF and NEVPT2 levels of theory

```
-------------------------------
AILFT MATRIX ELEMENTS (CASSCF)
-------------------------------

Ligand field one-electron matrix VLFT (a.u.) :
Orbital          dz2        dxz        dyz        dx2-y2     dxy
```

(continues on next page)

(continued from previous page)

```
dz2         -8.111733  -0.000000  -0.000000   0.000000  -0.000000
dxz         -0.000000  -8.111733  -0.000000  -0.000000   0.000000
dyz         -0.000000  -0.000000  -8.111733  -0.000000   0.000000
dx2-y2       0.000000  -0.000000  -0.000000  -8.111733  -0.000000
dxy         -0.000000   0.000000   0.000000  -0.000000  -8.111733

------------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
------------------------------------------------
F0dd(from 2el Ints)     =  0.980960738 a.u. =    26.693 eV =  215296.0 cm**-1 (fixed)
F2dd                    =  0.451725025 a.u. =    12.292 eV =   99142.2 cm**-1
F4dd                    =  0.280604669 a.u. =     7.636 eV =   61585.6 cm**-1
-------------------
Racah Parameters :
-------------------
A(F0dd from 2el Ints)   =  0.949782441 a.u. =    25.845 eV =  208453.2 cm**-1
B                       =  0.006037419 a.u. =     0.164 eV =    1325.1 cm**-1
C                       =  0.022270212 a.u. =     0.606 eV =    4887.7 cm**-1
C/B                     =  3.689
------------------------------------------------------------------------------


----------------------------------------
The ligand field one electron eigenfunctions:
----------------------------------------
Orbital    Energy (eV)  Energy(cm-1)      dz2         dxz         dyz         dx2-y2      dxy
1          0.000        0.0         -0.999978  -0.000164  -0.001934   0.005783  -0.002568
2          0.000        0.0         -0.005768  -0.000269  -0.000262  -0.999967  -0.005788
3          0.000        0.0          0.002600   0.000424   0.001046   0.005773  -0.999979
4          0.000        0.0          0.000241  -0.999246  -0.038831   0.000280  -0.000462
5          0.000        0.0          0.001930   0.038832  -0.999243   0.000246  -0.001022
Ligand field orbitals were stored in ni.3d.casscf.lft.gbw

...

-------------------------------
AILFT MATRIX ELEMENTS (NEVPT2)
-------------------------------

Ligand field one-electron matrix VLFT (a.u.) :
Orbital          dz2         dxz         dyz         dx2-y2      dxy
dz2         -8.118685   0.000000   0.000000   0.000005  -0.000000
dxz          0.000000  -8.118666  -0.000000  -0.000000   0.000000
dyz          0.000000  -0.000000  -8.118674  -0.000000   0.000000
dx2-y2       0.000005  -0.000000  -0.000000  -8.118676   0.000000
dxy         -0.000000   0.000000   0.000000   0.000000  -8.118667

------------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
------------------------------------------------
F2dd                    =  0.415943380 a.u. =    11.318 eV =   91289.0 cm**-1
F4dd                    =  0.259145554 a.u. =     7.052 eV =   56875.9 cm**-1
-------------------
Racah Parameters :
-------------------
B                       =  0.005550482 a.u. =     0.151 eV =    1218.2 cm**-1
C                       =  0.020567107 a.u. =     0.560 eV =    4514.0 cm**-1
C/B                     =  3.705
------------------------------------------------------------------------------


----------------------------------------
The ligand field one electron eigenfunctions:
```

(continues on next page)

(continued from previous page)

```
----------------------------------------
Orbital    Energy (eV)  Energy(cm-1)      dz2        dxz        dyz        dx2-y2     dxy
1          0.000        0.0         -0.927589   0.002893   0.009447   0.373452  -0.003963
2          0.000        3.0          0.337599   0.005897   0.449370   0.827017  -0.010128
3          0.000        3.0          0.160011  -0.005672  -0.893243   0.420094   0.001383
4          0.001        4.4          0.000644   0.105816  -0.006612  -0.009602  -0.994317
5          0.001        4.6          0.001541   0.994348  -0.007084  -0.002573   0.105893
Ligand field orbitals were stored in ni.3d.nevpt2.lft.gbw
```

Note that:

- At the CASSCF level F0 (and subsequently racah A) is computed from CASSCF 2-electron coulomb integrals
- On the other hand at the NEVPT2 level F0 is not defined hence F0 and racah A are not printed. Below an alternative using the effective Slater exponnets will be provided.
- The LFT orbitals are saved in *.lft.gbw files which can be processed by the orca_plot to generate orbital visualization files.

AILFT provides a Fit quality analysis (see the original paper [57, 490])

Note: That at the CASSCF level the AI matrix of free atoms and ions is exactly parameterized in the chosen LFT parameterization scheme. As a result the RMS AI-LFT fitting errors is expected to be practically zero. This is not the case when a correlation treatment is chosen like NEVPT2 and the errors are expected to be somewhat larger.

The above is shown below:

```
Calculating statistical parameters                    ... done

Reference energy AI-LFT =    -38.134150221 au
Reference energy AI     =    -38.134150221 au

-----------------------------------------------
COMPARISON OF AB INITIO AND LIGAND FIELD RESULTS
-----------------------------------------------

Block   1
---------
AI-Root   0: E(AI)=    0.000 eV -> LF-Root   0:     0.000 eV    S= 0.998 Delta=   -0.000 eV
AI-Root   1: E(AI)=    0.000 eV -> LF-Root   1:     0.000 eV    S= 0.981 Delta=   -0.000 eV
AI-Root   2: E(AI)=    0.000 eV -> LF-Root   2:     0.000 eV    S= 0.980 Delta=   -0.000 eV
AI-Root   3: E(AI)=    0.000 eV -> LF-Root   3:     0.000 eV    S= 0.773 Delta=    0.000 eV
AI-Root   4: E(AI)=    0.000 eV -> LF-Root   4:     0.000 eV    S= 0.774 Delta=   -0.000 eV
AI-Root   5: E(AI)=    0.000 eV -> LF-Root   5:     0.000 eV    S= 0.985 Delta=   -0.000 eV
AI-Root   6: E(AI)=    0.000 eV -> LF-Root   6:     0.000 eV    S= 0.986 Delta=   -0.000 eV
AI-Root   7: E(AI)=    2.464 eV -> LF-Root   7:     2.464 eV    S= 0.998 Delta=   -0.000 eV
AI-Root   8: E(AI)=    2.464 eV -> LF-Root   8:     2.464 eV    S= 0.998 Delta=   -0.000 eV
AI-Root   9: E(AI)=    2.464 eV -> LF-Root   9:     2.464 eV    S= 1.000 Delta=   -0.000 eV
RMS error for this block =     0.000 eV =       0.0 cm**-1

Block   2
---------
AI-Root   0: E(AI)=    2.033 eV -> LF-Root   0:     2.033 eV    S= 1.000 Delta=   -0.000 eV
AI-Root   1: E(AI)=    2.033 eV -> LF-Root   1:     2.033 eV    S= 1.000 Delta=   -0.000 eV
AI-Root   2: E(AI)=    2.033 eV -> LF-Root   2:     2.033 eV    S= 0.903 Delta=   -0.000 eV
AI-Root   3: E(AI)=    2.033 eV -> LF-Root   3:     2.033 eV    S= 0.967 Delta=   -0.000 eV
AI-Root   4: E(AI)=    2.033 eV -> LF-Root   4:     2.033 eV    S= 0.935 Delta=   -0.000 eV
AI-Root   5: E(AI)=    3.183 eV -> LF-Root   5:     3.183 eV    S= 0.996 Delta=   -0.000 eV
AI-Root   6: E(AI)=    3.183 eV -> LF-Root   6:     3.183 eV    S= 0.999 Delta=   -0.000 eV
AI-Root   7: E(AI)=    3.183 eV -> LF-Root   7:     3.183 eV    S= 0.996 Delta=   -0.000 eV
AI-Root   8: E(AI)=    3.183 eV -> LF-Root   8:     3.183 eV    S= 0.999 Delta=   -0.000 eV
AI-Root   9: E(AI)=    3.183 eV -> LF-Root   9:     3.183 eV    S= 0.999 Delta=   -0.000 eV
AI-Root  10: E(AI)=    3.183 eV -> LF-Root  10:     3.183 eV    S= 0.995 Delta=   -0.000 eV
```

(continues on next page)

(continued from previous page)

```
AI-Root  11: E(AI)=    3.183 eV -> LF-Root  11:     3.183 eV    S= 0.992 Delta=   -0.000 eV
AI-Root  12: E(AI)=    3.183 eV -> LF-Root  12:     3.183 eV    S= 0.999 Delta=   -0.000 eV
AI-Root  13: E(AI)=    3.183 eV -> LF-Root  13:     3.183 eV    S= 0.996 Delta=   -0.000 eV
AI-Root  14: E(AI)=    7.856 eV -> LF-Root  14:     7.856 eV    S= 1.000 Delta=   -0.000 eV
RMS error for this block =     0.000 eV =       0.0 cm**-1

Total RMS error g=     0.000 eV =       0.0 cm**-1
Note: Dropped RMS error for the reference energy.

Confidence intervals (95) computed from the square root of the
diagonal elements of the covariance matrix:
H(dz2   ,dz2   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dxz   ,dz2   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dxz   ,dxz   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dyz   ,dz2   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dyz   ,dxz   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dyz   ,dyz   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dx2-y2,dz2   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dx2-y2,dxz   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dx2-y2,dyz   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dx2-y2,dx2-y2)=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dxy   ,dz2   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dxy   ,dxz   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dxy   ,dyz   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dxy   ,dx2-y2)=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
H(dxy   ,dxy   )=  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
F2dd            =  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
F4dd            =  0.000000000 a.u. =     0.000 eV =       0.0 cm**-1
Pearson's correlation coefficient =    1.000 (should be close to 1)
```

```
Calculating statistical parameters                    ... done

Reference energy AI-LFT =    -38.134345028 au
Reference energy AI     =    -38.134150221 au

-------------------------------------------------
COMPARISON OF AB INITIO AND LIGAND FIELD RESULTS
-------------------------------------------------

Block   1
---------
AI-Root   0: E(AI)=   -0.000 eV -> LF-Root   0:     0.000 eV    S= 0.933 Delta=   -0.000 eV
AI-Root   1: E(AI)=    0.000 eV -> LF-Root   1:     0.000 eV    S= 0.769 Delta=   -0.000 eV
AI-Root   2: E(AI)=    0.001 eV -> LF-Root   2:     0.000 eV    S= 0.773 Delta=    0.000 eV
AI-Root   3: E(AI)=    0.001 eV -> LF-Root   3:     0.000 eV    S= 0.742 Delta=    0.001 eV
AI-Root   4: E(AI)=    0.002 eV -> LF-Root   4:     0.000 eV    S= 0.750 Delta=    0.001 eV
AI-Root   5: E(AI)=    0.003 eV -> LF-Root   5:     0.001 eV    S= 0.931 Delta=    0.002 eV
AI-Root   6: E(AI)=    0.003 eV -> LF-Root   6:     0.001 eV    S= 0.998 Delta=    0.003 eV
AI-Root   7: E(AI)=    2.195 eV -> LF-Root   7:     2.266 eV    S= 1.000 Delta=   -0.070 eV
AI-Root   8: E(AI)=    2.195 eV -> LF-Root   8:     2.266 eV    S= 0.998 Delta=   -0.070 eV
AI-Root   9: E(AI)=    2.195 eV -> LF-Root   9:     2.266 eV    S= 0.998 Delta=   -0.070 eV
RMS error for this block =     0.039 eV =     311.1 cm**-1

Block   2
---------
AI-Root   0: E(AI)=    1.812 eV -> LF-Root   0:     1.875 eV    S= 0.825 Delta=   -0.063 eV
AI-Root   1: E(AI)=    1.812 eV -> LF-Root   1:     1.875 eV    S= 0.938 Delta=   -0.063 eV
AI-Root   2: E(AI)=    1.812 eV -> LF-Root   2:     1.875 eV    S= 1.000 Delta=   -0.063 eV
AI-Root   3: E(AI)=    1.812 eV -> LF-Root   3:     1.875 eV    S= 1.000 Delta=   -0.063 eV
AI-Root   4: E(AI)=    1.812 eV -> LF-Root   4:     1.875 eV    S= 0.773 Delta=   -0.063 eV
AI-Root   5: E(AI)=    2.987 eV -> LF-Root   5:     2.932 eV    S= 0.955 Delta=    0.056 eV
```

(continues on next page)

(continued from previous page)

```
AI-Root   6: E(AI)=    2.987 eV -> LF-Root   6:     2.932 eV    S= 0.910 Delta=    0.055 eV
AI-Root   7: E(AI)=    2.988 eV -> LF-Root   7:     2.932 eV    S= 0.874 Delta=    0.056 eV
AI-Root   8: E(AI)=    2.988 eV -> LF-Root   8:     2.932 eV    S= 0.792 Delta=    0.056 eV
AI-Root   9: E(AI)=    2.988 eV -> LF-Root   9:     2.932 eV    S= 0.796 Delta=    0.056 eV
AI-Root  10: E(AI)=    2.988 eV -> LF-Root  10:     2.932 eV    S= 0.808 Delta=    0.056 eV
AI-Root  11: E(AI)=    2.988 eV -> LF-Root  11:     2.932 eV    S= 0.971 Delta=    0.056 eV
AI-Root  12: E(AI)=    2.988 eV -> LF-Root  12:     2.932 eV    S= 0.994 Delta=    0.056 eV
AI-Root  13: E(AI)=    2.989 eV -> LF-Root  13:     2.932 eV    S= 0.994 Delta=    0.057 eV
AI-Root  14: E(AI)=    7.122 eV -> LF-Root  14:     7.241 eV    S= 1.000 Delta=   -0.119 eV
RMS error for this block =     0.064 eV =     519.6 cm**-1

Total RMS error g=     0.057 eV =     457.2 cm**-1
Note: Dropped RMS error for the reference energy.

Confidence intervals (95) computed from the square root of the
diagonal elements of the covariance matrix:
H(dz2   ,dz2   )=  0.000523387 a.u. =     0.014 eV =     114.9 cm**-1
H(dxz   ,dz2   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dxz   ,dxz   )=  0.000523387 a.u. =     0.014 eV =     114.9 cm**-1
H(dyz   ,dz2   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dyz   ,dxz   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dyz   ,dyz   )=  0.000523387 a.u. =     0.014 eV =     114.9 cm**-1
H(dx2-y2,dz2   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dx2-y2,dxz   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dx2-y2,dyz   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dx2-y2,dx2-y2)=  0.000523387 a.u. =     0.014 eV =     114.9 cm**-1
H(dxy   ,dz2   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dxy   ,dxz   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dxy   ,dyz   )=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dxy   ,dx2-y2)=  0.000393965 a.u. =     0.011 eV =      86.5 cm**-1
H(dxy   ,dxy   )=  0.000523387 a.u. =     0.014 eV =     114.9 cm**-1
F2dd            =  0.002351095 a.u. =     0.064 eV =     516.0 cm**-1
F4dd            =  0.003038264 a.u. =     0.083 eV =     666.8 cm**-1
Pearson's correlation coefficient =    1.000 (should be close to 1)
```

Several utilities are offered for more specialized tasks that provide better control of the AILFT inputs and outputs:

- Skipping orbital otimization or reading in previously computed orbitals can be requested in two ways:
  1. By the !NoIter keyword in the command line
  2. By the AILFT_SkipOrbOpt in the ailft block (see example below)
- Estimating F0 SCPs or Racah A from single zeta Slater Exponents can be requested from the AILFT_EffectveSlaterExponents true keyword in the ailft block
- For the above task the knowledge of the principle quantum numbers is required. This can be specidied in two ways:
  1. By the AILFT_PQN x keyword in the ailft block (x=3 for 3d)
  2. By the LFTCase x keyword (LFTCase 3d, ommiting in this way the ActOrbs dOrbs keyword)

Let us see how all the above translates in the above example of the Ni^{2+} d^8 ion

```
!NoIter NEVPT2 def2-SVP def2-SVP/C

%casscf
 nel 8
 norb 5
 LFTCase 3d
 mult 3,1
 nroots 10,15
 ailft
```

(continues on next page)

(continued from previous page)

```
  #AILFT_SkipOrbOpt true (An alternative to NoIter)
  #AILFT_PQNs 3 (Works together with ActOrbs dOrbs as an alternative to LFTCase 3d)
  AILFT_SlaterExponents true
 end
 rel
  dosoc true
 end
end

*xyz 2 3
Ni      0.0000000000      0.0000000000      0.0000000000
*
```

By running the above input the fitted 1-electron energies and SCP parameters also Racah parameters for 1-shells will be printed at the CASSCF and NEVPT2 levels of theory, including F0s and Racah A as estimated from single zeta effective Slater exponents from the fitted F2dd SCPs.

```
------------------------------
AILFT MATRIX ELEMENTS (CASSCF)
------------------------------

Ligand field one-electron matrix VLFT (a.u.) :
Orbital          dz2        dxz        dyz        dx2-y2     dxy
dz2         -7.974440   0.000000  -0.000000   0.000000  -0.000000
dxz          0.000000  -7.974440  -0.000000  -0.000000   0.000000
dyz         -0.000000  -0.000000  -7.974440  -0.000000   0.000000
dx2-y2       0.000000  -0.000000  -0.000000  -7.974440   0.000000
dxy         -0.000000   0.000000   0.000000   0.000000  -7.974440

------------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
------------------------------------------------
--------------------------------------------------------------------------------
Computed Single Zeta Slater Effective Exponents                         ...
--------------------------------------------------------------------------------
kd(SZ)(from F2dd)       =  3.134434893 a.u.
--------------------------------------------------------------------------------
Computed F0s from Single Zeta Slater Effective Exponents                ...
--------------------------------------------------------------------------------
F0dd(from F2dd kd(SZ))  =  0.809116820 a.u. =    22.017 eV =  177580.6 cm**-1
F2dd                    =  0.427107567 a.u. =    11.622 eV =   93739.3 cm**-1
F4dd                    =  0.264174069 a.u. =     7.189 eV =   57979.5 cm**-1
--------------------------------------------------------------------------------
-------------------
Racah Parameters :
-------------------
A(from F2dd kd(SZ))     =  0.779764145 a.u. =    21.218 eV =  171138.4 cm**-1
B                       =  0.005721310 a.u. =     0.156 eV =    1255.7 cm**-1
C                       =  0.020966196 a.u. =     0.571 eV =    4601.5 cm**-1
C/B                     =  3.665
--------------------------------------------------------------------------------

----------------------------------------
The ligand field one electron eigenfunctions:
----------------------------------------
Orbital    Energy (eV)  Energy(cm-1)      dz2        dxz        dyz        dx2-y2     dxy
1          0.000        0.0        -0.999981   0.000787  -0.001735   0.004390  -0.003810
2          0.000        0.0         0.003811   0.000493   0.000955   0.000497  -0.999992
3          0.000        0.0         0.004388   0.000169   0.000080   0.999990   0.000514
4          0.000        0.0        -0.000750  -0.999810  -0.019460   0.000174  -0.000514
5          0.000        0.0        -0.001754  -0.019459   0.999809  -0.000069   0.000939
```

(continues on next page)

(continued from previous page)

```
Ligand field orbitals were stored in ni.3d.casscf.lft.gbw

...

------------------------------
AILFT MATRIX ELEMENTS (NEVPT2)
------------------------------

Ligand field one-electron matrix VLFT (a.u.) :
Orbital          dz2         dxz         dyz        dx2-y2      dxy
dz2         -7.974812   -0.000000   -0.000000    0.000002   -0.000000
dxz         -0.000000   -7.974860    0.000000    0.000000    0.000000
dyz         -0.000000    0.000000   -7.974856   -0.000000    0.000000
dx2-y2       0.000002    0.000000   -0.000000   -7.974837    0.000000
dxy         -0.000000    0.000000    0.000000    0.000000   -7.974837

------------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
------------------------------------------------
--------------------------------------------------------------------------------
Computed Single Zeta Slater Effective Exponents                          ...
--------------------------------------------------------------------------------
kd(SZ)(from F2dd)       =  3.126330433 a.u.
--------------------------------------------------------------------------------
Computed F0s from Single Zeta Slater Effective Exponents                 ...
--------------------------------------------------------------------------------
F0dd(from F2dd kd(SZ))  =  0.807024751 a.u. =    21.960 eV =  177121.5 cm**-1
F2dd                    =  0.426003229 a.u. =    11.592 eV =   93496.9 cm**-1
F4dd                    =  0.262001371 a.u. =     7.129 eV =   57502.7 cm**-1
--------------------------------------------------------------------------------
-------------------
Racah Parameters :
-------------------
A(from F2dd kd(SZ))     =  0.777913487 a.u. =    21.168 eV =  170732.3 cm**-1
B                       =  0.005723406 a.u. =     0.156 eV =    1256.1 cm**-1
C                       =  0.020793760 a.u. =     0.566 eV =    4563.7 cm**-1
C/B                     =  3.633
--------------------------------------------------------------------------------


----------------------------------------
The ligand field one electron eigenfunctions:
----------------------------------------
Orbital    Energy (eV)  Energy(cm-1)      dz2         dxz         dyz         dx2-y2      dxy
1          0.000        0.0         -0.000608   -0.999795    0.017518    0.010019    0.001244
2          0.000        0.9          0.000003   -0.017532   -0.999696   -0.003661    0.016920
3          0.001        4.9          0.065092   -0.009854    0.004898   -0.995782    0.063719
4          0.001        5.0         -0.004126    0.002173    0.016617    0.063640    0.997824
5          0.001       10.5          0.997871    0.000042   -0.000237    0.065225   -0.000030
Ligand field orbitals were stored in ni.3d.nevpt2.lft.gbw
```

It is also possible to treat only the High Spin states in the d- and f- 1-shell AILFT. Note that not all the cases can be treated as this renters the different SCP parameters undefined. In the beggining, AILFT will check whether such case is detected and will drop a warning message

For example in the case of the Fe^{2+} d^6 ion with an input like the following:

```
!NoIter def2-SVP def2-SVP/C


%casscf
 nel 6
```

(continues on next page)

(continued from previous page)

```
 norb 5
 mult 5
 nroots 5
 LFTCase 3d
end

*xyz 2 3
Fe      0.0000000000      0.0000000000      0.0000000000
*
```

the following Warning will be printed in the beggining of the calculation

```
WARNING: In AILFT F2dd remains undefined when considering only the HS Multiplicity block
Be Careful with the results!
TIP:     If possible use in addition a LS Multiplicity Block
```

Spin orbit coupling effects (SOC) can be introduced by parametrizing the effective SOC constant ζ. As long as SOC is requested in the rel CASSCF block the respective requested shell effective SOC constant ζ will be computed at the end of every AILFT calculation

Hence in the above examples one gets:

```
-------------------------------------------
SPIN ORBIT COUPLING (based on CASSCF orbitals)
-------------------------------------------

AI-SOC-X integrals (cm-1)
0          1          2          3          4
0       0.000000   0.000000 1078.568273  -0.000000  -0.000000
1      -0.000000   0.000000  -0.000000  -0.000000 622.711683
2   -1078.568273   0.000000   0.000000 -622.711683   0.000000
3       0.000000   0.000000 622.711683   0.000000  -0.000000
4       0.000000 -622.711683  -0.000000   0.000000   0.000000
AI-SOC-Y integrals (cm-1)
0          1          2          3          4
0      -0.000000 -1078.568273  -0.000000   0.000000   0.000000
1    1078.568273   0.000000  -0.000000 -622.711683  -0.000000
2       0.000000   0.000000   0.000000  -0.000000 -622.711683
3      -0.000000 622.711683   0.000000   0.000000  -0.000000
4      -0.000000   0.000000 622.711683   0.000000   0.000000
AI-SOC-Z integrals (cm-1)
0          1          2          3          4
0       0.000000  -0.000000   0.000000  -0.000000   0.000000
1       0.000000  -0.000000 -622.711683  -0.000000   0.000000
2      -0.000000 622.711683   0.000000  -0.000000  -0.000000
3       0.000000   0.000000   0.000000  -0.000000 -1245.423365
4      -0.000000  -0.000000   0.000000 1245.423365   0.000000

Fit to the SOC matrix elements
a = 15.000000
b =     1.158 eV =    9340.7 cm**-1
SOC constant zeta =     0.077 eV =     622.7 cm**-1


LF-SOC-X integrals (cm-1)
0          1          2          3          4
0       0.000000  -0.000000 1078.568273  -0.000000  -0.000000
1       0.000000   0.000000  -0.000000  -0.000000 622.711683
2    -1078.568273   0.000000   0.000000 -622.711683  -0.000000
3       0.000000   0.000000 622.711683   0.000000  -0.000000
4       0.000000 -622.711683   0.000000   0.000000   0.000000
```

(continues on next page)

(continued from previous page)

```
LF-SOC-Y integrals (cm-1)
0          1          2          3          4
0       0.000000 -1078.568273  -0.000000  -0.000000  -0.000000
1     1078.568273   0.000000  -0.000000 -622.711683  -0.000000
2       0.000000   0.000000   0.000000  -0.000000 -622.711683
3       0.000000 622.711683   0.000000   0.000000  -0.000000
4       0.000000   0.000000 622.711683   0.000000   0.000000
LF-SOC-Z integrals (cm-1)
0          1          2          3          4
0       0.000000  -0.000000  -0.000000  -0.000000  -0.000000
1       0.000000   0.000000 -622.711683  -0.000000  -0.000000
2       0.000000 622.711683   0.000000  -0.000000  -0.000000
3       0.000000   0.000000   0.000000   0.000000 -1245.423365
4       0.000000   0.000000   0.000000 1245.423365   0.000000

RMS error of nonzero matrix elements =    0.0 cm**-1


-----SOC-CONSTANTS-----
---All Values in cm-1---
ZETA_D = 622.71
------------------------
```

Starting fron ORCA 5.0 it is also possible in addition to CASSCF and NEVPT2 to employ DCDCAS(2) and Hermitian QD-NEVPT2 Abinitio Hamiltonians in AILFT Example inputs are provided below for DCDCAS(2):

```
!def2-SVP def2-SVP/C

%casscf
 nel 8
 norb 5
 actorbs dorbs
 mult 3,1
 nroots 10,15
 dcdcas true
 corrorder 2
 rel
  dosoc true
 end
end

*xyz 2 3
Ni      0.0000000000      0.0000000000      0.0000000000
*
```

and Hermitian QD-NEVPT2:

```
!def2-SVP def2-SVP/C

%casscf
 nel 8
 norb 5
 actorbs dorbs
 mult 3,1
 nroots 10,15
 PTMethod  sc_nevpt2
 PTSettings
   QDType QD_VanVleck
 end
 rel
  dosoc true
```

(continues on next page)

(continued from previous page)

```
 end
end

*xyz 2 3
Ni      0.0000000000      0.0000000000      0.0000000000
*
```

Running the above inputs the respective DCDCAS(2) and Hermitian QD-NEVPT2 Hamiltonians will be processed:

```
Calculating ab initio Hamiltonian matrices                ...
------------------------------------------------------------
DCDCAS correction for this block/order is calculated
DCDCAS Hamiltonian of block = 0 and order =  0 is passed
DCDCAS Hamiltonian diagonalized
------------------------------------------------------------
------------------------------------------------------------
DCDCAS correction for this block/order is calculated
DCDCAS Hamiltonian of block = 0 and order =  1 is passed
DCDCAS Hamiltonian diagonalized
------------------------------------------------------------
------------------------------------------------------------
DCDCAS correction for this block/order is calculated
DCDCAS Hamiltonian of block = 0 and order =  2 is passed
DCDCAS Hamiltonian diagonalized
------------------------------------------------------------
------------------------------------------------------------
DCDCAS correction for this block/order is calculated
DCDCAS Hamiltonian of block = 1 and order =  0 is passed
DCDCAS Hamiltonian diagonalized
------------------------------------------------------------
------------------------------------------------------------
DCDCAS correction for this block/order is calculated
DCDCAS Hamiltonian of block = 1 and order =  1 is passed
DCDCAS Hamiltonian diagonalized
------------------------------------------------------------
------------------------------------------------------------
DCDCAS correction for this block/order is calculated
DCDCAS Hamiltonian of block = 1 and order =  2 is passed
DCDCAS Hamiltonian diagonalized
------------------------------------------------------------

...

Calculating ab initio Hamiltonian matrices                ...
------------------------------------------------------------
Hermitian QD-NEVPT2 correction for this block is calculated
Hermitian QD-NEVPT2 Hamiltonian of block = 0 is passed
Hermitian QD-NEVPT2 Hamiltonian diagonalized
------------------------------------------------------------
------------------------------------------------------------
Hermitian QD-NEVPT2 correction for this block is calculated
Hermitian QD-NEVPT2 Hamiltonian of block = 1 is passed
Hermitian QD-NEVPT2 Hamiltonian diagonalized
------------------------------------------------------------
```

It should be noted that NEVPT2 and Hermitian QD-NEVPT2 AILFT require a complete saturation of the excitation space. This implies that if less roots than the required are requested the AILFT analyis will be skipped in these cases. This is on the contrary not the case in CASSCF or DCDCAS(2) in which AILFT can operate under incomplete saturation of the excitation space.

```
Calculating ab initio Hamiltonian matrices                ...
WARNING: Number of NEVPT2 roots for block 0 (5) is not equal to the number of CASCI CSFs (10)!
Skipping AILFT analysis with NEVPT2 energies!

WARNING: Number of NEVPT2 roots for block 1 (2) is not equal to the number of CASCI CSFs (15)!
Skipping AILFT analysis with NEVPT2 energies!
```

In a similar fashion one can request a 2-shell AILFT calulation.

For this purpose the recomended steps are the following:

- In a first step the valence active space orbitals are optimized in the framework of SA-CASSCF calculation.e.g. the 3d MOs in a core 1s3d or 2p3d AILFT calculation, or the f MOs in an 4f5d AILFT calculation)
- In a second step the relevant core or virtual orbitals are rotated into the active space and the chosen CASCI/AILFT problem is solved by saturating the excitation space with all the involved excitations/multiplicity.
- In the most of the cases the excitation space of two multiplicities the High-Spin one and the subsequent Low-Spin one are enouph for a succesfull fitting of the parameters

It should be noted that 2-shell AILFT ivolves a 2-step fitting process following a bottom up shell angular momentum approach :

1. At first when possible an intra-shell fitting is performed
2. In following the respective effective Slater exponents are derived
3. In a last step an inter-shell fitting is performed and all the computed/fitted parameters are printed

This implies that:

- the flag of computing effective Slater exponents is always on by default in 2-shell AILFT
- the desired LFT problem is best requested by the LFTCase keywords (e.g. LFTCase 1s3d)

Let look at the case of 1s3d LFT problem of the Ni^{2+} d^8 ion. A relevant input is provided below:

```
!NoIter NEVPT2 def2-SVP def2-SVP/C

%method
 frozencore fc_none
end

%scf
 rotate
  {0,8,90}
 end
end

%casscf
 nel 10
 norb 6
 mult 3,1
 nroots 100,100
 LFTCase 1s3d
 rel
  dosoc true
 end
end


*xyz 2 3
Ni      0.0000000000      0.0000000000      0.0000000000
*
```

Like in 1-shell AILFT, 2-shells AILFT starts with a sanity check

```
---- THE CAS-SCF GRADIENT HAS CONVERGED ----
--- FINALIZING ORBITALS ---
---- DOING ONE FINAL ITERATION FOR PRINTING ----
--- sd-orbitals (depends on the molecular axis frame)
L-Center:  0 Ni [0.000, 0.000, 0.000]
L-Center:  0 Ni Active Orbital 0 is the s orbital ; l = 0 ; active shell = 0
L-Center:  0 Ni Active Orbital 1 is one d orbital ; l = 2 ; active shell = 7
L-Center:  0 Ni Active Orbital 2 is one d orbital ; l = 2 ; active shell = 7
L-Center:  0 Ni Active Orbital 3 is one d orbital ; l = 2 ; active shell = 7
L-Center:  0 Ni Active Orbital 4 is one d orbital ; l = 2 ; active shell = 7
L-Center:  0 Ni Active Orbital 5 is one d orbital ; l = 2 ; active shell = 7
--- The active space contains 1 s orbitals and 5 d orbitals : OK
Setting 8 active MO to AO s       (0)
Setting 9 active MO to AO dz2     (11)
Setting 10 active MO to AO dxz     (12)
Setting 11 active MO to AO dyz     (13)
Setting 12 active MO to AO dx2y2   (14)
Setting 13 active MO to AO dxy     (15)
--- Canonicalize Internal Space
--- Canonicalize External Space
```

In following the AI-LFT Hamiltonians are constructed and the LFT parameters are fitted at the CASSCF and at the NEVPT2 levels of theory

```
------------------------------
AILFT MATRIX ELEMENTS (CASSCF)
------------------------------

Ligand field one-electron matrix VLFT (a.u.) with V(0,0) fixed :
0          1          2          3          4          5
0     -334.652557  -0.000000   0.000000  -0.000000  -0.000000  -0.000000
1      -0.000000 -10.085777   0.000000   0.000000  -0.000000  -0.000000
2       0.000000   0.000000 -10.085777  -0.000000  -0.000000  -0.000000
3      -0.000000   0.000000  -0.000000 -10.085777   0.000000   0.000000
4      -0.000000  -0.000000  -0.000000   0.000000 -10.085777  -0.000000
5      -0.000000  -0.000000  -0.000000   0.000000  -0.000000 -10.085777

-------------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
-------------------------------------------------
F0ss            = 17.137989284 a.u. =   466.348 eV = 3761353.9 cm**-1
F0dd            =  0.809116820 a.u. =    22.017 eV =  177580.6 cm**-1
F2dd            =  0.427107567 a.u. =    11.622 eV =   93739.3 cm**-1
F4dd            =  0.278548413 a.u. =     7.580 eV =   61134.3 cm**-1
F0sd            =  1.285327742 a.u. =    34.976 eV =  282096.8 cm**-1
G2sd            =  0.001928559 a.u. =     0.052 eV =     423.3 cm**-1
R2sddd          =  0.003748289 a.u. =     0.102 eV =     822.7 cm**-1

---------------------------------------
The ligand field one electron eigenfunctions:
---------------------------------------
Orbital    Energy (eV) Energy(cm-1)         s          dz2        dxz        dyz     dx2-y2    ␣
↪   dxz
1          0.000        0.0          -1.000000  -0.000000   0.000000  -0.000000  -0.000000   0.
↪000000
2       8831.911  71234174.2           0.000000  -0.006282   0.006583  -0.271975   0.962261  ␣
↪0.000000
3       8831.911  71234174.2           0.000000   0.000000   0.000000   0.000000   0.000000  ␣
↪1.000000
4       8831.911  71234174.2           0.000000  -0.075508  -0.016185  -0.959327  -0.271528  ␣
↪0.000000
```

(continues on next page)

(continued from previous page)

```
5       8831.911  71234174.2          -0.000000   0.071391  -0.997365   0.008467   0.009682  
↪0.000000
6       8831.911  71234174.2          -0.000000   0.994566   0.070404  -0.075159  -0.015232  -
↪0.000000
Ligand field orbitals were stored ni.1s3d.casscf.lft.gbw

...

------------------------------
AILFT MATRIX ELEMENTS (NEVPT2)
------------------------------

Ligand field one-electron matrix VLFT (a.u.) with V(0,0) fixed :
0          1          2          3          4          5
0     -334.652557  -0.000000   0.000000   0.000000   0.000000   0.000000
1       -0.000000 -10.298799   0.000008   0.000006  -0.000002  -0.000002
2        0.000000   0.000008 -10.293700  -0.000042  -0.000008  -0.000006
3        0.000000   0.000006  -0.000042 -10.293839  -0.000001   0.000008
4        0.000000  -0.000002  -0.000008  -0.000001 -10.293927  -0.000033
5        0.000000  -0.000002  -0.000006   0.000008  -0.000033 -10.293511

-----------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
-----------------------------------------------
F0ss            = 16.677683435 a.u. =   453.823 eV = 3660328.4 cm**-1
F0dd            =  0.806859685 a.u. =    21.956 eV =  177085.2 cm**-1
F2dd            =  0.425916096 a.u. =    11.590 eV =   93477.8 cm**-1
F4dd            =  0.277771367 a.u. =     7.559 eV =   60963.8 cm**-1
F0sd            =  1.424742585 a.u. =    38.769 eV =  312694.9 cm**-1
G2sd            =  0.025339297 a.u. =     0.690 eV =    5561.3 cm**-1
R2sddd          =  0.003739506 a.u. =     0.102 eV =     820.7 cm**-1

---------------------------------------
The ligand field one electron eigenfunctions:
---------------------------------------
Orbital    Energy (eV) Energy(cm-1)         s          dz2        dxz        dyz      dx2-y2   
↪   dxz
1          0.000        0.0           1.000000   0.000000  -0.000000  -0.000000  -0.000000  -0.
↪000000
2       8826.114  71187421.2          -0.000000   0.999998  -0.001607  -0.001229   0.000405  
↪0.000329
3       8826.247  71188489.9          -0.000000   0.000330  -0.040203  -0.027545  -0.995773  -
↪0.077852
4       8826.249  71188507.4          -0.000000   0.001631   0.264542   0.963492  -0.035726  -
↪0.020544
5       8826.254  71188542.9           0.000000  -0.001223  -0.962932   0.264866   0.034492  -
↪0.037635
6       8826.258  71188582.5           0.000000  -0.000317  -0.034070   0.027728  -0.077265  
↪0.996042
Ligand field orbitals were stored in ni.1s3d.nevpt2.lft.gbw
```

As discussed above saturation of the excitation space is a requirement also in the case of 2-shell AILFT. It is usually enough to specify a large number of roots for two multiplicites (e.g. 100 singlets and triplets in the above example) The exact number of roots will be automatically detected.

```
Multiplicity                    ...    3
#(Configurations)               ...   15
#(CSFs)                         ...   15
#(Roots)                        ...  100
WARNING (ORCA_CASSCF): NRoots > NCSFs. Adjusting to maximum number of roots. Please check the
↪output carefully!
```

(continues on next page)

(continued from previous page)

```
#(Roots)                            ...   15
ROOT=0 WEIGHT=    0.033333
ROOT=1 WEIGHT=    0.033333
ROOT=2 WEIGHT=    0.033333
ROOT=3 WEIGHT=    0.033333
ROOT=4 WEIGHT=    0.033333
ROOT=5 WEIGHT=    0.033333
ROOT=6 WEIGHT=    0.033333
ROOT=7 WEIGHT=    0.033333
ROOT=8 WEIGHT=    0.033333
ROOT=9 WEIGHT=    0.033333
ROOT=10 WEIGHT=    0.033333
ROOT=11 WEIGHT=    0.033333
ROOT=12 WEIGHT=    0.033333
ROOT=13 WEIGHT=    0.033333
ROOT=14 WEIGHT=    0.033333

BLOCK  2 WEIGHT=   0.5000
Multiplicity                        ...    1
#(Configurations)                   ...   21
#(CSFs)                             ...   21
#(Roots)                            ...  100
WARNING (ORCA_CASSCF): NRoots > NCSFs. Adjusting to maximum number of roots. Please check the␣
↪output carefully!
#(Roots)                            ...   21
ROOT=0 WEIGHT=    0.023810
ROOT=1 WEIGHT=    0.023810
ROOT=2 WEIGHT=    0.023810
ROOT=3 WEIGHT=    0.023810
ROOT=4 WEIGHT=    0.023810
ROOT=5 WEIGHT=    0.023810
ROOT=6 WEIGHT=    0.023810
ROOT=7 WEIGHT=    0.023810
ROOT=8 WEIGHT=    0.023810
ROOT=9 WEIGHT=    0.023810
ROOT=10 WEIGHT=    0.023810
ROOT=11 WEIGHT=    0.023810
ROOT=12 WEIGHT=    0.023810
ROOT=13 WEIGHT=    0.023810
ROOT=14 WEIGHT=    0.023810
ROOT=15 WEIGHT=    0.023810
ROOT=16 WEIGHT=    0.023810
ROOT=17 WEIGHT=    0.023810
ROOT=18 WEIGHT=    0.023810
ROOT=19 WEIGHT=    0.023810
ROOT=20 WEIGHT=    0.023810
```

However very often the required number of states to be computed in the framework of NEVPT2 type of calculations are quite large. In these cases a Hamiltonian reduction process on the basis of the Restrictive Active Space (RAS) is required. In fact all LFT parameters can be determined by considering up to double excitations from the donor-shell.

Let us consider the 2p3d case of the Fe^{2+} d^6 ion. Saturation of the active space requires to consider 70 triplet and 378 singlet states. Restriction of the active space to only up to double excitations from the 2p-shell results in 65 quintet and 330 triplet states. The Hamiltonian reduction can be requested in the ailft block:

```
ailft
 AILFT_Dimension 2 # Up to double excitations from the donor shell
                 Other options:
                 1   Up to single excitations from the donor shell
                 0   No excitations from the donor shell
```

(continues on next page)

(continued from previous page)

```
end
```

Hence the relevant input can be now formulated as:

```
!NoIter MOREAD DKH2 DKH-def2-TZVP def2-TZVP/C NEVPT2

%moinp "fe.3d.gbw"

%pal
 nprocs 16
end

%method
 frozencore fc_none
end

%scf
 rotate
  {2,6,90}
  {3,7,90}
  {4,8,90}
 end
end

%casscf
 nel 12
 norb 8
 mult 5,3
 nroots 65,330
 LFTCase 2p3d
 ailft
  AILFT_Dimension 2
 end
 rel
  dosoc true
  GTensor false
  DoDTensor false
 end
end

*xyz 2 5
Fe      0.0000000000      0.0000000000      0.0000000000
*
```

Note that before running the above calculation:

- An initial SA-CASSCF calculation has been performed on the valence states of Fe in the 3d active space. These orbitals (fe.3d.gbw) are read in
- The computation of the g- and D- tensors is switched off. This is recomended if the magnetism analysis is not required
- As core spectroscopy is targeted the frozen core is switched off

At the NEVPT2 part the reduced AI and LFT Hamiltonians will be constructed

```
Calculating  ab initio Hamiltonian matrices              ...
------------------------------------------------------------
NRoots (NEVPT2) for this block = 65
NEVPT2 correction for this block is calculated
Full NEVPT2 Hamiltonian constructed
Full NEVPT2 Hamiltonian diagonalized
```

(continues on next page)

(continued from previous page)

```
------------------------------------------------------------
------------------------------------------------------------
NRoots (NEVPT2) for this block = 330
NEVPT2 correction for this block is calculated
Full NEVPT2 Hamiltonian constructed
Full NEVPT2 Hamiltonian diagonalized
------------------------------------------------------------
```

As a result the CASSCF and NEVPT2 LFT parameters will be determined in the requested reduced basis

```
-------------------------------
AILFT MATRIX ELEMENTS (CASSCF)
-------------------------------

------------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
------------------------------------------------
F0pp            =  3.889665545 a.u. =   105.843 eV =  853682.9 cm**-1
F2pp            =  1.832901835 a.u. =    49.876 eV =  402275.5 cm**-1
F0dd            =  0.767706319 a.u. =    20.890 eV =  168492.1 cm**-1
F2dd            =  0.405248254 a.u. =    11.027 eV =   88941.7 cm**-1
F4dd            =  0.264292339 a.u. =     7.192 eV =   58005.5 cm**-1
F0pd            =  1.174187892 a.u. =    31.951 eV =  257704.5 cm**-1
F2pd            =  0.220959621 a.u. =     6.013 eV =   48495.0 cm**-1
G1pd            =  0.157243007 a.u. =     4.279 eV =   34510.9 cm**-1
G3pd            =  0.089473762 a.u. =     2.435 eV =   19637.2 cm**-1

...

-------------------------------
AILFT MATRIX ELEMENTS (NEVPT2)
-------------------------------

------------------------------------------------
Slater-Condon Parameters (electronic repulsion) :
------------------------------------------------
F0pp            =  3.527444471 a.u. =    95.987 eV =  774184.6 cm**-1
F2pp            =  1.906123325 a.u. =    51.868 eV =  418345.7 cm**-1
F0dd            =  0.808248242 a.u. =    21.994 eV =  177390.0 cm**-1
F2dd            =  0.426649072 a.u. =    11.610 eV =   93638.6 cm**-1
F4dd            =  0.278249395 a.u. =     7.572 eV =   61068.7 cm**-1
F0pd            =  1.657417509 a.u. =    45.101 eV =  363761.1 cm**-1
F2pd            =  0.198646995 a.u. =     5.405 eV =   43598.0 cm**-1
G1pd            =  0.206111579 a.u. =     5.609 eV =   45236.3 cm**-1
G3pd            =  0.128174221 a.u. =     3.488 eV =   28131.0 cm**-1

...
```

In the above example inclusion of SOC will result in the computation of the effective SOC ζ constants of both p and d shells:

```
-----SOC-CONSTANTS-----
---All Values in cm-1---
ZETA_P = 65018.19
ZETA_D = 453.53
------------------------
```

One important feature of 1- and in particular of the 2-shell AILFT is that it is connected to the standalone orca_lft multiplet program. Hence every succesfull AILFT calculation will automatically construct relevant inputs for the orca_lft.

For example in the avove 2p3d case of the Fe^{2+} d^6 ion the following inputs will be constructed

```
fe.2p3d.casscf.lft.inp
fe.2p3d.nevpt2.lft.inp
```

with the NEVPT2 one looking like this:

```
%lft

#-----Parameters------
NEl= 12
Shell_PQN= 0, 2, 3, 0
Mult= 5, 3
NRoots= 65, 330
#--------------------


#---Slater-Condon Parameters---
#---All Values in eV---
PARAMETERS
F0pp =  95.9866
F2pp =  51.8683
F0dd =  21.9936
F2dd =  11.6097
F4dd =  7.5716
F0pd =  45.1006
F2pd =  5.4055
G1pd =  5.6086
G3pd =  3.4878
end
#--------------------


#---LFT-Matrix Elelemnts---
#---All Values in eV---
FUNCTIONS
0   0   "  0.0000"
1   0   " -0.0146"
1   1   "  0.0947"
2   0   "  0.0210"
2   1   "  0.0061"
2   2   "  0.1155"
3   0   "  0.0000"
3   1   " -0.0000"
3   2   " -0.0000"
3   3   "1086.2398"
4   0   "  0.0000"
4   1   " -0.0000"
4   2   " -0.0000"
4   3   "  0.0323"
4   4   "1086.2181"
5   0   " -0.0000"
5   1   "  0.0000"
5   2   "  0.0000"
5   3   " -0.0356"
5   4   " -0.0154"
5   5   "1086.1183"
6   0   "  0.0000"
6   1   " -0.0000"
6   2   " -0.0000"
6   3   " -0.0767"
6   4   " -0.0080"
6   5   "  0.0341"
6   6   "1086.1219"
```

(continues on next page)

(continued from previous page)

```
7   0   " -0.0000"
7   1   "  0.0000"
7   2   "  0.0000"
7   3   "  0.0050"
7   4   " -0.0023"
7   5   "  0.0026"
7   6   " -0.0038"
7   7   "1086.0688"
end
#--------------------


#---SOC-CONSTANTS---
#---All Values in eV---
PARAMETERS
ZETA_P = 8.06
ZETA_D = 0.06
end
#--------------------


#---SPECTRA/PROPERTIES---
DoABS true
#------------------------
end


*xyz Charge Multiplicity
Atom 0.00 0.00 0.00
*
```

Further details regardin orca_lft can be found in the orca_lft section (*orca_lft*) and the orca_lft tutorial.

7.15.5 Core excited states with (C/R)ASCI/NEVPT2

Starting from ORCA 4.1, a CASCI/NEVPT2 protocol can be used to compute core excited spectra, namely X-ray absorption (XAS) and resonant inelastic scattering (RIXS) spectra. RASCI calculations can also be easily specified.

The XAS/RIXS spectra calculations requires two steps:

- In a first step one needs to optimize the valence active space orbitals in the framework of SA-CASSCF calculations, e.g. including valence excited states in the range between 6 to 15 eV.
- In a second step the relevant core orbitals are rotated into the active space and the (C/R)ASCI/NEVPT2 problem is solved by saturating the excitation space with singly core-excited electronic configurations using the previously optimized sets of orbitals

Further information can be found in reference[163]

A relevant input for Fe L-edge XAS calculation of a Fe(III) complex like $Fe(acac)_3$ is given below for CASCI/NEVPT2:

```
%moinp "Fe_acac3_casscf.gbw"

%scf
rotate
{4,89,90,0,0}
{3,88,90,0,0}
{2,87,90,0,0}
end
```

(continues on next page)

(continued from previous page)

```
end

%rel
picturechange true
FiniteNuc true
end

%method FrozenCore FC_NONE
end

# CASSCF/NEVPT2 on the valence and L-edge excited states
%casscf
nel 11
norb 8
mult 6,4
nroots 16,173
maxiter 1
# account for spin-orbit coupling
rel
  DoSOC true
end
# adding the dynamical correlation with NEVPT2
PTMethod SC\_NEVPT2
end

* xyz 0 6
...
*
```

For RASCI/NEVPT2 calculations the valence d AS is set to RAS2. The RAS3 space is usually set empty. The RAS1 space contains the previously rotated core orbitals. To generate a single core hole, the number of maximum holes in the RAS1 space must be set to 1. Accordingly, the maximum number of particles in the RAS3 space must be 0. The RASCI input should thus look the following

```
%casscf
 nel 11
 norb 8
 ...
 refs
  ras(11:3 1/5/0 0) # (Nel: NRAS1 MaxHoles / NRAS2 / NRAS3 MaxParticles)
 end
 ...
end
```

As it is explicitly described in the respective ROCIS section RIXS spectra can be requested by the following keywords:

```
RIXS    true   # Request RIXS calculation (NoSOC)
RIXSSOC true   # Request RIXS calculation (with SOC)
Elastic true   # Request RIXS calculation (Elastic)
```

Please consult section *Resonant Inelastic Scattering Spectroscopy* for processing and analyzing the generated spectra.

7.15.6 CASCI-XES

Starting from ORCA 5.0 likewise to RASCI-XES (see section *X-ray Spectroscopy*) orca features a CASCI-XES protocol.

Likewise to the RASCI-XES the CASCI-XES calculations requires two steps:

- In a first step one needs to optimize the valence active space orbitals in the framework of SA-CASSCF calculations, e.g. including valence excited states in the range between 6 to 15 eV.
- In a second step the relevant core orbitals e.g metal 1s and 3p are rotated into the active space and the CASCI problem is solved for the ionized system by saturating the excitation space with singly core-excited electronic configurations using the previously optimized sets of orbitals. In CASCI-XES this can be acheived by defining reference configurations. or via the RAS functionality.

The XESSOC calculation is called by speciyfing the following keywords in the rel block:

```
rel
XESSOC true
XASMOs Number of the rotated 1s MO
end
```

Following a SA-CASSCF calculation:

```
! def2-SVP def2-SVP/C ZORA CPCM PAL8
! NormalPrint
! NoLoewdin NoMulliken

%casscf
nel 5
norb 5
nroots 1, 24
mult 6, 4
#rel
#dosoc true
#end
end


* xyz -3 6
Fe      0       0       0
Cl      2.40    0       0
Cl      -2.40   0       0
Cl      0       2.40    0
Cl      0       -2.40   0
Cl      0       0       2.40
Cl      0       0       -2.40
*
```

A relevant input for Fe XES calculation of a Fe(III) complex like $FeCl_6$ is given below:

```
! def2-SVP def2-SVP/C ZORA CPCM PAL8
! NormalPrint
! MOREAD
! NoLoewdin NoMulliken


%moinp "FeCl6_casscf.gbw"

%scf
#Rotate the 1s and 3p orbitals below the SOMOs by using the rotate option
rotate {0,59,90} {36,60,90} {37,61,90} {38,62,90} end
end
```

(continues on next page)

(continued from previous page)

```
%casscf
nel 12
norb 9
nroots 1000, 1000
mult 7, 5
maxiter 1
refs
  1 2 2 2 0 0 1 2 2
  1 2 2 2 0 0 2 1 2
  1 2 2 2 0 0 2 2 1
  1 2 2 2 0 1 0 2 2
  1 2 2 2 0 1 1 1 2
  1 2 2 2 0 1 1 2 1
  1 2 2 2 0 1 2 0 2
...
  2 2 2 2 2 0 0 2 0
  2 2 2 2 2 0 1 0 1
  2 2 2 2 2 0 1 1 0
  2 2 2 2 2 0 2 0 0
  2 2 2 2 2 1 0 0 1
  2 2 2 2 2 1 0 1 0
  2 2 2 2 2 1 1 0 0
  2 2 2 2 2 2 0 0 0
end
DoDipoleVelocity true
DoHigherMoments true
rel
dosoc true
XESSOC true
XASMOs 59
DoDTensor false
DoVelocity true
end
end


%method
SpecialGridAtoms 26 #Increase the radial integration accuracy on Fe
SpecialGridIntAcc 7 #Requested radial integration accuracy values
end

* xyz -2 5
Fe      0       0       0
Cl      2.40    0       0
Cl      -2.40   0       0
Cl      0       2.40    0
Cl      0       -2.40   0
Cl      0       0       2.40
Cl      0       0       -2.40
*
```

In ORCA 6 all these inputs between like MRCI, CASSCF and LFT have been unified so like in *X-ray Spectroscopy* one can also specify the respective RAS excitation space as following:

```
%casscf
nel 12
norb 9
nroots 1000, 1000
mult 7, 5
maxiter 1
```

(continues on next page)

(continued from previous page)

```
refs ras(12:4 1/5/ 0 0) end
DoDipoleVelocity true
DoHigherMoments true
rel
dosoc true
XESSOC true
XASMOs 59
DoDTensor false
DoVelocity true
end
end
```

In the above inputs one notes that the exact knowledge of the states to saturate the excitations space is not required. One only need to specify a large number (e.g. 1000) and the program will automatically detect the required 19 septet and 270 quintet states:

```
BLOCK  1 WEIGHT=   0.5000
Multiplicity                        ...    7
#(Configurations)                   ...    4
#(CSFs)                             ...    4
#(Roots)                            ... 1000
WARNING (ORCA_CASSCF): NRoots > NCSFs. Adjusting to maximum number of roots. Please check the␣
↪output carefully!
#(Roots)                            ...    4

...

BLOCK  2 WEIGHT=   0.5000
Multiplicity                        ...    5
#(Configurations)                   ...   89
#(CSFs)                             ...  105
#(Roots)                            ... 1000
WARNING (ORCA_CASSCF): NRoots > NCSFs. Adjusting to maximum number of roots. Please check the␣
↪output carefully!
#(Roots)                            ...  105
```

By now running the above input for the 4 septet and the 81 quintet states the following output is generated

```
------------------------------

Printing the XES spectrum ...

------------------------------


-----------------------------------------------------------------------------------
SPIN-ORBIT X-RAY EMISSION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-----------------------------------------------------------------------------------
Transition          Energy           INT             TX        TY        TZ
1  421 ->    5      7231.106     0.000000000000     0.00000   0.00000   0.00000
2  422 ->    5      7231.106     0.000000000000     0.00000   0.00000   0.00000
3  423 ->    5      7231.106     0.000000000000     0.00000   0.00000   0.00000
4  424 ->    5      7231.106     0.000000000000     0.00000   0.00000   0.00000
5  425 ->    5      7231.106     0.000000000000     0.00000   0.00000   0.00000
6  426 ->    5      7231.106     0.000000000000     0.00000   0.00000   0.00000
7  427 ->    5      7231.106     0.000000000000     0.00000   0.00000   0.00000

...

2641  421 ->   25      7179.464     0.003027813485     0.00003   0.00001   0.00415
2642  422 ->   25      7179.464     0.003027769731     0.00114   0.00399   0.00001
2643  423 ->   25      7179.464     0.003027770766     0.00399   0.00114   0.00003
```

(continues on next page)

(continued from previous page)

```
2644  424 ->   25      7179.464     0.000000000035     0.00000   0.00000   0.00000
2645  425 ->   25      7179.464     0.000000000000     0.00000   0.00000   0.00000
2646  426 ->   25      7179.464     0.000000000063     0.00000   0.00000   0.00000
2647  427 ->   25      7179.464     0.000000000001     0.00000   0.00000   0.00000
2648  428 ->   25      7179.520     0.000000000000     0.00000   0.00000   0.00000
2649  429 ->   25      7179.520     0.000000000000     0.00000   0.00000   0.00000
2650  430 ->   25      7179.520     0.000000000000     0.00000   0.00000   0.00000
2651  431 ->   25      7179.520     0.000000000000     0.00000   0.00000   0.00000
2652  432 ->   25      7179.520     0.000000000000     0.00000   0.00000   0.00000
2653  433 ->   25      7181.156     0.000000887886     0.00000   0.00000   0.00007
2654  434 ->   25      7181.156     0.000000887873     0.00001   0.00007   0.00000
2655  435 ->   25      7181.156     0.000000887873     0.00007   0.00001   0.00000
...

54898  538 ->  420      7164.167     0.000077338518     0.00000   0.00000   0.00066
54899  539 ->  420      7164.167     0.000077338538     0.00047   0.00047   0.00000
54900  540 ->  420      7164.167     0.000077338525     0.00047   0.00047   0.00000
54901  541 ->  420      7164.186     0.000000000000     0.00000   0.00000   0.00000
54902  542 ->  420      7164.186     0.000000000000     0.00000   0.00000   0.00000
54903  543 ->  420      7164.186     0.000000000000     0.00000   0.00000   0.00000
54904  544 ->  420      7164.186     0.000000000000     0.00000   0.00000   0.00000
54905  545 ->  420      7164.186     0.000000000000     0.00000   0.00000   0.00000
54906  546 ->  420      7164.214     0.000000000000     0.00000   0.00000   0.00000
54907  547 ->  420      7164.214     0.000000000000     0.00000   0.00000   0.00000
54908  548 ->  420      7164.214     0.000000000000     0.00000   0.00000   0.00000
54909  549 ->  420      7164.214     0.000000000000     0.00000   0.00000   0.00000
54910  550 ->  420      7164.214     0.000050793492     0.00000   0.00000   0.00054
54911  551 ->  420      7164.214     0.000050793481     0.00020   0.00050   0.00000
54912  552 ->  420      7164.214     0.000050793480     0.00050   0.00020   0.00000

All Done
-----------------------------------------------------------------------------
```

Finally by processing the .out file with orca_mapspc:

```
orca_mapspc fecl6_xes.out XESSOC -x07140 -x17190 -w4.0 -eV -n10000
```

and by plotting the resulted XES spectrum one will get the respective RASCI-XES spectrum presented in Fig. 7.42

Since Orca 6.0 the computed transition moments in the presence of SOC can be taken beyond the dipole approximation by using the OPS tool, check section (*One Photon Spectroscopy*) for details.

All the possible approximations can be requested with the following commands:

```
#Non-Relativistic/Relativistic treatment

%casscf
 DoDipoleLength        true
 DoDipoleVelocity      true
 DoHigherMoments       true
 DecomposeFoscLength   true
 DecomposeFoscVelocity true
 DoFullSemiclassical   true
end
```

This wil generate a list of tables which for the case of XESSOC will look like the following:

```
-----------------------------------------------------------------------------------------
SOC COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE X-RAY EMISSION SPECTRUM
-----------------------------------------------------------------------------------------
                        INT (fosc)
-------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
Transition    Energy    D2     m2     Q2     D2+m2+Q2     D2/TOT    m2/TOT    Q2/TOT
              (eV)           (*1e6) (*1e6)
-----------------------------------------------------------------------------------------
...

-----------------------------------------------------------------------------------------
SOC COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE X-RAY EMISSION SPECTRUM
                                    (Origin Adjusted)
-----------------------------------------------------------------------------------------
                        INT (fosc)
-----------------------------------------------------------
Transition    Energy    D2     m2      Q2     D2+m2+Q2     D2/TOT    m2/TOT     Q2/TOT
              (eV)    (*1e6) (*1e6)
-----------------------------------------------------------------------------------------

...

-----------------------------------------------------------------------------------------
SOC COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE  X-RAY EMISSION SPECTRUM
                                  (Origin Independent, Velocity)
-----------------------------------------------------------------------------------------
                        INT (fosc)
-----------------------------------------------------------
Transition   Energy     D2     m2      Q2     D2+m2+Q2+DM+DO  D2/TOT     m2/TOT     Q2/TOT
              (eV)           (*1e6)  (*1e6)
-----------------------------------------------------------------------------------------

...

-----------------------------------------------------------------------------------------
SOC COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE  X-RAY EMISSION SPECTRUM
                                  (Exact Formulation, Velocity)
-----------------------------------------------------------------------------------------
↪-
                        INT (fosc)
-------------------------------------------------------------------------
Transition  Energy     D2     m2      Q2     Exact Osc. Str. D2/TOT      m2/TOT      Q2/TOT
             (eV)           (*1e6)  (*1e6)
-----------------------------------------------------------------------------------------
↪-
```

Orca_mapspc can process all these files. The list of the relevant keywords is:

```
ABSQ     #This will process non relativistic Origin Adjusted Absortion like spectra
ABSOI    #This will process the non relativistic Exact, Velocity Absortion like spectra
SOCABSQ  #This will process the SOC corrected Origin Adjusted Absortion like spectra
SOCABSOI #This will process the SOC corrected Exact Velocity Absortion like spectra
XESSOCQ  #This will process the SOC corrected Origin Adjusted X-ray Emission spectra
XESSOCOI #This will process the SOC corrected Exact Velocity X-ray Emission spectra
```

A more complete list can be found in (*orca_mapspc*)

7.16 CASSCF Linear Response

Similar to the SCF linear response (see *CP-SCF Options*), second-derivative properties can be calculated at the CASSCF level by solving the coupled perturbed (CP-)CASSCF equations for the linear response of the wavefunction parameters to a perturbation. These linear response equations are expressed as

$$\frac{\partial^2 E}{\partial \boldsymbol{\lambda}^2}\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{R}} = -\frac{\partial^2 E}{\partial \boldsymbol{\lambda}\partial \mathbf{R}}$$

where $\boldsymbol{\lambda}$ are the CASSCF wavefunction parameters and $\mathbf{R}$ are the perturbations for which the response equations are being solved [378, 860]. The property gradient on the right-hand side (RHS) is a known perturbation-dependent quantity, but the left-hand side (LHS) depends on the solution of the response of the wavefunction parameters to $\mathbf{R}$. Therefore, the response to the perturbation must be solved for iteratively, which is done using the trial vectors $\mathbf{X}$. This leads to the LHS being computed as a sigma-vector, which is reassembled at every iteration. The linear response equation for perturbation R_i can then be written as

$$\boldsymbol{\sigma} = \mathbf{H}\mathbf{X}^{(R_i)} = \mathbf{G}^{(R_i)}$$

where $\mathbf{H}$ is the perturbation-independent Hessian matrix and $\mathbf{G}$ is the perturbation-specific RHS matrix. Once the linear equation converges, the solution vector of the response parameters can then be contracted with the electron- and spin-density matrices to yield the AO response density matrices, $\left[\mathbf{P}^{\alpha\pm\beta}\right]^{(\mathbf{R})}$. The response densities can then be contracted with appropriate AO 1-electron property integrals to compute the second-order contributions to second-derivative properties.

In ORCA these equations are solved by the `orca_casscfresp` program and the underlying solver is `BHP22`, a Davidson-type linear equation solver. The RHS is built from the property integrals calculated in the `orca_propint` program. After the solution converges and the response densities are made and stored, the `orca_prop` program is called, wherein the appropriate densities and response densities are contracted with the necessary property integrals (also from `orca_propint`). This use of densities keeps the response property calculations in `orca_prop` generally applicable to all methods as the densities house the method-specific information.

7.16.1 Input Block

The input block `%casresp` is available for requesting the following options (given below with their default values):

```
%casresp
  TolR           1.e-5  # Convergence tolerance for the residual norm
                        # Note: TolR not affected by global keywords (TightSCF, etc.)
  MaxIter        64     # Maximum number of iterations
  MaxRed         24     # Maximum size of the reduced space per RHS

  PrintRHSVec    false  # Print significant contributions to the RHS vector
  PrintRspVec    false  # Print significant contributions to the response vector
  PrintWF        CFG    # print CI part in (CFG (default), CSF, or DET basis)
  TolPrintVec    1.e-2  # Threshold to print a contribution to the RHS/response vector
                        # Minimum value of the SQUARE of a vector element to print

  PreCondType    diag   # (default) Use diagonal preconditioner (requires auxiliary basis!)
                 none   # Do not use a preconditioner
  PreCondMaxRed  400    # Hessian precondition subset size
end
```

Note that the solver-related options in the `%elprop` and `%eprnmr` blocks do not affect the solution of the CP-CASSCF equations.

7.16.2 Technical Notes

It should be noted that CASSCF linear response uses the optimized CASSCF wavefunction as a starting point. Thus, a `%casscf` block with the appropriate inputs (see *Complete Active Space Self-Consistent Field Method*) must be provided in the input. State-specific (SS-)CASSCF response is run on SS (`NRoots 1`) CASSCF wavefunctions. If the CASSCF wavefunction is state-averaged (SA), the response is run over the averaged manifold of states and not on a specific root. In this case, one should analyze the output carefully. It is wise to only average states which are (nearly-)degenerate.

By default, a CASSCF calculation will be run from scratch before running the CASSCF Response. Alternatively, orbitals from a previously converged CASSCF calculation may be used with `!MOREAD NoIter` and supplying the appropriate `.gbw` file via `%moinp`. In this case, be sure that the input in the `%casscf` block matches that from the same input block of the previously converged calculation and be sure to check the orbitals well!

The appropriate property flags in the `%elprop` and `%eprnmr` blocks must be set to calculate the properties that are wanted (see *Electric Properties*, *EPR and NMR properties*, and *Calculation of Properties*). While all first-derivative properties work with CASSCF, not all second-derivative properties are currently available with CASSCF. The static properties currently available (sorted by perturbation taken for the response) include:

- Electric ($\mathbf{E}$) Field
  - Dipole/dipole Polarizability
  - Dipole/quadrupole Polarizability
- Quadrupole ($\mathbf{Q}$) Field
  - Quadrupole/Quadrupole Polarizability
- Magnetic ($\mathbf{B}$) Field (**without** GIAOs)
  - EPR g-Tensor
  - NMR Chemical Shieldings
- Velocity ($\mathbf{v}$)
  - Velocity Polarizability

The magnetic properties are currently implemented without gauge-including atomic orbitals (GIAOs). Thus, an appropriate gauge-origin must be provided in the `%eprnmr` block! For the EPR g-tensor, using a large basis set with a chemically relevant gauge origin is recommended. However, one should in general be wary of NMR chemical shieldings without GIAOs. GIAOs for CASSCF linear response are coming soon to ORCA!

7.16.3 Notes on Printing

The information on the types of property integrals computed can be found in the `ORCA PROPERTY INTEGRAL CALCULATIONS` section of the output file. The `orca_casscfresp` output begins with the header `ORCA CASSCF RESPONSE CALCULATION`. After information about the types and number of perturbations, the calculation splits into the major types of perturbations: real and imaginary. All response equations of the same type can be solved simultaneously.

Each of these types has its own section in the output file. These sections begin with information on the orbital ranges and the CI space, then go into the iterative solution of the equations. The printout here gives an overview of the iteration number, the residual norm of each response equation, and if that response equation has met the convergence criteria. The following is an example of this output (with iterations 3–7 removed for simplicity):

```
-------------------------------------------------------------------------------------------

           LINSOL Davidson-type linear equation solver

  Iter.               ||Error||_2   Conv.            (TolR = 1.000e-08)
-------------------------------------------------------------------------------------------
   0  (rhs    0)       1.645417e-01  No
```

(continues on next page)

(continued from previous page)

```
       (rhs    1)       1.826041e-01  No
       (rhs    2)       2.047543e-01  No
    1  (rhs    0)       7.069575e-02  No
       (rhs    1)       3.363295e-02  No
       (rhs    2)       7.229274e-02  No
    2  (rhs    0)       9.183136e-03  No
       (rhs    1)       3.922435e-03  No
       (rhs    2)       7.152764e-03  No
(...)
    8  (rhs    0)       4.070240e-09 Yes
       (rhs    1)       1.186846e-09 Yes
       (rhs    2)       4.407666e-09 Yes
```

Before going on to the necessary property modules, the significant contributions to the RHS and/or response vectors will be outputted if they were requested (via `PrintRHSVec` and `PrintRspVec`). Here, outputs such as the following can be seen.

```
  --------------------------------
   CASSCF Response Vector Analysis
  --------------------------------
    square of coefficients of particle-hole and CAS-CI excitations are printed if larger than␣
↪1.0e-02


IPERT: 0

   (-) I   9 (   9) -> V   0 (  17) :      0.07167 (c= -0.26772137)
(...)
   (-) A   1 (  14) -> V  13 (  30) :      0.07323 (c=  0.27061481)
(...)
   (-) I   4 (   4) -> A   0 (  13) :      0.02915 (c=  0.17074032)
(...)
   (-) CI                                  1.65592 [     7]: 1201
(...)
```

The majority of the output has been removed for the sake of simplicity. Here, the significant contributions to the response vector are printed if the square of the vector element is larger than `1.0e-02` (the default value for `TolPrintVec`). Each line begins with either a `(+)` or `(-)`, which denote the Hermiticity of the excitation operator. With static CASSCF linear response, `(+)` always denotes an imaginary perturbation and `(-)` a real perturbation. For the first vector (i.e. `IPERT: 0`), four contributions have been listed here.

The first three are from particle-hole excitations going from the left size of the arrow to the right. On each side of the arrow has three things: a letter designating if it is an inactive (`I`), active (`A`), or virtual (`V`) orbital; a number designating the index within that orbital subblock, and a number in parentheses designating the overall orbital number. Further to the right, the coefficient (i.e. the vector element) is given in parentheses and the square of that value is given to the left of it. This is how the relative contributions of each excitation can be analyzed.

The final contribution shown is from a CI excitation. Its weight is listed, followed by the index in brackets and the corresponding configuration of the active electrons among the active orbitals.

7.16.4 Troubleshooting

If the %casresp block is specified and the calculation does not run through the CASSCF Response section, then the properties requested are not second-derivatives and therefore do not require linear response equations to be solved. If, however, the job aborts, watch for these possibilities:

- At least one of the second-derivative properties requested is not available at the CASSCF level (see list above of those currently implemented)
- MaxIter may need to be increased or TolR decreased
- A magnetic property is requested without setting the gauge-origin in the %eprnmr block
- An appropriate auxiliary basis (or the !AutoAux keyword) may be required, especially for:
  - The diagonal preconditioner (which is on by default)
  - TrafoStep RI in the %casscf block
  - RIJCOSX, RIJONX
- RIJK was specified (this is **NOT** supported)

7.16.5 Minimal Input File

The following is an overview of the blocks required for a property calculation with

```
! def2-TZVP AutoAux    # (or other appropriate basis & auxiliary)

%casscf
  ...
end

%casresp     # only required to change solver, printing, or preconditioner options
  ...
end

%elprop      # only required if electric property requested
  ...
end

%eprnmr      # only required if magnetic property requested
  ...
  ORI ...
end

* xyzfile ...
```

7.17 Interface to SINGLE_ANISO module

7.17.1 General description

The SINGLE_ANISO program allows the non-perturbative calculation of effective spin (pseudospin) Hamiltonians and static magnetic properties of mononuclear complexes and fragments on the basis of an *ab initio*, including the spin-orbit interaction. As a starting point it uses the results of a CASSCF/NEVPT2/SOC calculation for the ground and several excited spin-orbit multiplets.

The following quantities can be computed:

- Parameters of pseudospin magnetic Hamiltonians (the methodology is described in [168]) :

1. First order (linear after pseudospin) Zeeman splitting tensor (g tensor), including the determination of the sign of the product $g_X \cdot g_Y \cdot g_Z$.
2. Second order (bilinear after pseudospin) zero-field splitting tensor (D tensor).
3. Higher order zero-field splitting tensors ($D^2, D^4, D^6, ..., etc.$)
4. Higher order Zeeman splitting tensors ($G^1, G^3, G^5, ..., etc.$)

- Crystal-Field parameters for the ground atomic $\tilde{J}$ multiplet for lanthanides. [414, 858]
- Crystal-Field parameters for the ground atomic $\tilde{L}$ term for lanthanides and transition metals.
- Static magnetic properties [856, 857]:
  1. Van Vleck susceptibility tensor $\chi_{\alpha\beta}(T)$.
  2. Powder magnetic susceptibility function $\chi(T)$.
  3. Magnetisation vector $\vec{M}(\vec{H})$ for specified directions of the applied magnetic field $\vec{H}$.
  4. Powder magnetisation $\overline{M_{mol}}(H,T)$.
  5. Magnetisation torque function $\vec{\tau}_{mol}(H,T)$.

The magnetic Hamiltonians are defined for a desired group of N low-lying electronic states obtained in `CASSCF/SOC` calculation to which a *pseudospin* $\tilde{S}$ is subscribed according to the relation $N = 2\tilde{S} + 1$. The *pseudospin* $\tilde{S}$ reduces to a true spin S in the absence of spin-orbit coupling. For instance, the two wave functions of a Kramers doublet correspond to the pseudospin $\tilde{S} = 1/2$. The implementation is done for any dimension of the pseudospin $\tilde{S}$, and controlled by the keyword `MLTP`.

The calculation of magnetic properties takes into account the contribution of excited states (the ligand-field and charge transfer states of the complex or mononuclear fragment which were included in the `CASSCF/CASPT2` calculation) via their thermal population and Zeeman admixture. The effect of intermolecular exchange interaction between magnetic molecules on the resulting magnetic properties in a crystal is described by a phenomenological parameter zJ specified by the user.

7.17.2 Running `SINGLE_ANISO` calculations

The `SINGLE_ANISO` is, in principle, a stand-alone utility (`otool_single_aniso`) that can be called directly from the shell with its own input file, provided that the *ab initio* datafile is available:

```
bash:$
bash:$  $ORCA/x86_64/otool_single_aniso < single_aniso.input > single_aniso.output
bash:$
```

However, this usage may not be so convenient, as the file `single_aniso.input` must include the true name of the datafile. For the user's convenience, a deeper integration between `SINGLE_ANISO` and CASSCF program in ORCA was implemented, as described below.

As a prerequisite for using the `SINGLE_ANISO` module to calculate the magnetic properties of the investigated compound, **spin-orbit coupling and other relativistic effects are already taken fully into account at the stage of quantum chemistry calculation of the investigated compound**. The necessary information of the *ab initio* calculation is provided in a form of a "datafile": energy spectra, angular momentum integrals, etc. The interface with ORCA generates the required `datafile` automatically. The following naming conventions were adopted for the `datafile` in function of the employed computational method:

- `CASSCF+SOC+SINGLE_ANISO` => `"$orca_input_name.CASSCF.anisofile"`
- `CASSCF+QD-NEVPT2+SOC+SINGLE_ANISO` => `"$orca_input_name.NEVPT2.anisofile"`

Note that if the `CASSCF+QD-NEVPT2+SOC+SINGLE_ANISO` calculation is requested, then the `SINGLE_ANISO` will be executed twice, and the above two datafiles will be generated. The interface will generate the `SINGLE_ANISO` input file with the keywords information provided in the `CASSCF/aniso` subblock. These filename of the `datafile`

is included automatically in the input file for the SINGLE_ANISO utility (keyword DATA), also generated automatically by the interface. The naming convention for the generated input files for the SINGLE_ANISO utility is "$orca_input_name.anisofile".

All keywords of the SINGLE_ANISO program are possible to be specified within the CASSCF/ANISO subblock. They are referenced in Section *Reference list of CASSCF/ANISO keywords*. Optionally, a working SINGLE_ANISO input file can be passed directly to the CASSCF module setting the filename with keyword InputNameOnDisk in the ANISO subblock.

An example of the full ORCA input for performing magnetic properties calculations within the CASSCF/SOC/SINGLE_ANISO methodology for a hypothetical Co(II) compound is provided below:

```
! 6-31G TightSCF   # basis set and other global ORCA settings

%maxcore 2000

%casscf nel 7
  norb 5           # 7 electrons in 5 orbitals (3d shell)
  mult 4, 2
  nroots 10, 40     # 10 quartet and 40 doublet states

  rel
  dosoc true       # include spin-orbit coupling
  end

  ANISO
    doaniso true     # generate datafile/input and call
                     # the SINGLE_ANISO module
    MLTP 2,2,2,2     # group of spin-orbit states for which the pseudospin
                     # Hamiltonian is computed: 4 low-lying spin-orbit doublet states.

    TINT 0, 300, 301 # 301 steps in the temperature interval [0-300]
                     # for magnetic susceptibility (in Kelvin)
    HINT 0, 7.0, 71  # 71 steps in the field interval [0-7]
                     # for molar magnetisation (in Tesla)
    TMAG 1.0,1.2,1.8,2.5,3.6  # temperature points for which molar magnetisation
                                            # is computed
     CRYS_element "Co"
     CRYS_charge 2
     PLOT true         # requires the ANISO to produce gnuplot scripts,
                              # datafiles and plots of various quantities

    # Alternative to the snippet above. Provide separate input file:
    # InputFile "$orca_input_name.anisoinput"
  end
end

*xyz 0 4   # charge is 0 for this neutral compound
Co   -2.80118000    9.91634000   19.40386000
O    -3.59660000   12.00284000   20.51731000
O    -5.12835000   10.85934000   19.53431000
O    -5.70975000   12.39302000   20.99406000
O    -1.30202341   11.67611386   19.17300658
O    -3.84191000    9.45315000   21.48634000
O    -1.27500262    8.12582233   19.18634310
O    -3.94611990    9.65426823   17.48476360
N    -4.85020000   11.78071000   20.36823000
H    -1.23636310   12.09677337   18.41017549
H    -1.07910455    7.59540828   19.85227241
H    -3.30514987    9.28034259   22.26327382
H    -4.79957696    9.43862752   21.55163236
H    -4.64801074    9.00163025   17.42987361
H    -3.73273676   10.19508893   16.72083912
```

(continues on next page)

(continued from previous page)

```
H    -0.75470916   11.94100908   19.91589125
*
```

The input above utilises the following keywords: `MLTP` keyword requires the computation of the g tensor for 4 groups of spin-orbit states, the dimensionality of each group being 2 (Kramers or Ising doublets). `TINT` requires computation of the magnetic susceptibility in the temperature interval 0 K - 300 K distributed equally in 300 temperature intervals. `TMAG` requires computation of powder molar magnetisation at 6 temperature points, in Kelvin (K): 1.0 K, 1.2 K, 1.8 K, 2.5 K, 2.9 K and 3.6 K. `HINT` defines the range for the magnetic field strength, in Tesla. `PLOT` keyword invokes the plotting function of the module. `CRYS_element + CRYS_charge` request for the computation of the crystal field parameters for the ground term of the Co^{2+} ion. For more information about the keywords in `SINGLE_ANISO` module, you can refer to section *Reference list of CASSCF/ANISO keywords*.

Please always check the obtained orbitals after `CASSCF` calculation. In this particular case, the active orbitals (45-49) are localised on the Co site and display dominant $3d$ character.

```
                45        46        47        48        49
           -0.37264  -0.36672  -0.36520  -0.36153  -0.35018
            1.40513   1.40270   1.39998   1.39823   1.39395
            --------  --------  --------  --------  --------
0 Co s        0.0       0.0       0.0       0.0       0.2
0 Co pz       0.1       0.1       0.1       0.0       0.0
0 Co px       0.0       0.2       0.0       0.0       0.1
0 Co py       0.1       0.1       0.0       0.0       0.1
0 Co dz2     23.6      39.5      10.0       2.0      23.5
0 Co dxz     12.8      30.8      48.6       4.4       2.2
0 Co dyz     46.5      10.4      19.3      20.8       2.4
0 Co dx2y2    6.9       0.3      21.2      69.2       1.6
0 Co dxy      9.5      17.6       0.2       2.8      68.0
```

We see that in the above output section, the five active orbitals have dominant contribution from the Co-$3d$ basis functions. This is OK and is expected for common transition metal compounds. For lanthanide compounds, the seven active orbitals should have dominant contribution from the $4f$ shell. Larger active spaces must be carefully inspected and analysed. We refer here to the respective section of this manual describing the CASSCF method and how to achieve convergence *The Complete Active Space Self-Consistent Field (CASSCF) Module*.

The results calculated by using `SINGLE_ANISO` module are placed after the `SOC` section in ORCA output. Here is the explanation for these results.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
CALCULATION OF PSEUDOSPIN HAMILTONIAN TENSORS FOR THE MULTIPLET 1 ( effective S =  1/2)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
The pseudospin is defined in the basis of the following spin-orbit states:
spin-orbit state 1. energy(1) =       0.000 cm-1.
spin-orbit state 2. energy(2) =       0.000 cm-1.

g TENSOR:
|-------------------------------------------------------|
|    MAIN VALUES    |             MAIN MAGNETIC AXES    |  x , y , z  -- initial Cartesian␣
↪axes
|-------------------|----|----- x ------- y ------- z ---|  Xm, Ym, Zm -- main magnetic axes
| gX =   0.09871069 | Xm | -0.456536 -0.363638  0.811998 |
| gY =   0.11729280 | Ym |  0.643532  0.495246  0.583605 |
| gZ =  11.21949040 | Zm | -0.614361  0.788985  0.007914 |
|-------------------------------------------------------|
The sign of the product gX * gY * gZ for multiplet 1: < 0.
```

The section above shows the g tensor for the ground Kramers doublet. Since the g_X and g_Y are much smaller than the g_Z component, the Zm axis is denoted as the $main magnetic axis$ of the computed molecule. The "Zm | -0.614361 0.788985 0.007914 |" denotes the Cartesian components of the Zm vector.

In the case the computation of the parameters of the crystal field is requested by `CRYS_element` and `CRYS_charge`, the following section will be found in the output:

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
     CALCULATION OF CRYSTAL-FIELD PARAMETERS OF THE GROUND ATOMIC TERM, L =  3.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

The parameters of the Crystal Field matrix are written in the coordinate system:
(Xm, Ym, Zm) --  the main magnetic axes of the ground pseudo-L = | 3> orbital multiplet.
Rotation matrix from the initial coordinate system to the employed coordinate system is:
-------------------------------------------------------------------|
x , y , z  -- initial Cartesian axes                               |
Xm, Ym, Zm -- main magnetic axes                                   |
                     x                 y                 z         |
      | Xm | -0.61155332461133 -0.79120321735748  0.00000000001900 |
 R =  | Ym |  0.79120321735748 -0.61155332461133 -0.00000000000264 |
      | Zm |  0.00000000001371  0.00000000001342  1.00000000000000 |
-------------------------------------------------------------------|
Quantization axis is Zm.


-----------------------------------------------------------------------------------------------|
   Ab Initio Crystal-Field Splitting Matrix written in the basis of Pseudo-L Eigenfunctions   |
-----------------------------------------------------------------------------------------------|
          |                    | -3 >                |                    | -2 >                |
----------|---------- REAL ----------- IMAGINARY ---|---------- REAL ----------- IMAGINARY ---|
   < -3 | |  -777.2218617165776     0.0000000000000 |    -0.0000000454613     0.0000000351393 |
   < -2 | |    -0.0000000454613    -0.0000000351393 |   285.4563720817839     0.0000000000000 |
   < -1 | |     0.0285952366255    -0.0055088548681 |     0.0000000002580    -0.0000000024089 |
   <  0 | |    -0.0000000160911     0.0000000604843 |     0.0001933353720    -0.0013498718536 |
   <  1 | |     0.0116130821579    -0.0096419531309 |    -0.0000000005651    -0.0000000026613 |
   <  2 | |     0.0000000070279    -0.0000000332291 |     0.0141002831881    -0.0117070160056 |
   <  3 | |    -0.0000002782243     0.0000001745761 |    -0.0000000070278     0.0000000332291 |
-----------------------------------------------------------------------------------------------|
-----------------------------------------------------------------------------------------------|
          |                    | -1 >                |                    |  0 >                |
----------|---------- REAL ----------- IMAGINARY ---|---------- REAL ----------- IMAGINARY ---|
   < -3 | |     0.0285952366255     0.0055088548681 |    -0.0000000160911    -0.0000000604843 |
   < -2 | |     0.0000000002580     0.0000000024089 |     0.0001933353720     0.0013498718536 |
   < -1 | |   347.8121781289439    -0.0000000000000 |     0.0000000069613    -0.0000000861891 |
   <  0 | |     0.0000000069613     0.0000000861891 |   287.9066230116981     0.0000000000000 |
   <  1 | |    -0.0106576508828     0.0003303251251 |    -0.0000000069613    -0.0000000861891 |
   <  2 | |     0.0000000005651     0.0000000026613 |     0.0001933353720    -0.0013498718536 |
   <  3 | |     0.0116130821579    -0.0096419531309 |     0.0000000160911    -0.0000000604843 |
-----------------------------------------------------------------------------------------------|
-----------------------------------------------------------------------------------------------|
          |                    |  1 >                |                    |  2 >                |
----------|---------- REAL ----------- IMAGINARY ---|---------- REAL ----------- IMAGINARY ---|
   < -3 | |     0.0116130821579     0.0096419531309 |     0.0000000070279     0.0000000332291 |
   < -2 | |    -0.0000000005651     0.0000000026613 |     0.0141002831881     0.0117070160056 |
   < -1 | |    -0.0106576508828    -0.0003303251251 |     0.0000000005651    -0.0000000026613 |
   <  0 | |    -0.0000000069613     0.0000000861891 |     0.0001933353720     0.0013498718536 |
   <  1 | |   347.8121781289439    -0.0000000000000 |    -0.0000000002580    -0.0000000024089 |
   <  2 | |    -0.0000000002580     0.0000000024089 |   285.4563720817837     0.0000000000000 |
   <  3 | |     0.0285952366255    -0.0055088548681 |     0.0000000454613     0.0000000351392 |
-----------------------------------------------------------------------------------------------|
----------------------------------------------------|
          |                    |  3 >                |
----------|---------- REAL ----------- IMAGINARY ---|
   < -3 | |    -0.0000002782241    -0.0000001745761 |
   < -2 | |    -0.0000000070278    -0.0000000332291 |
   < -1 | |     0.0116130821579     0.0096419531309 |
   <  0 | |     0.0000000160911     0.0000000604843 |
   <  1 | |     0.0285952366255     0.0055088548681 |
   <  2 | |     0.0000000454613    -0.0000000351392 |
```

(continues on next page)

(continued from previous page)

```
   <  3 | |  -777.2218617165773     0.0000000000000 |
-----------------------------------------------------|
```

In the above section, the low-lying CASSCF states of the Co^{2+} site originating from the free ion ^{4}F term are transformed towards the eigenstates of the ($\tilde{L} = 3$), and the low-lying CASSCF *diagonal* 7×7 energy matrix is re-written in this basis. The *non-diagonal* "Ab Initio Crystal-Field Splitting Matrix" is printed in the above section. In the subsequent output sections, the obtained crystal field matrix is decomposed in a linear combination of Irreducible Tensorial Operators (ITOs) and the obtained expansion coefficients are listed in the output.

$$\hat{H}_{CF} = \sum_{k}^{2L} \sum_{q}^{-k,+k} B_k^q \hat{O}_k^q(\tilde{L})$$

The parameters are given for several sets of ITOs.

```
*****************************************************************************
The Crystal-Field Hamiltonian:
   Hcf = SUM_{k,q} * [ B(k,q) * O(k,q) ];
where:
   O(k,q) =  Extended Stevens Operators (ESO)as defined in:
          1. Rudowicz, C.; J.Phys.C: Solid State Phys.,18(1985) 1415-1430.
          2. Implemented in the "EasySpin" function in MATLAB, www.easyspin.org.
   k - the rank of the ITO, = 2, 4, 6, 8, 10, 12.
   q - the component of the ITO, = -k, -k+1, ... 0, 1, ... k;
Knm are proportionality coefficients between the ESO and operators defined in
J. Chem. Phys., 137, 064112 (2012).
-----------------------------------------------|
  k |  q  |    (K)^2    |         B(k,q)        |
----|-----|-------------|-----------------------|
  2 |  -2 |       1.50  |  0.44029016547734E-03 |
  2 |  -1 |       6.00  |  0.24547763681975E-08 |
  2 |   0 |       1.00  | -0.43693326103120E+02 |
  2 |   1 |       6.00  | -0.84162317914775E-08 |
  2 |   2 |       1.50  |  0.12672200639220E-02 |
----|-----|-------------|-----------------------|
  4 |  -4 |      17.50  |  0.20185049671189E-03 |
  4 |  -3 |     140.00  | -0.24080325997038E-08 |
  4 |  -2 |      10.00  |  0.53646717565242E-04 |
  4 |  -1 |      20.00  |  0.99248109880376E-09 |
  4 |   0 |       1.00  | -0.67496280141952E+00 |
  4 |   1 |      20.00  | -0.53708624488205E-09 |
  4 |   2 |      10.00  |  0.33333801678482E-03 |
  4 |   3 |     140.00  | -0.66502585057483E-09 |
  4 |   4 |      17.50  |  0.24311509626065E-03 |
----|-----|-------------|-----------------------|
  6 |  -6 |      57.75  | -0.48493356946616E-09 |
  6 |  -5 |     693.00  |  0.45219078679397E-09 |
  6 |  -4 |      31.50  |  0.11222605476277E-05 |
  6 |  -3 |     105.00  | -0.14936428088413E-09 |
  6 |  -2 |      26.25  |  0.68538037767693E-06 |
  6 |  -1 |      42.00  | -0.24067665895440E-09 |
  6 |   0 |       1.00  | -0.18259217459128E-02 |
  6 |   1 |      42.00  |  0.11394843516111E-11 |
  6 |   2 |      26.25  |  0.48314159464149E-05 |
  6 |   3 |     105.00  | -0.33636623245296E-10 |
  6 |   4 |      31.50  |  0.13517294099051E-05 |
  6 |   5 |     693.00  |  0.95637007025194E-10 |
  6 |   6 |      57.75  | -0.77284500243776E-09 |
-----------------------------------------------|
```

In the sections below, the weight of various expansion terms on the total energy splitting of the corresponding term or multiplet is analysed.

```
CUMULATIVE WEIGHT OF INDIVIDUAL-RANK OPERATORS ON THE CRYSTAL FIELD SPLITTING:
  O2 :-----------------------------------------:  70.094642 %.
  O2 + O4 :------------------------------------:  99.417093 %.
  O2 + O4 + O6 :-------------------------------: 100.000000 %.

ENERGY SPLITTING INDUCED BY CUMULATIVE INDIVIDUAL-RANK OPERATORS (in cm-1).
----------|---------------|---------------|---------------|---------------|
  L = 3   |     RASSCF    |     ONLY      |     ONLY      |     ONLY      |
          |     INITIAL   |      O2       |     O2+O4     |   O2+O4+O6    |
----------|---------------|---------------|---------------|---------------|
 w.f.  1  |    0.00000000 |    0.00000000 |    0.00000000 |    0.00000000 |
 w.f.  2  |    0.00000220 |    0.00000000 |    0.00000149 |    0.00000220 |
 w.f.  3  | 1062.65990786 |  655.39989137 | 1058.22649805 | 1062.65990786 |
 w.f.  4  | 1062.69656233 |  655.39989157 | 1060.35861560 | 1062.69656233 |
 w.f.  5  | 1065.12848829 | 1048.63177735 | 1060.39653667 | 1065.12848829 |
 w.f.  6  | 1125.02338078 | 1048.64787570 | 1129.62295551 | 1125.02338078 |
 w.f.  7  | 1125.04470493 | 1179.71980502 | 1129.64777494 | 1125.04470493 |
----------|---------------|---------------|---------------|---------------|


WEIGHT OF INDIVIDUAL-RANK OPERATORS ON THE CRYSTAL FIELD SPLITTING:
  O2  :-----------------------------------------:  70.094642 %.
  O4  :-----------------------------------------:  29.322451 %.
  O6  :-----------------------------------------:   0.582907 %.


ENERGY SPLITTING INDUCED BY INDIVIDUAL-RANK OPERATORS (in cm-1).
----------|---------------|---------------|---------------|---------------|
  L = 3   |     RASSCF    |     ONLY      |     ONLY      |     ONLY      |
          |     INITIAL   |      O2       |      O4       |      O6       |
----------|---------------|---------------|---------------|---------------|
 w.f.  1  |    0.00000000 |    0.00000000 |    0.00000000 |    0.00000000 |
 w.f.  2  |    0.00000220 |    0.00000000 |  121.49328633 |    0.00351330 |
 w.f.  3  | 1062.65990786 |  655.39989137 |  121.49330341 |    4.60307936 |
 w.f.  4  | 1062.69656233 |  655.39989157 |  202.46860036 |    4.60307986 |
 w.f.  5  | 1065.12848829 | 1048.63177735 |  202.50909974 |    6.90310777 |
 w.f.  6  | 1125.02338078 | 1048.64787570 |  526.45202593 |    6.90437338 |
 w.f.  7  | 1125.04470493 | 1179.71980502 |  526.48994463 |   11.50506445 |
----------|---------------|---------------|---------------|---------------|


WEIGHT OF INDIVIDUAL CRYSTAL FIELD PARAMETERS ON THE CRYSTAL FIELD SPLITTING: (in descending␣
↪order):
CFP are given in ITO used in J. Chem. Phys. 137, 064112 (2012).
-------------------------------------------------------|
  k |  q  |          B(k,q)          |    Weight (in %)    |
----|-----|--------------------------|---------------------|
2 |   0 | -0.43693326103120E+02 |  70.08577012260780 |
4 |   0 | -0.67496280141952E+00 |  29.31873954450217 |
6 |   0 | -0.18259217459128E-02 |   0.58283348856981 |
4 |   2 |  0.10541073637635E-03 |   0.00457878555457 |
4 |   4 |  0.58115623682363E-04 |   0.00252440109385 |
4 |  -4 |  0.48251497695683E-04 |   0.00209592749496 |
2 |   2 |  0.10346808494753E-02 |   0.00165966774870 |
4 |  -2 |  0.16964581649793E-04 |   0.00073690009261 |
2 |  -2 |  0.35949541472835E-03 |   0.00057664442705 |
6 |   2 |  0.94299583491020E-06 |   0.00030100389209 |
6 |   4 |  0.24084325388052E-06 |   0.00007687706999 |
6 |  -4 |  0.19995783180553E-06 |   0.00006382645967 |
6 |  -2 |  0.13377255211448E-06 |   0.00004270014495 |
6 |   6 | -0.10169893585902E-09 |   0.00000003246226 |
6 |  -6 | -0.63812572794630E-10 |   0.00000002036895 |
```

(continues on next page)

(continued from previous page)

```
6 |  -1 | -0.37137214734348E-10 |   0.00000001185418 |
4 |  -1 |  0.22192552033089E-09 |   0.00000000963990 |
4 |  -3 | -0.20351589971646E-09 |   0.00000000884023 |
2 |   1 | -0.34359122410183E-08 |   0.00000000551133 |
6 |  -5 |  0.17177307577593E-10 |   0.00000000548299 |
4 |   1 | -0.12009613533364E-09 |   0.00000000521668 |
6 |  -3 | -0.14576461261073E-10 |   0.00000000465280 |
4 |   3 | -0.56204942711777E-10 |   0.00000000244141 |
2 |  -1 |  0.10021582557877E-08 |   0.00000000160750 |
6 |   5 |  0.36329494838219E-11 |   0.00000000115964 |
6 |   3 | -0.32825983078827E-11 |   0.00000000104780 |
6 |   1 |  0.17582625268298E-12 |   0.00000000005612 |
----|-----|-----------------------|--------------------|
```

In the case of lanthanide compounds, the same keywords (`CRYS_element` and `CRYS_charge`) trigger the energy decomposition of the lowest energy matrix corresponding to the ground $J-$ multiplet of the respective lanthanide ion.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
CALCULATION OF THE MAGNETIC SUSCEPTIBILITY
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

  Temperature dependence of the magnetic susceptibility will be calculated in 301 points,
  equally distributed in temperature range  0.0 --- 300.0 K.

|-----------------------------------------------------------------------------------------------------------|
|     |      T      | STATISTICAL |     X*T      |     X*T      |      X       |     1/X      |
|     |             |  SUM (Z)    |    (zJ=0)    |              |              |              |
|-----|-----------------------------------------------------------------------------------------------------|
|Units|   Kelvin    |    ---      |  cm3*K/mol   |  cm3*K/mol   |   cm3/mol    |   mol/cm3    |
|-----|-----------------------------------------------------------------------------------------------------|
|     |    0.000100 | 0.20000E+01 |   3.93592010 |   3.93592010 | 0.39359E+05 | 0.25407E-04 |
|     |    1.000000 | 0.20000E+01 |   3.94046247 |   3.94046247 | 0.39405E+01 | 0.25378E+00 |
|     |    2.000000 | 0.20000E+01 |   3.94500530 |   3.94500530 | 0.19725E+01 | 0.50697E+00 |
|     |    3.000000 | 0.20000E+01 |   3.94954814 |   3.94954814 | 0.13165E+01 | 0.75958E+00 |
|     |    4.000000 | 0.20000E+01 |   3.95409097 |   3.95409097 | 0.98852E+00 | 0.10116E+01 |
|     |    5.000000 | 0.20000E+01 |   3.95863380 |   3.95863380 | 0.79173E+00 | 0.12631E+01 |
|     |    6.000000 | 0.20000E+01 |   3.96317663 |   3.96317663 | 0.66053E+00 | 0.15139E+01 |
|     |    7.000000 | 0.20000E+01 |   3.96771946 |   3.96771946 | 0.56682E+00 | 0.17642E+01 |
|     |    8.000000 | 0.20000E+01 |   3.97226229 |   3.97226229 | 0.49653E+00 | 0.20140E+01 |
|     |    9.000000 | 0.20000E+01 |   3.97680512 |   3.97680512 | 0.44187E+00 | 0.22631E+01 |
|     |   10.000000 | 0.20000E+01 |   3.98134795 |   3.98134795 | 0.39813E+00 | 0.25117E+01 |
...
```

This section shows the computed magnetic susceptibility. The formula used for this calculation assumes the zero-field limit, $i.e. H = 0.0$ Tesla. A picture called "XT_no_field.png" using the above data will be created in the working directory whenever the `PLOT` keyword is included in the `SINGLE_ANISO` input. The picture shows the temperature dependence of the magnetic susceptibility.

```
|---------------------------------------------------------------------------------------------------------------
↪--------|
|                      VAN VLECK SUSCEPTIBILITY TENSOR FOR zJ = 0,  in cm3*K/mol                              ␣
↪        |
|---------------------------------------------------------------------------------------------------------------
↪--------|
|   T(K)    | |        SUSCEPTIBILITY TENSOR       |  MAIN VALUES  |               MAIN AXES                 ␣
↪        |
|----------|-|----- x --------- y --------- z ---|---------------|------ x --------- y --------
↪- z ----|
|          |x|  4.456611   -5.721848   -0.057261 | X:   0.000914 |  0.45654560  0.36364537 -0.
↪81199025 |
```

(continues on next page)

(continued from previous page)

```
| 0.000100 |y| -5.721848    7.349367    0.073827 | Y:   0.001291 |  0.64352653  0.49524178  0.
↪58361733 |
|          |z| -0.057261    0.073827    0.001782 | Z:  11.805555 | -0.61436123  0.78898519  0.
↪00791499 |
|----------|-|----- x --------- y --------- z ---|---------------|------ x --------- y --------
↪- z ----|
|          |x|  4.461142   -5.718873   -0.057275 | X:   0.007129 |  0.48578721  0.38619740 -0.
↪78413160 |
| 1.000000 |y| -5.718873    7.351927    0.074460 | Y:   0.008275 |  0.62173382  0.47788368  0.
↪62054351 |
|          |z| -0.057275    0.074460    0.008319 | Z:  11.805983 | -0.61437598  0.78897323  0.
↪00796220 |
|----------|-|----- x --------- y --------- z ---|---------------|------ x --------- y --------
↪- z ----|
|          |x|  4.465674   -5.715898   -0.057290 | X:   0.013344 | -0.49137357 -0.39055217  0.
↪77847352 |
| 2.000000 |y| -5.715898    7.354486    0.075093 | Y:   0.015261 |  0.61731357  0.47435127  0.
↪62762635 |
|          |z| -0.057290    0.075093    0.014856 | Z:  11.806411 | -0.61439073  0.78896126  0.
↪00800947 |
...
```

The section above shows how the main axes of the susceptibility tensor evolves with temperature.

```
        HIGH-FIELD POWDER MAGNETIZATION
              (Units: Bohr magneton)

|-----------|---------------|---------------|---------------|---------------|---------------|
|    H(T)   |STATISTICAL SUM|     1.000 K.  |     1.200 K.  |     1.800 K.  |     2.500 K.  |
|-----------|---------------|---------------|---------------|---------------|---------------|
|  0.000100 |     1.9995371 |  0.0007055560 |  0.0005880989 |  0.0003923371 |  0.0002827105 |
|  0.100000 |     1.6293687 |  0.6863212310 |  0.5768343572 |  0.3889480349 |  0.2814379254 |
|  0.200000 |     1.3961049 |  1.2730827904 |  1.0928431358 |  0.7585259147 |  0.5554332924 |
|  0.300000 |     1.2492960 |  1.7176941099 |  1.5137449727 |  1.0936561795 |  0.8154486026 |
|  0.400000 |     1.1568991 |  2.0312460704 |  1.8358752195 |  1.3858429640 |  1.0565149122 |
|  0.500000 |     1.0987474 |  2.2456644189 |  2.0736324473 |  1.6329635867 |  1.2755212118 |
|  0.600000 |     1.0621485 |  2.3917509695 |  2.2464760254 |  1.8374994655 |  1.4711430994 |
|  0.700000 |     1.0391143 |  2.4924803644 |  2.3720174309 |  2.0044495622 |  1.6435265950 |
|  0.800000 |     1.0246173 |  2.5633469179 |  2.4639356325 |  2.1396808073 |  1.7938697856 |
|  0.900000 |     1.0154934 |  2.6144012337 |  2.5321303741 |  2.2489088878 |  1.9240150206 |
|  1.000000 |     1.0097510 |  2.6521008670 |  2.5835380488 |  2.3371993536 |  2.0361143785 |
|  1.100000 |     1.0061370 |  2.6806180412 |  2.6229602467 |  2.4088035375 |  2.1323879997 |
|  1.200000 |     1.0038624 |  2.7026854246 |  2.6537186141 |  2.4671744880 |  2.2149684436 |
|  1.300000 |     1.0024309 |  2.7201250266 |  2.6781249799 |  2.5150622492 |  2.2858129074 |
|  1.400000 |     1.0015299 |  2.7341759016 |  2.6978046065 |  2.5546323752 |  2.3466634357 |
...
```

This section shows the field dependence of the powder molar magnetisation. A picture named "MH.png" can be created by using the `PLOT` keyword in the `SINGLE_ANISO` input file.

Running `CASSCF` calculations on lanthanides compounds in ORCA might be a bit more cumbersome compared to transition metal compounds, due to the convergence of this method. However, following the instructions in the *The Complete Active Space Self-Consistent Field (CASSCF) Module* section and the related tips in this manual and on the Forum, the calculations could be performed. From our experience, the main reason for the poor convergence of `CASSCF` calculation originates from the wrong orbitals occupying the active space. This issue can be overcame by performing a proper rotation of the molecular orbitals such that the seven orbitals with dominant $4f$ contribution are placed in the active space. As soon as the active orbitals acquire the dominant $4f$ weight, the convergence is quite straightforward.

Below we describe the calculation on a lanthanide fragment $[Ce(COT)_2]^-$ (COT=$(C_8H_8)^{2-}$) as an example:

```
!DKH DKH-DEF2-SVP slowconv KDIIS BP

%basis
newgto Ce "SARC2-DKH-QZVP" end
end

%scf
MaxIter 500
end

*xyz -1 2
Ce   5.97600100    5.09133100   13.17268800
C    5.47882500    2.98632700   11.42941100
C    4.38424700    3.88677900   11.27367600
H    4.21867900    4.06431900   10.35453600
C    3.47138800    4.59373300   12.09958800
H    2.87027600    5.12178400   11.58692300
C    3.21937800    4.72005000   13.49499100
C    3.84198900    4.08874200   14.61728000
H    3.46926600    4.38472800   15.43857500
C    4.86395700    3.14327700   14.81900800
H    4.97094200    2.92197400   15.73690500
C    5.77247800    2.44085400   13.98883400
H    6.34594800    1.86846400   14.48498500
C    6.03182900    2.38537000   12.60408600
H    6.75450600    1.80396200   12.40022700
C    6.40698800    7.65195200   12.14877700
C    6.11546300    7.83965300   13.53689900
H    5.47247100    8.52635500   13.66635300
C    6.52698400    7.27593700   14.77461400
H    6.07344100    7.67704500   15.50605500
C    7.44425200    6.26569500   15.21837900
C    8.37896000    5.47470700   14.49350900
H    8.88315100    4.90771300   15.06566300
C    8.74883600    5.31701900   13.13468700
H    9.45277500    4.68873000   13.02529900
C    8.32602800    5.86701900   11.90372100
H    8.81295200    5.51473700   11.16784000
C    7.36115600    6.80638600   11.50204100
H    7.33662500    6.90506000   10.55697400
H    5.93270067    2.68976505   10.50694264
H    5.83417492    8.22959522   11.45371334
H    2.43475960    5.39234559   13.77300645
H    7.40021961    6.07954201   16.27114118
*
```

This is the first step of the calculation. For heavier elements like lanthanides, we must consider relativistic effect by using DKH keyword. We explicitly use KDIIS in the calculation to smoothen out convergence. The orbital file called "CeCOT2_1.gbw" will be generated after this step. We further use this gbw file to do the CASSCF calculation.

```
!DKH DKH-DEF2-SVP TightSCF conv Moread

%moinp "CeCOT2_1.gbw"

%basis
newgto Ce "SARC2-DKH-QZVP" end
end

%casscf nel 1
   norb 7  # 1 electrons in 7 f orbitals
   mult 2
   nroots 7  # 7 doublet states
```

(continues on next page)

(continued from previous page)

```
   rel
     dosoc true  # include spin-orbit coupling
   end
end

*xyz -1 2
......
*
```

We need to check the orbitals after the CASSCF step with the orbital file named "CeCOT2_2.gbw" obtained.

```
              85        86        87        88        89        90        91
            0.45543   0.45310   0.27655   0.45085   0.45251   0.45760   0.45713
            0.14286   0.14286   0.14286   0.14286   0.14286   0.14286   0.14286
            --------  --------  --------  --------  --------  --------  --------
0 Ce f0       26.2       0.9       0.0       1.3       3.3       0.5       6.6
0 Ce f+1       5.4      20.2       0.0      24.0       3.4       5.6      39.1
0 Ce f-1       0.4      30.4       0.0      52.8      10.7       0.6       3.9
0 Ce f+2       3.8       0.2       0.7       1.4       4.7      76.7      10.7
0 Ce f-2      50.7       0.7       0.0       2.1       4.2       4.0       1.3
0 Ce f+3       8.2      22.0       0.0       1.2      55.8       0.0      11.6
0 Ce f-3       4.7      25.3       0.0      16.5      17.1      10.1      25.2
```

Orbitals 85, 86, 88-91 and 130 are occupied and strongly metal based $4f$-orbitals. For comparison, the converged CASSCF orbitals are pure $4f$-orbitals (99% metal-based). The orbitals need to be rotated in order to fit the active space (85-91). Then we can use the results of CASSCF/SOC calculation to call for the SINGLE_ANISO program.

```
!DKH DKH-DEF2-SVP TightSCF conv Moread
%moinp "CeCOT2_2.gbw"

%basis
newgto Ce "SARC2-DKH-QZVP" end
end

%scf rotate {87,130,90} end
end

%casscf nel 1
   norb 7
   mult 2
   nroots 7

   rel
   dosoc true
   end

   ANISO
     doaniso true
     MLTP 2,2,2    # 3 Kramers doublets, J=5/2
     MAVE 1, 12   # nsym=1, Lebedev grid number 12
     XFIE 0.1        # the applied magnetic field is 0.1 T
     CRYS_element "Ce"
     CRYS_charge 3
     NCUT 14
     ABCC_abc     11.0735, 12.6738, 22.4854, 84.436, 86.690, 83.969
     ABCC_center  0.82682, 0.31234, 0.78619
     ZJPR  -0.120
     HEXP_temp  2.0, 3.0
     HEXP_H  0.0, 1.0, 2.0, 3.0, 4.0
     HEXP_M[0]= 0.0, 2.46, 2.86, 2.95, 2.98
```

(continues on next page)

(continued from previous page)

```
    HEXP_M[1]= 0.0, 2.04, 2.68, 2.87, 2.94
    TEXP_temp  0.0, 10.0, 20.0, 30.0, 40.0, 50.0
    TEXP_chiT  4.5, 4.5, 4.58, 4.62, 4.66, 4.70
    UBAR true
    PLOT true
    ZEEM[0]=1.0, 0.0, 0.0
    ZEEM[1]=0.0, 0.0, 1.0
    ZEEM[2]=0.0, 1.0, 0.0
    ZEEM[3]=0.75, 0.0, 0.25
  end
end

*xyz -1 2
......
*
```

The order of the keywords listed in the CASSCF/ANISO subblock does not matter.

7.17.3 Reference list of CASSCF/ANISO keywords

The only required keyword for SINGLE_ANISO is the DATA, specifying the name of the datafile containing the *ab initio* information. The ORCA interface includes this keyword automatically and therefore it is not referenced here. All other keywords are extra and allow various customisation of the execution. For the computation of the EPR g-tensor, the only unknown variable for SINGLE_ANISO is the dimension (multiplicity) of the pseudospin(s). This information can be provided by the MLTP keyword. For example, in cases where spin-orbit coupling is weak, the multiplicity of the effective spin Hamiltonian is usually the same as the multiplicity of the lowest term (e.g. high spin Fe^{3+}: $S = \tilde{S} = 5/2$), while in the cases with strong anisotropy (lanthanide, actinide complexes, Co^{2+} complexes, cases with near-orbital degeneracy, etc.) the lowest energy levels form a group of states which may differ drastically from the spins of the lowest term. In these cases the user should specify the multiplicity corresponding to a chosen value of pseudospin ($2\tilde{S}+1$). For instance, in Dy^{3+} the spin of the ground state term is $S = 5/2$, but in most of real compounds only the ground Kramers doublet is considered. In such case, the multiplicity of the pseudospin equals to 2 (see MLTP keyword). For the calculation of the parameters of the crystal field corresponding to the ground atomic multiplet J for lanthanides should be requested with the keywords CRYS_element and CRYS_charge. Similarly, the parameters of the crystal field corresponding to the ground atomic term L for lanthanides and transition metals compounds can be requested with same keywords: CRYS_element and CRYS_charge.

Note that the keywords/syntax in the ORCA CASSCF/ANISO block are slightly different from the genuine SINGLE_ANISO input, where some of the keywords are grouped together. We aimed at keeping the control keywords as close as possible.

Optional general keywords to control the input within the ORCA interface (CASSCF/ANISO subblock):

InputNameOnDisk
: This keyword reads the name of the input file for SINGLE_ANISO, a string given between quotations. Example:

```
InputNameOnDisk "my_input_for_aniso.inp"
```

The interface with ORCA will add the DATA keyword with specific name of the datafile for the performed calculation. All the other keywords provided inside this file must follow their original names, as in MOLCAS.

MLTP
: The number of molecular multiplets ($i.e.$ groups of spin-orbital eigenstates) for which g, D and higher magnetic tensors will be calculated (default MLTP=1). With MLTP an comma separated list of numbers specifying the dimension of each multiplet is passed. The default is to select one multiplet which has the dimension equal to the multiplicity of the ground term. In cases of strong spin-orbit coupling the usage of this keyword is mandatory. Example:

```
MLTP 4, 4, 2, 2
```

SINGLE_ANISO will compute the g tensor for 4 groups of states: 2 groups having the effective spin $S = |3/2>$, and other 2 groups of states being Kramers doublets.

TINT

Specifies the temperature points for the evaluation of the magnetic susceptibility. The program will read three numbers: T_{min}, T_{max}, and nT.

- T_{min} - the minimal temperature (Default 0.0K)
- T_{max} - the maximal temperature (Default 300.0K)
- nT - number of temperature points (Default 101)

Example:

```
TINT 0.0, 330.0, 331
```

SINGLE_ANISO will compute temperature dependence of the magnetic susceptibility in 331 points evenly distributed in temperature interval: 0.0K - 330.0K.

HINT

Specifies the field points for the evaluation of the magnetisation in a certain direction. The program will read three numbers: H_{min}, H_{max} and nH.

- H_{min} - the minimal field (Default 0.0T)
- H_{max} - the maximal filed (Default 10.0T)
- nH - number of field points (Default 101)

Example:

```
HINT 0.0, 20.0, 201
```

SINGLE_ANISO will compute the molar magnetisation in 201 points evenly distributed in field interval: 0.0T - 20.0T.

TMAG

Specifies the temperature(s) at which the field-dependent magnetisation is calculated. The program will read the temperatures (in Kelvin) at which magnetisation is to be computed. Default is to compute magnetisation at one temperature point (2.0 K). Example:

```
TMAG 1.8, 2.0, 3.0, 4.0, 5.0
```

SINGLE_ANISO will compute the molar magnetisation at 5 temperature points (1.8 K, 2.0 K, 3.4 K, 4.0 K, and 5.0 K).

ENCU

The keyword expects to read two integer numbers. The two parameters (NK and MG) are used to define the cut-off energy for the lowest states for which Zeeman interaction is taken into account exactly. The contribution to the magnetisation coming from states that are higher in energy than E (see below) is done by second order perturbation theory. The program will read two integer numbers: NK and MG. Default values are: $NK = 100, MG = 100$.

$$E = NK \cdot k_{Boltz} \cdot \texttt{TMAG}_{max} + MG \cdot \mu_{Bohr} \cdot H_{max}$$

The field-dependent magnetisation is calculated at the maximal temperature value given by TMAG keyword. Example:

```
ENCU 250, 150
```

If H_{max} = 10 T and TMAG = 1.8 K, then the cut-off energy is: $$E = 250 \cdot k_{Boltz} \cdot 1.8 + 150 \cdot \mu_{Bohr} \cdot 10 = 1013.06258(cm^{-1})$$

This means that the magnetisation arising from all spin-orbit states with energy lower than $E = 1013.06258(cm^{-1})$ will be computed exactly (i.e. are included in the exact Zeeman diagonalisation) The

keywords NCUT, ERAT and ENCU have similar purpose. If two of them are used at the same time, the following priority is defined: NCUT > ENCU > ERAT.

NCUT

This flag is used to define the cut-off energy for the low-lying spin-orbit states for which Zeeman interaction is taken into account exactly. The contribution to the magnetisation arising from states that are higher in energy than lowest N_{CUT} states, is done by second-order perturbation theory. The program will read one integer number. In case the number is larger than the total number of spin-orbit states(N_{SS}, then the N_{CUT} is set to N_{SS} (which means that the molar magnetisation will be computed exactly, using full Zeeman diagonalisation for all field points). The field-dependent magnetisation is calculated at the temperature value(s) defined by TMAG. Example:

```
NCUT 32
```

The keywords NCUT, ERAT and ENCU have similar purpose. If two of them are used at the same time, the following priority is defined: NCUT > ENCU > ERAT.

ERAT

This flag is used to define the cut-off energy for the low-lying spin-orbit states for which Zeeman interaction is taken into account exactly. The program will read one single real number specifying the ratio of the energy states which are included in the exact Zeeman Hamiltonian. As example, a value of 0.5 means that the lowest half of the energy states included in the spin-orbit calculation are used for exact Zeeman diagonalisation.

Example:

```
ERAT 0.333
```

The keywords NCUT, ERAT and ENCU have similar purpose. If two of them are used at the same time, the following priority is defined: NCUT > ENCU > ERAT.

MVEC_x MVEC_y MVEC_z

MVEC_x, MVEC_y and MVEC_z define a number of directions for which the magnetisation vector will be computed. The directions are given as unitary vectors specifying the direction *i* of the applied magnetic field).

Example:

```
MVEC_x 0.00, 1.57, 1.57, 0.425
MVEC_y 0.00, 0.00, 1.57, 0.418
MVEC_z 0.00, 0.00, 1.57, 0.418
```

ZEEM

This keyword allows to compute Zeeman splitting spectra along certain directions of applied field. Directions of applied field are given as three real number for each direction, specifying the projections along each direction: Example:

```
ZEEM[0] 1.0, 0.0, 0.0
ZEEM[1] 0.0, 1.0, 0.0
ZEEM[2] 0.0, 0.0, 1.0
ZEEM[3] 0.0, 1.0, 1.0
ZEEM[4] 1.0, 0.0, 1.0
ZEEM[5] 1.0, 1.0, 0.0
```

The above input will request computation of the Zeeman spectra along six directions: Cartesian axes X, Y, Z (directions 1,2 and 3), and between any two Cartesian axes: YZ, XZ and XY, respectively. The program will re-normalise the input vectors according to unity length. In combination with PLOT keyword, the corresponding zeeman_energy_xxx.png images will be produced.

MAVE

The keyword requires two integer numbers, denoted MAVE_nsym and MAVE_ngrid. The parameters MAVE_nsym and MAVE_ngrid specify the grid density in the computation of powder molar magnetisation. The program uses Lebedev-Laikov distribution of points on the unit sphere. The parameters are integer numbers: n_{sym} and n_{grid}. The n_{sym} defines which part of the sphere is used for averaging. It takes one of

the three values: 1 (half-sphere), 2 (a quarter of a sphere) or 3 (an octant of the sphere). n_{grid} takes values from 1 (the smallest grid) till 32 (the largest grid, i.e. the densest). The default is to consider integration over a half-sphere (since $M(H) = -M(-H)$): $n_{sym} = 1$ and $n_{sym} = 15$ (i.e 185 points distributed over half-sphere). In case of symmetric compounds, powder magnetisation may be averaged over a smaller part of the sphere, reducing thus the number of points for the integration. The user is responsible to choose the appropriate integration scheme. Note that the program's default is rather conservative.

Example:

```
MAVE 1, 8
```

TEXP_temp TEXP_chiT

The parameters TEXP_temp and TEXP_chiT allow the computation of the magnetic susceptibility $\chi T(T)$ at experimental points. The experimental temperature (in K) and the experimental magnetic susceptibility (in $cm^3 K mol^{-1}$) are read as comma separated list. In the case both TEXP and TINT keywords are given, the TEXP will be used while the TINT input will be ignored.

Example:

```
TEXP_temp 0.0, 10.0, 20.0, 30.0, 40.0, 50.0
TEXP_chiT 4.5, 4.5, 4.58, 4.62, 4.66, 4.70
```

HEXP_temp HEXP_H HEXP_M

The three keywords HEXP_temp, HEXP_H and HEXP_M enable the computation of the molar magnetisation $M_{mol}(H)$ at experimental points. The experimental field strength (in Tesla) and the experimental magnetisation (in μ_{Bohr}) are read as a comma separated list. In the case both HEXP and HINT keywords are given, the HEXP will be used while the HINT input will be ignored. The magnetisation routine will print the standard deviation from the experiment. Example:

```
HEXP_temp 2.0, 3.0
HEXP_H 0.0, 1.0, 2.0, 3.0, 4.0
HEXP_M[0]= 0.0, 2.46, 2.86, 2.95, 2.98  # exp. M at T=2.0 K
HEXP_M[1]= 0.0, 2.04, 2.68, 2.87, 2.94  # exp. M at T=3.0 K
```

ZJPR

This keyword specifies the value (in cm^{-1}) of a phenomenological parameter of a mean molecular field acting on the spin of the complex (the average intermolecular exchange constant). It is used in the calculation of all magnetic properties (not for spin Hamiltonians) (Default is 0.0).

```
ZJPR -0.02
```

TORQ

This keyword specifies the number of angular points for the computation of the magnetisation torque function, $\vec{\tau}_\alpha$ as function of the temperature, field strength and field orientation.

```
TORQ 55
```

The torque is computed at all temperature given by TMAG or HEXP_temp inputs. Three rotations around Cartesian axes X, Y and Z are performed.

PrintLevel

This keyword controls the print level.

- 2 - normal. (Default)
- 3 or larger (debug)

CRYS_element CRYS_charge

The keywords CRYS_element and CRYS_charge request the computation of all 27 Crystal-Field parameters acting on the ground atomic multiplet of a lanthanide. With CRYS_element the chemical symbol of the lanthanide is set. **Note that the element symbol must be enclosed in quotation marks**. The charge is defined with CRYS_charge. By default the program will not compute the parameters of the Crystal-Field.

Example:

```
CRYS_element "Dy" CRYS_charge 3
```

QUAX

This keyword controls the quantisation axis for the computation of the Crystal-Field parameters acting on the ground atomic multiplet of a lanthanide. On the next line, the program will read one of the three values: 1, 2 or 3.

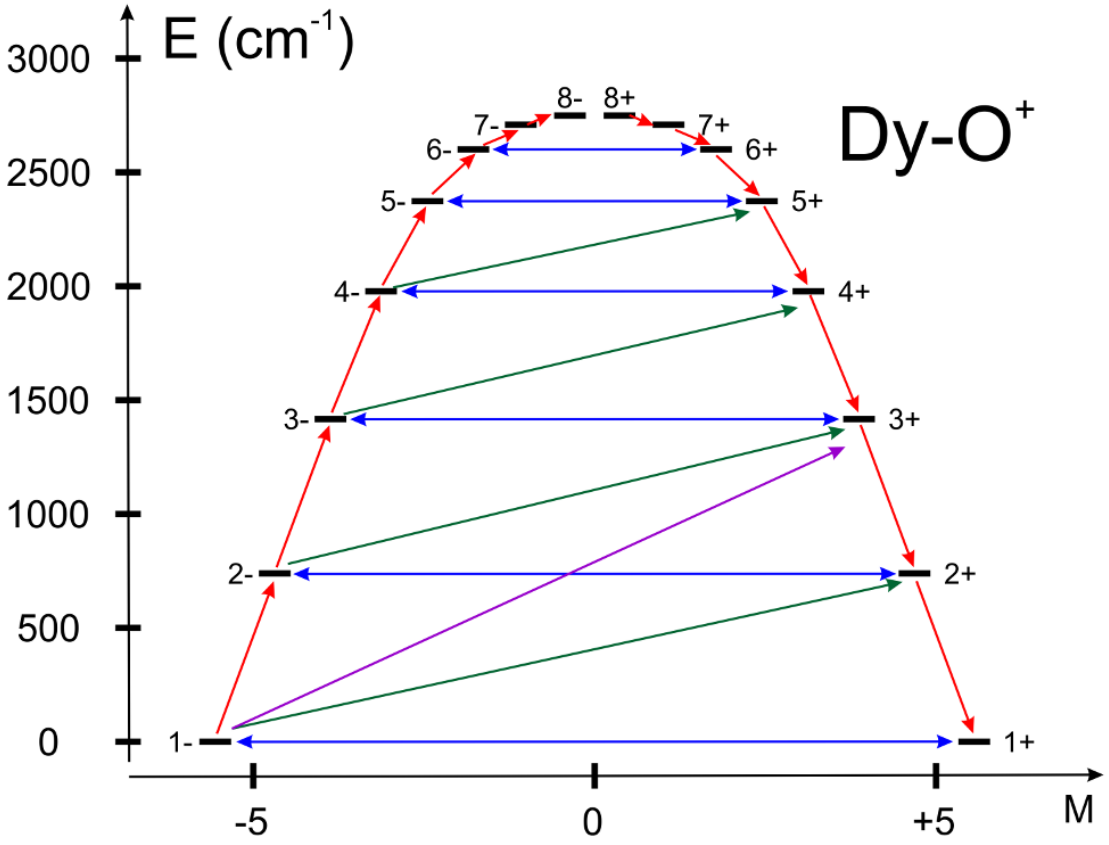
- 1 - quantisation axis is the main magnetic axis Zm of the ground pseudospin multiplet, whose size is specified within the `MLTP` keyword. (Default)
- 2 - quantisation axis is the main magnetic axis Zm of the entire atomic multiplet $|J, M_J>$.
- 3 - quantisation axis is the original Cartesian Z axis. Rotation matrix is unity.

Example:

```
QUAX 3
```

UBAR

With `UBAR` set to "true", the blocking barrier of a single-molecule magnet is estimated. The default is not to compute it. The method prints transition matrix elements of the magnetic moment according to the Figure below:



In this figure, a qualitative performance picture of the investigated single-molecular magnet is estimated by the strengths of the transition matrix elements of the magnetic moment connecting states with opposite magnetisaskytions ($n+ \rightarrow n-$). The height of the barrier is qualitatively estimated by the energy at which the matrix element ($n+ \rightarrow n-$) is large enough to induce significant tunnelling splitting at usual magnetic fields (internal) present in the magnetic crystals (0.01-0.1 Tesla). For the above example, the blocking barrier closes at the state ($8+ \rightarrow 8-$).

All transition matrix elements of the magnetic moment are given as $((|\mu_X| + |\mu_Y| + |\mu_Z|)/3)$. The data is given in Bohr magnetons (μ_{Bohr}).

Example:

```
UBAR true
```

ABCC_abc ABCC_center

The keywords ABCC_abc and ABCC_center set the computation of magnetic and anisotropy axes in the crystallographic abc system. With ABCC_abc, the program reads six real values, namely a, b, c, α, β, and γ, defining the crystal lattice. The values must be separated by a comma. With ABCC_center, the program reads the fractional coordinates of the magnetic center (from the CIF file) - again separated by comma. It is assumed that the XYZ coordinates used for the *ab initio* calculations did not rotate or translate the molecule from its crystallographic position. This input will ensure that all tensors computed by SINGLE_ANISO are given also in the abc system. The computed values in the output correspond to the crystallographic position of three "dummy atoms" located on the corresponding anisotropy axes, at the distance of 1.0 $\AA$ from the metal site. Example:

```
ABCC_abc 12.977, 12.977, 16.573, 90, 90, 120
ABCC_center 0.666667, 0.333333, 0.20413
```

XFIE

This keyword specifies the value (in T) of applied magnetic field for the computation of magnetic susceptibility by dM/dH and M/H formulas. A comparison with the usual formula (in the limit of zero applied field) is provided. (Default is 0.0). Example:

```
XFIE 0.35
```

This keyword together with the keyword PLOT will enable the generation of two additional plots: XT_with_field_dM_over_dH.png and XT_with_field_M_over_H.png, one for each of the two above formula used, alongside with respective gnuplot scripts and gnuplot datafiles.

PLOT

Set to "true", the program generates a few plots (png or eps format) via an interface to the linux program *gnuplot*. The interface generates a datafile, a *gnuplot* script and attempts execution of the script for generation of the image. The plots are generated only if the respective function is invoked. The magnetic susceptibility, molar magnetisation and blocking barrier (UBAR) plots are generated. The files are named: XT_no_field.dat, XT_no_field.plt, XT_no_field.png, MH.dat, MH.plt, MH.png, BARRIER_TME.dat, BARRIER_ENE.dat, BARRIER.plt and BARRIER.png, zeeman_energy_xxx.png etc. All files produced by SINGLE_ANISO are referenced in the corresponding output section. Example:

```
PLOT true
```

7.17.4 How to cite

We would appreciate if you cite the following papers in publications resulting from the use of SINGLE_ANISO:

- Chibotaru, L. F.; Ungur, L. *J. Chem. Phys.*, **2012**, *137*, 064112.
- Ungur, L. Chibotaru, L. F. *Chem. Eur. J.*, **2017**, *23*, 3708-3718.

In addition, useful information like the definition of pseudospin Hamiltonians and their derivation can be found in this paper.

7.18 Interface to POLY_ANISO module

7.18.1 General description

The POLY_ANISO is a stand-alone utility allowing for a semi-*ab initio* description of the (low-lying) electronic structure and magnetic properties of polynuclear compounds. The model behind it is based on the *localised* nature of the magnetic orbitals (i.e. the d and f orbitals containing unpaired electrons). For many compounds of interest, the localised character of the magnetic orbitals leads to very weak character of the exchange interaction between magnetic centers. Due to this weakness of the inter-site interaction, the molecular orbitals and corresponding localised ground and excited states may be optimized in the absence of the magnetic interaction at all. For this purpose, various fragmentation models may be applied. The most commonly used fragmentation model is exemplified in Fig. 7.5:

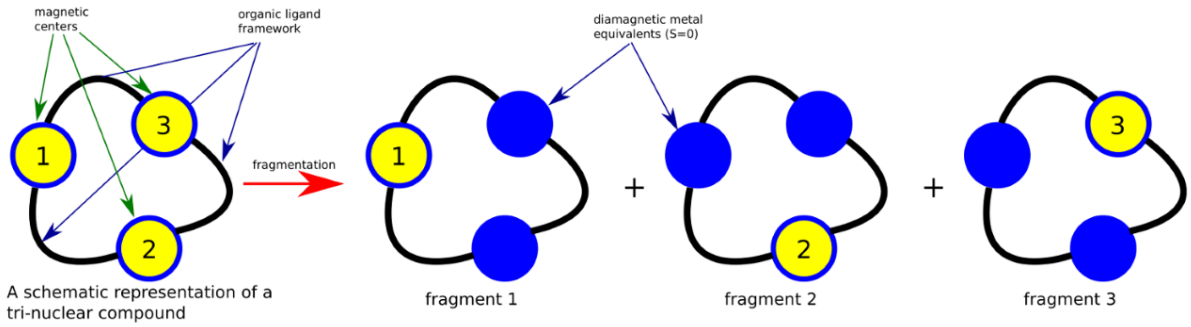


Fig. 7.5: Fragmentation model of a polynuclear compound.

The upper scheme shows a schematic overview of a tri-nuclear compound and the resulting three mononuclear fragments obtained by *diamagnetic atom substitution method*. By this scheme, the neighbouring magnetic centers, containing unpaired electrons are computationally replaced by their diamagnetic equivalents. As example, transition metal (TM) sites are best described by either a diamagnetic Zn(II) or Sc(III), in function of which one is the closest (in terms of charge and atomic radius). For lanthanides (LN), the same principle is applicable, La(III), Lu(III) or Y(III) are best suited to replace a given magnetic lanthanide. Individual mononuclear metal fragments are then investigated by the common CASSCF+SOC/NEVPT2+ SOC/SINGLE_ANISO computational method. A single datafile for each magnetic site, produced by the SINGLE_ANISO run, is needed by the POLY_ANISO code as input.

Magnetic interaction between metal sites is very important for accurate description of low-lying states and their properties. While the full exchange interaction is quite complex (e.g. requiring a multipolar description employing a large set of parameters [413, 870]), in a simplified model it can be viewed as a sum of various interaction mechanisms: magnetic exchange, dipole-dipole interaction, antisymmetric exchange, etc. In the POLY_ANISO code we have implemented several mechanisms, which can be invoked simultaneously for each interacting pair.

The description of the magnetic exchange interaction is done within the Lines model[527]. This model is exact in three cases:

1. interaction between two isotropic spins (Heisenberg),
2. interaction between one Ising spin (only S_Z component) and one isotropic (i.e. usual) spin, and
3. interaction between two Ising spins.

In all other cases when magnetic sites have intermediate anisotropy (i.e. when the spin-orbit coupling and crystal field effects are of comparable strengths), the Lines model represents an approximation. However, it was successfully applied for a wide variety of polynuclear compounds so far.

In addition to the magnetic exchange, magnetic dipole-dipole interaction can be accounted exactly, by using the *ab initio* computed magnetic moment for each metal site (as available inside the datafile). In the case of strongly anisotropic lanthanide compounds (like Ho^{3+} or Dy^{3+}), the magnetic dipole-dipole interaction is usually the dominant one. For example, a system containing two magnetic dipoles $\vec{\mu}_1$ and $\vec{\mu}_2$, separated by distance $\vec{r}$ have a total

energy:

$$E_{dip} = \frac{\mu^2_{Bohr}}{r^3}[\vec{\mu}_1 \cdot \vec{\mu}_2 - 3(\vec{\mu}_1\vec{n}_{1,2}) \cdot (\vec{\mu}_2\vec{n}_{1,2})],$$

where $\vec{\mu}_{1,2}$ are the magnetic moments of sites 1 and 2, respectively; r is the distance between the two magnetic dipoles, $\vec{n}_{1,2}$ is the directional vector connecting the two magnetic dipoles (of unit length). μ^2_{Bohr} is the square of the Bohr magneton; with an approximate value of 0.43297 in cm^{-1}/T. As inferred from the above Equation, the dipolar magnetic interaction depends on the distance and on the angle between the magnetic moments on magnetic centers. Therefore, the Cartesian coordinates of all non-equivalent magnetic centers must be provided in the input.

In brief, the POLY_ANISO is performing the following operations:

1. read the input and information from the datafiles
2. build the exchange coupled basis
3. compute the magnetic exchange, magnetic dipole-dipole, and other magnetic Hamiltonians using the *ab initio*-computed spin and orbital momenta of individual magnetic sites and the input parameters
4. sum up all the magnetic interaction Hamiltonians and diagonalise the total interaction Hamiltonian
5. rewrite the spin and magnetic moment in the exchange-coupled eigenstates basis
6. use the obtained spin and magnetic momenta for the computation of the magnetic properties of entire polynuclear compound

The actual values of the inter-site magnetic exchange could be derived from e.g. broken-symmetry DFT calculations. Alternatively, they could be regarded as fitting parameters, while their approximate values could be extracted by minimising the standard deviation between measured and calculated magnetic data.

7.18.2 Files

POLY_ANISO is called independently of ORCA for now. In the future versions of ORCA we will aim for a deeper integration, for a better experience.

```
bash:$
bash:$  $ORCA/x86_64/otool_poly_aniso  <  poly_aniso.input   >  poly_aniso.output
bash:$
```

The actual names of the poly_aniso.input and poly_aniso.output are not hard coded, and can take any names. A bash script for a more convenient usage of POLY_ANISO can be provided upon request or made available on the Forum.

Input files

The program POLY_ANISO needs the following files:

aniso_i.input
: This is an ASCII text file generated by the CASSCF/SOC/ SINGLE_ANISO run. It should be provided for POLY_ANISO as aniso_i.input (i=1,2,3, etc.): one file for each magnetic center. In cases when the entire polynuclear cluster or molecule has exact point group symmetry, only aniso_i.input files for crystallographically non-equivalent centers should be given. This saves computational time since equivalent metal sites do not need to be computed *ab initio*.

poly_aniso.input
: The standard input file defining the computed system and various input parameters. This file can take any name.

Output files

7.18.3 List of keywords

This section describes the keywords used to control the POLY_ANISO input file. Only two keywords NNEQ, PAIR (and SYMM if the polynuclear cluster has symmetry) are mandatory for a minimal execution of the program, while the other keywords allow customisation of the execution of the POLY_ANISO.

The format of the "poly_aniso.input" file resembles to a certain extent the input file for SINGLE_ANISO program. The input file must start with "&POLY_ANISO" text.

Mandatory keywords defining the calculation

Keywords defining the polynuclear cluster:

NNEQ This keyword defines several important parameters of the calculation. On the first line after the keyword the program reads 2 values: 1) the number of types of different magnetic centers (NON-EQ) of the cluster and 2) a letter T or F in the second position of the same line. The number of NON-EQ is the total number of magnetic centers of the cluster which cannot be related by point group symmetry. In the second position the answer to the question: "Have all NON-EQ centers been computed *ab initio*?" is given: T for True and F for False. On the following line the program will read NON-EQ values specifying the number of equivalent centers of each type. On the following line the program will read NON-EQ integer numbers specifying the number of low-lying spin-orbit functions from each center forming the local exchange basis.

Some examples valid for situations where all sites have been computed *ab initio* (case T, True):

```
NNEQ
  2  T
  1  2
  2  2
```

There are two kinds of magnetic centers in the cluster; both have been computed *ab initio*; the cluster consists of 3 magnetic centers: one center of the first kind and two centers of the second kind. From each center we take into the exchange coupling only the ground doublet. As a result, the $N_{exch} = 2^1 \times 2^2 = 8$, and the two datafiles aniso_1.input (for-type 1) and aniso_2.input (for-type 2) files must be present.

```
NNEQ
  3  T
  2  1  1
  4  2  3
```

There are three kinds of magnetic centers in the cluster; all three have been computed *ab initio*; the cluster consists of four magnetic centers: two centers of the first kind, one center of the second kind and one center of the third kind. From each of the centers of the first kind we take into exchange coupling four spin-orbit states, two states from the second kind and three states from the third center. As a result the $N_{exch} = 4^2 \times 2^1 \times 3^1 = 96$. Three files aniso_i.input for each center ($i = 1, 2, 3$) must be present.

```
NNEQ
  6  T
  1  1  1  1  1  1
  2  4  3  5  2  2
```

There are six kinds of magnetic centers in the cluster; all six have been computed *ab initio*; the cluster consists of 6 magnetic centers: one center of each kind. From the center of the first kind we take into exchange coupling two spin-orbit states, four states from the second center, three states from the third center, five states from the fourth center and two states from the fifth and sixth centers. As a result the $N_{exch} = 2^1 \times 4^1 \times 3^1 \times 5^1 \times 2^1 \times 2^1 = 480$. Six files aniso_i.input for each center ($i = 1, 2, ..., 6$) must be present.

Only in cases when some centers have NOT been computed *ab initio* (i.e. for which no aniso_i.input file exists), the program will read an additional line consisting of NON-EQ letters (A or B) specifying the type of each of the NON-EQ centers: A - the center is computed *ab initio* and B - the center is considered isotropic. On the following

number-of-B-centers line(s) the isotropic g factors of the center(s) defined as B are read. The spin of the B center(s) is defined: $S = (N-1)/2$, where N is the corresponding number of states to be taken into the exchange coupling for this particular center. Some examples valid for mixed situations: the system consists of centers computed *ab initio* and isotropic centers (case F, False):

```
NNEQ
  2  F
  1  2
  2  2
  A  B
  2.3   2.3   2.3
```

There are two kinds of magnetic centers in the cluster; the center of the first type has been computed *ab initio*, while the centers of the second type are considered isotropic with $g_X = g_Y = g_Z$=2.3; the cluster consists of three magnetic centers: one center of the first kind and two centers of the second kind. Only the ground doublet state from each center is considered for the exchange coupling. As a result the $N_{exch} = 2^1 \times 2^2 = 8$. File `aniso_i.input` (for-type 1) must be present.

```
NNEQ
  3  F
  2  1  1
  4  2  3
  A  B  B
  2.3  2.3  2.0
  2.0  2.0  2.5
```

There are three kinds of magnetic centers in the cluster; the first center type has been computed *ab initio*, while the centers of the second and third types are considered empirically with $g_X = g_Y$ =2.3; g_Z=2.0 (second type) and $g_X = g_Y$ =2.0; g_Z=2.5 (third type); the cluster consists of four magnetic centers: two centers of the first kind, one center of the second kind and one center of the third kind. From each of the centers of the first kind, four spin-orbit states are considered for the exchange coupling, two states from the second kind and three states from the center of the third kind. As a result the $N_{exch} = 4^2 \times 2^1 \times 3^1 = 96$. The file `aniso_i.input` must be present.

```
NNEQ
  6  T
  1  1  1  1  1  1
  2  4  3  5  2  2
  B  B  A  A  B  A
  2.12  2.12  2.12
  2.43  2.43  2.43
  2.00  2.00  2.00
```

There are six kinds of magnetic centers in the cluster; only three centers have been computed *ab initio*, while the other three centers are considered isotropic; the g factors of the first center is 2.12 (S=1/2); of the second center 2.43 (S=3/2); of the fifth center 2.00 (S=1/2); the entire cluster consists of six magnetic centers: one center of each kind. From the center of the first kind, two spin-orbit states are considered in the exchange coupling, four states from the second center, three states from the third center, five states from the fourth center and two states from the fifth and sixth centers. As a result the $N_{exch} = 2^1 \times 4^1 \times 3^1 \times 5^1 \times 2^1 \times 2^1 = 480$. Three files `aniso_3.input` and `aniso_4.input` and `aniso_6.input` must be present.

There is no maximal value for `NNEQ`, although the calculation becomes quite heavy in case the number of exchange functions is large.

`SYMM` Specifies rotation matrices to symmetry equivalent sites. This keyword is mandatory in the case more centers of a given type are present in the calculation. This keyword is mandatory when the calculated polynuclear compound has exact crystallographic point group symmetry. In other words, when the number of equivalent centers of any kind i is larger than 1, this keyword must be employed. Here the rotation matrices from the one center to all the other of the same type are declared. On the following line the program will read the number 1 followed on the next lines by as many 3×3 rotation matrices as the total number of equivalent centers of type 1. Then the rotation matrices of centers of type 2, 3 and so on, follow in the same format. When the rotation matrices contain irrational numbers (e.g. $\sin \frac{\pi}{6} = \frac{\sqrt{3}}{2}$), then more digits than presented in the examples below are advised to be

given: $\frac{\sqrt{3}}{2} = 0.8660254$. Examples:

```
NNEQ
  2  F
  1  2
  2  2
  A  B
  2.3  2.3  2.3
  SYMM
  1
  1.0 0.0 0.0
  0.0 1.0 0.0
  0.0 0.0 1.0
  2
  1.0 0.0 0.0
  0.0 1.0 0.0
  0.0 0.0 1.0
 -1.0 0.0 0.0
  0.0 -1.0 0.0
  0.0 0.0 -1.0
```

The cluster computed here is a tri-nuclear compound, with one center computed *ab initio*, while the other two centers, related to each other by inversion, are considered isotropic with $g_X = g_Y = g_Z = 2.3$. The rotation matrix for the first center is I (identity, unity) since the center is unique. For the centers of type 2, there are two matrices 3×3 since we have two centers in the cluster. The rotation matrix of the first center of type 2 is Identity while the rotation matrix for the equivalent center of type 2 is the inversion matrix.

```
NNEQ
  3  F
  2  1  1
  4  2  3
  A  B  B
  2.1 2.1 2.1
  2.0 2.0 2.0
  SYMM
  1
  1.0 0.0 0.0
  0.0 1.0 0.0
  0.0 0.0 1.0
  0.0 -1.0 0.0
  -1.0 0.0 0.0
  0.0 0.0 1.0
  2
  1.0 0.0 0.0
  0.0 1.0 0.0
  0.0 0.0 1.0
  3
  1.0 0.0 0.0
  0.0 1.0 0.0
  0.0 0.0 1.0
```

In this input a tetranuclear compound is defined, all centers are computed *ab initio*. There are two centers of type "1", related one to each other by C_2 symmetry around the Cartesian Z axis. Therefore the `SYMM` keyword is mandatory. There are two matrices for centers of type 1, and one matrix (identity) for the centers of type 2 and type 3.

```
NNEQ
  6  F
  1  1  1  1  1  1
  2  4  3  5  2  2
  B  B  A  A  B  A
  2.12 2.12 2.12
```

(continues on next page)

(continued from previous page)

```
 2.43 2.43 2.43
 2.00 2.00 2.00
```

In this case the computed system has no symmetry. Therefore, the SYMM keyword is not required. End of Input Specifies the end of the input file. No keywords after this one will be processed.

Keywords defining the magnetic exchange interactions

This section defines the keywords used to set up the interacting pairs of magnetic centers and the corresponding exchange interactions.

A few words about the numbering of the magnetic centers of the cluster in the POLY_ANISO. First all equivalent centers of the type 1 are numbered, then all equivalent centers of the type 2, etc. These labels of the magnetic centers are used further for the declaration of the magnetic coupling.

PAIR or LIN1

This keyword defines the interacting pairs of magnetic centers and the corresponding exchange interaction. A few words about the numbering of the magnetic centers of the cluster in the POLY_ANISO. First all equivalent centers of the type 1 are numbered, then all equivalent centers of the type 2, etc. These labels of the magnetic centers are used now for the declaration of the magnetic coupling. Interaction Hamiltonian is:

$$\hat{H}_{Lines} = -\sum_{p=1}^{N_{pairs}} J_p \hat{s}_i \hat{s}_j,$$

where i an j are the indices of the metal sites of the interacting pair p; J_p is the user-defined magnetic exchange interaction between the corresponding metal sites; $\hat{s}_i$ and $\hat{s}_j$ are the ab initio spin operators for the low-lying exchange eigenstates.

```
PAIR
3
1 2 -0.2
1 3 -0.2
2 3  0.4
```

The input above is applicable for a tri-nuclear molecule. Two interactions are antiferromagnetic while ferromagnetic interaction is given for the last interacting pair.

LIN3

This keyword defines a more involved exchange interaction, where the user is allowed to define 3 parameters for each interacting pair. The interaction Hamiltonian is given by:

$$\hat{H}_{Lines} = -\sum_{p=1}^{N_{pairs}} \sum_{\alpha} J_{p,\alpha} \hat{s}_{i,\alpha} \hat{s}_{j,\alpha},$$

where the α defines the Cartesian axis x, y, z.

```
LIN3
1
1 2 -0.2 -0.4 -0.6 # i, j, Jx, Jy, Jz
```

The input above is applicable for a mononuclear molecule.

LIN9

This keyword defines a more involved exchange interaction, where the user is allowed to define 9 parameters for each interacting pair. The interaction Hamiltonian is given by:

$$\hat{H}_{Lines} = -\sum_{p=1}^{N_{pairs}} \sum_{\alpha,\beta} J_{p,\alpha,\beta} \hat{s}_{i,\alpha} \hat{s}_{j,\beta},$$

where the α and β defines the Cartesian axis x, y, z.

```
LIN9
1
1 2 -0.1 -0.2 -0.3 -0.4 -0.5 -0.6 -0.7 -0.8 -0.9
# i,j,Jxx,Jxy,Jxz,Jyx,Jyy,Jyz,Jzx,Jzy,Jzz
```

The input above is applicable for a mononuclear molecule.

COOR

The COOR keyword turns ON the computation of the dipolar coupling for those interacting pairs which were declared under PAIR, LIN3 or LIN9 keywords. On the NON-EQ lines following the keyword the program will read the symmetrised Cartesian coordinates of NON-EQ magnetic centers: one set of symmetrised Cartesian coordinates for each type of magnetic centers of the system. The *symmetrized Cartesian coordinates* are obtained by translating the original coordinates to the origin of Coordinate system, such that by applying the corresponding SYMM rotation matrix onto the input COOR data, the position of all other sites are generated.

```
COOR
6.489149    3.745763    1.669546
5.372478    5.225861    0.505625
```

The magnetic dipole-dipole Hamiltonian is computed as follows:

$$\hat{H}_{dip} = \mu_{Bohr}^2 \sum_{p=1}^{N_{pairs}} \frac{\hat{\mu}_i \hat{\mu}_j - 3(\hat{\mu}_i \vec{n}_{i,j})(\hat{\mu}_j \vec{n}_{i,j})}{\vec{r_{i,j}^3}}$$

and is added to $\hat{H}_{exch}$ computed using other models. The $\hat{H}_{dip}$ is added for all magnetic pairs.

Optional general keywords to control the input

Normally POLY_ANISO runs without specifying any of the following keywords. However, some properties are only computed if it is requested by the respective keyword. Argument(s) to the keyword are always supplied on the next line of the input file.

MLTP

The number of molecular multiplets (i.e. groups of spin-orbital eigenstates) for which g, D and higher magnetic tensors will be calculated (default MLTP=1). The program reads two lines: the first is the number of multiplets (N_{MULT}) and the second the array of N_{MULT} numbers specifying the dimension (multiplicity) of each multiplet. Example:

```
MLTP
10
2 4 4 2 2 2 2 2 2 2
```

POLY_ANISO will compute the EPR g and D- tensors for 10 groups of states. The groups 1 and 4-10 are doublets ($\tilde{S} = |1/2\rangle$), while the groups 2 and 3 are quadruplets, having the effective spin $\tilde{S} = |3/2\rangle$. For the latter cases, the ZFS (D-) tensors will be computed. We note here that large degeneracies are quite common for exchange coupled systems, and the data for this keyword can only be rendered after the inspection of the exchange spectra.

TINT

Specifies the temperature points for the evaluation of the magnetic susceptibility. The program will read three numbers: T_{min}, T_{max}, and nT.

- T_{min} - the minimal temperature (Default 0.0 K)
- T_{max} - the maximal temperature (Default 300.0 K)
- nT - number of temperature points (Default 301)

Example:

```
TINT
0.0 300.0 331
```

POLY_ANISO will compute temperature dependence of the magnetic susceptibility in 331 points evenly distributed in temperature interval: 0.0 K - 330.0 K.

HINT

Specifies the field points for the evaluation of the molar magnetisation. The program will read three numbers: H_{min}, H_{max}, nH.

- H_{min} - the minimal field (Default 0.0 T)
- H_{max} - the maximal filed (Default 10.0 T)
- nH - number of field points (Default 101)

Example:

```
HINT
0.0 20.0 201
```

POLY_ANISO will compute the molar magnetisation in 201 points evenly distributed in field interval: 0.0 T - 20.0 T.

TMAG

Specifies the temperature(s) at which the field-dependent magnetisation is calculated. Default is one temperature point, T = 2.0 K.

Example:

```
TMAG
6   1.8 2.0 2.4 2.8 3.2 4.5
```

ENCU

The keyword expects to read two integer numbers. The two parameters (NK and MG) are used to define the cut-off energy for the lowest states for which Zeeman interaction is taken into account exactly. The contribution to the magnetisation coming from states that are higher in energy than E (see below) is done by second order perturbation theory. The program will read two integer numbers: NK and MG. Default values are: $NK = 100, MG = 100$.

$$E = NK \cdot k_{Boltz} \cdot \texttt{TMAG}_{max} + MG \cdot \mu_{Bohr} \cdot H_{max}$$

The field-dependent magnetisation is calculated at the maximal temperature value given by TMAG keyword. Example:

```
ENCU
250 150
```

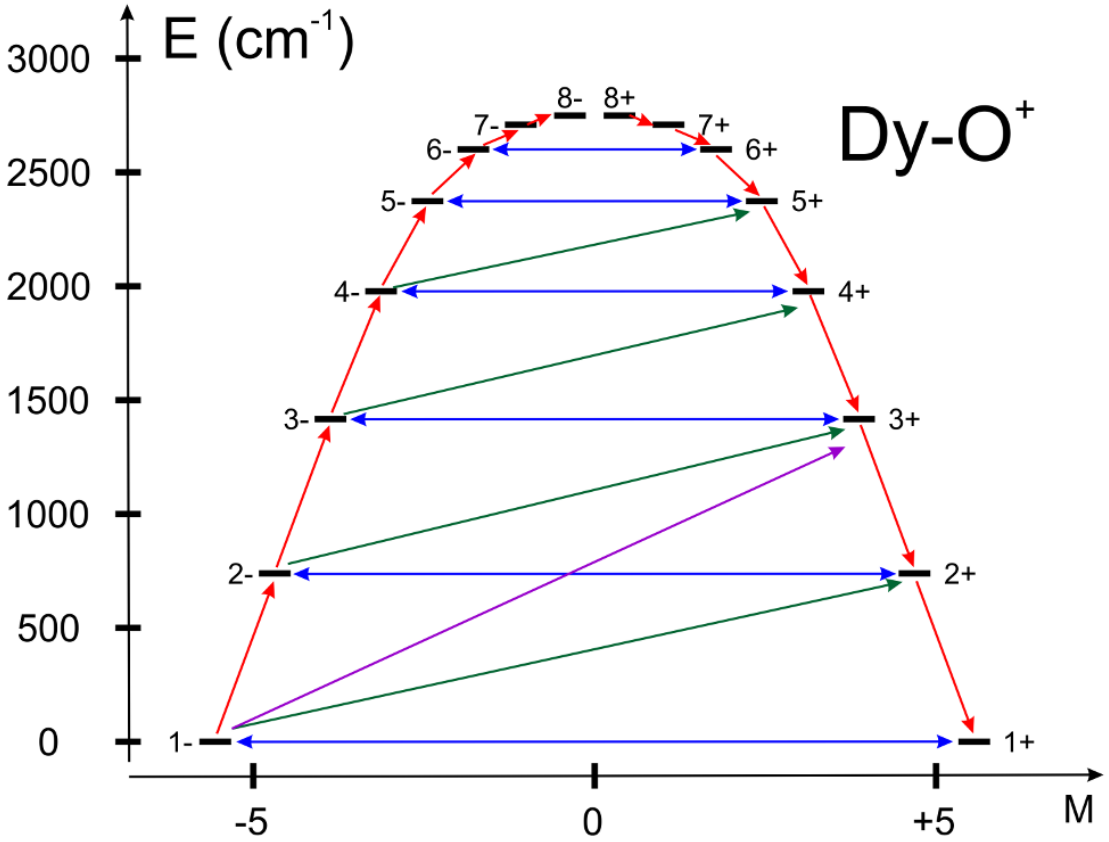
If H_{max} = 10 T and TMAG = 1.8 K, then the cut-off energy is:

$$E = 250 \cdot k_{Boltz} \cdot 1.8 + 150 \cdot \mu_{Bohr} \cdot 10 = 1013.06258(cm^{-1})$$

This means that the magnetisation arising from all exchange states with energy lower than $E = 1013.06258(cm^{-1})$ will be computed exactly (i.e. are included in the exact Zeeman diagonalisation) The keywords NCUT, ERAT and ENCU have similar purpose. If two of them are used at the same time, the following priority is defined: NCUT > ENCU > ERAT.

UBAR

With UBAR set to "true", the blocking barrier of a single-molecule magnet is estimated. The default is not to compute it. The method prints transition matrix elements of the magnetic moment according to the Figure below:



In this figure, a qualitative performance picture of the investigated single-molecular magnet is estimated by the strengths of the transition matrix elements of the magnetic moment connecting states with opposite magnetisations ($n+ \rightarrow n-$). The height of the barrier is qualitatively estimated by the energy at which the matrix element ($n+ \rightarrow n-$) is large enough to induce significant tunnelling splitting at usual magnetic fields (internal) present in the magnetic crystals (0.01-0.1 Tesla). For the above example, the blocking barrier closes at the state ($8+ \rightarrow 8-$). All transition matrix elements of the magnetic moment are given as $((|\mu_X| + |\mu_Y| + |\mu_Z|)/3)$. The data is given in Bohr magnetons (μ_{Bohr}). Example:

```
UBAR
```

ERAT

This flag is used to define the cut-off energy for the low-lying exchange-coupled states for which Zeeman interaction is taken into account exactly. The program will read one single real number specifying the ratio of the energy states which are included in the exact Zeeman Hamiltonian. As example, a value of 0.5 means that the lowest half of the energy states included in the spin-orbit calculation are used for exact Zeeman diagonalisation. Example:

```
ERAT
0.333
```

The keywords NCUT, ERAT and ENCU have similar purpose. If two of them are used at the same time, the following priority is defined: NCUT > ENCU > ERAT.

NCUT

This flag is used to define the cut-off energy for the low-lying exchange states for which Zeeman interaction is taken into account exactly. The contribution to the magnetisation arising from states that are higher in energy than lowest N_{CUT} states, is done by second-order perturbation theory. The program will read one integer number. In case the number is larger than the total number of exchange states(N_{exch}, then the N_{CUT} is set to N_{SS} (which means that the molar magnetisation will be computed exactly, using full Zeeman diagonalisation for all field points). The field-dependent magnetisation is calculated at the temperature value(s) defined by TMAG. Example:

```
NCUT
32
```

The keywords NCUT, ERAT and ENCU have similar purpose. If two of them are used at the same time, the following priority is defined: NCUT > ENCU > ERAT.

MVEC

MVEC, define a number of directions for which the magnetisation vector will be computed. The directions are given as vectors specifying the direction *i* of the applied magnetic field).

Example:

```
MVEC
4 # number of directions
1.0 0.0 0.0  # px, py, pz of each direction
0.0 1.0 0.0
0.0 0.0 1.0
1.0 1.0 1.0
```

ZEEM

This keyword allows to compute Zeeman splitting spectra along certain directions of applied field. Directions of applied field are given as three real number for each direction, specifying the projections along each direction: Example:

```
ZEEM
6
1.0 0.0 0.0
0.0 1.0 0.0
0.0 0.0 1.0
0.0 1.0 1.0
1.0 0.0 1.0
1.0 1.0 0.0
```

The above input will request computation of the Zeeman spectra along six directions: Cartesian axes X, Y, Z (directions 1,2 and 3), and between any two Cartesian axes: YZ, XZ and XY, respectively. The program will re-normalise the input vectors according to unity length. In combination with PLOT keyword, the corresponding zeeman_energy_xxx.png images will be produced.

MAVE

The keyword requires two integer numbers, denoted MAVE_nsym and MAVE_ngrid. The parameters MAVE_nsym and MAVE_ngrid specify the grid density in the computation of powder molar magnetisation. The program uses Lebedev-Laikov distribution of points on the unit sphere. The parameters are integer numbers: n_{sym} and n_{grid}. The n_{sym} defines which part of the sphere is used for averaging. It takes one of the three values: 1 (half-sphere), 2 (a quarter of a sphere) or 3 (an octant of the sphere). n_{grid} takes values from 1 (the smallest grid) till 32 (the largest grid, i.e. the densest). The default is to consider integration over a half-sphere (since $M(H) = -M(-H)$): $n_{sym} = 1$ and $n_{sym} = 15$ (i.e 185 points distributed over half-sphere). In case of symmetric compounds, powder magnetisation may be averaged over a smaller part of the sphere, reducing thus the number of points for the integration. The user is responsible to choose the appropriate integration scheme. Note that the program's default is rather conservative.

Example:

```
MAVE
1  8
```

TEXP

This keyword allows computation of the magnetic susceptibility $\chi T(T)$ at experimental points. On the line below the keyword, the number of experimental points N_T is defined, and on the next N_T lines the program reads the experimental temperature (in K) and the experimental magnetic susceptibility (in $cm^3 K mol^{-1}$). The magnetic susceptibility routine will also print the standard deviation from the experiment.

```
TEXP
54
299.9901   55.27433
290.4001   55.45209
279.7746   55.43682
269.6922   55.41198
259.7195   55.39274
249.7031   55.34379
239.735    55.29292
229.7646   55.23266
219.7354   55.15352
209.7544   55.06556
...
```

HEXP

This keyword allows computation of the molar magnetisation $M_{mol}(H)$ at experimental points. On the line below the keyword, the number of experimental points N_H is defined, and on the next N_H lines the program reads the experimental field intensity (in Tesla) and the experimental magnetisation (in μ_{Bohr}). The magnetisation routine will print the standard deviation from the experiment.

```
HEXP
3  1.0  5.3  2.4  # temperature values
10 # numer of field points
0.30 4.17 1.26 2.51  # H(T) and M for each temperature
1.00 5.47 3.57 4.82
1.88 5.79 4.54 5.30
2.67 5.92 4.96 5.54
3.46 5.97 5.20 5.70
4.24 6.00 5.36 5.81
5.03 6.01 5.48 5.88
5.82 6.02 5.57 5.93
6.61 6.02 5.65 5.97
7.40 6.03 5.72 5.99
```

ZJPR

This keyword specifies the value (in cm^{-1}) of a phenomenological parameter of a mean molecular field acting on the spin of the complex (the average intermolecular exchange constant). It is used in the calculation of all magnetic properties (not for spin Hamiltonians) (Default is 0.0).

```
ZJPR
-0.02
```

XFIE

This keyword specifies the value (in T) of applied magnetic field for the computation of magnetic susceptibility by dM/dH and M/H formulas. A comparison with the usual formula (in the limit of zero applied field) is provided. (Default is 0.0). Example:

```
XFIE
0.35
```

This keyword together with the keyword `PLOT` will enable the generation of two additional plots: `XT_with_field_dM_over_dH.png` and `XT_with_field_M_over_H.png`, one for each of the two above formula used, alongside with respective `gnuplot` scripts and gnuplot datafiles.

TORQ

This keyword specifies the number of angular points for the computation of the magnetisation torque function, $\vec{\tau}_\alpha$ as function of the temperature, field strength and field orientation.

```
TORQ
55
```

The torque is computed at all temperature given by `TMAG` or `HEXP_temp` inputs. Three rotations around Cartesian axes X, Y and Z are performed.

PRLV

This keyword controls the print level.

- 2 - normal. (Default)
- 3 or larger (debug)

PLOT

Set to "true", the program generates a few plots (png or eps format) via an interface to the linux program *gnuplot*. The interface generates a datafile, a *gnuplot* script and attempts execution of the script for generation of the image. The plots are generated only if the respective function is invoked. The magnetic susceptibility, molar magnetisation and blocking barrier (`UBAR`) plots are generated. The files are named: `XT_no_field.dat`, `XT_no_field.plt`, `XT_no_field.png`, `MH.dat`, `MH.plt`, `MH.png`, `BARRIER_TME.dat`, `BARRIER_ENE.dat`, `BARRIER.plt` and `BARRIER.png`, `zeeman_energy_xxx.png` etc. All files produced by `SINGLE_ANISO` are referenced in the corresponding output section. Example:

```
PLOT
```

7.19 N-Electron Valence State Pertubation Theory

CASPT2 and NEVPT2 belongs to the family of internally contracted perturbation theories with CASCI reference wavefunctions. Several studies indicate that CASPT2 and NEVPT2 produce energies of similar quality.[366, 756] The NEVPT2 methodology developed by Angeli et al exists in two formulations namely the strongly-contracted NEVPT2 (SC-NEVPT2) and the partially contracted NEVPT2 (PC-NEVPT2). [44, 45, 46] Irrespective of the name "partially contracted" coined by Angeli et al, the latter approach employs a fully internally contracted wavefunction (FIC). Hence, we use the term "FIC-NEVPT2" in place of PC-NEVPT2. ORCA features the fully internally contracted and the strongly contracted NEVPT2. The latter employs strongly contracted CSFs, which form a more compact and orthogonal basis making it computationally slightly more attractive. Hence, the SC-NEVPT2 has been our work horse a for long time. NEVPT2 has many desirable properties - among them:

- It is **intruder state free** due to the choice of the Dyall Hamiltonian [238] as the 0th order Hamiltonian.
- The **0th order Hamiltonian is diagonal** in the perturber space. Therefore no linear equation system needs to be solved.
- It is **strictly size consistent**. The total energy of two non-interacting systems is equal to the sum of two isolated systems.
- It is **invariant under unitary transformations** within the active subspaces.
- "**strongly contracted**": Perturber functions only interact via their active part. Different subspaces are orthogonal and hence no time is wasted on orthogonalization issues.
- "**fully internally contracted**": Invariant to rotations of the inactive and virtual subspaces.

As described in Section *N-Electron Valence State Perturbation Theory (NEVPT2)* of the manual, NEVPT2 requires a single keyword on top of a working CASSCF input. The methods are called within the CASSCF block and detailed settings can be adjusted in the `PTSettings` subblock.

We will go through some of the detailed setting in the next few subsections. For historical reasons, a few features, such as the quasi-degenerate NEVPT2, are only available for the strongly contracted NEVPT2. As shown elsewhere, the strong contraction is not a good starting point for linear scaling approaches.[805] Thus newer additions such as the DLPNO and the F12 correction rely on the FIC variant. [338, 344, 345] Note that ORCA by default employs the frozencore approximation, which can be disabled with the simple keyword `!NoFrozenCore`. A complete description of the frozecore settings can be found in section *Frozen Core Options*.

```
%casscf
  ...
  MULT   1,3 # multiplicity block
```

(continues on next page)

(continued from previous page)

```
  NRoots 2,2 # number of roots for the MULT blocks

  CIStep    DMRGCI  # optional to run DMRG-NEVPT2.
                    # default: CSFCI (recommended)
  trafostep ri      # RI approximation for CASSCF and NEVPT2

  # calling the PT2 method of choice
  PTMethod    SC_NEVPT2    # strongly contracted NEVPT2
              FIC_NEVPT2   # fully internally contracted / partially contracted NEVPT2
              DLPNO_NEVPT2 # FIC-NEVPT2 using the DLPNO framework for large molecules

  # detailed settings (optional) for the PT2 approaches
  PTSettings
    NThresh 1e-6     # FIC-NEVPT2 cut off for linear dependencies
    D4Step  Fly      # 4-pdm is constructed on the fly
    D4Tpre  1e-10    # truncation of the 4-pdm
    D3Tpre  1e-14    # trunaction of the 3-pdm
    EWIN  -3,1000    # Energy window for the frozencore setting fc_ewin
    TSMallDenom 1e-2 # printing thresh for small denominators

    # option to skip the PT2 correction for a selected multiplicity block and root
    # (same input structure as weights in %casscf)
    selectedRoots[0]=0,1 # skip the first roots of MULT=1
    selectedRoots[1]=0,0 # skip MULT=3 roots

    # SC-NEVPT2 specific features
    CanonStep 1   # default (exact):canonical orbitals for each state
    QDType    QD_VANVLECK   # QD-SC-NEVPT2: Van Vleck (recommended)

    # FIC-NEVPT2 specific features
    F12       true      # F12-Correction
    Density   unrelaxed # unrelaxed density generated for each state.
    NatOrbs   true      # Computes the natural orbitals

    # DLPNO specific settings
    TCutPNO 1e-8   # controls the accuracy (default 1e-8)

  end

end
```

NEVPT2 can also be set using the simple keywords on top of any valid CASSCF input.

```
!SC-NEVPT2       # for the strongly contracted NEVPT2
!FIC-NEVPT2   # for the fully internally contracted NEVPT2
!DLPNO-NEVPT2 # for the DLPNO variant of the FIC-NEVPT2
! ...
%casscf ...
```

The two computationally most demanding steps of the NEVPT2 calculation are the initial integral transformation involving the two-external labels and the formation of the fourth order density matrix (D4). Efficient approximations to both issues are available in ORCA.

If not otherwise specified (keyword `CIStep`), CASSCF and consequently NEVPT2 use a conventional CSF based solver for the CAS-CI problem. In principle, the NEVPT2 approach can be combined with approximate CI solution such as the DMRG approach described in section *Density Matrix Renormalization Group*. Starting with ORCA 4.0 it is possible to run NEVPT2-DMRG calculations for the FIC and SC type ansatz using the methodology developed by the Chan group.[336] Aside from the usual DMRG input, the program requires an additional parameter (`nevpt2_MaxM`) in the DMRG block. However, some of the features will be restricted to the default `CIStep`.

```
%casscf
  cistep DMRGCI
  %dmrg
    ...
    nevpt2_MaxM 2000 # (see Note below)
  end
  PTMethod SC_NEVPT2 # or FIC_NEVPT2
end
```

For the value `nevpt2_MaxM 2000` cf [336].

Using the RI approximation, large molecules with actives spaces of up to 20 orbitals should be computable. The DMRG extension can be combined with DLPNO and F12 variants. Future version might also support the `CIStep ACCCI` and `CIStep ICE`.

7.19.1 RI, RIJK and RIJCOSX Approximation

Setting the RI approximation on the CASSCF level, will set the RI options for NEVPT2 respectively. The three index integrals are computed and partially stored on disk. Three index integral with two internal labels are kept in main memory. The two-electron integrals are assembled on the fly. The auxiliary basis must be large enough to fit the integrals appearing in the CASSCF orbital gradient/Hessian and the NEVPT2 part. The auxiliary basis set of the type /J does not suffice here.

```
%casscf
  ...
  TrafoStep RI   # enable RI approximation in CASSCF and NEVPT2
  PTMethod  SC_NEVPT2 # or the NEVPT2 approach or your choice
end
```

Additional speedups can be obtained if the Fock operator formation is approximated using the `!RIJCOSX` or `!RIJK` techniques. In case of `RIJCOSX`, an additional auxiliary basis must be provided for the `AuxJ` auxiliary basis slot. For more information on the basis set slots see section *Built-in Basis Sets*.

```
#RIJCOSX one-liner
! def2-svp def2/J RIJCOSX def2-svp/C

# Commented out: Alternative definition via %basis block
#%basis
#auxJ "def2/J"     # or for example "AutoAux"
#auxC "def2-svp/C" # or for example "AutoAux"
#end
```

Whereas the `RIJK` requires a single auxiliary basis set (`AuxJK` slot), that is large enough to fit integrals in the Fock-matrix construction, orbital gradient/Hessian and the correlation part. In contrast to COSX, the calculation can also be carried out in `conv` mode (storing the AO integrals on disk).

```
#RIJK one-liner, conv is mandatory for RIJK in CASSCF
! def2-svp def2/JK RIJK

# Commented out: Alternative definition via %basis block
#%basis
#auxJK "def2/JK" # or "AutoAux"
#end
```

The described methodology allows the computation of systems with up to 2000 basis functions. Even larger molecules are accessible in the framework of DLPNO-NEVPT2 described in the next subsection. Several examples can be found in the CASSCF tutorial.

7.19.2 Beyond the RI approximation: DLPNO-NEVPT2

For systems with more than 80 atoms, we recommend the recently developed DLPNO-NEVPT2.[344] It is a successful combination of DLPNO strategy with the FIC-NEVPT2 method. As its single reference counterparts, DLPNO-NEVPT2 recovers 99.9% of the FIC-NEVPT2 correlation energies even for large system. The input structure is similar to the parenting FIC-NEVPT2 method. Below you find an input example for the Fe(II)-complex depicted in Fig. 7.6, where the active space consists of the metal-3d orbitals. The example takes about 9 hour (including 3 hour for one CASSCF iteration) using 8 cores (2.60GHz Intel E5-2670 CPU) for the calculation to finish. A detailed description of the DLPNO-NEVPT2 methodology can be found in our article.[344].

```
# DLPNO-NEVPT2 calculation for quintet state of FeC72N2H100
!PAL8 def2-TZVP def2/JK
!moread noiter
%moinp "FeC72N2H100.gbw-CASSCF"
%MaxCore 8000
%casscf
  nel 6
  norb 5
  mult 5
  TrafoStep RI # RI approximation is mandatory for DLPNO-NEVPT2

  PTMethod DLPNO_NEVPT2
  # detailed settings (optional)
  PTSettings
    TCutPNO 1e-8   # most important parameter controlling the accuracy (default 1e-8)
    MaxIter   20   # maximum for residual iterations
    MaxDIIS   7    # DIIS dimension
  end
end
*xyz 0 5 FeC72N2H100.xyz
```

Just like RI-NEVPT2, the calculations requires an auxiliary basis. The aux-basis should be of /C or /JK type (more accurate). Aside from the paper of Guo et al,[344] a concise report of the accuracy can be found in the CASSCF tutorial, where we compute exchange coupling parameters. Note that in the snippet above, we have repeated some of the default setting in the NEVPT sub-block. This is not mandatory and should be avoided to keep the input as simple as possible.

As mentioned earlier, the CASSCF step can be accelerated with the RIJK or RIJCOSX approximation. Both options are equally valid for the DLPNO-NEVPT2. The RIJK variant typically produces more accurate results than RIJCOSX. The input file is almost the same as before, except for the keyword line:

```
# The combination of RIJK with DLPNO-NEVPT2
!PAL8 def2-TZVP def2/JK conv RIJK
```

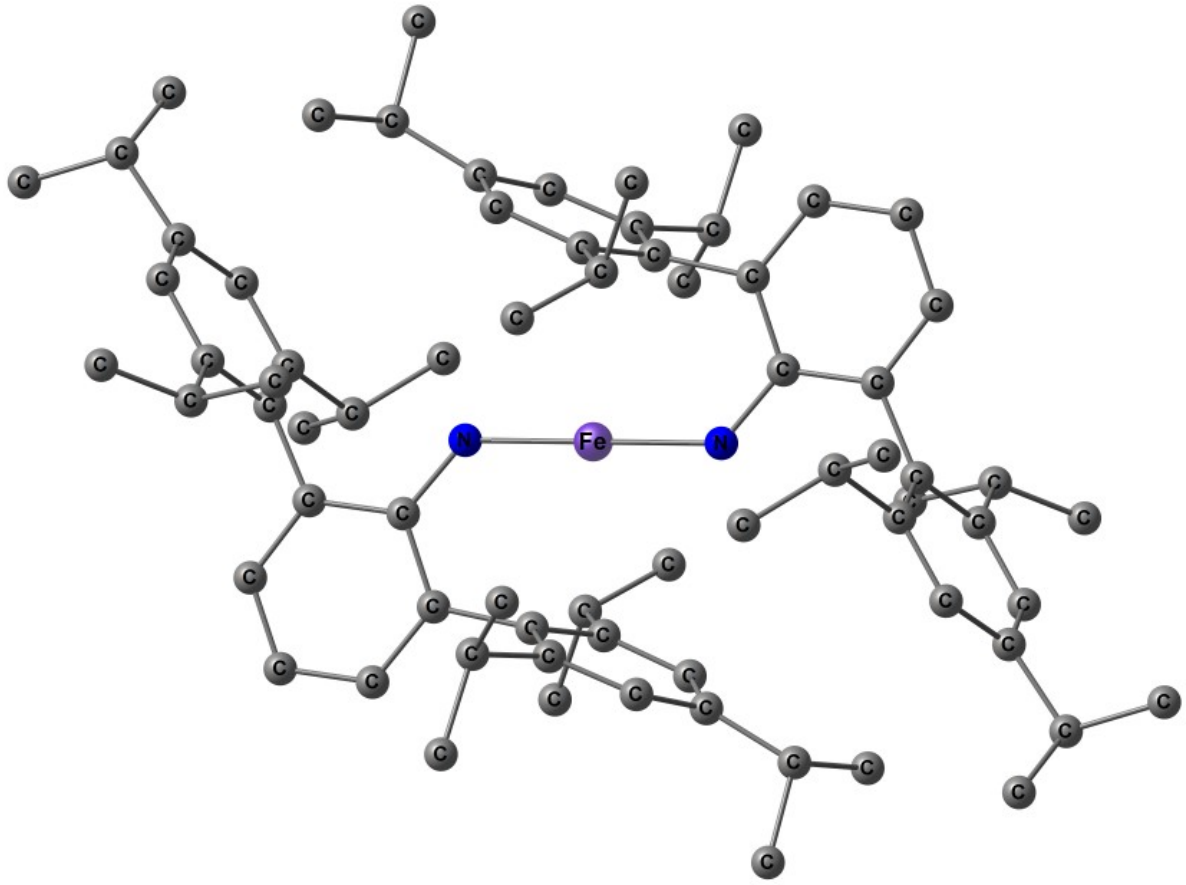


Fig. 7.6: Structure of the $FeC_{72}N_2H_{100}$

7.19.3 Explicitly correlated NEVPT2: NEVPT2-F12 and DLPNO-NEVPT2-F12

Like in the single-reference MP2 theory, the NEVPT2 correlation energy converges slowly with the basis set. Aside from basis set extrapolation, the R12/F12 method are popular methods to reach the basis set limit. For comparison of F12 and extrapolation techniques, we refer to the study of Liakos et al.[521] ORCA features an F12 correction for the FIC-NEVPT2 wavefunction using the RI approximation.[345] The RI approximation is mandatory as the involved integrals are expensive. In complete analogy to the single reference MP2-F12, the input requires an F12 basis, an F12-cabs basis and a sufficiently large RI basis (/JK or /C).

```
# aug-cc-pvdz/C used as RI basis
 ! cc-pvdz-F12 aug-cc-pvdz/C cc-pvdz-f12-cabs
 %casscf nel 8
         norb 6
         mult 3,1

         TrafoStep RI  #RI approximation must be on for F12
         PTMethod FIC_NEVPT2      # FIC-NEVPT2 or DLPNO_NEVPT2
         # detailed settings
         PTSettings
          F12 true     # Request the F12 correction
         end
 end
 *xyz 0 3
  O      0.0 0.0 0.0
  O      0.0 0.0 1.207
 *
```

A linear scaling version of NEVPT2-F12, the DLPNO-NEVPT2-F12, allows to tackle systems with several thou-

sand of basis functions.[338] With the exception of the DLPNO_NEVPT2 keyword, the input structure is otherwise identical to NEVPT2-F12 method.

```
# aug-cc-pvdz/C used as RI basis
! cc-pvdz-F12 aug-cc-pvdz/C cc-pvdz-f12-cabs
%casscf nel 8
        norb 6
        mult 3,1

        TrafoStep RI  #RI approximation must be on for F12
        PTMethod DLPNO_NEVPT2
                 # detailed settings
        PTSettings
         F12 true     #Do the F12 correction
        end
end
*xyz 0 3
 O      0.0 0.0 0.0
 O      0.0 0.0 1.207
*
```

Note that the DLPNO-NEVPT2-F12 algorithm is unitary invariant with respect to subspace rotation of inactive and active orbitals. By tightening the DLPNO truncation thresholds, the canonical NEVPT2-F12 can be reproduced, even with localized internal and active molecular orbitals.

```
# aug-cc-pvdz/C used as RI basis
! cc-pvdz-F12 aug-cc-pvdz/C cc-pvdz-f12-cabs
%casscf nel 8
        norb 6
        mult 3,1
        gtol 1e-6
        etol 1e-14

        TrafoStep RI  #RI approximation must be on for F12
        nevpt2 3 #DLPNO-NEVPT2
        actorbs locorbs #use localized active MOs.
        intorbs locorbs #use localized internal MOs.
        # detailed settings
        PTSettings
         F12 true     #Do the F12 correction
         TCutPNO 0.0
         TCutDO  0.0
         TCutCMO 0.0
         TCutDOij 0.0
        end
end
*xyz 0 3
 O      0.0 0.0 0.0
 O      0.0 0.0 1.207
*
```

7.19.4 Tackling large active CASSCF spaces

Large active spaces (CAS(10,10) and more) require special attention as the standard implementation involves the fourth order reduced density matrix (4-RDM).[46] The storage of the latter can easily reach several gigabytes and thus cannot be kept in core memory. ORCA thus by default constructs and contracts the 4-RDM on the fly (`D4Step fly`). Note that the program can be forced to keep the 4-RDM on disk (`D4Step disk`) or in memory (`D4Step core`).

Aside from the storage, the formation of the 4-RDM itself becomes the time dominating step of the NEVPT2 calculation for large active spaces. There are two set of approximations to tackle the challenge. The prescreening (PS) or the extended prescreening (EPS) approximation and the cumulant expansion.[343]
In addition, a reformulation of the canonical NEVPT2 is available, that avoids the 4-RDM.[458] The basic idea of the latter is similar to the recent development reported by Sokolov and coworkers.[162] In ORCA the reformulated "efficient" implementation is combined with the PS approximation. Note that the reformulation is presently restricted to the canonical NEVPT2 ansatz. An extension to the DLPNO variant will be available in the future. The new code is called setting "`D4Step efficient`". Irrespective of the formulation, ORCA by default truncates the CASSCF wave function prior computation of the fourth order reduced density matrix using the PS approximation.[342, 343] Only configurations with a weight larger than a given parameter `D4TPre` are taken into account. The same reduction is available for the third order density matrix using the keyword `D3TPre`. Both of the parameters can be adjusted within the `PTSettings` sub-block of the CASSCF module.

```
%casscf
  ...
  PTMethod     FIC_NEVPT2  # or SC_NEVPT2

  # detailed settings (optional)
  PTSettings
  D4Step efficient # calling the new NEVPT2 code.
                   # "fly" for the standard code
  D4TPre 1e-10     # default truncation 4-RDM
  D4TPre 1e-14     # default truncation 3-RDM
  imaginary 0.0    # imaginary shift (only for FIC-NEVPT2)
end
```

These approximations naturally affect the "configuration RI" as well. In this context, it should be noted that a configuration corresponds to a set of configuration state functions (CSF) with identical orbital occupation. For each state the dimension of the CI and and RI space is printed.

```
D3 Build          ... CI space truncated: 141 -> 82  CFGs
        ... RI space truncated: 141 -> 141 CFGs
        D4 Build          ... CI space truncated: 141 -> 82  CFGs
        ... RI space truncated: 141 -> 141 CFGs
```

The default values usually produce errors of less than 1 mEh. However, the error introduced by the `D4TPre` is system dependent and should be double checked. The exact NEVPT2 energy is recovered with the parameters set to zero. The approximation is available for all variants of NEVPT2 (SC, FIC and DPLNO-FIC). For crude cut-offs, the approximation may lead to so called *false intruder states*.[342, 343, 912] The behavior shows up as unreasonably large correlation energy contributions of the 1h (V_i) or 1p (V_a) excitations class e.g. positive or large correlation energies compared to the 2h-2p (V_ijab) excitation class. This is a system specific issue, which is avoided with tighter thresholds (D4TPre=1e-14). The default settings is chosen conservative and rarely produces artifacts. As last resort, an imaginary shift can be added to mitigate intruder states. Note that imaginary shifts (default=0.0)are restricted the canonical FIC-NEVPT2 - not DLPNO.

The PS approximation completely neglects CFGs with a small weight. This is contrasted by the EPS approximation, where the small weights (up to thresh `D4TQuad`) are still accounted for (first order correction).[343]. The results are more robust but also more expensive compared to the PS approximation.

```
PTSettings
  D4Step  D4PT  # running the standard code withthe EPS approximation
  D4TPre  1e-10 # default truncation 4-RDM
```

(continues on next page)

(continued from previous page)

```
  D4TQuad 1e-14 # selects CFGs for the first order correction.
end
```

Huge computational savings can be achieved with the cumulant expansion, which have been recently re-evaluated.[343]. The results should be treated with care as false intruder states can emerge.[912] In these cases, the imaginary level shift is the only mitigation tool. Note that the imaginary shift is implemented only for FIC-NEVPT2.

```
PTSettings
  D4Step Cu4    # 4-RDM approximated with cumulants
                # "Cu34" to approximate 3-RDM and 4-RDM
  imaginary 0.0 # imaginary shift (only for FIC-NEVPT2)
end
```

7.19.5 Selecting or Specific States for NEVPT2

ORCA by default computes all states defined in the CASSCF block input with the NEVPT2 approach. There are cases, where this is not desired and the user wants to skip some of these states. The input mask of `SelectedRoots` is equivalent to the `weights` keyword in the %casscf block: The enumeration `SelectedRoots[0]` refers to the numbering of the multiplicity blocks and the respective roots defined in CASSCF.

```
!NEVPT2 ...
%casscf
  ...
  MULT   1,3 # multiplicity block
  NRoots 2,2 # number of roots for the MULT blocks

  # detailed settings (optional) for the PT2 approaches
  PTSettings

    # option to skip the PT2 correction for a selected multiplicity block and root
    # (same input structure as weights in %casscf)
    selectedRoots[0]=0,1 # skip the first roots of MULT=1
    selectedRoots[1]=0,0 # skip MULT=3 roots

  end
end
```

7.19.6 Unrelaxed Densities and Natural Orbitals

With the FIC-NEVPT2 ansatz, it is possible to request state-specific unrelaxed densities

$$\gamma(p,q) =< \Psi_I|E_q^p|\Psi_I >,$$

where Ψ_I refers to NEVPT2 wave function of the I'th state. The code is implemented using the ORCA AGE tool-chains.[474] In its present form the code runs serial. Note that the density can be used to generate natural orbitals.

```
%casscf
  ...
  PTMethod     FIC_NEVPT2

  # detailed settings (optional)
  PTSettings
    # densities are disabled by default
    Density Unrelaxed  # unrelaxed density <0+1|E(p,q)|0+1>
            Cu4        # cumulant 4-rdm approximated unrel. density
```

(continues on next page)

(continued from previous page)

```
            Cu34       # cumulant 3/4-rdm approximated unrel. density
            FirstOrder # approximate unrel. density <0|E(p,q)|1>

    NatOrbs True       # off by default
  end
end
```

The density as well the natural orbitals are state-specific. Thus, ORCA repeats the population analysis for each state. With the added keyword `!KeepDens` the NEVPT2 density is stored in the density container (**.densities** file on disk). The latter can be used to create density plots interactively (see Section *orca_plot*). Natural orbitals are stored in the gbw file-format as **.nat** file with a prefix corresponding to the jobname, multiplicity and root. The density can be used to generate natural orbitals.

A typical output takes the following form:

```
Unrelaxed Density     ...
Incorizing ADC        ... done in      0.6 sec
Norm <Psi|Psi>        ... done in      0.1 sec (NORM=     1.064186836)
RDM1 <Psi|E|Psi>      ... done in      0.7 sec
Reference Weight      ...     0.939684618
Trace RDM1            ...    20.000000000 (prior correction)

*** Repeating the population analysis with unrelaxed density.
Orbital energies/occupations assumed diagonal. ***
(Note: Temporarily storing unrelaxed densities as cas.scfp)
-------------------------------------------------------------------------------
ORCA POPULATION ANALYSIS
-------------------------------------------------------------------------------


...
Natural Orbital Occupation Numbers:
...
N[  4] =   1.98812992
N[  5] =   1.98308480
N[  6] =   1.93858508
N[  7] =   1.49303660
N[  8] =   1.49303660
N[  9] =   1.48519842
N[ 10] =   1.48519842
N[ 11] =   0.05922342
N[ 12] =   0.00921465
N[ 13] =   0.00921465
N[ 14] =   0.00794869
N[ 15] =   0.00620254
...
 ============================================================
 NEVPT2 Results
 ============================================================
 ...
```

NEVPT2 natural orbital can be used to do natural orbital iterations (`!MORead NoIter`). They might also be a useful tool to find suitable orbital to extend the active space.[443]

7.19.7 State-averaged NEVPT2

In the definition of the Dyall Hamiltonian [238] the CASSCF orbitals are chosen to diagonalize the Fock operator (pseudo-canonicalized). Therefore, using a state-averaged CASSCF wave function, the NEVPT2 procedure involves the construction and diagonalization of the "state-specific" Fock operators and is thus resulting in a unique set of orbitals for each state. This becomes quickly inefficient for large number of states or large molecular systems since each orbital set implies an integral-transformation. This is the **default** setting for NEVPT2 and is printed in the output

```
NEVPT2-SETTINGS:
Orbitals      ... canonical for each state
```

Other orbital options can be set using the keyword `canonstep`.

```
%casscf
  ...
  # detailed settings (optional)
  PTSettings
  canonstep  0 # state-averaged orbitals and specific orbital energies
               1 # canonical for each state
               2 # state-averaged orbitals and orbital energies
               3 # 1-step orbital relaxation
                 # and canonical for each state (partially relaxed)
  end
end
```

The final orbitals of the state-averaged CASSCF diagonalize the state-averaged Fock operator. Large computational savings can be made if these orbitals are employed for all of the states. `canonstep 0` chooses orbital energies as diagonal elements of the state-specific Fock operators. In release version ORCA 3.0 and older, this has been the default setting. These options work best if the averaged states are similar in nature. For SC-NEVPT2, we have implemented two more `canonsteps`, which trade accuracy for speed and *vice versa*. `canonstep 2` is more approximate and employs orbital energies from the state-averaged calculation. Thus there is no contribution to excitation energies from the perturber class V_{ij}^{ab} at this level of approximation.

If the states under consideration are substantially different, these approximations will be of poor quality and should be turned off. Better results can be achieved, if the state-averaged orbitals are partially relaxed for each state before the actual SC-NEVPT2 calculation. [224] Often it is not possible to optimize the excited states separately. Thus `canonstep 3` will try a single steepest descent step for each state before running the actual SC-NEVPT2 calculation with canonicalized orbitals. Optionally, instead of a steepest descent using an approximate diagonal Hessian, a single Newton-Raphson step can be made.

```
%casscf
  ...
  PTMethod SC_NEVPT2
  # detailed settings (optional)
  PTSettings
  gstep SOSCF true  # steepest descent step
  NR    false       # Newton-Raphson   step
  end
end
```

Despite a converged state-averaged calculation, the gradient for the individual states can be surprisingly large. As a consequence, the orbital relaxation might fail as both methods might be outside their convergence radius. ORCA will retry the relaxation with an increased damping. If the orbital update still fails, the program will stick with the initial orbitals. Setting an overall damping manually, might help the relaxation procedure.

```
PTMethod SC_NEVPT2
PTSettings
  gscaling 0.5 # damp gradient with a pre-factor
end
```

7.19.8 Quasi-Degenerate SC-NEVPT2

NEVPT2 as it is presented in the previous subsections follows the recipe of "diagonalize and perturb". The 0th order wavefunction is determined by the diagonalization of the CAS-CI matrix. The space spanned by the CAS-CI vectors is often referred to as "model space". The subsequent perturbation theory is constructed based on the assumption that the states under consideration are well described within the model space. Consequently, the first order correction to the wavefunction $\Psi_I^{(1)}$ does not affect the composition of the reference state $|I\rangle$. Corrections to the wavefunction and energy arise from the interaction of the reference state with the functions $|k\rangle$ of the contributing first order interacting space

$$\Psi_I^{(1)} = \sum_k C_k \left|k\right\rangle$$

$$E_I^{(2)} = \sum_k \frac{\left\langle I\right|H\left|k\right\rangle \left\langle k\right|H\left|I\right\rangle}{E_I^{(0)} - E_k}$$

This is problematic, when the interaction/mixing of states are falsely described at the CASSCF level. A typical example is the dissociation of lithium fluoride.

```
!def2-tzvp nevpt2 nofrozencore
%casscf
  nel 2
  norb 2 #Li(2s), F(2pz)
  mult 1
  nroots 2
end
%paras
r = 3,7,200
end
*xyz 0 1
Li 0 0 0
F 0 0 {r}
*
```

Here, the ground and first excited state of Σ^+ should not cross. However, at the NEVPT2 level, an erratic double crossing is observed instead.

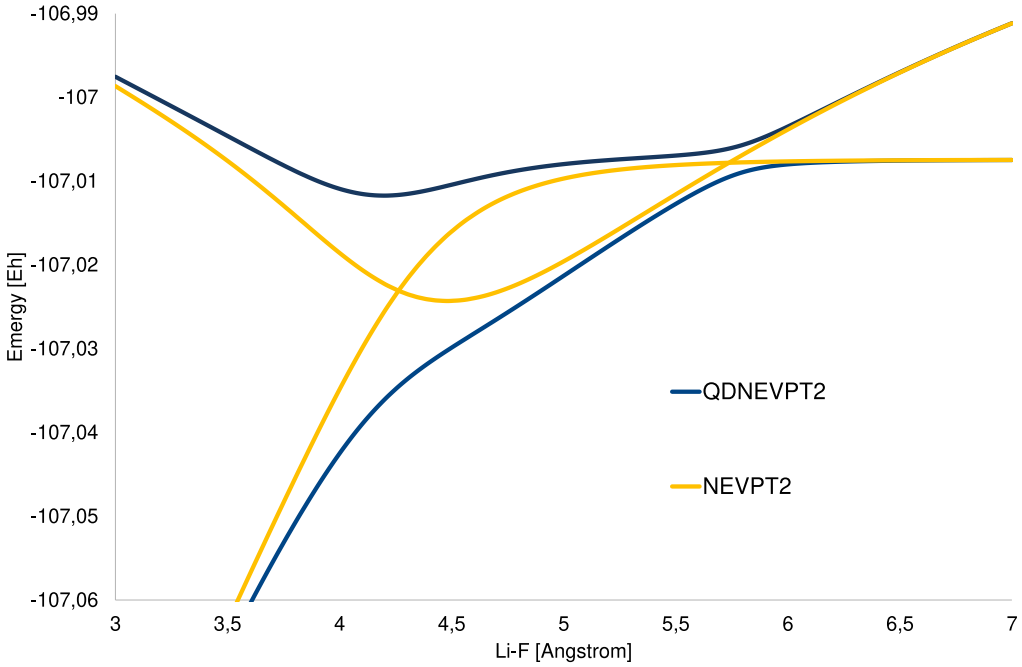


Fig. 7.7: SC-NEVPT2 and QD-SC-NEVPT2 Li-F dissociation curves of the ground and first excited states for a CAS(2,2) reference

A re-organizing of the reference states can be introduced in the framework of quasi-degenerate perturbation theory. In practice, an effective Hamiltonian is constructed allowing "off-diagonal" corrections to the second order energy

$$H_{IJ} = \delta_{IJ} E_I^{(0)} + \sum_k \frac{\langle I|H|k\rangle \langle k|H|J\rangle}{E_I^{(0)} - E_k}$$

Diagonalization of this eff. Hamiltonian yields improved energies and rotation matrix (right eigenvectors) that introduces the desired re-mixing of the reference states. The quasi-degenerate extension to SC-NEVPT2 [43] can be switched on with the keyword `QDType`.

```
%casscf
  ...
  PTMethod SC_NEVPT2     # Must be SC_NEVPT2
  PTSettings
    QDType 0             # disabled (default)
           QD_VanVleck   # Hermitian  eff. Hamiltonian (recommended)
           QD_Bloch      # non-Hermitian eff. Hamiltonian
           QD_Cloiszeaux # Hermitian eff. Hamiltonian

  end
end
```

ORCA will print the eff. Hamiltonian matrix and its eigenvectors at the end of the calculation.

```
============================================================
                      QD-NEVPT2 Results
============================================================
  ***********************
```

(continues on next page)

(continued from previous page)

```
  MULT 1,
 *********************

 Total Hamiltonian to be diagonalized
                0          1
   0    -107.074594  -0.012574
   1      -0.011748-107.003810

 Right Eigenvectors
               0          1
  0      -0.987232   0.170171
  1      -0.159292  -0.985414

---------------------------
       ROOT = 0
---------------------------
       Total Energy Correction : dE = -0.25309172934720
       Zero Order Energy       : E0 = -106.82353108218946
       Total Energy (E0+dE)    : E  = -107.07662281153667

---------------------------
       ROOT = 1
---------------------------
       Total Energy Correction : dE = -0.23103459727281
       Zero Order Energy       : E0 = -106.77074682157986
       Total Energy (E0+dE)    : E  = -107.00178141885267
```

By construction the Hamiltonian is non-Hermitian (`QDType QD_Bloch`). Hence the computation of properties with the revised wave function e.g. expectation values require left- and right eigenvectors. A single set of eigenvectors ("right") can be constructed using the Des Cloizeaux scheme (`QDType QD_Cloiszeaux`) leading to an Hermitian effective Hamiltonian.[204] The transformation does not change the energies but affects the mixing of states. Note that actual eff. Hamiltonian is printed with a `PrintLevel` larger 4 in the PTSettings subblock. The diagonalization of the general matrices appearing in both formulations may occasionally lead to complex eigenpairs - an undesired artifact. Although, the eigenvalues have typically only a small imaginary component, the results are not reliable and ORCA prints a warning.

```
--- complex eigenvalues and eigenvectors
  WARNING! QD-Matrix has eigenvalues with imaginary component! iE(0)=-0.000016
  WARNING! QD-Matrix has eigenvalues with imaginary component! iE(1)=0.000016
```

The `QD_VanVleck` option avoids the general eigenvalue decomposition. The equations are derived from second order Van Vleck perturbation theory, which results in a Hermitian eff. Hamiltonian.[493] The methodology is equivalent to the symmetrization of the Bloch Hamiltonian. The solution is always real and properties are easily accessible. Thus, **`QD_VanVleck` is the recommended approach in ORCA**. For a more detailed comparison of the different eff. Hamiltonian theories, we refer to the literature.[123, 789]
In all three formulations, the energy denominator in the quasi-degenerate NEVPT2 is very sensitive to approximations. The canonicalization options with averaged orbitals and orbitals energies (`canonstep 0/2`) have the tendency to lessen the energy-denominator. To avoid artifacts, the calculation is restricted to `canonstep 1` — each state has its own orbitals.

If properties are requested within the casscf module i.e. zero-field splitting, there will be an additional printing with the "improved" CI vectors and energies. For technical reasons, properties that are not computed in CASSCF such as the Mössbauer parameters do not benefit from the QD-NEVPT2 correction.

7.20 Complete Active Space Peturbation Theory : CASPT2 and CASPT2-K

The fully internally contracted CASPT2 (FIC-CASPT2) approach is available with real, imaginary and IPEA shifts.[40, 270, 731]. The ORCA implementation employs a reformulation of the CASPT2, that completely avoids the fourth order reduced density matrix, that would appear in the canonical implementation.[458] Some concepts are shared by a recent development reported by Sokolov and coworkers.[162] The modification allows calculations with large active spaces without approximating the results e.g. with the cumulant expansion.

It should be noted that the IPEA shift in OpenMOLCAS slightly deviates from ORCA.[253]. Here, the IPEA shift, λ, is added to the matrix elements of the internally contracted CSFs $\Phi^{pr}_{qs} = E^p_q E^r_s|\Psi^0>$ with the generalized Fock operator

$$< \Phi^{p'r'}_{q's'}|\hat{F}|\Phi^{pr}_{qs} > + =< \Phi^{p'r'}_{q's'}|\Phi^{pr}_{qs} > \cdot \frac{\lambda}{2} \cdot (4 + \gamma^p_p - \gamma^q_q + \gamma^r_r - \gamma^s_s),$$

where $\gamma^p_q =< \Psi^0|E^p_q|\Psi^0 >$ is the expectation value of the spin-traced excitation operator.[441] The labels p,q,r,s refer to general molecular orbitals (inactive, active and virtual). Irrespective of the ORCA implementation, the validity of the IPEA shift in general remains questionable and is thus by default disabled.[921] ORCA features an alternative formulation, named **CASPT2-K**, that revises the zeroth order Hamiltonian itself.[460] Here, two additional Fock matrices are introduced for excitation classes that add or remove electrons from the active space. The new Fock matrices are derived from the generalized Koopmans' matrices corresponding to electron ionization and attachment processes. The resulting method is less prone to intruder states and the same time more accurate compared to the canonical CASPT2 approach. For more a detailed discussion, we refer to the paper by Kollmar et al.[460]

The CASPT2 and CASPT2-K approaches are called in complete analogy to the FIC-NEVPT2 approach. Note that the methodology can be combined with the RI approximation. A detailed example with comments on the output is given in Section *Complete Active Space Perturbation Theory: CASPT2 and CASPT2-K*. Below is concise list with the accessible options.

```
%casscf
  ...

  MULT   1,3 # multiplicity block
  NRoots 2,2 # number of roots for the MULT blocks

  TrafoStep RI      # optional for RI approximation for CASSCF and CASPT2
  PTMethod  FIC_CASPT2  # canonical CASPT2 approach
            FIC_CASPT2K # CASPT2-K with revised H0

  # Detailed settings (this is optional)
  PTSettings
    CASPT2_ishift 0.0    # imaginary level-shift
    CASPT2_rshift 0.0    # real      level-shift
    CASPT2_IPEAshift 0.0 # IPEA shift.
    MaxIter       20     # Maximum for the CASPT2 iterations
    TSmallDenom   1e-2   # printing thresh for small denominators

    # general settings
    NThresh 1e-6  # FIC-CASPT2 cut off for linear dependencies
    D4Tpre  1e-10 # truncation of the 4-pdm
    D3Tpre  1e-14 # trunaction of the 3-pdm
    EWIN  -3,1000 # Energy window for the frozencore setting fc_ewin

    # Option to skip the PT2 correction for a selected multiplicity blocks and roots
    # (same input structure as weights in %casscf)
    selectedRoots[0]=0,1 # skip the first roots of MULT=1
    selectedRoots[1]=0,0 # skip MULT=3 roots

    #CASPT2-K specific options
```

(continues on next page)

(continued from previous page)

```
    TReg 1e-2 # default for the Tikhonov reguralization
  end

end
```

CASPT2 can also be set using the simple keywords on top of any valid CASSCF input.

```
!CASPT2       # FIC-CASPT2
!CASPT2K      # FIC-CASPT2-K
!RI-CASPT2    # FIC-CASPT2 with RI approximation
!RI-CASPT2-K  # FIC-CASPT2-K with RI approximation
%casscf ...
```

7.21 Dynamic Correlation Dressed CAS

DCD-CAS(2) is a post-CASSCF MRPT method of the perturb-then-diagonalize kind, i.e. it can modify the CAS wavefunction compared to the previous CASSCF.[652] In cases where CASSCF already provides a good 0th order wavefunction, DCD-CAS(2) energies are comparable to NEVPT2.

7.21.1 Theory of Nonrelativistic DCD-CAS(2)

The DCD-CAS(2) method is based on solving the eigenvalue problem of an effective Hamiltonian of the form

$$H_{IJ}^{\mathrm{DCD},S} = \langle \Phi_I^{SS}|H|\Phi_J^{SS}\rangle - \sum_{K\in\mathrm{FOIS}} \frac{\langle \Phi_I^{SS}|H|\tilde{\Phi}_K^{SS}\rangle\langle \tilde{\Phi}_K^{SS}|H|\Phi_J^{SS}\rangle}{E_K^S - E_0^S}$$

for each total spin S separately. The 0th order energies E_K^S of the perturbers $|\tilde{\Phi}_K^{SS}\rangle$ are obtained by diagonalizing the Dyall's Hamiltonian in the first-order interacting space (FOIS). The effective Hamiltonian has the form of a CASCI Hamiltonian that is dressed with the effect of dynamic correlation (dynamic correlation dressed, DCD), hence the name for the method. E_0^S is chosen to be the ground state CASSCF energy for the respective total spin S. Since this choice is worse for excited states than for the ground state, excitation energies suffer from a "ground state bias".

For the contribution coming from perturbers in which electrons are excited from two inactive (ij) to two virtual (ab) orbitals, we use (when writing the DCD Hamiltonian in a basis of CASCI states) the alternative expression

$$\langle \Psi_I^{SS}|H^{\mathrm{DCD}}(ij\rightarrow ab)|\Psi_J^{SS}\rangle = -\delta_{IJ}E_{\mathrm{MP2}}$$

$$E_{\mathrm{MP2}} = \sum_{ijab} \frac{(ib|ja)^2 - (ib|ja)(ia|jb) + (ia|jb)^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}$$

Since in this version the $ij\rightarrow ab$ perturber class does not contribute at all to excitation energies (like it is assumed in the difference-dedicated configuration interaction method), we call this the difference-dedicated DCD-CAS(2) method. Since the $ij\rightarrow ab$ class contributes the largest part of the dynamic correlation energy, this also removes the largest part of the ground state bias. This option is used as default in DCD-CAS(2) calculations. In order to also remove the ground state bias from the other perturber classes, we furthermore apply a perturbative correction to the final energies. At first order (which is chosen as default), it takes the form

$$\Delta E_I = -\Delta_I \sum_{K\in\mathrm{FOIS}} \frac{\langle \tilde{\Psi}_I|H|\tilde{\Phi}_K\rangle\langle \tilde{\Phi}_K|H|\tilde{\Psi}_I\rangle}{(E_K - E_0)^2}$$

$$\Delta_I = \langle \tilde{\Psi}_I|H|\tilde{\Psi}_I\rangle - E_0$$

for the correction ΔE_I to the total energy of the Ith DCD-CAS(2) root $|\tilde{\Psi}_I\rangle$.

7.21.2 Treatment of spin-dependent effects

The theory so far is valid for a nonrelativistic or scalar-relativistic Hamiltonian H. If we modify it to a Hamiltonian $H+V$, where V contains effects that are possibly spin-dependent, this leads us to a theory [491] which has a similar form as QDPT with all CAS roots included. The form of the spin-dependent DCD-CAS(2) effective Hamiltonian is

$$\langle \Phi_I^{SM} | H^{\text{DCD}} | \Phi_J^{S'M'} \rangle = \delta_{SS'}\delta_{MM'} H_{IJ}^{\text{DCD},S,\text{corr}} + \langle \Phi_I^{SM} | V | \Phi_J^{S'M'} \rangle.$$

$$\mathbf{H}^{\text{DCD},S,\text{corr}} = \mathbf{C}^{\text{DCD}} \mathbf{E} (\mathbf{C}^{\text{DCD}})^T.$$

In order to construct it, we first need to solve the scalar-relativistic DCD-CAS(2) problem to construct the matrix $\mathbf{H}^{\text{DCD},S,\text{corr}}$ from the bias corrected energies $\mathbf{E}$ and DCD-CAS(2) CI coefficients $\mathbf{C}$ and then calculate the matrix elements of the operators contributing to V in the basis of CSFs $|\Phi_I^{SM}\rangle$.

Zero field splitting D tensors are extracted using the effective Hamiltonian technique, i.e. fitting the model Hamiltonian to a des-Cloiseaux effective Hamiltonian that is constructed from the relativistic states and energies by projection onto the nonrelativistic multiplet (see section *Zero-Field Splitting* and the reference [565]). There are limitations to this approach if spin orbit coupling becomes so strong that the relativistic states cannot uniquely be assigned to a single nonrelativistic spin multiplet.

Hyperfine A-matrices and Zeeman g-matrices for individual Kramers doublets consisting of states $|\Phi\rangle, |\overline{\Phi}\rangle$ are extracted by comparing the spin Hamiltonians

$$H_{\text{Zeeman}} = \mu_B \vec{B} \cdot g \cdot \vec{S}$$

$$H_{\text{HFC}} = \sum_A \vec{I}^A \cdot A^A \cdot \vec{S}$$

to the matrix representation of the many-electron Zeeman and HFC operators in the basis of the Kramers doublet. This yields [491]

$$\begin{aligned} g_{k1} &= 2\Re\langle\overline{\Phi}|L_k + g_e S_k|\Phi\rangle \\ g_{k2} &= 2\Im\langle\overline{\Phi}|L_k + g_e S_k|\Phi\rangle \\ g_{k3} &= 2\langle\Phi|L_k + g_e S_k|\Phi\rangle \end{aligned}$$

$$\begin{aligned} A_{k1} &= -2\gamma_A \Re\langle\overline{\Phi}|B_k^{\text{HFC}}(\vec{R}_A)|\Phi\rangle \\ A_{k2} &= -2\gamma_A \Im\langle\overline{\Phi}|B_k^{\text{HFC}}(\vec{R}_A)|\Phi\rangle \\ A_{k3} &= -2\gamma_A \langle\Phi|B_k^{\text{HFC}}(\vec{R}_A)|\Phi\rangle \end{aligned}$$

where $B_k^{\text{HFC}}(\vec{R}_A)$ is the kth component of the magnetic hyperfine field vector at the position of nucleus A and γ_A is the gyromagnetic ratio.

7.21.3 List of keywords

The following keywords can be used in conjunction with the DCD-CAS(2) method:

```
%casscf
  dcdcas true               # Do a DCD-CAS(2) calculation
  diffded true              # Use difference-dedicated DCD-CAS(2) for the
                            # ij->ab class
  corrorder 1               # Maximum order for the perturbative bias correction
                            # (lower orders come for free)
  dcd_ritrafo false         # Use RI approximation for electron repulsion integrals

  dcd_soc false             # Relativistic DCD-CAS(2) with spin orbit coupling
  dcd_ssc false             # Relativistic DCD-CAS(2) with direct electronic
                            # spin-spin coupling
  dcd_domagfield 0          # Number of user-specified finite magnetic fields
```

(continues on next page)

(continued from previous page)

```
  dcd_dtensor false          # Calculate an effective Hamiltonian D-tensor
  dcd_nmultd 1               # The number of nonrelativistic multiplets for which the
                             # D-tensor is calculated
  dcd_gmatrix false          # Calculate an effective Kramers pair Zeeman g-matrix
  dcd_amatrix false          # Calculate an effective Kramers pair Hyperfine A-matrix
  dcd_kramerspairs 1         # The number of Kramers pairs for which g and/or A
                             # is calculated
  dcd_magnetization false    # Calculate the magnetization of the molecule in an
                             # external magnetic field

  dcd_cascimode false        # Run relativistic calculation in CASCI mode, i.e.
                             # without the dynamic correlation dressing
  dcd_natorbs false          # Calculate natural orbitals for each state and write
                             # them to disk
  dcd_populations false      # Perform population analysis on the DCD-CAS(2) states
end
```

Note that the calculation of SSC requires the definition of an auxiliary basis set, since it is only implemented in conjunction with RI integrals. If `dcd_magnetization` is requested, the values for magnetic flux density and temperature to be used can be specified via the keywords `MAGTemperatureMIN`, `MAGTemperatureMAX`, `MAGTemperatureNPoints`, `MAGFieldMIN`, `MAGFieldMAX`, `MAGNpoints` of the `rel` subblock of the `%casscf` block (see section *Magnetization and Magnetic Susceptibility*). If the keyword `dcd_domagfield` is set to a number different than 0, the magnetic fields can be entered as a matrix of xyz coordinates (in Gauss), e.g.

```
%casscf
  dcdcas true
  ...
  dcd_domagfield 2
  dcd_magneticfields[0] = 10000, 0, 0
  dcd_magneticfields[1] = 0, 10000, 0
end
```

Furthermore, there is the keyword `DCD_EDIAG` that when running the DCD-CAS(2) code in CASCI mode works analogously to the keyword `EDiag` in the `soc` subblock of the `%mrci` block (see section *Zero-Field Splitting*). The only difference is that the energies should be entered in atomic units, not in wavenumbers, e.g.

```
%casscf
  ...
  dcdcas true
  dcd_cascimode true
  dcd_soc true
  DCD_EDIAG[0] -2220.920028
  DCD_EDIAG[1] -2220.834377
  DCD_EDIAG[2] -2220.835871
  DCD_EDIAG[3] -2220.810290
  DCD_EDIAG[4] -2220.812293
  DCD_EDIAG[5] -2220.756732
end
```

7.22 Density Matrix Renormalization Group

The BLOCK code in ORCA is only available on the Linux platform!

BLOCK is an implementation of the density matrix renormalization group (DMRG) algorithm from the Chan group. [156, 157, 158, 296, 788] The references given should be cited when using this part of the program.

The DMRG is a variational wavefunction method. It can be viewed as (i) an efficient method for strong correlation in large complete active spaces, (ii) a brute force method to systematically approach FCI for a large number of electrons and orbitals, (iii) a polynomial cost route to exact correlation in pseudo-one-dimensional molecules, such as chains and rings.

Although the algorithm is somewhat complicated compared to many quantum chemistry methods, significant effort has been devoted in BLOCK to ensure that it can be run in a simple black-box fashion. In most cases, only a single keyword needs to be specified.

To provide an idea of how the DMRG can be used, here are some examples. The timings will vary depending on your computational setup, but the following are calculations that run in a few hours to a day, on a single 12-core Xeon Westmere cluster node:

- Complete active space (CAS) CI calculations for active spaces with up to 30 electrons in 30 active orbitals, targetting up to 1–10 states, e.g. Jacobsen's catalyst in a 32 electron, 25 orbital active space,
- One-dimensional chain molecules, with "widths" of up to 4 orbitals, and about 100 orbitals in total, e.g. the π-active space of a 4×25 graphene nanoribbon,
- FCI benchmark solutions in molecules with fewer than 20 electrons, and up to 100 orbitals, e.g. C_2 in a cc-pVTZ basis, D_{2h} symmetry (12 electrons in 60 orbitals),
- Accuracies in energy differences or total energies of about 1 kcal/mol.

The following are calculations which are possible with the BLOCK code, but which are challenging, and require large memory (e.g. up to 8 GB per core) and computational time (e.g. from a day to more than a week on up to 6 12-core Xeon Westmere nodes),

- Complete active space (CAS) CI calculations in active spaces with around 40 electrons in 40 active orbitals, targetting a few states, for example, an Fe(II)-porphine (40 electrons in 38 orbitals) with an active space of Fe 3d, 4d and all porphine π and σ donor orbitals, or an Fe 3d, S 3p active space calculation for $[Fe_4S_4(SCH_3)_4]^{2-}$,
- One-dimensional chain molecules, with "widths" of up to 6 orbitals, and about 100 total orbitals,
- Champion FCI benchmark solutions in small molecules, such as butadiene in a cc-pVDZ basis (22 electrons in 82 orbitals),
- Accuracies in energy differences or total energies of about 1 kcal/mol.

If any these calculations interest you, then you might want to try a DMRG calculation with BLOCK!

7.22.1 Technical capabilities

Currently, BLOCK implements the following

- An efficient DMRG algorithm for quantum chemistry Hamiltonians
- Full spin-adaptation (SU(2) symmetry) and Abelian point-group symmetries
- State-averaged excited states

Note that the standalone version of BLOCK may provide more capabilities than are available through the external interface. See the BLOCK website for details [155].

7.22.2 How to cite

We would appreciate if you cite the following papers in publications resulting from the use of BLOCK :

- G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002),
- G. K.-L. Chan, *J. Chem. Phys.* **120**, 3172 (2004),
- D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.*, **128**, 144117 (2008),
- S. Sharma and G. K-.L. Chan, *J. Chem. Phys.* **136**, 124121 (2012).

In addition, useful DMRG references relevant to quantum chemistry can be found in the review below by Chan and Sharma.

- G. K-.L. Chan and S. Sharma, *Ann. Rev. Phys. Chem.* **62**, 465 (2011),

7.22.3 Overview of BLOCK input and calculations

Within ORCA, the BLOCK program is accessed as part of the CASSCF module. BLOCK can be run in two modes: CASCI mode (no orbital optimization) or CASSCF mode. To enable CASCI mode, set `maxiter 1`.

```
%casscf
   maxiter 1 # remove if doing CASSCF
   CIStep DMRGCI
   ...
   end
```

For small molecule CASCI it may be possible to correlate all orbitals. In general, similar to a standard CASSCF calculation, it is necessary to select a sensible active space to correlate. (See Section *Orbital optimization* on CASSCF). This is the responsibility of the user.

7.22.4 Standard commands

Once the orbitals to correlate have been chosen, and the wavefunction symmetries and quantum numbers are specified, the accuracy of the DMRG calculation is governed by two parameters: the maximum number of renormalized states M; and, the order and localization of the orbitals.

The most important parameter in the DMRG calculation is M, the number of renormalized states. This defines the maximum size of the wave-function expansion, which is $O(M^2)$ in length in the renormalized basis. As M is increased, the DMRG energy converges to the exact (FCI or CASCI) limit.

The DMRG maps orbitals onto a 1D lattice, thus the best results are achieved if strongly interacting orbitals are placed next to each other. For this reason, the DMRG energy is not generally invariant to orbital rotations within the active space, and orbital rotation and ordering can improve the DMRG energy for a given M. As M is increased, the DMRG energy becomes less and less sensitive to the orbital ordering and localization.

To minimize the number of wavefunction optimization steps, it is often advantageous to perform DMRG calculations at small M, then increase M to the final maximum value. This sequence of optimizations is governed by the *sweep schedule*, which specifies how many optimization steps (sweeps) to perform at each intermediate value of M.

The above may seem to make running a DMRG calculation more complicated than a usual quantum chemistry calculation, however, BLOCK provides a set of default settings which eliminate the need to specify the above parameters by hand. We highly recommend that you first learn to use the BLOCK program *with these default settings*. In the default mode, the orbitals are ordered automatically (Fiedler vector method [58, 72, 259, 260]) and a default sweep schedule is set.

An example of a default CASCI calculation on the C~2~ molecule correlating all electrons in a VTZ basis, is given here:

```
!cc-pvtz pal4
%MaxCore 16000
%casscf
    nel 8
    norb 58
    nroots 1
    mult 1
    maxiter 1
    CIStep DMRGCI
    DMRG
        maxM 5000
    end
end

* xyz 0 1
C 0 0 -0.621265
C 0 0  0.621265
*
```

Once you are familiar with the default mode, we recommend exploring the localization of orbitals. In general, DMRG benefits from the use of localized orbitals, and these should be used unless the high-symmetry of the molecule (e.g., D_{2h} symmetry) provides compensating computational benefits. We recommend using "split-localized" orbitals, which correspond to localizing the occupied and virtual orbitals separately. An example of a split-localized default DMRG calculation on the porphine molecule, correlating the full π-space (26 electrons in 24 orbitals), in a cc-pVDZ basis is given in Sec. *Appendix: Porphine \pi-active space calculation*.

For a given `maxM`, it can take a long time to tightly converge DMRG calculations (e.g. to the default 1e-9 tolerance). To decrease computation time, you may wish to loosen the default tight sweep tolerance or control the maximum number of sweep iterations with the commands `sweeptol` and `maxIter`.

Orbital optimization

Orbital optimization (mixing the external/internal space with the active space, not to be confused with orbital rotation and ordering in the active space) in DMRG calculation can be performed by using the `BLOCK` program as the "CIStep" within a CASSCF calculation, as described above. For the moment, spin-densities and related properties are not available for this `CIStep`.

During the optimization iterations it is important that the active orbitals maintain their overlap and ordering with previous iterations. This is done using `actConstrains`. This flag is set by default.

```
%casscf
ActConstrains  1 # maintain shape and ordering of active orbitals
...
end
```

In general, performing a DMRG calculation with orbital optimization is quite expensive. Therefore, it is often best to carry out the orbital optimization using a small value of `maxM` (enabled by the default parameters `maxM=25` and the resulting sweep schedule), and to carry out a final single-point calculation using a larger value of `maxM`.

Advanced options

There may be times when one wants finer control of the DMRG calculation. All keywords are shown in the *complete set of BLOCK options* *Complete set of BLOCK options* below. The `startM` command allows to change the starting number of states in DMRG calculations. It is also possible to specify the entire sweep schedule manually. A sweep schedule example follows:

```
%casscf
   ...
  dmrg

    MaxIter          14
    switch_rst       1e-3
    TwoDot_to_oneDot 12
    NSchedule        3
    sche_iteration   0,     4,    8
    sche_M          50,   100,  500
    sche_sweeptol 1e-4,  1e-6, 1e-9
    sche_noise    1e-8, 1e-11,  0.0

  end
end
```

The commands above are:

- `MaxIter`, corresponds to the maximum number of sweeps done by DMRG;
- `NSchedule`, specifies the total number of schedule parameters we will specify;
- `Sche_iteration`, details the sweep number at which to change the parameters of the calculation. Notice count begins at 0;
- `Sche_M`, is the number of renormalized states at each sweep;
- `Sche_sweeptol`, is the tolerance of the Davidson algorithm;
- `Sche_noise`, is the amount of perturbative noise we add each sweep;
- `Twodot_to_onedot`, specifies the sweep at which the switch is made from a twodot to a onedot algorithm. The recommended choice is to start with twodot algorithm and then switch to onedot algorithm a few sweeps after the maximum M has been reached. To do a calculation entirely with the twodot or the onedot algorithm, replace the `twodot_to_onedot` line with `twodot 1` or `onedot 1`;
- `switch_rst`, defines the switching threshold of orbital gradient below which DMRG turns to onedot algorithm and restarts from previous operators and wavefunction. This is essential to avoid oscillation of energy values in the orbital optimization.

The default DMRG sweep schedule is selected automatically according to the choice of computational mode. By default two different sets of predefined schedules are supported for CASCI and CASSCF computations, respectively.

In CASCI mode, the default schedule corresponds to the following: starting from a given `startM` (where the default is 250 and 8 sweeps), increase to a value of 1000 (8 sweeps) and increment by 1000 every 4 iterations until `maxM` is reached. The algorithm switches from twodot to onedot two sweeps after the `maxM` has been reached.

In CASSCF mode, the orbital optimization requires much fewer renormalized states to converge the wavefunction with respect to orbital rotations. The default schedule therefore starts with `startM` (where the default is 25 and 2 sweeps), and increments by a factor of 2 every 2 sweeps util `maxM` is reached. The algorithm continues the sweep at `maxM` by decreasing the Davison tolerance `sche_sweeptol` and noise level `sche_noise` every 2 cycles by a factor of 10, until `sche_sweeptol` becomes smaller than `sweeptol`.

For better control of the orbital ordering, we also provide a genetic algorithm minimization method of a weighted exchange matrix. The genetic algorithm usually provides a superior orbital ordering to the default ordering, but can itself take some time to run for large numbers of orbitals. The genetic algorithm can be enabled by

```
%casscf
   ...
   DMRG
     auto_ordering GAOPT
   end
end
```

within the %casscf input.

Troubleshooting

The two most common problems with DMRG calculations are that (i) convergence with maxM is slower than desired, or (ii) the DMRG sweeps get stuck in a local minimum. (i) is governed by the orbital ordering / choice of orbitals. To improve convergence, turn on the genetic algorithm orbital ordering.

If you suspect (ii) is occurring, the simplest thing to do is to increase the starting number of states with the startM (e.g. from 500 to 1000 states). Local minima can also sometimes be avoided by increasing the noise in the DMRG schedule, e.g. by a factor of 10. To check that you are stuck in a local minimum, you can carry out a DMRG extrapolation (see extended Manual in the BLOCK website).

Note that the present DMRG-SCF establishes the input order of active space orbitals according to their Hartree-Fock occupancy, even if these orbitals are ultimately canonical or split-localized canonical in nature. This is specified by hf_occ in which the Hartree-Fock occupancy is derived by default from the one-electron integrals. Other options for obtaining the occupancy are available (see *Complete set of BLOCK options*).

Somet times the energy values produced from one SCF cycle to another may oscillate. Such a nonlinear numerical behaviour may occur typically by the last few iterations, most likely caused by the loss of a certain distribution of quantum numbers (eg, particle number, irrep symmetry and spin) in the blocking and decimation procedure due to incomplete many-body basis. On the other hand, the loss of quantum numbers is the main source of energy discontinuities on potential energy curves calculated by DMRG-SCF using a small number of renormalized states.

In the current release of DMRG-SCF implementation, the number of quantum states is locked to avoid these problems. The locking mechanism is turned on when the orbital gradient falls below a certain threshold defined by the keyword switch_rst (default: 0.001). The DMRG calculation then starts from previous operators and wavefunction in which a perturbative noise is not added. Locking quantum states and restaring DMRG wavefunction not only ensures a smooth convergence towards the final energy but also minimizes the number of iterations. Note that the locking procedure introduces an arbitrariness to the final energy, when a very small M is used, since the final digits of energy depend on where the locking begins. It is therefore not recommended to start locking too early in iterations which could trap the orbital solution in a local mimimum. Finally the quality of resulting orbitals can be checked by carrying out a DMRG calculation with sufficient renormalized states. Using the default value of switch_rst DMRG-SCF usually results in the orbitals that are good enough to reproduce the CASSCF energy.

Complete set of BLOCK options

```
%casscf
 ...
 dmrg

 startM   25     # CASSCF mode: number of re-normalized states for a singlee root
          250    # CASCI mode: number of re-normalized states for a single root
 maxM     25     # CASSCF mode: number of re-normalized states for a singlee root
          250    # CASCI mode: number of re-normalized states for a single root
 DryRun   false  # just create an input for Block
 SweepTol 1e-9   # energy tolerance for the sweeps
 auto_ordering NOREORDER # auto_ordering is an int. If set to 0
                         # or the alias NOREORDER, the reordering is skipped.
               FIEDLER   # (default) let Block optimize the active orbital ordering
               GAOPT     # let Block optimize the active orbital ordering
                         # using genetic algorithm
```

(continues on next page)

(continued from previous page)

```
 hf_occ   0 # user-defined initial Hartree-Fock occupancy manually
          1 # default: initial Hartree-Fock occupancy based on the values of
              the one-electron integrals
          2 # initial Hartree-Fock occupancy based on the energy ordering
              of canonical orbitals

 TwoDot_to_OneDot  1  # Switch from two-dot expressions to one-dot
 OneDot  0   # Only one-dot expressions. %In CASCI mode only.
 TwoDot  0   # Only two-dot expressions. %In CASCI mode only.
 switch_rst 1e-3   # Specify the threshold of orbital gradient below which DMRG
                     swithches to one-dot expression by reading previous wavefunction.
 warmup  1   # wilson warm-up type
         2, 3 or 4   # n=3 is the default option.
                       The full configuration space of the n sites next to the system
                       constitutes the environment states in the warm-up.
                       The remaining sites use the Hartree-Fock guess occupation
 nonspinadapted 0   # default: spin-adapted DMRG
                1   # non-spin-adatped DMRG in which the spin-density calculation
                      is available


 # Define a schedule for DMRG
 MaxIter   14   # Specify maximum number of iterations
 NSchedule -1   # default sweep schedule in CASSCF mode
            0   # default sweep schedule in CASCI mode
           >0   # Number of manual sweep schedule parameters
                # All schedule parameters must be set if this flag is set manually!
 sche_iteration 0, 4, 8   # vector with sweep-number to execute changes
                          # (schedule parameter)

 sche_M        50,100,500   # vector with corresponding M values (schedule parameter)
 sche_sweeptol 1e-4,1e-6,1e-9  # vector with sweep tolerances (schedule parameter)
 sche_noise    1e-8, 1e-11,0.0 # vector with the noise level (schedule parameter)

 # Define a separate maxM for DMRG-NEVPT2
 nevpt2_maxm 25   # set maximum number of renormalized states
                    for DMRG-NEVPT2 calculation (default: MaxM)

 end
end
```

7.22.5 Appendix: Porphine π-active space calculation

We provide a step-by-step basis on localizing the π-orbitals of the porphine molecules and running a CASSCF-DMRG calculation on this system. It will be important to obtain an initial set of orbitals, rotate the orbitals which are going to be localized, localize them, and finally run the CASSCF calculation. We will abbreviate the coordinates as $[\ldots]$ after showing the coordinates in the first input file, but please note they always need to be included.

1. First obtain RHF orbitals:

```
# To obtain RHF orbitals
!cc-pvdz
* xyz 0 1
N      2.10524     -0.00000     0.00000
N     -0.00114      1.95475    -0.00000
N     -2.14882      0.00000    -0.00000
N     -0.00114     -1.95475     0.00000
C      2.85587     -1.13749    -0.00000
C      2.85587      1.13749     0.00000
C      1.02499      2.75869    -0.00000
```

(continues on next page)

(continued from previous page)

```
C     -1.10180      2.78036     0.00000
C     -2.93934      1.13019    -0.00000
C     -2.93934     -1.13019     0.00000
C     -1.10180     -2.78036    -0.00000
C      1.02499     -2.75869     0.00000
C      4.23561     -0.67410    -0.00000
C      4.23561      0.67410     0.00000
C      0.69482      4.18829    -0.00000
C     -0.63686      4.14584    -0.00000
C     -4.25427      0.70589    -0.00000
C     -4.25427     -0.70589     0.00000
C     -0.63686     -4.14584     0.00000
C      0.69482     -4.18829     0.00000
H      5.10469     -1.31153     0.00000
H      5.10469      1.31153    -0.00000
H      1.36066      5.02946     0.00000
H     -1.28917      5.00543     0.00000
H     -5.12454      1.34852     0.00000
H     -5.12454     -1.34852    -0.00000
H     -1.28917     -5.00543    -0.00000
H      1.36066     -5.02946    -0.00000
C      2.46219      2.41307     0.00000
C     -2.39783      2.44193     0.00000
C     -2.39783     -2.44193    -0.00000
C      2.46219     -2.41307    -0.00000
H      3.18114      3.22163    -0.00000
H     -3.13041      3.24594    -0.00000
H     -3.13041     -3.24594     0.00000
H      3.18114     -3.22163     0.00000
H      1.08819      0.00000    -0.00000
H     -1.13385     -0.00000     0.00000
*
```

2. We then swap orbitals with π-character so they are adjacent to each other in the active space. (π orbitals are identified by looking at the MO coefficients). When they are adjacent in the active space, they can be easily localized in the next step.

```
#To rotate the orbitals (so that we can localize them in the next step)
!cc-pvdz moread noiter
%moinp "porphine.gbw"
%scf
    rotate
    # Swap orbitals
    {70, 72}
    {65, 71}
    {61, 70}
    {59, 69}
    {56, 68}
    {88, 84}
    {92, 85}
    {93, 86}
    {96, 87}
    {99, 88}
    {102, 89}
    {103, 90}
    {104, 91}
    end
end
* xyz 0 1
[...]
*
```

3. After rotating the orbitals, we localize the 13 occupied π-orbitals. This is performed using the `orca_loc` code. The input file follows.

```
porphine_rot.gbw
porphine_loc.gbw
0
68
80
120
1e-3
0.9
0.9
1
```

4. After localizing the occuppied orbitals, we localize the 11 virtual π-orbitals using the `orca_loc` code once again. The input file follows.

```
porphine_loc.gbw
porphine_loc2.gbw
0
81
91
120
1e-3
0.9
0.9
1
```

5. After these steps are complete, we run a CASSCF-DMRG calculation. The standard input file is shown below

```
!cc-pvdz moread pal4
%moinp "porphine_loc2.gbw"
%MaxCore 16000

%casscf nel 26
norb 24
nroots 1
CIStep DMRGCI
end
* xyz 0 1
[...]
*
```

7.23 Relativistic Options

The relativistic methods in ORCA are implemented in a fairly straightforward way but do require some caution from the user. The options are controlled through the `%rel` block which features the following variables:

```
%rel
      #-----------------------------------------------------
      # Basic scalar relativistic method
      #-----------------------------------------------------
      method    DKH       # Douglas-Kroll-Hess
                ZORA      # ZORA (numerical integration)
                IORA      # IORA (numerical integration)
                IORAmm    # IORA with van Wuellens
                          #  modified metric
                X2C       # Exact two-component
     # -----------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
# Choice of the model potential for ALL methods
# ----------------------------------------------------
ModelPot  VeN, VC, VXa, VLDA, VPC
         # Flags for terms in the model potential
         # =0 not included =1 included
         # WARNING: default is currently 1,1,1,1 for ZORA and IORA and
         # VeN = nuclear attraction term
         # VC  = model Coulomb potential (ZORA/IORA only)
         # VXa = model Xalpha potential (ZORA/IORA only)
         # VLDA= VWN-5 local correlation model pot. (ZORA/IORA only)
         # VPC = external point charges (X2C only)
Xalpha 0.7 # default value for the X-Alpha potential,
           # only has an effect when VXa is part of the model potential
# ----------------------------------------------------
# This variable determines the type of fitted atomic
# density that enters the Coulomb potential part of the
# model potential (has no effect when using DKH):
# ----------------------------------------------------
ModelDens     rhoDKH  # DKH4 model densities (default)
              rhoZORA # ZORA model densities
              rhoHF   # Hartree-Fock model densities
# ----------------------------------------------------
# This flag controls whether only one center terms
# are retained. If this is true an approximate treat-
# ment of relativistic effects is the result, but
# geometry optimizations CAN BE PERFORMED WITH ALL
# METHODS AND MODEL POTENTIALS
# In addition one gets NO gauge noninvariance
# errors in ZORA or IORA
# ----------------------------------------------------
OneCenter false     # default value
# ----------------------------------------------------
# Flag for the diagonal approximation to the unitary
# decoupling matrix (DLU) in X2C. Mutually exclusive
# with OneCenter. See section "DLU approximation".
# ----------------------------------------------------
DLU       false     # default value
# ----------------------------------------------------
# Specify the speed of light used in relativistic
# calculations
# ----------------------------------------------------
C       137.0359895 # speed of light used (137.0359895 is the default value)
                    # synonyms for C are VELIT, VELOCITY
# ----------------------------------------------------
# Picture change for properties
# ----------------------------------------------------
PictureChange 0 # (or false): no picturechange (default)
              1 # (or true):  include picturechange
              2 # for DKH:    use second-order DKH transformation
                #             (see section "Picture-Change Effects")
              2 # for X2C:    include the response of the unitary
                #             decoupling transformation (see section
                #             "Exact Two-Component Theory")
# ----------------------------------------------------
# Order of DKH treatment (this has no effect on ZORA calculations)
# ----------------------------------------------------
order 1 # first-order DKH Hamiltonian
      2 # second-order DKH Hamiltonian
# ----------------------------------------------------
# Kind of Foldy-Wouthuysen transformation for picturechange effects
# in g tensors (see section "Picture-Change Effects")
```

(continues on next page)

(continued from previous page)

```
    # ----------------------------------------------------
    fpFWtrafo true   # do not include vector potential into momentum (default)
              false  # include vector potential
    # ----------------------------------------------------
    # Finite Nucleus Model: (see section "Finite Nucleus Model")
    # ----------------------------------------------------
    FiniteNuc false  # Use point-charge nuclei (default)
              true   # Use finite nucleus model
    # ----------------------------------------------------
    # X2C intermediate storage level: (see section "X2C derivatives and properties")
    # ----------------------------------------------------
    StorageLevel 0-5 # default: 2
    end
```

Note

It is important to recognize that in the one-center approximation (`OneCenter true`) ALL methods can be used for geometry optimization. Several papers in the literature show that this approximation is fairly accurate for the calculation of structural parameters and vibrational frequencies. Since this approximation is associated with negligible computational effort relative to the nonrelativistic calculation it is a recommended procedure.

7.23.1 Approximate Relativistic Hamiltonians

In the relativistic domain, calculations are based on the one-electron, stationary Dirac equation in atomic units (rest mass subtracted)

$$h_D\Psi = \left(\left(\beta - 1\right)c^2 + c\boldsymbol{\alpha}\cdot\boldsymbol{p} + V\right)\Psi = E\Psi. \tag{7.162}$$

The spinor Ψ can be decomposed in its so-called large and small components

$$\Psi = \begin{pmatrix}\Psi_L\\ \Psi_S\end{pmatrix} \tag{7.163}$$

These are obviously coupled through the Dirac equation. More precisely, upon solving for Ψ_S, the following relation is obtained:

$$\Psi_S = \frac{1}{2c}\left(1 + \frac{E-V}{2c^2}\right)^{-1}\boldsymbol{\sigma}\cdot\boldsymbol{p}\Psi_L = R\Psi_L \tag{7.164}$$

Through the unitary transformation

$U = \begin{pmatrix}\Omega_+ & -R^+\Omega\\ R\Omega_+ & \Omega_-\end{pmatrix}$ with $\Omega_+ = \frac{1}{\sqrt{1+R^+R}}, \Omega_- = \frac{1}{\sqrt{1+RR^+}}$,

the Hamiltonian can be brought into block-diagonal form

$$U^+h_DU = \begin{pmatrix}\tilde{h}_{++} & 0\\ 0 & \tilde{h}_{--}\end{pmatrix} \tag{7.165}$$

The (electronic) large component thus has to satisfy the following relation

$$h_{++}\Psi_L = \Omega_+\left(h_{++} + h_{\pm}R + R^+\left(h_{\mp} + h_{--}R\right)\right)\Omega_+\Psi_L = E_+\Psi_L. \tag{7.166}$$

The approximate relativistic schemes implemented in ORCA use different methods to substitute the exact relation (7.166) with approximate ones.

Two approximation schemes are available in ORCA: the regular approximation and the Douglas-Kroll-Hess (DKH) approach.

7.23.2 The Regular Approximation

In the regular approximation, (7.166) is approximated by

$$R = \frac{c}{2c^2 - V}\boldsymbol{\sigma} \cdot \boldsymbol{p}. \tag{7.167}$$

At the zeroth-order level (ZORA), $\Omega_{\pm} = 1$, so that the ZORA transformation is simply

$$U_{\mathrm{ZORA}} = \begin{pmatrix} 1 & -R^{+} \\ R & 1 \end{pmatrix} \tag{7.168}$$

and the corresponding Hamiltonian given by

$$\tilde{h}_{++}^{\mathrm{ZORA}} = V + c\boldsymbol{\sigma} \cdot \boldsymbol{p}\frac{1}{2c^2 - V}c\boldsymbol{\sigma} \cdot \boldsymbol{p}. \tag{7.169}$$

At the infinite-order level (IORA), $\Omega_{\pm}$is taken into account, so that

$$U_{\mathrm{IORA}} = U_{\mathrm{ZORA}} \begin{pmatrix} \Omega_{+} & 0 \\ 0 & \Omega_{-} \end{pmatrix} \tag{7.170}$$

and

$$\tilde{h}_{++}^{\mathrm{IORA}} = \Omega_{+} \left(V + c\boldsymbol{\sigma} \cdot \boldsymbol{p}\frac{1}{2c^2 - V}c\boldsymbol{\sigma} \cdot \boldsymbol{p} \right) \Omega_{+} \tag{7.171}$$

is the corresponding Hamiltonian. Note that despite the name – *infinite-order* regular approximation – this is still *not* exact.

In ORCA, the spin-free (scalar-relativistic) variant of ZORA and IORA are implemented. These are obtained from those above through the replacement

$$\boldsymbol{\sigma} \cdot \boldsymbol{p}\frac{1}{2c^2 - V}\boldsymbol{\sigma} \cdot \boldsymbol{p} \rightarrow \boldsymbol{p}\frac{1}{2c^2 - V}\boldsymbol{p}. \tag{7.172}$$

The regular Hamiltonians contain only part of the Darwin term and no mass-velocity term. A problem with relations (7.171) and (7.169) is that due to the non-linear dependence of the resulting regular Hamiltonians on V, a constant change of V, which in the Dirac and Schrödinger equations will result in a corresponding change of energy

$$E \rightarrow E + \mathrm{const}$$

does not so in the regular approximation. Several attempts have been made to circumvent this problem. The scaled ZORA variant is one such procedure. Another one is given through the introduction of model potentials replacing V. Both approaches are available in ORCA.

The scaled ZORA variant

This variant goes back to van Lenthe et al. [864]. The central observation is that the Hamiltonian

$$h_{\mathrm{scaledZORA}} = \frac{h_{\mathrm{ZORA}}}{1 + \left\langle \Psi_L \left| c\boldsymbol{\sigma} \cdot \boldsymbol{p}\frac{1}{(2c^2 - V)^2}c\boldsymbol{\sigma} \cdot \boldsymbol{p} \right| \Psi_L \right\rangle} \tag{7.173}$$

produces constant energy-shifts $E \rightarrow E + \mathrm{const}$ when the potential V is changed by a constant – *for hydrogenic ions*. For many-electron systems, the scaled-ZORA Hamiltonian still does not yield simple, constant energy shift for $V \rightarrow V + \mathrm{const}$. But it produces the exact Dirac energy for hydrogen-like atoms and performs better than the first-order regular approximation for atomic ionization energies.

The regular approximation with model potential

The idea of this approach goes back to Van Wüllen [866], who suggested the procedure for DFT. However we also use it for other methods. The scalar relativistic ZORA self-consistent field equation is in our implementation (in atomic units):

$$\left[\mathbf{p}\frac{c^2}{2c^2-V}\mathbf{p}+V_{\text{eff}}\right]\psi_i=\varepsilon_i\psi_i \tag{7.174}$$

where c is the speed of light. It looks like the normal nonrelativistic Kohn–Sham equation with the KS potential V_{eff}:

$$V_{\text{eff}}\left(\mathbf{r}\right)=-\sum_A\frac{Z_A}{\left|\mathbf{r}-\mathbf{R}_A\right|}+\int\frac{\rho\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}d\mathbf{r}'+V_{\text{xc}}\left[\rho\right]\left(\mathbf{r}\right) \tag{7.175}$$

(Z_A is the charge of nucleus A and R_A is its position; $\rho(r)$ is the total electron density and $V_{xc}\left[\rho\right]$ the exchange-correlation potential – the functional derivative of the exchange-correlation energy with respect to the density). The kinetic energy operator $T=-\frac{1}{2}\nabla^2$ of the nonrelativistic treatment is simply replaced by the ZORA kinetic energy operator:

$$T^{\text{ZORA}}=\mathbf{p}\frac{c^2}{2c^2-V}\mathbf{p} \tag{7.176}$$

Clearly, in the regions where the potential V is small compared to c^2, this operator reduces to the nonrelativistic kinetic energy. V could be the actual KS potential. However, this would require to solve the ZORA equations in a special way which demands recalculation of the kinetic energy in every SCF cycle. This becomes expensive and is also undesirable since the ZORA method is not gauge invariant and one obtains fairly large errors from such a procedure unless special precaution is taken. Van Wüllen [866] has therefore argued that it is a reasonable approximation to replace the potential V with a model potential $\tilde{V}_{\text{model}}$ which is constructed as follows:

$$\tilde{V}_{\text{model}}=-\sum_A\frac{Z_A}{\left|\mathbf{r}-\mathbf{R}_A\right|}+\int\frac{\rho_{\text{model}}\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}d\mathbf{r}'+V_{\text{xc}}^{\text{LDA}}\left[\rho^{\text{model}}\right]\left(\mathbf{r}\right) \tag{7.177}$$

The model density is constructed as a sum over spherically symmetric (neutral) atomic densities:

$$\rho_{\text{model}}\left(\mathbf{r}\right)=\sum_A\rho^A\left(\mathbf{r}\right) \tag{7.178}$$

Thus, this density neither has the correct number of electrons (for charged species) nor any spin polarization. Yet, in the regions close to the nucleus, where the relativistic effects matter, it is a reasonable approximation. The atomic density is expanded in a sum of s-type Gaussian functions like:

$$\rho^A\left(\mathbf{r}\right)=\sum_i d_i\exp\left(-\alpha_i\left|\mathbf{r}-\mathbf{R}_A\right|^2\right) \tag{7.179}$$

The fit coefficients were determined in three different ways by near basis set limit scalar relativistic atomic HF calculations and are stored as a library in the program. The fitting densities are available for elements up to Rn, as well as the actinoids. Through the variable `ModelDens` (*vide supra*) the user can choose between these fits and study the dependence of the results in this choice (it should be fairly small except, perhaps, with the heavier elements and the HF densities which are not recommended). The individual components of the model potential (eq. (7.177)) can be turned on or off through the use of the variable `ModelPot` (*vide supra*).

Van Wüllen has also shown that the calculation of analytical gradients with this approximation becomes close to trivial and therefore scalar relativistic all electron geometry optimizations become easily feasible within the ZORA approach. However, since T^{ZORA} is constructed by numerical integration it is very important that the user takes appropriate precaution in the use of a suitable integration grid and also the use of appropriate basis sets! In the case of `OneCenter true` the numerical integration is done accurately along the radial coordinate and analytically along the angular variables such that too large grids are not necessary unless your basis set is highly decontracted and contains very steep functions.

7.23.3 The Douglas-Kroll-Hess Method

The Douglas-Kroll-Hess (DKH) method expands the exact relation (7.166) in the external potential V. In ORCA the first- and second-order DKH methods are implemented. The first-order DKH Hamiltonian is given by

$$\tilde{h}^{(1)}_{++} = E_p + A_p V A_p + B_p V^{(p)} B_p, \tag{7.180}$$

with

$$E_P = \sqrt{c^4 + c^2 p^2},\ A_p = \sqrt{\frac{E_p + c^2}{2E_p}},\ B_p = \frac{c}{\sqrt{2E_p(E_p + c^2)}} \tag{7.181}$$

At second order, it reads

$$\tilde{h}^{(2)}_{++} = \tilde{h}^{(1)}_{++} + \frac{1}{2}\left[W_p, O\right] \tag{7.182}$$

where

$$\{W_p, E_p\} = \beta O,\ O = A_p\left[R_p, V\right]A_p,\ R_p = \frac{c\sigma p}{E_p + c^2} \tag{7.183}$$

define the second-order contribution. In ORCA, the spin-free part of $\tilde{h}^{(2)}_{++}$ is implemented.

The occurrence of the relativistic kinetic energy, E_P, which is not well-defined in position space, makes a transformation to the p^2-eigenspace necessary. Thus any DKH calculation will start with a decontraction of the basis set, to ensure a good resolution of the identity. Then the non-relativistic kinetic energy is diagonalized and the E_P-dependent operators calculated in that space. The potential V and $V^{(p)}$ are transformed to p^2-eigenspace. After all contributions are multiplied to yield the (first- or second-order) Hamiltonian, the transformation back to AO space is carried out and the basis is recontracted.

The (spin-free) DKH-Hamiltonians contain all spin-free, relativistic correction terms, e.g. the mass-velocity and Darwin terms. As the potential enters linearly, no scaling or model potential is necessary to introduce the correct behaviour of the energy under a change

$$V \to V + \text{const.}$$

In all these respects the DKH Hamiltonians are much cleaner than the regular Hamiltonians.

7.23.4 Picture-Change Effects

Irrespective of which Hamiltonian has been used in the determination of the wave function, the calculation of properties requires some special care. This can be understood in two ways: First of all, we changed from the ordinary Schrödinger Hamiltonian to a more complicated Hamiltonian. As properties are defined as derivatives of the energy, it is clear that a new Hamiltonian will yield a new expression for the energy and thus a new and different expression for the property in question. Another way of seeing this is that through the transformation U, we changed not only the Hamiltonian but also the wave function. To obtain the property at hand as the expectation value of the property operator with the wave function, we have to make sure that property operator and wave function are actually given in the same space. This is done through a transformation of either the property operator or the wave function.

In any case, the difference between the non-relativistic and (quasi) relativistic property operator evaluated between the (quasi) relativistic wave function is called the picture-change effect. From what was said above, this is clearly not a physical effect. It describes how consistent the quasi relativistic calculation is carried out. A fully consistent calculation requires the determination of the wave function on the (quasi) relativistic level as well as the use of the (quasi) relativistic property operator. This is obtained through the choice

```
%rel PictureChange 1 end  # or 2 - see below
```

in the %rel block. It may be that the (quasi) relativistic and non-relativistic property operator do produce similar results. In this case, a calculation with picture-change turned off (PictureChange=0) may be a good approximation. This is, however, not the rule and cannot be predicted before carrying out the calculation. It is therefore highly recommended to turn on picture-change in all (quasi) relativistic property calculations!

For DKH2, the fully consistent picture-change effects are obtained using the same transformation order for the property operator as for the one-electron Hamiltonian, i.e. setting

```
%rel PictureChange 2 end
```

while with PictureChange=1 only first-order changes on the property operators are taken into account, which reduces the computational cost. However, since this is in no way a significant reduction, this choice is not recommended. A similar argument applies to X2C (see *X2C derivatives and properties*).

For magnetic properties, the DKH transformation and consequently the DKH Hamiltonian and the corresponding property operators are not unique. Depending on whether the magnetic field is included in the free-particle Foldy–Wouthuysen (fpFW) transformation carried out in the first step of the DKH protocol or not, two different Hamiltonians result. If the magnetic field is included in the fpFW transformation, the resulting Hamiltonian is a function of the gauge invariant momentum

$$\boldsymbol{\pi} = \mathbf{p} + \mathbf{A}.$$

It is therefore gauge invariant under gauge transformations of the magnetic vector potential $\mathbf{A}$ and thus are the property operators derived from it. This is referred to as fπFW DKH Hamiltonian. If the magnetic field is not included in the FW transformation, the resulting Hamiltonian is a function of the kinetic momentum $\mathbf{p}$ only and thus is not gauge invariant. The latter Hamiltonian is referred to as fpFW DKH Hamiltonian. A comparison of both Hamiltonians is given in Table Table 7.20.

Table 7.20: Comparison of the properties of the fpFW and fπFW DKH Hamiltonians. For details see Ref. [745].

Criterion	fpFW Hamiltonian	fπFW Hamiltonian
Convergence of Eigenvalues to Dirac Eigenvalues	?	yes
1st order is bounded	no	yes
Reproduces Pauli Hamiltonian	no	yes
Gauge invariance	no	yes
Lorentz invariance	no	no

From this Table, it becomes clear that the fπFW DKH Hamiltonian is clearly preferred over the fpFW Hamiltonian. To obtain the property operators, it is however necessary to take the derivatives of these Hamiltonians. It turns out that in the case of the hyperfine-coupling tensor, the necessary derivatives produce divergent property operators in the case of the fπFW DKH Hamiltonian. This may be due to the unphysical assumption of a point-dipole as a source of the magnetic field of the nucleus. As a physical description of the magnetization distribution of the nucleus is not available due to a lack of experimental data, the magnetization distribution is assumed to be the same as the charge distribution of the nucleus, see Section *Finite Nucleus Model*. This is unphysical as the magnetization is caused by the one unpaired nucleon in the nucleus whereas the charge distribution is generated by the protons in the nucleus. So, physically, the magnetization should occupy a larger volume in space than the charge. This might also be the reason why the resulting finite-nucleus model is insufficient to remedy the divergences in the fπFW hyperfine-coupling tensor. Consequently, the hyperfine-coupling tensor is only implemented in the version resulting from the fpFW DKH Hamiltonian. In the case of the g-tensor both versions are implemented and accessible via the keyword

```
%rel fpFWtrafo true/false end
```

By default, this keyword is set to true. A detailed form of the property operators used for the g-tensor and hyperfine-tensors can be found in Ref. [745].

7.23.5 Finite Nucleus Model

Composite particles like nuclei have, as opposed to elementary particles, a certain spatial extent. While the point-charge approximation for nuclei is in general very good in nonrelativistic calculations, in relativistic calculations it might lead to non-negligible errors. A finite-nucleus model is available for all calculations in the ORCA program package. It is accessible from the `%rel` block via

```
%rel FiniteNuc true/false end
```

By default, this keyword is set to `false`. If the keyword is set to `true`, finite-nucleus effects are considered in the following integrals:

- nucleus potential V
- DKH-integral $V^{(p)}$
- one-electron spin-orbit integrals SOC (also in one-electron part of SOMF)
- electric-field gradient EFG (and thus, as a consequence in the Fermi-contact and spin-dipole terms of the HFC tensor)
- nucleus-orbit integral NUC
- angular-momentum integral l

The finite-nucleus model implemented in ORCA is the Gaussian nucleus model of Ref. [871].

7.23.6 Exact Two-Component Theory (X2C)

The X2C implementation in ORCA closely follows that of Franzke, Weigend, and their coworkers, described in the references: [658] (energy), [274] (gradient), [276] (EPR hyperfine coupling), [273] (NMR spin–spin coupling), [272] (NMR shielding). These are also consistent with the work of Gauss and coworkers in refs: [165] (gradient, electric properties), [166] (Hessian), [167] (NMR shielding). However, despite the name, only a scalar relativistic (spin-free) version is available at present, resulting effectively in a one-component method (more aptly called "SF-X2C-1e"), very similar to the DKH and ZORA approaches described above. The main difference to the latter two is that the decoupling of the one-electron Dirac Hamiltonian is exact, rather than approximate. SF-X2C-1e is implemented in ORCA for energies, gradients, and various properties (see below) with both a point- and finite nucleus model.

We briefly describe the working equations here, using the notation of the aforementioned references, which differs somewhat from that in the previous sections. The one-electron Dirac equation is solved directly:

$$\mathbb{DC} = \mathbb{MCE}, \quad \mathbb{C} = \begin{pmatrix} \mathbf{C}_+^{\mathrm{L}} & \mathbf{C}_-^{\mathrm{L}} \\ \mathbf{C}_+^{\mathrm{S}} & \mathbf{C}_-^{\mathrm{S}} \end{pmatrix} \tag{7.184}$$

to obtain the unitary transformation matrix:

$$\begin{aligned} \mathbb{U} &= \begin{pmatrix} \mathbf{U}_{++} & \mathbf{U}_{+-} \\ \mathbf{U}_{-+} & \mathbf{U}_{--} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & -\mathbf{X}^\dagger \\ \mathbf{X} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}' \end{pmatrix} \\ \mathbf{X} &= \mathbf{C}_+^{\mathrm{S}} \left(\mathbf{C}_+^{\mathrm{L}}\right)^{-1} \\ \mathbf{R} &= \mathbf{S}^{-\frac{1}{2}} \left(\mathbf{S}^{-\frac{1}{2}} \tilde{\mathbf{S}} \mathbf{S}^{-\frac{1}{2}}\right)^{-\frac{1}{2}} \mathbf{S}^{\frac{1}{2}} \\ \tilde{\mathbf{S}} &= \mathbf{S} + \frac{1}{2c^2} \mathbf{X}^\dagger \mathbf{T} \mathbf{X} \end{aligned}$$

which exactly block-diagonalizes the Hamiltonian:

$$\mathbb{U}^\dagger \mathbb{D} \mathbb{U} = \begin{pmatrix} \mathbf{h}^+ & \mathbf{0} \\ \mathbf{0} & \mathbf{h}^- \end{pmatrix} \tag{7.185}$$

$$\mathbf{h}^+ = \mathbf{R}^\dagger \mathbf{L} \mathbf{R} \tag{7.186}$$

$$\mathbf{L} = \mathbf{V} + \mathbf{T}\mathbf{X} + \mathbf{X}^\dagger\mathbf{T} + \mathbf{X}^\dagger\left(\frac{1}{4c^2}\mathbf{W} - \mathbf{T}\right)\mathbf{X} \tag{7.187}$$

where $\mathbf{V}$, $\mathbf{T}$, $\mathbf{S}$, and $\mathbf{W}$ are the potential, kinetic, overlap, and relativistic potential ($\hat{p}\,\hat{V}\,\hat{p}$) integral matrices. $\mathbf{h}^+$ is thus the matrix form of the relativistically-corrected one-electron Hamiltonian used for the rest of the calculation.

Note that the potential operator $\hat{V}$ used in the above equations only includes the electron–nuclear Coulomb interaction. External point charges may optionally be included via `%rel ModelPot[4]=1` (default 0). This option affects energy and NMR shielding calculations but it is presently not available for gradients or Hessians.

DLU approximation

The diagonal local approximation to the unitary transformation matrix (DLU), as introduced by Peng and Reiher,[659] reduces the computational cost of the X2C transformation by approximating $\mathbb{U}$ (i.e., $\mathbf{R}$ and $\mathbf{X}$) as an atomic-block-diagonal matrix.

$$\mathbf{R} \approx \bigoplus_A \mathbf{R}_A \quad \mathbf{X} \approx \bigoplus_A \mathbf{X}_A$$

where $\bigoplus_A$ denotes a direct sum of atomic diagonal blocks. Eqs (7.184)–(7.187) can then be solved independently for diagonal atomic blocks $\mathbf{h}^+_{AA}$. Off-diagonal blocks $\mathbf{h}^+_{AB}$ are obtained as:

$$\mathbf{h}^+_{AB} = \mathbf{R}^\dagger_A \mathbf{L}_{AB} \mathbf{R}_B$$
$$\mathbf{L}_{AB} = \mathbf{V}_{AB} + \mathbf{T}_{AB}\mathbf{X}_B + \mathbf{X}^\dagger_A\mathbf{T}_{AB} + \mathbf{X}^\dagger_A\left(\frac{1}{4c^2}\mathbf{W}_{AB} - \mathbf{T}_{AB}\right)\mathbf{X}_B$$

It is also possible to approximate light atoms a, b (below a certain atomic number) as non-relativistic:

$$\mathbf{X}_a = \mathbf{R}_a = \mathbf{I}$$
$$\mathbf{W}_{ab} = \mathbf{W}_{aB} = \mathbf{0}$$
$$\mathbf{h}^+_{ab} = \mathbf{V}_{ab} + \mathbf{T}_{ab}$$
$$\mathbf{h}^+_{aB} = \left(\mathbf{V}_{aB} + \mathbf{T}_{aB}\mathbf{X}_B\right)\mathbf{R}_B$$

Unlike the one center approximation (which is also available for X2C), all nuclei are included in the potential operator $\hat{V}$. Use of the DLU approximation is controlled via:

```
%rel
  DLU         false # default - not used
               true # turn DLU on
  LightAtomThresh 0 # (default) highest atomic number treated as non-relativistic
end
```

X2C derivatives and properties

As discussed in section *Picture-Change Effects*, when computing properties with relativistic methods, derivatives need to be taken of the correct Hamiltonian, namely $\mathbf{h}^+$ (eq (7.186)) in the X2C case. These include contributions due to the derivatives of $\mathbf{R}$ and $\mathbf{X}$, which are often small. Therefore, it is possible to neglect them and save some computational time via the following option:

```
%rel
  PictureChange 2 # compute the full relativistic Hamiltonian derivative
                1 # neglect derivatives of R and X
                0 # (default) use the non-relativistic operator
end
```

The same setting is applied to all properties for which the X2C correction is implemented. Currently, these are: geometric gradients, electric dipoles, quadrupoles, and polarizabilities, electric field gradients, EPR hyperfine couplings, and NMR shieldings and spin–spin couplings. For the Hessian, the X2C correction is implemented in a

semi-numeric fashion. The DLU approximation is applied throughout, if requested, and reduces the computational effort dramatically. Note that for magnetic properties, the restricted magnetic balance (RMB) for the small component basis functions is used whenever GIAOs are requested (e.g. NMR shielding), while the restricted kinetic balance (RKB) is used otherwise (e.g. NMR coupling). The spin–orbit coupling integrals used for various properties only include a relativistic correction to the one-electron term, as is the case for DKH.

For second derivative properties whose number is proportional to the number of atoms, e.g. the DSO term of NMR couplings, first derivatives of various intermediate quantities required for the full X2C second derivative are stored on disk. The storage requirements can be reduced via the `StorageLevel` keyword and the missing intermediates will be recomputed on-the-fly, which of course increases the computation time.

```
%rel
  StorageLevel 0 # do not store anything - not always available
               1 # store integral derivatives
               2 # 1 + derivatives of R and X (default, minimum for DLU)
               3 # 2 + derivatives of L
               4 # 3 + derivatives of S-tilde
               5 # 4 + derivatives of D and M
end
```

7.23.7 Basis Sets in Relativistic Calculations

For relativistic calculations, special basis sets have been designed, both as DKH and ZORA recontractions of the non-relativistic Ahlrichs basis sets (in their all-electron versions) for elements up to Kr, and as purpose-built segmented all-electron relativistically contracted (SARC) basis sets for elements beyond Kr [62, 640, 641, 642, 643, 728]. Their names are "ZORA-" or "DKH-" followed by the conventional basis set name. For X2C calculations, the "x2c-XVPall" basis sets and their variants are available.[275, 689] See section *Choice of Basis Set* for a complete list of basis sets.

7.24 Approximate Full CI Calculations in Subspace: ICE-CI

7.24.1 Introduction

In many circumstances, one would like to generate a wavefunction that is as close as possible to the full-CI result, but Full CI itself is out of the question for computational reasons. Situations in which that may be desirable include a) one wants to generate highly accurate energies for small molecules or b) one wants to sort out a number of low-lying states or c) one wants to run CASSCF calculations with larger active spaces than the about fourteen orbitals that have been the state of the art for a long time.

ORCA features a method that has been termed Iterative-Configuration Expansion Configuration Interaction (ICE-CI).[171, 172] It is based on much older ideas brought forward by Jean-Paul Malrieu and his co-workers in the framework of the CIPSI (an abbreviation for a method with a rather bulky name *Configuration Interaction by Perturbation with multiconfigurational zeroth-order wave functions Selected by Iterative process*) in the early 1970s.

The goal of the ICE-CI is to provide compact wavefunction(s) (e.g. one or several states) close to the full-CI limit at a small fraction of the computational cost. However, ICE-CI itself is not designed to deal with hundreds of atoms or thousands of basis functions. Thus, unlike, say DLPNO-CCSD(T) which is a high accuracy method for treating large sytems, ICE-CI is either a highly robust high accuracy method for very small systems or a "building block" for large systems. By itself it can treat a few dozen electrons and orbitals – e.g. *much* more than full CI – but it cannot do wonders. Its scope is similar to the density matrix renormalization group (DMRG) or Quantum Monte Carlo Full CI (QMCFCI) procedures.

ICE-CI should be viewed as a multireference approach. It is self-adaptive and robust, even in the presence of near or perfect degeneracies. It yields orthogonal states (when applied to several states) and spin eigenfunctions. It also yields a density and a spin density.

7.24.2 The ICE-CI and CIPSI Algorithms

The general idea of ICE-CI is straightforward: Consider a many-particle state that has at least a sizeable contribution from a given configuration $\mathbf{n}^0$ (this is a set of occupation numbers for the active orbitals that are $n_p^0 = 0, 1$ *or* 2 (p = any active orbital). By nature of the non-relativisitic Hamiltonian only configurations that differ by at most two orbital occupations from $\mathbf{n}^0$ will interact with it. We can use perturbation theory to select the subset of singles and doubles that interact most strongly with $\mathbf{n}^0$ and then solve the variational problem. We can then analyze the CI vector for configurations that make a dominant contribution to the ground state. Say, we single out the configurations with $C_I^2 > T_{\text{gen}}$. This defines the "generator" set of configurations. The other configurations are called "variational" configurations. They are treated to infinite order by the variational principle, but are not important enough to bring in their single and double excitations. In the next iteration, we perform singles and doubles relative to these general configurations and select according to their interaction with the dominant part of the previous CI vector (truncated to the generators). This procedure can be repeated until no new important configurations are found and the total energy converges (See Fig. 7.8).

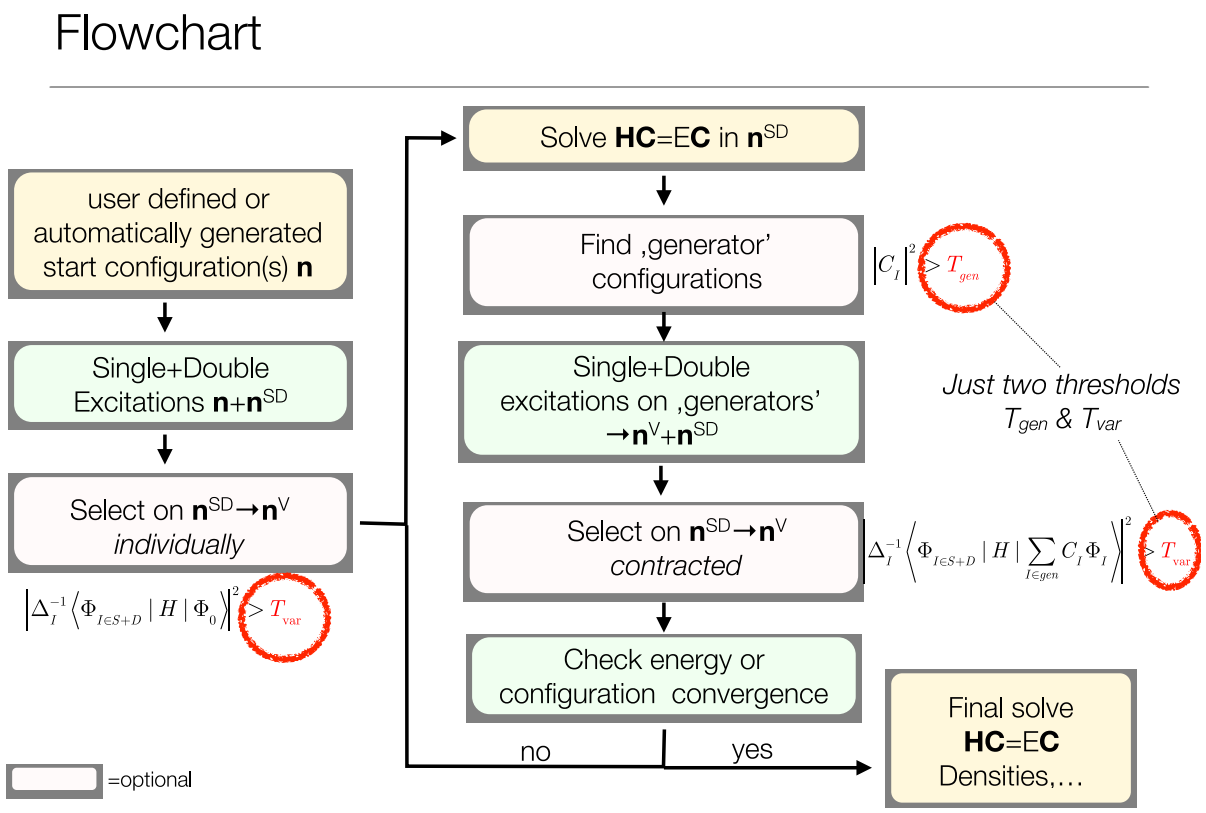


Fig. 7.8: Flowchart of the ICE-CI procedure.

The described procedure is very similar to Malrieu's three level CIPSI procedure. One major technical difference, is that ICE is centered around *configurations* and *configuration state functions* rather than *determinants*. A configuration is a set of occupation numbers 0, 1 or 2 that describes how the electrons are distributed among the available spatial orbitals. A *configuration state function* (CSF) is created by coupling the unpaired spins in a given configuration to a given total spin S. In general there are several, if not many ways to construct a linearly independent set of CSFs. CSFs on the other hand can be expanded in terms of Slater determinants, but there are more Slater determinants to a given configuration than CSFs. For example for a CAS(14,18) calculation one has about 10^9 determinants, but only about $3\text{x}10^8$ CSFs and $3\text{x}10^7$ configurations. In the configuration based ICE (CFG-ICE) all logic happens at the level of configurations. That is, it is the relationship between two configurations that determines whether and if yes, by which integrals the CSFs or determinants of two given configurations interact. Since the configuration space is so much more compact than the determinant space substantial computational benefit can

be realized by organizing the calculation around the concept of a configuration. In general, in CFG-ICE all CSFs that belong to a given configuration are included and all selection quantities are summed over all CSFs of a given configuration before it is decided whether this CSFs is included or not. In the configuration state functions based ICE (CSF-ICE) the logic of generation and selection occurs at the level of individual CSFs and therefore we get rid of the requirement to carry around all the CSFs for a given configuration. This provides substantial gains in the case of molecules containing a large number of transition metal atoms, where each atom contains a high-spin center. In such cases only a few CSFs of a the dominant CFG play a dominant role and other show negligible contribution to the wavefunction. Finally, in some cases the original determinant based CIPSI procedure could be preferred. Such cases can be handled by the determinant based ICE termed DET-ICE. The three variants of ICE therefore cover all the possible types of multi-reference systems that one encounters in quantum chemistry.

It should be noted that although the procedure contains a perturbative element, the final energy is strongly dominated by the variational energy and hence, for all intents and purposes, the ICE-CI procedure is variational (but not rigorously size consistent – size consistency errors are on the same order of magnitude as the error in absolute energy).

7.24.3 A Simple Example Calculation

Let us look at a simple calculation on the water molecule:

```
#
# Check the ICECI implementation
#

! SV

%ice  nel 10           # number of active electrons
      norb 13          # number of active orbitals
      nroots 1         # number of requested roots
      integrals exact  # exact 4-index transformation
                       # can be set to RI to avoid bottlenecks

      icetype   CFGs   # The configuration based ICE-CI
                CSFs   # The CSF based ICE-CI
                DETs   # The determinnat based ICE-CI

      Tgen 1e-04       # value for Tgen. Default is 1e-4
      Tvar 1e-11       # value for Tvar. Default is 1e-11 (1e-7*Tgen)

      etol 1e-06       # energy convergence tolerance
end

* int 0 1
O 0 0 0 0.0   0.000 0.000
H 1 0 0 1.0   0.000 0.000
H 1 2 0 1.0 104.060 0.000
*
```

Let us look at the output:

```
-----------------------------------------------------------------------------
                ORCA Iterative Configuration Expansion
             - a configuration driven CIPSI type approach -
-----------------------------------------------------------------------------

(some startup information)

Making an initial 'Aufbau' configuration      ... done
Performing S+D excitations from     1   configs ... done (  0.0 sec) NCFG=581
Performing perturbative selection             ... done (  0.0 sec)
   # of configurations before selection       ... 581
```

(continues on next page)

(continued from previous page)

```
   # of configurations after selection        ... 191
   'rest' energy (probably not very physical) ... -3.736391e-10

                    ******************************
                    *  ICECI MACROITERATION   1  *
                    ******************************

   # of active configurations =    191
   Now calling CI solver (269 CSFs)
(...)
CI SOLUTION :
STATE   0 MULT= 1: E=    -76.0463127108 Eh W=  1.0000 DE= 0.000 eV      0.0 cm**-1
      0.95752  : 2222200000000
   Selecting new configurations         ... done (  0.0 sec)
   # of selected configurations          ...    191
   # of generator configurations         ...     69
   Performing single and double excitations relative to generators ... done (  0.0 sec)
   # of configurations after S+D         ...  13174
   Selecting from the generated configurations  ... done (  0.1 sec)
   # of configurations after Selection  ...   3827
    Root   0:    -76.046312711       -0.000000063     -76.046312773
(...)
                    ******************************
                    *  ICECI MACROITERATION   3  *
                    ******************************

   # of active configurations =   3866
   Now calling CI solver (9606 CSFs)

CI SOLUTION :
STATE   0 MULT= 1: E=    -76.0539542296 Eh W=  1.0000 DE= 0.000 eV      0.0 cm**-1
      0.95097  : 2222200000000
(...)

             ********* ICECI IS CONVERGED *********
(one final CI)

           ********************************************
           **   ICECI Problem solved in    2.6 sec   **
           ********************************************


FINAL CIPSI ENERGIES
  Final CIPSI Energy Root   0:    -76.053954291 EH
```

From the output the individual steps in the calculation are readily appreciated. The program keeps cycling between variational solution of the CI problem, generation of new configurations and perturbative selection until convergence of the energy is achieved. Normally, this occurs rapidly and rarely requires more than five iterations. The result will be close to the Full CI result.

Let us look at a H_2O/cc-pVDZ calculation in a bit more detail (See Fig. 7.9). The calculation starts out with a single Hartree-Fock configuration. The first iteration of ICE-CI creates the singles and doubles and altogether 544 configurations are selected. These singles and doubles bring in about half of the correlation energy. Already the second iteration, which leads to 73000 selected CSFs provides a result close to the full CI. At this point up to quadruple excitations from the Hartree-Fock reference have been included. It is well known that such quadruple excitations are important for the correct behavior of the CI procedure (near size consistency will come from the part of the quadruple excitations that are products of doubles). However, only a very small fraction of quadruples will be necessary for achieving the desired accuracy. In the first iteration the procedure is already converged and provides 99.8% of the correlation energy, using 0.5% of the CSFs in the full CI space and at less than 0.2% the calculation time required for solving the full CI problem. Hence, it is clear that near exact results can be obtained while realizing spectacular savings.

H_2O, cc-pVDZ, 1s frozen, 8 electrons, 23 orbitals $T_{Gen}=10^{-4}$, $T_{var}=10^{-11}$

	Energy	N(CFG)	N(CSF)		time(sec)	
Hartree-Fock	-76.0260966	1	1		2	
	1977 CFG's generated, 544 selected				<0.1	
Iter=1	-76.2292278	544	805	S+D	0.3	
	98717 CFG's generated, 22471 selected				1.1	
Iter=2	-76.2412460	22471	73296	T+Q	17.1	
	10677 CFG's generated, 23241 selected				2.2	
Iter=3	-76.2412769	23241	75705	P+H	20.1	
	10677 CFG's generated, 23241 selected				3.3	
Iter=4	-76.2412769	23241	75705	S+O	20.1	
					83.5	(1 final CI)
	99.75% of EC	0.57%	0.4%		0.18% of FCI	
Full-CI	-76.2418240	4065963	18818646		44529.0	

Fig. 7.9: An ICE-CI calculation on the water molecule in the cc-pVDZ basis (1s frozen)

7.24.4 Accuracy

The accuracy of the procedure is controlled by two parameters T_{gen} and T_{var} Since we have found that $T_{var} = 10^{-7}$ T_{gen} always provides converged results, this choice is the default. However, T_{var} can be set manually. It can be reduced considerably in order to speed up the calculations at the expense of some accuracy. Our default values are $T_{gen} = 10^{-4}$ and $T_{var} = 10^{-11}$. This provides results within about 1 mEh of the full CI results (roughly speaking, a bit better than CCSDT for genuine closed-shell systems).

During the development of ICE-CI systematic test calculations have been performed using a reference set of 21 full CI energies on small molecules. The convergence pattern of the mean absolute error is shown in Fig. 7.10. It is evident from the figure that the convergence of ICE-CI towards the FCI result is very smooth and that high accuracy can be obtained. In fact, the default settings lead to an accuracy of <1 mEh deviation to the full-CI result. μEh accuracy can be achieved by further tightening. The achieved accuracy relative to accurate coupled-cluster results shows that the accuracy of even CCSDTQ can be surpassed by ICE-CI. The achievable accuracy is only limited by the value of T_{gen} and much less so by the value of T_{var}. Hence, it is advisable to use a value for T_{var} that is essentially converged and control the accuracy of the procedure by T_{gen}.

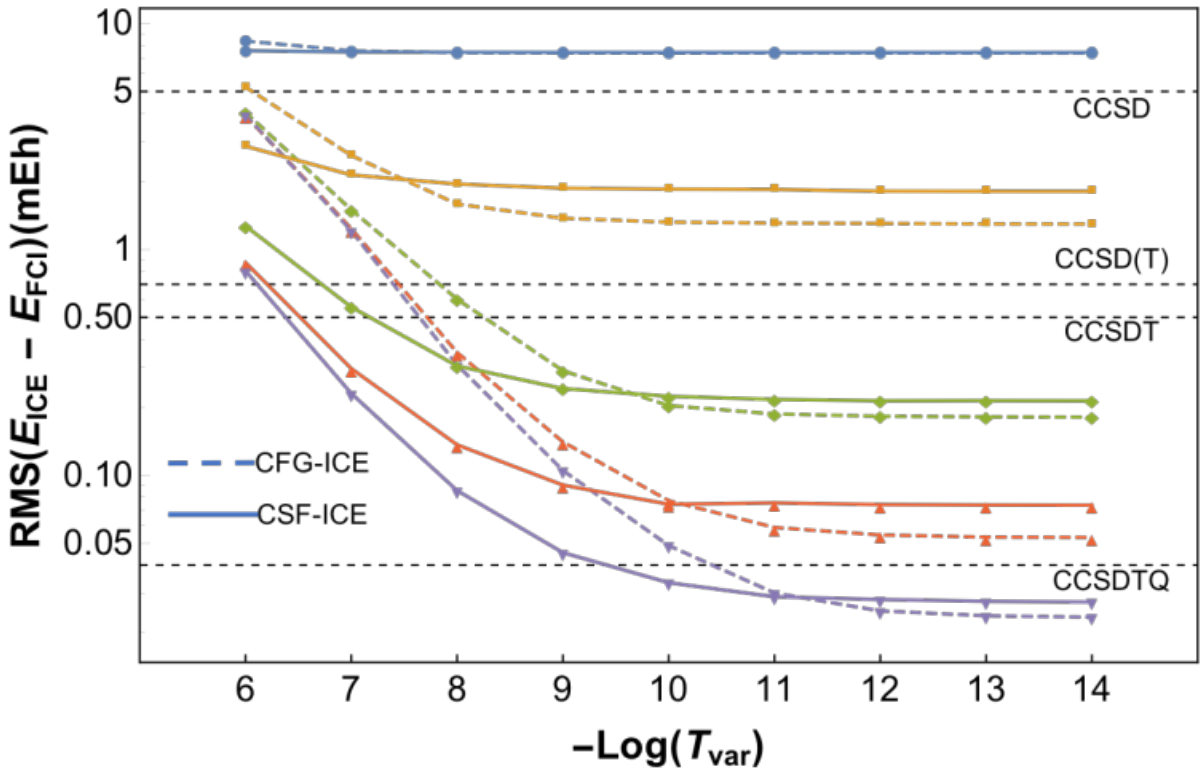


Fig. 7.10: Convergence of the ICE-CI procedure towards the full CI results for a test set of 21 full CI energy. Shown is the RMS error relative to the Full CI results. The corresponding errors for various coupled-cluster variants is shown by broken horizontal lines.

7.24.5 Scaling behavior

ICE-CI will break the factorial scaling of the full CI problem and scale polynomially. The actual order of the polynomial scaling is system dependent and accuracy dependent. In order to provide some impression, consider some calculations on linear polyene chains.

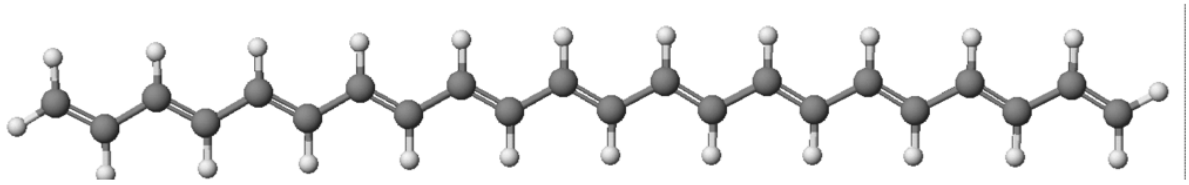


Fig. 7.11: Polyene chains used for scaling calculations.

The results are displayed in the Fig. 7.12. It is evident from Fig. 7.12 that ICE-CI breaks the factorial scaling of the full CI problem. In fact, for a thresholds of T_{gen} $=10^{-4}$, 10^{-3} and 10^{-2} the observed scalings are approximately $O(N^8)$, $O(N^7)$ and $O(N^6)$ respectively. These numbers will obviously be very system dependent but should serve as a rough guide. The calculations become quickly much more expensive if T_{gen} is tightened. A rule of thumb is that each order of magnitude tightening of T_{gen} increases the computation time by a factor of 10. The above calculations have been performed on a simple desktop computer and it was already possible to solve a CAS(30,30) problem in less than one day of elapsed time using the default thresholds. Large active spaces will require either loosening of the tresholds or large, more powerful machines.

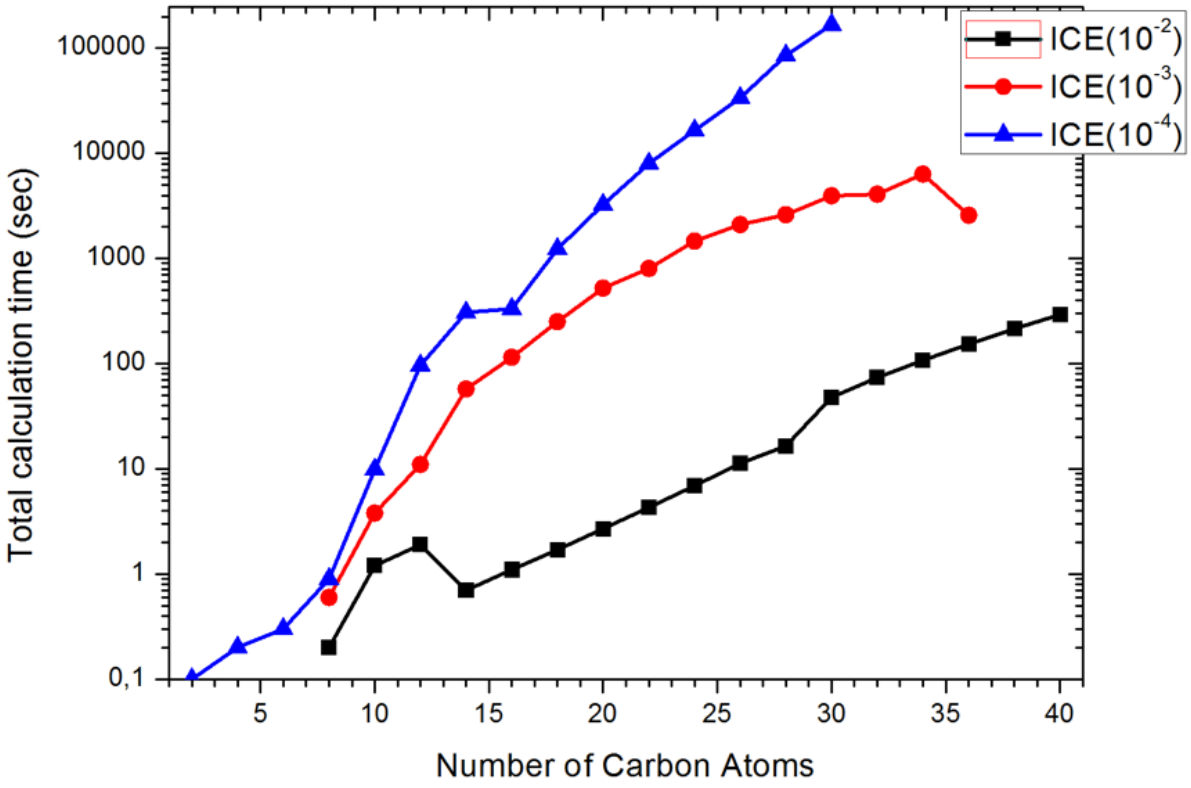


Fig. 7.12: Scaling behavior of ICE-CI for linear polyene chains (Full π-electron active space) as a functions of system size for different generator thresholds.

7.24.6 Accuracy of the Wavefunction

The accuracy of the many particle wavefunction is not straightforward to check. A reasonable measure, however, is how well it converges towards the exact result for one-electron expectation values. Since every expectation value can be written in terms of natural orbitals of the one-particle density as:

$$\left\langle \hat{O} \right\rangle = \left\langle \Psi \left| \sum\nolimits_o \hat{o}(\mathbf{x}_i) \right| \Psi \right\rangle = \sum_{pq} D_{pq} \left\langle \psi_p | \hat{o} | \psi_q \right\rangle = \sum_p n_p \left\langle \tilde{\psi}_p | \hat{o} | \tilde{\psi}_p \right\rangle$$

where $\hat{o}(\mathbf{x}_i)$ is an arbitrary one-particle operator, D_{pq} is the density matrix of the ICE-CI wavefunction, $\tilde{\psi}_p$ are the natural orbitals of the ICE-CI wavefunction and n_p are their occupations numbers. It is reasonable to take the deviation of the natural orbital occupation numbers as a measure for wavefunction convergence.

For example, we treat the H_2O/cc-pVDZ problem again. From the results in Fig. 7.13 it becomes evident that the ICE-CI wavefunction is fairly accurate. At the default threshold the occupation numbers agree to within 10^{-3} with the full CI reference numbers, which means that expectation values will be of similar accuracy. Interestingly, the largest errors occur in the region of the HOMO-LUMO gap, where apparently all approximate wavefunction approaches tend to depopulate the high lying orbitals too much and put too much electron density in the low lying empty orbitals. From comparison, it is seen, that the CCSD natural occupation numbers for this problem are significantly less accurate. Hence, this is evidence that the ICE-CI wavefunction is properly converging to the right result.

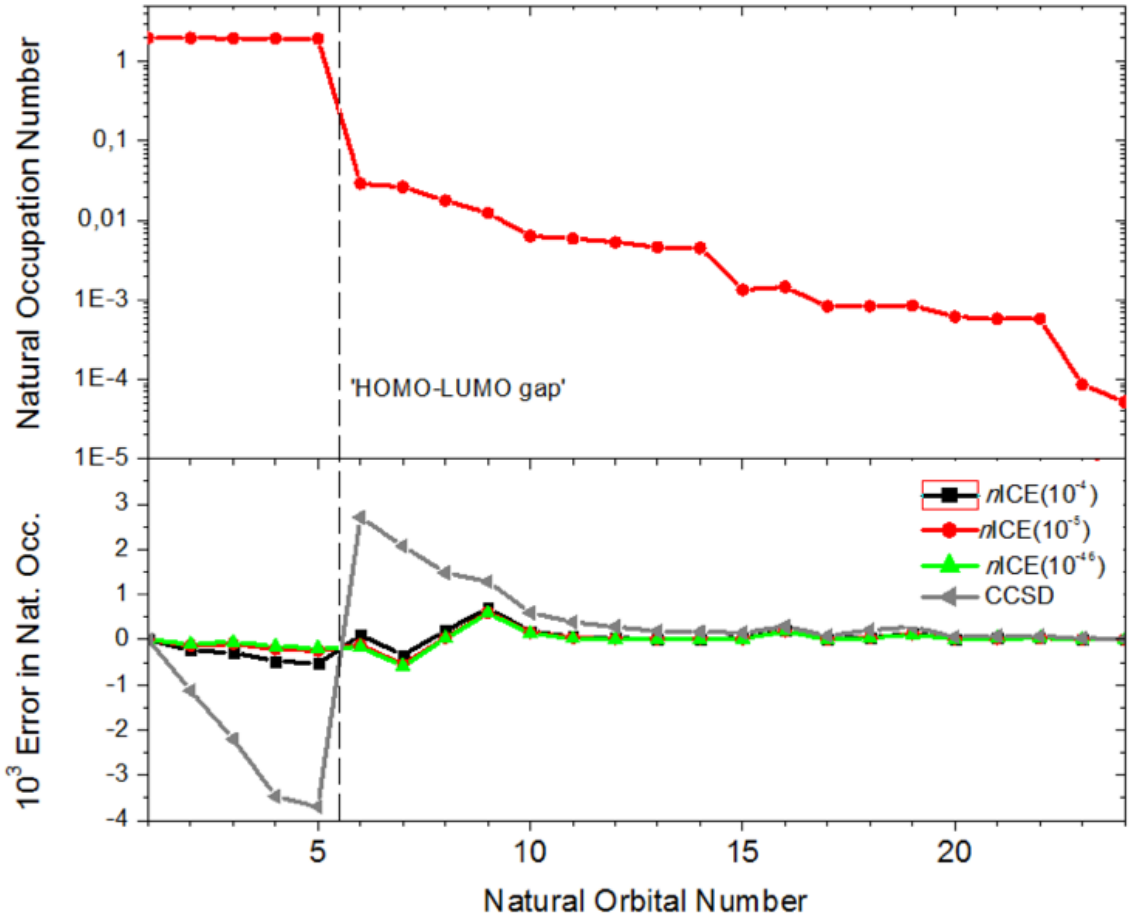


Fig. 7.13: Convergence of the ICE-CI natural orbital occupation numbers. The upper panel is showing the Full CI occupation numbers, the lower panel the deviation of the ICE-CI values from these exact values. For comparison, the CCSD natural orbital occupation numbers are also provided.

7.24.7 Potential Energy Surfaces

You can use ICE-CI to scan entire potential energy surfaces. In general, the non-parallelity error along a potential energy surface is very small. Thus, ICE-CI yields consistent quality throughout the surface.

For example, let us look at the potential energy surface of the N_2 molecule (Fig. 7.14) – a common test case for quantum chemical methods. There are not too many methods that would disscociate the triple bond of N_2 correctly – ICE-CI is one of them. The potential energy surface is entirely smooth and also correctly behaves in the dissociation limit. Near the minimum it is very close to high-level coupled-cluster methods that, however, all fail badly as the bond is stretched.

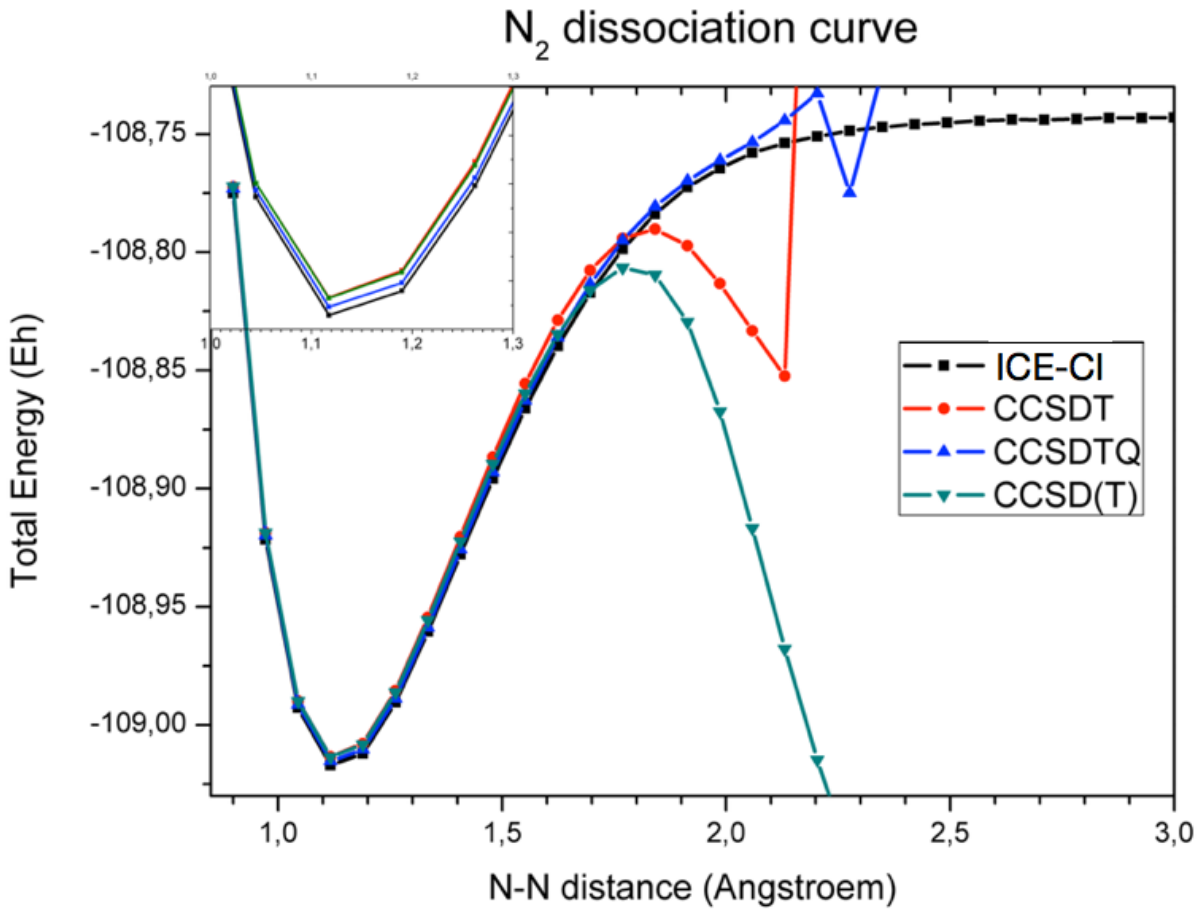


Fig. 7.14: Potential energy surface of the N_2 molecule in the SV basis. For comparison higher level coupled-cluster results are also shown.

It is interesting to observe the variations of the ICE-CI wavefunction along the dissociation potential energy surface. As an example, we look at the dissociation curve of H_2O where both O-H bonds are simultaneously stretched (Fig. 7.15). It is seen that the ICE-CI method is extremely parallel to the full CI curve at all distances. Hence, the description of the bond remains consistent, even when Hartree-Fock becomes a bad approximation. The agreement is particularly good if MP2 natural orbitals are used in the ICE-CI procedure. With the default value of $T_{gen} = 10^{-4}$ and MP2 natural orbitals the error is consistently below 0.2 mEh. For tighter thresholds, the error is below 0.05 mEh. By contrast, the CCSD(T) method shows relatively large deviations from the full CI results and also behaves very non-parallel as a function of O-H distance.

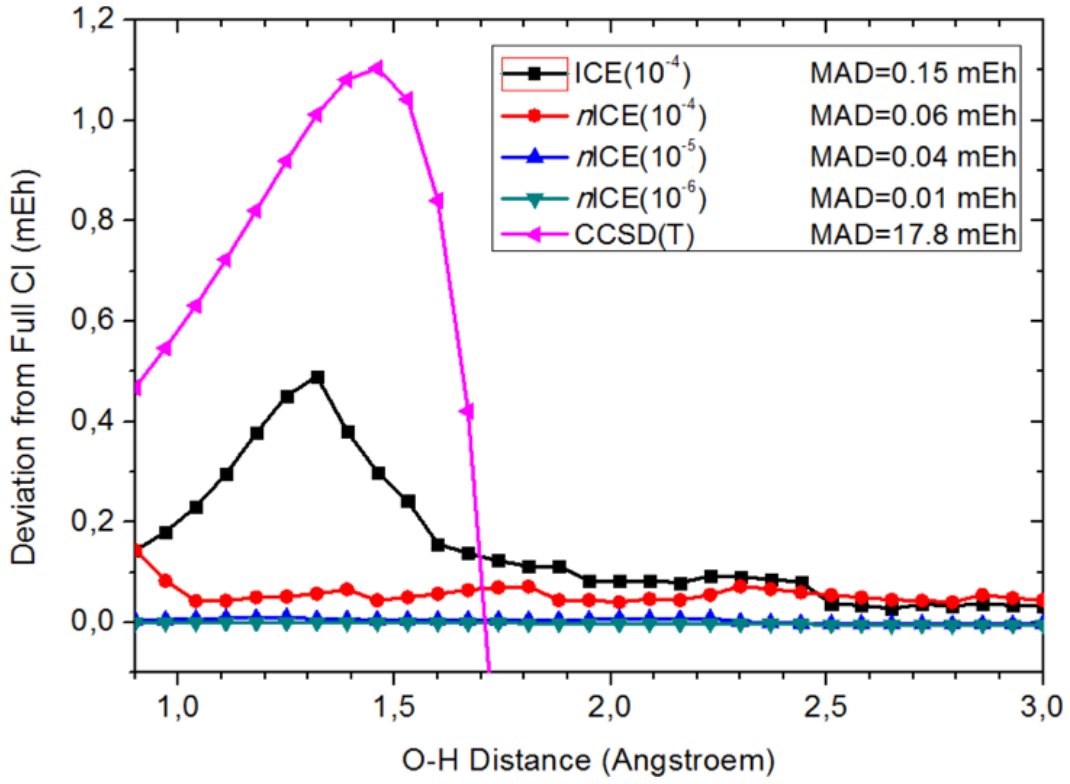


Fig. 7.15: Non-parallelity error of ICE-CI for the H_2O molecule in the SV basis. Shown is the deviation from the full CI value as a function of O-H distance (both bonds stretched). For comparison, the CCSD(T) curve is also shown

It is instructive to analyze the ICE-CI wavefunction along the dissociation pathway (Fig. 7.16). It becomes apparent that the wavefunctions stays compact along the entire surface, even in the dissociation limit, where the weight of the Hartree-Fock wavefunction drops to less than 25%. Even in this drastic limit, the ICE-CI wavefunction consists of only about 60000 CSFs, which is very similar to the size of the wavefunction at equilibrium geometry. As the wavefunction becomes more multiconfigurational, the number of generator configurations goes slightly up from the equilibrium value of 77 to a maximum of 118 and finally 112 at dissociation. It is also interesting to note that along the entire dissociation pathway no configuration with more than 8 open shells is generated, which means that no more than quadruple excitations are contained in the ICE-CI wavefucntion. The number of iterations required in the ICE-CI procedure also stays constant along the surface at 4 iterations, which impressively shows that a dominant configuration is not necessary for a successful ICE-CI calculation.

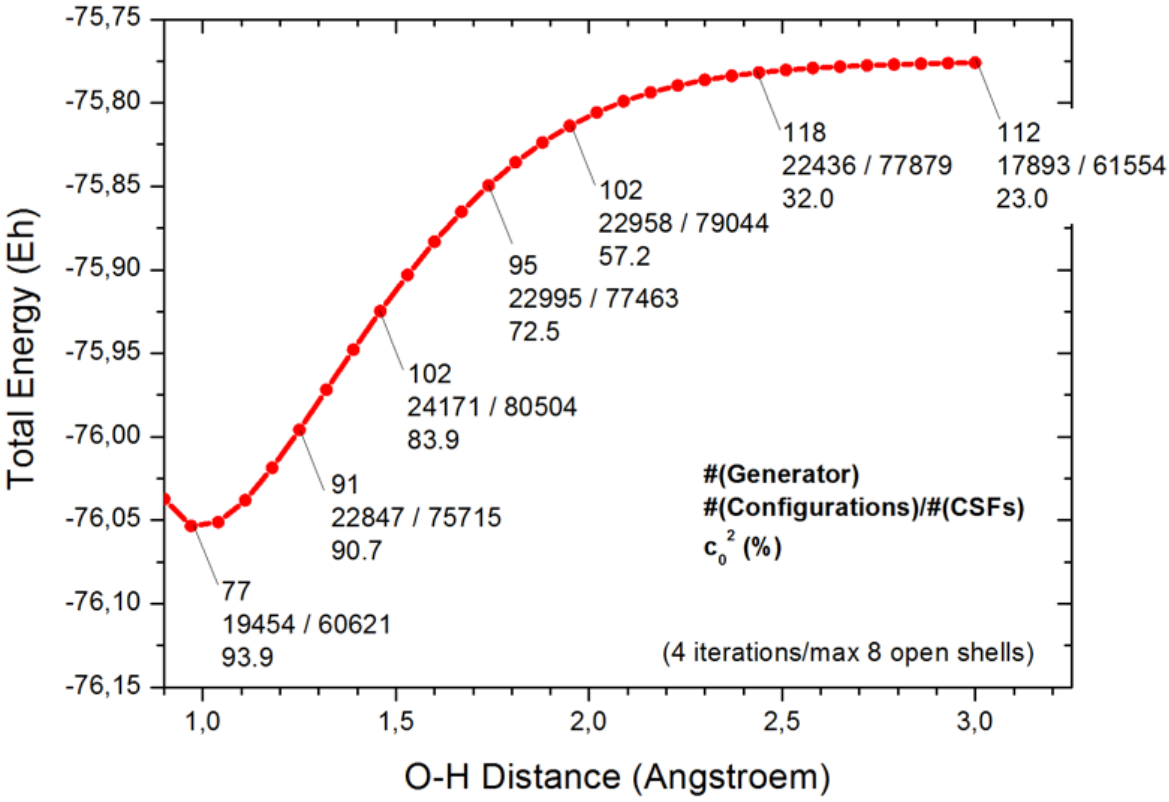


Fig. 7.16: Analysis of the ICE-CI wavefunction along the O-H dissociation pathway.

7.24.8 Excited States

ICE-CI can be used to obtain some insight into excited states starting from no knowledge at all. Of course, the best was to start an excited state calculation is to have some idea which configurations are important for the low-lying states of the system. If this is not the case, an automated procedure is used. The program will first generate an "Aufbau" configuration using the orbitals that are provided on input. Starting from this Aufbau configuration, single excitations at the configuration level are performed an the Hamiltonian is diagonalized for the required number of roots. These roots are then analyzed for the leading configurations and the regular ICE-CI procedure is started from those configurations. For example, look at a calculation on the CN radical. In this case, we know the relevant orbitals and leading configurations for the lowest four roots (a doublet Σ ground state, a doublet Π excited state and a doublet Σ excited state) and hence can provide them in the input file as shown below.

```
#
! cc-pVDZ VeryTightSCF

%maxcore 4096

%casscf nel 7
        norb 4
        nroots 4
        mult 2
        end

%ice   nel 9
        norb 26
        nroots 4
        cimode 3
        tvar 1e-11
        tgen 1e-4
```

(continues on next page)

(continued from previous page)

```
        refs { 2 2 2 2 1 }
             { 2 2 2 1 2 }
             { 2 2 1 2 2 }
             { 2 1 2 2 2 }
             end
        end

* xyz 0 2
C 0 0 0
N 0 0 1.07
*
```

The result is shown below. The excitation energies are reasonable but not highly accurate due to the limitations of the basis set (experimentally the doublet Π state is at 1.32 eV and the doublet Σ state at 3.22 eV). There is a very slight symmetry breaking In the doublet Π state that arises from the selection procedure. It should be noted that the state averaged CASSCF excitation energies are 0.25 eV and 3.18 eV.

```
STATE   0 MULT= 2: E=    -92.4542949092 Eh W=  0.2500 DE= 0.000 eV      0.0 cm**-1
      0.67408  : 22221000000000000000000000
STATE   1 MULT= 2: E=    -92.3777338429 Eh W=  0.2500 DE= 2.083 eV  16803.2 cm**-1
      0.65997  : 22212000000000000000000000
      0.12092  : 22122000000000000000000000
      0.11545  : 22221000000000000000000000
STATE   2 MULT= 2: E=    -92.3776916150 Eh W=  0.2500 DE= 2.084 eV  16812.5 cm**-1
      0.69860  : 21222000000000000000000000
      0.16565  : 22122000000000000000000000
STATE   3 MULT= 2: E=    -92.3412510872 Eh W=  0.2500 DE= 3.076 eV  24810.3 cm**-1
      0.51251  : 22122000000000000000000000
      0.16825  : 22212000000000000000000000
      0.11151  : 21222000000000000000000000
```

Below, it is described how to do ICE-CI calculations on excited states if the dominant configurations are not known.

7.24.9 Tips and Tricks

ICE-CI can be used very fruitfully together with, say, MP2 natural orbitals. This usually results in results that are closer to full CI results and at the same lead to more compact wavefunctions (it may be called nICE). The use of MP2 natural orbitals is requested by choosing `UseMP2nat true` inside the %ice block. Alternatively, improved virtual orbitals can be used (requested by `UseIVOs true`). A comparison is shown in Scheme *Comparison of MP2 natural orbitals and improved virtual orbitals for the ICE-CI procedure (H2O molecule, cc-pVDZ basis, equilibrium geometry)*. It is evident that the calculations based on the MP2 natural orbitals show an error relative to full CI that is almost a factor of two smaller than the corresponding result with canonical orbitals while at the same time the wavefunction is more compact by more than 30%. Hence, the use of MP2 natural orbitals appears to be a very good idea in conjunction with the ICE-CI procedure. This also holds when MP2 itself is a bad approximation (for example in the dissociation limit of the H2O molecule as shown above). On the other hand, the IVOs behave very similar to canonical orbitals and hence, seem to offer fewer advantages.

	#(Iterations)	#(CSF)	Energy/Eh	Error/μEh
Canonical orbitals:	4	85220	-76.241 306 7	518
Improved virtual orbitals:	4	85064	-76.241 316 6	508
MP2 natural orbitals:	4	59447 *(0.3% of FCI!)*	-76.241 434 0 *(99.82% of E_{cl}!)*	390
Full CI result		18818646	-76.241 824 0	

Fig. 7.17: Comparison of MP2 natural orbitals and improved virtual orbitals for the ICE-CI procedure (H2O molecule, cc-pVDZ basis, equilibrium geometry)

If ICE-CI is used in conjunction with MP2 natural orbitals, there also is the possibility of letting the program automatically choose the active space (this is called auto-ICE). The general idea is simple – we base the active space on the MP2 natural orbitals and their occupation numbers. All orbitals between occupation number say 1.98 down to 0.02 will be included in the active space. A relevant input is shown below.

```
! cc-pVDZ aug-cc-pV6Z/C Auto-ICE
%ice nmin 1.99 nmax 0.01 end

%paras R= 1.0 end

* int 0 1
O 0 0 0 0 0 0
H 1 0 0 { R} 0 0
H 1 2 0 { R} 104 0
*
```

If we scan along the H_2O dissociation surface one can see that despite changing active spaces, the dissociation curves are smooth and remain fairly parallel to the full CI dissociation curve. Depending on the tightness of the thresholds the active space may change from a small 6 electrons in 5 orbitals to a larger 8 electrons in 7 or 8 orbitals upon dissociation. This is the expected behavior as the σ-antibonding orbital becomes more stable along the bond stretching coordinate. Hence, these results are encouraging in as far as in many situations the program will be able to select a sensible active space without extended input from the user.

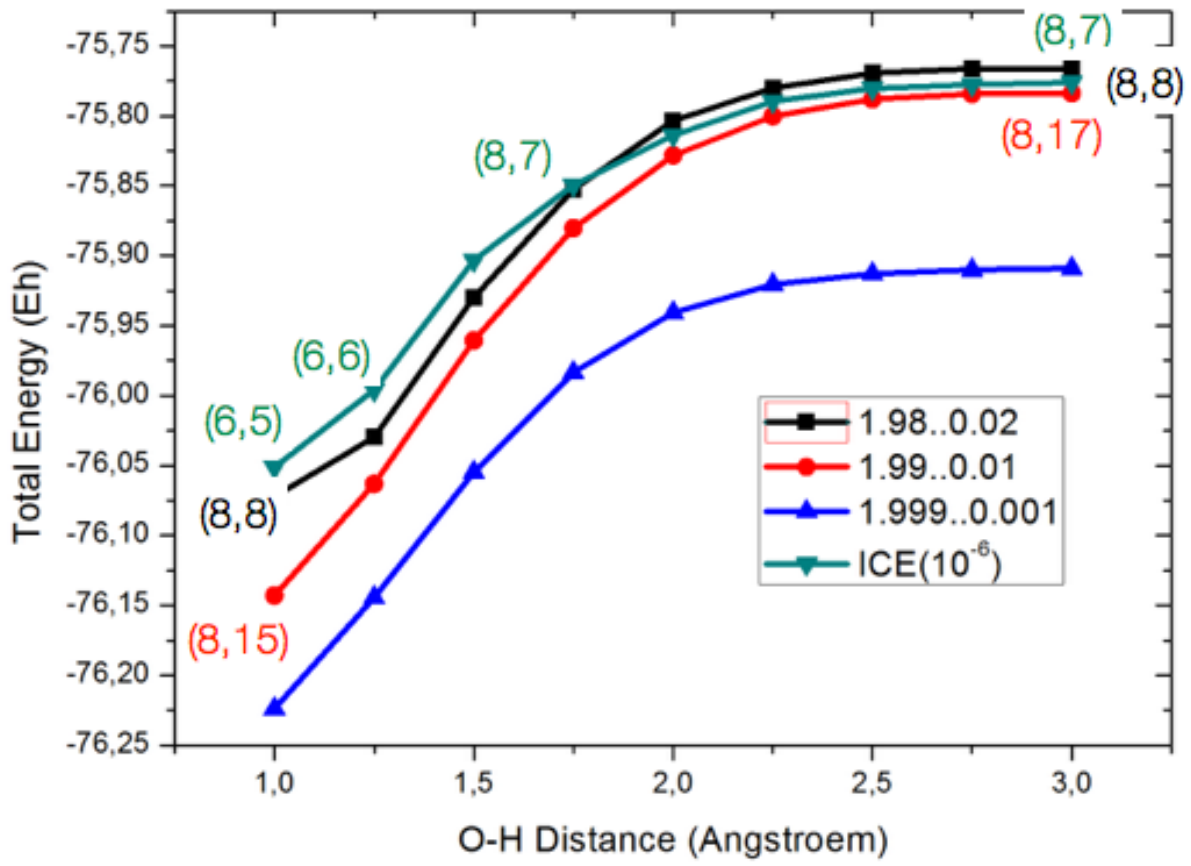


Fig. 7.18: Automatic active space selection along the H_2O dissociation surface. The reference curve (blue triangles) is the ICE-CI method for the full orbital space with the default parameters.

Another place, where automatic selection comes in conveniently is in the calculation of excited states. If there are no user supplied configurations, what happens is that the program will first choose an Aufbau "reference" configuration and then perform all single excitations relative to this configuration. The program will then diagonalize the Hamiltonian over the this set of configurations to create 0^{th} order approximations for the chosen number of roots of interest and then initiate the ICE-CI procedure starting from the leading configurations of these states. Here is an example for the benzene molecule:

```
! RHF def2-SVPD def2-SVP/C Auto-ICE
%cclib "/Users/neese/prog_c/orca/cclib/orcacc"
%ice nroots 5
     nmin 1.98
     nmax 0.02
     integrals ri
     end

* int 0 1
C 0 0 0 0.000000 0.000 0.000
C 1 0 0 1.389437 0.000 0.000
C 2 1 0 1.389437 120.000 0.000
C 3 2 1 1.389437 120.000 0.000
C 4 3 2 1.389437 120.000 0.000
C 5 4 3 1.389437 120.000 0.000
H 1 2 3 1.082921 120.000 180.000
H 2 1 3 1.082921 120.000 180.000
H 3 2 1 1.082921 120.000 180.000
H 4 3 2 1.082921 120.000 180.000
H 5 4 3 1.082921 120.000 180.000
```

(continues on next page)

(continued from previous page)

```
H 6 5 4 1.082921 120.000 180.000
*
```

(The %cclib statement is explained below and is not mandatory here). The Auto-ICE procedure comes up with as many as 24 electrons in 19 orbitals, which already is a fairly heavy calculation. The procedure converges in five iterations and provides indeed the correct states: the gorund state, the $^1B_{2u}$ state at 6.4 eV, the $^1B_{1u}$ state at 8.9 eV and a degenerate $^1E_{1u}$ state at 10.0 eV. These excitation energies are still in error by about 2 eV relative to experiment, which is mainly due to missing dynamic correlation. However, the correct states and their sequence has been found.

The ICE-CI can be used to find the ground state if the actual ground state is not known. To this end, one simply has to turn off the selection steps. This makes the calculations more expensive, but they will converge to the lowest state. In the example below (again, the H_2O molecule) we start from a random quintuply excited configuration – the ICE-CI still finds the ground state after four iterations:

```
%ice
  nel     8
  norb    23
  nroots   1
  tvar 1e-11
  tgen 1e-04
  etol 1e-06

  # selection
  SelStart false
  SelIter  false
  # algorithm details
  useivos   false
  integrals exact
  cimaxdim     5 #Davidson expansion space = MaxDim * NRoots
  cimode        3

  # spatial sym (buggy)
  irrep 0
  # startup (optional)
  refs { 2 1 0 1 0 2 1 0 1 }
       end
end
```

However, if one wants to converge to an excited state, one should turn on the selection. In the example below (once more the water molecule) one can converge to the second excitated singlet state by judicious choice of the start configuration:

```
%ice
  nel 8
  norb 23
  nroots 1
  tvar 1e-11
  tgen 1e-04
  etol 1e-06
  # selection
  SelStart true
  SelIter true
  # algorithm details
  useivos false
  integrals exact
  cimaxdim 5
  cimode 3
  # spatial sym (buggy)
  irrep 0
  { #} startup (optional)
```

(continues on next page)

(continued from previous page)

```
  refs { 2 2 1 2 0 1 }
       end
end
```

7.24.10 Large-scale approximate CASSCF: ICE-SCF

ICE-CI can be used as a replacement for the CI step in a CASSCF framework. In this way, much larger CASSCF calculations than previously possible can be envisioned. In using the ICE-CI in this way, the active orbitals should be chosen as natural orbitals in order to ensure a proper canonicalization. In general, ICE-CI results will not be invariant with respect to the choice of orbitals. However, in practice we have not found this to be problematic. We refer to this as ICE-SCF.

The use is simple: in the %casscf block choose:

```
%casscf
...
cistep ice
# optional input with refined settings
  ci
     tgen 1e-4   # controls accuracy (default = 1e-4)
     tvar 1e-11  # default = 1e-7 * TGen
     maxiter 100 # number of allowed cycles (default = 64)
  end
end
```

The entire remaining input is the one for standard CASSCF calculations. In this way one can do CASSCF calculations with very large active space in reasonable turnaround times. We have not observed convergence problems that are worse than in the standard CASSCF procedure. The results in Fig. 7.19 show that the deviations from regular CASSCF energies are very small. The largest deviation observed for C_2H_4 is on the order of 0.2 mEh, which appears acceptable. Note that the CASSCF tutorial also covers larger examples and excitations energies computed with the ICE-CI as CI solver. As mentioned in the CASSCF section *The Complete Active Space Self-Consistent Field (CASSCF) Module*, some feature are not supported for ICE-CI e.g. magnetic properties as well NEVPT2 corrections are not yet available.

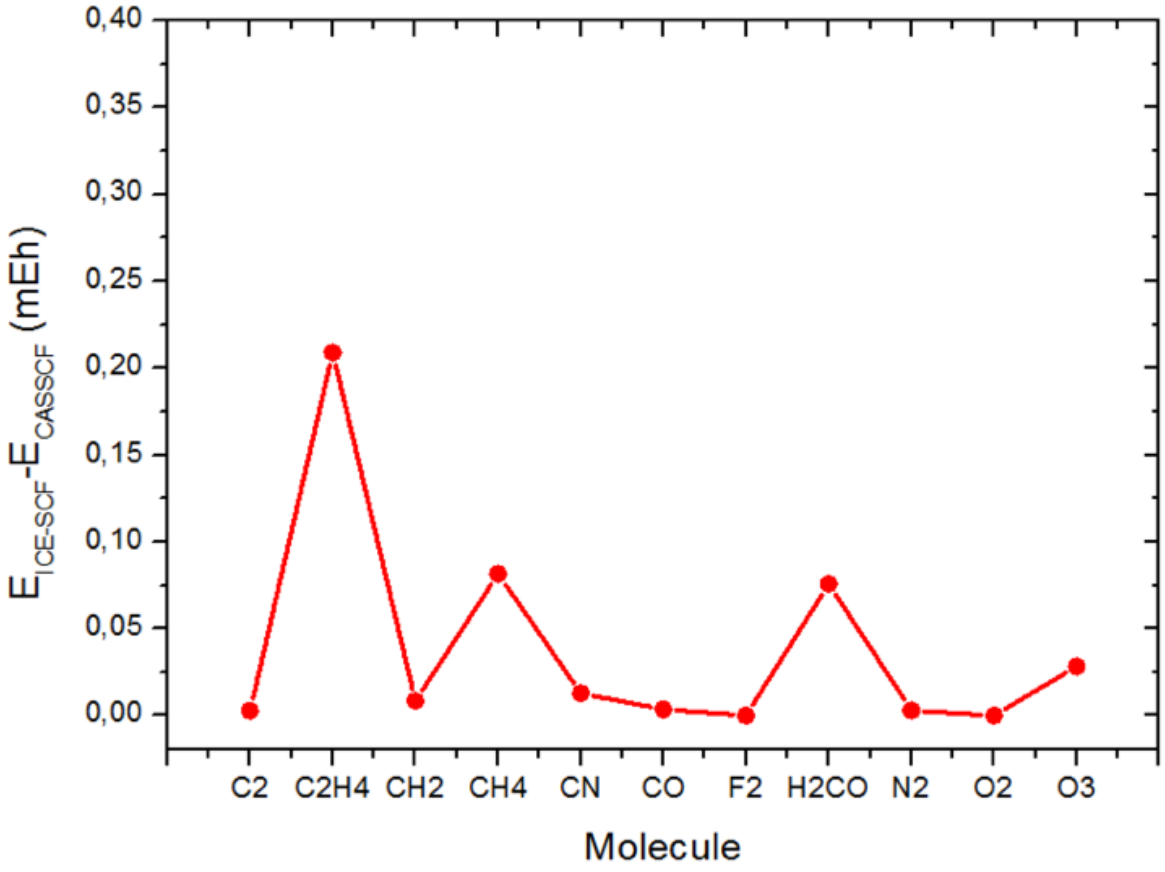


Fig. 7.19: Deviations of ICE-SCF from CASSCF energies for a selection of molecules (standard truncation parameters $T_{\text{gen}} = 10^{-4}$ and $T_{\text{var}} = 10^{-11}$)

Since CASSCF is fully variational, it is possible to optimize geometries with that procedure. It is our experience so far, that the ICE-SCF geometries are virtually indistinguishable from CASSCF geometries (an example is shown in Fig. 7.20).

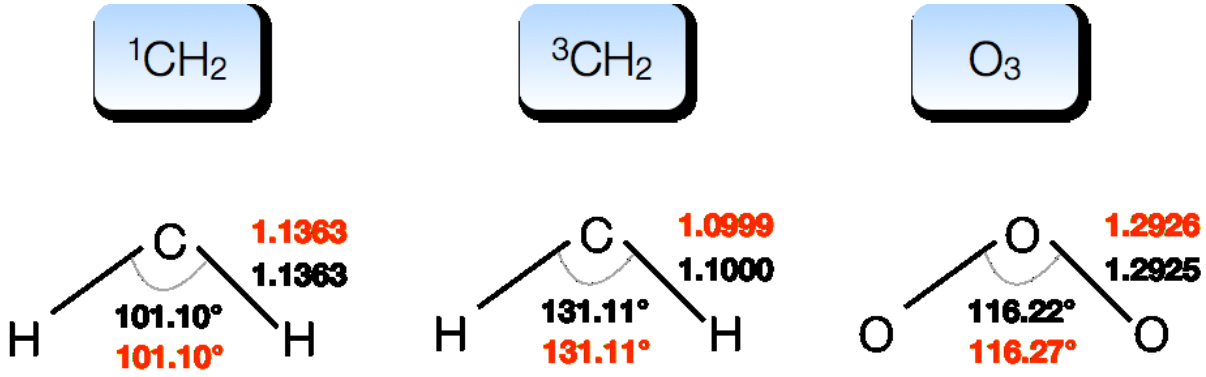


Fig. 7.20: CASSCF and ICE-SCF optimized geometries for methylene and ozone (cc-pVDZ basis set, default parameters).

7.24.11 The entire input block explained

For completeness, the parameters that can be specified in the input block are summarized below:

```
%ice
        nel            8 # number of active electrons
        norb          23 # number of active orbitals
        nroots         1 # number of roots
        mult           1 # requested multiplicity
        irrep          0 # requested irrep (buggy :-()
        tgen       1e-04 # generator threshold
        tvar       -1e-7 # negative -> 1e-7*tgen
        etol       1e-06 # convergence tolerance

        icetype     CFGs # The configuration based ICE-CI
                    CSFs # The CSF based ICE-CI
                    DETs # The CSF determinant based ICE-CI

        # algorithm details
        useMP2nat  false   # use MP2 natural orbitals
        useIVOs    false # use improved virtual orbitals
        useQROs    false # For UHF: use quasi-restricted MOs?
        integrals  exact # exact or ri transformation
        CIMaxIter     64 # max. number of CI iterations in the Davidson procedure
        CINGuessMat   512 # size of the CI guess matrix in the Davidson procedure
        CIMaxDim      10 # max. size of expansion space in the Davidson procedure
        CIMode         3 # default=accelerated CI, other settings not recommended
        CSFCIwithRI    1 # 1=with RI, 0=without RI, use withour RI when norb >> nel (e.g. (30e,
↪ 120o))
        CIBufferLength 10240 # Size (in elements) of the buffer list. Should be increased␣
↪according to
                             # the size of RI space.
        CIPSIOrbSweepWindow 1 # Use in case of large size of the S+D list and small available␣
↪memory.
                              # MOs can be divided into chunks (e.g. MOblock = MO/
↪CIPSIOrbSweepWindow)

        # startup configurations(optional)
        refs { 2 2 2 2 0 }
             { 2 2 2 0 2 }
             end
        end

        # natural orbitals with the ICE ansatz
        NatOrbs 1 # generates the .cipsi.nat GWB file containing the natural orbitals
```

7.24.12 A Technical Note: `orcacclib`

We should finally mention a technical aspect. The CI procedure in ICE-CI is based around the so-called one particle coupling coefficients

$$A_{pq}^{IJ} = \langle I | E_p^q | J \rangle$$

where A_{pq}^{IJ} is a coupling coefficient, I and J are configuration state functions (CSFs) and E_p^q is the spin-free excitation operator that promotes an electron from orbital p to orbital q. The values of these coupling coefficients only depend on the logical relationship between the CSFs I and J but not on the absolute values of I, J, p, q. In fact, they only depend on the number of unpaired electrons in I and the total spin S that both CSFs refer to. Hence, prototype coefficients can be pre-tabulated. This is normally done in a CI run at the beginning of the run. However, in ICE-CI it may have to be repeated several dozen times and for large numbers of open shells (say 14), the process is time and memory consuming.

In order to ease the computational burden, we have provided a small utility program that tabulates the coupling coefficients for a given total spin S (rather the multiplicity $M = 2S + 1$) and maximum number of open shells. This program is called `orcacclib`. It is called like:

```
orca_cclib Mult MaxNOpen
```

or - if you want to speed up the generation of the cclib:

```
mpirun -np 4 /full_path/orca_cclib_mpi Mult MaxNOpen cclib  # using 4 processes
```

It will produce a series of files orcacc.el.mult.nopen (electron density coupling coefficients) and orcacc.sp.mult.nopen (spin-density coupling coefficients) in the current directory. These files are binary files. They can be copied to an arbitrary directory. You instruct the program to read these coefficients (rather than to recalculate them all the time) by setting the path to this directory:

```
# My Job
! def2-SVP Auto-ICE
%cclib "/user/me/orca/cclib/orcacc"
```

The remaining part of the filename will be automatically added by the program. This option can save humongous amounts of time. The coupling coefficient library needs to be made for the desired multiplicities only once. The practical limit will be 14-16 open shells. If you are running the calculation on a cluster using some submit script, you have to ensure that the provided `cclib` path is accessible from the compute node.

7.25 CI methods using generated code

The AUTOCI module is a replacement of the `orca_mdci` for cases, where manual implementation of the method would be tedious or practically impossible. The module works with all types of reference wave function available in ORCA, i.e., RHF, ROHF, UHF and CASSCF and offers CI and related methods. All the methods are implemented in canonical orbital basis and storing all integrals on disk.

7.25.1 Introduction

All the theories are obtained by the means of automated programming within the ORCA-AGE (Automated Generator Environment for ORCA).[474, 500] The CI module reads in the SCF wavefunction and optimizes the coefficient of the CI expansion. Conceptually, the module is similar to `orca_mdci`, therefore the input and output do have a lot in common.

7.25.2 Input

All parameters applicable to the AUTOCI module are shown below.

```
%autoci
    # Algorithm selection
    citype      # Type of the CI expansion to be applied (one of following)
        CID     # Configuration Interaction with double substitutions
        CISD    # Configuration Interaction with singles and doubles
        CISDT   # Configuration Interaction with singles, doubles and triples
        CCD     # Coupled Cluster with double substitutions
        CCSD    # Coupled Cluster with single and double substitutions
        CCSDT   # Coupled Cluster with single, double and triple substitutions
        CEPA(0) # Linearized CCSD
        QCISD   # Quadratic CISD
        CC2     # Approximate CCSD
        CC3     # Approximate CCSDT
        CCSD[T]    # CCSD with perturbative [T] correction
```

(continues on next page)

(continued from previous page)

```
        CCSD(T)    # CCSD with perturbative (T) correction
        CCSDT-1    # Approximate CCSDT, CCSDT-1
        CCSDT-2    # Approximate CCSDT, CCSDT-2
        CCSDT-3    # Approximate CCSDT, CCSDT-3
        CCSDT-4    # Approximate CCSDT, CCSDT-4
        FIC-MRCI   # Fully internally contracted MRCI
        FIC-MRCEPA(0)   # Fully internally contracted CEPA(0)
        FIC-MRACPF      # Fully internally contracted ACPF
        FIC-MRAQCC      # Fully internally contracted AQCC
        FIC-DDCI3       # FIC-MRCI without the IJAB excitation
        FIC-MRCC        # Fully internally contracted MRCC
        MP2             # Second order Moeller-Plesset perturbation theory
        MP3             # Third order Moeller-Plesset perturbation theory
        MP4(SDQ)        # MP4 without triple substitutions
        MP4             # Fourth order Moeller-Plesset perturbation theory
        MP5             # Fifth order Moeller-Plesset perturbation theory

    # converger details
    stol       1e-06 # residue convergence tolerance
    maxiter    50    # maximum number of iterations
    maxdiis    5     # depth of the DIIS memory
    diisstartiter 2  # Apply DIIS starting at iteration 1
    ExcludeHigherExcDIIS false # exclude triples and higher excitations from DIIS procedure

    # CAS settings similar to the CASSCF input
    nel        6         # number of active electrons (for CAS)
    norb       7         # number of active orbitals (for CAS)
    mult           1 # requested multiplicity block
    nroots         1 # number of roots for mult block
    irrep          0 # requested irrep for mult block
    nthresh       1e-6 # Threshold for lin. dependencies in the IC-CSFs basis
    D3TPre        1e-14 # Density truncation in D3
    D4TPre        1e-14 # Density truncation in D4

    # Algorithm details
    printlevel   3    # Amount of printing
    trafotype  0      # Type of integral transformation
        0             # Full canonical
        1             # Full using RI (RI basis needed)

    keepints                # Keep the transformed integrals on disk
    useoldints                # Use the transformed integrals found on disk
    RunROHFasUHF    # Invokes AUTOCI UHF modules with orbitals from ROHF calculation

    # Property calculations
    density           # type of density requested
         none         # No density calculation
         linearized   # Linear part of the density
         unrelaxed    # Unrelaxed 1-body density matrix
         relaxed      # Relaxed 1-body density matrix

end
```

N.B. For a ROHF reference, only CISD calculations can be performed in the current version. However, it is possible to run UHF calculations with an ROHF reference by setting the `RunROHFasUHF` flag to true. Note that this only makes sense when the reference is indeed ROHF, e.g. (calculating the isotropic part of the HFC at CCSD level, running AutoCI UHF CCSD module, with orbitals obtained from the ROHF SCF calculation):

```
! ROHF def2-svp tightscf pmodel AUTOCI-CCSD

%autoci
```

(continues on next page)

(continued from previous page)

```
    RunROHFasUHF true
end

* xyz 0 2
Cu  0.0  0.0  0.0
*

%eprnmr nuclei = all Cu {aiso} end
```

If one wishes to experiment with the module itself and the reference wavefunction stays constant, it is possible to store the transformed MO integrals on disk (`keepints`) and reuse them (`useoldints`). The program checks only whether the dimension of the integrals on disk match the problem actually solved, i.e. the user is responsible for valid data.

7.25.3 Available Properties

The following single-reference methods are currently implemented in the AUTOCI.

Reference	Correlation	Energy	Gradient
RHF	CID	✓	✓
RHF	CISD	✓	✓
RHF	CISDT	✓	
RHF	CCD	✓	✓
RHF	CCSD	✓	✓
RHF	CCSD[T]	✓	✓
RHF	CCSD(T)	✓	✓
RHF	CCSDT	✓	
RHF	CEPA(0)	✓	✓
RHF	CC2	✓	
RHF	QCISD	✓	
RHF	MP2	✓	✓
RHF	MP3	✓	✓
RHF	MP4	✓	✓
RHF	MP4(SDQ)	✓	✓
UHF	CID	✓	✓
UHF	CISD	✓	✓
UHF	CISDT	✓	✓
UHF	CCD	✓	✓
UHF	CCSD	✓	✓
UHF	CCSD[T]	✓	✓
UHF	CCSD(T)	✓	✓
UHF	CCSDT	✓	✓
UHF	CCSDT-1	✓	
UHF	CCSDT-2	✓	
UHF	CCSDT-3	✓	
UHF	CCSDT-4	✓	
UHF	CEPA(0)	✓	✓
UHF	CC2	✓	
UHF	CC3	✓	
UHF	QCISD	✓	
UHF	MP2	✓	✓
UHF	MP3	✓	✓
UHF	MP4	✓	✓
UHF	MP4(SDQ)	✓	✓
UHF	MP5	✓	✓

continues on next page

Table 7.21 – continued from previous page

Reference	Correlation	Energy	Gradient
ROHF	CISD	✓	

Any AUTOCI method can be called from the simple keyword by prepending `AUTOCI-` to the correlation method, for instance

```
! AUTOCI-CCSD
```

7.25.4 Analytic Nuclear Gradients with AUTOCI

Obtaining accurate geometries is crucial to computing molecular properties accurately. In order to perform geometry optimisations, the nuclear gradient is necessary and while this can easily be obtained using numerical finite difference methods, it is also quite costly. More importantly, perhaps, is the fact that numeric derivatives tend to become unstable. Therefore, being able to evaluate analytic gradients is of vital importance. Using the AGE, a general framework has been built that supports arbitrary-order CI, CC and MPn nuclear gradients (and other derivatives).[500]

An example is shown below how to optimise a geometry using AUTOCI's gradients at the CCSD level of theory

```
! RHF cc-pVTZ AUTOCI-CCSD VerytightSCF Opt

%maxcore 10000

*xyz 0 1
...
*
```

The analytic gradients can even be used to perform semi-numerical frequency calculations

```
! AUTOCI-CCSD NumFreq
```

Besides nuclear gradients, all other first-order properties available in ORCA are available for the respective methods, such as dipole/quadrupole moments, hyperfine couplings or quadrupole splittings. As discussed above, (un)relaxed densities can be requested via

```
%autoci
  density relaxed
end
```

For geometry optimisations, both the unrelaxed and relaxed densities are computed automatically and do not need to be requested explicitly.

7.25.5 AUTOCI Response Properties via Analytic Derivatives

For single-reference methods (currently limited to `CCSD` and `MP2`), some response properties could be calculated via taking the analytic derivative of the wavefunctions computed by AUTOCI.

The input parameters for the wavefunctions are controlled by the `%autoci` block, while the property-related parameters are controlled by respective property blocks, like `elprop` and `eprnmr`. Some useful options are shown below.

```
%autoci
    citype          # only CCSD and MP2 available
        CCSD
        MP2
    STol      1e-06 # residue convergence tolerance
```

(continues on next page)

(continued from previous page)

```
    MaxIter   50    # maximum number of iterations
    MaxDIIS   5     # depth of the DIIS memory
    density         # need at least unrelaxed density (see details below)
        unrelaxed
        relaxed
end

%elprop
    polar   true    # polarizability via analytic derivative
end

%eprnmr
    NMRShielding 2  # NMR shielding for all nuclei, equivalent to the 'NMR' simple input
end
```

N.B. For the response property calculations, the electron density and electron response density needs to be calculated. Currently for the analytic polarizability at CCSD level, only the unrelaxed density version is implemented:

```
! UHF AUTOCI-CCSD Def2-SVP NoFrozenCore

%elprop
    polar true
end

* xyz 0 1
O         0.0000000000  0.0000000000 -0.1190150726
H         0.7685504811  0.0000000000  0.4760602904
H        -0.7685504811  0.0000000000  0.4760602904
*
```

For MP2, both the polarizability and NMR shielding needs the relaxed densities:

```
! UHF AUTOCI-MP2 Def2-SVP NoFrozenCore

%elprop
    polar true
end

%autoci
    density relaxed
end

* xyz 0 1
O         0.0000000000  0.0000000000 -0.1190150726
H         0.7685504811  0.0000000000  0.4760602904
H        -0.7685504811  0.0000000000  0.4760602904
*
```

```
! UHF AUTOCI-MP2 Def2-SVP NMR NoFrozenCore

%autoci
    density relaxed
end

* xyz 0 1
O         0.0000000000  0.0000000000 -0.1190150726
H         0.7685504811  0.0000000000  0.4760602904
H        -0.7685504811  0.0000000000  0.4760602904
*
```

Please also note that for AUTOCI NMR calculations, the GIAOs (gauge-including atomic orbitals) are necessary

(turned on by default). Also, there is no fronzen core options implemented yet (`NoFrozenCore` keyword is needed).

7.25.6 Fully Internally Contracted MRCI

Starting point for any multireference approach is a reference wavefunction that consists of multiple determinants or configurations state functions (CSFs). In many instances this is the complete active space SCF (CASSCF) wavefunction. In the uncontracted MRCI approach, as implemented in the `orca_mrci` module, the wavefunction is expanded in terms of excited CSFs that are generated by considering excitations with respect to all reference CSFs. The methodology scales with the number of reference CSFs and hence is restricted to small reference spaces. Moreover, the configuration driven algorithm used in `orca_mrci` keeps all integrals in memory, which further limits the overall size of the molecule.

Internal contraction as proposed by Meyer and Siegbahn avoids these bottlenecks [584, 796]. Here, excited CSFs are generated by applying the excitation operator to the reference wavefunction as whole. The fully internally contracted MRCI presented here (FIC-MRCI) uses the same internal contraction scheme as the FIC-NEVPT2 (aka PC-NEVPT2). The entire methodology as well as a comparison with the conventional uncontracted MRCI is reported in our article [805]. The CEPA0, ACPF and AQCC variants are straight forward adoptions [741]. The residue of the FIC-MRCI ansatz

$$R_K = \langle \Phi_{qs}^{pr} | \hat{H} - E^{\mathrm{CAS}} - \lambda E_c | \Psi^{\mathrm{FIC}} \rangle \,, \tag{7.188}$$

is modified by the factor

$$\lambda = \begin{cases} 1 & \text{MRCI} \\ 0 & \text{CEPA0} \\ \frac{2}{N_e} & \text{ACPF} \\ \frac{1-(N_e-3)(N_e-2)}{N_e \cdot (N_e-1)} & \text{AQCC} \end{cases}$$

Here, E_c is the correlation energy and Φ_{qs}^{pr} denote the internally contracted CSF that arise from the action of the spin-traced excitation operators on the CAS-CI reference wave function

$$\Phi_{qs}^{pr} = E_q^p E_s^r \, |\Psi^{\mathrm{CAS}}\rangle \,.$$

In case of ACPF and AQCC the λ factor explicitly depends on the number of correlated electrons, N_e.

The general input structure is like that of the CASSCF module, e.g., the following example input reads an arbitrary set of orbitals and starts the FIC-MRCI calculation. The internal contracted formalism requires CAS-CI reduced densities up to fourth order, which can be expensive to construct. By default, the density construction is speed up using the prescreening (PS) approximation reported in Section *N-Electron Valence State Pertubation Theory*.

```
!def2-tzvp moread allowrhf noiter nofrozencore
%moinp "start.gbw"  # could be from CASSCF

%autoci
citype   FIC-MRCI      # Fully internally contracted MRCI (singles, doubles)
         FIC-MRCEPA(0) # CEPA0 version of FIC-MRCI
         FIC-MRACPF    # ACPF  version of FIC-MRCI
         FIC-MRAQCC    # AQCC  version of FIC-MRCI
         FIC-DDCI3     # FIC-MRCI without the IJAB excitation

# CAS-CI reference wavefunction
nel 2
norb 2
mult 1,3
nroots 3,1

nthresh 1e-6  # removal of linear dependencies in the IC-CSFs
D3TPre        1e-14 # default density truncation of the 3-RDM
D4TPre        1e-10 # default density truncation of the 4-RDM
```

(continues on next page)

(continued from previous page)

```
# Davidson correction for the FIC-MRCI
DavidsonOpt 0  # none (default)
            1  # Davidson correction
end
```

Currently, the program is capable of computing total energies and vertical excitation energies. More features will be available with future releases.

7.25.7 Fully Internally Contracted MRCC

Several approaches have been taken towards extending CC theory to work with genuinely multiconfigurational reference wave functions [538], yet none of these approaches has found widespread adoption. As of 2011, the internally contracted MRCC theory has had a revival, with a rigorous theoretical investigation of several approximations that also proved its orbital invariance [249] and a first report of a polynomial-scaling code obtained through automatic equation generation [357].

Our implementation in ORCA is akin to the previously published formulations in Refs. [249, 357], although everything is formulated rigorously in terms of the spin-free excitation operators $\hat{E}^p_q = \hat{a}^{p\alpha}\hat{a}_{q\alpha} + \hat{a}^{p\beta}\hat{a}_{q\beta}$, using an improved version of the ORCA-AGE code generator.[500] To begin with, the *ansatz* for the wave function is

$$|\Psi\rangle = \mathrm{e}^{\hat{T}}|0\rangle \ , \tag{7.189}$$

where $|0\rangle$ denotes a zeroth order CASSCF reference wave function and the cluster operator can be written as (Einstein's summation convention implied)

$$\begin{aligned}\hat{T} = &\frac{1}{2}t^{ij}_{ab}\hat{E}^a_i\hat{E}^b_j + t^{it}_{ab}\hat{E}^a_i\hat{E}^b_t + \frac{1}{2}t^{tu}_{ab}\hat{E}^a_t\hat{E}^b_u + t^{ij}_{at}\hat{E}^a_i\hat{E}^t_j + t^{it}_{au}\hat{E}^a_i\hat{E}^u_t \\ &+t^{it}_{ua}\hat{E}^u_i\hat{E}^a_t + t^{tu}_{av}\hat{E}^a_t\hat{E}^v_u + \frac{1}{2}t^{ij}_{tu}\hat{E}^t_i\hat{E}^u_j + t^{it}_{uv}\hat{E}^u_i\hat{E}^v_t \ .\end{aligned}$$

Note that we do not use normal order in the cluster operator.

Inserting the *ansatz* from Eq. (7.189) into the Schrödinger equation and pre-multiplying with the inverse exponential, we obtain the similarity transformed Hamiltonian and the energy expression,

$$\bar{H}|0\rangle = \mathrm{e}^{-\hat{T}}\hat{H}\mathrm{e}^{\hat{T}}|0\rangle = E|0\rangle \ .$$

In our code, the similarity-transformed Hamiltonian is truncated after the quadratic terms since that approximation has been found to only have minor impact on the accuracy of the method [249],

$$\bar{H} \approx \hat{H} + [\hat{H},\hat{T}] + \frac{1}{2}[[\hat{H},\hat{T}],\hat{T}] \ .$$

The residual conditions are subsequently obtained by projecting with contravariant excited functions $\langle\tilde{\Phi}_P|$ onto the Schrödinger equation,

$$r_P = \langle\tilde{\Phi}_P|\bar{H}|0\rangle \ .$$

For a definition of the contravariant projection functions, we refer to Ref. [805] since this fic-MRCC implementation uses the same contravariant functions as the published fic-MRCI implementation. Despite using contravariant projection functions, this is not sufficient to remove all linear dependencies from the set of projection functions $\{\tilde{\Phi}_P\}$, i.e. the metric matrix

$$S_{PQ} = \langle\tilde{\Phi}_P|\tilde{\Phi}_Q\rangle \neq \delta_{PQ}$$

has off-diagonal elements *within* excitation classes and between classes with the same number of inactive and virtual indices (ITAU and ITUA). Hence, the set of projection functions needs to be orthonormalized, which is achieved with Löwdin's canonic orthogonalization in the AUTOCI module.[1]

[1] This is similar to scheme A from Ref. [357].

Input Example

The fic-MRCC module can be started by specifying the `CIType` keyword in the `%autoci` block or by adding `fic-MRCC` to the simple input line of an ORCA input file. The following example computes the singlet ground state energy of four hydrogen atoms arranged as a square with a side length of $2a_0$, which is commonly known as the H4 model [418].

```
! cc-pVTZ Bohrs  # it is possible to add the `fic-MRCC' keyword here
                 # and omit the %autoci block below

%maxcore 10000

%casscf
  nel 2
  norb 2
  mult 1
  nroots 1
end

%autoci  # CAS settings are automatically copied from the CASSCF block!
  citype fic-mrcc
end

* int 0 1
  H 0 0 0 0.0  0.0 0.0
  H 1 0 0 2.0  0.0 0.0
  H 2 1 0 2.0 90.0 0.0
  H 1 2 3 2.0 90.0 0.0
*
```

In this example, ORCA will first run a state-specific CASSCF calculation, and then immediately continue with the fic-MRCC calculation on top of the CASSCF solution from the first step. It is, however, not required to always run a CASSCF calculation before the `autoci` module. Any ORCA `gbw/mp2nat/...` file is accepted through `%moinp`, although that route requires the user to specify the active space in the `autoci` block. `autoci` will then compute a CASCI solution with the provided input orbitals and use that information to drive the correlated calculations.

Note that in the current implementation, `autoci_fic_mrcc` requires large stack sizes, which may lead to segmentation faults immediately after starting the program. We recommend a stack size of at least 16 MiB. On Unix-like systems, this can be achieved by issuing the command `ulimit -s 16384` in the terminal or in the submission script prior to running the fic-MRCC module.

Please be aware that fic-MRCC is a very extensive theory, which leads to long run times. The computational effort depends mainly on the number of orbitals, the number of total electrons and the size of the active space. On modestly modern hardware, calculations of ~ 300 orbitals with a CAS(2,2) should be readily achievable. For larger active spaces, such as a CAS(6,6), calculations with a total of ~ 200 orbitals will also complete within a day.

7.26 Geometry Optimization

ORCA is able to calculate equilibrium structures (minima and transition states) using the quasi Newton update procedure with the well known BFGS update [67, 241, 399, 762, 763, 764], the Powell or the Bofill update. The optimization can be carried out in either redundant internal (recommended in most cases) or Cartesian displacement coordinates. As initial Hessian the user can choose between a diagonal initial Hessian, several model Hessians (Swart, Lindh, Almloef, Schlegel), an exact Hessian and a partially exact Hessian (both recommended for transition state optimization) for both coordinate types. In redundant internal coordinates several options for the type of step to be taken exist. The user can define constraints via two different paths. He can either define them directly (as bond length, angle, dihedral or Cartesian constraints) or he can define several fragments and constrain the fragments internally and with respect to other fragments. The ORCA optimizer can be used as an external optimizer, i.e.the energy and gradient calculations done by ORCA.

7.26.1 Input Options and General Considerations

The use of the geometry optimization module is relatively straightforward.[1]

```
%method RunTyp Opt # use geometry optimization.
                   #(equivalent is RunTyp=Geom)
        end

# or simply "! Opt" in the keyword line

# details of the optimization are controlled here
%geom
      MaxIter 50  # max. number of geometry iterations
                  #  (default is 3N (N = number of atoms), at least 50 )
      # coordinate type control
      coordsys redundant     # redundant internal coords (2022)
               cartesian     # Cartesian coordinates
      # fallback option to Cartesian step if internals fail
      cartfallback true
      # transition state (TS) optimization
      TS_search EF      # Switch on TS search, EF means
                        #  "eigenvector following"
                        #  alternatively use "! OptTS"
      TS_Mode {M 0} end # Choose the mode to follow uphill in the
                        #  TS optimization. {M X}: eigenvector of
                        #  the Hessian with X. lowest eigenvalue
                        #  (start counting at zero) (default: X=0)
      # Instead of a mode choose an internal coordinate strongly
      #  involved in the eigenmode followed uphill
      TS_Mode {B 0 1} end      # bond between atoms 0 and 1  or
      TS_Mode {A 2 1 0} end    # angle between atoms 2, 1 and 0  or
      TS_Mode {D 3 2 1 0} end  # dihedral of atoms 3, 2, 1 and 0
      # add or remove internal coordinates from the automatically
      #  generated set of redundant internal coords
      modify_internal
          { B 10 0 A }      # add a bond between atoms 0 and 10
          { A 8 9 10 R }    # remove the angle defined
                            #  by atoms 8, 9 and 10
          { D 7 8 9 10 R }  # remove the dihedral angle defined
          end               #  by atoms 7, 8, 9 and 10
      # constrain internal coordinates:
      Constraints
          { B N1 N2 value C }       # the bond between N1 and N2
          { A N1 N2 N1 value C }    # the angle defined by N1, N2
                                    #  and N3
          { D N1 N2 N3 N4 value C } # the dihedral defined by N1,
                                    #  N2, N3 and N4
          { C N1 C }                # the cartesian position of N1
          { B N1 * C}       # all bonds involving N1
          { B * * C}        # all bonds
          { A * N2 * C }    # all angles with N2 as central atom
          { A * * * C }     # all angles
          { D * N2 N3 * C } # all dihedrals with N2 and N3 as
                            #  central atoms
          { D * * * * C }   # all dihedrals
      end
      # scan an internal coordinate:
      Scan B N1 N2 = value1, value2, N end
```

(continues on next page)

[1] But that doesn't mean that geometry optimization itself is straightforward! Sometimes, even when it is not expected the convergence can be pretty bad and it may take a better starting structure to come to a stationary point. In particular floppy structures with many possible rotations around single bonds and soft dihedral angle modes are tricky. It may sometimes be advantageous to compute a Hessian matrix at a "cheap" level of theory and then do the optimization in Cartesian coordinates starting from the calculated Hessian.

(continued from previous page)

```
  # perform constrained optimizations with varying N1-N2-
  #  distance from value1 up to value2 in N steps;
  #  works as well for angles (use A N1 N2 N3) and for
  #  dihedrals (use D N1 N2 N3 N4)
Scan B N1 N2 [value1 value2 value3 ... valueN] end
  # perform constrained optimizations with N1-N2-distances
  #  as given in the list;
  #  works as well for angles (use A N1 N2 N3) and for
  #  dihedrals (use D N1 N2 N3 N4)
# perform a simultaneous multidimensional scan
# possible for up to 3 different scan parameters. They each must
# have an identical number of scan steps
simul_Scan true # default false
fullScan true   # if !ScanTS is requested, fullScan assures
                #  that the relaxed surface scan is fully
                #  carried out before the TS optimization is
                #  started (Default is false)
# fragment optimization:
#  1. all atoms have to belong to a fragment
#  2. you have to connect the fragments
ConnectFragments
    {1 2 C}      # constrain the internal coordinates
                 #  connecting fragments 1 and 2
    {1 2 C N1 N2}# constrain the internal coordinates
                 #  connecting fragments 1 and 2, the
                 #  fragments are connected via atoms
                 #  N1 and N2
    {1 3 O}      # optimize the internal coordinates
                 #  connecting fragments 1 and 3
    {1 3 O N1 N2}# optimize the internal coordinates
                 #  connecting fragments 1 and 3, the
                 #  fragments are connected via atoms
                 #  N1 and N2
end
#  3. you can constrain the fragment internally
ConstrainFragments  # constrain all internal coordinates
    { 1 }           #  containing only atoms of fragment 1
end
# optimize hydrogens
optimizeHydrogens true
  # in the context of a normal optimization all internal
  #  coordinates not involving any hydrogens are constrained
  # in the context of a fragment optimization all internal
  #  coordinates involving hydrogens are optimized (also in a
  #  constrained fragment)
# freeze the hydrogen positions with respect to the
#  heteroatoms
freezeHydrogens true
# invert the defined constraints, i.e. optimize the
#  constraints and constrain the remaining coordinates
# this only works for the redundant internal coordinates
# Cartesian coordinates are not affected by invertConstraints
invertConstraints true      # step type control
Step     qn       # quasi-Newton step
         rfo      # Rational function step (Default for !Opt)
         prfo     # partitioned RFO step (Default for !OptTS)
# Step size control
MaxStep 0.3 # maximum step length in internal coordi-
            #  nates. Default is 0.3 au
Trust -0.3 # Initial trust radius. Default is -0.3 au
           # Trust <0 - use fixed trust radius
```

(continues on next page)

(continued from previous page)

```
                #  of size -trust. I.e. -0.3 means fix
                #  the trust radius at 0.3
                # Trust >0 - use trust radius update. I.e. 0.3
                #  means start with trust radius 0.3 and update
                #  the trust radius after each optimization step
    # Convergence tolerances. Note that the calculation is
    # only converged if all criteria are fullfilled. All
    # values given are default values.
    TolE 5e-6      # Energy change (a.u.)
    TolRMSG  1e-4  # RMS gradient (a.u.)
    TolMaxG  3e-4  # Max. element of gradient (a.u.)
    TolRMSD  2e-3  # RMS displacement (a.u.)
    TolMaxD  4e-3  # Max. displacement (a.u.)
    # keyword for frequently used sets of convergence thresholds
    Convergence normal    # Default
                loose
                tight
    ProjectTR false # project translation and rotation
                    # default is false. MUST be false for
                    # redundant internals
 end
```

Keywords for the control of the Hessian (especially important for the TS optimization):

```
# initial Hessian control
     inhess   unit      # unit matrix
              Read      # Hessian in a .hess file (e.g. from
                        #  a previous NumFreq run), this command
                        #  comes with the following:
              InHessName "filename.hess" # filename of
                                         # Hessian input file
                                         # a previous .opt file
                                         # or a previous .carthess file
              # these only for redundants
              Lindh     # Lindh's model Hessian
              Almloef   # Almloef's model Hessian
              Schlegel  # Schlegel's model Hessian
              Swart     # Swart and Bickelhaupt`s model Hessian
              XTB0      # GFN0-xTB Hessian
              XTB1      # GFN1-xTB Hessian
              XTB2      # GFN2-xTB Hessian
              GFNFF     # GFN-FF Hessian
     # additional Hessian control for TS optimization
     Calc_Hess true  # calculate the Hessian numerically at the beginning
     Recalc_Hess  5  # calculate the Hessian at the beginning
                     #  and recalculate it after 5,10,.. cycles
     Hybrid_Hess {0 1 5 6} end # calculates a Hybrid Hessian
                               # exact calculation for
                               # atoms 0, 1, 5 and 6; works also
                               # with Calc_Hess and Recalc_Hess
     NumHess true      # requests use of numerical Hessian
     # modification of the internal Hessian
     Hess_Internal
       {A 3 2 1 D 2.0} # define a diagonal Hessian value of
                       #  2 Eh/Bohr2 for the angle between
                       #  atoms 3 2 1. This can also be done for
                       #  bonds, dihedrals and Cartesian
                       #  coordinates.) The Hessian values of
                       #  multiple coordinates can be modified
       reset 5    # reset the modified internal Hessian values
                  #  after 5 cycles
```

(continues on next page)

(continued from previous page)

```
       # The following is only recommended
       #  after a relaxed surface scan
       #  in this example of the scan coordinate B 1 0;
       #  "basename.004.xyz" contains the optimized structure
       #  of the scan step with highest energy
       {B 1 0 C}
       XYZ1 "scanName.003.xyz" # the xyz-files of the structures
       XYZ2 "ScanName.005.xyz" #  next to the highest energy point
       GBW1 "ScanName.003.gbw" # the gbw-files of the structures
       GBW2 "ScanName.005.xyz" #  next to the highest energy
                               # the gbw-files are optional
      end
     # Hessian update procedure
     Update   Powell
              Bofill  # default for TS optimization
              BFGS    # default for geometry optimization
     # Hessian modification (only for P-RFO step)
     HESS_Modification Shift_Diag # shift the diagonal elements
                                  #  (default)
                       EV_Reverse # reverse the
                                  #  diagonal elements
     # Minimal value of Hessian eigenvalues (only P-RFO step)
     HESS_MinEV 0.0001  # if an absolute Hessian eigenvalue
                        #  is smaller than this value, it is
                        #  set to HESS_MinEV
     # Rebuilding the model Hessian after a number of cycles can
     #  accelerate the convergene of the optimization
     NResetHess 20 # Set the number of geometry steps after which
                   #  a new model Hessian is built (only with BFGS
                   #  update)
     NStepsInResetHess 5  # since previous steps and gradients are
                          #  available, it is possible to include
                          #  information about the PES in the
                          #  newly built Hessian (via a BFGS
                          #  update). This number should be
                          #  smaller than NResetHess
  end
```

As for parameter scan runs ORCA has some special options that may help to speed up the optimization:

```
%geom UseSOSCF false # switches the converger to SOSCF
                     # after the first point. SOSCF may
                     # converge better than DIIS if the
                     # starting orbitals are good.
                     # default = false
      ReducePrint true # reduce printout after the first
                       # point default=true
 # the initial guess can be changed after the first
 # point. The default is MORead. The MOs of the pre-
 # vious point will in many cases be a very good guess
 # for the next point. In some cases however, you may
 # want to be more conservative and use a general guess.
 OptGuess   = OneElec  # the one electron matrix
            = Hueckel  # the extended Hueckel guess
            = PAtom;   # the PAtom guess
            = Pmodel   # the PModel guess
            = MORead   # MOs of the prev. point (default)
 end
```

Redundant Internal Coordinates

There are three types of internal coordinates: redundant internals, old redundant internals (redundant_old) and a new set of redundant internals (redundant_new, with improved internals for nonbonded systems). All three sets work with the same "primitive" space of internal coordinates (stretches, bends, dihedral angles and improper torsions). Only the redundant internals works with one more type of bends in cases where a normal bend would have been approximately 180°. In redundant internal coordinates the full primitive set is kept and the Hessian and gradient are transformed into this – potentially large – space. A geometry optimization step requires, depending on the method used for the geometry update, perhaps a diagonalization or inversion of the Hessian of dimension equal to the number of variables in the optimization. In redundant internal coordinates this space may be 2-4 times larger than the nonredundant subspace which is of dimension $3N_{\mathrm{atoms}} - 6(5)$. Since the diagonalization or inversion scales cubically the computational overhead over nonredundant spaces may easily reach a factor of 8–64. Thus, in redundant internal coordinates there are many unnecessary steps which may take some real time if the number of primitive internals is greater than 2000 or so (which is not so unusual). The timing problem may become acute in semiempirical calculations where the energy and gradient evaluations are cheap.

We briefly outline the theoretical background which is not difficult to understand:

Suppose, we have a set of n_I (redundant) primitive internal coordinates $\mathbf{q}$ constructed by some recipe and a set of $n_C = 3N_{\mathrm{atoms}}$ Cartesian coordinates $\mathbf{x}$. The B-matrix is defined as:

$$B_{ij} = \frac{\partial q_i}{\partial x_j} \tag{7.190}$$

This matrix is rectangular. In order to compute the internal gradient one needs to compute the "generalized inverse" of $\mathbf{B}$. However, since the set of primitive internals is redundant the matrix is rank-deficient and one has to be careful. In pratice one first computes the $n_I \times n_I$ matrix $\mathbf{G}$:

$$\mathbf{G} = \mathbf{B}\mathbf{B}^T \tag{7.191}$$

The generalized inverse of $\mathbf{G}$ is denoted $\mathbf{G}^-$ and is defined in terms of the eigenvalues and eigenvectors of $\mathbf{G}$:

$$\mathbf{G}^- = \begin{pmatrix} \mathbf{U} \\ \mathbf{R} \end{pmatrix}^T \begin{pmatrix} \mathbf{\Lambda^{-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{U} \\ \mathbf{R} \end{pmatrix} \tag{7.192}$$

Here $\mathbf{U}$ are the eigenvectors belonging to the nonzero eigenvalues $\mathbf{\Lambda}$ which span the nonredundant space and $\mathbf{R}$ are the eigenvectors of the redundant subspace of the primitive internal space. If the set of primitive internals is carefully chosen, then there are exactly $3N_{\mathrm{atoms}} - 6(5)$ nonzero eigenvalues of $\mathbf{G}$. Using this matrix, the gradient in internal coordinates can be readily computed from the (known) Cartesian gradient:

$$\mathbf{g}_q = \mathbf{G}^-\mathbf{B}\mathbf{g}_x \tag{7.193}$$

The initial Hessian is formed directly in the redundant internal space and then itself or its inverse is updated during the geometry optimization.

Before generating the Newton step we have to ensure that the displacements take place only in the nonredundant part of the internal coordinate space. For this purpose a projector P':

$$\mathbf{P}' = \mathbf{G}\mathbf{G}^- = \mathbf{G}^-\mathbf{G} \tag{7.194}$$

is applied on both the gradient and the Hessian:

$$\tilde{\mathbf{g}}_\mathbf{q} = \mathbf{P}'\mathbf{g}_\mathbf{q} \tag{7.195}$$

$$\tilde{\mathbf{H}}_\mathbf{q} = \mathbf{P}'\mathbf{H}_\mathbf{q}\mathbf{P}' + \alpha\left(\mathbf{1} - \mathbf{P}'\right) \tag{7.196}$$

The second term for $\tilde{H}$ sets the matrix elements of the redundant part of the internal coordinate space to very large values ($\alpha = 1000$).

Coordinate steps

A Quasi-Newton (QN) step is the simplest choice to update the coordinates and is given by:

$$\mathbf{\Delta q} = -\tilde{\mathbf{H}}_{\mathbf{q}}^{-\mathbf{1}}\tilde{\mathbf{g}}_{\mathbf{q}} \tag{7.197}$$

A more sophisticated step is the rational function optimization step which proceeds by diagonalizing the augmented Hessian:

$$\begin{pmatrix} \mathbf{H}_q & \mathbf{g}_q \\ \mathbf{g}_q & 0 \end{pmatrix} \begin{pmatrix} \Delta\mathbf{q} \\ 1 \end{pmatrix} = v \begin{pmatrix} \Delta\mathbf{q} \\ 1 \end{pmatrix} \tag{7.198}$$

The lowest eigenvalue ν_0 approaches zero as the equilibrium geometry is approached and the nice side effect of the optimization is a step size control. Towards convergence, the RFO step is approaching the quasi-Newton step and before it leads to a damped step is taken. In any case, each individual element of $\Delta\mathbf{q}$ is restricted to magnitude `MaxStep` and the total length of the step is restricted to `Trust`. In the RFO case, this is achieved by minimizing the predicted energy on the hypersphere of radius `Trust` which also modifies the direction of the step while in the quasi-Newton step, the step vector is simply scaled down.

Thus, the new geometry is given by:

$$\mathbf{q}_{\text{new}} = \mathbf{q}_{\text{old}} + \Delta\mathbf{q} \tag{7.199}$$

However, which Cartesian coordinates belong to the new redundant internal set? This is a somewhat complicated problem since the relation between internals and Cartesians is very nonlinear and the step in internal coordinates is not infinitesimal. Thus, an iterative procedure is taken to update the Cartesian coordinates. First of all consider the first (linear) step:

$$\Delta\mathbf{x} = \mathbf{A}\Delta\mathbf{q} \tag{7.200}$$

with $\mathbf{A} = \mathbf{B}^T\mathbf{G}^-$. With the new Cartesian coordinates $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}$ a trial set of internals $\mathbf{q}_{k+1}$ can be computed. This new set should ideally coincide with $\mathbf{q}_{\text{new}}$ but in fact it usually will not. Thus, one can refine the Cartesian step by forming

$$\Delta\Delta\mathbf{q} = \mathbf{q}_{\text{new}} - \mathbf{q}_{k+1} \tag{7.201}$$

which should approach zero. This leads to a new set of Cartesians $\Delta\mathbf{x}' = \mathbf{A}\Delta\Delta\mathbf{q}$ which in turn leads to a new set of internals and the procedure is iterated until the Cartesians do not change and the output internals equal $\mathbf{q}_{new}$ within a given tolerance (10^{-7} RMS deviation in both quantities is imposed in ORCA).

Constrained Optimization

Constraints on the redundant internal coordinates can be imposed by modifying the above projector P' with a projector for the constraints C:

$$\mathbf{P} = \mathbf{P}' - \mathbf{P}'\mathbf{C}\left(\mathbf{CPC}\right)^{-\mathbf{1}}\mathbf{CP}' \tag{7.202}$$

C is a diagonal matrix with 1's for the constraints and 0's elsewhere. The gradient and the Hessian are projected with the modified projector:

$$\tilde{g}_q = Pg_q \tag{7.203}$$

$$\tilde{H}_q = PH_qP + \alpha\left(1 - P\right) \tag{7.204}$$

Constrained Fragments Optimization

The constrained fragments option was implemented in order to provide a convenient way to handle constraints for systems consisting of several molecules. The difference to a common optimization lies in the coordinate setup. In a common coordinate setup the internal coordinates are built up as described in the following:

In a first step, bonds are constructed between atom pairs which fulfill certain (atom type specific) distance criteria. If there are fragments in the system, which are not connected to each other (this is the case when there are two or more separate molecules), an additional bond is assigned to the nearest atom pair between the nonbonded fragments. All other internal coordinates are constructed on the basis of this set of bonds. Here, in a second step, bond angles are constructed between the atoms of directly neighboured bonds. If such an angle reaches more than 175°, a special type of linear angles is constructed. In a third step, dihedral angles (and improper torsions) are constructed between the atoms of directly neighbouring angles.

If the constrained fragments option is switched on, the set of bonds is constructed in a different way. The user defines a number of fragments. For each fragment a full set of bonds (not seeing the atoms of the other fragments) is constructed as described above. When using this option, the user also has to define which fragments are to be connected. The connection between these fragments can either be user-defined or automatically chosen. If the user defines the connecting atoms N1 and N2, then the interfragmental bond is the one between N1 and N2. If the user does not define the interfragmental bond, it is constructed between the atom pair with nearest distance between the two fragments. Then the angles and dihedrals are constructed upon this (different) set of bonds in the already described fashion.

Now let us regard the definition of the fragment constraints: A fragment is constrained internally by constraining all internal coordinates that contain only atoms of the respective fragment. The connection between two fragments A and B is constrained by constraining specific internal coordinates that contain atoms of both fragments. For bonds, one atom has to belong to fragment A and the other atom has to belong to fragment B. Regarding angles, two atoms have to belong to fragment A and one to fragment B and vice versa. With respect to dihedrals, only those are constrained where two atoms belong to fragment A and the other two belong to fragment B.

7.26.2 Transition State Optimization

As transition state finder we implemented the well-established eigenvector following algorithm using a P-RFO step as implemented by Baker [67]. This algorithm is a quasi-Newton like algorithm.

The Taylor series of the energy, truncated after the quadratic term, is:

$$E = E_0 + g_q{}^{+}\Delta q_q + \frac{1}{2}\Delta q{}^{+}\mathrm{H}_q\Delta q \tag{7.205}$$

The Newton-Raphson step to get from the actual point to a stationary point is:

$$\Delta q = -\mathrm{H}_q^{-1} g_q = \sum -\frac{V_i^{+} g_q V_i}{b_i} \tag{7.206}$$

with V_i and b_i as eigenvectors and eigenvalues of the Hessian H_q. This step leads to the nearest stationary point on the PES. This stationary point can be a minimum or a saddle point, according to the curvature of the PES at the actual point.

With a simple shift of the Hessian eigenvalues b_i in this equation one can guide the step to a stationary point with the required characteristics (Hessian with exactly one negative eigenvalue). The transition state search is separated into two different optimization problems. The energy is maximized along one Hessian eigenmode and minimized along the remaining $3N - 7(6)$ eigenmodes. We introduce two different shift parameters λ_p and λ_n, where λ_p is the shift parameter for the eigenmode being maximized and λ_n shifts the Hessian eigenvalues of the modes being minimized. This method allows us to maximize along any mode, not only the one with smallest eigenvalue. Starting from two different RFO-matrices for the different optimization problems (see description above) we get for λ_p and λ_n:

$$\lambda_p = \frac{1}{2}b_k + \frac{1}{2}\sqrt{b_k^2 + 4F_k^2} \quad \text{and} \quad \sum_{i \neq k} \frac{F_i^2}{\lambda_n - b_i} = \lambda_n \tag{7.207}$$

whereas $F_i = V_i^+ g$ is the component of g along the Hessian eigenmode V_i and λ_n has to get solved iteratively. The solution for λ_n has to be negative and lower than b_2 (or lower than b_1, if not the lowest mode is being followed). If the Hessian has more than one negative eigenvalue, these properties might not be fulfilled, and the Hessian would have to be modified. In our implementation the Hessian diagonal elements are either shifted or reversed in such a case.

Once the shift parameters are known the P-RFO step h is calculated as follows:

$$\Delta q_k = -\frac{\bar{F}_k V_k}{b_k - \lambda_p} \quad \text{and} \quad \Delta q_i = -\frac{\bar{F}_i V_i}{b_i - \lambda_n} \quad \text{with} \quad i = 1 \ldots n, \;\; i \neq k \tag{7.208}$$

$$\Delta q = \sum_{j=1}^{n} \Delta q_j \tag{7.209}$$

ScanTS option

For TS modes of rather local nature (involving only one bond or an angle; no concerted movements over multiple atoms) we implemented the ScanTS feature. Here the user can carry out a relaxed surface scan and a TS optimization in one calculation. After the relaxed surface scan the algorithm chooses the optimized structure of the scan with highest energy as initial guess structure and the two neighbouring structures for the calculation of the second derivative of the scanned coordinate (e.g., if scan step number 4 gives the structure with highest energy, then structure `basename.004.xyz` is the initial guess for the TS optimization; the structures `basename.003.xyz` and `basename.005.xyz` are used for the calculation of the second derivative). Before the first step of the subsequent TS optimization the energies and gradients for all three structures are calculated. The gradients are then transformed to internal coordinates. The diagonal Hessian value of the scanned coordinate is then calculated via finite difference of the internal gradients of the two given structures (003 and 005 in our example).

For the construction of the initial Hessian a model force field Hessian is built up (this Hessian has got only diagonal entries and zeros as off-diagonal elements). The exactly calculated diagonal Hessian value replaces the model force field Hessian entry for the respective internal coordinate.

If the user already performed a regular relaxed surface scan without the subsequent TS optimization, then he can nevertheless use these structures for the same procedure. A relaxed surface scan always gives you the xyz-files and gbw-files of the optimized structures of each scan step. A separate TS optimization can be carried out where the structure with highest energy is the starting structure. Additionally the two files with the two adjacent structures (as explained above) have to be provided (via the Hess_Internal keyword, see below). Furthermore, the internal coordinate, for which the diagonal Hessian value has to be calculated, has to be given (the previously scanned coordinate). This exact Hessian calculation is only possible for one internal coordinate:

```
%geom
  Hess_Internal
     {B 1 0 C}               # previously scanned coordinate
     XYZ1 "scanName.003.xyz" # the xyz-files of the structures
     XYZ2 "ScanName.005.xyz" #  next to the highest energy point
     GBW1 "ScanName.003.gbw" # the gbw-files of the structures
     GBW2 "ScanName.005.xyz" #  next to the highest energy
                             # the gbw-files are optional
   end
end
```

Additionally the manipulation of the diagonal Hessian values of the internal Hessian is possible for further internal coordinates, but without an extra calculation. Here the user can just define a value (in Eh/Bohr^2).

```
Hess_Internal
  {A 3 2 1 D 2.0} # define a diagonal Hessian value of
                  #  2 Eh/Bohr2 for the angle between
                  #  atoms 3 2 1
  {B 1 0 D -0.5}  # define a diagonal Hessian value of
                  #  -0.5 Eh/Bohr2 for the bond between
                  #  atoms 1 and 0
end
```

The definition of such Hessian (diagonal) elements is possible for multiple internal coordinates. These just replace the values of the force field model Hessian.

Hybrid Hessian

We implemented the calculation of a "Hybrid Hessian" as an alternative to the full Hessian calculation for TS optimization. Here only those parts of the Hessian, that are important for the TS optimization, are calculated exactly. For this calculation we define two kinds of atoms: atoms whose couplings with the other atoms are treated exactly (E) and atoms whose couplings are treated approximately (A).

In a first step an Almloef model Hessian is built up in redundant internal coordinates and transformed to Cartesian coordinates. This Hessian gives the second derivative elements for atom pairs A/A. In a second step the second derivative elements between pairs E/E and E/A are calculated numerically as in a numerical frequency calculation:

$$\frac{\mathbf{\Delta E}}{\Delta i_B \Delta j_C} = \frac{\mathbf{\Delta E}}{\Delta j_C \Delta i_B} = \frac{g_{j,C}^{i,B} - g_{j,C}^{eq.}}{displ.} \tag{7.210}$$

with:

i, j	x-, y- or z-direction
B, C	pairs of E/E, E/A, A/E
$displ.$	magnitude of displacement
$g_{j,C}^{eq.}$	force on atom C in direction j in current geometry
$g_{j,C}^{i,B}$	force on atom C in direction j after displacement of atom B in direction i

Partial Hessian Vibrational Analysis

We implemented the Partial Hessian Vibrational Analysis (PHVA), as published by Li, Jensen in [513], for the analysis of the nature of stationary points of structures obtained with QM/MM optimizations.

```
# PHVA after a QM/MM optimization in the (dispersion-/PC-) field
# caused by the MM-atoms
! NumFreq
%LJCoefficients "temp.LJ" # file with the Lennard Jones
                          #  coefficients for dispersion interaction
                          # obtained from last QM/MM run
%pointcharges "temp.pc"   # file with the point charges for
                          #  electrostatic interaction
                          # obtained from last QM/MM run
 #
%freq
    PARTIAL_Hess {0 1 2}  # atoms which are "frozen" and which make
                          #  the boundary to the MM-system
     end
   end
```

NOTE

- This procedure should be used for QM/MM optimized structures only to verify the nature of the stationary point and have an estimate of the ZPE.

Here we shortly describe the procedure: In PHVA we divide the system into two parts B (of size n atoms) and C (size $N - n$). Let the atom set B belong to the region where the chemical change is localized. The Partial Hessian matrix is built up as follows:

$$\begin{pmatrix} K_{BB} & 0 \\ 0 & K_{CC}^{\varepsilon} \end{pmatrix} \tag{7.211}$$

With:

$$K_{BB} \quad : x, y, z \quad \text{direction}$$

$$K^{\varepsilon}_{CC} = \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}, \varepsilon = 10^{-8}\,\text{au}, \tag{7.212}$$

this corresponds to using near-infinite masses for the atoms in C.

With this procedure we get the following eigenvalue structure:

- Six zero eigenvalues with modes corresponding to translational and rotational motion of the entire molecule.
- $3(N-n)-6$ small (less than 1 cm^{-1}) eigenvalues with modes corresponding mainly to internal motion within region C.
- Three eigenvalues (typically less than 10 cm^{-1}) with modes corresponding mainly to motion of region C relative to region B.
- $(3n-3)$ eigenvalues with modes corresponding mainly to relative motion of B and C as well as internal motion within region B.

7.26.3 Minimum Energy Crossing Points

The MECP optimization allows the user to optimize to structures where two different potential energy surfaces (PES1 and PES2) cross each other. In this optimization two conditions apply: the energy E_1 of PES1 is minimized while at the same time the energy difference $(E_1 - E_2)^2$ of both surfaces is minimized. For the implementation we follow in principle the suggestions of Harvey et al. in [365].

For the minimization two different gradients are constructed:

The first gradient chosen for the minimization is

$$\mathbf{f} = \frac{\partial}{\partial q}\left(E_1 - E_2\right)^2 = 2\left(E_1 - E_2\right) \cdot x_1 \tag{7.213}$$

where x_1 is the gradient difference vector

$$x_1 = \left[\left(\frac{\partial E_1}{\partial q}\right) - \left(\frac{\partial E_2}{\partial q}\right)\right] \tag{7.214}$$

which is orthogonal to the crossing hyperline near the MECP.

The gradient

$$\mathbf{g} = \left(\frac{\partial E_1}{\partial q}\right) - \frac{x_1}{|x_1|}\left[\left(\frac{\partial E_1}{\partial q}\right) \cdot \frac{x_1}{|x_1|}\right] \tag{7.215}$$

is orthogonal to x_1.

Both gradients are combined to yield the effective surface crossing gradient

$$\mathbf{g_{SC}} = \mathbf{g} + \mathbf{f} \tag{7.216}$$

The crossing hyperline is defined as the $3N-7$ dimensional subspace of PES1, which is orthogonal to x_1. In the MECP optimization we want to find the point of lowest energy within this subspace.

Our calculation of normal modes and force constants for movements along the crossing hyperline differ from the one proposed by Harvey et al. A standard frequency analysis can not be performed, but a similar procedure is applied:

Let us regard the second-order Taylor expansion for the energy of both surfaces near the MECP for a displacement along the crossing hyperline (orthogonal to x_1):

$$\mathbf{E}_A = E_{\text{MECP}} + \frac{1}{2}\Delta q^T H_{\text{eff},A} \Delta q \tag{7.217}$$

with:

$\mathbf{E}_A$	Energy E_1 on PES1 or E_2 on PES2
$H_{\text{eff},A}$	effective Hessian for PES1 or PES2
Δq	displacement along the crossing hyperline

Diagonalization of this effective Hessian gives us the normal modes of the crossing hyperline and thus allows us to decide whether the MECP optimization converged to a minimum in the $3N - 7$ dimensional subspace of the crossing hyperline.

The procedure for the calculation of the effective Hessian is now as follows: For each of both surfaces the second derivative matrix is calculated. Then the 6 rotations and translations and additionally the direction of the gradient difference vector x_1 (this ensures that movement orthogonal to the crossing hyperline, for which we do NOT satisfy the conditions of a stationary point, is excluded) are projected out from the Hessian matrix.

For MECP optimizations the following options exist:

```
%mecp
   SurfCrossOpt true  # switches on the MECP optimization
                      #  alternatively use: ! SurfCrossOpt
   SurfCrossNumFreq true # switches on the MECP effective Hessian
                         #  calculation
                         #  alternatively use: ! SurfCrossNumFreq
 # separate MO input for the second spin state (PES2)
   moinp "Myfile.gbw"# MO input for PES2
 # information on the electronic structure of PES 2
   Mult 3            # multiplicity of PES2
   brokenSym 1,1     # broken symmetry for PES2
 # CASSCF options for PES2 (also see the CASSCF chapter)
   casscf_nel 6       # number of active space electrons
   casscf_norb 6      # number of active orbitals
   casscf_mult 1,3    # multiplicities singlet and triplet
   casscf_nroots 4,2  # four singlets, two triplets
   casscf_bweight 2,1 # singlets and triplets weighted 2:1
   casscf_weights[0] = 0.5,0.2,0.2,0.2 # singlet weights
   casscf_weights[1] = 0.7,0.3         # triplet weights
 end
```

7.26.4 Conical Intersections

The minima in the conical intersection seam-space between two states (named here I and J) can be found by using regular geometry optimization algorithms, except that the gradient to be optimized is [544]:

$$\mathbf{g} = \mathbf{g'}_{diff} + \mathbf{P}\mathbf{g}_{mean}$$

where $\mathbf{g'}_{diff} = 2(E_I - E_J)(\partial E_I/\partial q - \partial E_J/\partial q)$ is parallel to the gradient difference vector; $\mathbf{g}_{mean}$ is the gradient mean and $\mathbf{P}$ is a projection matrix that projects out the gradient difference ($\mathbf{x}$) and non-adiabatic coupling ($\mathbf{y}$) direction components:

$$\mathbf{P} = \mathbf{1} - \mathbf{x}\mathbf{x}^\mathbf{T} - \mathbf{y}\mathbf{y}^\mathbf{T}$$

Now we have three approaches to solve this problem in ORCA, that will be explained next.

Gradient Projection: This is exactly what has been described above, and will be chosen as default whenever NACMEs between I and J are available. It is in principle the faster and most accurate method. It can be invoked by setting:

```
%CONICAL
  METHOD GRADIENT_PROJECTION  #or simply GP
END
```

OBS.: Turning on the ETF (see Sec. *NACMEs with built-in electron-translation factor*) can improve the optimization when using the full Gradient Projection method.

Gradient Projection (without NACME) : It is an approximation to the method above, that one gets by completely neglecting the NACMEs. It is essentially equivalent to finding a surface crossing point, and will not necessarily find minima inside the CI seam-space, although the final ΔE_{IJ} should be zero.

```
%CONICAL
  METHOD GP_NONACME
END
```

Updated Branching Plane: Here the idea is to start from a guess NACME, which is any unit vector perpendicular to $\mathbf{x}$, and do an progressive update on it, similar to the BFGS update on the Hessian [544]. The "Branching Plane" defined by $\mathbf{x}$ and $\mathbf{y}$ gets then iteratively more accurate until covergence is achieved. It has been shown to be quite accurate and is the default whenever NACMEs are not available. Can be used with:

```
%CONICAL
  METHOD UBP
END
```

Finally, the ΔE_{IJ} energy threshold for the optimization can be altered with:

```
%CONICAL
  ETOL 1e-4 #default
END
```

7.26.5 Numerical Gradients

If you want to use numerical instead of analytic gradients you have to use

```
! NumGrad
```

in your input file. Additionally the settings for the numerical differentiation can be changed:

```
%numgrad
  CentralDiff true # You should use two-sided numerical differentiation, but it
                   # is possible to switch to one-sided numerical differentiation.
  DX 0.005         # Increment in Bohr for the differentiation.
  TransInvar  true # Take advantage of translation invariance
 end
```

7.26.6 ORCA as External Optimizer

If you want to make use of ORCA's routines for optimization, TS optimization, NEB, IRC, GOAT, etc., but not use ORCA's built-in electronic structure methods, you can use the keyword:

```
! ExtOpt
```

in your input file. All information that you give on the electronic structure is discarded. In each optimization/NEB/IRC step ORCA writes an input file called `basename_EXT.extinp.tmp` with the following info:

```
basename_EXT.xyz # xyz filename: string, ending in '.xyz'
0 # charge: integer
1 # multiplicity: positive integer
```

(continues on next page)

(continued from previous page)

```
1 # NCores: positive integer
0 # do gradient: 0 or 1
pointcharges.pc # point charge filename: string (optional)
```

Comments from # until the end of the line should be ignored. The file basename_EXT.xyz will also be present in the working directory. ORCA always requests the energy, but a gradient only if needed for the chosen calculation type.

ORCA then calls:

```
scriptname basename_EXT.extinp.tmp [args]
```

where args are optional command line arguments, which can be provided in the ORCA input file (see below) and are directly passed to the command line for the external program.

scriptname is the name of the external program or wrapper script, which is not distributed with the ORCA binaries and must be supplied by the user in one of the following ways:

1. as a file or link named otool_external in the same directory as the ORCA executables;
2. by assigning the EXTOPTEXE environment variable to the *full path* to the external program;
3. via the ORCA input:

```
%method
  ProgExt "/full/path/to/script"
  Ext_Params "optional command line arguments"
end
```

Regardless of which option is used, the keyword Ext_Params can be used to specify the additional command line arguments as a single string.

The external script starts the energy (and gradient) calculation and finally provides the results in a file called basename_EXT.engrad using the same basename as the XYZ file. This file must have the following format:

```
#
# Number of atoms: must match the XYZ
#
3
#
# The current total energy in Eh
#
-5.504066223730
#
# The current gradient in Eh/bohr: Atom1X, Atom1Y, Atom1Z, Atom2X, etc.
#
-0.000123241583
0.000000000160
-0.000000000160
0.000215247283
-0.000000001861
0.000000001861
-0.000092005700
0.000000001701
-0.000000001701
```

Comments from # until the end of the line are ignored, as are any comment-only lines.

The script may also print relevant output to STDOUT and/or STDERR. STDOUT will either be printed in the ORCA standard output, or redirected to a temporary file and removed afterwards, depending on the type of job (e.g., for NumFreq the output for the displaced geometries is always redirected) and ORCA output settings:

```
%output
  Print[P_EXT_OUT]  1 # (default) print the external program output
  Print[P_EXT_GRAD] 1 # (default) print the gradient from the ext. program
end
```

7.26.7 Gaussian as External Optimizer

To use the external optimizer from Gaussian in ORCA, the following keywords were provided in the past:

```
%geom
  UseGaussian true # Use the external Gaussian optimizer instead
                   #  of the ORCA optimizer.
  GaussianName "GAU" # String defining the name of the Gaussian
                     #  optimizer
  GauOptFlags   # String indicating the optimization flags
  Gaussian Constraints # List defining the constraints for
                       #  the Gaussian optimizer.
end
```

Since the ORCA team got banned by Gaussian in January 2007 we can no longer support these option flags. They have not been removed from the code and may or may not work. If there is trouble with it we can – unfortunately – not offer any help since we do not have access to the Gaussian code any longer.

7.27 Frequency calculations - numerical and analytical

In the ORCA program package, the calculation of frequencies through the numerical or analytical Hessian is done via the `orca_numfreq` module and the combination of the `orca_propint`, `orca_scfresp`, and `orca_prop` modules, respectively.

The parameters to control these frequency calculations can be specified in the %freq-block.

```
%freq

  # Flags to switch frequencies calculation on/off

  NumFreq    false       # numerical frequencies (available for all methods)
  AnFreq     false       # analytical frequencies (available for HF, DFT)
                         # (One of these options has to be set to true,
                         #  to request a freq calculation)

  ScalFreq   1.0         # Scaling factor for frequencies (default = 1.0)
                         # NOTE: Scaling is applied to the frequencies after they are
                         #       calculated. SCALED frequencies will be stored in the
                         #       .hess file and printed in the output file.
                         #       In the .hess file you have accesss to the frequency
                         #       scaling factor (see below).


  # Flags to control NumFreq calculation:

  CentralDiff  true      # use central differences [f(x+h)-f(x-h)]/2h - or -
                         # use one-sided differences [f(x+h)-f(x)]/h
  Restart      false     # restart a (numerical) frequency calculation
  DX           0.005     # increment h
  Increment    0.005     # increment h
```

(continues on next page)

(continued from previous page)

```
  Hybrid_Hess  {...} end # calculate (numerical) Hybrid Hessian
  Partial_Hess [...] end # calculate (numerical) Partial Hessian


  # Flags to control subsequent vibrational analysis:

  QuasiRRHO       true   # Evaluate Vibrational Entropy with
                         # Quasi-Rigid Rotor Harmonic Oscillator
  QRRHORefFreq    25     # reference frequency used in the QuasiRRHO in cm-1
                         # default value is 100 from original paper
  CutOffFreq      1.0    # Threshold for frequencies to be considered
                         # in spectra, thermochemistry and printout (cm-1)
  Temp        298.15     # run the thermochemistry calculations at user defined
                         # temperatures (max 16 temperatures, separated by ',')
  T    290, 292, 295     # same as Temp

  XTBVPT2     True       # use XTB for the VPT2 correction of the IR
  Delq           0.5     # the displacement in dimensionless coordinates used
                         # during the VPT2

  TransInvar  True       # enforce translation invariance while calculating the Hessian?
  ProjectTR   True       # project out translation and rotation degrees of freedom
                         # in frequency calculation and thermochemistry analysis?



end
```

At present, analytical Hessians can be calculated for SCF only. However, there are some additional restrictions. Analytical Hessians cannot be performed for

- Double-Hybrid functionals
- RI-JK approximation

Here is what you would do, if you ran a frequency calculation and have a .hess file on disk and want to try different scaling factors for the frequencies

```
$frequency_scale_factor
  0.90      <<<---- you change this to whatever you want
```

```
orca_vib  myjob_scaled_freq.hess
```

The program will then read the Hessian, diagonalize it and apply your scaling factor. Whatever scaling factor was used in the actual input that generated the Hessian is irrelevant since the Hessian is re-diagonalized. To avoid confusion, we recommend that if the goal is to play with the scaling factor, then to leave the scaling factor in the input at 1.0. Nothing bad happens if you don't though.

7.27.1 Restarting Numerical Frequency calculations

To restart a numerical frequencies calculation, use:

```
%FREQ
  restart true
END
```

and ORCA will look for basename.res.{ } files in the same folder where the calculation is being run, check for what already has been done and restart where it is needed.

7.28 Intrinsic Reaction Coordinate

The Intrinsic Reaction Coordinate (IRC) method finds a path connecting a transition state (TS) with its downhill-nearest intermediates. The implementation in ORCA follows the method suggested by Morokuma and coworkers.[412]

The IRC method follows the gradient of the nuclear coordinates. As the gradient is negligible at a TS, first an initial displacement from the TS structure has to be carried out, based on the eigenmodes of the Hessian, in order to get to a region with nonnegligible gradient. For the initial displacement the eigenvector of the eigenmode with lowest frequency (hessMode=0) is normalized and then scaled by Scale_Displ_SD (which by default is chosen such that an energy change of Init_Displ_DE can be expected). Two initial displacements, forward and backward, are taken by adding the resulting displacement vector (multiplied with +1 and -1, respectively) to the initial structure. If the user requests the downhill direction (e.g. from a previous unconverged IRC run), it is assumed that the gradient is nonzero and thus no initial displacement is carried out.

After the initial displacement the iterations of the IRC method begin. Each iteration consists of two main steps, which each consist again of multiple SP and gradient runs:

1. Initial steepest descent (SD) step:
    1. The gradient (grad0) of the starting geometry (G0) is normalized, scaled by Scale_Displ_SD, and the resulting displacement vector (SD1) is applied to G0.
    2. Optional (if SD_ParabolicFit is true): If SD1 increases the energy, a linear search is taken along the direction of the displacement vector:
        1. The displacement vector SD1 is scaled by 0.5 (SD2 = 0.5 x SD1) and again added to G0.
        2. A parabolic fit for finding the displacement vector (SD3) which leads to minimal energy is carried out using the three SP energies (G0, geometry after SD1 and after SD2 step). SD3 has the same direction as SD1 and SD2, but can have a different length.
        3. The keyword Interpolate_only controls whether the length of SD3 has to be in between 0 and and the length of SD1. If that is the case, the maximum length is determined by SD1, the minimum length is zero.
    3. At the resulting geometry G1 (G0+SD1 or G0+SD3) the gradient is calculated (grad1).
2. Optional (if Do_SD_Corr is true): Correction to the steepest descent step:
    1. Based on grad0 and grad1 a vector is computed which represents a correction to the first SD (SD1 or SD3) step. This correction brings the geometry closer to the IRC.
    2. This vector is normalized, scaled by Scale_Displ_SD_Corr times the length of SD1 or SD3, and the resulting displacement vector (SDC1) is applied to G1.
    3. Optional (if SD_Corr_ParabolicFit is true):
        1. If the energy increases after applying step SDC1, SDC1 is scaled by 0.5 (SDC2 = 0.5 x SD1), if the energy decreases, SDC1 is scaled by 2 (SDC2 = 2 x SD1). SDC2 is then added to G1.

2. A parabolic fit for finding the displacement vector (SDC3) which leads to minimal energy is carried out using the three SP energies (G1, geometry after SDC1 and after SDC2 step). SDC3 has the same direction as SDC1 and SDC2, but can have a different length.
3. The keyword Interpolate_only controls whether the length of SDC3 has to be in between 0 and and the length of SDC1. If that is the case, the maximum length is determined by SDC1, the minimum length is zero.

4. At the resulting geometry G2 (G1+SDC1 or G1+SDC3) the gradient is calculated (grad2).

3. The gradient at the new geometry is checked for convergence.
4. Optional (if Adapt_Scale_Displ is true):
   1. If the resulting overall step size is smaller than 0.5 times Scale_Displ_SD, Scale_Displ_SD is multiplied by 0.5.
   2. If the resulting overall step size is larger than 2 times Scale_Displ_SD, Scale_Displ_SD is multiplied by 2.
   3. Scale_Displ_SD may not become smaller than 1/16 times the initial Scale_Displ_SD and not larger than 4 times the initial Scale_Displ_SD.

The following keywords are available:

```
! IRC
%irc
   MaxIter    20
   PrintLevel 1
   Direction  both # both - default
                   # forward
                   # backward
                   # down
 # Initial displacement
   InitHess   read # by default ORCA uses the Hessian from AnFreq or NumFreq, or
                   # computes a new one
                   # read    - reads the Hessian that is defined via Hess_Filename
                   # calc_anfreq  - computes the analytic Hessian
                   # calc_numfreq - computes the numeric Hessian
   Hess_Filename "h2o.hess"  # input Hessian for initial displacement, must be used
                             # together with InitHess = read
   hessMode   0  # Hessian mode that is used for the initial displacement. Default 0
   Init_Displ DE      # DE (default) - energy difference
                      # length       - step size
   Scale_Init_Displ 0.1 # step size for initial displacement from TS. Default 0.1 a.u.
   DE_Init_Displ    2.0 # energy difference that is expected for initial displacement
                        #  based on provided Hessian (Default: 2 mEh)
 # Steps
   Follow_CoordType cartesian # default and only option
   Scale_Displ_SD       0.15  # Scaling factor for scaling the 1st SD step
   Adapt_Scale_Displ    true  # modify Scale_Displ_SD when the step size becomes
                              # smaller or larger
   SD_ParabolicFit      true  # Do a parabolic fit for finding an optimal SD step
                              # length
   Interpolate_only     true  # Only allow interpolation for parabolic fit, not
                              # extrapolation
   Do_SD_Corr           true  # Apply a correction to the 1st SD step
   Scale_Displ_SD_Corr  0.333 # Scaling factor for scaling the correction step to
                              # the SD step. It is multiplied by the length of the
                              # final 1st SD step
   SD_Corr_ParabolicFit true  # Do a parabolic fit for finding an optimal correction
                              # step length
 # Convergence thresholds - similar to LooseOpt
   TolRMSG   5.e-4      # RMS gradient (a.u.)
   TolMaxG   2.e-3      # Max. element of gradient (a.u.)
```

(continues on next page)

(continued from previous page)

```
# Output options
  Monitor_Internals # Up to three internal coordinates can be defined
     {B 0 1}        # for which the values are printed during the IRC run.
     {B 1 5}        # Possible are (B)onds, (A)ngles, (D)ihedrals and (I)mpropers
  end
end
```

NOTE

- For direction=down (downhill) no initial displacement is necessary, and thus no Hessian is needed.

7.29 Nudged Elastic Band Method

The Nudged Elastic Band (NEB)[384, 416, 585] method is used to find a minimum energy path (MEP) connecting two local energy minima on the potential energy surface (PES) and thereby an estimate of the activation energy for the transition. The two minima are referred to as the reactant and product in the following discussion. The path can have one or more maxima, each one corresponding to a first order saddle point on the energy surface. The NEB method offers an advantage over eigenvector-following methods in that it is guaranteed to find saddle points that connect the given reactant and product states. The minimum energy path is often used to represent the reaction coordinate of the transition between the two states. The methods and implementation outlined below are discussed further in Ref. [4].

The user needs to specify the reactant and product configurations. The reactant energy minimum is inserted into the regular ORCA input file while the product should be in a separate .xyz file (keyword: *neb_end_xyzfile*). The reactant and product configurations should be optimized *a priori* by relaxing to energy minima on the PES, see section *Geometry Optimizations*. This can also be achieved via the 'preopt_ends' keyword. It is important to carefully prepare the reactant and product such that the position (or index) of the atoms is the same in the two configuration files, i.e. there should be one-to-one mapping between the reactant and product configurations.

The discretized path between the two minima is represented by a set of M system configurations that are referred to as images, i.e. $\mathbf{R} = [r_0, r_1, ..., r_{M-1}]$. The number of intermediate images between the end points (i.e. reactant and product) is specified by the user. The general rule of thumb is to include $M = 7 - 12$, or $5 - 10$ intermediate images per energy maximum in order to obtain a high enough resolution of the path and the saddle point. However, calculations can often converge and give accurate results with fewer images but complex paths with multiple maxima or long tails may require more images. During an NEB calculation the intermediate images are iteratively shifted towards the MEP using the component of the atomic force that is perpendicular to the current path as estimated from the tangent to the path at each image. While, the end point images are typically kept fixed. In each step of the iterative process, the energy and atomic forces of each intermediate image need to be computed. One of the main advantages of the NEB method is that the calculations of the images are carried out in parallel, where the electronic structure computations can be distributed over multiple processors (see discussion below for more details on the parallelization). While the CPU time is proportional to the number of images, the number of iterations needed for convergence to the MEP can become smaller when more images are included in the discrete representation of the path.

The tangent to the path at each image, $\hat{\tau}_i$, can be estimated in two ways (keyword: *tangent*), either by the original method [416] or by the more numerically stable improved [384] estimate (default option). In the former, the tangent at an image is taken to be a linear combination of the two line segments that connect to image i,

$$\tau_i = \frac{r_{i+1} - r_i}{|r_{i+1} - r_i|} + \frac{r_i - r_{i-1}}{|r_i - r_{i-1}|}$$

In the improved tangent estimate, the line segment from the current image to the adjacent image with higher energy is used, i.e.

$$\tau_i = \begin{cases} \tau^+ & \text{if } \mathrm{E}_{i+1} > \mathrm{E}_i > \mathrm{E}_{i-1} \\ \tau^- & \text{if } \mathrm{E}_{i+1} < \mathrm{E}_i < \mathrm{E}_{i-1} \end{cases}$$

where $\tau^+ = r_{i+1} - r_i$ and $\tau^- = r_i - r_{i-1}$. With the exception of image i being at an energy extremum along the path, then an energy-weighted average of the two line segments to adjacent images is used,

$$\tau_i = \begin{cases} \tau^+ \Delta V_i^{\max} + \tau^- \Delta V_i^{\min} & \text{if } \mathrm{E}_{i+1} > \mathrm{E}_{i-1} \\ \tau^+ \Delta V_i^{\min} + \tau^- \Delta V_i^{\max} & \text{if } \mathrm{E}_{i+1} < \mathrm{E}_{i-1} \end{cases}$$

where

$$\Delta V_i^{\max} = \max(|E_{i+1} - E_i|, |E_{i-1} - E_i|)$$
$$\Delta V_i^{\max} = \min(|E_{i+1} - E_i|, |E_{i-1} - E_i|)$$

Then the tangent is normalized, $\hat{\tau}_i = \tau_i / |\tau_i|$. With an accurate estimate of the tangent, the perpendicular component of the atom force is computed by,

$$F_i^{\perp} = F_i - (F_i \cdot \hat{\tau}_i)\, \hat{\tau}_i$$

In the free-end NEB method, the endpoint images are optimized simultaneously along with the intermediate images, i.e. M movable images are used. Three different variants of the free-end NEB method (keywords:*free_end*, *free_end_type*) are included in ORCA: where (i) the endpoints are restrained to move along the same (or separate) energy isocontour[5, 920], (ii) according to the atomic force acting perpendicular to the path and (iii) the full atomic force. For the first variant, the 'reactant' and 'product' endpont images, can be restrained to move along two different energy contours, E_0 (given by keyword:*free_end_ec*) and E_1 (given by keyword: *free_end_ec_end*), to keep the path bounded. Then, if the images drift away from the isocontour because of curvature of the energy surface, an harmonic restraint term is used to pull the image back to the contour.[5] The stiffness of the harmonic restraint is given by a user-supplied parameter (keyword:*free_end_kappa*). It is important to carefully select the energy values and the strength of the harmonic restraint, otherwise the path may become kinked. For the second variant, the endpoint images are unbounded but displaced directly towards the MEP. This is typically acceptable if the endpoint images are already in vicinity to the MEP and is less prone to kinks developing along the path. For the third variant, the endpoint images become optimized to the reactant and product energy minimum.

7.29.1 Spring forces

In order to control the distribution of the images along the path, spring forces are included to act between adjacent images, tangential to the path[416],

$$F_i^{\mathrm{sp},\parallel} = \left(k_i^{\mathrm{sp}} |r_{i+1} - r_i| - k_{i-1}^{\mathrm{sp}} |r_i - r_{i-1}|\right) \hat{\tau}_i$$

The magnitude of the spring forces (i.e. stiffness) is controlled by spring constants, k^{sp}. The typical values of the spring constant (keyword: *springconst*) can be taken to be within the range of 0.01 Eh/Bohr^2 to 1.0 Eh/Bohr^2. If the spring constants are choosen to be the same for all pairs of adjacent images, the images will be equally distributed along the path. However, it is also possible to choose energy-weighted spring constants (keyword: *energy_weighted*) so as to increase the density of images in the higher energy regions[4, 385]. In an energy-weighted NEB calculation the spring constants are scaled according to the relative energy of the images, from a lower-bound value (keyword:*springconst*) to an upper-bound value (keyword:*springconst2*), by

$$k_i^{\mathrm{sp}} = \begin{cases} (1 - \alpha_i) k_{\mathrm{u}} + \alpha_i k_{\mathrm{l}}, & \text{if} E_i > E_{\mathrm{ref}} \\ k_l, & \text{otherwise} \end{cases}$$
$$\alpha_i = \frac{E_{\max} - E_i}{E_{\max} - E_{\mathrm{ref}}}$$

where k_{u} and k_{l} are the upper- and lower-bound values for the spring constant. $E_{\max}$ is the current estimate of the maximum energy along the path, E_i is the higher energy image of the pair of images connected by line segment i and E_{ref} is the energy of the higher energy minimum. The energy-weighted springs will typically serve to improve the tangent estimate in the barrier region and hence stabilize the calculations. The inclusion of energy-weighted springs can be important in reactions where the energy barrier is narrow and/or the pathway is characterized by a long 'energy tail', e.g., in rearrangements or dissociation reactions. The choice of spring constants will affect the behavior of a calculation, especially the number of iterations needed to reach convergence. Other formulations for spring forces are also available since ORCA 4.2 (keyword: *springtype*). These are referred to as the original[416] and ideal[553] spring forces. The original spring forces are estimated by a spring acting on each degree of freedom

between a pair of adjacent images. For ideal springs each image is assigned an ideal position along the path based on a linear interpolation of the current location of the images and the individual images interact with the ideal locations via spring forces. The ideal springs are currently not implemented to work with energy weighted spring constants.

While only the component of the spring force parallel to the path is included in an NEB calculation (by default) the user can choose to include a fraction of the spring force acting perpendicular to the path to stabilize calculations (keyword: *perpspring*. Since, the perpendicular component of the spring force can serve to straighten out the path and is useful for complex pathways with multiple energy extrema and low resolution of the path. The inclusion of the perpendicular component of the spring force is always accompanied by a switching function that is used to scale it according to (i) convergence to the MEP (the 'tan' function)[791] (ii) the angle between adjacent images (the 'cos' function)[416] or a combination of both. The inclusion of the perpendicular spring force can help to reduce the number of iterations and eliminate kinks on the path. However, it is important to stress that by inclusion of the perpendicular spring forces, the images may not converge on the MEP. Alternatively, the user can choose to use the modified DNEB method[791, 853].

7.29.2 Optimization and convergence of the NEB method

The effective force used in a standard NEB calculations is the sum of the atomic force component perpendicular to the path and the spring force component parallel to the path,

$$F_i^{\mathrm{NEB}} = F_i^{\perp} + F_i^{\mathrm{sp},\parallel}$$

The path can be brought to the MEP by moving according to the effective force, as is shown in Fig. 7.21.

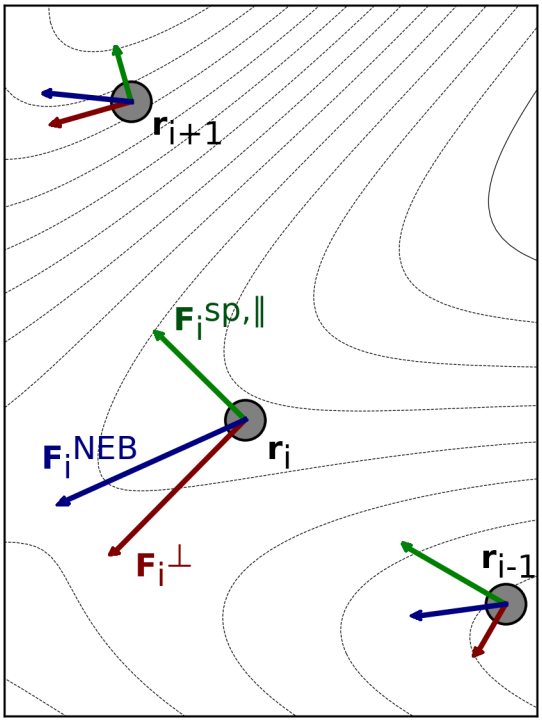


Fig. 7.21: Visualization of the effective force, F^{NEB} and its two components: $F_i^{\perp}$ and $F_i^{\mathrm{sp},\parallel}$ for three images along an intermediate path in a NEB optimization. The figure is taken from Ref. [3]

This can be achieved by using any of the three optimization methods implemented (keyword:*opt_method*): velocity projection optimization (VPO)[416], fast inertial relxation engine (FIRE)[108] and L-BFGS[635]. VPO and FIRE are more robust for regions that are far from the MEP, while L-BFGS converges faster when the images are close to the MEP. FIRE and VPO both have a local and global implementation (keyword:*local*). In the local variant, all of the images are treated individually when taking an optimization step, while in the latter the whole band is treated as a single point. A constant 'trust-radius' is used for all optimization methods, where if the magnitude of

the maximum Cartesian component of an optimization step exceeds a user-supplied threshold (keyword:*maxmove*), the whole displacement is scaled down accordingly. The number of steps stored in the L-BFGS optimization for the construction of the approximate Hessian matrix can be adjusted by the user (keyword:*lbfgs_memory*). For a more conservative optimization with L-BFGS, the memory can also be erased if a large step is attempted (keyword:*lbfgs_restart_on_maxmove*).

The convergence of the intermediate images is gauged from the maximum Cartesian component of the perpendicular atom force as well as the root-mean-square, i.e. $\max(|\mathbf{F}^{\perp}|), \mathrm{RMS}(\mathbf{F}^{\perp})$. The atomic force on the images perpendicular to the path vanishes as the images are located on the MEP. A typical value of the tolerance for the maximum component of the atomic force perpendicular to the path is $1 \cdot 10^{-3}$ Eh/Bohr. Typically the tolerance for the root-mean-square value is chosen to be smaller by a factor of 1/2 or 1/3. Sometimes a tighter tolerance for the maximum component of the force is needed, for example $5 \cdot 10^{-4}$ Eh/Bohr or even $2 \cdot 10^{-4}$ Eh/Bohr. The maximum number of optimization steps allowed is set by the keyword 'maxiter'.

The configuration of each image after each iteration is written to a '_trj.xyz' file (see file: basename_MEP_ALL_trj.xyz). This file is useful for troubleshooting non-convergent calculations.

7.29.3 Climbing image NEB

In order make the highest energy image converge more accurately to the (highest) energy maximum along the MEP, the climbing image variant of the NEB method (CI-NEB) can be used. [385] In the CI-NEB method, the effective force F^{NEB} acting on the climbing image (i.e. $i = \mathrm{ci}$) is transformed to:

$$F_{\mathrm{ci}}^{\mathrm{NEB}} = F_{\mathrm{ci}} - 2\left(F_{\mathrm{ci}} \cdot \hat{\tau}_{\mathrm{ci}}\right)\hat{\tau}_{\mathrm{ci}}$$

where the climbing image is pushed up-hill along the path and relaxed down-hill perpendicular to the path. That is, the energy is maximized with respect to one degree of freedom corresponding to the direction of the tangent while the energy is minimized with respect to all other degrees of freedom. The effective force on the climbing image does not include any spring force and the density of images then becomes different on either side of the climbing image. As long as the tangent estimate is accurate enough the climbing image will converge rigorously to the point of highest energy along the path. An illustration of how the climbing image NEB method works is shown in Fig. 7.22 for a simple two-dimensional energy function.

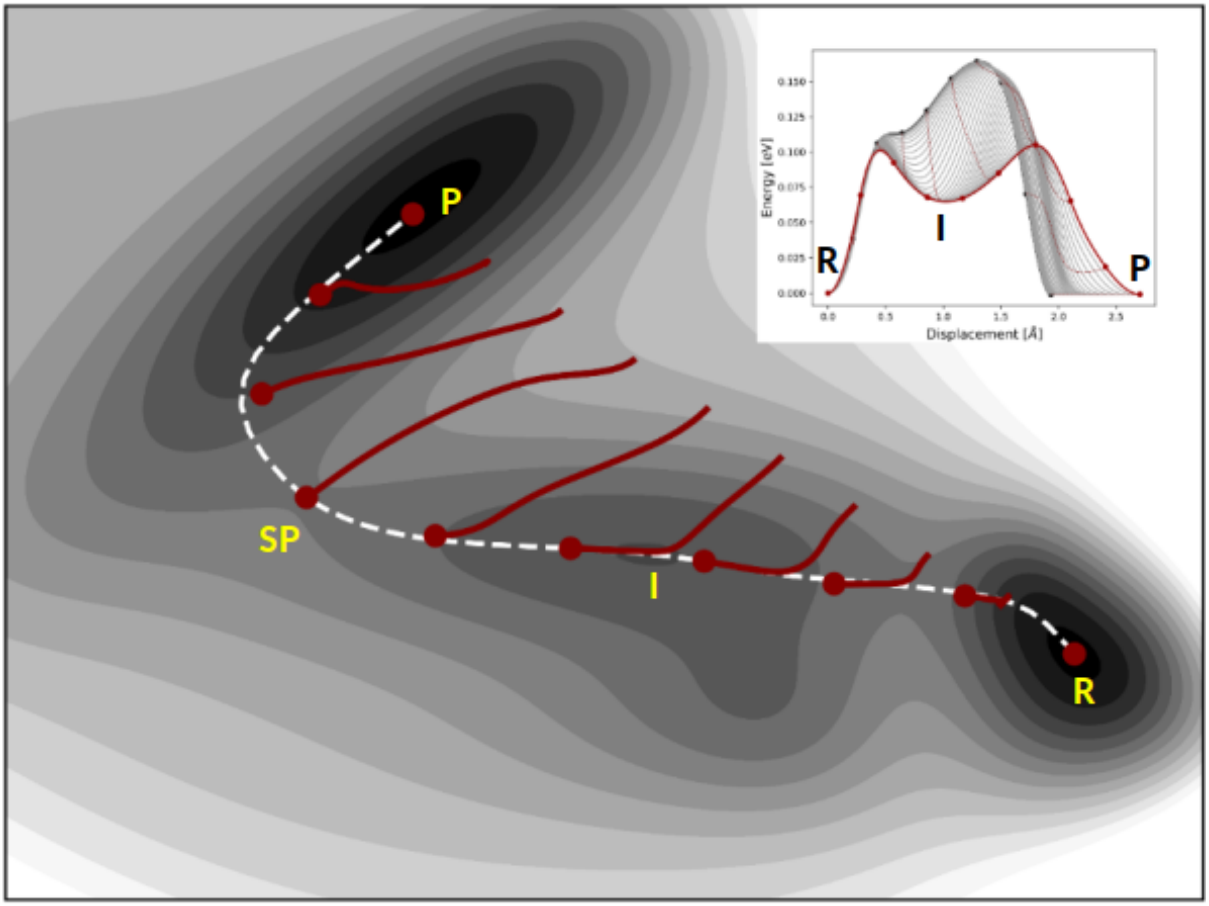


Fig. 7.22: Illustration of how the CI-NEB method works on a two-dimensional Müller-Brown energy surface, $E(x, y)$.[539] The calculation is started from a linear interpolation between the reactant (**R**) and product (**P**) energy minima, using $M = 10$. The images are displaced in the orthogonal direction to the path (red curves), until they converge to the minimum energy path (white dashed curve). The climbing image accurately locates the higher energy first order saddle point along the path (denoted by **SP**). The optimization profile is shown as an inset. In such a plot the interpolated energy along the path is plotted as a function of displacement, for each (or selected) optimization step. The figure is taken from Ref. [3]

It can be useful to start the climbing image after the magnitude of the atomic forces perpendicular to the path drops below a given user supplied threshold (keyword: *tol_turn_on_ci*). Then, the highest energy image along the current path is converted to a climbing image. It is usually most efficient to initiate the climbing image early on or even from the start of the NEB calculation. This applies when using the VPO optimization method. For L-BFGS it is recommended to start CI-NEB when the path has partially converged to the MEP, e.g., around 0.01-0.02 Eh/Bohr. If there are two or more maxima in the energy along the MEP it is possible that the image near the highest maximum is not chosen as the climbing image at an early stage of the NEB calculation. Then, later on the choice of the climbing image can be switched automatically. Also, for barrierless reactions, the climbing image is not turned on. The atom coordinates of the climbing image (in a CI-NEB calculation) or the highest energy image (in an NEB calculation) are written to files '_NEB-CI_converged.xyz' and '_NEB-HEI_converged.xyz', respectively, when a calculation has successfully completed.

The convergence of a CI-NEB calculation can either be gauged by monitoring the forces on all images or only on the climbing image, (keyword: *convtype*). The latter option may be used when the objective of the calculation is to locate the highest energy saddle point connecting the reactant and product states and can save significant computational effort. To gauge for convergence on the climbing image, both the root-mean-square and magnitude of the maximum (Cartesian) component of the atom force are monitored, i.e. $\max(|\mathbf{F}_{\mathrm{ci}}|), \mathrm{RMS}(\mathbf{F}_{\mathrm{ci}})$. When gauging the convergence of all images in a CI-NEB calculation it is typically acceptable to converge the regular images more loosely than the climbing image. By default, the tolerance for the regular images is a factor of 10 larger than that of the climbing image. This scaling of the tolerances is a parameter that can be set by the user (keyword: *tol_scale*). Typically for an acceptable convergence to the saddle point, the tolerance threshold for the maximum magnitude

of an atomic force component of the climbing image can be set to $5 \cdot 10^{-4}$ Eh/Bohr.

7.29.4 Generation of the initial path

One of the most important aspects of any NEB or CI-NEB calculation is the generation of the initial path between the reactant and product energy minima (keyword: *interpolation*. The recommended method is the image-dependent pair potential (IDPP) method [807]. An alternative and simpler method is linear interpolation in Cartesian coordinates. In either case, the user should always inspect the initial path (see file: basename_initial_path_trj.xyz) to make certain that it is acceptable.

By default, the endpoint structures are scanned for covalently linked fragments. If two fragments are detected in either of the endpoint structures, the structure is automatically prepared such that the distance between the fragments is no larger than a maximum distance defaulted to 3.5 Angstrom.

The linear interpolation in Cartesian coordinates may result in overlap of atoms leading to large, initial, atomic forces and even divergence in the SCF cycles. The IDPP method solves these problems by interpolating pairwise distances between neighboring atoms in the reactant and product energy minimum. Then, a path is generated to match those distances as closely as possible. Since there are many more pairwise distances than atom coordinates, the initial path is found by minimizing the objective function,

$$S_{\kappa}^{\text{IDPP}}(r_i) = \sum_{A}^{M} \sum_{B>A}^{M} w(d_{AB}) \left(d_{AB}^{\kappa} - d_{AB}\right)^2 \tag{7.218}$$

where d_{AB} is the pairwise distance between atoms A and B for intermediate image i. d_{AB}^{κ} is the ideal interpolated distance between atoms A and B of the same image. w is a weight-function to give shorter bond distances more weight and make unnecessary bond-breaking unfavorable. The weight function is given as $w = (d_{\text{AB}})^{-4}$. The IDPP path avoids the overlap of atoms and can also generate a path that is closer to the MEP than the linear interpolation in Cartesian coordinates[807]. The IDPP path is obtained from an NEB calculation using the IDPP objective function starting from a linear interpolation of the Cartesian coordinates between the endpoint structures, but this calculation requires little computational effort since it does not require any electronic structure computations. Note that it is possible that the initial path breaks covalent bonds and that it can be far from the optimal MEP, so the user should always inspect the initial path before starting an NEB calculation. The user can adjust the settings of the IDPP calculations using the 'idpp' related keywords, but the default values should suffice for most applications. Note that the units of the IDPP are in Ångströms instead of atomic units. The matrix of ideal interpolated distance between the atoms A and B at image i, $d_{AB,\,i}^{\kappa}$, is most simply obtained by linear interpolation as

$$d_{AB,\,i}^{\kappa} = d_{AB}^{initial} + P_i * \left(d_{AB}^{final} - d_{AB}^{initial}\right) ,$$

where $P_i = \frac{i}{n_{im}-1}$ is the position of image i on the path between 0, corresponding to the initial image, and 1, corresponding to the final image, and n_{im} is the total number of images.

As this linear distance interpolation may not be ideal, one can alternatively choose a bilinear distance interpolation scheme that takes atomic bonds into account. Bonds between the atoms A and B in the initial and final images are identified if their distance $d_{AB} < t_{bond} = t_c \left(R_A^{cov} + R_B^{cov}\right)$, where t_c is an input factor impacting at what distances a bond is detected by default taken to be 1.2, and R_X^{cov} is the covalent radius of atom X. If a bond is detected in one of the endpoint structures but not the other, two linear distance interpolations are performed to obtain $d_{AB,i}^{\kappa}$. The first interpolates the distance between the endpoint structure in which the atoms are bonded and t_{bond}, the second interpolates the distance between t_{bond} and the endpoint structure in which the bond is broken. The two linear distance interpolations are joined at the path position P_{join} defaulted to 0.5, corresponding to the midpoint on the path between the endpoint images. If the atoms are bonded in the initial structure but not in the final structure, the interpolation is

$$d_{AB,\,i}^{\kappa} = \begin{cases} d_{AB}^{initial} + \frac{P_i\left(t_{bond} - d_{AB}^{initial}\right)}{P_{join}} , & \text{if } P_i < P_{join} \\ d_{AB}^{final} - \frac{(1-P_i)\left(d_{AB}^{final} - t_{bond}\right)}{(1.0-P_{join})} , & \text{if } P_i \geq P_{join} . \end{cases}$$

If the atoms are bonded in the final structure but not in the initial structure, the interpolation is

$$d_{AB,\,i}^{\kappa} = \begin{cases} d_{AB}^{initial} - \frac{P_i\left(d_{AB}^{initial} - t_{bond}\right)}{1-P_{join}} , & \text{if } P_i < 1 - P_{join} \\ d_{AB}^{final} + \frac{(1-P_i)\left(t_{bond} - d_{AB}^{final}\right)}{P_{join}} , & \text{if } P_i \geq 1 - P_{join} . \end{cases}$$

This bilinear interpolation leads to a slower increase of the atom distance in the bonded region compared to the non-bonded region in most cases, emphesizing the bond breaking process and thus the region where the energy is changed most by a change in the bond distance. If the change in the bond distance in the non-bonded region is slower than the change in the bonded region, the implementation falls back to the regular linear distance interpolation.

In challenging cases, even the IDPP NEB calculation starting from a linear interpolation of the Cartesian coordinates between the endpoint structures may not provide a reasonable initial path. This may happen when the linear interpolation path has atoms come very close to each other in intermediate images, which leads to bond breaking even after NEB optimization using the IDPP objective function. In such cases, a reasonable initial path can often be obtained by avoiding the linear interpolation by sequentially adding images to the path starting from both endpoint structures instead. This method is referred to as sequential IDPP (S-IDPP).[765] The two sets of images close to each endpoint structure are separated by a larger distance than the images in each set by choosing a smaller spring constant than between the images in each set where equal spring constants are used. The images closest to the center are converged in an NEB calculation and an image added at an appropriate distance from the converged image between the two images closest to the center. This process is repeated until the requested amount of images has been added. The tangent of the images closest to the center follows the improved tangent definition and always treats them as extrema on the path, i.e. an average of the normalized distance vectors to the adjacent images is used. The number of images used to form the S-IDPP initial path can be different from the number of images used in the subsequent NEB calculation involving electronic structure calculations. It can be benficial to use an even number of images in the S-IDPP computation since no image is placed at the center of the initial path then and moving to the requested odd number of images in an additional IDPP NEB calculation afterwards. This option is used by default, but can be deactivated. For systems involving very heavy rotations of large groups, the method becomes more robust when twice the amount of images is used in the S-IDPP calculation. Half the images are then removed automatically for the subsequent NEB calculation involving electronic structure calculations. This option is available, but not used by default.

The user may have a preconceived notion of the saddle point configuration or have an estimate of the path from a calculation carried out at a lower level of theory. The initial path can be generated in such a way as to include an intermediate configuration as one of the images using the 'NEB_TS_XYZFile' keyword. Since this image will be optimized along with the other intermediate images during the NEB calculation the guess does not have to be accurate.

If inspection of the initial path reveals problems, e.g., unnecessary bond breaking, it is often a good idea to insert a reasonable configuration into the initial interpolation to avoid such problems. Moreover, if an NEB calculation is unable to converge to the MEP (or saddle point) with the given maximum number of iterations, the user can restart the calculation from the 'allxyz' file (see file: basename_MEP.allxyz) which is written to the disk after each iteration during the optimization. Note, when starting an NEB calculation from an output from a previous CI-NEB calculation and vice-versa the band may require a few iterations to adjust the distribution of images to achieve the desired distribution, depending on the selected spring type and the choice of NEB method.

If the system can be modeled reasonably well using the GFN-xTB method, another possible choice is the generation of an initial path on XTB level (keywords 'XTB0', 'XTB1' or 'XTB2' for GFN0-xTB[698] GFN-xTB[332] or GFN2-xTB[70]). In this case the initial path on IDPP level is refined using an NEB calculation on the chosen XTB level. If this NEB run is successful, the entire MEP on XTB level is used as the initial path. If the NEB run on XTB level is not successful, the initial path on IDPP level is used instead.

Another keyword that makes use of the XTB method is the 'XTBTS' keyword ('XTB0TS', 'XTB1TS' or 'XTB2TS'). In this case the initial path on IDPP level is refined using an NEB-CI calculation on XTB level. If the NEB-CI run is successful, the resulting CI structure is chosen as TS guess structure, and the final initial path is generated using an IDPP path from reactant to TS guess and from TS guess to product. If the NEB-CI run on XTB level is not successful, the initial path on IDPP level is used instead.

7.29.5 Removal of translational and rotational degrees of freedom

For NEB and CI-NEB calculations of molecular systems it is important to project out the six (or five in the case of linear molecules) degrees of freedom corresponding to the global rotation and translation of the system. This can be done either at the start of a calculation or for each optimization step (keyword: *quatern*). For the latter, the center-of-geometry of each image is translated to origin and a rotational matrix is constructed using the quaternion approach[576] to minimize the root-mean-square deviation between any two adjacent images. Depending on whether a fixed center is used (keyword: *fix_center*), the images are either kept in that place or transferred back to their original position. This procedure is repeated for any pair of adjacent images in each step of the optimization. Also, the net effective NEB force is set to zero (keyword: *remove_extern_force*).[434]

7.29.6 Reparametrization of the path

In some cases it may be beneficial to enable redistribution of the images along the path every N iterations (i.e. analogous to the string method[240]) (keyword: *reparam*). The path is then interpolated using either a linear or cubic polynomial fitted to both the coordinates and the tangent to the path, and the images are redistributed evenly along the interpolated path. Both N and the type of interpolation are specified by the user (keyword: *reparam_type*). The cubic interpolation method should be better in calculations where the resolution of the path and hence the estimate of the tangent is good, while the linear interpolation is generally more robust as it does not as dependent on the tangent as much.

7.29.7 Useful output

After each iteration, the energy profile along the path is obtained by making a piecewise cubic polynomial interpolation using both the energy and the component of the atom force tangentinal to the path[384]. The interpolation can reveal important information about the MEP in locations between the intermediate images. The interpolation is written to the file 'interp' (see file: basename.interp) in each step of the optimization. Moreover, as NEB and CI-NEB calculations can be quite computationally demanding and in order to properly analyze what may have gone wrong in such a calculation, the user can inspect the '.log' file, which includes information about the type of calculation, energy, length of the path, spring forces, atomic forces etc.

7.29.8 Important warning messages

Some tests are carried out during the optimization in order to detect problems on the fly. The angle between the two straight lines going through an image and its neighbors on each side is calculated. If the angle becomes large e.g. exceeding 90° the estimate of the tangent has likely become inaccurate and a better resolution of the path is required. If the angle is close to 180° the ordering of the images may have become incorrect. Especially in the latter case, it may be a good choice of action to terminate the calculation and include a larger number of intermediate images in the subsequent calculation. Some information from the calculation, e.g, a guess for the saddle point, could be incorporated in the new calculation.

Another issue is the identification of an intermediate minimum along the path. If an intermediate minimum is observed in M subsequent iterations (M is supplied by the user) a warning is issued to the user that a possible intermediate minimum may exist along the path. It is a good idea to check the status of the calculation, the path and the convergence behaviour. If the calculation appears to be proceeding normally and heading for convergence, the best course of action is to allow the calculation to finish. However, it is in general better to carry out separate NEB calculations for segments of the MEP on either side of the intermediate minimum. Especially, if the intermediate minimum is deep w.r.t to the reactant and product state energy minima.

In such cases, the image closest to the apparent intermediate minimum is selected and a structural minimization carried out. The resulting configuration is then used as the initial state or final state in the subsequent CI-NEB calculations.

7.29.9 Parallel execution

If the number of processes (NProcs) specified in the input is larger than 1, NEB will automatically start up in multi-processes mode:

NProcs <= NImages
: NProcs processes will handle NProcs images independently with 1 process per image. Choose NProcs = X*NImages (e.g. X = 1 or 0.5)

NProcs > NImages
: NProcs processes will handle NImage images, each image being treated by (NProcs/ NImages) processes. If you want to dedicate more than 1 process to each single image-calculation, choose NProcs = X*NImages (e.g. X = 2, 3, 4, ...).

Note: If in the second case multiple compute-nodes are involved, the user will need to define the ORCA specific environment variable RSH_COMMAND, which tells the NEB driver how to connect to the individual nodes (set it to either 'rsh' or 'ssh'). However, this may not work with all queueing systems.

If the energy and force calculations are fast (e.g. with semiempirical methods), there is no gain in using multiple processes per image. Starting up and finalizing MPI may consume more time than the gain from parallel processing.

7.29.10 zoomNEB

A preliminary version of the Zoom-NEB (Z-NEB) method has been included this implementation, where the objective of the method is to locate a saddle point more accurately with a better resolution compared to CI-NEB calculations. The Z-NEB method is an automatic two step procedure, where in the first step a CI-NEB calculation is carried out to obtain a rough convergence towards the MEP. Then, the region surrounding the highest energy maximum along the path is identified and a new set of images distributed along this region. In the second step, a free-end CI-NEB calculation is started from this new path. In this calculation, the endpoints are optimized according to the atom force acting perpendicular to the path. This will ensure that the endpoints of the second CI-NEB calculation will converge to the MEP, as well. Furthermore, to maintain the parallelization of the CI-NEB method, same number of movable images are used in the first and second CI-NEB calculations.

7.29.11 NEB-TS

Probably the most efficient saddle point search methods are obtained when double ended methods (e.g. NEB) are combined with single ended methods (e.g., eigenvector-following). In the current implementation, a combination of EW-CI-NEB and EF P-RFO (eigenvector-following partitioned rational function optimization) methods is presented and referred to as the NEB-TS method [4].

In NEB-TS, the EW-CI-NEB method is used to partially converge to the MEP and hence saddle point, i.e., the optimization of the images along the MEP is halted once the climbing image is converged to a prescribed tolerance. Then, the climbing image is used to provide a subsequent eigenvector-following calculation with an accurate initial guess configuration, as is shown in Fig. 7.23 and the tangent at the CI is used to select the eigenvector to be followed. The tangent estimate should already provide an accurate approximation to the unstable mode at the saddle point. The initial Hessian matrix needed for the eigenvector-following calculation can either be computed analytically (if available) or taken as the Almlöf empirical Hessian matrix [264]. If the Almlöf Hessian matrix is used, the curvature at the CI is estimated by using a finite difference approximation (i.e. using the atom force acting on the neighboring images to CI) and used to scale the corresponding eigenvalue of the selected eigenvector, allowing the eigenvector-following calculation to be started from a Hessian matrix of correct form. The typical '%geom' block can be used to modify the settings of the eigenvector-following calculation.

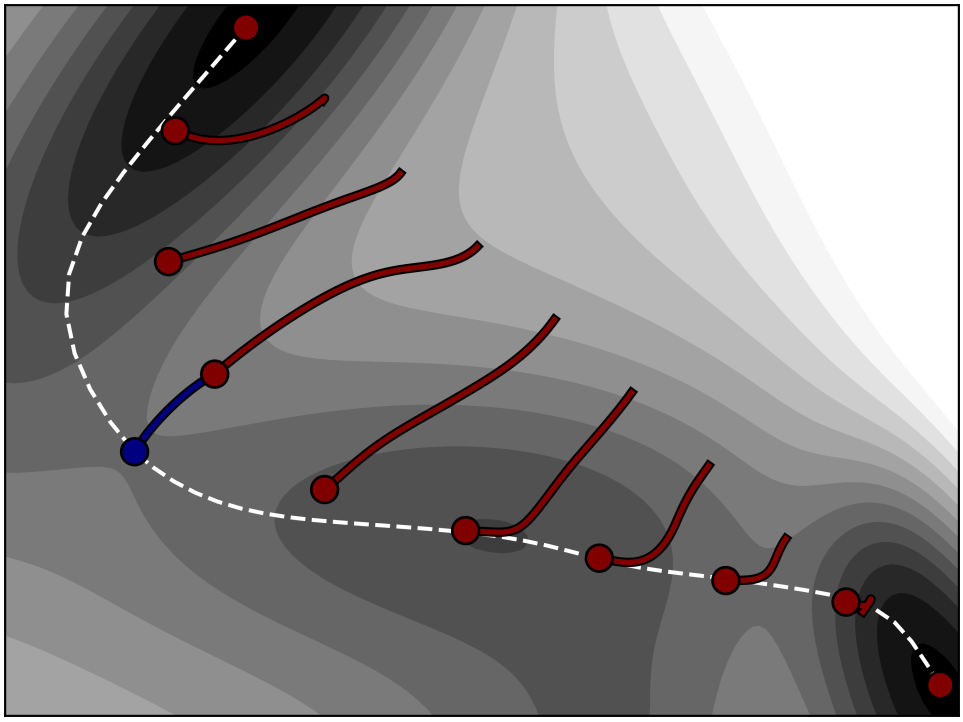


Fig. 7.23: Illustration of how the NEB-TS method works on a two-dimensional Müller-Brown energy surface, $E(x, y)$ [539]. The calculation is started from a linear interpolation between the reactant and product energy minima, using $M = 10$. The images are displaced in the orthogonal direction to the path (red curves) using the CI-NEB algorithm, until a rough convergence to the minimum energy path (white dashed curve) is obtained. The climbing image then provides an approximate saddle point configuration that can be used to start eigenvector-following partitioned rational function optimization to accurately (and swiftly) identify the (higher energy) first order saddle point. The figure is taken from Ref. [4]

7.29.12 FAST-NEB-TS and LOOSE-NEB-TS

Since our first NEB-TS implementation, we investigated a lot more settings and variants, see [4]. Based on those findings, two new algorithms and convergence threshold settings have been implemented into ORCA. FAST-NEB-TS corresponds to the IDPP-TS in the paper, and LOOSE-NEB-TS corresponds to the actual NEB-TS defaults, which are defined in the paper. Both features, FAST- and LOOSE-NEB-TS, show slightly lower robustness, but need significantly less NEB cycles.

7.29.13 NEB / NEB-TS and TD-DFT

The NEB and NEB-TS algorithm now also works in combination with TD-DFT. The input:

```
! NEB-TS
%neb
 product "product.xyz"
end
%tddft
 NRoots 1
 IRoot 1
end
```

not only computes the MEP and TS of the first excited state, but it also prints out (after NEB convergence) the excited state as well as ground state energies over the MEP:

```
----------------------------------------
Image    E0      Root  1
```

(continues on next page)

(continued from previous page)

```
----------------------------------------
  0   -67.203     0.000
  1   -66.443     1.517
  2   -54.382     5.774
  3   -42.483    10.473
  4   -32.798    14.323
  5   -28.056    16.108
  6   -34.102    13.915
  7   -48.125     8.249
  8   -64.188     2.212
  9   -67.066     0.018
```

You can even request an NEB-TS calculation on the ground state PES, and at the same time gain information on the excited PES via the input:

```
! NEB-TS
%neb
 product "product.xyz"
end
%tddft
 NRoots 1
 IRoot 0
end
```

During this NEB calculation, only ground state energies and gradients are computed. Only after NEB convergence, the additional excited state energies are computed on each image, in order to yield the ordering of the states on the MEP:

```
----------------------------------------
Image    E0      Root  1
----------------------------------------
  0     0.000   121.545
  1     6.281   128.198
  2    21.434   143.835
  3    35.437   151.581
  4    43.223   155.579
  5    43.235   155.501
  6    35.472   151.495
  7    21.498   143.715
  8     6.395   128.060
  9     0.166   121.425
```

7.29.14 Summary of Keywords

The following keywords are available:

```
! NEB           # NEB calculation
! NEB-CI        # Climbing Image NEB calculation
! NEB-TS        # NEB calculation plus subsequent TS optimization
! FAST-NEB-TS   # NEB calculation with one iteration only plus subsequent TS opt.
! LOOSE-NEB-TS  # NEB calculation with default convergence criteria from NEB-TS paper
! TIGHT-NEB-TS  # NEB calculation with tighter convergence criteria plus
                # subsequent TS optimization
! ZOOM-NEB      # NEB calculation plus zoomed NEB calculation
! ZOOM-NEB-CI   # Climbing Image NEB plus zoomed climbing image NEB calculation
! ZOOM-NEB-TS   # NEB calculation plus zoomed NEB calculation plus subsequent
                #  TS optimization
! NEB-IDPP      # IDPP (Initial Path) NEB calculation - for estimation of path
                #  length
```

(continues on next page)

(continued from previous page)

```
%neb
   Product "product.xyz" # product structure. Input is mandatory.
   NImages  8         # default 8. Number of images without fixed endpoints,
                      #  for free_end total number of images
   PrintLevel   1     # default 1. Normal printout. Use 0 for no printout, higher
                      #  numbers (<=4) for more detailed printout.

   TS "TSGuess.xyz"   # Provide guess for the TS structure. Images
                                 #  are interpolated between reactant and TS guess
                                 #  and between TS guess and product.
   NEB_TS_Image     3 # default -1. Number of the image the TS guess is used for.
                      #  If not defined (=-1), the image which gives lowest RMSD
                      #  for all image distances is used.

 # Restart option: After each iteration the NEB method stores all image
 # structures in an .allxyz file. In case of an abort this file can be used
 # for a restart. File should contain the structures for all images.
   Restart_ALLXYZFile "NEB1.allxyz"# use the trajectory from file if filename is
                                   #  provided

 # Alternatively NEB can be started on user prepared wavefunctions for each image.
 # The names of of these wavefunction files should consist of a user-chosen basename
 # and the extension '_imN.gbw', where N is the image number.
 # The basename should be provided in the input, ORCA will add extension '_imN.gbw'
   Restart_GBW_BaseName "NEB2"     # use the wavefunctions from file NEB2_imN.gbw

 # Check SCF convergence: If true, SCF convergence is checked for and
 # calculation aborts if:
 # -any of the images does not show SCF convergence in four subsequent cycles.
 # -any of the images does not show SCF convergence in two subsequent cycles
 #  after the gradient is converged.
   CheckSCFConv     true           # default true

 # PDB file input format:
   Product_PDBFile  "product.pdb"  # Product structure in pdb format. If this is
                                   #  given, xyz does not need to be given.
   TS_PDBFile       "TSGuess.pdb"  # TS guess structure in pdb format. If this is
                                   #  given, xyz does not need to be given.
   Free_End         false          # Use free-end NEB. In this case the NImages
                                   #  corresponds to the total number of images.
   PreOpt           false          # do optimization of reactant and product in
                                   #  internal coordinates before NEB starts
   NSteps_FoundIntermediate  30    # Number of steps the intermediate has to be
                                   #  present
   AbortIf_FoundIntermediate false # If an intermediate is found abort the run.
   NPTS_Interpol    10             # Number of abscissa in cubic polynomial
                                   #  interpolation
   Interpolation    IDPP           # Method to generate the images based on the
                                   #  reactant, product (and potentially TS guess)
                                   # linear
                                   # IDPP
                                   # XTB0TS  - TS on GFN0-xTB level
                                   # XTB0    - entire path on GFN0-xTB level
                                   # XTB1TS  - TS on GFN1-xTB level
                                   # XTB1    - entire path on GFN1-xTB level
                                   # XTB2TS   - TS on GFN2-xTB level
                                   # XTB2     - entire path on GFN2-xTB level
   Prepare_Frags    true           # Analyze endpoint structures for fragments.
                                   #  If two fragments are detected in an
                                   #  endpoint structure, reduce distance to
                                   #  maximum distance given by Max_Frag_Dist
```

(continues on next page)

(continued from previous page)

```
 Max_Frag_Dist    3.5            # Maximum allowed fragment distance. If they
                                 #  are farther apart, reduce the distance to
                                 #  this value (Ang.)
 Bond_Cutoff      1.2            # Factor to multiply sum of covalent radii of
                                 #  two atoms by. If the distance is smaller
                                 #  than the result, the atoms are considered
                                 #  bonded.

# The formulation used to estimate the tangent to the path
 Tangent improved                # improved (default)
                                 # original

# The type of the spring interaction parallel to the path. Original springs apply
# spring interaction between each degree of freedom of adjacent images, while
# 'image' springs apply a spring interaction between the images
# Spring type
 SpringType image                # image / distance (default)
                                 # dof / original
                                 # ideal
 SpringConst  0.01               # The spring constant used to scale the spring
                                 #  forces parallel to the path. If energy-weighted
                                 #  springs are used. This parameter gives the
                                 #  lower bound value of the spring constant
 SpringConst2 0.1                # If energy-weighted spring forces are used.
                                 #  This parameters give the value for the upper
                                 #  bound value of the spring constant.
 Energy_Weighted true            # Employ energy-weighted springs. When
                                 #  energy-weighted springs are used, the
                                 #  images tend to accumlate in higher energy
                                 #  regions of the path.

# The type of the spring interaction perpendicular to the path. The perpendicular
# spring is introduced via a scaling function: cos, tan, costan, which all use
# the spring component perpendicular to the path.
# DNEB is the doubly nudged elastic band method.
 PerpSpring    no                # no      (default)
                                 # cos
                                 # tan
                                 # cosTan
                                 # DNEB

  LLT_Cos true                   # Enables the cos-type spring force
                                 #  acting perpendicular to the band.

# Translational and rotational degrees of freedom
Quatern     always           # no,
                             # startonly
                             # always (default)
# Fix_center specifies whether the centroid of each image should be
# constrained to the origin of the coordinate system or to the center
# of each image individually.
Fix_center  True

# Fix_center specifies whether the centroid of each image should be
# constrained to the origin of the coordinate system or to the center
# of each image individually.
Remove_extern_Force  True    # Removes the net effective NEB force before
                             #  translation of the path

# Options for Free-End NEB
 Free_End_Type    Perp       # Type of optimization of endpoints in free-end
```

(continues on next page)

(continued from previous page)

```
                              #  NEB.
                              # contour - constrain end points to a fixed
                              #  contour with energy EC, see below
                              # perp - allow end points to move according to
                              #  perp. spring force
                              # full - allow to move according to full force,
                              #  i.e. relax to energy minimum
  Free_End_EC                 # Energy contour value for image 0 - needed for
                              #  free_end_type = contour
  Free_End_EC_End             # Energy contour value for image N - needed for
                              #  free_end_type = contour
  Free_End_Kappa              # harmonic restraint term - needed for
                              #  free_end_type = contour


# Monitor convergence for all images or only the CI.
# Convergence type
  ConvType all                    # all   (default)
                                  # CIOnly

  CI false                        # Do Climbing image NEB

  NEB_TS false                    # Do CI NEB and subsequent TS opt.

# Convergence tolerance. In Eh / Bohr (except Tol_Scale ).
  Tol_MaxFP_I      1.e-3  # Default. The convergence tolerance for the
                          #  maximum component of the atomic force
                          #  perpendicular to the path.
  Tol_RMSFP_I      5.e-4  # Default. The convergence tolerance for the rms
                          #  atomic force perpendicular to the path. Only
                          #  applies to regular images.
  Tol_MaxF_CI      2.e-3  # The convergence tolerance for the maxmimum
                          #  component of the atomic force acting on the CI.
                          #  Only applies to (ZOOM-)NEB-CI/-TS calculations.
                          #  Default is 5.e-4 (-CI) and 2.e-3 (-TS)
  Tol_RMSF_CI      1.e-3  # The convergence tolerance for the rms atomic
                          #  force acting on the CI. Only applies to (ZOOM-)NEB-CI.
                          #  Default is 2.5e-4 (-CI) and 1.e-3 (-TS)
  Tol_Turn_On_CI   2.e-2  # Thresholds for max. atomic force for switching on
                          #  CI in (ZOOM-)NEB-CI and (ZOOM-)NEB-TS.
                          #  Defaults: 0.02 for LBFGS, 0.2 for VPO and FIRE
  Tol_Scale        10.0   # For convergence type 'all' the user can scale
                          #  the convergence tolerance of the regular images
                          #  relative to the CI values using this
                          #  multiplicative factor. Only applies to (ZOOM-)NEB-CI
                          #  and (ZOOM-)NEB-TS calculations.

 # Interpolation and redistribution of the path is performed every 'reparam'
 # iterations. The type of interpolation is set by reparam_type.
  Reparam_type linear   # Cubic
                        # Linear (default)
  Reparam 0             # No. of iterations after which the path should be
                        #  reparametrized
                        #  0 (default) means: reparametrization is off
  Tol_Reparam 0.0       # User-defined threshold at which the path should be
                        #  reparametrized
                        # 0.0 (default) means: reparametrization is off
 # The optimization method used to converge the band on the MEP / saddle point.
 # The L-BFGS is more aggressive and efficient, but also more error-prone.
 # VPO is conservative and robust.
  Opt_Method LBFGS      # LBFGS   (default)
```

(continues on next page)

(continued from previous page)

```
                        # VPO
                        # FIRE
                        # BFGS  - TODO Villi correct?

# Options Optim. Method
  Maxmove  0.1    # maximum component allowed per step. Default is 0.1 (LBFGS)
                  #  and 0.2 (VPO / FIRE)
  Stepsize 1.0    # multiplicative factor to scale the size of the step in each
                  # optimization cycle.
                  #  Default is 1.0 (LBFGS) and 0.5 (VPO / FIRE)
  MaxIter  500    # Maximum number of iterations. 500 for LBFGS, 1000 for VPO / FIRE.
  Local    false  # Use local optimization.
                  #  Default is false for NEB, but true for (ZOOM-)NEB-CI/-TS.

# Options LBFGS
  LBFGS_Mem      20    # the number of previous steps to be kept in memory and used
                       #  to construct the approximate Hessian matrix.
  LBFGS_DR 1.e-3       # Size of the finite difference step taken at the
                       #  initialization of L-BFGS
  LBFGS_Restart_On_Maxmove true  # Re-initialize L-BFGS for the next step when
                                 #  the 'max-move' limit is reached.
  LBFGS_Reparam_On_Restart false # Re-parametrize when L-BFGS is re-initialized
  LBFGS_Precondition true        # If true, then after initialization, the curvature
                                 #  along direction of the force is estimated and
                                 #  used to determine the first step

# FIRE parameters
  FIRE_INITIAL_DAMP 0.1   # Initial value for the damping factor
  FIRE_DAMP_DECR"   0.99  # Decrease of the damping factor
  FIRE_STEP_INCR"   1.1   # Factor to increase the stepsize
  FIRE_STEP_DECR"   0.5   # Factor to decrease the stepsize
  FIRE_MAX_STEP"    5.0   # Default is 10 x Stepsize
  FIRE_RETENTION"   5     # Retention before starting acceleration

# Options Zoom
  Tol_Turn_On_Zoom  0.1   # use ZOOM-NEB(-CI/TS)
  Zoom_Offset       1     # if manual selection is chosen, how many
                          #  images away from CI should be chosen
  Zoom_Auto         true  # automatically select zoom region
  Zoom_Alpha        0.5   # determines how much of the barrier
                          #  zoom-auto should select
  Zoom_Interpolation      # linear (default)
                          # cubic
  Zoom_PrintFullTrj       # print full trajectory including fixed region during Zoom

# Set of parameters to adjust the IDPP pre-optimization when generating the initial
# path.
# Options IDPP
  IDPP_NMax                    7000  # maximum number of cycles allowed in IDPP
  IDPP_Tol_MaxF                0.01  # tolerance on the maximum component of the
                                     #  atomic force perpendicular to the path.
                                     #  For S-IDPP, this setting is used in the
                                     #  final IDPP optimization after all images
                                     #  have been added to the path
  IDPP_ksp                     1.0   # spring constant used to scale the spring
                                     #  force parallel to the path.
  IDPP_Alpha                   0.01  # multiplicative factor to scale the size
                                     #  of the step in each opt. cycle
  IDPP_MaxMove                 0.05  # maximum component allowed per step
  IDPP_Debug                   false # will print out the convergence of IDPP
                                     #  and also the optimization trajectory and
```

(continues on next page)

(continued from previous page)

```
                                    #  the log file for the IDPP run.
  IDPP_Quatern                  true  # Whether quaternions should be used in the
                                      #  IDPP optimization
 # Interpolation scheme between the endpoint images to use for the pairwise atom
 # distances to obtain the ideal distances. Bilinear identifies bonded atoms in
 # one endpoint structure that are not bonded in the other and leads to a smaller
 # change in the bond distance in the bonded region than in the non-bonded one.
  IDPP_Dist_Interpolation             # Linear (default)
                                      # Bilinear
  IDPP_Bilinear_Partition       0.5   # Path position at which to join the two
                                      # linear distance interpolations. The default
                                      # corresponds to the midpoint of the path.
  SIDPP                         false # Whether the IDPP optimization should use a
                                      #  Cartesian linear interpolation initial
                                      #  guess or sequentially add images to the
                                      #  path
  SIDPP_Tol_MaxF                0.01  # Tolerance on the maximum component of the
                                      #  atomic force perpendicular to the path
                                      #  acting on the images closest to the center
                                      #  of the path. This setting defines when
                                      #  new images are added to the path
  SIDPP_Reparam                 true  # Whether to reparameterize the path once
                                      #  after all images have been added to it
                                      #  (uses Reparam_type setting) before the
                                      #  final IDPP optimization of the full path
  SIDPP_Energy_Weighted_Tangent false # Whether to use an energy weight in the
                                      #  tangent definition of the images closest
                                      #  to the center of the path
  SIDPP_Even_NIm                true  # Whether to perform S-IDPP with one image
                                      #  less than requested until all images
                                      #  have been added to the path and then
                                      #  add the last image afterwards. This
                                      #  setting increases robustness of the
                                      #  method since no image is placed in the
                                      #  center of the path during image addition
  SIDPP_Double_NIm              false # Whether to perform S-IDPP with twice the
                                      #  amount of images and then remove half of
                                      #  them automatically. This setting
                                      #  increases robustness of the method for
                                      #  systems involving very heavy rotations of
                                      #  large groups
  SIDPP_Ideal_Springconst       false # Whether the spring constant between the
                                      #  images closest to the center of the path
                                      #  should be scaled according to the ratio
                                      #  of the number of images that have been
                                      #  placed on the path already and the
                                      #  requested number of images. This is a
                                      #  more aggressive setting pulling the
                                      #  images closest to the center of the path
                                      #  together more.


 # Extra Output options
   Monitor_Internals # Up to three internal coordinates can be defined
      {B 0 1}        # for which the values are printed during the NEB run.
      {B 1 5}        #  Possible are (B)onds, (A)ngles, (D)ihedrals and (I)mpropers
   end
 end
```

Output files:

- Configuration and trajectory files:

  - basename_initial_path_trj.xyz: The initial path generated at the start of the NEB run and after minimization of RMSD between the reactant and product states.
  - basename_MEP_trj.xyz: The final converged MEP trajectory.
  - basename_MEP_ALL_trj.xyz: The configurations of each image is appended to this file for each step of the NEB optimization.
  - basename_trj.xyz: The trajectory of TS optimization.
  - basename_MEP.allxyz: Restart file that includes the configuration of each image from the last iteration of an NEB or NEB-CI iteration.
  - basename_NEB-CI_converged.xyz: The configuration of the climbing image after a successful NEB-CI calculation.
  - basename_NEB-HEI_converged.xyz: The configuration of the highest energy image after a successful NEB calculation.
  - basename.xyz: The configuration of the optimized saddle point using the TS optimization.

- Log files:
  - basename.interp: The interpolated energy profile of the path for each iteration during the NEB/NEB-CI optimization.
  - basename.interp.final: The energy profile for the converged path of an NEB/NEB-CI optimization.
  - basename.log: A general log file containing essential information regarding the run e.g., energy, forces and step size.

7.30 Excited States via RPA, CIS, TD-DFT and SF-TDA

ORCA features a relatively efficient single-excitation CI (CIS), "random-phase approximation" (RPA) and time-dependent DFT module that can be used to calculate excitation energies, absorption intensities and CD intensities. Especially TD-DFT became very popular for excited state calculations as it offers significantly better results than HF-CIS at about the same cost. However, there are also many pitfalls of TD-DFT, some of which are discussed in reviews[612][614]. TD-DFT methods are available for closed-shell and spin-unrestricted reference states, together with its collinear spin-flip variant. Analytic gradients are available for all these cases. There also is a doubles correction implemented that improves the results (but also the computational cost). It is often used together with double-hybrid functionals as explained below. The TD-DFT module of ORCA is also extensively used for the calculation of X-ray absorption spectra at the K-edge of a given element.

Starting from version 6.0.0, the output format of the absorption wavelength, oscillator strength etc. has changed compared to the 5.0.x version. For more details on the interpretation of the output, please refer to *One Photon Spectroscopy*.

7.30.1 General Features

The module is invoked with the block:

```
%cis  end

# or equivalently

%tddft end
```

There are a variety of options. The most important one is the number of excited states that you want to have calculated:

```
%cis NRoots    10  end
```

The convergence tolerances are given by:

```
%cis
  ...
  ETol  1e-6
  RTol  1e-6
end
```

The variable `ETol` gives the required convergence of the energies of the excited states (in Eh) and `RTol` is the required convergence on the norm of the residual vectors. Under normal ciorcumstances the calculations need about 5-10 iterations to converge to the default convergence tolerances.

Once converged, the program prints the wave function composition. To keep the printing concise, coefficients smaller than 0.01 are omitted. The threshold can be adjusted with the keyword `TPrint`.

```
%cis
  ...
  TPrint 0.0001 # cut-off for the wave function printing, default= 0.01
end
```

If closed-shell references are used the program can calculate the singlet and spin-adapted triplet excited states at the same time by using:

```
%cis
  ...
  triplets  true
end
```

This is available for all combinations of methods, including analytic gradients, and for double-hybrids.

In order to control the orbitals that should be taken into account in the calculation two mechanisms are available. The first mechanism is the default mechanism and consists of specifying and orbital energy window within which all single excitations will be considered:

```
%cis
  ...
  EWin  -3,3  # (orbital energy window in Eh)
end
```

Thus, the default is to keep core orbitals frozen and to neglect very high lying virtual orbitals which is a sensible approximation. However, you may want to consider to include all virtual orbitals by choosing for example `EWin -3,10000`. The second mechanism is to explicitly give an orbital energy window for each operator, i.e.

```
%cis
  ...
  OrbWin[0] = 2,-1,-1,14 # orbital window for spin-up MOs
  OrbWin[1] = 2,-1,-1,16 # orbital window for spin-down MOs
end
```

The "-1"'s in the above example mean that the HOMO and LUMO for the spin-.up and spin-down orbitals will be automatically determined by the program. In other words, in the above example, only the following excitations are included in the TDDFT calculation:

- Excitations from any occupied alpha orbital whose index is between 2 (inclusive) and that of the alpha HOMO (inclusive), to any virtual alpha orbital whose index is between that of alpha LUMO (inclusive) and 14 (inclusive)
- Excitations from any occupied beta orbital whose index is between 2 (inclusive) and that of the beta HOMO (inclusive), to any virtual beta orbital whose index is between that of beta LUMO (inclusive) and 16 (inclusive)

For calculations based on a restricted reference, `OrbWin[1]` will be ignored.

In using the CIS/TD-DFT module five different types of calculations should be distinguished:

- Semiempirical methods
- Hartree-Fock calculations
- DFT calculations without HF exchange (non-hybrid functionals)
- DFT calculations with HF exchange (hybrid functionals)
- DFT calculations with HF exchange and MP2 correlation (double-hybrid functionals)

7.30.2 Semiempirical Methods

The semiempirical INDO/S method is very suitable to calculate absorption spectra of medium sized to large organic and inorganic molecules. It has been parameterized by the late M. C. Zerner for optical spectroscopy and in my experience at least, it tends to work nicely for many systems. With the semiempirical approach it is easy to calculate many states of large molecules. For example, consider the following calculation on a bis-histidine ligated iron-porphyrin model (in the Fe(II) state) that includes 92 atoms and $\approx$ 16,500 CSFs in the single excitation space. Yet the calculation requires only a few minutes on an ordinary computer for the prediction of the first 40 excited states.

The calculated spectrum is in essentially reasonable agreement with experiment in showing a huge band around 400 nm (the famous Soret band) and a smaller but still intense band between 500 and 550 nm (the Q-band). There are no predicted absorptions below $\approx$ 10,000 cm^{-1}.

The input for the job is shown below:

```
# Test CIS in conjunction with INDO/S

! ZINDO/S TightSCF DIIS NoMOPrint
%cis  NRoots  40
      end
* xyz 0 1
Fe   -0.01736     0.71832    -0.30714
C     2.65779     4.03195    -0.13175
C     3.51572     3.02488    -0.24101
C     2.66971     1.82027    -0.30891
C     3.30062     0.51609    -0.42755
C     2.61022    -0.60434    -0.47131
C     3.32146    -1.89491    -0.57434
C     2.35504    -2.79836    -0.57179
C     1.11740    -1.99868    -0.46878
C    -0.04908    -2.61205    -0.44672
C    -1.30967    -1.89127    -0.38984
C    -2.58423    -2.63345    -0.40868
C    -3.50492    -1.68283    -0.37930
C    -2.72946    -0.42418    -0.33711
C    -3.35747     0.73319    -0.28970
C    -2.66935     2.01561    -0.22869
C    -3.31167     3.19745    -0.16277
C    -4.72835     3.62642    -0.14517
C    -5.84825     2.89828    -0.20597
C    -2.21443     4.15731    -0.09763
C    -1.11572     3.39398    -0.14235
C     0.19578     4.02696    -0.10122
C     1.33370     3.36290    -0.15370
C     3.09165     5.44413    -0.02579
C     2.35656     6.55323     0.10940
N     1.43216     2.09428    -0.24815
N     1.34670    -0.74673    -0.42368
N    -1.39885     2.15649    -0.21891
N    -1.47620    -0.63353    -0.34705
C     5.03025     3.02708    -0.28544
C     4.81527    -2.12157    -0.66646
C    -5.01065    -1.83771    -0.38886
```

(continues on next page)

(continued from previous page)

```
C    -2.28137     5.66820    -0.00321
C    -2.73691    -4.14249    -0.43699
C    -2.42579    -4.72805    -1.83259
C     2.45978    -4.31073    -0.64869
C     2.19678    -4.82182    -2.08201
C     1.60835    -6.22722    -2.10748
C    -1.90102    -6.15737    -1.82447
O    -1.96736    -6.92519    -2.75599
O     1.60982    -7.01844    -1.19330
O    -1.15355    -6.41323    -0.74427
O     0.89871    -6.41433    -3.22828
H     4.17823     5.62170    -0.05623
H     2.86221     7.53117     0.17503
H     1.26303     6.57673     0.17212
H     0.21799     5.11603    -0.03468
H    -1.78003     6.14426    -0.87498
H    -3.32281     6.05139     0.01906
H    -1.78374     6.03115     0.92347
H    -4.89690     4.71221    -0.07358
H    -6.82566     3.40843    -0.18007
H    -5.88239     1.80643    -0.28628
H    -4.44893     0.70720    -0.28575
H    -5.32107    -2.89387    -0.54251
H    -5.45075    -1.49552     0.57400
H    -5.46788    -1.24144    -1.20929
H    -2.05997    -4.55939     0.34045
H    -3.76430    -4.43895    -0.12880
H    -3.33638    -4.66246    -2.47119
H    -1.65517    -4.10119    -2.33605
H    -0.56422    -7.14866    -1.00437
H     0.26056    -7.12181    -3.00953
H     1.48118    -4.13253    -2.58671
H     3.13949    -4.79028    -2.67491
H     3.46153    -4.65168    -0.30336
H     1.73023    -4.75206     0.06633
H     5.26172    -1.51540    -1.48550
H     5.31767    -1.84036     0.28550
H     5.06416    -3.18438    -0.87628
H    -0.07991    -3.70928    -0.48866
H     4.39835     0.46775    -0.47078
H     5.39550     2.59422    -1.24309
H     5.47197     4.04179    -0.19892
H     5.44914     2.41988     0.54738
N     0.01831     0.60829     1.68951
C     0.02054     1.64472     2.54371
C     0.04593    -0.50152     2.45186
N     0.04934     1.20474     3.84418
C     0.06582    -0.16578     3.80848
H     0.00322     2.72212     2.31829
N    -0.05051     0.81937    -2.30431
H     0.05251    -1.53704     2.08183
C     0.11803     1.92670    -3.04495
H     0.05712     1.81091     4.70485
H     0.08982    -0.83278     4.68627
C    -0.24302    -0.18840    -3.17641
C    -0.19749     0.28568    -4.49059
N     0.03407     1.63309    -4.38373
H     0.30109     2.95786    -2.70479
H    -0.41432    -1.24242    -2.91290
H    -0.31761    -0.27403    -5.43315
H     0.12975     2.31943    -5.17616
```

(continues on next page)

(continued from previous page)

```
*
```

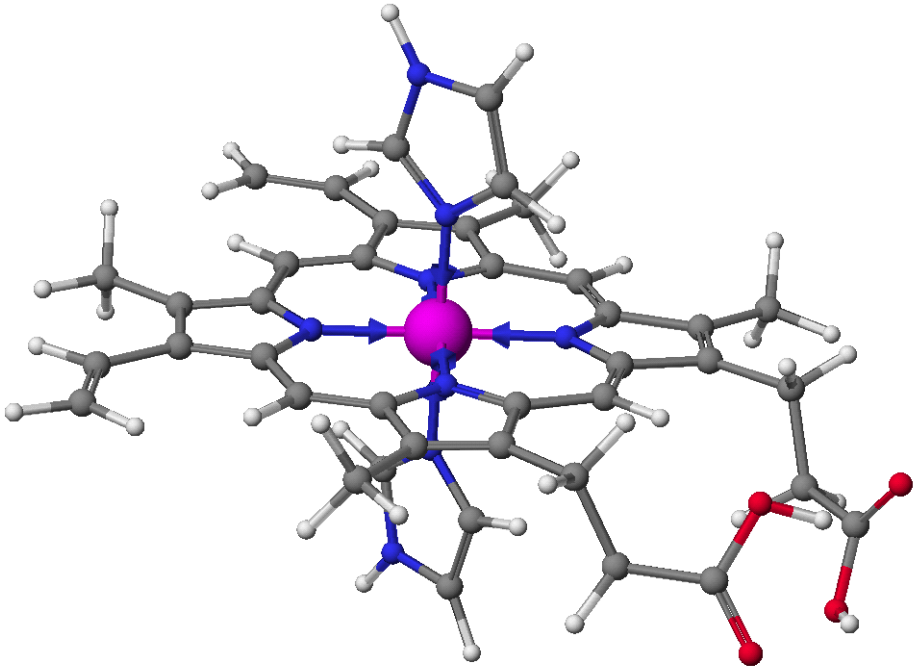


Fig. 7.24: Structure of the iron-porphyrin used for the prediction of its absorption spectrum (the structure was obtained from a molecular mechanics calculation and the iron-imidazole bondlength was set to 2.0 Å).

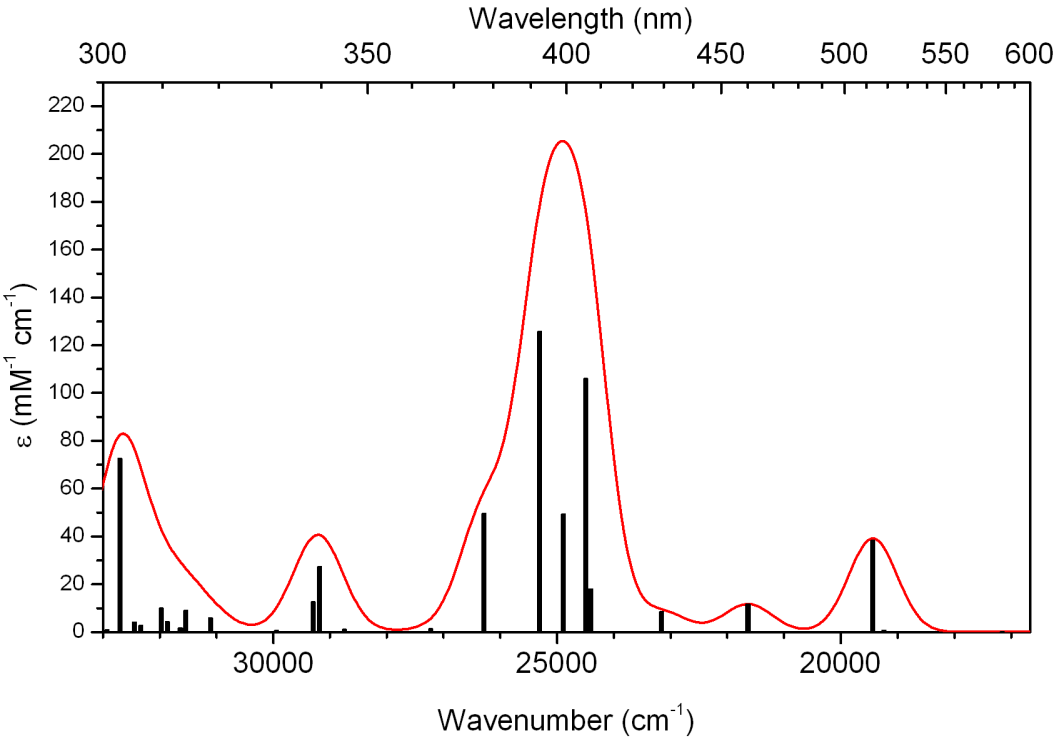


Fig. 7.25: The ZINDO/S predicted absorption spectrum of the model iron porphyrin shown above. The spectrum has been plotted using the `orca_mapspc` tool.

Note that ORCA slightly departs from standard ZINDO/S in using dipole integrals in the intensity calculations that include all one- and two-center terms which are calculated via a STO-3G expansion of the Slater basis orbitals. The calculated intensities are not highly accurate anyways. In the present case they are overestimated by a factor of ≈ 2.

7.30.3 Hartree-Fock Wavefunctions

When applying the procedures outlined above to pure Hartree-Fock, one obtains the "random-phase approximation" (RPA) or the CI singles (CIS) model (when effectively using the Tamm-Dancoff Approximation, TDA). In general, RPA and CIS calculations do not lead to good agreement with experimental excitation energies and errors of 1-5 eV are common. Therefore HF/CIS is mostly a qualitative tool or can be used with caution for larger molecules if more extensive and more well balanced CI calculations are not computationally tractable.

7.30.4 Non-Hybrid and Hybrid DFT

For DFT functionals there is the choice between the full TD-DFT (eq. (7.219)) treatment and the so-called Tamm-Dancoff approximation (TDA).

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{7.219}$$

The TDA is the same approximation that leads from RPA to CIS (i.e. neglect of the so-called "B" matrix, see eq. (7.220)). The results for vertical excitation energies are usually very similar between the two approaches.

$$\mathbf{A}\mathbf{X}_{\mathrm{TDA}} = \omega_{\mathrm{TDA}}\mathbf{X}_{\mathrm{TDA}} \tag{7.220}$$

In general, the elements of matrix "A" and "B" for singlet-singlet excitations in the spin-restricted case are given by eqs. (7.221) and (7.222).

$$\begin{aligned} A_{ia,jb} =& \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + 2(ia|jb) - a_{\mathrm{X}}(ij|ab) \\ &+ (1 - a_{\mathrm{X}})(ia|f_{\mathrm{XC}}|jb) \end{aligned} \tag{7.221}$$

and

$$B_{ia,jb} = 2(ia|bj) - a_{\mathrm{X}}(ib|aj) + (1 - a_x)(ia|f_{\mathrm{XC}}|bj). \tag{7.222}$$

Here, i, j denote occupied and a, b virtual orbitals. a_{X} is the amount of non-local Fock exchange in the density functional. If a_{X} is equal to one, eqs. (7.219) and (7.220) correspond to the RPA and CIS case, based on a Hartree-Fock ground state determinant.

The TDA is actually the default method for TD-DFT, and can be turned off by:

```
%tddft
  TDA  false
end
```

There are situations where hybrid functionals give significantly better results than pure functionals since they suffer less from the self-interaction error. In those cases, the RIJCOSX procedure[623] [415][383] leads to very large speedups in such calculations at virtually no loss in accuracy[675], and is turned on by default whenever the SCF uses that too.

7.30.5 Collinear Spin-Flip TDA (SF-TD-DFT)

Another approach to obtain excited states via CIS/TD-DFT are the so called spin-flip methods (for a good review, please check ref [144]). The idea is to start from an UHF state, and then "flip" one of the alpha electrons to generate states with $MS_{SF} = MS_{UHF} - 1$. In order to do that, we look for excitations from alpha-to-beta orbitals only, and that makes the A matrix from TDA even simpler:

$$A^{SF}_{i\bar{a},j\bar{b}} = \delta_{ij}\delta_{\bar{a}\bar{b}}(\epsilon_{\bar{a}} - \epsilon_i) - a_{\mathrm{X}}(ij|\bar{a}\bar{b}) \tag{7.223}$$

where the overbar represent beta orbitals, and no-overbars alpha orbitals.

OBS.: Please note that for pure DFT (with $a_X = 0$, and no HF contribution), the A matrix is based simply in the orbital energies, and thus it is always good to have a good amount of HF on the functional!

In order to facilitate the discussion on the results one gets from the SF-TDA, let's take a closer look at the picture representing some possible excitations:

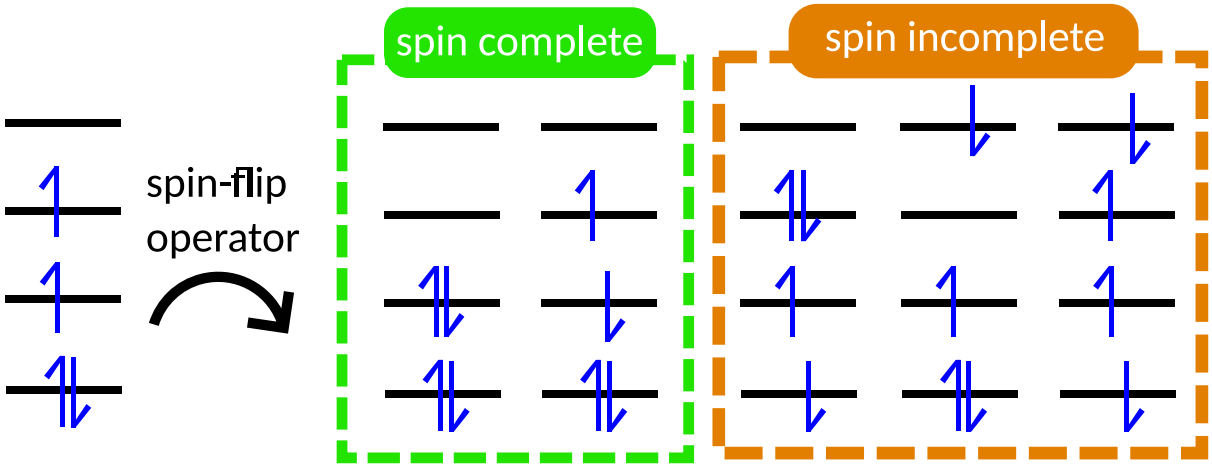


Fig. 7.26: Effect of the spin-flip operator on a UHF ($MS = 3$) wavefunction. The "spin-complete" states are eigenvectors of the S^2 operator, while the "spin-incomplete" are not. Alpha and beta orbitals here are represented with the same energy, just to simplify the image. Adapted from the previously mentioned review.

It is important to note that no all SF-excitations lead to determinants that are eigenvalues of the S^2 operator. That means, depending on how much of these "spin incomplete" excitations are present in the final SF-state, the spin-contamination could be high, and in this case, states with $\langle S^2 \rangle \simeq 1$ would be predicted. These are undefined states within the SF theory and should be treated carefully.

OBS.: Any SF method can only be used starting from a UHF wavefunction, with a multiplicity of at least 3!

First example: methylene and SF-CIS

One simple example is the calculation of the vertical singlet-triplet splitting of the methylene radical within CIS, using the following input with symmetry included:

```
!HF 6-31G USESYM
%TDDFT SF TRUE END
* XYZ 0 3
C 0 0 0.1058
H 0 0.9910 -0.3174
H 0 -.9910 -.3174
*
```

The geometry was taken from a high-level CCSD(T)/cc-pVQZ (X^3B_1) optimized geometry, and after the regular UHF SCF, the SF-CIS result is:

```
---------------------
SF-CIS EXCITED STATES
---------------------

the weight of the individual excitations are printed if larger than 1.0e-02

(SPIN-FLIP GROUND STATE)
STATE  1:  E=   0.004953 au      0.135 eV     1087.0 cm**-1 <S**2> =   2.044208 Sym: B1
     1a ->   3b  :     0.018853 (c=  0.13730475)
     3a ->   3b  :     0.474153 (c= -0.68858776)
     3a ->  10b  :     0.015096 (c= -0.12286571)
     4a ->   4b  :     0.451519 (c=  0.67195159)
     4a ->   9b  :     0.023981 (c=  0.15485668)

STATE  2:  E=   0.065212 au      1.774 eV    14312.3 cm**-1 <S**2> =   0.019616 Sym: A1
     3a ->   4b  :     0.126253 (c=  0.35532096)
     4a ->   3b  :     0.833446 (c= -0.91293269)
     4a ->  10b  :     0.017089 (c= -0.13072354)

STATE  3:  E=   0.085608 au      2.330 eV    18788.7 cm**-1 <S**2> =   0.028873 Sym: B1
     3a ->   3b  :     0.461538 (c=  0.67936623)
     3a ->  10b  :     0.010687 (c=  0.10337584)
     4a ->   4b  :     0.497210 (c=  0.70513090)
     4a ->   9b  :     0.018632 (c=  0.13649832)
```

Now, **it is very important to consider that the SF ground state is not the UHF ground state anymore**, the "new" ground state within the SF scheme is actually STATE 1. You can think of the UHF as being only an initial model, on the basis of which the SF states are built. The final energy of the new ground state is actually the SCF energy + energy of the STATE 1 (which is the one given as the FINAL SINGLE POINT ENERGY is no IROOT is given). This last contribution can be either positive or negative, depending on the case.

Anyway, the ground state is predicted to be a triplet state (here with $M_S = 0$), as expected for this carbene, and the S-T spiting energy is $1.774 - 0.135$ eV $= 1.639$ eV. The full CI results for that is 1.50 eV, so it is already almost there! Of course, in this case computing the RHF singlet - UHF triplet makes no sense, since the RHF singlet would not have the necessary open-shell singlet character.

Benzyne and SF-TDA

Benzyne is a classic diradical that can be generated from benzene by hydrogen abstraction (Fig. 7.27). It is known to have an open-shell singlet ground state, and has its adiabatic sinlget-triplet splitting measured experimentally. Let's try to compute this value using SF-TDA with ORCA.

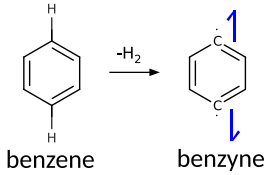


Fig. 7.27: Lewis representation of the benzene and benzyne molecules, indicating the diradical character of the later.

First, we optimize the open-shell singlet by using SF, and the input that follows. Here we use now DFT, in particular the BHANDHLYP functional, which uses 50% of HF correlation, and is recommended for this kind of application. By default, the IROOT to be optimized is 1, which in this case corresponds to the SF ground state.

```
!BHANDHLYP DEF2-TZVPD OPT
%TDDFT
  SF TRUE
  NROOTS  3
END
* xyz 0 3
C         -1.39113        0.00000        0.00000
C          0.69557        1.20476        0.00000
C         -0.69557        1.20476        0.00000
C         -0.69557       -1.20476        0.00000
C          0.69557       -1.20476        0.00000
C          1.39113        0.00000        0.00000
H         -1.24291        2.15278        0.00000
H         -1.24291       -2.15278        0.00000
H          1.24291       -2.15278        0.00000
H          1.24291        2.15278        0.00000
*
```

And after the optimization of IROOT 1, the final SF-TDA result is:

```
---------------------
SF-TDA EXCITED STATES
---------------------

the weight of the individual excitations are printed if larger than 1.0e-02

(SPIN-FLIP GROUND STATE)
STATE  1:  E=   0.024231 au      0.659 eV     5318.2 cm**-1 <S**2> =   0.023398
    11a ->  19b  :     0.018546 (c= -0.13618298)
    17a ->  20b  :     0.245671 (c= -0.49565233)
    17a ->  27b  :     0.016834 (c=  0.12974401)
    20a ->  19b  :     0.666596 (c= -0.81645317)
    20a ->  25b  :     0.020096 (c=  0.14176006)

STATE  2:  E=   0.032598 au      0.887 eV     7154.5 cm**-1 <S**2> =   2.018033
    11a ->  20b  :     0.015438 (c= -0.12424884)
    17a ->  19b  :     0.448627 (c= -0.66979616)
    17a ->  25b  :     0.017992 (c=  0.13413494)
    20a ->  20b  :     0.460929 (c= -0.67891734)
    20a ->  27b  :     0.024021 (c=  0.15498845)
```

(continues on next page)

(continued from previous page)

```
STATE  3:  E=   0.106572 au      2.900 eV    23389.9 cm**-1 <S**2> =   1.029619
    15a ->  20b  :     0.051524 (c=  0.22698827)
    18a ->  19b  :     0.910478 (c= -0.95418975)
    18a ->  25b  :     0.017481 (c=  0.13221571)
```

confirming the singlet ground state, with an upper triplet excited state.

Now to optimize the triplet state using SF-TDA, one has to use a similar input, except that now IROOT 2 has to be chosen as the one to be optimized:

```
!BHANDHLYP DEF2-TZVPD OPT
%TDDFT  SF      TRUE
        NROOTS  3
        IROOT   2
END
* xyz 0 3
C          -1.39113        0.00000        0.00000
C           0.69557        1.20476        0.00000
C          -0.69557        1.20476        0.00000
C          -0.69557       -1.20476        0.00000
C           0.69557       -1.20476        0.00000
C           1.39113        0.00000        0.00000
H          -1.24291        2.15278        0.00000
H          -1.24291       -2.15278        0.00000
H           1.24291       -2.15278        0.00000
H           1.24291        2.15278        0.00000
*
```

After the optimization, the final predicted adiabatic singlet-triplet gap is 0.163 eV, very close to the experimental value of 0.165 eV [178], and even better than what the Broken-Symmetry (BS) result would be (0.074 eV).

method	$\Delta_{ST}^{ad}(eV)$
Exp	0.165 ± 0.016
SF-TDA	0.163
CCSD(dT)	0.172
ΔUKS	1.477
BS	0.074

7.30.6 Including solvation effects via LR-CPCM theory

The LR-CPCM theory, as developed by Cammi and Tomasi [134], is implemented for both energies and gradients of excited states. It is turned on by default, whenever CPCM is also requested for the ground state.

The major change is that now there is a $G_{ia,jb}$ term in the $\mathbf{A}$ part of Eq. (7.219), related to solvation effects.

$$\begin{aligned} A_{ia,jb} =& \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + 2(ia|jb) + 2G_{ia,jb} \\ & - a_{\mathrm{X}}(ij|ab) + (1 - a_{\mathrm{X}})(ia|f_{\mathrm{XC}}|jb) \end{aligned} \tag{7.224}$$

where $G_{ia,bj}$ is defined as:

$$G_{ia,jb} = (\mathbf{V}_{ia})^T \mathbf{q}_{jb} \tag{7.225}$$

Equilibrium and non-equilibrium conditions

These charges q_{jb} are calculated in the same way as described in *The Conductor-like Polarizable Continuum Model (C-PCM)*, but for excited states, two different values of ε can be used, depending on the dynamics of the system:

- **Non-equilibrium**: If the calculation assumes that the electronic excitation is so fast, that there is no time for the solvent to reorganize around the solute, then the ε_{inf} of the solvent is used, which is equivalent to the square of the refractive index. That is the case if one wants to compute the vertical excitation energy, and it is the default in that case.
- **Equilibrium**: If the excited state is assumed to be completely solvated, then the true dielectric constant ε of the solvent should be used. That is the case for geometry optimizations, frequencies or inside ORCA_ESD. This is turned on by default whenever analytic gradients are requested.

In any case, these conditions can be controlled by the flag CPCMEQ, that can be set to TRUE or FALSE by the user, and will then override the defaults.

These are available to all CIS/TD-DFT options: singlets, spin-adapted triplets, UHF and spin-flip variants. It works inclusive for double-hybrids and whenever SOC is requested.

Population Analysis of Excited States

If you want to print a population analysis for the excited state using CIS/TD-DFT, there are two options available: using **unrelaxed** or **relaxed** densities. For the unrelaxed densities, simply use UPOP TRUE:

```
!B3LYP DEF2-SVP
%TDDFT   NROOTS  5
         UPOP    TRUE
END
* XYZ 0 1
O          -1.88199        1.42016       -0.00000
C          -1.80947        0.20286        0.00000
H          -2.50488       -0.38174       -0.59212
H          -1.04956       -0.29504        0.59212
*
```

and the atomic changes and bond orders will be printed for the chosen IROOT (default 1):

```
-------------------------------------------------------------------------------
                    UNRELAXED CIS/TDA DENSITY POPULATION ANALYSIS
                                  IROOT 1
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
                              ORCA POPULATION ANALYSIS
-------------------------------------------------------------------------------
Input electron density              ... form.cisp
BaseName (.gbw .S,...)              ... form

********************************
* MULLIKEN POPULATION ANALYSIS *
********************************

-----------------------
MULLIKEN ATOMIC CHARGES
-----------------------
  0 O :    0.166776
  1 C :   -0.402481
  2 H :    0.117828
  3 H :    0.117876

(...)
```

(continues on next page)

(continued from previous page)

```
****************************
* MAYER POPULATION ANALYSIS *
****************************

NA   - Mulliken gross atomic population
ZA   - Total nuclear charge
QA   - Mulliken gross atomic charge
VA   - Mayer's total valence
BVA  - Mayer's bonded valence
FA   - Mayer's free valence

ATOM       NA         ZA         QA         VA         BVA        FA
0 O      7.8332     8.0000     0.1668     2.5573     1.4373     1.1200
1 C      6.4025     6.0000    -0.4025     3.8545     3.1963     0.6582
2 H      0.8822     1.0000     0.1178     0.9737     0.8626     0.1112
3 H      0.8821     1.0000     0.1179     0.9737     0.8625     0.1112

Mayer bond orders larger than 0.100000
B(  0-O ,  1-C ) :   1.4379 B(  1-C ,  2-H ) :   0.8792 B(  1-C ,  3-H ) :   0.8792
```

To get the analysis from the relaxed density, simply use !ENGRAD to a run a gradient calculation:

```
!B3LYP DEF2-SVP ENGRAD
%TDDFT   NROOTS  5
END
* XYZ 0 1
O         -1.88199        1.42016       -0.00000
C         -1.80947        0.20286        0.00000
H         -2.50488       -0.38174       -0.59212
H         -1.04956       -0.29504        0.59212
*
```

and the printout is:

```
-------------------------------------------------------------------------------
               RELAXED CIS/TDA DENSITY POPULATION ANALYSIS
                                 IROOT 1
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
                         ORCA POPULATION ANALYSIS
-------------------------------------------------------------------------------
Input electron density              ... form.cisp
BaseName (.gbw .S,...)              ... form

*******************************
* MULLIKEN POPULATION ANALYSIS *
*******************************

-----------------------
MULLIKEN ATOMIC CHARGES
-----------------------
0 O :   -0.094934
1 C :   -0.074730
2 H :    0.084824
3 H :    0.084840
Sum of atomic charges:    0.0000000

(...)
```

In order to print the analysis for multiple states, simply use IROOTLIST and TROOTLIST:

```
!B3LYP DEF2-SVP
%TDDFT   NROOTS    5
         IROOTLIST 1,2,3
         TROOTLIST 1,2,3
         UPOP      TRUE
END
* XYZ 0 1
O          -1.88199        1.42016       -0.00000
C          -1.80947        0.20286        0.00000
H          -2.50488       -0.38174       -0.59212
H          -1.04956       -0.29504        0.59212
*
```

7.30.7 Simplified TDA and TD-DFT

ORCA also supports calculations of excited states using the simplified Tamm-Dancoff approach (sTDA) by S. Grimme[323]. The sTDA is particularly suited to calculate absorption spectra of very large systems. sTDA as well as the simplified time-dependent density functional theory (sTD-DFT)[69] approach require a (hybrid) DFT ground state calculation. For large systems, using range-separated hybrid functionals (e.g. ωB97X) is recommended.[724] The sTD-DFT approach in particular yields much better electronic circular dichroism (ECD) spectra and should be used for this purpose.

Theoretical Background

A brief outline of the theory will be given in the following. For more details, please refer to the original papers[69, 323]. In the sTDA, the TDA eigenvalue problem from eq. (7.220) is solved using a truncated and semi-empirically simplified A' matrix. The trunctation negelects all excitations that are beyond the energy range of interest, except a few strongly coupled ones. The matrix elements from eq. (7.221) are simplified by neglecting the response of the density functional and by approximating the remaining two-electron integrals as damped Coulomb interactions between transition/charge density monopoles. In the following, the indices i, j denote occupied, a, b virtual and p, q either kind of orbitals.

$$A'_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + \sum_{A,B}^{N_{\text{atoms}}} (2q_{ia}^A \gamma_{AB}^K q_{jb}^B - q_{ij}^A \gamma_{AB}^J q_{ab}^B) \tag{7.226}$$

q_{pq}^A and q_{pq}^B are the transition/charge density monopoles located on atom A and B, respectively. These are obtained from Löwdin population analysis (see Sec. *Löwdin Population Analysis*). ϵ_p is the Kohn-Sham orbital energy of orbital p. γ_{AB}^K and γ_{AB}^J are the Mataga-Nishimoto-Ohno-Klopman damped Coulomb operators for exchange-type (K) and Coulomb-type (J) integrals, respectively.

$$\gamma_{AB}^J = \left(\frac{1}{(R_{AB})^\beta + (a_X \eta)^{-\beta}} \right)^{\frac{1}{\beta}} \tag{7.227}$$

$$\gamma_{AB}^K = \left(\frac{1}{(R_{AB})^\alpha + \eta^{-\alpha}} \right)^{\frac{1}{\alpha}} \tag{7.228}$$

Here, η is the arithmetic mean of the chemical hardness of atom A and B. α and β are the parameters of the method and are given by:

$$\alpha = \alpha_1 + a_x \alpha_2 \tag{7.229}$$

$$\beta = \beta_1 + a_x \beta_2 \tag{7.230}$$

For any global hybrid functional, α_1, α_2, β_1 and β_2 are identical. α and β then depend on the amount of Fock exchange (a_X) only. This is different for range-separated hybrid functionals where α_2 and β_2 are set to zero. α_1 and β_1 along with a value a_x for the sTDA treatment are individually fitted for each range-separated hybrid functional.[724] It can bee seen from eq. (7.226) that the method is asymptotically correct which is crucial for excitations of charge transfer type.

In sTD-DFT, eq. (7.219) is solved using the simplified matrices A' (see above) and B'.

$$B'_{ia,jb} = \sum_{A,B}^{N_{\text{atoms}}} (2q_{ia}^{A}\gamma_{AB}^{K}q_{bj}^{B} - a_{\text{X}}q_{ib}^{A}\gamma_{AB}^{K}q_{aj}^{B}) \tag{7.231}$$

This approach yields better transition dipole moments and therefore spectra but the method is more costly than sTDA (a factor of 2–5 for typical systems). The parameters used in sTDA and sTD-DFT are identical. There are **no** additional parameters fitted for this method.

Calculation Set-up

sTDA and sTD-DFT can be combined with any (restricted or unrestricted) hybrid DFT singlepoint calculation. Gradients and frequencies are **not** implemented! The methods can be invoked via the `%tddft` block. Table *Keyword list for sTDA and sTD-DFT.* gives a list of the possible keywords.

Table 7.22: Keyword list for sTDA and sTD-DFT.

Mode sTDA	Invokes a sTDA calculation
Mode sTDDFT	Invokes a sTD-DFT calculation
EThresh *value*	Energy threshold up to which CSFs are included (in eV)
PTLimit *value*	Energy threshold up to which CSFs beyond EThresh may be selected (in eV)
PThresh *value*	Selection criterion to include CSF beyond EThresh (in Eh)
axstda *value*	Fock exchange parameter used in sTDA/sTD-DFT calculation (for range-separated hybrids)
beta1 *value*	Constant part of J integral parameter β
beta2 *value*	a_{X} scaled part of J integral parameter β
alpha1 *value*	Constant part of K integral parameter α
alpha2 *value*	a_{X} scaled part of K integral parameter α
triplets true	Calculate singlet-triplet excitations (default: singlet-singlet)

The following example shows how to run such a sTDA calculation using the BHLYP functional if one is interested in all excitations up to 10 eV.

```
! bhlyp def2-SV(P) nososcf tightscf
! smallprint printgap nopop
%maxcore 5000
%tddft
 Mode sTDA
 Ethresh 10.0
 maxcore 5000
end

* xyzfile 0 1 coord.xyz
```

Replacing `Mode sTDA` by `Mode sTDDFT` will invoke a sTD-DFT calculation instead. This is shown in the next example in combination with the ωB97X functional and user specified parameters:

```
! wb97x def2-SV(P) nososcf tightscf
! smallprint printgap nopop
%maxcore 5000
%tddft
 Mode sTDDFT
 Ethresh 10.0
 axstda  0.56
 beta1  8.00
 beta2  0.00
 alpha1 4.58
```

(continues on next page)

(continued from previous page)

```
 alpha2 0.00
 maxcore 5000
end

* xyzfile 0 1 coord.xyz
```

For the range-separated hybrid functionals LC-BLYP, CAM-B3LYP, ωB97, ωB97X, ωB97X-D3 and ωB97X-D3BJ, parameters are available and will be used by default if one of these functionals is used. The way of specifying parameters as shown above is useful if there is a range-separated hybrid functional that has not been parametrized for sTDA yet. For very large systems (e.g. $>$ 500 atoms), it may be useful to define an upper boundary `PTLimit` for the selection of configurations that are beyond `EThresh` (otherwise the whole configuration space will be scanned). This can be done as shown below:

```
! cam-b3lyp def2-SV(P) nori tightscf
! nososcf smallprint printgap nopop
%pal nprocs 4
end
%maxcore 5000
%tddft
 Mode sTDDFT
 Ethresh 10.0
 PThresh 1e-4
 PTLimit 30
 maxcore 20000
end
%method
  runtyp energy
end
* xyzfile 0 1 coord.xyz
```

In this case, all excitations up to 7 eV are considered from the very beginning. Configurations between 7 and 14 eV are included if their coupling to the configurations below 7 eV is strong enough (in total larger than `PThresh`). All configurations beyond 14 eV are neglected. Since the sTDA/sTD-DFT calculations run in serial mode, it is recommended to reset the maxcore within the `%tddft` block (as done in the above examples). In the latter sample input, the ground state procedure runs in parallel mode on 4 cores with a maxcore of 5000 MB set for each node. The subsequent sTD-DFT calculation then runs on a single core, but in order to use all the available memory, the maxcore is reset to a larger value (i.e., 20000 MB). If the maxcore statement within the `%tddft` block was missing, only 5000 MB of memory would be available in the sTD-DFT calculation. Note furthermore that for very large systems, using a functional with the correct asymptotic behaviour is very important (due to the fixed amount of GGA exchange, CAM-B3LYP does **not** provide this property).

The ORCA output will summarize the important properties of your calculation which allows you to check your input:

```
-------------------------------------------------------------------------------

                        ORCA sTDA CALCULATION

                        please cite in your paper
 orginal sTDA method:   S. Grimme, J. Chem. Phys. 138, 244104 (2013)
 range-separated sTDA:  T. Risthaus, A. Hansen, S. Grimme, Phys. Chem. Chem. Phys.
                        16, 14408-14419 (2014)
 sTD-DFT approach:      C. Bannwarth, S. Grimme, Comp. Theor. Chem.
                        1040-1041, 45-53 (2014)
-------------------------------------------------------------------------------


spectral range up to (eV)      ...   10.000000
occ. MO cut-off (eV)           ...   -24.052589
virt. MO cut-off (eV)          ...   17.726088
```

(continues on next page)

(continued from previous page)

```
perturbation threshold         ...   1.000e-04
CSF selection range up to (eV) ...   30.000000
MOs in sTD-DFT                 ...          37
occ. MOs in sTD-DFT            ...          14
virt. in sTD-DFT               ...          23
calculate triplets           ...   no

Calculating the dipole lengths integrals ...
Transforming integrals ...
Calculating the dipole velocity integrals ...
Transforming integrals ...
Calculating magnetic dipole integrals ...
Transforming integrals ...

SCF atom population (using active MOs):

   4.009   4.182   4.182   4.318   4.318   0.867   0.867   0.876   0.876   0.876
   0.876   0.876   0.876

Number of electrons in sTDA:   28.000

ax(DF)   :  0.3800
s_k      :  2.0000
beta  (J):  1.8600
alpha (K):  0.9000
```

The spectroscopic data is also printed out after the calculation has finished:

```
14 roots found, lowest/highest eigenvalue :     6.627     9.945

excitation energies, transition moments and amplitudes

molecular weight:   68.119
state   eV        nm        fL         fV        Rl        RV
   0   6.627     187.1   0.000000   0.000001   0.002400   0.033014     0.71 (  12->  14)   ...
   1   6.637     186.8   0.000188   0.000233  -6.595360  -6.544674    -0.71 (  13->  14)   ...
   2   8.162     151.9   0.000022   0.000113  -0.169704  -0.383021    -0.65 (  12->  16)   ...
   3   8.185     151.5   0.708166   0.559459 -33.378989 -33.157817     0.62 (  13->  16) ...
   4   8.514     145.6   0.461396   0.349012  64.100474  55.364958    -0.63 (  12->  17)   ...
   5   8.531     145.3   0.000004   0.000282   0.539213   4.637973    -0.72 (  13->  17)   ...
   6   8.927     138.9   0.000080   0.001340   0.439265   1.794914     0.70 (  13->  18)   ...
   7   8.929     138.9   0.002612   0.003077  -5.590091  -7.144206    -0.69 (  12->  18)   ...
   8   9.156     135.4   0.432008   0.300685 -30.271745 -29.351033    -0.74 (  12->  17) ...
   9   9.347     132.6   0.058500   0.054136 -37.502752 -36.077121    -0.53 (  12->  19) ...
  10   9.534     130.0   0.338851   0.235400  59.709273  68.042758     0.66 (  12->  18)   ...
  11   9.624     128.8   0.007213   0.004968  25.554619  21.208832    -0.49 (  13->  18)   ...
  12   9.922     125.0   0.021172   0.019486 -22.874039 -23.258574     0.81 (  13->  20) ...
  13   9.945     124.7   0.001403   0.001498   6.301469   6.510456     0.79 (  12->  20)   ...

sTD-DFT done

Total run time:        0.326 sec

           *** ORCA-CIS/TD-DFT FINISHED WITHOUT ERROR ***
```

`fL`, `fV`, `RL` and `RV` are the length and velocity expressions of the oscillator and rotatory strengths, respectively. They may be convoluted by a spectrum processing program to yield the UV/Vis absorption and ECD spectra.

7.30.8 Double-hybrid functionals and Doubles Correction

The program can compute a doubles correction to the CIS excitation energies. The theory is due to Head-Gordon and co-workers.[371] The basic idea is to compute a perturbative estimate (inspired by EOM-CCSD theory) to the CIS excited states that is compatible with the MP2 ground state energy. In many cases this is a significant improvement over CIS itself and comes at a reasonable cost since the correction is computed *a posteriori*. Of course, if the CIS prediction of the excited state is poor, the (D) correction – being perturbative in nature – cannot compensate for qualitatively wrong excited state wavefunctions.

In addition – and perhaps more importantly – the (D) correction is compatible with the philosophy of the double-hybrid functionals and should be used if excited states are to be computed with these functionals. The results are usually much better than those from TD-DFT since due to the large fraction HF exchange, the self-interaction error is much smaller than for other functionals and after the (D) correction the results do not suffer from the overestimation of transition energies that usually comes with increased amounts of HF exchange in TD-DFT calculations.

Since the calculations would require a fairly substantial integral transformation that would limit it to fairly small molecules if no approximation are introduced we have decided to only implement a RI version of it. With this approximation systems with more than 1000 basis functions are readily within the reach of the implementation.

Since one always has a triad of computational steps: MP2-CIS solution-(D) correction, we have implemented several algorithms that may each become the method of choice under certain circumstances. The choice depends on the size of the system, the number of roots, the available main memory and the available disk space together with the I/O rate of the system. The formal cost of the (D) correction is $O(N^5)$ and its prefactor is higher than that of RI-MP2. In the best case scenario, the rate limiting step would be the calculation of the pair-contribution in the "U-term" which requires (for a closed-shell system) twice the effort of a RI-MP2 calculation *per state*.

The use of the (D)-correction is simple. Simply write:

```
! HF DEF2-SVP DEF2-SVP/C TightSCF
%cis dcorr n # n=1-4. The meaning of the four algorithms is
             # explained below.
   # algorithm 1 Is perhaps the best for small systems. May use a
   #             lot of disk space
   # algorithm 2 Stores less integrals
   # algorithm 3 Is good if the system is large and only a few
   #             states are calculated. Saves disk and main
   #             memory.
   # algorithm 4 Uses only transformed RI integrals. May be the
   #             fastest for large systems and a larger number
   #             of states
end
```

Table 7.23: Integral handling in various implementations of the (D) correction (i,j=occupied MOs, a,b=virtual MOs, Q=aux function; NumInt=numerical integration).

DCORR =	**1**	**2**	**3**	**4**
(ia|jb) integrals	Stored	Stored	Not stored	Not stored
(ij|ab) integrals	Stored	Not made	Not made	Not made
(ab|Q) integrals	Stored	Not made	Not made	Stored
(ij|Q) integrals	Stored	Stored	Stored	Stored
(ia|Q) integrals	Stored	Stored	Stored	Stored
Coulomb CIS	From (ia|jb)	From (ia|jb)	From (ia|Q)	From (ia|Q)
Exchange CIS	From (ij|ab)	RI-AO-direct	RI-AO-direct	From (ab|Q)
XC-CIS	Num. Int.	Num. Int.	Num. Int.	Num. Int.
V-term in (D)	From (ia|jb)	From (ia|jb)	From (ia|Q)	From (ia|Q)
U-term in (D)	From (ab|Q)	RI-AO-direct	RI-AO-direct	From (ab|Q)

NOTE:

- In all three involved code sections (MP2, CIS, (D)) the storage format FLOAT is respected. It cuts down use of disk and main memory by a factor of two compared the default double precision version. The loss of

accuracy should be negligible; however it is – as always in science – better to double check.

- The (ab|Q) list of integrals may be the largest for many systems and easily occupies several GB of disk space (hence algorithms 2 and 3). However, that disk-space is often well invested unless you run into I/O bottlenecks.
- The (ia|jb) and (ij|ab) lists of integrals is also quite large but is relatively efficiently handled. Nevertheless, I/O may be a problem.
- Making the exchange contribution to the CIS residual vector in an RI-AO direct fashion becomes quite expensive for a larger number of states. It may be a good choice if only one or two excited states are to be calculated for a larger system.
- Calculations are possible with the full TD-DFT and the TDA-DFT versions.
- Usage of time-dependent double-hybrids should be cited as follows: For TD or TDA with any double hybrid,[328] TD-B2GPLYP,[310] TDA-PBE0-DH or TDA-PBE0-2,[578] TD-PBE0-DH, TD-PBE0-2, or TDA-B2GP-PLYP [771], TD-ωB2PLYP or TD-ωB2GPPLYP [146], TDA-ωB2PLYP or TDA-ωB2GPPLYP [145], TD(A)-RSX-QIDH or TD(A)-RSX-0DH [145], TDA-PBE-QIDH [119], TD-PBE-QIDH [386], TD(A)-DSD-BLYP or TD(A)-DSD-PBEP86 or many other spin-component-scaled double-hybrid functionals with TD(A)-DFT from 2017 [771], TD(A) ωB88PP86 or TD(A) ωPBEPP86 or many other spin-component and opposite scaled double hybrids with TD(A)-DFT from 2021 [147].
- For instructions on how to employ spin-component-scaling, spin-opposite-scaling, and the calculation of singlet-triplet excitation energies with double hybrids, see Sec. *Doubles Correction*. Note that SCS/SOS-CIS(D) is only automatically used when a TD(A)-DFT calculation is requested for the functionals from 2021 by Casanova-Páez and Goerigk. [147] In those instances, "doscs" has not to be set. SCS/SOS-CIS(D) is not automatically used for PWPB95, ωwB97X-2, or the DSD functionals.
- Cite Ref. [145] when singlet-triplet excitations are calculated with double hybrids.

7.30.9 Natural Transition Orbitals

Results of TD-DFT or CIS calculations can be tedious to interprete as many individual MO pairs may contribute to a given excited state. In order to facilitate the analysis while keeping the familiar picture of an excited state originating from essentially an electron being promoted from a donor orbital to an acceptor orbital, the concept of "natural transition orbitals" can be used.

The procedure is quite straightforward. For example, consider the following job on the pyridine molecule:

```
! PBE D3ZERO def2-SVPD tightscf

%tddft  nroots 5
        DoNTO  true      # flag to turn on generation of natural transition orbitals
        NTOStates 1,2,3  # States to consider for NTO analysis;
                                  #if empty all will be done
        NTOThresh 1e-4   # threshold for printing occupation numbers
        end

* xyz 0 1
  N      0.000000    0.000000    1.401146
  C      0.000000    1.146916    0.702130
  C      0.000000   -1.146916    0.702130
  C     -0.000000    1.205574   -0.702848
  C     -0.000000   -1.205574   -0.702848
  C      0.000000   -0.000000   -1.421344
  H     -0.000000    2.079900    1.297897
  H     -0.000000   -2.079900    1.297897
  H     -0.000000    2.179600   -1.219940
  H     -0.000000   -2.179600   -1.219940
  H      0.000000    0.000000   -2.525017
*
```

which results in:

```
-----------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE    1
-----------------------------------------

Making the (pseudo)densities                        ... done
Solving eigenvalue problem for the occupied space  ... done
Solving eigenvalue problem for the virtual space   ... done
Natural Transition Orbitals were saved in TD-DFT-Example-6.s1.nto
Threshold for printing occupation numbers 0.000100

 E=   0.158709 au      4.319 eV    34832.6 cm**-1
    20a ->  21a  : n=  0.99824359
    19a ->  22a  : n=  0.00067784
    18a ->  23a  : n=  0.00051644
    17a ->  24a  : n=  0.00030975


-----------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE    2
-----------------------------------------

Making the (pseudo)densities                        ... done
Solving eigenvalue problem for the occupied space  ... done
Solving eigenvalue problem for the virtual space   ... done
Natural Transition Orbitals were saved in TD-DFT-Example-6.s2.nto
Threshold for printing occupation numbers 0.000100

 E=   0.159970 au      4.353 eV    35109.3 cm**-1
    20a ->  21a  : n=  0.99941615
    19a ->  22a  : n=  0.00019849
    18a ->  23a  : n=  0.00019659


-----------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE    3
-----------------------------------------

Making the (pseudo)densities                        ... done
Solving eigenvalue problem for the occupied space  ... done
Solving eigenvalue problem for the virtual space   ... done
Natural Transition Orbitals were saved in TD-DFT-Example-6.s3.nto
Threshold for printing occupation numbers 0.000100

 E=   0.197236 au      5.367 eV    43288.3 cm**-1
    20a ->  21a  : n=  0.64398585
    19a ->  22a  : n=  0.35061220
    18a ->  23a  : n=  0.00163202
    17a ->  24a  : n=  0.00112466
    16a ->  25a  : n=  0.00073130
    15a ->  26a  : n=  0.00062628
    14a ->  27a  : n=  0.00045034
    13a ->  28a  : n=  0.00022996
    12a ->  29a  : n=  0.00019819
    11a ->  30a  : n=  0.00017291
    10a ->  31a  : n=  0.00011514


------------------------------
TD-DFT/TDA-EXCITATION SPECTRA
------------------------------
```

(continues on next page)

(continued from previous page)

```
Center of mass = (  0.0000,  0.0000,  0.0036)
Generating CIS transition densities           ... done
------------------------------------------------------------------
Using One-Photon Spectroscopy module
------------------------------------------------------------------


-------------------------------------------------------------------------------------------------
↪-----
                     ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
↪-----
     Transition      Energy     Energy  Wavelength   fosc         T2        TX        TY       ␣
↪TZ
                      (eV)       (cm-1)    (nm)                (au**2)    (au)      (au)     ␣
↪(au)
-------------------------------------------------------------------------------------------------
↪-----
  0-1A  ->  1-1A    4.318686   34832.6   287.1   0.004072319   0.03849   0.19619   0.00000   0.
↪00000
  0-1A  ->  2-1A    4.352995   35109.3   284.8   0.000000000   0.00000   0.00000  -0.00000  -0.
↪00000
  0-1A  ->  3-1A    5.367066   43288.3   231.0   0.024724713   0.18803   0.00000  -0.43363   0.
↪00000
  0-1A  ->  4-1A    6.156133   49652.6   201.4   0.000029118   0.00019  -0.00000  -0.00000  -0.
↪01389
  0-1A  ->  5-1A    6.746055   54410.6   183.8   0.027331420   0.16537  -0.00000  -0.40666   0.
↪00001
```

We see that there is a weakly allowed transition (S1) that is essentially totally composed of a single NTO pair (20a→21a : n= 0.99825296), while the third excited state (S3) is strongly allowed and requires two NTO pairs for its description (20a→21a : n= 0.64493520 and 19a→22a : n= 0.34962356).

These orbitals are shown below. It is evident that the S1 state donor orbital (NTO20) is a nitrogen lone pair and the acceptor orbital is a $\pi*$ orbital of the ring. For the S3 state the two NTO donor orbitals are comprised of a nearly degenerate set of π orbitals (they would be degenerate in the parent benzene) and the acceptor orbitals are a pair of nearly degenerate $\pi*$ orbitals. It is evident from this example that by looking at the NTOs one can obtain a nicely pictorial view of the transition process, even if many orbital pairs contribute to a given excited state in the canonical basis.

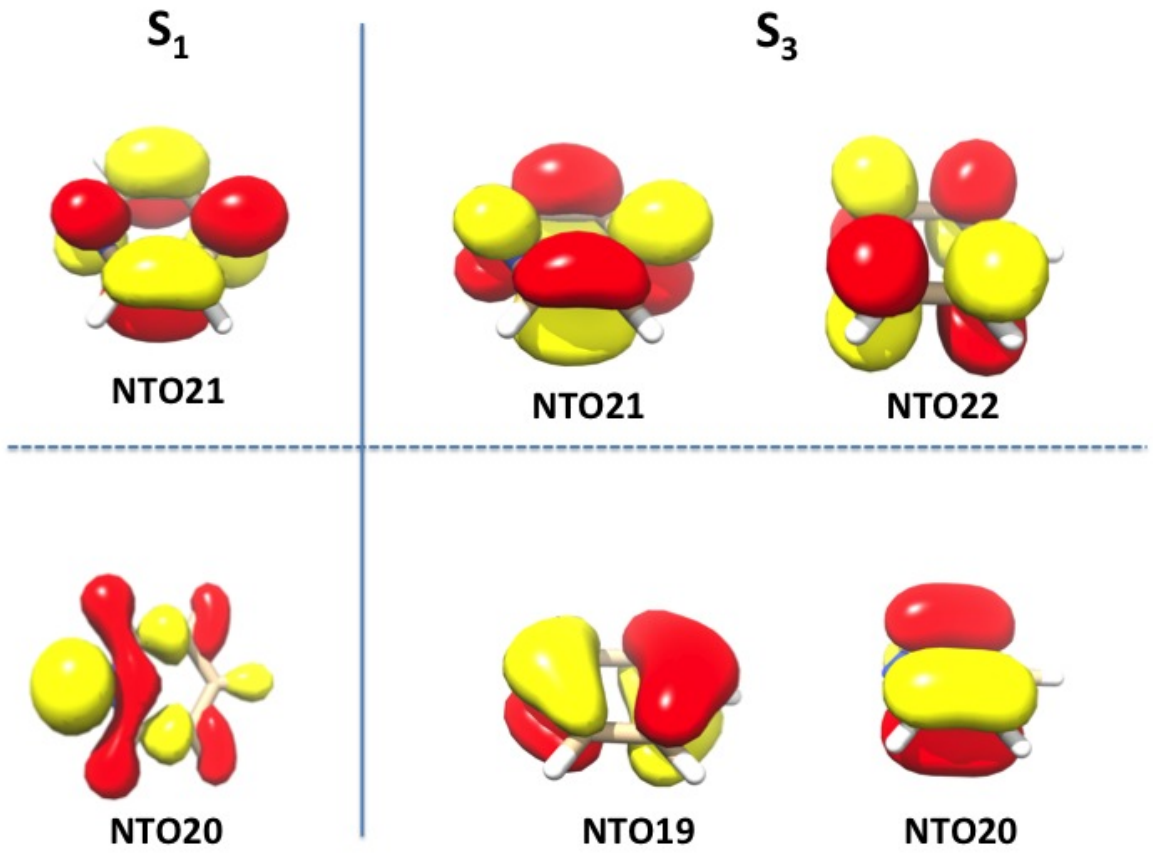


Fig. 7.28: Natural transition orbitals for the pyridine molecule in the S1 and S3 states.

Similar analysis can be performed in the case of ROCIS and DFT/ROCIS calculations as it will be described in section *Natural Transition Orbitals/ Natural Difference Orbitals*.

7.30.10 Computational Aspects

RI Approximation (AO-Basis)

If the SCF calculation used the RI approximation it will also be used in the TD-DFT calculation. The RI approximation saves a large amount of time while giving close to identical results (the errors will usually be $<$0.1 eV) and is generally recommended. If the functional is a hybrid functional the RI-approximation will only be applied to the Coulomb term while the exchange will be treated as before. In the SCF you can use this feature with the keyword (`! RIJONX`). It will then also be used in the TD-DFT calculation. Again, the RIJCOSX approximation can be used in TD-DFT and CIS calculations and leads to very large speedups at virtually no loss in accuracy.

RI Approximation (MO-Basis)

As an alternative to the direct AO-basis computation ORCA allows to use RI-integrals transformed to the MO basis to generate the CI matrix. This algorithm is more disk-intensive. However, for medium sized molecules we have observed speedups on the order of 15 or more with this method. It is particularly benefitial together with hybrid functionals.

In order to use this method you have to specify `mode riints` in the `%tddft` block and you also have to assign an auxiliary basis set (for example `def2-TZVP/C`). There is a second algorithm of this kind that is labelled `mode riints_disk`

Note that the auxiliary basis set has to be valid for correlation treatments in case that you have a hybrid functional. Thus the basis sets developed for RI-MP2 are suitable (`def2-SVP/C`, `def2-TZVP/C` and `def2-TZVPP/C`). If you have a non-hybrid functional the normal RI-J auxiliary basis sets are fine.

An example that uses the B3LYP functional is given below:

```
! RKS B3LYP/G SV(P) def2-SVP/C TightSCF

%tddft
    mode    riints   # or riints_disk (often faster but requires more disk space)
    nroots  8
end

* int 0 1
    C  0 0 0  0.00   0.0   0.0
    O  1 0 0  1.20   0.0   0.0
    H  1 2 0  1.08 120.0   0.0
    H  1 2 3  1.08 120.0 180.0
*
```

Note

- Do not forget to assign a suitable auxiliary basis set! If Hartree-Fock exchange is present (HF or hybrid-DFT) these are the auxiliary bases optimized for correlation while for non-hybrid functionals the standard RI-J bases are suitable.
- The standard auxiliary basis sets may not be suitable if you have diffuse functions present and want to study Rydberg states. You have to augment the axuliary basis with diffuse functions yourself in this case.
- Be prepared that the transformed integrals take up significant amounts of disk space.

Integral Handling

If the SCF calculation is carried out in an integral direct fashion this will also be done in the CIS/TD-DFT calculation. Thus, no bottlenecks arising from large integral transformations or large disk space requirement arise in the calculations. An exception is the MO based RI approximations described in the previous section.

Valence versus Rydberg States

For valence excited states the usual orbital basis sets are reasonable. Thus, with polarized double-zeta basis sets sensible results are obtained. Especially DFT calculations have the nice feature of not being overly basis set dependent.

If Rydberg states are desired, you should make sure that diffuse functions are present in your basis set. You could always use the augmented-specific basis, e.g. DEF2-TZVPD, ma-DEF2-TZVP, or aug-cc-pVTZ, or add some extra diffuse basis to your regular basis. These can be added to any "normal" basis set. For example, the following example provides a rather high quality basis for excited state calculations that is based on the Ahlrichs basis set:

```
%basis
  # augment the carbon basis set by diffuse functions
  addgto 6
    s 1
       1 0.01 1.0
    p 1
       1 0.01 1.0
    d 1
       1 0.07 1.0
    end
end
```

Tip

If you want to augment a given basis set it is sensible to run a preliminary SCF calculation and use `%output print[p_basis] 2 end`. This will provide you with a detailed listing of basis functions and their exponents. You can then add additional s, p and perhaps d-functions with the `AddGTO` command as in the example above. It is sensible to decrease the exponent of the diffuse functions by roughly a factor of 3 from the smallest exponent in the original basis.

Restrictions for Range-Separated Density Functionals

Several restrictions apply for range-separated (hybrid as well as double-hybrid) density functionals. They are currently only implemented to work with the AO-based algorithm within the RIJONX, RIJCOSX, and NORI integral schemes. Additionally, the asymptotic correction has been disabled. However, the nuclear gradient for the excited states is now available, including for the triplets. Please no that the IROOTMULT flag must be set to TRIPLET under %CIS or %TDDFT in order to obtain that.

Potential Energy Surface Scans

ORCA allows the combination the scan feature with CIS or TD-DFT. This can be used to map out the excited state potential energy surfaces as a function of one- two- or three parameters. The output of the "trajectory" run automatically contains the excited state energies in addition to the ground state energy. For example consider the following simple job.

```
! def2-TZVPD
%method scanguess pmodel  # this assignment forces a PModel guess at each step
                          # which is often better if diffuse functions are present
        end
%cis  NRoots        7
      end
%paras rCO = 0.85,1.45,21;
       end
* xyz 0 1
  O  0  0  0
  C  0  0  {rCO}
*
```

The output file from this job contains the *total* energies (i.e. the ground state energy plus the excitation energy) for each excited state as a function of C-O bondlength as shown below. Howerver, the assignment of the individual states will change with geometry due to curve crossings. Thus, the state-to-state correlation must be worked out "by hand". These calculations are nevertheless very helpful in obtaining at least a rough idea about excited state energy surfaces.

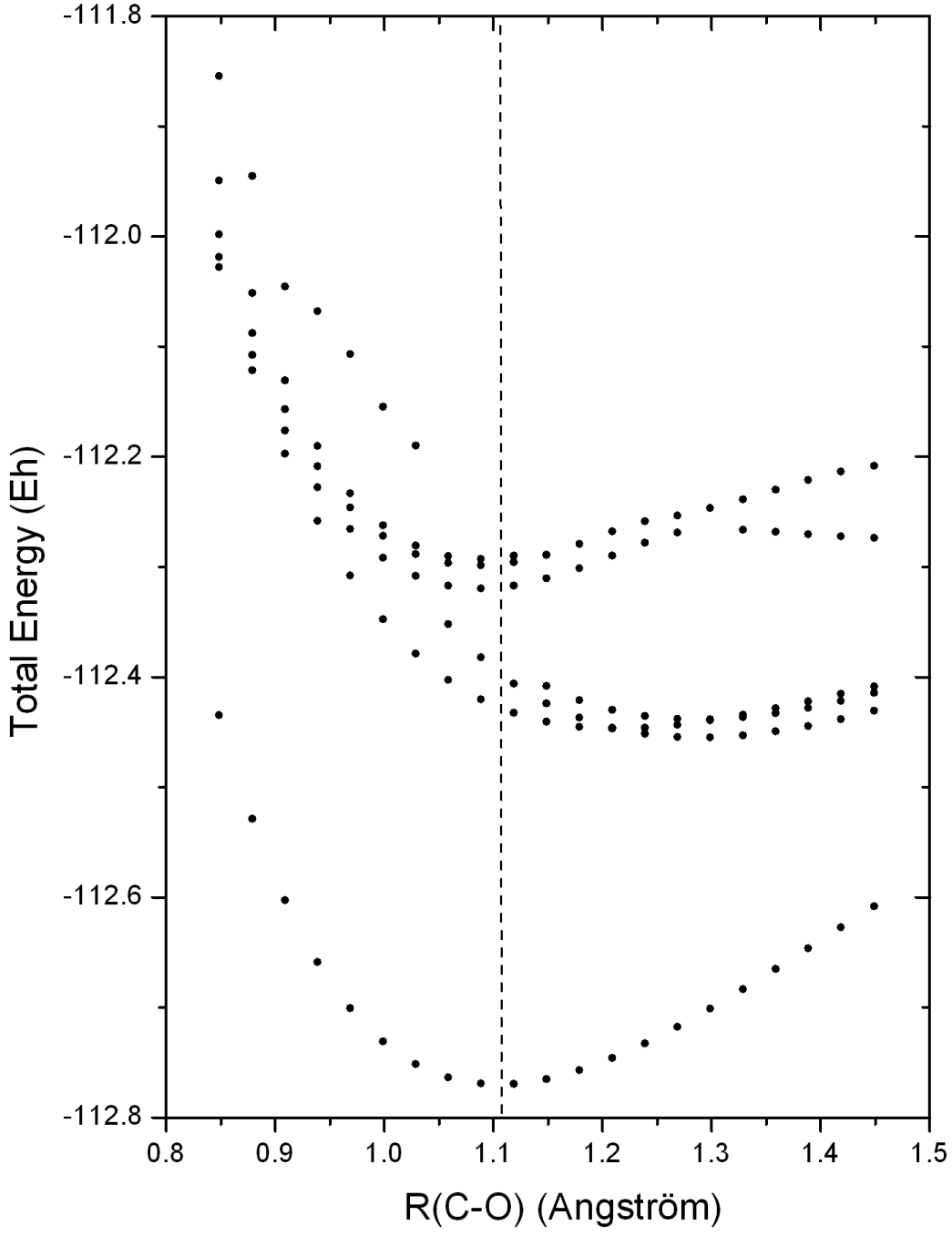


Fig. 7.29: Result of a potential energy surface scan for the excited states of the CO molecule using the `orca_cis` module.

Potential Energy Surface Scans along Normal Coordinates

The ground and excited state potential energy surfaces can also be mapped as a function of normal coordinates. The normal mode trajectory run is invoked by the keyword `!MTR`. In addition several parameters have to be specified in the block `%mtr`. The following example illustrates the use:

First you run a frequency job:

```
#
! BP86 def2-SV(P) def2/J TightSCF AnFreq

* xyz 0 1
  C       0.000001     -0.000000     -0.671602
  C       0.000000      0.000000      0.671602
  H      -0.000000     -0.940772     -1.252732
  H      -0.000000     -0.940772      1.252732
  H      -0.000000      0.940772     -1.252732
  H      -0.000000      0.940772      1.252732
*
```

and then:

```
! BP86 def2-SV(P) def2/J TightSCF MTR

%tddft
  NRoots 3
  triplets false
 end

%mtr
   HessName "ethene.hess"
   modetype normal
   MList 9,13
   RSteps   4,5
   LSteps   4,5
   ddnc     1.0, 0.5
end

* xyz 0 1
  C       0.000001     -0.000000     -0.671602
  C       0.000000      0.000000      0.671602
  H      -0.000000     -0.940772     -1.252732
  H      -0.000000     -0.940772      1.252732
  H      -0.000000      0.940772     -1.252732
  H      -0.000000      0.940772      1.252732
*
```

The `HessName` parameter specifies the name of the file which contains nuclear Hessian matrix calculated in the frequency run. The Hessian matrix is used to construct normal mode trajectories. The keyword `MList` provides the list of the normal modes to be scanned. The parameters `RSteps` and `LSteps` specify the number of steps in positive and negative direction along each mode in the list. In general, for a given set of parameters

```
mlist m1,m2,...mn
rsteps rm1,rm2,...rmn
lsteps lm1,lm2,...lmn
```

the total number of the displaced geometries for which single point calculations will be performed is equal to $\prod_{m_i} (r_{m_i} + l_{m_i} + 1)$. Thus, in the present case this number is equal to $(4 + 4 + 1)\,(5 + 5 + 1) = 99$.

The `ddnc` parameter specifies increments δq_α for respective normal modes in the list in terms of dimensionless normal coordinates (DNC's). The trajectories are constructed so that corresponding normal coordinates are varied in the range from $-l_\alpha \delta q_\alpha$ to $r_\alpha \delta q_\alpha$. The measure of normal mode displacements in terms DNC's is appropriate

choice since in spectroscopical applications the potential energy function U is usually expressed in terms of the DNC's. In particular, in the harmonic approximation $U(q_\alpha)$ has a very simple form around equilibrium geometry:

$$U = U_0 + \sum_{\alpha}^{3N-6} \frac{\hbar\omega_\alpha}{2} q_\alpha^2 \tag{7.232}$$

where $\omega_\alpha$is the vibrational frequency of the α-th mode.

Dimensionless normal coordinate q_α can be related to the vector of atomic Cartesian displacements $\delta\mathbf{X}$ as follows:

$$q_\alpha = \left(\frac{\omega_\alpha}{\hbar}\right)^{\frac{1}{2}} \sum_{k=1}^{3N} L_{k\alpha}\delta X_k \sqrt{M_k} \tag{7.233}$$

where $\{L_{k\alpha}\}$ is the orthogonal matrix obtained upon numerical diagonalization of the mass-weighted Hessian matrix, and $\mathbf{M}$ is the vector of atomic masses. Accordingly, the atomic Cartesian displacements corresponding to a given dimensionless normal coordinate q_α are given by:

$$\delta X_k = \left(\frac{\hbar}{\omega_\alpha}\right)^{\frac{1}{2}} L_{k\alpha} q_\alpha \left(M_k\right)^{-\frac{1}{2}} \tag{7.234}$$

Alternatively, it is possible to specify in the input the Cartesian increment for each normal mode. In such a case, instead of the `ddnc` parameter one should use the `dxyz` keyword followed by the values of Cartesian displacements, for example:

```
%mtr
   HessName "ethene.hess"
   modetype normal
   MList 9,13
   RSteps   4,5
   LSteps   4,5
   dxyz     0.01, 0.02 # increments in the Cartesian basis
                       # are given in angstrom units
end
```

For a given Cartesian increment $d_{X,\alpha}$ along the α–th normal mode the atomic displacements are calculated as follows:

$$\delta X_k = \frac{d_{X,\alpha}}{\|\mathbf{T}_\alpha\|} L_{k\alpha} \left(M_k\right)^{-\frac{1}{2}} \tag{7.235}$$

The vector $\mathbf{T}_\alpha$ in the Cartesian basis has components $T_{i\alpha} = L_{k\alpha} \left(M_k\right)^{-\frac{1}{2}}$ and length (norm) $\|\mathbf{T}_k\|$.

The increment length can also be selected on the basis of an estimate for the expected change in the total energy ΔE due to the displacement according to eq.(7.118). The value of ΔE can be specified via the `EnStep` parameter:

```
%mtr
   HessName "ethene.hess"
   modetype normal
   MList 9,13
   RSteps   4,5
   LSteps   4,5
   EnStep   0.001, 0.001 # the values are given in Eh
end
```

All quantum chemical methods have to tolerate a certain amount of numerical noise that results from finite convergence tolerances or other cutoffs that are introduced into the theoretical procedures. Hence, it is reasonable to choose ΔE such that it is above the characteristic numerical noise level for the given method of calculation.

At the beginning of the program run the following trajectory files which can be visualized in gOpenMol will be created:

- `BaseName.m9.xyz` and `BaseName.m13.xyz` contain trajectories along normal modes 9 and 13, respectively.

- `BaseName.m13s1.m9.xyz - BaseName.m13s5.m9.xyz` contain trajectories along normal mode 9 for different fixed displacements along mode 13, so that the file `BaseName.m13sn.m9.xyz` corresponds to the n-th step in the positive direction along mode 13.
- `BaseName.m13s-1.m9.xyz - BaseName.m13s-5.m9.xyz` contain trajectories along normal mode 9 for different fixed displacements along mode 13, so that the file `BaseName.m13s-n.m9.xyz` corresponds to the n-th step in the negative direction along mode 13.
- `BaseName.m9s1.m13.xyz - BaseName.m9s4.m13.xyz` contain trajectories along normal mode 13 for different fixed displacements along mode 9, so that the file `BaseName.m9sn.m13.xyz` corresponds to the n-th step in the positive direction along mode 9.
- `BaseName.m9s-1.m13.xyz - BaseName.m9s-4.m13.xyz` contain trajectories along normal mode 13 for different fixed displacements along mode 9, so that the file `BaseName.m9s-n.m13.xyz` corresponds to the n-th step in the negative direction along mode 9.

The results of energy single point calculations along the trajectories will be collected in files `BaseName.mtr.escf.S.dat` (for the SCF total energies) and files `BaseName.mtr.ecis.S.dat` (for the CIS/TDDFT total energies), where "S" in the suffix of `*.S.dat` filenames provides specification of the corresponding trajectory in the same way as it was done for the case of trajectory files `*.xyz` (e.g. `S="m9s-1.m13"`). Likewise, the calculated total energies along the trajectories will be collected in files `BaseName.mtr.emp2.S.dat` in the case of MP2 calculations, `BaseName.mtr.emdci.S.dat` (MDCI), `BaseName.mtr.ecasscf.S.dat` (CASSCF), `BaseName.mtr.emrci.S.dat` (MRCI).

Note, that in principle normal coordinate trajectories can be performed for an arbitrary number normal modes. This implies that in general trajectories will contain geometries which involve simultataneous displacement along several (>2) modes. However, trajectory files `*.xyz` and corresponding `*.dat` files will be generated only for the structures which are simultaneously displaced along not more than 2 normal coordinates.

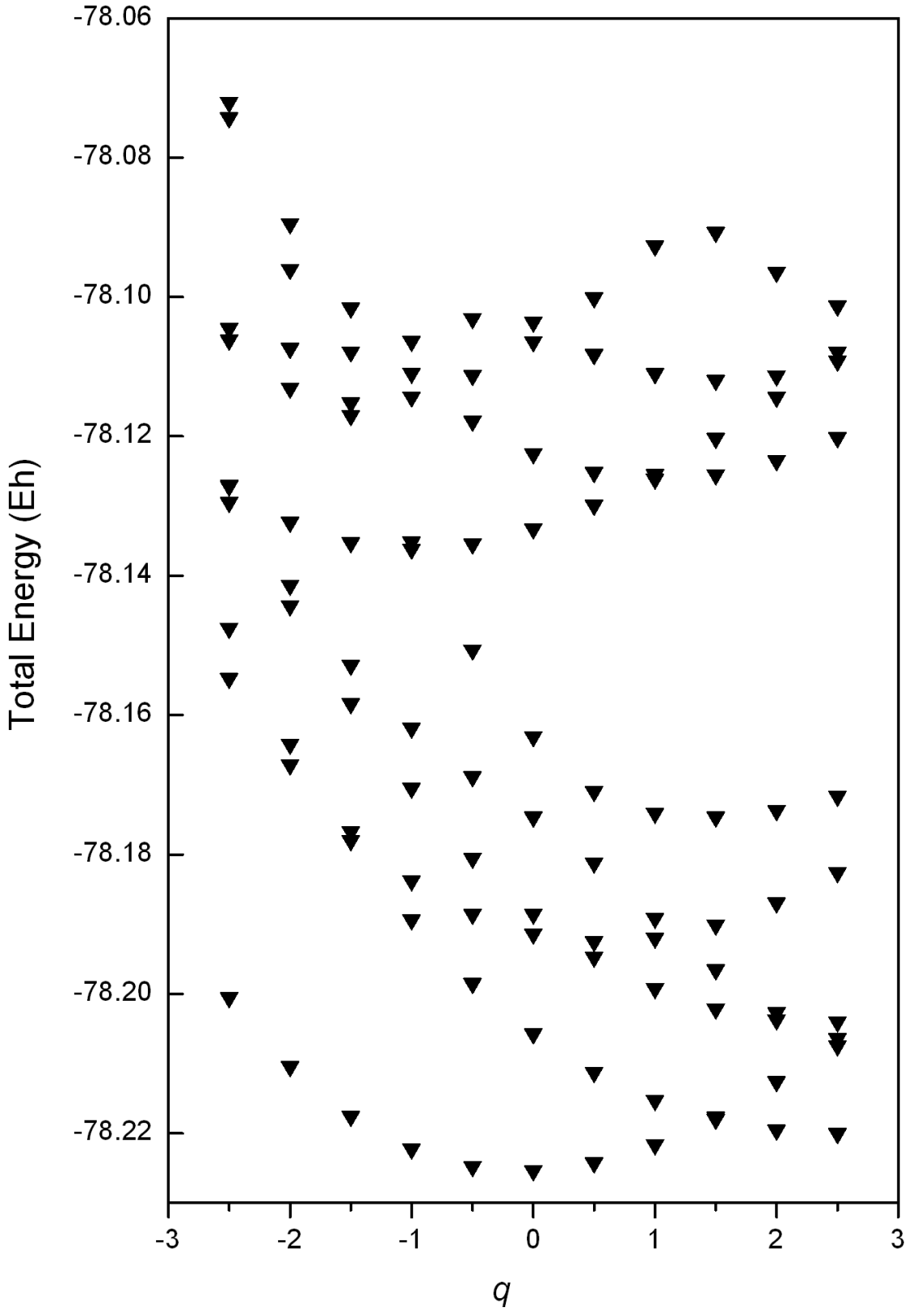


Fig. 7.30: Result of a potential energy surface scan along C-C stretching normal coordinate (mode 13 in the present example) for the excited states of the ethene molecule using the `orca_cis` module.

Normal Mode Scan Calculations Between Different Structures

This type of job allows to map PES between two different structures as a function of normal coordinates. The H_2O molecule represent a trivial case which has formally 2 equivalent equilibrium structures which differ by angle H_1—O—H_2 (103.5° and 256.5°, respectively, as follows from the BP86/SV(P) calculations). In such a case the input for the nomal mode trajectory run would require the calculation of geometry difference between both structures in terms of the dimensionless normal coordinates. This can be done in `orca_vib` run as follows :

```
> orca_vib water.hess ddnc geom.xyz
```

The second parameter ddnc in the command line invokes the calculation of geometry difference in terms of the DNC's. Both structures are specified in the file `geom.xyz` which has a strict format:

```
2  3
0
        0.000000      0.000000      0.000000
        0.000000      0.607566      0.770693
        0.000000      0.607566     -0.770693
1
        0.000000      0.000000      0.000000
        0.000000     -0.607566      0.770693
        0.000000     -0.607566     -0.770693
```

The first line of the input specifies the number of the structures and total number of atoms (2 and 3, respectively). Specification of each structure in sequence starts with a new line containing the number of the structure. The number 0 in the second line is used to denote the reference structure. Note that atomic coordinates should be given in units of Å and in the same order as in the ORCA input for the frequency run from which the file `water.hess` was calculated.

At the end of the `orca_vib` run the file `geom.ddnc` is generated. It contains the geometry difference in terms of the dimensionless normal coordinates between the structures with nonzero numbers and the reference one in `geom.xyz` :

```
1
1 9
    0       0.000000
    1       0.000000
    2       0.000000
    3       0.000000
    4       0.000000
    5       0.000000
    6       9.091932
    7      -9.723073
    8       0.000000
```

The output file indicates that the structural difference occurs along 2 normal coordinates: 6 (bending mode) and 7 (totally symmetric O—H stretching mode). On the basis of the calculated displacement pattern the following input for the normal mode trajectory run between two structures can be designed:

```
! RKS BP86 SV(P) def2/J  RI TightScf MTR

%mtr
   HessName "water.hess"
   modetype normal
   mlist 6,7
   rsteps   10,0
   lsteps   0, 10
   ddnc     0.9091932, 0.9723073
end

* xyz 0 1
```

(continues on next page)

(continued from previous page)

```
O        0.000000      0.000000      0.000000
H        0.000000      0.607566      0.770693
H        0.000000      0.607566     -0.770693
*
```

Here the parameters `RSteps`, `LSteps` and `ddnc` are chosen in such a way that in the scan along modes 6 and 7 the corresponding dimensionless normal coordinates will be varied in the range 0 − 9.091932 and -9.723073 − 0, respectively, in accordance with the projection pattern indicated in the file `geom.ddnc`. Note that normal modes are only defined up to an arbitrary choice of sign. Consequently, the absolute sign of the dimensionless displacements is ambiguous and in principle can vary in different `orca_vib` runs. It is important that the normal mode scan between different structures exemplified above is performed using the same sign of normal modes as in the calculation of normal mode displacements. This condition is fulfilled if the same normal modes are used in `orca_vib` run and trajectory calculation. Thus, since in `orca_vib` calculation normal modes are stored in `.hess` file it is necessary to use the same Hessian file in the trajectory calculation.

Printing Extra Gradients Sequentially

If you want to print extra gradients for external applications or any other reason, you can use the keywords SGRADLIST and TGRADLIST, for singlets and triplets. This will print the gradients sequentially after the CIS/TDDFT run. If you put 0 on the singlet list, the ground state gradient will also be added, always at the end.

```
%TDDFT SGRADLIST  0, 1, 2
       TGRADLIST  2, 3
END
```

In order to save this gradients in a text file, please use:

```
%METHOD STORECISGRAD TRUE END
```

7.30.11 Keyword List

```
%cis or %tddft

       NRoots 3            #The number of desired roots
       IRoot  1            #The root to be optimized
       IRootMult Singlet   #or Triplet to optimize it

       MaxDim 5            #Davidson expansion space = MaxDim * NRoots
       MaxIter 35          #Maximum CI Iterations
       NGuessMat 512       #The dimension of the guess matrix
       MaxCore 1024        #The maximum memory to be used on this calculation
       ETol 1e-6           #Energy convergence tolerance
       RTol 1e-6           #Residual Convergence tolerance

       TDA    false        #Switch off for full TDDFT

       LRCPCM true         #Use LRCPCM
       CPCMEQ false        #Which epsilon is used to compute the charges.

       DoNTO               #Generate Natural Transition Orbitals
       NTOStates 1,2,3     #States to consider for NTO analysis. If empty, all will be done.
       NTOThresh 1e-4      #Threshold for printing occupation numbers

       SaveUnrNatOrb       #Saves natural orbitals (not NTO) from unrelaxed densities
                           #for the IROOT chosen (including IROOTLISTs)

       DoSoc     false     #Include spin-orbit coupling?
```

(continues on next page)

(continued from previous page)

```
      SocGrad   false     #Set true to compute the SOC gradient for a given IROOT

      DOTRANS   false     #Transient spectra - starting from IROOT
                ALL       #Compute all possible transitions
```

7.31 Excited States via ROCIS and DFT/ROCIS

The ORCA program package includes the `orca_rocis` module to perform configuration interaction with single excitations (CIS) calculations using a restricted open-shell Hartee-Fock (ROHF) reference function. It produces excitation energies, absorption energies and CD intensities. It was designed with the aim to reproduce and - even more importantly - reliably predict transition metal L-edges as observed in X-ray absorption spectroscopy (XAS).

7.31.1 General Use

In the present implementation the `orca_rocis` module is only able to perform CIS calculations on top of a high-spin ROHF reference function. All spins of the unpaired electrons have to be coupled ferrmoagnetically to give a total spin of $S = \frac{1}{2}N$, where N is the number of unpaired electrons. Other ROHF functions such as Zerner's configuration averaged or spin averaged ROHF cannot be used as reference. The input for a high spin ROHF calculation is done in the `%scf` block.

```
%scf
  HFTyp ROHF          # Flag for ROHF
  ROHF_Case HighSpin  # selects the high-spin case
  ROHF_NEl[1] = 4     # the number of unpaired electrons
end
```

In our experience ROHF calculations suffer a lot from convergence problems. UHF calculations generally exhibit better convergence properties. In most cases the quasi-restricted orbitals (qro's) of a UHF calculation resemble the ROHF orbitals. Thus the program features the ability to start a ROCIS calculation on top of a UHF calculation. It will automatically create the qro's and build the reference determinant with them. If one wants to avoid the (small) errors that are introduced by this procedure, one may take the qro's of a UHF calculation as starting orbitals for a subsequent ROHF calculation. Furthermore it is possible to invoke the `orca_rocis` module for closed-shell molecules. The program will then perform a CI calculation with the provided RHF reference function. In this case it will yield the same result as the `orca_cis` program.

A number of basic variables in the `%rocis` block control the settings of the Davidson procedure that is used to solve the CI problem:

```
%rocis
  NRoots 6      # number of desired roots
  MaxDim 5      # Davidson expansion space = MaxDim * NRoots
  ETol  1e-6    # energy convergence tolerance
  RTol  1e-6    # residual vector convergence tolerance
  MaxIter 35    # maxmimum number of iterations
  NGuessMat 512 # dimension of the guess matrix: 512x512
end
```

The dimension of the iterative subspace is given by MaxDim · NRoots. The lowest possible choice for MaxDim is a value of 2. In general, by choosing `MaxDim` ≈ 5-10 times `NRoots` you will achieve a more favorable convergence by the cost of an increased disk space requirement. Increasing the `NGuessMat` variable will improve the convergence of the iterative CI procedure. The amount of output produced during the calculation is controlled via the `PrintLevel` variable

```
%rocis NRoots 3
       PrintLevel 3
       end
```

Note, that this does not influence which spectra are calculated or printed. The absorption spectrum calculated on the basis of the pure dipole approximation for your calculation is always printed. In addition, it is possible to allow for electric quadrupole and magnetic dipole contributions to the absorption spectrum as well as to calculate the CD spectrum check section (*One Photon Spectroscopy*) for details. By defining in the %rocis block:

```
%rocis
  NRoots 6
  DoDipoleLength        true
  DoDipoleVelocity      true
  DoHigherMoments       true
  DecomposeFoscLength   true
  DecomposeFoscVelocity true
  DoFullSemiclassical   true
  DoCD                  true
end
```

The printed spectra look like this:

```
-------------------------------------------------------------------------
         ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------
State   Energy  Wavelength   fosc         T2         TX         TY         TZ
        (cm-1)    (nm)                 (au**2)     (au)       (au)       (au)
-------------------------------------------------------------------------
   1    2635.0   3795.1   0.000000001   0.00000   0.00001  -0.00001   0.00029
   2    4365.5   2290.7   0.000011416   0.00086   0.01200  -0.00864   0.02534
   3    4368.2   2289.3   0.000011174   0.00084  -0.02006   0.01442   0.01523
   4    5977.9   1672.8   0.000093897   0.00517  -0.04164  -0.05863   0.00000
   5   65245.3    153.3   0.027669631   0.13961  -0.20555  -0.31203  -0.00023

-------------------------------------------------------------------------
         ABSORPTION SPECTRUM VIA TRANSITION VELOCITY DIPOLE MOMENTS
-------------------------------------------------------------------------
State   Energy  Wavelength   fosc         P2         PX         PY         PZ
        (cm-1)    (nm)                 (au**2)     (au)       (au)       (au)
-------------------------------------------------------------------------
   1    2635.0   3795.1   0.000000085   0.00000  -0.00000   0.00000  -0.00004
   2    4365.5   2290.7   0.001777771   0.00005  -0.00315   0.00223  -0.00618
   3    4368.2   2289.3   0.001850956   0.00006   0.00526  -0.00372  -0.00371
   4    5977.9   1672.8   0.003237195   0.00013   0.00667   0.00937   0.00000
   5   65245.3    153.3   0.057301314   0.02555   0.08779   0.13358   0.00010

-----------------------------------------------------------------
                          CD SPECTRUM
-----------------------------------------------------------------
State   Energy Wavelength       R          MX         MY         MZ
        (cm-1)   (nm)       (1e40*sgs)   (au)       (au)       (au)
-----------------------------------------------------------------
   1    2635.0   3795.1      0.00007  -0.00511  -0.01539   0.00021
   2    4365.5   2290.7     10.02484   0.57434  -0.40490   0.42899
   3    4368.2   2289.3    -10.03730   0.34432  -0.24269  -0.71470
   4    5977.9   1672.8      0.01537  -0.00033  -0.00032  -0.00286
   5   65245.3    153.3     -0.00865   0.00004   0.00003  -0.00005

-------------------------------------------------------------------------------------------------
↪------
               COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE SPECTRUM
-------------------------------------------------------------------------------------------------
↪------
State   Energy Wavelength    D2        m2        Q2         D2+m2+Q2       D2/TOT    m2/TOT    ␣
↪Q2/TOT
        (cm-1)   (nm)                (*1e6)    (*1e6)
```

(continues on next page)

(continued from previous page)

```
-----------------------------------------------------------------------------------------------
↪------
  1    2635.0   3795.1   0.00000   0.00011   0.00000   0.00000000080469   0.86010   0.13938  ␣
↪0.00052
  2    4365.5   2290.7   0.00001   0.47866   0.00000   0.00001189497194   0.95976   0.04024  ␣
↪0.00000
  3    4368.2   2289.3   0.00001   0.48629   0.00000   0.00001166062671   0.95830   0.04170  ␣
↪0.00000
  4    5977.9   1672.8   0.00009   0.00001   0.00001   0.00009389664707   1.00000   0.00000  ␣
↪0.00000
  5   65245.3    153.3   0.02767   0.00000   0.06183   0.02766969236508   1.00000   0.00000  ␣
↪0.00000


-----------------------------------------------------------------------------------------------
↪------
      COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE SPECTRUM (origin␣
↪adjusted)
-----------------------------------------------------------------------------------------------
↪------
State   Energy Wavelength    D2        m2        Q2         D2+m2+Q2       D2/TOT    M2/TOT   ␣
↪Q2/TOT
        (cm-1)   (nm)                (*1e6)    (*1e6)
-----------------------------------------------------------------------------------------------
↪------
  1    2635.0   3795.1   0.00000   0.00000   0.00000   0.00000000069409   0.99716   0.00016  ␣
↪0.00268
  2    4365.5   2290.7   0.00001   0.38277   0.00039   0.00001179947536   0.96753   0.03244  ␣
↪0.00003
  3    4368.2   2289.3   0.00001   0.36798   0.00045   0.00001154275975   0.96808   0.03188  ␣
↪0.00004
  4    5977.9   1672.8   0.00009   0.00000   0.00001   0.00009389663928   1.00000   0.00000  ␣
↪0.00000
  5   65245.3    153.3   0.02767   0.00003   0.06176   0.02766969232228   1.00000   0.00000  ␣
↪0.00000
```

Furthermore like in TD-DFT (section *Use of TD-DFT for the Calculation of X-ray Absorption Spectra*) or CASSCF one may obtain intensities by evaluating the 2nd order oscillation strengths, or the full semi-classical oscillation strengths.

- The exact oscillation strengths behave like the multipole expansion in the velocity representation.
- They are by definition origin independent they do not suffer from artificial negative values like the multipole moments beyond 1st order.
- They are used with the multipole moments up to 2nd order to regenerate the electric dipole, electric quadrupole and magnetic dipole contributions in either length or the velocity representation.

For the Fe K-edge XAS spectrum of $[FeCl_4]^{2-}$. This will result in addition to the following tables for the velocity representation:

```
      -----------------------------------------------------------------------------------------
↪---------------------
      COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE SPECTRUM (Origin␣
↪Independent, Velocity)
      -----------------------------------------------------------------------------------------
↪---------------------
      State   Energy    Wavelength  P2        m2        Q2     P2+m2+Q2+PM+PO     P2/TOT    m2/
↪TOT    Q2/TOT
      (cm-1)      (nm)              (*1e6)    (*1e6)
      -----------------------------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
↪----------------------
      1 57131638.5      0.2   0.00000   0.00000   3.75184   0.00000375184371   0.00000   0.
↪00000   1.00000
      2 57131638.5      0.2   0.00000   0.00000   3.75184   0.00000375184267   0.00000   0.
↪00000   1.00000
      3 57145543.6      0.2   0.00007   0.00000   3.46619   0.00007086820341   0.95853   0.
↪00000   0.04891
      4 57145543.6      0.2   0.00007   0.00000   3.46620   0.00007078008474   0.95972   0.
↪00000   0.04897
      5 57145543.6      0.2   0.00007   0.00000   3.46620   0.00007084079919   0.95889   0.
↪00000   0.04893
     11 57351031.6      0.2   0.00000   0.00000   0.00000   0.00000000000002   0.99463   0.
↪00618   0.00216
     12 57351031.6      0.2   0.00000   0.00000   0.00000   0.00000000000001   0.00000   0.
↪00000   0.00000
     13 57351031.6      0.2   0.00000   0.00000   0.00000   0.00000000000002   0.99414   0.
↪00692   0.00217
     15 57354687.7      0.2   0.00000   0.00000   0.00000   0.00000000000888   0.00898   0.
↪00000   0.00002


      ------------------------------------------------------------------------------------
↪----------------------
      COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC QUADRUPOLE SPECTRUM (Exact␣
↪Formulation, Velocity)
      ------------------------------------------------------------------------------------
↪----------------------
      State   Energy    Wavelength  P2        m2        Q2    Exact Osc. Strength  P2/TOT   ␣
↪m2/TOT    Q2/TOT
      (cm-1)      (nm)              (*1e6)    (*1e6)
      ------------------------------------------------------------------------------------
↪----------------------
      1 57131638.5      0.2   0.00000   0.00000   3.02719   0.00000302719471   0.00000   0.
↪00000   1.00000
      2 57131638.5      0.2   0.00000   0.00000   2.66225   0.00000266224706   0.00000   0.
↪00000   1.00000
      3 57145543.6      0.2   0.00007   0.00000   3.46619   0.00007092969904   0.95853   0.
↪00000   0.04891
      4 57145543.6      0.2   0.00007   0.00000   3.46620   0.00007074406444   0.95972   0.
↪00000   0.04897
      5 57145543.6      0.2   0.00007   0.00000   3.46620   0.00007075200792   0.95889   0.
↪00000   0.04893
     11 57351031.6      0.2   0.00000   0.00000   0.00000   0.00000000000002   0.99463   0.
↪00618   0.00216
     12 57351031.6      0.2   0.00000   0.00000   0.00000   0.00000000000001   0.98256   0.
↪01631   0.00209
     13 57351031.6      0.2   0.00000   0.00000   0.00000   0.00000000000002   0.99414   0.
↪00692   0.00217
     15 57354687.7      0.2   0.00000   0.00000   0.00000   0.00000000001200   0.00898   0.
↪00000   0.00002
      ....
```

These spectra are plotted by calling:

```
orca_mapspc  MyOutput.out ABS/ABSV/CD/ABSQ/ABSOI/ABSVOI -eV -x0(start) -x1(stop)
                    -w(width) -n(points)
```

In particular ABSOI and ABSVOI will plot the exact transition moments spectra at the Length and Velocity representations (For the multiple expansion contributions).

If calculations on large molecules are conducted, the integral transformation will be the most time-consuming part. Therefore it is strongly recommended to use the resolution of the identity (RI) approximation in those cases.

It effectively reduces the computational costs of the transformation step by only introducing minor errors to the calculation. It has to be kept in mind that in order to keep the introduced errors small, one has to provide a reasonable auxiliary basis sets along with your normal basis set input.

Starting from ORCA 4.0 the basis set definition on ORCA has changed. This also affects the definition of the auxiliary basis set when the DoRI keyword is set. ROCIS will then only allow in the mainline `/C` auxiliary basis sets to be set (i.e. def2-TZVP/C). As these basis are usually optimised on the presence of effective core potentials (ECPs) they are generally not recommended for core-electron calculations. The `/J` auxiliary basis set need to be used and they are specified in the following way.

```
%basis
AuxC "def2/J"
end
```

```
! def2-TZVP def2-TZVP/C TightSCF SlowConv

%SCF         HFTyp ROHF
               ROHF_Case HighSpin
               ROHF_Nel[1] = 1
               End

%ROCIS        NROOTS 5
               DoRI true                # invokes the RI approximation
               DoHigherMoments true
               end

* xyz 0 2
N 0 0 0
O 0 0 1.15
*
```

The `orca_rocis` module provides two ways of choosing the orbital excitation space: by orbital energy or orbital number. In the former case an energy window has to be specified and the program will then take all orbitals, whose orbital energies lie within this window, into account. Note, that one actually has to define two orbital windows: One for the donor and the second for the acceptor orbital. The input of the windows is done as an array: The first two numbers define the donor space while the last two numbers define the acceptor space.

```
%rocis
  NRoots 3
  EWin = -5,5,-5,5
end
```

The default is to keep core orbitals and very high lying virtual orbitals out of their respective orbital excitation spaces. Since these orbitals span a space that is usually not reachable with regular UV/Vis spectroscopy, this is a reasonable approximation. One has to keep in mind that an orbital energy window makes only sense if the orbitals used in the calculation have a well-defined orbital energy. As a consequence one cannot use an orbital energy window for a calculation with localized orbitals. The second way to specify the excitation space is by orbital numbering.

```
%rocis
  NRoots 3
  OrbWin = 1,13,9,22
end
```

In restricted calculations only one set of spatial orbitals is created. Hence it is not necessary to provide orbital windows for α and β electrons separately. Of course, only doubly or singly occupied orbitals can act as donor orbitals and only singly and nonoccupied orbitals can act as acceptor orbitals. The program recognises nonoccupied orbitals in the donor space and doubly occupied orbitals in the acceptor space and removes both.

The many-electron expansion space of a ROCIS calculation in ORCA is divided into five classes. Using second

quantised replacement operators $E_p^q = \hat{a}_{q\alpha}^{\uparrow}\hat{a}_{p\alpha} + \hat{a}_{q\beta}^{\uparrow}\hat{a}_{p\beta}$ they take the form[726].

$$\begin{aligned}
|\Phi_i^s\rangle &= E_i^s\,|0\rangle \\
|\Phi_s^a\rangle &= E_s^a\,|0\rangle \\
|\Phi_i^a\rangle &= \tfrac{1}{\sqrt{2}}E_i^a\,|0\rangle \\
|\Phi_{ti}^{as}\rangle &= E_t^a E_i^s\,|0\rangle \\
|\Phi_{ti}^{as}\rangle &= \tfrac{1}{\sqrt{6}}\left(E_i^a - 2E_s^a E_i^s\right)|0\rangle
\end{aligned} \tag{7.236}$$

The orbital label i denotes a doubly occupied orbital, s and t refer to singly occupied orbitals and orbital label a corresponds to a virtual orbital. The form of the excitation classes ensures that all excited states are eigenfunctions of the $\hat{S}^2$-operator and have the same total spin S as the electronic ground state. Each of the five excitation classes can be switched on or off manually.

```
%rocis
  NRoots 3
  Do_is true     # Include DOMO->SOMO excitations
  Do_sa true     # Include SOMO->Virtual excitations
  Do_ia true     # Include DOMO->Virtual excitations
  Do_ista true   # Include DOMO->SOMO coupled to
                 # SOMO->Virtual excitations with s not equal t
  Do_isa true    # Include DOMO->SOMO coupled to
                 # SOMO->Virtual excitations with s = t
                 # ---------------------------------
                 # by default all switches for the
                 # excitation classes are set to
                 # ``true''
                 # ---------------------------------
end
```

Formally, the $|\Phi_{ti}^{as}\rangle$ and $|\Phi_{ti}^{at}\rangle$ excitation classes can be regarded as double excitations. When the program finishes the ROCIS calculation it gives the excitation energy together with the composition for each root. According to the number of labels of the respective functions $|\Phi\rangle$, contributions from excited configuration state functions belonging to the different excitation classes are given by two, three or four numbers.

```
STATE   5   Exc. Energy: 297.279mEh   8.089eV       65245.3cm**-1
      47->50               :   0.2196
      47->51               :   0.0138
      37->50               :   0.1165
      41->50               :   0.0960
      38->46   ;   47->50  :   0.0103
      37->46  ->50         :   0.0150
      37->47  ->50         :   0.0938
      37->48  ->50         :   0.0179
      37->49  ->50         :   0.0179
      41->46  ->50         :   0.0174
      41->47  ->50         :   0.0585
      41->48  ->50         :   0.0213
      41->49  ->50         :   0.0211
```

Furthermore the `orca_rocis` module is able to calculate the effect of spin-orbit coupling (SOC) on the calculated ground and excited states. It introduces SOC in the framework of quasi-degenerate perturbation theory (QDPT). The SOC Hamiltonian is diagonalized in the basis of the calculated ROCIS states $\left|\Psi_I^{SM}\right\rangle$, where I is the root label and S and M are the spin and magnetic spin quantum numbers, respectively[621], [726].

```
%rocis
  NRoots 3
  OrbWin = 1, 3 ,9 ,22
  SOC true        # invokes the calculation of SOC effects
  TEMPERATURE 10  # temperature for SOC corrected spectra in Kelvin
end
```

After the SOC calculation the program will produce additional spectra for the SOC corrected results. The spectra contain transitions from the $2S + 1$ lowest lying states into all excited states, where S is the spin quantum number

of the electronic ground state. These $2S+1$ lowest states may be split up in the order of 1-100 cm^{-1}. Due to the small magnitude of the splitting, all of the $2S+1$ states can be significantly populated even at low temperatures. Experimentally, the intensity of a given transition is dependent on the population of the corresponding initial state. With the `TEMPERATURE` keyword the population of the theoretically calculated states can be manipulated by the varying the fictive temperature of the system. It has to be mentioned that the electric quadrupole transitions between spin-orbit coupled states are not well defined and are likely to give unreasonable results. Hence it is recommended to use the `DoHigherMoments` keyword only for calculations that do not include SOC.

```
-------------------------------------------------------------------------------
SPIN ORBIT CORRECTED ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------
States    Energy  Wavelength   fosc         T2         TX         TY         TZ
          (cm-1)    (nm)                  (au**2)     (au)       (au)       (au)
-------------------------------------------------------------------------------
 0  1       5.6      0.0   0.000000000   0.00000   0.00003   0.00002   0.00000
 0  2       6.2      0.0   0.000000000   0.00000   0.00000   0.00000   0.00005
 0  3      23.7   422287.3   0.000000000   0.00000   0.00000   0.00000   0.00000
 0  4      23.7   421562.8   0.000000000   0.00000   0.00018   0.00025   0.00000
 0  5    2621.7   3814.3   0.000000000   0.00000   0.00000   0.00001   0.00005
 0  6    2622.0   3813.9   0.000000000   0.00000   0.00003   0.00012   0.00000
 0  7    2634.7   3795.5   0.000000095   0.00002   0.00388   0.00273   0.00049
 0  8    2634.9   3795.2   0.000000103   0.00002   0.00039   0.00027   0.00495
 0  9    2639.5   3788.6   0.000000001   0.00000   0.00001   0.00001   0.00036
 0 10    4223.6   2367.6   0.000000103   0.00002   0.00043   0.00029   0.00390
 0 11    4223.9   2367.5   0.000000120   0.00002   0.00348   0.00236   0.00046
 0 12    4296.3   2327.6   0.000000696   0.00010   0.00562   0.00842   0.00000
 0 13    4357.6   2294.8   0.000000002   0.00000   0.00001   0.00001   0.00049
 0 14    4418.1   2263.4   0.000005778   0.00083   0.00653   0.00468   0.02762
 0 15    4422.1   2261.4   0.000005517   0.00079   0.02184   0.01559   0.00832
 0 16    4488.2   2228.0   0.000000001   0.00000   0.00004   0.00006   0.00038
 0 17    4524.2   2210.3   0.000000001   0.00000   0.00030   0.00018   0.00000
 0 18    4597.2   2175.2   0.000000027   0.00000   0.00023   0.00016   0.00191
 0 19    4597.4   2175.2   0.000000051   0.00001   0.00213   0.00153   0.00023
 0 20    6043.6   1654.6   0.000047989   0.00502   0.04104   0.05779   0.00000
 0 21    6049.5   1653.0   0.000000014   0.00000   0.00109   0.00057   0.00001
 0 22    6051.3   1652.5   0.000000021   0.00000   0.00001   0.00004   0.00150
 0 23    6069.7   1647.5   0.000000000   0.00000   0.00005   0.00007   0.00000
 0 24    6069.9   1647.5   0.000000028   0.00000   0.00098   0.00138   0.00000
 0 25   65281.7    153.2   0.014223474   0.13787   0.20423   0.31010   0.00023
 0 26   65281.7    153.2   0.000000035   0.00000   0.00032   0.00048   0.00011
 0 27   65281.7    153.2   0.000009000   0.00009   0.00522   0.00774   0.00001
 0 28   65281.7    153.2   0.000007207   0.00007   0.00460   0.00698   0.00000
 0 29   65281.7    153.2   0.000047448   0.00046   0.01179   0.01791   0.00001
 1  2       0.6      0.0   0.000000000   0.00000   0.00001   0.00001   0.00000
 1  3      18.1   553477.5   0.000000000   0.00000   0.00000   0.00000   0.00009
 1  4      18.1   552233.6   0.000000000   0.00000   0.00006   0.00004   0.00000
 1  5    2616.1   3822.5   0.000000063   0.00001   0.00006   0.00003   0.00261
 1  6    2616.4   3822.1   0.000000060   0.00001   0.00211   0.00144   0.00006
 1  7    2629.1   3803.6   0.000000143   0.00002   0.00225   0.00321   0.00003
 1  8    2629.3   3803.3   0.000000002   0.00000   0.00015   0.00025   0.00040
 1  9    2633.9   3796.7   0.000000271   0.00003   0.00011   0.00008   0.00538
 1 10    4218.0   2370.8   0.000000005   0.00000   0.00031   0.00046   0.00019
...
```

If the `PrintLevel` value is set to 3 or higher, the program will print out the composition of the SOC corrected states in the basis of states $\left|\Psi_I^{SM}\right\rangle$.

```
Eigenvectors of SOC calculation:
the threshold for printing is: 0.010000
        weight : Root Spin Ms
State 0:       0.00 cm**-1    0.00000 eV
```

(continues on next page)

(continued from previous page)

```
    0.378045   :  0    2    2
    0.235825   :  0    2    0
    0.378045   :  0    2   -2

State 1:       5.61 cm**-1    0.00070 eV
    0.496236   :  0    2    2
    0.496236   :  0    2   -2

State 2:       6.20 cm**-1    0.00077 eV
    0.496291   :  0    2    1
    0.496291   :  0    2   -1
```

Further details of the SOC calculation such as the procedure of SOC integral calculation can be controlled via the `%rel` block (section *Relativistic Options*.

7.31.2 Transition Metal L-Edges with ROCIS or DFT/ROCIS

The `orca_rocis` program was designed to calculate transition metal L-edge spectra of large molecules as they are observed in X-ray absorption spectroscopy (XAS). An L-edge results when an electron is promoted from the 2p shell of a transition metal ion into the valence d shell by an X-ray photon. Strong spin-orbit coupling in the 2p shell and p-d coupling phenomena complicate the interpretation and even more so the prediction of these spectra. It has to be kept in mind that the present program applies a variety of approximations which might lead to observable deviations from experimentally determined spectra. However, we believe that the results obtained from the program are in general qualitatively correct and in most cases accurate close to the experimental uncertainty. In cases where quantitative accuracy is not met, the provided results might still give some insight into the mechanisms of intensity distribution in the spectra.

The special input structure for orbital windows described in *General Use* allows the user to restrict the donor orbital space to the transition metal 2p shell. The acceptor orbital space is the same as in regular UV/Vis spectroscopy. It should include all singly occupied molecular orbitals and as many virtual orbitals as one can afford in the calculation. The number of roots should be chosen large enough so that at least all 2p-3d single excitations are calculated. In many cases even more roots are required since doubly excited or charge transfer states may become important. Moreover the strong SOC apparent in the 2p shell of transition metal ions necessitates the additional calculation of excited states with a total spin of $S' = S + 1$ and $S' = S - 1$ where S is the total spin of the electronic ground state. Accordingly four additional excitation classes introduce excited configuration state functions with a lower and higher spin multiplicity. They feature the second quantized spin raising and lowering operators $\hat{S}^{+}_{pq} = \hat{a}^{\uparrow}_{q\alpha}\hat{a}_{p\beta}$, $\hat{S}^{-}_{pq} = \hat{a}^{\uparrow}_{q\beta}\hat{a}_{p\alpha}$.

$$\left.\begin{aligned}
\left|\Phi_i^{(t-)}\right\rangle &= \sqrt{\tfrac{2S'+1}{2S'+2}}S^-_{ti}\left|0\right\rangle - \sum_{u\neq t}^{\text{SOMO}} \tfrac{1}{\sqrt{2S'+1}}\tfrac{1}{\sqrt{2S'+2}}S^-_{uu}E^t_i\left|0\right\rangle \\
\left|\Phi_i^{(t-)}\right\rangle &= \sqrt{\tfrac{2S'+1}{2S'+2}}S^-_{ti}\left|0\right\rangle - \sum_{u\neq t}^{\text{SOMO}} \tfrac{1}{\sqrt{2S'+1}}\tfrac{1}{\sqrt{2S'+2}}S^-_{uu}E^t_i\left|0\right\rangle \\
\left|\Phi_i^{(a-)}\right\rangle &= \sqrt{\tfrac{2S'+1}{2S'+3}}S^-_{ai}\left|0\right\rangle - \sum_{t}^{\text{SOMO}} \sqrt{\tfrac{(S'+1)^2-S'^2}{(S'+1)(2S'+3)}}\tfrac{1}{\sqrt{2(2S'+2)}}S^-_{tt}E^a_i\left|0\right\rangle \\
&\quad + \sum_{t,u\neq t}^{\text{SOMO}} \sqrt{\tfrac{2}{(2S'+2)(2S'+3)}}\sqrt{\tfrac{1}{(2S'+2)2(2S'+1)}}S^-_{tt}S^-_{uu}S^+_{ai}\left|0\right\rangle
\end{aligned}\right\} \quad S' = S-1$$

$$\left|\Phi_i^{a+}\right\rangle = S^+_{ai}\left|0\right\rangle \quad S' = S+1 \tag{7.237}$$

Inclusion of configuration state functions with higher or lower multiplicity is invoked with the keywords `DoLowerMult` and `DoHigherMult`, respectively.

```
%rocis
  NRoots 20
  SOC true
  DoRI true
  PrintLevel 3
```

(continues on next page)

(continued from previous page)

```
  DoLowerMult true     #Invokes a CI calculation  #with S'=S-1
  DoHigherMult true    #Invokes a CI calculation  #with S'=S+1
  OrbWin = 6,8,0,2000
end
```

The program will conduct a separate Davidson procedure for each multiplicity. Subsequently it gives the excitation energies and compositions of the calculated excited states for all included multiplicities. After all CI calculations are finished, the program gives a list of all calculated roots with their excitation energies and their multiplicities. It is this number that will be referred to as label I in the decomposition of spin-orbit coupled states in the basis $\left|\Psi_I^{SM}\right\rangle$. It is very important to note, that when states with different multiplicities are calculated this number might deviate from the number that appears in the respective CI part of the output. If one gets confused about the numbering of the states, the state energies might act as a guideline through the output of the program.

Without SOC the spin exclusion rule applies which means that only excited states with a total spin equal to the ground state spin ($S' = S$) give rise to non-vanishing intensities. Hence, only these transitions are listed in the spectra before SOC.

```
-------------------------------------------------------------------------------
    ROOT    Mult      Excitation energy[Eh]           [cm-1]          [eV]
-------------------------------------------------------------------------------
     0       5            0.00000000                    0.00          0.000
     1       5            26.24822856                5760820.28       714.251
     2       5            26.24833619                5760843.90       714.254
     3       5            26.27159871                5765949.43       714.887
     4       5            26.27982129                5767754.08       715.110
     5       5            26.30321870                5772889.22       715.747
     6       5            26.30458669                5773189.46       715.784
     7       5            26.33143414                5779081.79       716.515
     8       5            26.33600432                5780084.83       716.639
     9       5            26.33865219                5780665.97       716.711
    10       5            26.34522494                5782108.52       716.890
    11       5            26.34577552                5782229.36       716.905
    12       5            26.35183534                5783559.34       717.070
    13       3            26.42121780                5798787.03       718.958
    14       3            26.42122881                5798789.45       718.958

...

    42       7            27.22926558                5976133.02       740.946
    43       7            27.23201078                5976735.52       741.021
    44       7            27.23280499                5976909.83       741.042
    45       7            27.23594814                5977599.67       741.128
    46       7            27.23865050                5978192.77       741.201
    47       7            27.26590445                5984174.32       741.943
    48       7            27.26597947                5984190.78       741.945
    49       7            27.26604364                5984204.87       741.947
    50       3            27.29447169                5990444.10       742.720
    51       3            27.30121861                5991924.88       742.904
    52       3            27.30655497                5993096.08       743.049
    53       3            27.30685328                5993161.55       743.057
    54       3            27.31274496                5994454.62       743.218
    55       7            27.52164817                6040303.58       748.902
    56       7            27.52433114                6040892.42       748.975
    57       7            27.52448641                6040926.50       748.979
    58       7            27.53903479                6044119.50       749.375
    59       7            27.53935644                6044190.10       749.384

------------------------
ROCIS-EXCITATION SPECTRA
------------------------
```

(continues on next page)

(continued from previous page)

```
NOTE: At this point no SOC is included!!!
Hence only transitions to states with the same spin multiplicity
as the ground state are observed!!!

Center of mass = ( -0.0011, -0.0021,  0.0000)
Calculating the Dipole integrals                 ... done
Transforming integrals                           ... done
Calculating the Linear Momentum integrals        ... done
Transforming integrals                           ... done

-------------------------------------------------------------------------------
         ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------
State   Energy  Wavelength   fosc         T2         TX         TY         TZ
        (cm-1)    (nm)                 (au**2)     (au)       (au)       (au)
-------------------------------------------------------------------------------
   1 5760820.3      1.7   0.000985130   0.00006   0.00612  -0.00434   0.00011
   2 5760843.9      1.7   0.000777158   0.00004  -0.00008   0.00006   0.00666
   3 5765949.4      1.7   0.000000036   0.00000   0.00000   0.00001  -0.00004
   4 5767754.1      1.7   0.000007564   0.00000   0.00033   0.00057  -0.00000
   5 5772889.2      1.7   0.025379335   0.00145  -0.00031   0.00021  -0.03804
   6 5773189.5      1.7   0.026898175   0.00153   0.03203  -0.02254  -0.00039
   7 5779081.8      1.7   0.000000323   0.00000  -0.00006  -0.00009  -0.00008
   8 5780084.8      1.7   0.001711738   0.00010  -0.00572  -0.00805   0.00001
   9 5780666.0      1.7   0.113054940   0.00644  -0.04616  -0.06564  -0.00001
  10 5782108.5      1.7   0.151287595   0.00861   0.00073  -0.00052   0.09281
  11 5782229.4      1.7   0.147199895   0.00838   0.07488  -0.05266  -0.00088
  12 5783559.3      1.7   0.000000026   0.00000   0.00001  -0.00001   0.00004
  28 5960986.7      1.7   0.004292708   0.00024  -0.00881  -0.01263  -0.00000
  29 5963084.1      1.7   0.001638281   0.00009  -0.00774   0.00553   0.00006
  30 5963136.7      1.7   0.001369356   0.00008  -0.00005   0.00003  -0.00869
  31 5963484.9      1.7   0.000935993   0.00005   0.00415   0.00587  -0.00000
  32 5968477.0      1.7   0.000661255   0.00004   0.00493  -0.00349  -0.00007
  33 5968705.6      1.7   0.000607238   0.00003   0.00006  -0.00004   0.00579
  35 5970943.7      1.7   0.000000001   0.00000   0.00000   0.00000  -0.00001
```

After calculation of SOC in the basis of all calculated ROCIS roots, the program prints out the composition of the spin-orbit coupled states (if `PrintLevel` >2) and the corresponding absorption spectrum.

```
Eigenvectors of SOC calculation:
the threshold for printing is: 0.010000
        weight : Root Spin Ms
State 0:       0.00 cm**-1    0.00000 eV
    0.129027   :  0    2    2
    0.741116   :  0    2    0
    0.129027   :  0    2   -2
```

```
-------------------------------------------------------------------------------
SPIN ORBIT CORRECTED ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------
States    Energy  Wavelength   fosc         T2         TX         TY         TZ
          (cm-1)    (nm)                 (au**2)     (au)       (au)       (au)
-------------------------------------------------------------------------------
 0  1       0.0      0.0   0.000000000   0.00000   0.00000   0.00000   0.00000
 0  2       0.8      0.0   0.000000000   0.00000   0.00000   0.00000   0.00000
 0  3       0.8      0.0   0.000000000   0.00000   0.00000   0.00000   0.00000
 0  4       1.0      0.0   0.000000000   0.00000   0.00000   0.00000   0.00000
 0  5 5729330.4      1.7   0.000080556   0.00002   0.00013   0.00009   0.00464
 0  6 5729330.4      1.7   0.000096984   0.00003   0.00415   0.00295   0.00013
 0  7 5731365.3      1.7   0.000000001   0.00000   0.00001   0.00000   0.00000
 0  8 5731365.4      1.7   0.000000000   0.00000   0.00000   0.00000   0.00001
```

(continues on next page)

(continued from previous page)

```
0  9 5733452.5      1.7   0.000058329   0.00002   0.00323   0.00227   0.00004
0 10 5733477.2      1.7   0.000066389   0.00002   0.00003   0.00002   0.00421
0 11 5734964.4      1.7   0.000000034   0.00000   0.00005   0.00007   0.00004
0 12 5737151.2      1.7   0.000047769   0.00001   0.00208   0.00291   0.00000
```

With the aid of the `orca_mapspc` program it is possible to extract a `.plt` file from the printed spectra, which then can be used to generate a plot of the intensity vs the excitation energy. The `orca_mapspc` program applies Gaussian type lineshape functions to the calculated transitions with a user-defined FWHM. One has to provide some information for the program such as the name of the output file, the type of spectrum you wish to plot, the energy range and the like. It is invoked in the command line and the parameters are given as arguments:

```
orca_mapspc FeIICl4.out socabs -eV -w1 -n3000 -x0710 -x1740
```

The first argument has to be the output file of your calculation followed by the type of spectrum that should be plotted. In the case of transition metal L-edges it is an absorption spectrum after the SOC correction. The arguments "-eV" (use electron Volt as energy unit), "-w1" (FWHM = 1eV), "-n3000" (use 3000 grid points), "-x0710" and "-x1740" (energy range: 710 to 740 eV) have to be adapted to the specific calculation. As a result, one obtains a `.plt` and a `.stk` file. The `.plt` file contains five columns. In the first column one finds the energy and in the second the total intensity. Columns three to five contain the x-,y- and z-components of the transition moment. Note, that the distribution of the transition moment among its spatial components depends on the orientation of your molecular axis system. The `.stk` file contains a list of all transitions with their respective transition energy and intensity. A more detailed description of the `orca_mapspc` program and its usage can be found in chapter *orca_mapspc*.

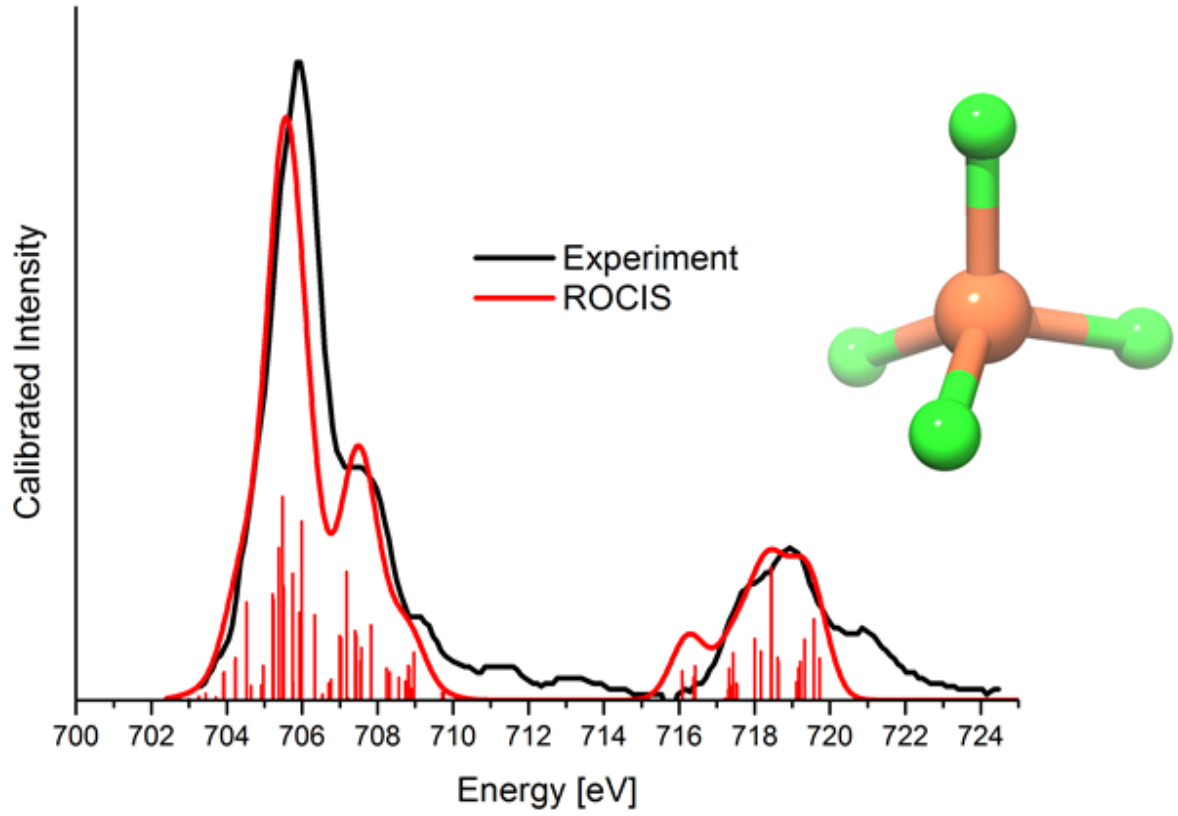


Fig. 7.31: Comparison of the experimentally observed (black) and calculated ROCIS (red) Fe L-edge of $[FeCl_4]^{2-}$. The red bars highlight the contribution of individual states to the total spectrum. The calculation was performed using the TZVP basis set.

For many transition metal compounds the description of the electronic ground and excited states by Hartree-Fock theory and CIS is of rather poor quality. Especially covalency and relative spin state energetics are not reproduced

correctly. This in turn might lead to wrong intensity distributions in the calculated L-edge spectra. In the majority of these cases the quality of the description and hence the predicted L-edge spectra can be significantly improved with the DFT/ROCIS method[726]. It features the usage of a restricted open-shell Kohn-Sham matrix as reference and also uses the DFT orbitals for setting up the excited configuration state functions in the CI expansion. The two electron integrals that include the DFT orbitals are scaled according to their nature and their position in the CI matrix by the parameters c_1, c_2 and c_3. They all lie in the interval [0;1]. Parameters c_1 and c_2 scale coulomb- and exchange- like terms in the diagonal part of the CI matrix, whereas c_3 reduces the size of all off-diagonal elements of the CI matrix. For example:

$$\begin{aligned} H_{ia,ia}^{\text{DFT/ROCIS}} &= F_{aa}^{C(\text{KS})} - F_{ii}^{C(\text{KS})} - c_1\,(ii|aa) + 2c_2\,(ia|ia) \\ H_{ia,jb}^{\text{DFT/ROCIS}} &= c_3\left\{\delta_{ij}F_{ab}^{C(\text{KS})} - \delta_{ab}F_{ji}^{C(\text{KS})} - (ij|ab) + 2\,(ia|jb)\right\} \end{aligned} \tag{7.238}$$

The three default parameters $c_1 = 0.18$, $c_2 = 0.20$ and $c_3 = 0.40$ have been optimized for a test set of molecules and their excited states on a B3LYP/def2-TZVP(-f) level of theory but can be freely chosen[726]. It is most likely that for a different combination of test molecules, functional and basis set, a different set of parameters gives better results. Since the parameters are chosen with regard of a good "balance" between orbital energies, Coulomb and exchange integrals, a new set of parameters should at least crudely resemble their relative proportions.

```
! B3LYP def2-TZVP(-f) TightSCF

%Basis
  AuxC "def2/J"
end

%ROCIS
  NRoots 20
  DoRI true
  SOC true
  DoHigherMult true
  PrintLevel 3
  OrbWin = 5,7,50,60
  DoDFTCIS true   #switches on the DFT/ROCIS method
  DFTCIS_c = 0.18, 0.20, 0.40 #Array input of the three parameters
end
```

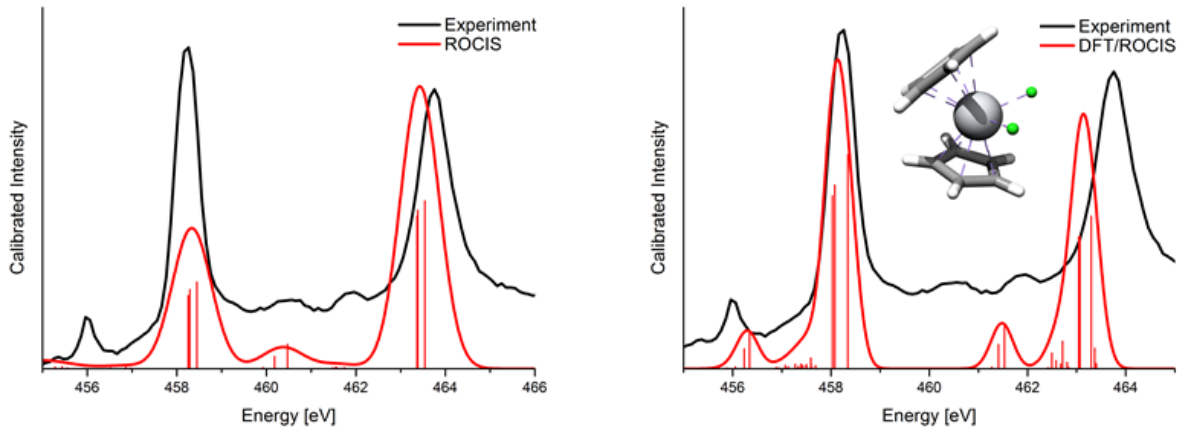


Fig. 7.32: Comparison of the experimentally observed (black) and calculated (red) Ti L-edge of $[Cp_2TiCl_2]$. The red bars highlight the contribution of the individual states to the total spectrum. The pure ROCIS method (left) predicts a wrong L_3-L_2 intensity ratio and strongly overestimates the splitting of the satellite features to the main bands. Better results are obtained with the DFT/ROCIS method (right).

7.31.3 Natural Transition Orbitals/ Natural Difference Orbitals

Likewise to CIS and TD-DFT (section *Natural Transition Orbitals*) The nature of the calculated excited states in ROCIS and DFT/ROCIS can be analyzed by using the Natural Transition Orbitals (NTO) or Natural Difference Orbitals (NDO) machineries.[687] Note that:

- The NTO analysis is based on the transition density between ground and excited states. Hence is valid for singly excited states and for states of the same multiplicity.
- The NDO analysis on the otherhand is somewhat more flexible in this respect as it is based on the difference density between ground and excited states.
- Presently, only one analysis (NTO or NDO) can be performed at a time while when both flags are on the NTO analysis switches off.

An example is given below for $[FeCl_4]^{2-}$:

```
!B3LYP def2-TZVP Conv TightSCF LargePrint PAL4

%Basis
  AuxC "def2/J"
end

%ROCIS
  NRoots 40
  PrintLevel 3
  MaxCore 4000
  MaxDim 360
  SOC true
  DoRI true
  DoNTO true
  DoNDO true
  NDOThresh/NTOThresh 1e-4
  NDOStates/NTOStates= 1,2,3,4,5,6,7,8,9,10,13,14,15
  DoLowerMult true
  DoHigherMult true
  DoDFTCIS true
  DFTCIS_c = 0.18, 0.20, 0.40
  OrbWin = 6,8,0,2000
end


* xyz -2 5
Fe  -17.84299991694815     -0.53096694321123      6.09104775508499
Cl  -19.84288422845700      0.31089495619796      7.04101319789001
Cl  -17.84298666758073      0.11868125024595      3.81067954087770
Cl  -17.84301352218429     -2.87052442818457      6.45826391412877
Cl  -15.84311566482982      0.31091516495189      7.04099559201853
*
```

Then the respective NTO and NDO analysis for state 15 is given below:

```
----------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE   14
----------------------------------------

done
Solving eigenvalue problem for the Occupied space  ... done
Solving eigenvalue problem for the Acceptor space   ... done
Natural Transition Orbitals were saved in nto.14.nto
Threshold for printing occupation numbers 1.0e-04

E=  25.447756 au    692.469 eV  5585137.0 cm**-1
```

(continues on next page)

(continued from previous page)

```
49[0] ->  46[1]  : n=  0.39056909
48[0] ->  47[1]  : n=  0.08619374
47[0] ->  48[1]  : n=  0.00441125
```

```
------------------------------------------------
NATURAL DIFFERENCE ORBITALS FOR STATE   14
------------------------------------------------

done
Solving eigenvalue problem for the Occupied space  ... done
Solving eigenvalue problem for the Acceptor space   ... done
Natural Difference Orbitals were saved in ndo.14.ndo
Threshold for printing occupation numbers 1.0e-04

E=  25.447756 au    692.469 eV  5585137.0 cm**-1
49[0] ->  46[1]  : n=  0.81173217
48[0] ->  47[1]  : n=  0.17903699
47[0] ->  48[1]  : n=  0.01165859
46[0] ->  49[1]  : n=  0.00922738
45[0] ->  50[1]  : n=  0.00112567
```

For closed shell cases the orbitals are save in similar way to TDDFT and CIS (section *Natural Transition Orbitals*). In the case of open shell cases for convenience donor orbitals are saved with orbital operator 0 while acceptor orbitals with orbital operator 1. This needs to be specified in the `orca_plot` program and should not be confused with the `spin-up` and `spin-down` orbitals in the UHF and UKS cases.

In practice one can use this machinery to analyze for example the relativistically corrected states located at 705.5 eV (when shifted with respect to experiment). It can be seen that these states contain for example significant contributions from state 14. NTO or NDO analysis then shows that this state is dominated by the spin conserving DOMO-SOMO $2p_z - 3d_{yz}$ single electron excitation.

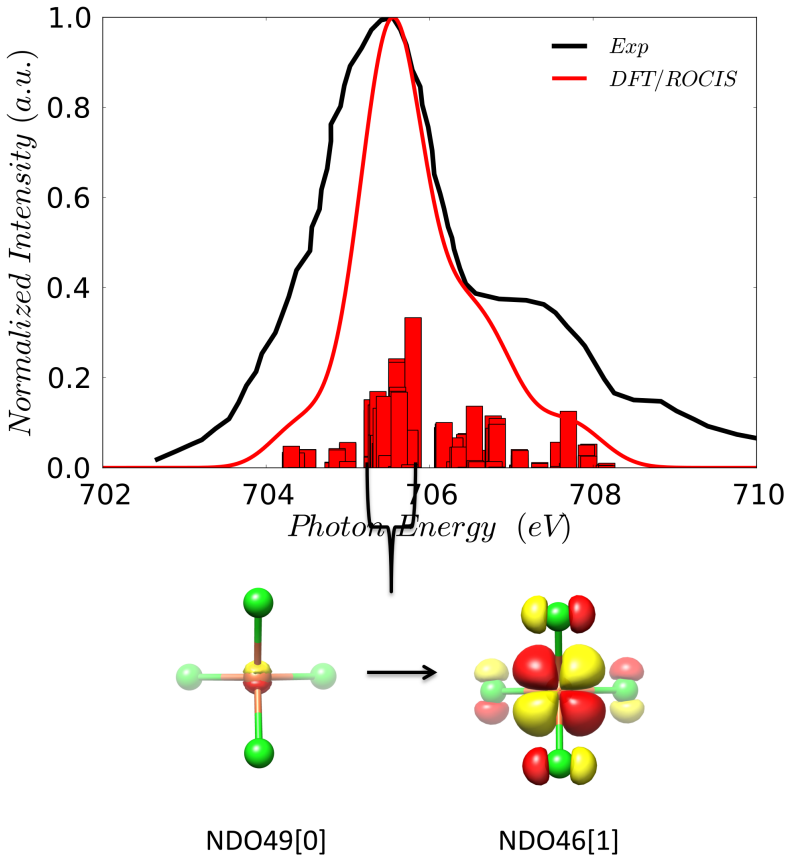


Fig. 7.33: DFT/ROCIS calculated L3 XAS spectrum of $[Fe(Cl)_4]^{2-}$ together with NDO analysis for state 14. Constant broadening 0.5 eV and isovalue for the orbital plots 0.03 a.u. is used throughout

7.31.4 Resonant Inelastic Scattering Spectroscopy

General

Starting from ORCA version 4.0 ROCIS module can be used to calculate RIXS spectra

The present implementation is directly based on the Kramers Heisenerg Dirac (KDH) expression formula for near resonant and resonant conditions

$$|\alpha_{\rho\lambda}(E_{ex},E_{sc})|^2_{Total} = \sum_F \left| \sum_V \frac{\langle F|\, m_\rho\, |V\rangle \langle V|\, m_\lambda\, |I\rangle}{E_{VI} - E_{ex} - i\frac{1}{2}\Gamma_V} \right|^2 \left\{ \frac{\Gamma_F}{(E_{FV} - E_{ex} + E_{sc})^2 + \frac{1}{4}{\Gamma_F}^2} \right\}$$

$$|\alpha_{\rho\lambda}(E_{ex},E_{sc},V)|^2_{resonant} = \sum_F |\langle F|\, m_\rho\, |V\rangle|^2 |\langle V|\, m_\lambda\, |I\rangle|^2 f(E_{VI},E_{FV},E_{ex},E_{sc},\Gamma_V,\Gamma_F)$$

$$|\alpha_{\rho\lambda}(E_{VI},E_{sc})|^2_{Direct} = \sum_V |\alpha_{\rho\lambda}(E_{VI},E_{sc},V)|^2_{resonant}$$

The resonance scattering cross section for total and direct cases, averaged over all orientations of the molecule and integrated over all directions and polarizations of scattered radiation is given in equations:

$$\sigma^{Total}_{RXES}(E_{ex},E_{sc}) = \frac{8\pi E^3_{sc} E_{ex}}{9c^4} \sum_{\rho,\lambda=x,y,z} |\alpha_{\rho\lambda}(E_{ex},E_{sc})|^2_{Total}$$

$$\sigma^{Direct}_{RXES}(E_{ex},E_{sc}) = \frac{8\pi E^3_{sc} E_{ex}}{9c^4} \sum_{\rho,\lambda=x,y,z} |\alpha_{\rho\lambda}(E_{ex},E_{sc})|^2_{Direct}$$

Interference effects can be then derived in a straightforward way from equation:

$$\sigma_{RXES}^{interference}(E_{ex}, E_{sc}) = \sigma_{RXES}^{Total}(E_{ex}, E_{sc}) - \sigma_{RXES}^{Direct}(E_{ex}, E_{sc})$$

In order to access RIXS spectroscopy in the ROCIS module one needs in addition to specify a 2nd donor space. This is specified by defining an OrbWin array with 6 elements: The first four elements define the ranges of the two donor spaces while the last two elements the respective acceptor space range.

```
OrbWin = 0,0,2,4,45,60
```

An important difference with respect to the conventional ROCIS or DFT/ROCIS calculations is the fact that two donor spaces of very different energy ranges are involved (e.g. K-edge, L-edge) which requires to restrict somewhat the acceptor space and saturate it with as many states as possible.

The main calling commands in order to perform a RIXS calculation within both ROCIS and CASSCF blocks are the following:

- RIXS true. Similar to absorption spectroscopy, this requests the RIXS calculation to be performed based on the calculated non-relativistic ground state multiplicity States
- RIXSSOC true. By turning-on this flag the RIXS is calculated by taking in account the relativistically corrected Ms States.
- Elastic true. This flag indicates whether the resonant condition in which the initial and Final states coincide should be taken into account. Note that the intensity of this spectral feature might be overestimated as presently the non resonant terms are not treated

The respective ROCIS input reads then as follows:

```
!B3LYP def2-TZVP  SlowConv

%Basis
  AuxC "def2/J"
end

%ROCIS
  NRoots 200
  PrintLevel 3
  MaxCore 4000
  DoRI true
  DoHigherMult true
  SOC true
  RIXS true    # Request RIXS calculation (NoSOC)
  RIXSSOC true # Request RIXS calculation (with SOC)
  Elastic true # Request RIXS calculation (Elastic)
  DoDFTCIS true
  DFTCIS_c =0.18,0.20,0.40
  OrbWin = 2,4,25,33,0,100
end
* xyzfile 2 2 test.xyz
```

When running the calculation one can monitor if the requested NRoots were sufficient enough to select the states dominated by both the donor orbital spaces

```
-------------------------------------------------------------------------------
ROOT    Mult      Excitation energy[Eh]           [cm-1]          [eV]
-------------------------------------------------------------------------------
0       2            0.00000000                     0.00          0.000
1       2            0.06611737                 14511.08          1.799
2       2            0.07728471                 16962.03          2.103
3       2            0.07732428                 16970.72          2.104
...
84       2            33.75471831               7408304.35        918.513
```

(continues on next page)

(continued from previous page)

```
85       2            33.77073325              7411819.22        918.948
86       2            33.77076955              7411827.19        918.949
87       4            34.06882971              7477243.83        927.060
88       2            34.07021441              7477547.74        927.098
...
```

If that is not the case the respective RIXS calculations will not be performed and a Warning Message will be generated:

```
Making the RIXS files ...
WARNING!: Flag for RIXS property calculation was identified but
there is zero number of Intermediate  and/or Final states:
No Cross-Section properties will be evaluated ...Skipping this part
TIP: Increase the number of NRoots and/or decrease or increase
the acceptor orbital space
...Done
```

A successful run on the other hand will generate the following messages for RIXS and RIXSSOC calculations.

```
-------------------------------------------------------------------------------
ROCIS RIXS SPECTRUM
-------------------------------------------------------------------------------

Making the RIXS data files for Inelastic and Elastic Scattering
Ground State:                   1
Intermediate States:           21
Final States:                  59
The RIXS cross section will be generated from:
60 Ground-Final State Pairs and   21 Intermediate States/Pair
Calculating Intensities...
10% done
20% done
30% done
40% done
50% done
60% done
70% done
80% done
90% done
100% done
Storing the files...All Done
-------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------
ROCIS RIXSSOC SPECTRUM
-------------------------------------------------------------------------------

Making the RIXS-SOC data files for Inelastic and Elastic Scattering
Ms States:                  2
Intermediate States:        78
Final States:               214
The RIXS cross section will be generated from:
432 Ground-Final State Pairs and   78 Intermediate States/Pair

Calculating Intensities...
10% done
20% done
30% done
40% done
50% done
60% done
```

(continues on next page)

(continued from previous page)

```
70% done
80% done
90% done
100% done

Storing the files...All Done
-----------------------------------------------------------------------------
```

In both cases the number of involved Initial, Final and Intermediate states is specified explicitly.

For example in the case of RIXSSOC 2 Ms Ground states, 78 Intermediate states and 214 Final states are involved. Then the RIXS cross section for elastic and inelastic scattering will be generated by 432 (2*(2+214)) Ground-Final State-Pairs and 78 Intermediate States per Ground-Final state pair.

Processing the spectra with `orca_mapspc`

By calling `orca_mapspc` with the following keywords:

```
orca_mapspc test.el_inel.rocis.rixssoc RIXS -x0871 -x1876 -x2-1 -x34 -w0.4 -g0.4
        -l -n125 -m125 -dx20 -eaxis1
```

The program will process the `test.el_inel.rocis.rixssoc` file with the following parameters:

Energy range along x : 871-876 eV

Energy range along y: -1-34 eV

-l indicates Lorentzian broadening

Width along x (gamma): 0.4 eV

Width along y (gamma): 0.4 eV

Points along x: 125

Points along y:125

Shift to be applied along Incident energy/Emission axis: 20 eV

The y axis will be Energy Transfer axis. If -eaxis2 is the y axis will be then Emission Energy axis

All this information is printed during the data processing:

```
Mode is RIXS
Using Lorentzian shape
Cannot read the paras.inp file ...
taking the line width parameter from the command line
Cannot read the udex.inp file ...
taking the excitation energy ranges from the command line
Cannot read the udem.inp file ...
taking the emission energy ranges from the command line
Cannot read the gfsp.inp file ...
No Ground-Final State Pairs will be evaluated
-----------------------------------------------------------------------------
PLOTTING RIXS SPECTRA
-----------------------------------------------------------------------------
Input File : test.el_inel.rocis.rixssoc
Incident Energy Excitation axis :   871.000 ...   876.000 eV  125 points
Energy transfer axis            :    -1.000 ...     4.000 eV  125 points
Incident Energy Shift         :    20.000 eV
Lorenzian Linewidth along Incident Axis                  :     0.400 eV
Lorenzian Linewidth along Energy Transfer/Emission Axis  :     0.400 eV
y axis  :     1 -> Energy transfer
Number of user defined cuts at constant Excitation Energy axis:    0
```

(continues on next page)

(continued from previous page)

```
Number of user defined cuts at constant Emission/Energy Transfer Energy axis  : 0


Making checks...Done

Proccessing data...
10% done
20% done
...
100% done

RIXS-plotting done
Incident Energy range:            845.800 ...   869.249
Emission/Energy-transfer range:     0.000 ...     4.853

Now storing the 2D file...
Done

Making the Integrated spectra along Energy Transfer/Emission axis... Done

Making the Integrated spectra along Incident axis... Done

All Done
-------------------------------------------------------------------------------
```

Successful run will generate the following files: The RIXS planes of the Total, Direct and Interference RIXS intensity as indicated in the above equations:

```
test.el_inel.rocis.rixssoc.total_rixs.dat
test.el_inel.rocis.rixssoc.direct_rixs.dat
test.el_inel.rocis.rixssoc.interference_rixs.dat
```

In addition one obtains the integrated spectra at constant Incident energies (CIE):

```
test.el_inel.rocis.rixssoc.dw.dat
```

as well as at constant Emission/Energy Transfer energies (CEE/CET):

```
test.el_inel.rocis.rixssoc.wex.dat
```

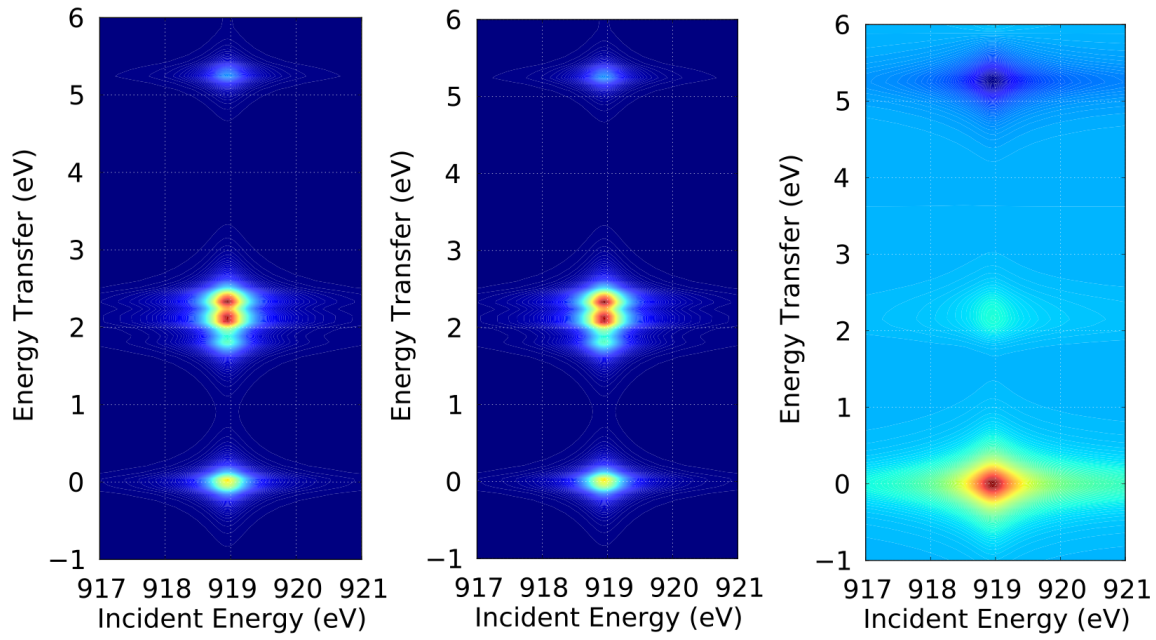


Fig. 7.34: DFT/ROCIS calculated RIXS planes for $[Cu(NH_3)_4]^{2-}$. Left: Total RIXS Intensity, Middle: Direct RIXS intensity and Right: Interference RIXS intensity. Lorentzian lineshape broadening with constant widths along Incident and Energy Transfer axis (0.5 and 0.2 eV respectively) were used throughout.

Generating Cuts

Cuts along x and y axis can be generated with two ways:

1) At first, this action can be performed by adding the following keywords: `uex` and `udw` accounting for generating cuts at constant Incident Energies (CIE) and at constant Emission (CEE)/or at constant Energy Transfer (CET) respectively, together with the desired number of cuts.

2) Alternatively, the energies of the desired cuts can be specified as lists in the files udex.inp (user defined excitations) udem.inp (user defined emissions)

For example if in udex.inp one specifies:

```
872.5
874.2
```

and for the cuts along Energy Transfer axis one just specify -udw3

```
orca_mapspc test.el_inel.rocis.rixssoc RIXS -x0871 -x1876 -x2-1 -x34 -w0.4 -g0.4
       -l -n125 -m125 -dx20 -eaxis1 -udw3
```

Then at the end one gets:

```
Making the specified  cuts (2) at constant Excitation Energy axis...
       Writing file: test.el_inel.rocis.rixssoc_872.50.rxes_vs.dat  ...Done
       Writing file: test.el_inel.rocis.rixssoc_872.50.rxes_fs.dat  ...Done
```

(continues on next page)

(continued from previous page)

```
        Writing file: test.el_inel.rocis.rixssoc_874.20.rxes_vs.dat  ...Done
        Writing file: test.el_inel.rocis.rixssoc_874.20.rxes_fs.dat  ...Done
        Done

        Making the specified cuts (3) at constant Emission/Energy Transfer axis...
        Writing file: test.el_inel.rocis.rixssoc_-1.00.xas_vs.dat  ...Done
        Writing file: test.el_inel.rocis.rixssoc_-1.00.xas_fs.dat  ...Done
        Writing file: test.el_inel.rocis.rixssoc_1.50.xas_vs.dat  ...Done
        Writing file: test.el_inel.rocis.rixssoc_1.50.xas_fs.dat  ...Done
        Writing file: test.el_inel.rocis.rixssoc_4.00.xas_vs.dat  ...Done
        Writing file: test.el_inel.rocis.rixssoc_4.00.xas_fs.dat  ...Done
        Done
        All Done
        ------------------------------------------------------------------------------
```

The files `*_rxes_fs.dat` are RXES spectra containing all individual contributions from all Final states together with the Direct, the Total and the Interference contributions at the given constant Incident Energy.

Similarly, the `*_rxes_vs.dat` are RXES spectra containing individual contributions of the Intermediate states, together with the Direct the Total and the Interference contributions at the given constant Incident Energy

Likewise, the respective `*_xas_fs.dat` and `*_xas_vs.dat` are XAS type spectra with individual contributions at a given constant Emission or Energy transfer Energy

These files are Energy vs Intensity files and read like:

1) for `*fs.dat`

```
X     S-  1(  0- 0)      S-  2(  0- 1)       DIRECT       TOT       INTERFERENCE
```

2) for `*vs.dat`

```
X     S-  1(  45)        S-  2(  47)         DIRECT       TOT       INTERFERENCE
```

In the first case S -1(0-0) represents the individual contribution of a given Ground-Final state pair. The numbering follows the numbering of the output file e.g.:

```
Eigenvalues:     cm-1         eV      Boltzmann populations at T =  300.000 K
        0:          0.0000     0.0000         3.44e-01
        1:          8.3818     0.0010         3.31e-01
```

Hence, in this case S -1 represents the elastic scattering intensity.

In the second case S -1(45) represents the individual contribution of a given Intermediate state.

```
44:      66918.6071     8.2968         1.43e-140
45:    6996678.8061   867.4775         0.00e+00
46:    6996693.0276   867.4793         0.00e+00
```

In this case S -1 represents the intensity contribution of the first Intermediate state.

Starting from ORCA 4.2 in every RIXS requested calculation the Off resonant XES spectrum is automatically generated in every RIXS requested calculation.

```
------------------------------------------------------------------------------
ROCIS RIXS SPECTRUM
------------------------------------------------------------------------------

Making the RIXS data files for Inelastic and Elastic Scattering
Ground State:                  1
Intermediate States:           28
Final States:                  588
The RIXS cross section will be generated from:
```

(continues on next page)

(continued from previous page)

```
589 Ground-Final State Pairs and   28 Intermediate States/Pair
The Off-Resonance XES spectrum will be printed

Calculating Intensities...
10% done
20% done
30% done
40% done
50% done
60% done
70% done
80% done
90% done
100% done

Printing the XES spectrum and Storing the files...
-------------------------------------------------------------------------------
X-RAY EMISSION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------
Transition          Energy           INT             TX        TY        TZ
(eV)           (fosc)          (au)      (au)      (au)
-------------------------------------------------------------------------------
1  589 ->    0      6403.377     0.000000000721     0.00000   0.00000   0.00000
2  590 ->    0      6403.380     0.000000000083    -0.00000   0.00000   0.00000
3  591 ->    0      6403.685     0.000873238810     0.00236   0.00000   0.00000
4  592 ->    0      6404.766     0.000000000154     0.00000   0.00000   0.00000
5  593 ->    0      6408.288     0.000000006850    -0.00001   0.00000   0.00000
6  594 ->    0      6408.295     0.000034710300    -0.00047   0.00000   0.00000
...
16490  614 ->  588      6387.989     0.000000000000     0.00000   0.00000   0.00000
16491  615 ->  588      6388.222     0.000000000000     0.00000   0.00000   0.00000
16492  616 ->  588      6388.881     0.000000000000     0.00000   0.00000   0.00000
All Done
-------------------------------------------------------------------------------
```

Hence also the myfile-rixs.out file can also be processed with the `orca_mapspc` to generate the respective XES spectra:

```
orca_mapspc myfile_rixs.out XES/XESSOC -x06000 -x16500 -w2.0 -eV -n10000
```

7.31.5 Core PNO-ROCIS, PNO-ROCIS/DFT

It has been shown recently[545] that it is possible to combine the powerful machinery of the PNOs with the ROCIS and ROCIS/DFT methods to formulate the core PNO-ROCIS and PNO-ROCIS/DFT methods. The usage of PNOs here is somewhat unconventional since they are not used to treat electron correlation effects in a state specific manner. Rather, the PNOs are used to identify the relevant part of the virtual space that can be reached by excitation out of local core orbitals. This subspace of the virtual space is local, thus leading to a linear scaling, state universal method.

The PNO-ROCIS calculations can be requested with the following keywords:

```
...
DoPNO true #Flag to call the PNO truncation
TCutPNO 1e-11#Threshold to cutout the PNO populations
XASElems 0 #Number of the involved element to the calculated core XAS calculation
OrbWin = 0,0,0,2000
...
```

As has been shown in reference[545] a universal TCutPNO 1e-11 threshhold can be defined for all edges provided that the PNOs are constructed by taking into account all the availiable core orbitals in the systems. For example in

the case of a 1st row transition metal this will be the 9 1s, 2s, 2p, 3s and 3p MOs. These orbitals will be identified automatically by the program provided that the element or the elements for which the XAS calculation will be performed are specified within the XASElems keyword. In the following example these correspond to Core MOs 36-44. Note that the CoreMOs list should not be confused with the OrbWin which is used to specify the excitation space that will be actually used in the actual calculation.

```
============================================
Core PNO/ROCIS truncation
============================================

--------------------------------------------
Calculating Integrals...
--------------------------------------------

...

--------------------------------------------
Calculating Guess Amplitudes and Densities...
--------------------------------------------

-----------------------------------------------------------------
The densities will be generated from the Detected Core MOs:
-----------------------------------------------------------------

MO= 36, E= -261.246087 Eh
MO= 37, E=  -31.777896 Eh
MO= 38, E=  -27.263122 Eh
MO= 39, E=  -27.263122 Eh
MO= 40, E=  -27.263122 Eh
MO= 41, E=   -3.914132 Eh
MO= 42, E=   -2.457405 Eh
MO= 43, E=   -2.457405 Eh
MO= 44, E=   -2.457405 Eh
```

Alternatively one can also use the CoreMOs keyword to individual select the respective CoreMOs

```
...
DoPNO true #Flag to call the PNO truncation
TCutPNO 1e-11#Threshold to cutout the PNO populations
CoreMOs 0,1,6,7,8,29,30,31,32  #The core MOs for the selected element
                               #to perform the XAS calculation
OrbWin = 0,0,0,2000
...
```

A complete list of CoreMOs of the different atoms can be found in reference[545] The program will then proceed and generate the Core PNOs and use the TCutPNO threshold to reduce the Virtual MO space. In the following example only virtual orbitals are selected out of the total 1445 virtual MOs

```
TCutPNO:                              1.000e-11
Virtual orbitals before selection:   368 ... 1812 (1445 MO's)
Virtual orbitals after selection:    368 ...  447 ( 80 MO's)
PNO transformation completed in:    177.09 sec
```

From this point and on the programm will proceed the usual way. This will result in extraordinary computation speeding ups without loss in accuracy.

7.31.6 ROCIS Magnetic Properties

Several magnetic properies are availiable in the ROCIS method Including g-tensors (G-Matrix), zero field splittings (ZFS), hyperfine couplings (HFCs) and electric field gradients (EFGs).

The g-tensors as well as the zfs are calculated on the basis of the Effective Hamiltonian as well in the sum over states (SOS) framework. HFCs are calculated in the SOS framework while EFGs are calculated as expectation values. Please consult also the respective discussion in the MRCI chapter (section *The Multireference Correlation Module*)

```
...
DoHeff true    # Requests calculation of G-tenosrs and ZFS
               # in the effective Hamiltonian framework
DoEPR true     # Requests calculation of G-tenosrs, ZFS and HFCs
               # in the Sum over states (SOS) framework
AtensorNuc 0   # Nuclei to account for the HFCs calculation
NAtensors 1    # How many Nuclei are included in the HFCs calculation
ATensor 0      # Nucleus to calculate HFCs and EFGs
NDoubGtensor 1 # Kramers doublets to account for the g tensor calculations
...
```

This will enter the calculation in the ROCIS Spin Hamiltonian section

```
-----------------------------------------------------
ROCIS SPIN HAMILTONIAN PROPERTIES
-----------------------------------------------------
```

7.31.7 Keyword List

```
%rocis
#------------------------------------------------------------
# GENERAL KEYWORDS
#------------------------------------------------------------
NRoots 3          # The number of desired roots
MaxDim 5          # Davidson expansion space = MaxDim * NRoots
MaxIter 35        # Maximum CI Iterations
NGuessMat 512     # The dimension of the guess matrix
ETol 1e-6         # Energy convergence tolerance
RTol 1e-6         # Residual Convergence tolerance
MaxCore 2000      # Maximum memory used during the calculation in MB
EWin= -5,5,-5,5   # Energy Window that defines orbital excitation space
OrbWin=6,8,0,2000 # Orbital Window that defines orbital excitation space
                  # (overrides EWin)
DoRI false        # Switch for the RI approximation
DoLoc false       # Switch for localization of Donor orbital space
LocMet PipekMezey # chooses the localization method:
                  # PipekMezey or FosterBoys.
                  # Abbreviations "PM" and "FB"
                  # are equivalent to full names.
SOC false         # Switch for inclusion of SOC

TEMPERATURE 10    # The fictive temperature for the
                  # SOC corrected spectra
DoDFTCIS false    # Switch for the DFT/ROCIS method
DFTCIS_C = 0.18, 0.20, 0.40 #Array Input of the
                  # three DFT/ROCIS parameters

#------------------------------------------------------------
# FLAGS FOR EXCITATION SPACES
#------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
Do_is true          # Include DOMO->SOMO excitations
Do_sa true          # Include SOMO->Virtual excitation
Do_ia true          # Include DOMO->Virtual excitations
Do_ista true        # Include DOMO->SOMO excitations
                    # coupled to SOMO->Virtual
                    # excitations with s not equal t
Do_isa true         # Include DOMO->SOMO excitations
                    # coupled to SOMO->Virtual
                    # excitations with s = t
DoLowerMult false   # Switch for excitation with S'=S-1
Do_LM_is true       # Include DOMO->SOMO excitations
                    # with S'=S-1
Do_LM_sa true       # Include SOMO->Virtual excitations
                    # with S'=S-1
Do_LM_ia true       # Include DOMO->Virtual excitations
                    # with S'=S-1
Do_LM_ss true       # Include SOMO->SOMO excitations
                    # with S'=S-1
DoHigherMult false  # Switch for DOMO->Virtual
                    # excitations with S'=S+1


#------------------------------------------------------------
OUTPUT KEYWORDS
#------------------------------------------------------------
PrintLevel 3                # Controls the amount of output
                            # produced during the calculation
RIXS false                  # Perform a RIXS calculation
RIXSSOC false               # Perform a RIXS calculation on the basis
                            # of relativistically corrected states
Elastic false               # Include the elastic line in the generation
                            # of the RIXS or RIXSSOC spectra
PlotDiffDens = 1,2          # Array input for plotting
                            # difference densities of CI roots
                            # 1 and 2 to the ground state.
PlotSOCDiffDens=1,2         # Array input for plotting
                            # difference densities of SOC
                            # states 1 and 2 to the ground state
DoNTO false                 # Request Natural Transition Orbital Analysis
DoNDO false                 # Request Natural Difference Orbital Analysis
                            # (if true it switches off the NTO analysis)

NDOThresh 1e-4              # Threshold for printing occupation numbers
NTOThresh 1e-4              # Threshold for printing occupation numbers
NDOStates = 1,2             # Array input for states to be taken into account
NTOStates = 1,2             # Array input for states to be taken into account
TPrint 0.01                 # Threshold for contributions to CI
                            # and SOC states to be printed
DoPNO false                 # Performs the calculation in the PNO-ROCIS framework

DoCD true                   # Request circular dichroism calculation
DoDipoleLength true         # Request the use of electric moments in a length formulation
DoDipoleVelocity true       # Request the use of electric moments in a velocity formulation
DoHigherMoments true        # Request the calculation of electric quadrupole and magnetic
                            # dipole moments contributions
DoFullSemiclassical true    # Request the calculation of complete semiclassical
                            # multipolar moments
DecomposeFoscLength true    # Request the decomposition of the oscillator strengths
                            # in a multipolar expansion under a length formulation
DecomposeFoscVelocity true  # Request the decomposition of the oscillator strengths
                            # in a multipolar expansion under a velocity formulation
```

7.32 Excited States via MC-RPA

MC-RPA excitation energies and transition moments are computed from the poles and residues of the linear response function of CASSCF a wave function.[380, 417, 901] By following similar lines, it is in principle possible to compute any kind of static and dynamic molecular property that is based on analytic derivatives of the CASSCF energy, which may be available in future releases of ORCA.

7.32.1 General Description

The starting point of response theory for variational wave functions like CASSCF is the time-dependent (TD) Schrödinger equation in its phase-isolated form[174]

$$\Big(\hat{H} - i\frac{\partial}{\partial t} - Q\Big)|\tilde{0}\rangle = 0$$

with the TD quasi energy

$$Q(t) = \langle\tilde{0}|\Big(\hat{H} - i\frac{\partial}{\partial t}\Big)|\tilde{0}\rangle.$$

The Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}^t$ consists of the unperturbed time-independent Hamiltonian $\hat{H}_0$ and a TD perturbation

$$\hat{V}^t = \sum_k e^{-i\,\omega_k\,t} \sum_x \varepsilon_x(\omega_k)\,\hat{X}$$

which is described as a sum of periodic perturbations, i.e. a Fourier series. TD molecular properties are obtained by applying the TD variational principal

$$\delta\{Q(t)\}_T = 0$$

up to a certain order in the time-averaged quasi energy

$$\{Q(t)\}_T = \frac{1}{T}\int_{-T/2}^{T/2} Q(t)\,dt$$

while $\{Q(t)\}_T$ is expanded by the perturbation strengths ε_X at vanishing frequencies $\omega_k = 0$. Applying the TD variational principle for the second-order quasi energy leads to

$$0 = \delta\{Q(t)\}_T^{(2)} \Longrightarrow \frac{\partial Q^{XY}(-\omega_Y,\omega_Y)}{\partial \boldsymbol{\lambda}_X(-\omega_Y)} = \Big(\mathbf{E}^{(2)} - \omega_Y\,\mathbf{S}^{(2)}\Big)\boldsymbol{\lambda}^Y - \mathbf{V}^Y = 0. \tag{7.239}$$

The first-order response equations *(7.239)* become singular if the perturbation frequency ω_Y approaches any eigenvalue

$$\mathbf{E}^{(2)}\,\boldsymbol{\lambda} = \mathbf{S}^{(2)}\,\boldsymbol{\lambda}\,\omega,$$

of second-derivative matrices $\mathbf{E}^{(2)}$ and $\mathbf{S}^{(2)}$. The eigenvalues ω correspond to the electronic excitation energies. The second-derivative matrices $\mathbf{E}^{(2)}$ and $\mathbf{S}^{(2)}$ have a paired structure as both kind of operators that express orbital excitation and de-excitations are involved:

$$\left[\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}^* & \mathbf{A}^*\end{pmatrix} - \omega\begin{pmatrix}\boldsymbol{\Sigma} & \boldsymbol{\Delta}\\ -\boldsymbol{\Delta}^* & -\boldsymbol{\Sigma}^*\end{pmatrix}\right]\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}^*\end{pmatrix} = \begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}$$

The eigenvalue equations above are valid for all variational wave functions methods, e.g. DFT, HF, CASSCF etc. The only difference is the operator manifold and the unperturbed Hamiltonian $\hat{H}_0$ that is used. For the CASSCF linear response and eigen value equations, the super matrices $\mathbf{A}$, $\mathbf{B}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{\Delta}$ have the following structure:

$$\mathbf{A} = \begin{pmatrix}\langle 0|\,[q_i,[\hat{H}_0,q_j^\dagger]]\,|0\rangle & \langle 0|\,[[q_i,\hat{H}_0],R_j^\dagger]\,|0\rangle\\ \langle 0|\,[R_i,[\hat{H}_0,q_j^\dagger]]\,|0\rangle & \langle 0|\,[R_i,[\hat{H}_0,R_j^\dagger]]\,|0\rangle\end{pmatrix}\quad \boldsymbol{\Sigma} = \begin{pmatrix}\langle 0|\,[q_i,q_j^\dagger]\,|0\rangle & \langle 0|\,[q_i,R_j^\dagger]\,|0\rangle\\ \langle 0|\,[R_i,q_j^\dagger]\,|0\rangle & \langle 0|\,[R_i,R_j^\dagger]\,|0\rangle\end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix}\langle 0|\,[q_i,[\hat{H}_0,q_j]]\,|0\rangle & \langle 0|\,[[q_i,\hat{H}_0],R_j]\,|0\rangle\\ \langle 0|\,[R_i,[\hat{H}_0,q_j]]\,|0\rangle & \langle 0|\,[R_i,[\hat{H}_0,R_j]]\,|0\rangle\end{pmatrix}\quad \boldsymbol{\Delta} = \begin{pmatrix}\langle 0|\,[q_i,q_j]\,|0\rangle & \langle 0|\,[q_i,R_j]\,|0\rangle\\ \langle 0|\,[R_i,q_j]\,|0\rangle & \langle 0|\,[R_i,R_j]\,|0\rangle\end{pmatrix}$$

The TD CASSCF wave function is expressed in terms of orbital excitation $q_i^\dagger$ and de-excitation operators q_i,

$$q_i^\dagger = E_{pq} = a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta}, \quad q_i = E_{qp}, \quad \text{with} \quad p > q$$

as well as so called state transfer operators

$$R_i^\dagger = |i\rangle \langle 0|, \quad R_i = |0\rangle \langle i|, \quad \text{with } i \neq 0$$

that account for relaxation of orbitals and CI coefficients when perturbed by an electromagnetic field, respectively.

The eigenvalue (and response) equations are solved iteratively by a customized version of the Davidson algorithm that simultaneously determines the N lowest lying roots. The most time-consuming step is the transformation of the trial vectors that contain an orbital and CI coefficient part with the electronic Hessian matrix $\mathbf{E}^{(2)}$. The working equations are very similar to those of the CASSCF electronic gradient that is computed when minimizing the CASSCF ground state energy.

As a show case example the UV/Vis spectrum of a Nickel dimethylglyoximato complex (Ni(dmg) $_2$) was simulated with both SA-CASSCF and MC-RPA. A CAS (12/9) with the 3d electrons on Ni and 4 π orbitals and electrons from the ligands was selected; the def2-SVP basis set was used. For SA-CASSCF we have averaged over 21 states while for MC-RPA the 20 lowest roots were determined. Though both UV/Vis spectra have two intense peaks, their excitation energies and oscillator strengths differ quite substantially. This can be attributed to the lack of state-specific orbital relaxation that is only available in MC-RPA. In subsection *Natural Transition Orbitals* the most important natural transition orbitals[380, 561] and active natural orbitals of MC-RPA and SA-CASSCF are shown, respectively.

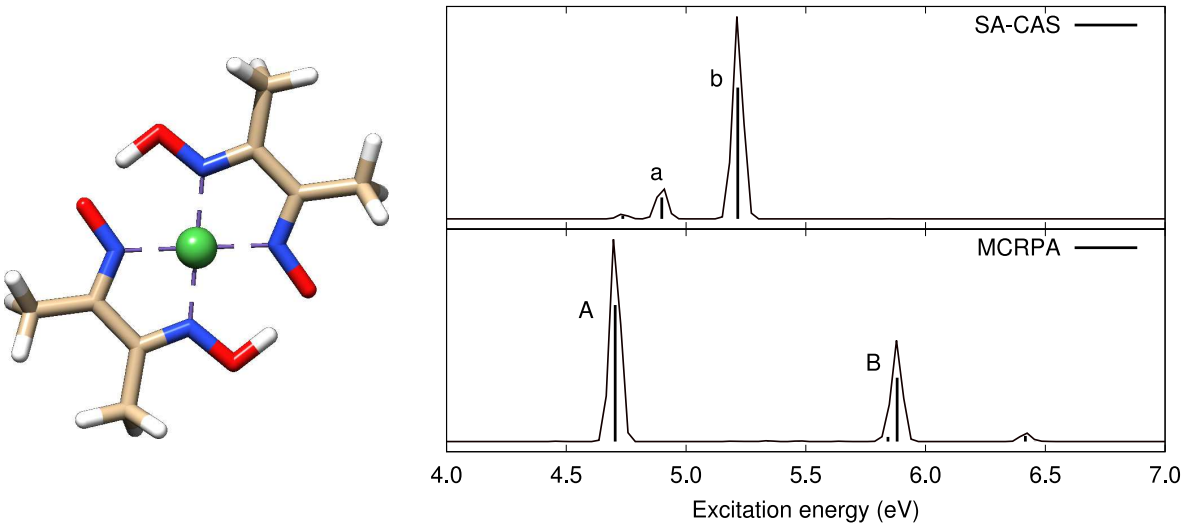


Fig. 7.35: Calculated UV/Vis spectra of Ni(dmg)$_2$.

7.32.2 Detecting CASSCF Instabilities

Selecting the right orbitals for the active space is not always an easy task. A wrong selection may lead to convergence to excited states or saddle points when minimizing the CASSCF energy. Such an instability in the wave function can be detected by computing the lowest excitation energy, i.e. the lowest root of the electronic Hessian with MC-RPA.

Instabilities may occur even for the simplest cases if the starting orbitals for CASSCF energy calculation were inappropriate. Let us look at a benzene CAS(6/6) calculation where we started from the `model potential` initial guess for the MOs.

```
! TZVP Def2-TZVP/C VeryTightSCF
```

(continues on next page)

(continued from previous page)

```
%pal
 nprocs = 20
end

%casscf
 nel           6
 norb          6
 mult          1
 nroots        1
 TrafoStep     RI
 gtol  1e-8
 etol  1e-12
end

%mcrpa
 nroots        1
 TolR        1e-4
 MaxRed        80
end

*xyz 0 1
H           0.000000       2.484212       0.000000
H           0.000000      -2.484212       0.000000
H           2.151390       1.242106       0.000000
H          -2.151390      -1.242106       0.000000
H          -2.151390       1.242106       0.000000
H           2.151390      -1.242106       0.000000
C           0.000000       1.396792       0.000000
C           0.000000      -1.396792       0.000000
C           1.209657       0.698396       0.000000
C          -1.209657      -0.698396       0.000000
C          -1.209657       0.698396       0.000000
C           1.209657      -0.698396       0.000000
*
```

The energy converges smoothly

```
E(CAS)=  -230.560657053 Eh DE=     0.000000000
E(CAS)=  -230.767928524 Eh DE=    -0.207271471
E(CAS)=  -230.810472828 Eh DE=    -0.042544304
E(CAS)=  -230.811818980 Eh DE=    -0.001346152
E(CAS)=  -230.812866285 Eh DE=    -0.001047305
E(CAS)=  -230.812930081 Eh DE=    -0.000063796
E(CAS)=  -230.812944302 Eh DE=    -0.000014221
E(CAS)=  -230.812944635 Eh DE=    -0.000000333
E(CAS)=  -230.812943979 Eh DE=     0.000000656
E(CAS)=  -230.812944834 Eh DE=    -0.000000856
E(CAS)=  -230.812944944 Eh DE=    -0.000000110
E(CAS)=  -230.812944943 Eh DE=     0.000000001
E(CAS)=  -230.812944952 Eh DE=    -0.000000009
E(CAS)=  -230.812944953 Eh DE=    -0.000000000
```

as the gradient norm does

```
||g|| =      2.538097040 Max(G)=    -0.486818498 Rot=144,0
||g|| =      0.850498225 Max(G)=     0.219916319 Rot=23,11
||g|| =      0.192712320 Max(G)=    -0.055714161 Rot=23,11
||g|| =      0.144524323 Max(G)=     0.035741221 Rot=23,11
||g|| =      0.034101354 Max(G)=    -0.011346113 Rot=23,17
||g|| =      0.016336825 Max(G)=     0.005232633 Rot=122,17
||g|| =      0.003090776 Max(G)=     0.000935457 Rot=122,17
```

(continues on next page)

(continued from previous page)

```
||g|| =      0.002539517 Max(G)=    -0.000597246 Rot=21,16
||g|| =      0.004134603 Max(G)=    -0.000980627 Rot=21,16
||g|| =      0.001410672 Max(G)=     0.000353494 Rot=21,16
||g|| =      0.000506790 Max(G)=    -0.000172518 Rot=23,17
||g|| =      0.000408528 Max(G)=    -0.000191970 Rot=122,17
||g|| =      0.000087961 Max(G)=    -0.000042271 Rot=23,17
||g|| =      0.000107447 Max(G)=    -0.000041858 Rot=23,17
||g|| =      0.000139470 Max(G)=    -0.000058693 Rot=23,11
||g|| =      0.000094867 Max(G)=     0.000020639 Rot=23,11
||g|| =      0.000033056 Max(G)=    -0.000011307 Rot=23,11
||g|| =      0.000009908 Max(G)=    -0.000005192 Rot=23,11
||g|| =      0.000013678 Max(G)=    -0.000007007 Rot=23,11
||g|| =      0.000011838 Max(G)=     0.000004991 Rot=23,17
||g|| =      0.000006431 Max(G)=    -0.000002076 Rot=122,17
||g|| =      0.000008817 Max(G)=    -0.000002828 Rot=122,17
||g|| =      0.000012070 Max(G)=     0.000004189 Rot=23,17
||g|| =      0.000001891 Max(G)=     0.000001284 Rot=23,17
||g|| =      0.000003505 Max(G)=     0.000001187 Rot=23,17
||g|| =      0.000002121 Max(G)=    -0.000000447 Rot=122,17
||g|| =      0.000002233 Max(G)=    -0.000000508 Rot=122,17
||g|| =      0.000000933 Max(G)=     0.000000494 Rot=23,17
||g|| =      0.000000711 Max(G)=     0.000000369 Rot=23,17
||g|| =      0.000000430 Max(G)=     0.000000230 Rot=23,17
||g|| =      0.000000200 Max(G)=     0.000000047 Rot=23,11
||g|| =      0.000000103 Max(G)=     0.000000030 Rot=23,11
||g|| =      0.000000025 Max(G)=     0.000000005 Rot=20,16
```

Though we have reached convergence for a CASSCF ground state energy calculation, the MC-RPA calculation however detects an instability

```
Davidson Eigenvalue solver (Iteration 10)

 State      Eigenvalue     RMSD error        Converged
   0          0.2405996888      1.4767724243e-01 F

 WARNING
 1 null space vectors in reduced space Hessians!
 This indicates an instability in your reference wave function!

 Davidson Eigenvalue solver (Iteration 11)

 State      Eigenvalue     RMSD error        Converged
   0          0.0000000000      0.0000000000e+00 T
```

by finding positive-indefiniteness by a Cholesky decomposition of the reduced space Hessians.

Instabilities in the CASSCF wavefunction can usually be avoided by carefully monitoring the active space orbitals in the

```
----------------------------
LOEWDIN REDUCED ACTIVE MOs
----------------------------
```

section of the CASSCF output.

```
18        19        20        21        22        23
                  -0.62274  -0.32864  -0.32863   0.15983   0.16011   0.77971
                   1.99910   1.93810   1.93808   0.06203   0.06195   0.00075
                  --------  --------  --------  --------  --------  --------
 2 H  s              11.6       0.0       0.0       0.0       0.0      12.8
 3 H  s              11.6       0.0       0.0       0.0       0.0      12.8
```

(continues on next page)

(continued from previous page)

```
 4 H  s             11.6       0.0       0.0       0.0       0.0      12.8
 5 H  s             11.6       0.0       0.0       0.0       0.0      12.8
 6 C  pz             0.0       0.0      32.2       0.0      30.9       0.0
 7 C  pz             0.0       0.0      32.2       0.0      30.9       0.0
 8 C  pz             0.0      24.1       8.0      23.2       7.7       0.0
 8 C  px             6.7       0.0       0.0       0.0       0.0       3.0
 9 C  pz             0.0      24.1       8.0      23.2       7.7       0.0
 9 C  px             6.7       0.0       0.0       0.0       0.0       3.0
10 C  pz             0.0      24.1       8.0      23.2       7.7       0.0
10 C  px             6.7       0.0       0.0       0.0       0.0       3.0
11 C  pz             0.0      24.1       8.0      23.2       7.7       0.0
11 C  px             6.7       0.0       0.0       0.0       0.0       3.0
```

In this particular example, MOs 18 and 23 are not part of the π system and have to be rotated with orbitals 16 and 31, respectively. After rotating all π orbitals into the active space, the CASSCF converges to a lower energy.

```
--------------
CASSCF RESULTS
--------------

Final CASSCF energy       : -230.844448647 Eh   -6281.5968 eV
```

The electronic CASSCF Hessian is now positive definite and the lowest MC-RPA excitation energy becomes

```
STATE  1:  E=   0.171023 au      4.654 eV    37535.3 cm**-1
```

7.32.3 Natural Transition Orbitals

Natural transition orbitals[380, 561] (NTO) are obtained from a singular value decomposition of the MC-RPA ground-to-excited state (f) transition density matrices $\rho_{pq}^{0\to f}$. As for TD-DFT and ROCIS one obtains two sets of orbitals for each state that describe the donation (occupied and active) and acceptance (active and virtual) of an electron in the electronic transition. The orbital structure of $\rho_{pq}^{0\to f}$ for CASSCF wave functions is illustrated in Fig. 7.36.

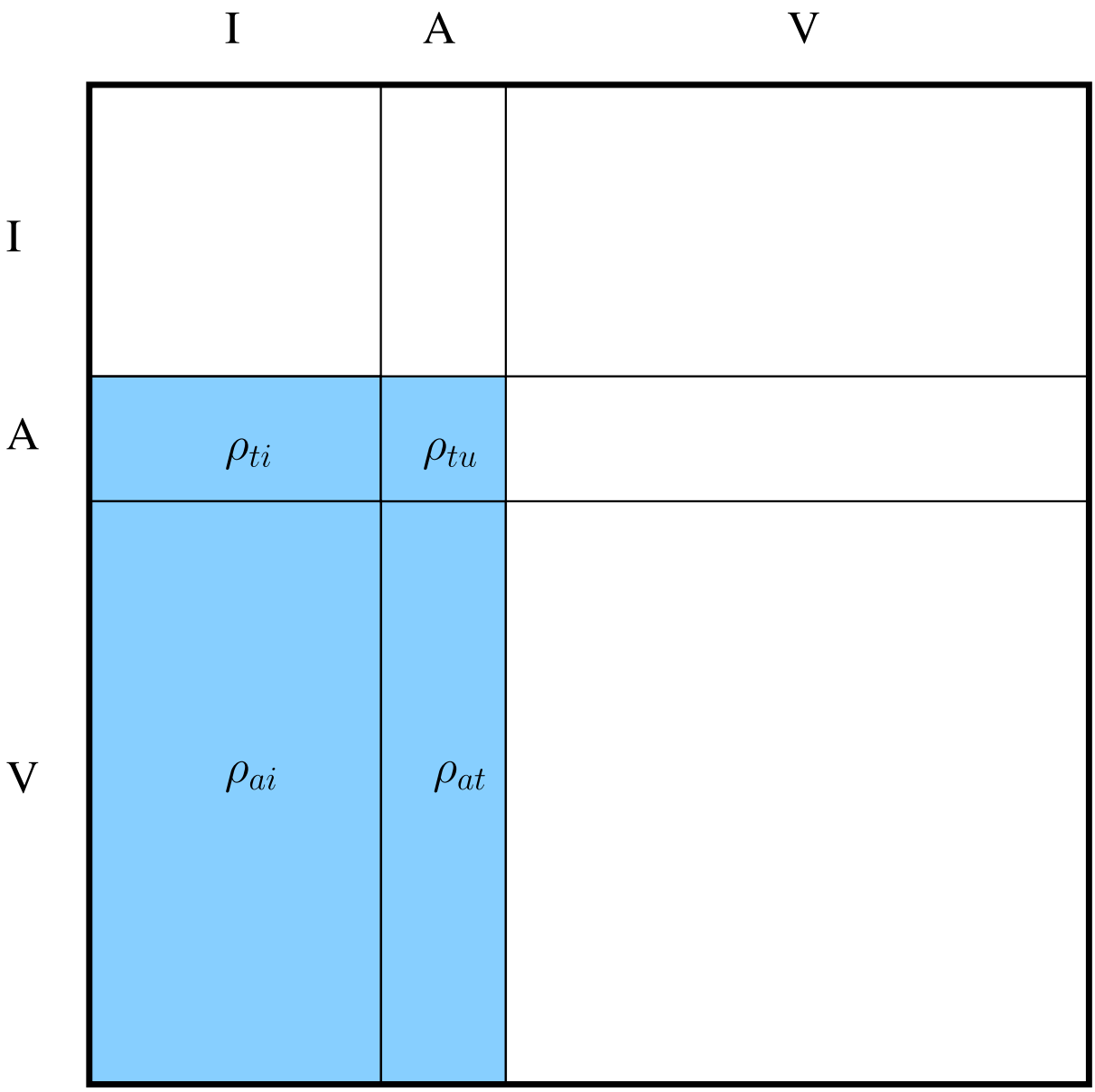


Fig. 7.36: Structure of MC-RPA transition density matrix $\rho_{pq}^{0\to f}$

The compute NTOS only the following flag in the input has to switched ON:

```
%mcrpa
 nroots 20
 DoNTO true
 NTOThresh 5e-5
end
```

This will compute all NTOs with a singular value larger then the `NTOThresh` threshold for ALL roots.

```
----------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE   12
----------------------------------------

Natural Transition Orbitals were saved in ni-dmg-2-svp-cas-12-9-mcrpa.s12.nto-donor
Threshold for printing occupation numbers 1.0000e-03
Natural Transition Orbitals were saved in ni-dmg-2-svp-cas-12-9-mcrpa.s12.nto-acceptor
Threshold for printing occupation numbers 1.0000e-03
```

(continues on next page)

(continued from previous page)

```
STATE 13:  E=   0.214726 au      5.843 eV    47126.9 cm**-1

    77 ->  69  : n=  0.16680786
    76 ->  70  : n=  0.06575768
    75 ->  71  : n=  0.02841330
    74 ->  72  : n=  0.01485889
    73 ->  73  : n=  0.01138840
    72 ->  74  : n=  0.01099324
    71 ->  75  : n=  0.00899487
    70 ->  76  : n=  0.00764206
    69 ->  77  : n=  0.00546103
    68 ->  78  : n=  0.00517046
    67 ->  79  : n=  0.00511544
    66 ->  80  : n=  0.00485839
    65 ->  81  : n=  0.00405297
    64 ->  82  : n=  0.00379270
    63 ->  83  : n=  0.00325052
    62 ->  84  : n=  0.00315477
    61 ->  85  : n=  0.00297172
    60 ->  86  : n=  0.00291081
    59 ->  87  : n=  0.00268810
    58 ->  88  : n=  0.00243609
    57 ->  89  : n=  0.00240536
    56 ->  90  : n=  0.00238574
    55 ->  91  : n=  0.00183946
    54 ->  92  : n=  0.00181492
    53 ->  93  : n=  0.00165943
    52 ->  94  : n=  0.00154428
    51 ->  95  : n=  0.00146299
    50 ->  96  : n=  0.00137434
    49 ->  97  : n=  0.00136656
    48 ->  98  : n=  0.00128274
    47 ->  99  : n=  0.00121332
    46 -> 100  : n=  0.00117752
    45 -> 101  : n=  0.00107962
    44 -> 102  : n=  0.00105733
    43 -> 103  : n=  0.00104762
```

For the above example, the most important (controlled by `NTOThresh`) donating and accepting NTOs of state 13 are written to the `gbw`-type files

```
ni-dmg-2-svp-cas-12-9-mcrpa.s12.nto-donor
ni-dmg-2-svp-cas-12-9-mcrpa.s12.nto-acceptor
```

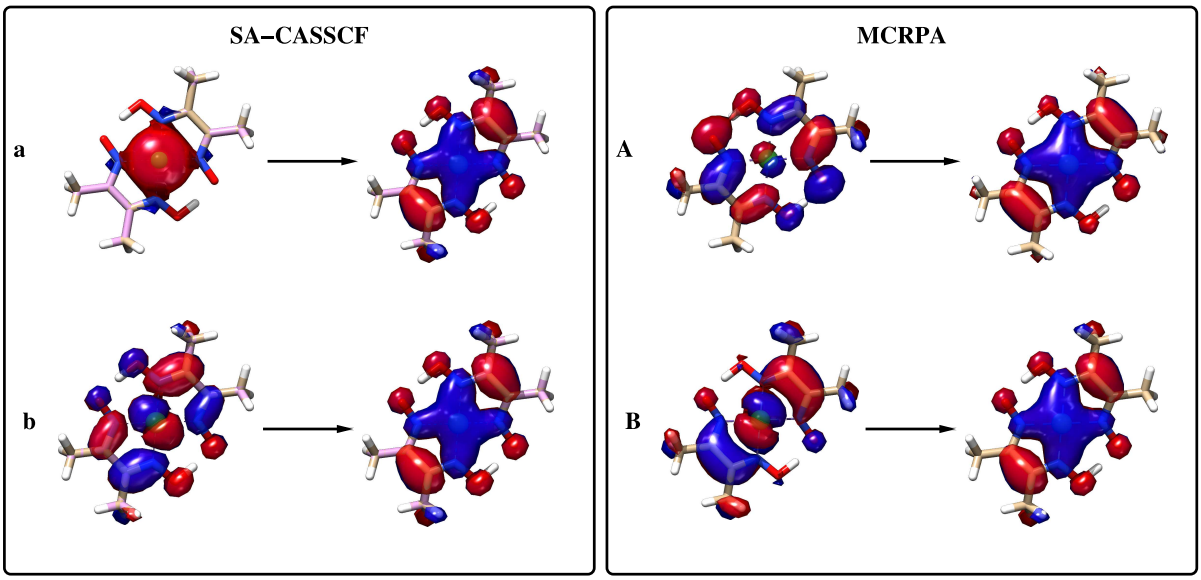
and can be plotted with the `orca_plot` program (see Sec. *orca_plot*)

```
orca_plot ni-dmg-2-svp-cas-12-9-mcrpa.s12.nto-donor -i
```

Please be aware of the different indices for states in the in and output!

To compute less or more NTOs the threshold `NTOThresh` can be adapted accordingly.

Let us come back to the UV/Vis spectra of Ni(dmg)$_2$. For the two most intense peaks the natural orbitals and NTOs of MC-RPA and SA-CASSCF, respectively, are shown in Fig. 7.37. While the most intense peak in each spectrum (b and A) correspond to the same $\pi \rightarrow \pi^*$ excitation, transition a and B are complete different, i.e. $d \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$.



(a) (a) SA natural orbitals

(b) (b) NTOs

Fig. 7.37: Calculated UV/Vis spectra of Ni(dmg)$_2$.

7.32.4 Computational Aspects

The code is intended to be used for medium-sized and larger open-shell molecules. It has the same scaling as ORCA's first-order CASSCF energy implementation though a larger pre-factor as the computational cost grow "in principle" linearly with the number of roots.

The implementation is AO-driven meaning that the computational bottleneck is the Fock matrix construction for the several state-specific pseudo AO densities. Note that there are up to 6 pseudo AO densities for each state. The computational costs can be reduced significantly if the `RIJCOSX` approximation is employed, which is highly recommended.

The second most expensive part of the MC-RPA computation are the two-electron integrals with 3 active indices g_{ptuv}. As we aim for running calculations on larger systems, there is only an implementation of the integral transformation that uses the resolution-of-the-identity (RI) approximation.

The restrictions on the auxiliary basis sets are the same as for the CASSCF code (Sec. *General Description*). That is

- If the Fock matrices are constructed in `Direct` or `Conventional` mode, the /C bases are used for the RI approximation of the g_{ptuv} integrals.
- If the `RIJCOSX` approximation for the Fock matrices is employed, the /JK bases are used for both the Fock matrices and the g_{ptuv} integrals.

Note that MC-RPA implementation can be run in parallel with MPI which allows for computing UV/Vis and ECD spectra large open-shell molecules in a limited amount of time.

Before starting running MC-RPA, it is recommended to converge the state specific CASSCF energy calculation until you hit the point of stagnating convergence. Note that property calculations in general assume vanishing electronic gradients otherwise numerical issues in the eigenvalue / response equations may occur.

7.32.5 Keyword List

```
%mcrpa

       NRoots    0         # The number of desired roots

       TolR      1e-5      # Convergence threshold for residual norm
       TolLDP    1e-5      # linear dependency threshold for generalized eigenvalue problem
       PrintWF   CFG       # print CI part in (CFG (default), CSF, or DET basis)
       TolPrint  1e-2      # Threshold for printing elements of excitation vector

       MaxRed    200       # maximum size of reduces space for ALL VECTORS IN TOTAL
       MaxIter   100       # maximum number of (Davidson) iterations

       TDA       false     # Switch off full TD-CASSCF (Tamm-Dancoff approximation)

       DoNTO     false     # Generate Natural Transition Orbitals
       NTOThresh 1e-3      # Threshold for printing occupation numbers

       DoCD      true      #Request circular dichroism calculation
       DoDipoleLength true #Request the use of electric moments in a length formulation
       DoDipoleVelocity true #Request the use of electric moments in a velocity formulation
       DoHigherMoments true  #Request the calculation of electric quadrupole and magnetic␣
↪dipole moments contributions
       DoFullSemiclassical true #Request the calculation of complete semiclassical multipolar␣
↪moments
       DecomposeFoscLength true #Request the decomposition of the oscillator strengths in a␣
↪multipolar expansion under a length formulation
       DecomposeFoscVelocity true #Request the decomposition of the oscillator strengths in a␣
↪multipolar expansion under a velocity formulation
```

7.33 Excited States via EOM-CCSD

The EOM-CCSD routine is part of the `orca_mdci` module of the ORCA program package. It is called after a successful coupled-cluster calculation, if the appropriate flags and the number of roots have been set. In the following chapter the general program flow and all input parameters of the EOM routine will be described in detail (for typical use, see *Excited States with EOM-CCSD*). For an RHF or UHF reference, the EE-, IP- and EA-EOM-CCSD approaches are available for the computation of excitation energies, ionization potentials and electron affinities, respectively. Currently, the following simple input keywords are available:

```
!EOM-CCSD      # same as !EE-EOM-CCSD
!EE-EOM-CCSD   # EOM for electronically excited states
!IP-EOM-CCSD   # IP version
!EA-EOM-CCSD   # EA version
```

7.33.1 General Description

The EOM wave function is parametrized in the following manner

$$\mathcal{R}|\Psi_{CC}\rangle,$$

i.e. via the action of a linear excitation operator $\mathcal{R}$ on the coupled-cluster ground state wave function Ψ_{CC}. Here, $\mathcal{R}$ is a particle conserving operator, in the case of the excitation energy problem. However, this is not true for the ionization or electron attachment problems, where $\mathcal{R}$ is a net annihilation or net creation operator, respectively. The ground state coupled-cluster T-amplitudes are obtained from a CCSD calculation, and our task is to obtain $\mathcal{R}$. Note that since the CC Hamiltonian is nonsymmetric, a left hand solution ($\mathcal{L}$) would also be needed to evaluate properties. For the calculation of excitation, ionization or electron attachment energies, however, it is enough to

obtain the right hand solutions ($\mathcal{R}$). In principle, this is done by building the Hamiltonian and diagonalizing it in order to obtain energy expectation values.

In practice, the size of the CCSD Hamiltonian matrix is prohibitively large and thus, various methods have been devised to obtain its lowest few eigenvalues and eigenstates. One of the most popular of these approaches is the Davidson method, which can be summarized as follows:

- Construct an initial guess of orthogonal trial vectors, C.
- Evaluate the sigma vectors $\sigma = HC$.
- Build model Hamiltonian $\mathcal{H} = C^T\sigma$.
- Diagonalize $\mathcal{H}$: $\mathcal{E} = \mathcal{U}^T\mathcal{H}\mathcal{U}$.
- Compute Ritz vectors $X = C\mathcal{U}$.
- Compute residuals $R = X\mathcal{E} - \sigma\mathcal{U}$, check convergence: if yes, pass $X, \mathcal{E}$ as solutions.
- Preconditioning: $T = MR$ (many possible choices for the preconditioner M).
- Check if adding new trial vectors would exceed the maximum number of trial vectors:
  - if no, add T to C, and orthonormalize the united set
  - if yes, then set X as C (orthonormalize if H is nonsymmetric); then add T and orthonormalize

The advantage of the above method is that, instead of the full Hamiltonian, only the sigma vectors have to be explicitly evaluated and stored.

It is also possible to use a lower scaling version of the EOM-CCSD methods, which relies on the perturbative truncation of the coupled-cluster similarity transformed Hamiltonian. Presently, only the second order truncated version (CCSD(2) approximation) is available for closed-shell molecules (RHF). However, it is better to use the PNO based implementation, as it has the cost of EOM-CCSD(2), but its accuracy is comparable to canonical EOM-CCSD.

Below are all the parameters that influence the RHF EOM routine. In the following sections, these parameters will be explained following the solver algorithm described above.

```
%mdci
#EOM parameters - defaults displayed
  DoEOM false          # whether to perform EOM
  UseEOMOptS true      # use optimized sigma routines for singles
  UseEOMOptD true      # use optimized sigma routines for doubles
  NDav 20              # maximum size of reduced space (i.e. 20*NRoots)
  CCSD2 false          # Use the lower scaling CCSD(2) approximation
  CheckEachRoot true   # check convergence for each root separately
  RootHoming true      # apply root homing
  DoLanczos false      # use the Lanczos procedure rather than Davidson
  UseCISUpdate true    # use diagonal CIS for updating
  NInits 0             # number of roots in the initial guess, if 0, use preset value
  DRESS3ES true        # construct the external dressing to singles
                               # or calculate on the fly
  DRESS3ED false       # construct the external dressing to doubles
                               #or calculate on the fly
  DOCOSXEOM false      # use COSX approximation for external exchange term in EOM
  DOAOX3E false        # use COSX approximation for 4 external terms contribution
                               # to 3 external intermediate
  DoRootwise false     # solves for each root separately,
                               # more stable for large number of roots
  DoTDM false          # option for calculation of default transition moment
  Doleft  false           # calculation of exact left vector
  NRootsPerBatch  1  # no of roots calculated together
  FOLLOWCIS false      # follows the initial singles guess
  DoCore true          # initiates ionization or excitation from core orbital
  CoreHole 0           # core orbital from which ionization or excitation is needed
  CVSEP false          # separates core orbital from valence,
```

(continues on next page)

(continued from previous page)

```
  DTol 1e-5            # default for EOM residual threshold
#keywords which affect EOM parameters, but do not belong to the routine itself
  NRoots 9             # number of roots
  OTol 1e-14           # orthogonalization threshold
  KCOpt KC_MO          # method for external exchange formation
        KC_AOX         # when asked for exact TDM calculation
        KC_AOBLAS      # most efficient
  PrintLevel 3         # the amount of information to be printed
  MaxCore 500          # total amount of memory
end
```

In the case of the UHF EOM-CCSD implementation, the parameters that influence a given calculation are provided below.

```
%mdci
#UHF EOM parameters - defaults displayed
  DoEOM false          # whether to perform EOM
  DoAlpha false        # removal/attachment of an alpha electron (IP/EA calculations)
  DoBeta false         # removal/attachment of a beta electron (IP/EA calculations)
  NDav 20              # maximum size of reduced space (i.e. 20*NRoots)
  UseQROs false
  CheckEachRoot  true    # check convergence for each root separately
  RootHoming        true    # apply root homing
  DoLanczos          false   # use the Lanczos procedure rather than Davidson
  DoOlsen              false   # use the Olsen procedure rather than Davidson
  UseCISUpdate    true    # use diagonal CIS for updating
  NInits 0             # number of roots in the initial guess, if 0, use preset value
  DOCOSXEOM false      # use COSX approximation for external exchange term in EOM
  DOAOX3E false        # use COSX approximation for 4 external terms contribution
                       # to 3 external intermediate
  DoRootwise false     # solves for each root separately,
                       # more stable for large number of roots
  NRootsPerBatch  1  # no of roots calculated together
  FOLLOWCIS true       # follows the initial singles guess
  DTol 1e-5            # default for EOM residual threshold
#keywords which affect EOM parameters, but do not belong to the routine itself
  NRoots 9             # number of roots
  OTol 1e-14           # orthogonalization threshold
  KCOpt KC_AOX         # AO exchange for the four external contributions
                       # (the only option available at present)
  PrintLevel 3         # the amount of information to be printed
  MaxCore 500          # total amount of memory
end
```

7.33.2 Memory Management

The most important data coming from the coupled-cluster routine are the ground state energy and wave function, and the molecular integrals. The integrals are then used to create "dressed" integral containers, which allows for an efficient factorization of the EOM equations, since these dressed quantities do not change during the calculation. Most of these are written on disk, with the possible exception of the integral container which has three external labels. This, and the solver files may remain in core if enough memory is available. The program sequentially tries to allocate memory for the files in the order of their importance, and what cannot be kept in core, goes on disk. The order of allocation is as follows: 1. residual vectors, 2. Ritz vectors, 3. three external integrals, 4. sigma vectors and 5. state (trial) vectors, as seen in the example below:

```
--------------------------------
AUTOMATIC CHOICE OF INCORE LEVEL
--------------------------------
```

(continues on next page)

(continued from previous page)

```
Memory available                           ...   6512.00 MB
Memory needed for Residual-vectors         ...     71.27 MB
Memory needed for Ritz-vectors             ...     71.27 MB
Memory needed for 3-ext integrals          ...     92.05 MB
Memory needed for Sigma-vectors            ...   1425.31 MB
Memory needed for State-vectors            ...   1425.31 MB
 -> Final InCoreLevel    ... 5
```

Half of the memory specified with the keyword `MaxCore` is distributed among the five candidates. In the above case, everything fits in memory. Note that these are only the largest contributors to memory consumption, and there should ideally be a safety margin when allocating memory.

In order to estimate the amount of necessary memory, it should be kept in mind that, in the closed shell case, the memory requirements of the residual and Ritz vectors are proportional to $N_R N_P N_V^2$, the three external integrals to $N_R N_O N_V^3$ and the sigma and trial vectors to $N_D N_R N_P N_V^2$, where N_O and N_V are the number of occupied and virtual orbitals, $N_P = \frac{N_O(N_O+1)}{2}$ is the number of occupied pairs, N_R is the number of roots, and N_D is the maximum size of the reduced space. The keyword `NRoots` sets N_R, while `NDav` determines N_D. Luckily, the contributions that, in our experience, are the most important to keep in memory, are also the ones that require the smallest amount of it. It is advisable to use `KCOpt AOBLAS`, as it has the lowest memory requirements.

Note that in the UHF EE-EOM-CCSD implementation, the memory requirements of the residual and Ritz vectors are proportional to $N_R(N_{P_\alpha} N_{V_\alpha}^2 + N_{P_\beta} N_{V_\beta}^2 + N_{O_\alpha} N_{O_\beta} N_{V_\alpha} N_{V_\beta})$, the three external integrals to $N_R(N_{O_\alpha} N_{V_\alpha}^2 + N_{O_\beta} N_{V_\beta}^2 + N_{O_\alpha} N_{V_\alpha} N_{V_\beta}^2 + N_{O_\beta} N_{V_\beta} N_{V_\alpha}^2)$ and the sigma and trial vectors memory requirements are proportional to $N_D N_R(N_{P_\alpha} N_{V_\alpha}^2 + N_{P_\beta} N_{V_\beta}^2 + N_{O_\alpha} N_{O_\beta} N_{V_\alpha} N_{V_\beta})$, where N_{O_α}, N_{O_β}, N_{V_α} and N_{V_β} are respectively, the number of occupied alpha, occupied beta, virtual alpha and virtual beta orbitals and $N_{P_\alpha} = \frac{N_{O_\alpha}(N_{O_\alpha}-1)}{2}$ and $N_{P_\beta} = \frac{N_{O_\beta}(N_{O_\beta}-1)}{2}$ are the number of alpha and beta occupied pairs, respectively.

7.33.3 Initial Guess

The present initial guess in the RHF EOM implementation consists of constructing a CIS Hamiltonian of a certain dimension, and diagonalizing it. The roots are preselected based on the energetic ordering of the diagonal elements of the Hamiltonian. In the UHF case, the guess is constructed from the solutions of a UHF CIS calculation. The number of roots in the initial guess is determined as 20 times the number of roots desired in EOM (`NRoots`) if `NDav` is 20 or smaller, otherwise it is set to `NDav` times the number of EOM roots. If the parameter `NInits` is larger than zero, then the number of initial guess roots will be set to this parameter times `NRoots`. The maximum possible number of roots is the full CIS dimension, ($N_O N_V$ (RHF) or $N_{O_\alpha} N_{V_\alpha} + N_{O_\beta} N_{V_\beta}$ (UHF)) . One should keep in mind, while increasing the number of initial guess vectors, that this corresponds to diagonalizing a matrix of increasing dimension. If, for example `NRoots` is 10, then by default 200 roots are considered in the initial guess (unless it exceeds the size of the CIS space), or if `NInits` is set to 100, then there will be 1000 roots in the guess. In some cases, the roots calculated using EOM may not be the lowest ones, but a few of these may be replaced by some higher roots which are "easier" to find. In such cases, it may help to increase `NRoots` or `NInits` to converge to the proper roots. The program can be made to follow the initial CIS guess by setting `FollowCIS` to true and is necessary if we wish to ionize or excite from inner-valence or core orbitals. In the RHF implementation, the core orbital, from which the ionization or excitation originates, can be specified using the keyword `CoreHole`, in addition to setting `DoCore` and `FollowCIS` to true. The `CoreHole` keyword is quite general and in principle, ionization or excitation processes from any occupied orbital can be specified using this keyword.

7.33.4 Hamiltonian Construction

The Hamiltonian construction begins by calling the sigma routines. In the case of the closed-shell code, the logical variables `UseEOMOptS` and `UseEOMOptD` choose the routines to be used in the evaluation of the singles and doubles sigma vectors, respectively. If true, the optimized sigma routine, using dressed integrals, will be used. This should not be changed, the option is there mainly for debugging purposes. If set to false, an automatically generated, and much slower serial code will be used instead. In the case of the open-shell UHF implementation, optimized sigma routines have been generated using the ORCA Automated Generator Environment (AGE) [474]. In each early iteration, N_R sigma vectors will be determined, except in the case of a restart, where the number of sigma vectors is $2N_R$. For further details on convergence, see *Convergence, Restart, Preconditioning and Subspace Expansion* below.

The most time consuming part of the sigma vector construction is the formation of the external exchange contribution, which can be influenced via the CC keyword `KCOpt`. Currently, there are three options that are compatible with the RHF EOM implementation: `KC_MO`,`KC_AOX` and `KC_AOBLAS` (see the MDCI documentation) and `KC_AOX` is the only option available in the UHF EOM code. The external exchange term can be treated most efficiently using COSX, which in the closed-shell case, leads to average speed ups of 10x for the external exchange term and an overall speedup of 3x for the EOM calculation. This is accompanied by a drastic reduction of the storage cost[233]. The error introduced is below 1 meV, which is 200-fold less than the error bar of the method[233] itself. It is the default for `KCOpt` `KC_AOX` and `KC_AOBLAS` and can be controlled by the keyword `DOCOSXEOM`. The default grid settings for EOM are `GridX 1` and `IntAccX 2.68`.

Once the sigma vectors are available, they are multiplied with the trial vectors to yield the reduced space Hamiltonian. The Hamiltonian is built in a way that, in each iteration, only the new vector products are added to the "edge" of the old Hamiltonian, so that a full build is avoided. It should be clear that the parameter `NDav` plays an important role here, since it determines the maximum size of the Hamiltonian ($N_D N_R$), and also controls how much memory is needed for the trial and sigma vectors, as seen above. Since the choice of this parameter influences convergence properties, it will be discussed further in *Convergence, Restart, Preconditioning and Subspace Expansion*.

7.33.5 Solution of the (Nonsymmetric) Eigenproblem

Following the construction of the Hamiltonian, a nonsymmetric eigensolver is called. In this case, it is possible to have complex eigenvalues. In practice, this is rarely the case, and indicates a problem of some kind. A warning will be given if this happens, however, one may get away with this if it only happens in an isolated iteration step.

Once the eigenvectors are available, they are compared with those of the previous iteration, if root homing is turned on, i.e. if the `RootHoming` keyword is set to true. This means evaluating the overlap of the old and new eigenvectors, in order to keep track of the possible movement of the eigenvectors if root flipping occurs. If converged roots are removed from further iterations (see next section), it is important to keep track of changes in ordering, especially if a converged and a non-converged root is swapped. After diagonalization, the Ritz vectors and residuals can be evaluated.

7.33.6 Convergence, Restart, Preconditioning and Subspace Expansion

Convergence is signaled once a residual square norm based criteria is fulfilled. This criteria is determined by the `CheckEachRoot` keyword. If it is true (default), the convergence of the residual square norm of each root is checked separately. This is due to the fact that different roots converge at a different rate. Once a root is converged, no new trial vectors will be generated, belonging to that vector. This means that the EOM iterations will progressively become faster (until restart). Turning off the rootwise convergence check is possible, but not recommended. In this case, the maximum of all residual square norms is checked for convergence, and all iterations will take roughly the same amount of time since no vectors are removed in any iteration. However, this procedure can be numerically unstable, since the residuals of some roots might become very close to zero, and trying to generate new vectors, which are orthogonal to these, may lead to numerical disaster. In short, the recommended default is having both `CheckEachRoot` and `RootHoming` set to true. If `CheckEachRoot` is false, then `RootHoming` should also be set to false, as it may cause problems if `NDav` is too small. The convergence threshold of the residual in Davidson's method can be larger than that for the ground state CC residual threshold in order to obtain converged results. Namely, a value of `DTol` of 1e-5 is almost always enough to get well converged energies.

At this point it is worth discussing the role of the keyword `NDav`. This keyword determines at what point the Davidson algorithm should be restarted. If it is chosen too small, it may cause slow convergence. If this value is too large, this may result in overwhelming demands on memory/disk space requirements. The default value (20) is chosen with the hope that no, or maybe one restart will be required. It should only be changed if computational resources demand it. However, the treatment of core ionization or core excitation processes often requires a large value of `NDav`. At restart, Ritz vectors are copied as new trial vectors for all roots, which will then be orthonormalized, while new vectors will only be generated for the non-converged roots. This means that the step after the rebuilding of the expansion space will be 1-2 times as expensive as one of the initial steps.

New directions (trial vectors) are generated from the preconditioned residual vectors. If no preconditioning is applied (the preconditioner is taken to be a unit matrix), one falls back to the Lanczos algorithm, which is inferior to the Davidson algorithm. This happens if the keyword `DoLanczos` is true. This is not recommended, as the Lanczos algorithm converges several times slower than Davidson's, and is there for debugging mainly. The original Davidson preconditioner is the inverse of a diagonal matrix which contains the difference of the diagonal elements of the Hamiltonian and the current approximation to the eigenvalue belonging to the given root. Let us consider the closed-shell RHF implementation for simplicity. If R_{ia} and R_{ijab} are elements of the singles and doubles amplitudes, respectively, then the updated vectors (T_{ia}, T_{ijab}) have the form

$$T_{ia} = \frac{R_{ia}}{D_{ia} + \mathcal{E}_R}$$

for singles, and

$$T_{ijab} = \frac{R_{ia}}{D_{ijab} + \mathcal{E}_R}$$

for doubles. Here, D_{ia} and D_{ijab} are related to, and possibly approximations of, the respective diagonal Hamiltonian elements. The simplest approximation is just to construct these from diagonal Fock matrix elements (i.e. orbital energies) as $D_{ia} = \varepsilon_a - \varepsilon_i$ and $D_{ijab} = \varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j$. A slightly better preconditioning can be obtained as follows. For singles, take the exact CIS diagonal elements, $D_{ia} = \varepsilon_a - \varepsilon_i + \overline{g}_{iiaa}$, where the last term is the respective antisymmetrized integral; and construct the doubles as $D_{ijab} = D_{ia} + D_{jb}$. This is the default, and can be changed back to the simple Fock matrix guess by setting `UseCISUpdate` to false.

Following the preconditioning step, the resulting vectors are orthogonalized to the previous set of trial vectors, and orthonormalized among themselves. Since, the trial vectors do not change once they are generated (unless a restart occurs), only the new elements of the overlap matrix need to be generated for the orthonormalization. The numerical threshold for the inversion (and other division steps) is controlled by the parameter `OTol`. Finally, the amount of printed information can be controlled via the `PrintLevel` keyword. If not given or equal to 2, only basic iteration information will be printed. If set to 3, detailed iteration information will be printed (recommended if timing results for individual steps are required), while 4 or higher triggers additional (and very verbose) information from other subroutines as well.

The default solver is a multi-root Davidson procedure. The single-root solver can be initiated by setting `DoRootwise` and `FollowCIS` to true. The latter is more stable when a large number of roots are requested.

7.33.7 Properties in the RHF EOM implementation

The only property that can be calculated with the current RHF EOM implementation is the transition moment. It is calculated as a CI-like expectation value, as proposed by Stanton and Bartlett. The right and left transition density are defined as

$$\rho_{pq}^{Gr\to Ex} = \langle\phi_0|(1+\Lambda)[e^{-T}\{p^+q^-\}e^T, R]|\phi_0\rangle$$

$$\rho_{pq}^{Ex\to Gr} = \langle\phi_0|Le^{-T}\{p^+q^-\}e^T|\phi_0\rangle$$

In the above equation, Λ corresponds to the ground state left vector, which needs to solved once and L is the left vector , which needs to be solved separately for each root. Once the right and left vectors have been obtained, the left and right transition densities are constructed and the oscillator strength is calculated using following formula

$$f = \frac{2}{3}\varepsilon|\mu_{pq}\rho_{pq}^{Ex\to Gr}||\mu_{pq}\rho_{pq}^{Gr\to Ex}|$$

The oscillator strength, calculated by default, employs a linear approximation for Λ. The L vectors are, on the other hand, calculated as a general inverse of the corresponding R vectors. This approximation requires no additional effort over the energy calculation and gives similar accuracy as that of the exact oscillator strength calculation, which is at least twice the cost of the energy calculation. Exact EOM-CC transition moments can, however, be calculated by setting `DoLeft` and `DoTDM` to true. Please note that transition moments have not yet been implemented for the UHF EOM-CCSD approach.

7.33.8 Some tips and tricks for EOM-CC calculation

- The COSX approximation gives significant savings in terms of memory use, disk space use and computational timings without almost no loss of accuracy[233]. Therefore, the preferred setting for large scale calculations should include `DoCOSXEOM true`, `DoAOX3e true` and `KCOpt KC_AOBLAS` (Note that `KC_AOX` is the only option available for `KCOpt` in the UHF implementation).
- The EOM-CC code in ORCA has three version of the Davidson's solver. The default one is multi-root solver which does optimization of all the roots together. It gives the fastest convergence and is more suitable when one is interested only in a few roots of a big molecule. However, the multi-root solver can land into numerical issues, if more than 10 root are desired. In that case, one can invoke the root-wise solver by setting `DoRootwise true`. The single root solver is very stable and should be used when large number of roots are desired. However, the convergence of the single root solver is slower than the multi-root one. In the RHF implementation, there is also a batchwise solver, where a subset of the total number of roots is optimized together. This can be invoked by setting `NRootsPerBatch` to true and is intermediate between the multi-root and single-root solver in terms of stability and convergence.
- If the EOM iterations do not converge within 50 cycles, one can try to increase the number of iterations by setting `MaxIter` in the `%mdci` block to a larger value. One can also try to increase the dimension of the Davidson's space by increasing the `NDav` value and this generally helps in convergence acceleration. However, setting `NDav` to a value larger than 200 can make the calculation prohibitively costly .
- Convergence thresholds of `DTol 1e-5` (Davidson convergence) and `STol 1e-7` (ground state CCSD convergence) generally yield sufficiently converged energies, and are suitable for most purposes.
- The normal Davidson solver generally leads to the lowest energy solutions. This procedure can also yield roots dominated by double excitations (the so-called satellite states) for the IP and EA variants of EOM-CC, when one asks for a large number of roots. If one interested in the low lying Koopman's type of IP and EA states, they can be obtained by setting `FollowCIS` to true. This will follow the initial guess provided by the Fock operators.

7.34 Excited States via STEOM-CCSD

The EOM-CCSD approach for excitation energies becomes prohibitively costly for large systems because of its O(N^6) scaling. Therefore, one needs a more compact form of the wave-function ansatz. A second similarity transformation can compress the final matrix diagonalization step to the CIS space only. The resulting STEOM-CCSD method of Marcel Nooijen and co-workers [638] is an efficient way for accurate calculations of excitation energies.

7.34.1 General Description

In the standard EOM-CC method, the transformed Hamiltonian is diagonalized over a singles and doubles space to obtain ionized, attached, or excited states of the reference state. In STEOM-CC, one performs a second similarity transformation

$$\hat{G} = \{e^{\hat{S}}\}^{-1}\hat{\bar{H}}\{e^{\hat{S}}\}$$

The transformation operator $\hat{S}$, including singles and doubles, is defined as

$$\hat{S} = \hat{S}^{IP} + \hat{S}^{EA},$$

$$\hat{S}^{IP} = S^{m}_{i'}\hat{E}^{m}_{i'} + \frac{1}{2}S^{mb}_{ij}\hat{E}^{mb}_{ij},$$

$$\hat{S}^{EA} = S^{a'}_{e}\hat{E}^{a'}_{e} + \frac{1}{2}S^{ab}_{ej}\hat{E}^{ab}_{ej}.$$

In the above equations, m and e denote active indices of the hole and particle type respectively, while a prime denotes a restriction to orbitals that are not active. The amplitudes of the operator $\hat{S}$ are defined in such a way that matrix elements of the transformed Hamiltonian, in second quantized notation, become equal to zero.

$$g^{m}_{i'} = g^{mb}_{ij} = g^{a'}_{e} = g^{ab}_{ej} = 0$$

In addition, the zeros which pre-existed in $\bar{H}$, after solving the CCSD equations, remain preserved. The above equations are linear in $\hat{S}$ and are equivalent to the Fock space multireference coupled cluster equations for the one valence problem. However, to ensure numerical stability, the equations are re-casted as matrix diagonalization problem and solved as IP-EOM-CCSD and EA-EOM-CCSD problems. The $\hat{S}^{IP}$ and $\hat{S}^{EA}$ are then extracted from the converged previous calculations, respectively, by invoking intermediate normalization on the suitably chosen eigenvectors corresponding to active holes and active particles. The total process can be described as following

- Solution of the ground state coupled cluster equations
- Construct the first similarity transformed Hamiltonian as $\hat{\bar{H}} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$
- Solution of the IP-EOM and EA-EOM equations
- Extraction of the $\hat{S}$ amplitudes
- Construct the second similarity transformed Hamiltonian as $\hat{\bar{G}} = e^{-\hat{S}}\hat{\bar{H}}e^{\hat{S}}$
- Diagonalization of $\hat{\bar{G}}$ in CIS space

The advantage of the above method is that, instead of one iterative O(N^6) scaling diagonalization step, it requires two iterative O(N^5) scaling steps, one non-iterative O(N^5) scaling step and one iterative O(N^4) scaling matrix diagonalization step. The presence of so-called 'implicit triples excitation' term ensures the charge transfer separability of the excited states, which is absent in EOM-CCSD. In addition, since the final diagonalization step is performed in a CIS space, the spin adaption is trivial and excited states of triplet multiplicity can be obtained without going through the complications of a spin orbital based implementation.

The STEOMCC approach has also recently been extended for applications to open-shell systems within the UHF formalism [403]. In this case, the expressions for the operators $\hat{S}^{IP}$ and $\hat{S}^{EA}$ take the form,

$$\begin{aligned}
\hat{S}^{IP} &= \frac{1}{2}\sum_{i,e,a,b} s^{ab}_{ie}\{\hat{a}^\dagger\,\hat{b}^\dagger\,\hat{e}\,\hat{i}\} + \sum_{\bar{i},e,\bar{a},b} s^{\bar{a}b}_{\bar{i}e}\{\hat{\bar{a}}^\dagger\,\hat{b}^\dagger\,\hat{e}\,\hat{\bar{i}}\} \\
&+ \frac{1}{2}\sum_{\bar{i},\bar{e},\bar{a},\bar{b}} s^{\bar{a}\bar{b}}_{\bar{i}\bar{e}}\{\hat{\bar{a}}^\dagger\,\hat{\bar{b}}^\dagger\,\hat{\bar{e}}\,\hat{\bar{i}}\} + \sum_{i,\bar{e},a,\bar{b}} s^{a\bar{b}}_{i\bar{e}}\{\hat{a}^\dagger\,\hat{\bar{b}}^\dagger\,\hat{\bar{e}}\,\hat{i}\}, \\
\widehat{S}_{-} &= \frac{1}{2}\sum_{i,j,a,m} s^{am}_{ij}\{\hat{a}^\dagger\,\hat{m}^\dagger\,\hat{j}\,\hat{i}\} + \sum_{\bar{i},j,\bar{a},m} s^{\bar{a}m}_{\bar{i}j}\{\hat{\bar{a}}^\dagger\,\hat{m}^\dagger\,\hat{j}\,\hat{\bar{i}}\} \\
&+ \frac{1}{2}\sum_{\bar{i},\bar{j},\bar{a},\bar{m}} s^{\bar{a}\bar{m}}_{\bar{i}\bar{j}}\{\hat{\bar{a}}^\dagger\,\hat{\bar{m}}^\dagger\,\hat{\bar{j}}\,\hat{\bar{i}}\} + \sum_{i,\bar{j},a,\bar{m}} s^{a\bar{m}}_{i\bar{j}}\{\hat{a}^\dagger\,\hat{\bar{m}}^\dagger\,\hat{\bar{j}}\,\hat{i}\}.
\end{aligned}$$

where we use overbars to distinguish the β orbitals from the α orbitals. The amplitudes $\left\{s^{ab}_{ie}, s^{\bar{a}b}_{\bar{i}e}\right\}$ are determined by solving the UHF EA-EOM-CCSD equations for the attachment of an α electron, while the $\left\{s^{\bar{a}\bar{b}}_{\bar{i}\bar{e}}, s^{a\bar{b}}_{i\bar{e}}\right\}$ amplitudes are extracted from a UHF EA-EOM-CCSD calculation for the attachment of a β electron. Similarly, the sets of amplitudes $\left\{s^{am}_{ij}, s^{\bar{a}m}_{\bar{i}j}\right\}$ and $\left\{s^{\bar{a}\bar{b}}_{\bar{i}\bar{j}}, s^{a\bar{m}}_{i\bar{j}}\right\}$ are determined by solving the decoupled UHF IP-EOM-CCSD problems for the ionization of an α electron and the ionization of a β electron, respectively. Hence, an UHF STEOMCC calculation involves two separate IP calculations (O(N^5) scaling) and two separate EA calculations (O(N^5) scaling steps).

All the speed up options, including CCSD(2) (only available in RHF implementation) and COSX, which are available for EOM-CCSD are also available for STEOMCC. The most important steps in a STEOMCC calculation are the IP-EOM and EA-EOM calculations. These steps are performed using the EOM-CCSD module and the relevant keywords are the same as that described in *Excited States via EOM-CCSD*. The keywords which are exclusive to the RHF STEOM module are:

```
%mdci
#RHF STEOM parameters - defaults displayed
DoCISNat true        # automatic selection of active space
NActIP  3            # number of states defined as active in the IP calculation
NActEA  2            # number of states defined as active in the EA calculation
DoTriplet false      # target state of triplet multiplicity
DoDbFilter true      # filters out states with doubles excitation  character
DoNewActSch true    # new active space selection scheme for STEOM-CCSD
DoSOLV               # perturbative correction for solvation effects (experimental)
#Default values for automatic active space selection scheme
OThresh 0.001        # Cut off occupation of CIS natural orbitals in IP calculation
VThresh 0.001        # Cut off occupation of CIS natural orbitals in EA calculation
IPSThrs 80            #  The percentage singles threshold for the IP calculation
EASThrs 80           #  The percentage singles threshold for the EA calculation
end
```

The keywords pertaining to the UHF STEOM module are:

```
%mdci
#UHF STEOM parameters - defaults displayed
DoCISNat true        # automatic selection of active space
NActIP_a  3          # number of states defined as active in the IP calculation
                     # for the removal of an α electron
NActIP_b  3          # number of states defined as active in the IP calculation
                     # for the removal of a β electron
NActEA_a  2          # number of states defined as active in the EA calculation
                     # for the attachment of an α electron
NActEA_b 2           # number of states defined as active in the EA calculation
                     # for the attachment of a β electron
DoDbFilter true      # filters out states with doubles excitation  character
UseQROs  false       # use QROs or not
DoNewActSch true     # new active space selection scheme for STEOM-CCSD
#Default values for automatic active space selection scheme
OThresh 0.001        # Cut off occupation of CIS natural orbitals in tIP calculations
VThresh 0.001        # Cut off occupation of CIS natural orbitals in EA calculations
IPSThrs 80           #  The percentage singles threshold for the IP calculations
EASThrs 80           #  The percentage singles threshold for the EA calculations
end
```

7.34.2 Selection of Active space

The results of a STEOM-CC calculation depend upon the number of roots selected as active in the EOMIP and EOMEA calculations. In ORCA, they are chosen automatically, by using state-averaged CIS natural transition orbitals (NTO). By default, the number of roots included in this initial CIS computation is equal to the number of roots requested in STEOM (`NRoots`). However, this can be modified setting `NRootsCISNAT` to higher values. The orbitals up to a predefined occupation are then chosen to be active in the EOMIP and EOMEA calculations, and this is controlled by the keywords `OThresh` and `VThresh` respectively. Now, there are two possible ways to chose active space. One is to use the criteria of percentage occupation of NTO's as described in ref [236]. However, a newer and more robust approach is to use the criteria of absolute occupation, which is default in the current implementation. One can switch on the old percentage occupation based active space selection by setting `DoNewActSch` to false (not recommended).

One can also select the active spaces manually by turning the `DoCISNat` to false and setting the `NActIP` and `NActEA` (RHF STEOM calculation) or the `NActIP_a`, `NActIP_b`, `NActEA_a` and `NActEA_b` (UHF STEOM calculation) to desired values. However, this is not recommended for general uses. The following snippet shows the output of the active orbital selection procedure on a closed-shell molecule:

```
-----------------------------------------
STATE AVERAGED NATURAL ORBITALS FOR ACTIVE SPACE SELECTION
```

(continues on next page)

(continued from previous page)

```
----------------------------------------

Solving eigenvalue problem for the occupied space  ... Occupied block occupation :
     0   0.000478
     1   0.002266
     2   0.169928
     3   0.171663
     4   0.310125
     5   0.345541
Orbital taken as active for IP roots:
     0   0.345541
     1   0.310125
     2   0.171663
     3   0.169928
done
Solving eigenvalue problem for the virtual space   ... Virtual block occupation :
     6   0.640886
     7   0.332262
     8   0.017272
     9   0.005326
    10   0.001752
    11   0.000667
    12   0.000574
    13   0.000540
    14   0.000160
    15   0.000150
    16   0.000139
    17   0.000086
    18   0.000082
    19   0.000037
    20   0.000023
    21   0.000016
    22   0.000013
    23   0.000008
    24   0.000003
    25   0.000002
    26   0.000001
    27   0.000000
    28   0.000000
    29   0.000000
    30   0.000000
    31   0.000000
    32   0.000000
    33   0.000000
    34   0.000000
    35  -0.000000
Orbital taken as active for EA roots :
     0   0.640886
     1   0.332262
     2   0.017272
done
 No of roots active in IP calculation:    4
 No of roots active in EA calculation:    3
```

7.34.3 Active space selection using TD-DFT densities

Instead of using a CIS calculation for selected the Active Space roots, a TD-DFT based one can also be considered. Be aware that using DFT Kohn-Sham orbitals for computing the CCSD GS energy can lead to some instabilities and give incorrect results.

The main interest of this approach is to start the STEOM-CCSD calculation with TD-DFT electronic densities which are in general better than the CIS one, especially for some specific compounds (metallic complexes for example). The computed TD-DFT densities are also often more stable than the CIS one. It will however slow down the calculation.

The input has to be written like this:

```
!RHF/UHF BHANDHLYP STEOM-CCSD TZVP

%mdci
nroots 10
tddftguess true
end

%tddft
nroots 10
end

*xyz 0 1
```

Any DFT functional can be used but we recommend one with a decent amount of HF exchange. On top of this, the keyword `TDDFTGuess` has to be set to true in mdci block and the tddft has to be added together with the `NRoots` keyword. In both input blocks (`%mdci` and `%tddft`) the same number of roots has to be given. Starting from ORCA 6, using TD-DFT guess with an UHF reference is also possible.

7.34.4 The reliability of the calculated excitation energy

The excitation energy for any states calculated in STEOM-CC are only reliable when the dominant excitation for that states are confined within the active space. This can be verified from the percentage active character of the calculated states, an *a posteriori* diagnostic which is defined as

$$\%activecharacter = \frac{\sum\limits_{m,e} C(m,e) * C(m,e)}{\sum\limits_{i,a} C(i,a) * C(i,a)} * 100$$

for closed-shell systems and takes the form,

$$\%activecharacter = \frac{\sum\limits_{m,e} C(m,e) * C(m,e) + \sum\limits_{\bar{m},\bar{e}} C(\bar{m},\bar{e}) * C(\bar{m},\bar{e})}{\sum\limits_{i,a} C(i,a) * C(i,a) + \sum\limits_{\bar{i},\bar{a}} C(\bar{i},\bar{a}) * C(\bar{i},\bar{a})} * 100.$$

within the UHF formalism. The roots which have $\%activecharacter$ higher than 98.0 are considered to be converged with respect to the active space.

```
------------------
STEOM-CCSD RESULTS
------------------

IROOT=  1:  0.145412 au      3.957 eV   31914.3 cm**-1
  Amplitude    Excitation
  -0.169361     4 ->   8
  -0.984822     7 ->   8

Percentage Active Character     99.86
```

(continues on next page)

(continued from previous page)

```
  Amplitude    Excitation in Canonical Basis
  -0.166580     4 ->   8
  -0.975432     7 ->   8
  -0.124356     7 ->  13

IROOT=  2:  0.309409 au     8.419 eV   67907.5 cm**-1
  Amplitude    Excitation
   0.994141     7 ->   9

Percentage Active Character     99.78

  Amplitude    Excitation in Canonical Basis
  -0.990029     7 ->   9

IROOT=  3:  0.336993 au     9.170 eV   73961.4 cm**-1
  Amplitude    Excitation
  -0.994078     5 ->   8

Percentage Active Character     99.10

  Amplitude    Excitation in Canonical Basis
  -0.984116     5 ->   8
  -0.136769     5 ->  13

IROOT=  4:  0.357473 au     9.727 eV   78456.2 cm**-1
  Amplitude    Excitation
   0.181761     4 ->  10
   0.728209     6 ->   8
   0.611668     7 ->  10
  -0.191540     7 ->  12

Percentage Active Character     94.10

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
  -0.184144     4 ->  10
  -0.725183     6 ->   8
  -0.633718     7 ->  10

IROOT=  5:  0.386654 au    10.521 eV   84860.8 cm**-1
  Amplitude    Excitation
   0.980406     4 ->   8
  -0.178551     7 ->   8

Percentage Active Character     99.79

  Amplitude    Excitation in Canonical Basis
   0.971678     4 ->   8
   0.122877     4 ->  13
  -0.179242     7 ->   8

IROOT=  6:  0.444881 au    12.106 eV   97640.1 cm**-1
  Amplitude    Excitation
  -0.995150     6 ->   9

Percentage Active Character     99.69

  Amplitude    Excitation in Canonical Basis
```

(continues on next page)

(continued from previous page)

```
  -0.989966     6 ->   9
```

If the $\%activecharacter$ for any calculated state is less than 98, that state may have not converged with respect to active space and the excitation energy for that particular state is less reliable. The user should request a larger number of roots under those conditions.

7.34.5 Removal of IP and EA states with double excitation character

To obtain accurate results with STEOM-CCSD, only the $\hat{S}$ amplitudes corresponding to the states dominated by single excitations should be included in the second similarity transformation. This is ensured in ORCA in two ways. First, the root following (`FollowCIS`) is activated by default so that it converges to the states dominated by singly excited guess vectors. This avoids the calculation of so called 'satellite states', which are of double excitation character with respect to the ground state. Secondly, among the converged IP and EA roots, the states which have %singles character below a certain predefined threshold (i.e. controlled by the keywords `IPThresh` and `EAThresh`) are automatically excluded from the second similarity transformation.

```
EOM-CCSD RESULTS
----------------

IROOT=  1:  0.105316 au      2.866 eV   23114.2 cm**-1
  Amplitude    Excitation
   0.697547     x ->   8
IROOT=  2:  0.217925 au      5.930 eV   47829.1 cm**-1
  Amplitude    Excitation
  -0.701454     x ->   9
IROOT=  3:  0.304098 au      8.275 eV   66741.8 cm**-1
  Amplitude    Excitation
  -0.700458     x ->  10
IROOT=  4:  0.350387 au      9.535 eV   76901.1 cm**-1
  Amplitude    Excitation
   0.702705     x ->  11
IROOT=  5:  0.651462 au     17.727 eV  142979.4 cm**-1
  Amplitude    Excitation
   0.637352     x ->  12
   0.121747     x ->   8     4 ->  10
   0.177039     x ->   8     5 ->   9
   0.109987     x ->   9     5 ->   8
  -0.206789     x ->   8     7 ->  10
  -0.109870     x ->  10     7 ->   8

EA STATE=  0:  percentage singles     95.282
EA STATE=  1:  percentage singles     96.981
EA STATE=  2:  percentage singles     96.540
EA STATE=  3:  percentage singles     97.844
EA STATE=  4:  percentage singles     68.884

Warning: high double excitation character, excluding from the STEOM transformation
 Final no active EA roots:    4
```

Note that the use of CIS natural transition orbitals can lead to convergence issues for the IP and EA states which are dominated by double excitation character. This can be remedied by setting `DoDbFilter` to true.

7.34.6 Transition and difference densities

At the end of a STEOM computation, it is possible to store the final eigenvectors in a file "*job.cis*", in analogy with what is done for CIS and TD-DFT computations. This file can be obtained by setting `DoStoreSTEOM true` in the input. This file can then be processed by `orca_plot` to obtain the difference and / or the transition densities.

An Natural Transition Orbitals analysis can also be performed within the STEOM-CCSD scheme, as described in *Natural Transition Orbitals*. It can be performed by setting the keyword `DoSTEOMNatTransOrb` to true.

7.34.7 Properties

The dipolar and transition moments (as well as the oscillator strength) can be computed within the STEOM module using different kinds of approximations. Please cite our paper on these corrected STEOM transition densities [297]! Starting from ORCA 5, new defaults (`DoSimpleDens false`) are used that are much better than the previous CIS-like approximation, and the full option is of CC3-like quality.

```
%mdci

DosimpleDens false      # Default, using the STEOM-CCSD density + some doubles effect.
AddL2term true

DosimpleDens false      # using the STEOM-CCSD density + some doubles effect.
AddL2term true          # + neglected GS double
UpdateL1 true

DosimpleDens false      # using the STEOM-CCSD density + some doubles effect.
AddL2term true          # + neglected GS double + doubles from EOM-CCSD
UpdateL1 true           # (expensive, but of CC3 quality - see reference)
AddDDTerm true

end
```

By default, the STEOM-CCSD densities with `AddL2term true` should be used for all calculation as discussed in ref. [297].

7.34.8 Solvation (Experimental)

In STEOM-CCSD, the excitation energies and densities can be corrected using the CPCM solvation scheme in ORCA.

To use it, the keyword `DoSolv` has to be set to true in the `%mdci` block and the simple keyword `CPCM` (or `SMD`) + name of the solvent has to be given.

```
!CPCM(ethanol) STEOM-CCSD TightSCF def2-TZVP def2-TZVP/C def2/J

%mdci
Nroots 5
DoSolv true
end
```

7.34.9 Spin-Orbit Coupling (Experimental)

You can compute the spin-orbit coupling between singlet and triplets states in STEOM-CCSD using the keyword `STEOMSOC true`. Please note that all SOC matrix elements and properties are currently computed from the **right vector only**!

7.34.10 Core excitation

The STEOM-CCSD (and bt-PNO-STEOM-CCSD) method can also be used to compute the K-edge core-excitation energy of molecules. See *Core-Excitation* for more details.

7.34.11 Transient absorption

Transient absorption spectra can be computed using the keyword `DoTrans true`. The `IRoot` keyword will select the targeted excited state.

7.35 Excited States via IH-FSMR-CCSD

An alternative approach for decoupling the singles excitation space from the space of double and higher excitations is to use the so called Fock space multi-reference coupled cluster (FSMRCC) method. The method is similar to STEOM-CCSD, but much more flexible in terms of formulation.

7.35.1 General Description

FSMRCC is originally based on an effective Hamiltonian (EH). The basic idea of EH theory is to obtain some selective eigenvalues of the Hamiltonian operator from the total eigenvalue spectrum. For this purpose, the entire configuration space is divided into a model and an outer space with projection operators P_M and Q_M, respectively (see Fig. 7.38). The diagonalization of the EH takes care of the non-dynamic correlation coming from the interactions between the model space configurations. On the other hand, the dynamic correlation arises due to the interactions of the model space configurations with the outer space configurations. This interaction is introduced through a universal wave operator Ω, which is parametrized such that it generates the exact wave function when acting on the model space. The valence universal wave operator Ω has the form

$$\Omega = e^{\tilde{S}^{(p,h)}}$$

where the braces indicate normal ordering of the cluster operators and $\tilde{S}^{(p,h)}$ is defined as

$$\tilde{S}^{(p,h)} = \sum_{k=0}^{p} \sum_{l=0}^{h} \tilde{S}^{(k,l)}$$

The cluster operator $\tilde{S}^{(k,l)}$ is capable of destroying exactly k active particles and l active holes, in addition to creation of holes and particles. The $\tilde{S}^{(p,h)}$ subsumes all lower sector Fock space $\tilde{S}^{(k,l)}$ operators. The $\tilde{S}^{(0,0)}$ is equivalent to standard single-reference coupled cluster $\hat{T}$ operator. The EH for (p,h) valence system can be defined as

$$\hat{H}_{eff} = P_M^{(p,h)} \Omega^{-1} \hat{H} \Omega P_M^{(p,h)}$$

However, Ω^{-1} may not be well defined in all the cases. Therefore, the above definition for the EH is seldom used. Instead, the Block-Lindgren approach is generally used for solving the equations, which is defined by

$$P_M^{(p,h)} \left[\hat{H}\Omega - \Omega \hat{H}_{eff} \right] P_M^{(p,h)} = 0$$

$$Q_M^{(p,h)} \left[\hat{H}\Omega - \Omega \hat{H}_{eff} \right] P_M^{(p,h)} = 0$$

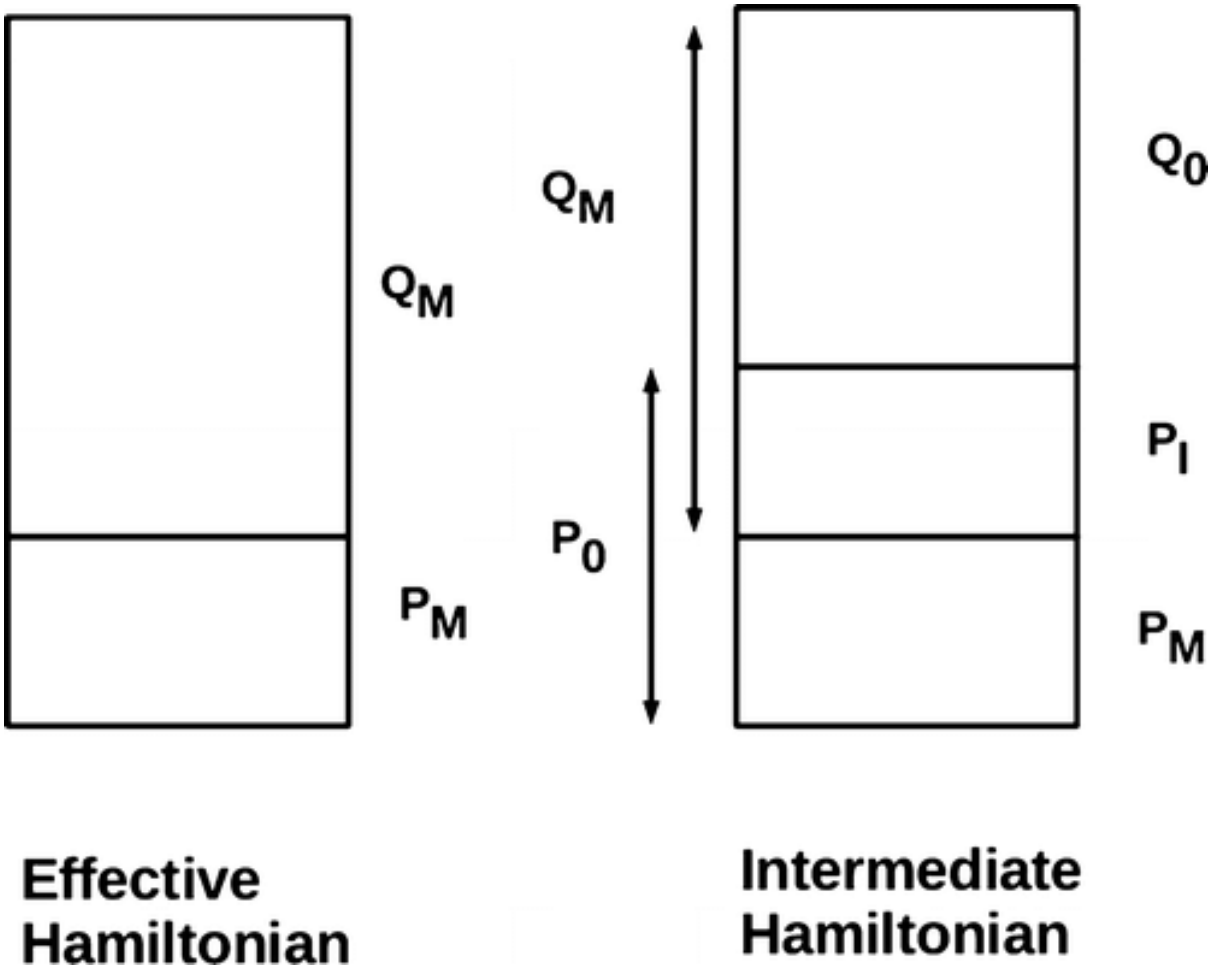


Fig. 7.38: Division of the configuration space into model and outer space in effective Hamiltonian (EH) theory and into model, intermediate, and outer space in intermediate Hamiltonian (IH) theory. P and Q denote the respective projection operators.

When the model space is not energetically well separated from the outer space, this method faces convergence problems. This is commonly termed as the intruder state problem. In the intermediate Hamiltonian (IH) formulation, configuration space is divided into three subspaces, namely, the main(M), the intermediate(I), and the outer(O) space (see Fig. 7.38) with projection operators P_M, P_I and Q_O, respectively. The intermediate space acts as a buffer between the model and the outer space. When diagonalization the IH, a subset of the eigenvalues correspond to the main space obtained through EH theory. The IH is for the singly excited state sector (1,1) is defined as

$$H_I^{(1,1)} = P_O^{(1,1)} \bar{H} P_O^{(1,1)} + P_O^{(1,1)} \bar{H} Y^{(1,1)} P_M^{(1,1)}$$

where

$$Y^{(1,1)} = Q_O^{(1,1)} \left\{ S_2^{(0,1)} + S_2^{(1,0)} + S_2^{(0,1)} S_1^{(1,0)} + S_2^{(1,0)} S_1^{(0,1)} + S_2^{(1,0)} S_2^{(0,1)} \right\} P_M^{(1,1)}$$

The $S^{(1,0)}$ and $S^{(0,1)}$ are extracted from converged EOMIP-CCSD and EOMEA-CCSD calculations, respectively, by invoking intermediate normalization on the suitably chosen eigenvectors corresponding to active holes and active particles. The total procedure can be described as following

- solve the ground state coupled cluster equations
- construct $\hat{\bar{H}} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$
- solve the EOMIP and EOMEA equations
- extract the $\hat{S}$ amplitudes

- construct the second similarity transformed Hamiltonian as $H_I^{(1,1)}$
- diagonalize the $H_I^{(1,1)}$ in CIS space

The automatic active space selection scheme and all the speed up options which are available for STEOM-CCSD, including bt-PNO and COSX, are also available for IH-FSMR-CCSD. All the keywords controlling the IH-FSMR-CCSD are similar to STEOM-CCSD as described in *Excited States via STEOM-CCSD*.

No UHF variant of IH-FSMR-CCSD is currently available.

7.35.2 Properties

The transition properties can be calculated using a simple CIS-like formulation, employing the converged IH-FSMR-CC eigenvectors. The transition moments are computed by default in an IH-FSMR-CCSD calculation.

7.35.3 Solvation Correction

Solvent effects can be approximated by a simple perturbative correction to the IH-FSMR-CCSD via

$$\omega_k = \omega_k^0 + \frac{1}{2}\bar{V}^{\Delta}\bar{Q}^{\Delta}$$

where

$$\omega_k^0 = \hat{L_K} H_I^{(1,1)} \hat{R_K}$$

The CPCM correction directly enters the $H_I^{(1,1)}$, the modified Hatree-Fock orbitals. In the non-equilibrium regime, one can simply write the perturbative correction as

$$\omega_k^{neq} = \omega_k^{0,neq} + \frac{1}{2}\bar{V}\left(P_{\Delta}^{neq}\right)\bar{Q}\left(P_{\Delta}^{neq}\right)$$

where

$$P_{\Delta}^{neq} = L_k R_k$$

A typical input file looks like

```
! aug-cc-pVDZ IH-FSMR-CCSD
!CPCM(water)
%mdci
NROOTS 8
DoSOLV true
DTol 1e-10
end
*xyz 0 1
O       0.0000  0.0000  0.1173
H       0.0000  0.7572 -0.4692
H       0.0000 -0.7572 -0.4692
*
```

For the above input, the following output is obtained:

```
--------------------------------
     CALCULATED SOLVENT SHIFTS
            CPCM  MODEL
--------------------------------

Contributions of the 'fast' term to the solvent shift

 State Shift(Eh) Shift(eV) Shift(cm**-1) Shift(nm) E_FSMRCC(eV)  E_FSMRCC+SHIFT(eV)
```

(continues on next page)

(continued from previous page)

```
-------------------------------------------------------------------------------
   0: -0.0058000 -0.158    -1273.0        3.3      7.814          7.656
   1: -0.0133523 -0.363    -2930.5        5.0      9.650          9.287
   2: -0.0053622 -0.146    -1176.9        1.8     10.144          9.998
   3: -0.0078092 -0.212    -1713.9        2.2     10.958         10.746
   4: -0.0040294 -0.110     -884.4        1.0     11.534         11.424
   5: -0.0137147 -0.373    -3010.0        3.3     12.003         11.630
   6: -0.0093172 -0.254    -2044.9        2.2     12.131         11.878
   7: -0.0077514 -0.211    -1701.2        1.7     12.562         12.352
```

Thee perturbative correction only changes the transition energies and neither the wave function nor the transition moment.

7.36 Excited States using PNO-based coupled cluster

Despite the successes of the DLPNO-CC approximation for ground states, the use of PNOs for excited states has been less fruitful. It is not straightforward to define a PNO-based scheme for excited states, which will maintain the balance between speed and accuracy, as observed for the ground state. As an intermediate solution, the basis for ground state DLPNO quantities is transformed back to the canonical basis and are used within the canonical EOM routine. This procedure is justified, as the main bottle neck of the EOM-CCSD or STEOM-CCSD methods comes from the ground state calculation. Approximating the ground state CCSD amplitudes with MP2 amplitudes is also possible, as done in the EOM-CCSD(2) approach. However, it is not reliable and can lead to large errors, when the reference HF wave function does not provide a reasonable zeroth order approximation to the ground state wave function. Note that the back-transformed PNO scheme (bt-PNO) described here is available for both open-shell (UHF (QROs) or ROHF reference) and closed-shell (RHF reference) systems.

7.36.1 General Description

The back transformation of the ground state DLPNO-CCSD amplitudes to the virtual space involves three steps. The T amplitudes in the PNO basis are first converted into the PAO basis, then subsequently to the atomic orbital (AO) basis, and finally to the canonical MO basis[234]. For example, in the closed-shell case, we have

$$d^{ij}_{\tilde{\mu}\tilde{a}_{ij}} T^{ij}_{\tilde{a}_{ij}\tilde{b}_{ij}} d^{ij}_{\tilde{b}_{ij}\tilde{\mu}} \Rightarrow L^{ij}_{\mu\tilde{\mu}} T^{ij}_{\tilde{\mu}\tilde{\nu}} L^{ij}_{\tilde{\nu}\nu} \Rightarrow C^{ij}_{a\mu} T^{ij}_{\mu\nu} C^{ij}_{\nu b} \Rightarrow T^{ij}_{ab},$$

$$d^{i}_{\tilde{\mu}\tilde{a}_{ii}} T^{i}_{\tilde{a}_{ii}} \Rightarrow L^{i}_{\mu\tilde{\mu}} T^{i}_{\tilde{\mu}} \Rightarrow C^{i}_{a\mu} T^{i}_{\mu} \Rightarrow T^{i}_{a},$$

The AO basis functions are denoted as $\mu, \nu, \ldots$, while $\tilde{\mu}, \tilde{\nu}, \ldots$ refers to PAOs. The missing pairs are treated using MP2 amplitudes. If all the thresholds are set to zero, the back-transformed amplitudes match exactly with the canonical RI-EOM-CCSD ones. On the other hand, when all the thresholds are made infinitely tight, one obtains the EOM-CCSD(2) results. This PNO-based excited state approach is available for all the flavors of EOM-CCSD and for STEOM-CCSD in both open- and closed-shell systems.

Below, we list all the parameters that influence the DLPNO-CCSD-based excited state calculations

```
%mdci
#bt-PNO-EOM and STEOM parameters - defaults displayed
DoEOMMP2 true        # MP2 correction for missing pairs
DoRECAN  true        # recanonicalization of the occupied
                     # orbitals before the excited state calculation
                     #(only relevant for the RHF implementation)
end
```

7.36.2 Reference State Energy

Here, it should be noted that the reference energy for PNO-based EOM-CCSD or STEOM-CCSD is slightly different from that printed for a converged ground state DLPNO-CCSD calculation, as it includes the perturbative correction for different truncated quantities.

```
----------------------
COUPLED CLUSTER ENERGY
----------------------

E(0)                                       ...   -113.913498239
E(CORR)(strong-pairs)                      ...     -0.401457078
E(CORR)(weak-pairs)                        ...     -0.000339627
E(CORR)(corrected)                         ...     -0.401796705
E(TOT)                                     ...   -114.315294944
```

In the bt-PNO-EOM-CCSD scheme, the CI-like excited state treatment of the reference state is defined by back-transformed DLPNO amplitudes (or MP2 amplitudes for the weak pairs). The energy corresponding to this set of amplitude is printed at the beginning of the EOM calculations.

```
Dressing integrals for EOM-CCSD            ...
Reference state energy for EOM-DLPNO-CCSD     ...   -114.314954945

done (    0.4)
```

Therefore, to calculate the total energy of an excited (ionized or electron attached) state, one needs to add the excitation energy to the reference state energy in bt-PNO-EOM-CCSD.

7.36.3 Use of Local Orbitals

The use of local orbitals makes it difficult to follow a particular guess vector in the Davidson digonalization process in EOM-CC and STEOM-CC. Therefore, it is advisable to recanonicalize the occupied orbitals after the ground state DLPNO-CCSD calculation by setting DoRECAN to true (i.e. only relevant for the closed-shell RHF implementation). It should be noted that the recanonicalization does not change the EOM-CCSD energies. However, the STEOM-CC energies are not invariant to orbital rotations and differ slightly for local and canonical orbitals. In the open-shell bt-PNO implementation, we follow a different procedure in that all quantities are transformed to the delocalized basis before proceeding with the back transformation and the excited state calculation.

7.36.4 Some tips and tricks for bt-PNO calculations

- The bt-PNO scheme with tightPNO settings gives results, which are within 0.01 eV of the canonical EOM-CCSD numbers, at a fraction of the computational cost[234]. So, use of bt-PNO scheme is always preferable over canonical calculations.
- In the case of an RHF reference, one should set 'DLPNOLINEAR true' and 'NEWDOMAINS true' in the mdci block input to use the 2015 fully linear scaling implementation, which is more robust than the 2013 implementation used as default in bt-PNO scheme.
- The transition moment in bt-PNO-EOM (RHF only) and bt-PNO-STEOM (RHF, UHF (QROs) or ROHF) is only available using the linear approximation.

7.37 Excited States via DLPNO-STEOM-CCSD

7.37.1 PNO dressing (experimental keyword)

In STEOM-DLPNO-CCSD method most of the steps are done using the powerful PNO approximation except the last one corresponding to the STEOM-CCSD calculation itself. In the canonical version of the dressing (default in ORCA) all the amplitudes previously computed at the PNO level are recanonicalized which increases the calculation cost.PNO dressing uses some of the PNO intermediates to reduce the dressing time. It has however no effect on the storage size.

Using this PNO dressing will reduce the calculation timing but at a reduce accuracy for the excitation energies. We thus recommend this option only for specific large systems (>2500 basis functions).

The keyword **DirectDressSTEOM** set to true enable this option.

7.37.2 Keywords from STEOM-CCSD

Most of the keywords from STEOM-CCSD can be used within the DLPNO version, except for core excitation. More information on the *Excited States via STEOM-CCSD* section.

7.37.3 Tips and Tricks

As written in the typical calculation section of the manual we would recommend this input for standard organic molecules.

As a general guideline we will discuss some of the keywords used there:

```
! STEOM-DLPNO-CCSD def2-TZVP def2-TZVP/C def2/J TightSCF

%mdci
  NRoots 6
  DoRootWise true
  OThresh 0.005
  VThresh 0.005
  TCutPNOSingles 1e-11
  NDAV 400
  DoStoreSTEOM true
  DoSimpleDens false
  AddL2Term True
  DTol 1e-5
end

* xyz 0 1
  C     0.016227   -0.000000    0.000000
  O     1.236847    0.000000   -0.000000
  H    -0.576537    0.951580   -0.000000
  H    -0.576537   -0.951580   -0.000000
*
```

- TIGHTSCF is a must for any CCSD calculation.
- We will recommend using TIGHTPNO for all molecules as it is not a lot more expensive and helps achieving a better convergence.
- The OTHRESH, VTHRESH and TCutPNOSingles keywords help with converging the calculations, increasing the percentage active of each root. In contrary to standard STEOM-CCSD, we would acknowledge that the roots are converged when the percentage active character is at least 96%. Of course you have to check that the amplitude and orbitals associated with the excitation are correct (and what you expect). Tightening the 3 keywords mentioned will increase this percentage active character. The most important being the TCutPNOSingles one. Be careful, the computational cost increases exponentially when tightened. In more

case 10^{-12} or 5×10^{-12} should be enough. For OTHRESH and VTHRESH you should not go below 10^{-3}, as the benefits are not so obvious.A trick to achieve a better convergence is to play with the number of roots. It is often not necessary to compute a lot of roots if you are only interested in the first 3 for example. If some high energy roots have some low percentage active character, removing them can help for the convergence of other roots.

- As shown in the related paper[298], the STEOM density with the L2term is now the default for computing the excited state properties.
- The choice of the basis set can also speed up the calculation without a significant loss of accuracy. For most organic molecules def2-TZVP(-f) is enough. Trying def2-SVP is also a good idea for preliminary tests.
- The STEOM-CCSD excitation energies are very dependent from the starting geometry. Geometries optimised with several DFT functional can yield significant differences for the excitation energies (about 0.1 eV).
- It has happened that some roots were missing with the keyword **Dorootwise** set to true. Turning it off solved the issue but this keyword should be on by default.

7.38 Core-level spectroscopy with coupled cluster methods

The equation of motion coupled cluster method and its similarity transformed version provides an easy way to directly calculate core-ionization and core-excitation energies. Currently, the core-level spectroscopy with EOM-CCSD and STEOM-CCSD is only available for closed shell systems.

7.38.1 Core-ionization

One can obtain core-ionized states if one calculates a large no of roots. The ORCA implementation of IP-EOM-CCSD, however, allows one to directly target the ionization from the core-orbitals. A typical IP-EOM-CCSD input file for the acetic acid will look like

```
!IP-EOM-CCSD ExtremeSCF cc-pvtz
!NoFrozencore
%maxcore 5000

%mdci
nroots 4      #no of roots
CVSORB 0,3    #orbital considered for core-valence separation
FollowCIS true  # Follow  the initial guess orbital
CVSEP true      # Core valence separation
DoCVS true      # Core valence separation (currently both the option needs to be true)
DoCore true     # Directly target the core
corehole 0      # The state from which it will count the roots
printlevel 3    # the printing options
maxiter 500     # no of iteration, generally requires larger no of roots
end

*xyz 0 1
 C             -6.7624010562   0.1328615492   0.0389382700
 C             -5.3564667033   0.2819965475  -0.5188248498
 H             -6.9983743824   1.0019615710   0.6510029634
 H             -7.4924880320   0.0542210905  -0.7741766747
 H             -6.8380664832  -0.7720291637   0.6519904379
 O             -4.9303467983  -0.7518088469  -1.3223158759
 H             -5.6257914271  -1.4265892921  -1.4015111180
 O             -4.6208051175   1.2132365445  -0.3081931529
*
```

The output of it will be

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT=  1: 19.902202 au   541.566 eV 4368028.3 cm**-1
  Amplitude    Excitation
  -0.673297     0 -> x
Percentage singles character=     82.93

IROOT=  2: 19.842487 au   539.942 eV 4354922.4 cm**-1
  Amplitude    Excitation
   0.672818     1 -> x
Percentage singles character=     82.71

IROOT=  3: 10.891843 au   296.382 eV 2390483.3 cm**-1
  Amplitude    Excitation
  -0.669218     2 -> x
Percentage singles character=     81.11

IROOT=  4: 10.754926 au   292.656 eV 2360433.5 cm**-1
  Amplitude    Excitation
  -0.670254     3 -> x
Percentage singles character=     81.57
```

The option 'DoCore true' starts the counting of the roots from 'corehole' upwards. The default is 'DoCore false' and it counts the root from the HOMO downwards. The 'corehole 0' starts the counting from the first occupied orbital which is the oxygen K-edge in this case. One can directly target the carbon K-edge in this case by putting 'corehole 2'.

```
!IP-EOM-CCSD ExtremeSCF cc-pvtz
!NoFrozencore
%maxcore 5000

%mdci
nroots 2      #no of roots
CVSORB 2,3    #orbital considered for core-valence separation
FollowCIS true  # Follow  the initial guess orbital
CVSEP true      # Core valence separation
DoCVS true      # Core valence separation (currently both the option needs to be true)
DoCore true     # Directly target the core
corehole 2      # The state from which it will count the roots
printlevel 3    # the printing options
maxiter 500     # no of iteration, generally requires larger no of roots
end

*xyz 0 1
 C             -6.7624010562   0.1328615492   0.0389382700
 C             -5.3564667033   0.2819965475  -0.5188248498
 H             -6.9983743824   1.0019615710   0.6510029634
 H             -7.4924880320   0.0542210905  -0.7741766747
 H             -6.8380664832  -0.7720291637   0.6519904379
 O             -4.9303467983  -0.7518088469  -1.3223158759
 H             -5.6257914271  -1.4265892921  -1.4015111180
 O             -4.6208051175   1.2132365445  -0.3081931529
*
```

The output of it will be

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------
```

(continues on next page)

(continued from previous page)

```
IROOT=  1: 10.891843 au   296.382 eV 2390483.3 cm**-1
  Amplitude    Excitation
   0.669218     2 -> x
Percentage singles character=     81.11

IROOT=  2: 10.754926 au   292.656 eV 2360433.5 cm**-1
  Amplitude    Excitation
  -0.670254     3 -> x
Percentage singles character=     81.57
```

Now, the core-ionized states remains embedded in the high density of doubly ionized valence states that form the continuum. This leads to severe convergence problems. One easy way to overcome this is to use the core-valence separation approximation which is turned on by the two keywords 'CVSEP true' and 'DoCVS true'. The orbitals from which the contributions are not neglected for the core-valence separation are set by 'CVSORB initial,final'. It is generally a good idea to include all the core orbitals corresponding to a particular element if one is interested in the ionization from any of the core orbitals for the particular element. In the second example both the carbon core-orbitals are included in the 'CVSORB 2,3'. A 'bt-PNO-IP-EOM-CCSD' input file for the same example will look like

```
!bt-PNO-IP-EOM-CCSD ExtremeSCF cc-pvtz cc-pvtz/c
!NoFrozencore
%maxcore 5000

%mdci
nroots 4
CVSORB 0,3
FollowCIS true
CVSEP true
DoCVS true
DoCore true
DoRECAN true # recanonilize the occupied space before the EOM step
corehole 0
printlevel 3
maxiter 500
end

*xyz 0 1
 C             -6.7624010562   0.1328615492   0.0389382700
 C             -5.3564667033   0.2819965475  -0.5188248498
 H             -6.9983743824   1.0019615710   0.6510029634
 H             -7.4924880320   0.0542210905  -0.7741766747
 H             -6.8380664832  -0.7720291637   0.6519904379
 O             -4.9303467983  -0.7518088469  -1.3223158759
 H             -5.6257914271  -1.4265892921  -1.4015111180
 O             -4.6208051175   1.2132365445  -0.3081931529
*
```

The output of it will be

```
IROOT=  1: 19.901845 au   541.557 eV 4367950.0 cm**-1
  Amplitude    Excitation
  -0.673298     0 -> x
Percentage singles character=     82.93

IROOT=  2: 19.842152 au   539.932 eV 4354848.9 cm**-1
  Amplitude    Excitation
   0.672832     1 -> x
Percentage singles character=     82.72

IROOT=  3: 10.892369 au   296.396 eV 2390598.7 cm**-1
  Amplitude    Excitation
```

(continues on next page)

(continued from previous page)

```
  -0.669213     2 -> x
Percentage singles character=     81.11

IROOT=  4: 10.754951 au   292.657 eV 2360438.9 cm**-1
  Amplitude    Excitation
  -0.670244     3 -> x
Percentage singles character=     81.56
```

The results are in excellent agreement with the canonical one. A DLPNO variant for the same example will look like

```
!IP-EOM-DLPNO-CCSD ExtremeSCF cc-pvtz cc-pvtz/c autoaux def2/J TightPNO pal16
!NoFrozencore
%maxcore 5000

%mdci
nroots 4
CVSORB 0,3
FollowCIS true
CVSEP true
DoCVS true
DoCore true
corehole 0
printlevel 3
maxiter 500
end

*xyz 0 1
 C             -6.7624010562   0.1328615492   0.0389382700
 C             -5.3564667033   0.2819965475  -0.5188248498
 H             -6.9983743824   1.0019615710   0.6510029634
 H             -7.4924880320   0.0542210905  -0.7741766747
 H             -6.8380664832  -0.7720291637   0.6519904379
 O             -4.9303467983  -0.7518088469  -1.3223158759
 H             -5.6257914271  -1.4265892921  -1.4015111180
 O             -4.6208051175   1.2132365445  -0.3081931529
*
```

The output of it will be

```
----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT=  1: 19.945319 au   542.740 eV 4377491.5 cm**-1
  Amplitude    Excitation
   0.678788     0 -> x
Percentage singles character=    101.06

IROOT=  2: 19.890529 au   541.249 eV 4365466.5 cm**-1
  Amplitude    Excitation
   0.679716     1 -> x
Percentage singles character=    101.02

IROOT=  3: 10.912292 au   296.939 eV 2394971.3 cm**-1
  Amplitude    Excitation
   0.672646     2 -> x
Percentage singles character=    101.22

IROOT=  4: 10.792478 au   293.678 eV 2368675.1 cm**-1
  Amplitude    Excitation
```

(continues on next page)

(continued from previous page)

```
   0.674795     3 -> x
Percentage singles character=    101.14
```

Although the error in the absolute IP values are as large as 1 eV, the so-called 'chemical shift' i.e. the difference between the IP value of two different atoms of the same elements are reasonably correct.

7.38.2 Core-Excitation

The STEOM-CCSD approach provides an efficient and accurate way to do the K-edge core-excitation spectroscopy. A typical input file for the STEOM-CCSD will look like

```
!STEOM-CCSD ExtremeSCF aug-cc-pCVQZ Bohrs NoFrozencore

%mdci
nroots 10
CVSORB 6,6  # should always be HOMO
FollowCIS true
CVSEP true
DoCVS true
DoCore true
DoSimpleDens False # use exact STEOM transition moment
corehole 0
maxiter 500
NDAV 80
printlevel 3
end

*xyz 0 1
O 0 0 0.913973
C 0 0 -1.218243
*
```

The output will be

```
------------------
STEOM-CCSD RESULTS
------------------

IROOT=  1: 19.686174 au   535.688 eV 4320615.8 cm**-1
  Amplitude    Excitation
  -0.215317     6 ->   7
  -0.211332     6 ->   8
  -0.442546     6 ->  11
   0.330607     6 ->  12
   0.568497     6 ->  15
  -0.506486     6 ->  16
  Ground state amplitude:  0.000000

Percentage Active Character     97.59

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
  -0.177019     0 ->   8
   0.346366     0 ->   9
   0.590044     0 ->  11
   0.464534     0 ->  12
  -0.214399     0 ->  14
   0.338336     0 ->  15
```

(continues on next page)

(continued from previous page)

```
  -0.146649     0 ->  20
   0.169124     0 ->  21
  -0.192930     0 ->  24

IROOT=  2: 19.686174 au   535.688 eV 4320615.8 cm**-1
  Amplitude    Excitation
   0.211332     6 ->   7
  -0.215317     6 ->   8
   0.330607     6 ->  11
   0.442546     6 ->  12
  -0.506486     6 ->  15
  -0.568497     6 ->  16
  Ground state amplitude:  0.000000

Percentage Active Character     97.59

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
   0.346366     0 ->   8
   0.177019     0 ->   9
   0.464534     0 ->  11
  -0.590044     0 ->  12
   0.338336     0 ->  14
   0.214399     0 ->  15
   0.169124     0 ->  20
   0.146649     0 ->  21
  -0.192930     0 ->  23

IROOT=  3: 19.865373 au   540.564 eV 4359945.5 cm**-1
  Amplitude    Excitation
  -0.571289     6 ->   9
   0.792679     6 ->  10
   0.137627     6 ->  13
  -0.112257     6 ->  17
  Ground state amplitude: -0.000591

Percentage Active Character     97.37

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
  -0.900242     0 ->   7
  -0.116863     0 ->  13
  -0.375672     0 ->  18
   0.128317     0 ->  19

IROOT=  4: 19.909335 au   541.761 eV 4369594.0 cm**-1
  Amplitude    Excitation
   0.340300     6 ->   7
   0.704671     6 ->   8
  -0.338179     6 ->  11
   0.511324     6 ->  12
  Ground state amplitude:  0.000000

Percentage Active Character     99.71

  Amplitude    Excitation in Canonical Basis
   0.101796     0 ->   8
```

(continues on next page)

(continued from previous page)

```
  -0.793309     0 ->   9
   0.482364     0 ->  11
   0.160972     0 ->  12
  -0.209491     0 ->  15
  -0.128207     0 ->  23
  -0.188543     0 ->  24

IROOT=  5: 19.909335 au   541.761 eV 4369594.0 cm**-1
  Amplitude    Excitation
  -0.704671     6 ->   7
   0.340300     6 ->   8
   0.511324     6 ->  11
   0.338179     6 ->  12
  Ground state amplitude:  0.000000

Percentage Active Character     99.71

  Amplitude    Excitation in Canonical Basis
  -0.793309     0 ->   8
  -0.101796     0 ->   9
   0.160972     0 ->  11
  -0.482364     0 ->  12
  -0.209491     0 ->  14
  -0.188543     0 ->  23
   0.128207     0 ->  24

IROOT=  6: 19.914772 au   541.909 eV 4370787.3 cm**-1
  Amplitude    Excitation
  -0.804799     6 ->   9
  -0.557108     6 ->  10
  -0.125228     6 ->  13
   0.119745     6 ->  17
  Ground state amplitude:  0.000364

Percentage Active Character     97.38

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
   0.934227     0 ->  10
  -0.273222     0 ->  13
   0.144269     0 ->  18
  -0.159846     0 ->  22

IROOT=  7: 19.966983 au   543.329 eV 4382246.2 cm**-1
  Amplitude    Excitation
   0.190413     6 ->  10
  -0.954987     6 ->  13
   0.113662     6 ->  22
  Ground state amplitude:  0.000138

Percentage Active Character     94.88

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
  -0.190758     0 ->   7
   0.246282     0 ->  10
   0.890287     0 ->  13
```

(continues on next page)

(continued from previous page)

```
   0.198835     0 ->  18
   0.170782     0 ->  19
   0.180954     0 ->  25

IROOT=  8: 19.981194 au   543.716 eV 4385365.2 cm**-1
  Amplitude    Excitation
   0.513702     6 ->   7
   0.587528     6 ->  11
   0.608754     6 ->  15
  Ground state amplitude: -0.000000

Percentage Active Character     98.92

  Amplitude    Excitation in Canonical Basis
   0.391272     0 ->   8
  -0.123811     0 ->   9
  -0.308174     0 ->  12
  -0.796992     0 ->  14
   0.173278     0 ->  15
  -0.202678     0 ->  20
   0.130010     0 ->  29

IROOT=  9: 19.981194 au   543.716 eV 4385365.2 cm**-1
  Amplitude    Excitation
  -0.513700     6 ->   8
   0.587528     6 ->  12
   0.608756     6 ->  16
  Ground state amplitude: -0.000000

Percentage Active Character     98.92

  Amplitude    Excitation in Canonical Basis
   0.123796     0 ->   8
   0.391276     0 ->   9
   0.308177     0 ->  11
  -0.173290     0 ->  14
  -0.796989     0 ->  15
  -0.202680     0 ->  21
   0.130002     0 ->  28

IROOT= 10: 20.005026 au   544.364 eV 4390595.8 cm**-1
  Amplitude    Excitation
   0.997209     6 ->  14
  Ground state amplitude:  0.000000

Percentage Active Character     99.44

  Amplitude    Excitation in Canonical Basis
   0.980462     0 ->  16
  -0.142106     0 ->  26
```

Excitation from a particular core orbital can be considered currently. In the present case it is the 1S orbital of oxygen. The required orbital can be specified using the keyword 'corehole'. For the oxygen 1S it should be 'corehole 0'. The carbon 1S can be specified with 'corehole 1'

```
!STEOM-CCSD ExtremeSCF aug-cc-pCVQZ Bohrs NoFrozencore

%mdci
nroots 10
CVSORB 6,6  # should always be HOMO
FollowCIS true
```

(continues on next page)

(continued from previous page)

```
CVSEP true
DoCVS true
DoCore true
DoSimpleDens False # use exact STEOM transition moment
corehole 1
maxiter 500
NDAV 80
printlevel 3
end

*xyz 0 1
O 0 0 0.913973
C 0 0 -1.218243
*
```

It will give the carbon K-edge spectra as follows

```
------------------
STEOM-CCSD RESULTS
------------------

IROOT=  1: 10.569902 au   287.622 eV 2319825.3 cm**-1
  Amplitude    Excitation
   0.429677     6 ->   8
   0.132429     6 ->  11
   0.880021     6 ->  16
  Ground state amplitude: -0.000000

Percentage Active Character     98.73

  Amplitude    Excitation in Canonical Basis
   0.158027     1 ->   8
   0.273564     1 ->   9
  -0.354592     1 ->  11
  -0.668565     1 ->  12
   0.184004     1 ->  14
   0.308132     1 ->  15
   0.112105     1 ->  20
   0.195715     1 ->  21
   0.134711     1 ->  23
   0.275743     1 ->  24
  -0.140527     1 ->  29

IROOT=  2: 10.569902 au   287.622 eV 2319825.3 cm**-1
  Amplitude    Excitation
   0.429686     6 ->   7
  -0.132345     6 ->  12
   0.880029     6 ->  15
  Ground state amplitude: -0.000000

Percentage Active Character     98.73

  Amplitude    Excitation in Canonical Basis
  -0.273564     1 ->   8
   0.158027     1 ->   9
   0.668565     1 ->  11
  -0.354592     1 ->  12
  -0.308132     1 ->  14
   0.184004     1 ->  15
  -0.195715     1 ->  20
   0.112105     1 ->  21
```

(continues on next page)

(continued from previous page)

```
  -0.275744     1 ->  23
   0.134711     1 ->  24
   0.140527     1 ->  28

IROOT=  3: 10.807563 au   294.089 eV 2371985.9 cm**-1
  Amplitude    Excitation
  -0.759965     6 ->   9
  -0.366007     6 ->  10
   0.514219     6 ->  13
  Ground state amplitude:  0.000746

Percentage Active Character     97.59

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
   0.865305     1 ->   7
  -0.246321     1 ->  10
  -0.230712     1 ->  13
   0.312193     1 ->  18
   0.113059     1 ->  25

IROOT=  4: 10.840510 au   294.985 eV 2379217.0 cm**-1
  Amplitude    Excitation
  -0.813493     6 ->   7
   0.345752     6 ->  12
   0.449353     6 ->  15
  Ground state amplitude:  0.000000

Percentage Active Character     98.33

  Amplitude    Excitation in Canonical Basis
   0.803886     1 ->   8
  -0.349815     1 ->   9
   0.318183     1 ->  11
  -0.124975     1 ->  12
   0.186068     1 ->  14
  -0.239028     1 ->  23

IROOT=  5: 10.840510 au   294.985 eV 2379217.0 cm**-1
  Amplitude    Excitation
   0.813492     6 ->   8
   0.345726     6 ->  11
  -0.449377     6 ->  16
  Ground state amplitude:  0.000000

Percentage Active Character     98.33

  Amplitude    Excitation in Canonical Basis
   0.349809     1 ->   8
   0.803873     1 ->   9
   0.124977     1 ->  11
   0.318188     1 ->  12
   0.186103     1 ->  15
  -0.239038     1 ->  24

IROOT=  6: 10.845730 au   295.127 eV 2380362.5 cm**-1
  Amplitude    Excitation
   0.116483     2 -> 136
   0.438618     6 ->   9
```

(continues on next page)

(continued from previous page)

```
  -0.879925     6 ->  10
   0.106779     6 ->  19
  Ground state amplitude:  0.000706

Percentage Active Character     96.67

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
  -0.244947     1 ->   7
  -0.900149     1 ->  10
  -0.170451     1 ->  13
  -0.205137     1 ->  18
  -0.192328     1 ->  22
   0.116486     3 -> 136

IROOT=  7: 10.906409 au   296.778 eV 2393680.2 cm**-1
  Amplitude    Excitation
   0.113733     2 -> 136
  -0.397156     6 ->   9
  -0.234927     6 ->  10
  -0.815420     6 ->  13
  -0.293892     6 ->  17
   0.126290     6 ->  30
  Ground state amplitude: -0.000417

Percentage Active Character     87.78

Warning:: the state may have not converged with respect to active space
-------------------- Handle with Care --------------------

  Amplitude    Excitation in Canonical Basis
   0.219961     1 ->   7
  -0.178103     1 ->  10
   0.883509     1 ->  13
  -0.239553     1 ->  18
  -0.188730     1 ->  19
  -0.161616     1 ->  25
   0.113733     3 -> 136

IROOT=  8: 10.926332 au   297.321 eV 2398052.6 cm**-1
  Amplitude    Excitation
  -0.365429     6 ->   8
   0.927126     6 ->  11
  Ground state amplitude:  0.000000

Percentage Active Character     99.66

  Amplitude    Excitation in Canonical Basis
  -0.129108     1 ->   8
  -0.295033     1 ->   9
   0.119506     1 ->  11
   0.302428     1 ->  12
   0.353687     1 ->  14
   0.777349     1 ->  15
   0.200807     1 ->  21

IROOT=  9: 10.926332 au   297.321 eV 2398052.8 cm**-1
  Amplitude    Excitation
   0.365449     6 ->   7
```

(continues on next page)

(continued from previous page)

```
   0.927109     6 ->  12
  Ground state amplitude:  0.000000

Percentage Active Character     99.66

  Amplitude    Excitation in Canonical Basis
  -0.295006     1 ->   8
   0.129096     1 ->   9
   0.302490     1 ->  11
  -0.119530     1 ->  12
   0.777245     1 ->  14
  -0.353640     1 ->  15
   0.201146     1 ->  20

IROOT= 10: 10.944288 au   297.809 eV 2401993.5 cm**-1
  Amplitude    Excitation
   0.998859     6 ->  14
  Ground state amplitude: -0.000000

Percentage Active Character     99.77

  Amplitude    Excitation in Canonical Basis
  -0.983120     1 ->  16
  -0.104304     1 ->  33
```

The core-valence separation should be used similar to that in the core-ionization. The only difference is that the natural orbital based active space selection scheme in STEOM-CCSD always rotate the particular core orbital to the HOMO. Therefore, CVSORB should always be HOMO in STEOM-CCSD irrespective of the core-hole. One should use the exact STEOM-CCSD transition moment by using DoSimpleDens False. Fig. 7.40 presents the STEOM-CCSD oxygen K-edge spectra in thymine.

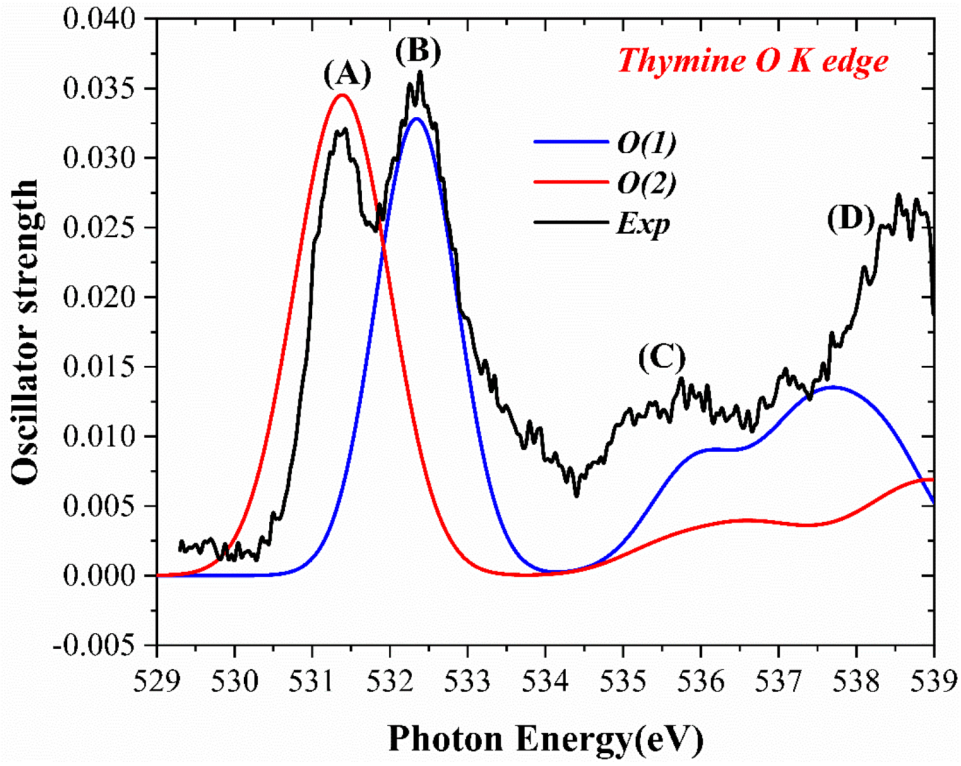


Fig. 7.39: Comparison of theoretical and experimental X-ray absorption spectra of oxygen K-edge in thymine. The simulated spectrum is shifted by -3.7 eV to align with the experimental spectrum.

One can interpret the results in terms of NTOs caculated from STEOM-CC eigen vectors

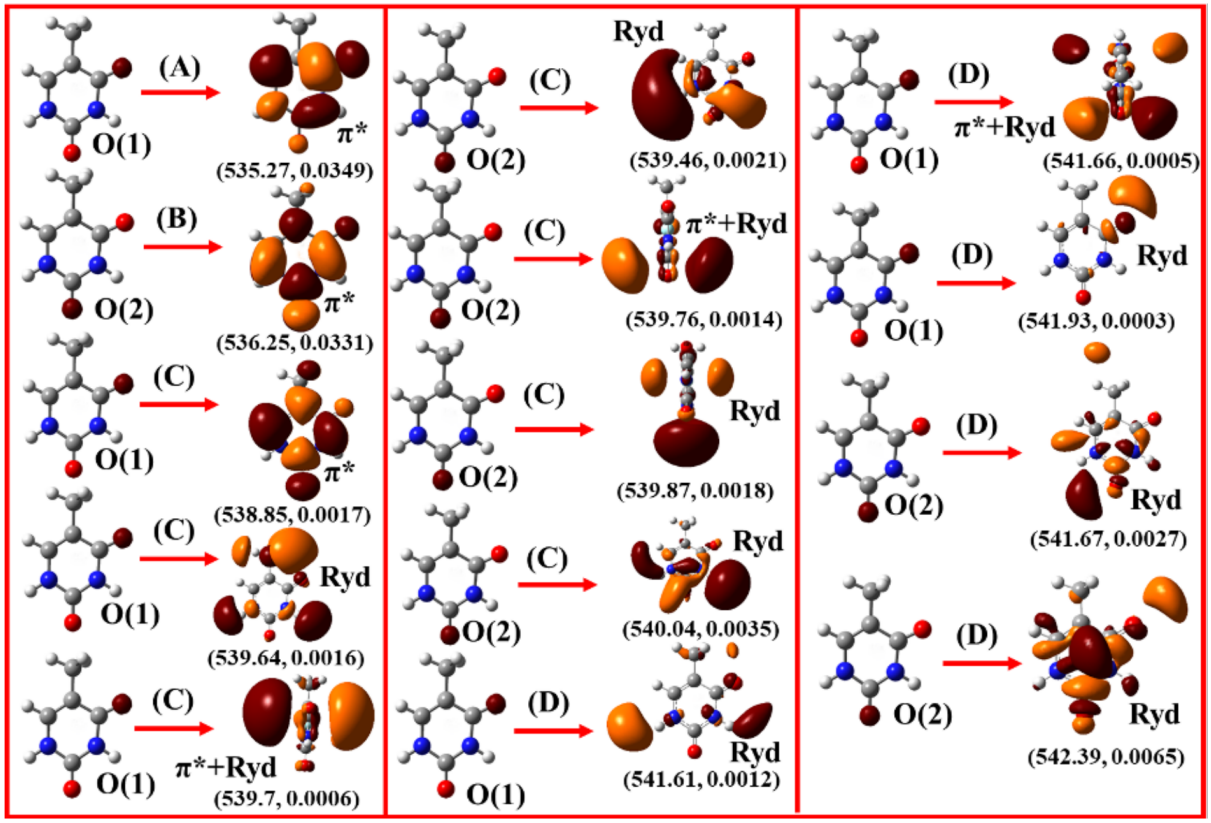


Fig. 7.40: Natural transition orbitals (ntos) for the oxygen K edge spectrum of thymine. All the core EE values mentioned are in eV and provided in the format (EE,Oscillator Strength).

7.39 The Multireference Correlation Module

7.39.1 General Description

A number of **uncontracted** multireference approaches are implemented in ORCA and reside in the `orca_mrci` module. All of these approaches start with a reference wavefunction that consists of multiple configurations (orbital occupation patterns). The reference wavefunction defined in the `ref` subblock can be a complete active space (CAS), restricted active space (RAS) or an arbitrary list of configurations. The total wavefunction is constructed by considering single and double excitations out of the reference configurations. These excited configurations are then used to generate configuration state functions (CSF) that have the proper spin and spatial symmetry. The number of wavefunction parameters rapidly grows with the number of reference functions. The orca_mrci module features a set of truncation criteria (`TSel, TPre, TNat`) that help to reduce the number of wavefunction parameters. Furthermore, by default, the program only considers reference configurations that already have the target spin and spatial symmetry. There are situations, where this is undesired and the restrictions can be lifted with the keyword `rejectinvalidrefs false`. For more information on the theory, the program module as well as its usage we recommend the review article by Neese et al.[628]. A tutorial type introduction to the subject is presented in chapter *The Multireference Correlation Module* of the manual and more examples in the CASSCF tutorial. The detailed documentation of all features of the MR-CI and MR-PT module is somewhat premature and at this point only a summary of keywords is given below. A thorough description of all technical and theoretical subtleties must wait for a later version of the manual.

The overall scaling of uncontracted approaches is steep. Hence, the methodology is restricted to small reference spaces and small molecules in general. **Note that all integrals must be kept in memory!** Internally contracted multireference approaches such as NEVPT2 do not share these bottlenecks. Aside from NEVPT2, ORCA features

a fully internally contracted MRCI (FIC-MRCI) that resides in the `orca_autoci` module. For more details on the FIC-MRCI we refer to section *CI methods using generated code*.

```
%mrci
       # ------------------------------------------------------------
       # Orbital selection
       # NOTE: The orbitals are used as supplied. Thus, the ORDER of
       # orbitals is critical. Say you have
       #   nact    electrons in the active space
       #   nint    electrons in the internal space
       #   nfrozen electrons
       # * The first nfrozen/2 orbitals will not be included in the CI
       # * The next nint/2 orbitals will be doubly occupied in all
       #   references
       # * the nact electrons are distributed over the,say, mact
       #   orbitals according to the active space definitions.
       # The remaining orbitals are external.
       # IT IS YOUR RESPONSIBILITY THAT THE ORBITAL ORDERING MAKES
       # SENSE!
       # A sensible two-step procedure is:
       #  * generate some orbitals and LOOK AT THEM. Decide which ones
       #    to include in the CI.
       #  * re-read these orbitals with ! MORead NoIter. Perhaps use
       #    the "rotate" feature to reorder the MOs
       #    Then jump right into the CI which is defined in this se-
       #    cond job
       #
       # NOTE: the MRCI module respects the %method FrozenCore settings
       # ------------------------------------------------------------
       Loc  0,0,0
           # Localize orbitals in the internal (first flag), active
           # (second flag) and external space (third flag).
       UseIVOs  false
           # Use improved virtual orbitals in the CI

       # ---------------------------------
       # Method selection
       # ---------------------------------
       CIType   MRCI      # Multireference CI (default)
                MRDDCI1   # Difference dedicated CI 1-degree of freedom
                MRDDCI2   # Difference dedicated CI 2-degrees of freedom
                MRDDCI3   # Difference dedicated CI 3-degrees of freedom
                MRACPF    # Average coupled-pair functional
                MRACPF2   # Modified version of ACPF
                MRACPF2a  # A slightly modified version of ACPF-2a
                MRAQCC    # Average quadratic coupled-cluster
                MRCEPA_R  # Multireference CEPA due to Ruttink
                MRCEPA_0  # CEPA-0 approximation
                SORCI     # Spectroscopy oriented CI
                SORCP     # Spectroscopy oriented couplet pair approx.
                MRMP2     # Multireference Moeller-Plesset at second order
                MRMP3     # Multireference Moeller-Plesset at third order
                MRMP4     # Multireference Moeller-Plesset at fourth order
                          # but keeping only singles and doubles relative to
                          # the reference configurations.

       # ---------------------------------
       # Selection thresholds
       # ---------------------------------
       Tsel    1e-6   # Selection threshold for inclusion in the CI based
                      # 2nd order MP perturbation theory <0|H|I>/DE(MP)
       Tpre    1e-4   # Selection of configurations in the reference space
                      # after the initial diagonalization of the reference
```

(continues on next page)

(continued from previous page)

```
                # space only configurations with a weight large>Tpre
                # to any root are included
    AllSingles   false
                # include ALL SINGLES in the CI. Default is now TRUE!!!

    # perturbative estimate of the effect of the rejected configurations
    EunselOpt  0   # no correction
               1   # based on the overlap with the 0th order wavefunction
               2   # calculation with the relaxed reference space
                   # coefficients. This is the most accurate and only
                   # slightly more expensive

    # For CIType=MRCI,MRDDCI and SORCI the approximate correction for
    # higher excitations
    DavidsonOpt  Davidson1  # default
                 Davidson2  # modified version
                 Siegbahn   # Siegbahn's approximation
                 Pople      # Pople's approximation

    # For MRACPF,MRACPF2,MRAQCC and SORCP
    NelCorr      0
       # Number of electrons used for computing the average coupled-
       # pair correction.
       # =0 : set equal to ALL electrons in the CI
       # =-1: set equal to all ACTIVE SPACE electrons
       # =-2: set equal to ACTIVE SPACE electrons IF inactive doubles
       #      are excluded (as in MRDDCI)
       # >0 : set equal to user defined input value
    LinearResponse false
       # Use ground state correlation energy to compute the shift for
       # higher roots (not recommended)

    # ---------------------------------
    # Natural Orbital Iterations
    # ---------------------------------
    NatOrbIters    0 # default
       # number of average natural orbital iterations
    Tnat        1e-4
       # cutoff of natural orbitals. NOs with an occupation number less
       # then Tnat will not be included in the next iteration
       # Also, orbitals with occupation number closer than Tnat to 2.0
       # will be frozen in the next iteration
    Tnat2       -1
       # if chosen >0 then Tnat2 is the threshold for freezing the
       # almost doubly occupied orbitals. Otherwise it is set equal
       # to Tnat

    # ---------------------------------
    # Additional flags and algorithmic
    # details
    # ---------------------------------
    PrintLevel   2  # default. Values between 1 and 4 are possible

    DoDDCIMP2  false
       # for DDCI calculations: if set to true the program computes
       # a MP2 like correction for the effect of inactive double
       # excitations which are not explicitly included in the CI. This
       # is necessary if you compare molecules at different geometries
       # or compute potential energy surfaces.
    # ---------------------------------
    # The SORCP model
```

(continues on next page)

(continued from previous page)

```
      # ---------------------------------
      CIType_in  # First step CIType
      CIType_fi  # Second step CIType
      Exc_in     # First step excitation scheme
      Exc_fi     # Second step excitation scheme
      Tsel_in    # First step Tsel
      Tsel_fi    # Second step Tsel
      Tpre_in    # First step Tsel
      Tpre_fi    # Second step Tpre

        # Thus, the SORCI model corresponds to CIType=SORCP with
        # CIType_in MRCI  CIType_fi MRCI
        # Exc_in    DDCI2 Cexc_fi   DDCI3
        # Tsel_in   1e-5  Tsel_fi   1e-5
        # Tpre_in   1e-2  Tpre_fi   1e-2

      # ---------------------------------
      # Multirerence perturbation theory
      # ---------------------------------
      MRPT_b  0.02   # Intruder state avoidance PT after Hirao (default 0.0)
                     # with this flag individual intruders are shifted away to
                     # to some extent from the reference space
      MRPT_shift 0.3 # Level shift introduced by Roos which shifts the entire
                     # excited manifold away in order to avoid intruder states.
                     # A correction is applied afterwards but results do depend
                     # on this (arbitrary) value to some extent.
      H0Opt  projected  # use an off-diagonal definition of H0
             Diagonal   # use a diagonal definition of H0 (much faster but maybe
                        # a little less reliable
      Partitioning MP   # Moeller plesset partitioning
                   EN   # Epstein-Nesbet partitioning (not recommended)
      Fopt  Standard    # Standard definition of MR Fock operators
            G3          # uses Anderson's g3 correction also used in CASPT2


      #---------------------------------------
      # restrict reference configurations
      #---------------------------------------
      RejectInvalidRefs true # by default reference CSFs are restricted
                             # to target spin and spatial symmetry

      # ====================================
      # Definitions of blocks of the CI Matrix
      # ====================================
      NewBlock 2 *    # generate a Block with doublet(=2) multiplicity
        Nroots  1     # number of roots to be generated
        Excitations cis   # CI with single excitations
                    cid   # CI with double excitations
                    cisd  # CI with single and double excitations
                    ddci1 # DDCI list with one degree of freedom
                    ddci2 # DDCI list with two degrees of freedom
                    ddci3 # DDCI list with three degrees of freedom
        Flags[_class_] 0 or 1
                  # Turn excitation classes on or off individually
                  # ``s" stands for any SOMO, ``i",``j" for internal orbitals and
                  # ``a",``b" for external orbitals
                  #  Singles _class_ = ss, sa, is, ia
                  #  Doubles _class_ = ijss, ijsa, ijab,
                  #                    isss, issa, isab,
                  #                    ssss, sssa, ssab
                  # ``Flags" takes priority over ``Excitations". In fact ``Excitations"
```

(continues on next page)

(continued from previous page)

```
            # does nothing but to set ``Flags". So, you can use ``Excitations"
            # to provide initial values for ``Flags" and then modify them
            # with subsequent ``Flags" assignments
      refs
            #
            # First choice - complete active space
            #
            CAS(nel,norb)  # CAS-CI reference with nel electrons in
                           # Norb orbitals
            #
            # Second choice - restricted active space
            #
            RAS(nel: m1 h/ m2 / m3 p)
                           # RAS-reference with nel electrons
                  # m1= number orbitals in RAS-1
                  # h = max. number of holes in RAS-1
                  # m2= number of orbitals in RAS-2 (any number of
                  #     electrons or holes)
                  # m3= number of orbitals in RAS-3
                  # p = max. number of particles in RAS-3
            #
            # Third choice - individually defined configurations
            #
             \{ 2 0 1 0\}
             \{ 1 1 1 0\}
               etc.
               # define as many configurations as you want. Doubly occupied MOs
               # singly occupied MOs and empty MOs. Important notes:
               #  a) the number of electrons must be the same in all references
               #  b) the number of orbitals is determined from the number of
               #     definitions. Thus, in the example above we have three active
               #     electrons and four active orbitals despite the fact that the
               #     highest orbital is not occupied in any reference.
               #  The program determines the internal, active and external spaces
               #  automatically from the number of active electrons and orbitals
            end
      end
      # there can be as many blocks as you want!!!

    # ---------------------------------
    # Density matrix generation flags
    #  First Key= State densities <I|D|I>
    #     =0: none
    #     =1: Ground state only (lowest root of all blocks; Electron only)
    #     =2: Ground state only (Electron and spin density)
    #     =3: Lowest root from each block (Electron density)
    #     =4: Lowest root from each block (Electron and spin density)
    #     =5: All states (Electron density)
    #     =6: All states (Electron and spin density)
    #  Second Key= Transition densities <I|D|J>
    #     needed for all transition intensities, g-tensor etc
    #     =0: none
    #     =1: from the ground state into all excited states (el)
    #     =2: from the ground state into all excited states (el+spin)
    #     =3: from all lowest states into all excited states (el)
    #     =4: from all lowest states into all excited states (el+spin)
    #     =5: all state pairs (el)
    #     =6: all state pairs (el+spin)
    # Note that for perturbation theory the density is computed as
    # an expectation value over the first (second) order wavefunction.
    # which is renormalized for this purpose
```

(continues on next page)

(continued from previous page)

```
# ----------------------------------
Densities 1,1

# ----------------------------------
# Complete printing of the wavefunction
# ----------------------------------
PrintWF    1      # CFG based printing (default)
           det    # Determinant based wavefunction printing
TPrintWF   3e-3   # Threshold for the printing of the CFGs/Dets

# ----------------------------------
# Algorithm for the solver
# ----------------------------------
Solver   Diag   # Davidson like solver
         DIIS   # DIIS like solver
  # both solvers have their pros and cons. The DIIS may converge
  # better or use less time since it only recomputes the vectors that
  # have not yet converged; The DIIS may be less sensitive to root flipping
  # effects but occasionally it converges poorly and states of the same
  # symmetry are occasionally a little problematic
  # For perturbation theory DIIS is always used.
# For both solvers
MaxIter   100  # the maximum number of iterations
Etol      1e-6 # convergence tolerance for energies in Eh
Rtol      1e-6 # convergence tolerance for residual

# For Solver=Diag (Davidson solver)
Otol      1e-16 # Orthogonality threshold for Schmidt process
NGuessMat 512   # Dimension of the guess matrix 512x512
                # be used to compute the initial guess of the actual MRCI calculation
NGuessMatRefCI 512 # Dimension of the guess matrix
                   # for the reference CI

MaxDim       3  # Davidson expansion space = MaxDim * NRoots
# For the Solver=DIIS. Particularly recommended for anything else but
# straightforward CI and also for calculations in direct2 mode!
MaxDIIS       5 # Maximum number of old vectors to be used in DIIS
RelaxRefs  true # Relax reference space coefficients in the CI or
                # freeze them to their zeroth order values
LevelShift  0.4 # Level Shift for stabilizing the DIIS procedure

# ----------------------------------
# RI Approximation
# ----------------------------------
IntMode RITrafo #Use RI integrals
        FullTrafo #No RI (default)

# ----------------------------------
# Integral storage, memory and files
# ----------------------------------
IntStorage    FloatVals
              DoubleVals  (default)
   # store integrals with float (4 byte) or double (8 byte)
   # accuracy in main memory
FourIndexInts false (default)
              True
   # Store ALL four index integrals over Mos in main memory
   # only possible for relatively small systems, perhaps up
   # to 150-200 MOs included in the CI
MaxMemInt     256
   # Maximum amount of core memory devoted to the storage of
```

(continues on next page)

(continued from previous page)

```
        # integrals. If NOT all three index integrals fit into main
        # memory the program fails
    MaxMemVec      16
        # Maximum amount of memory used for section of the trial and
        # sigma vectors. This is not a particularly critical variable
    KeepFiles    false
        # Keep integrals and CI program input file (.mrciinp). Then
        # you can manually edit the .mrciinp file which is a standard
        # ASCII file and run the MRCI program directly. The only thing
        # you cannot change is the orbital window.
    end
```

7.39.2 Properties Calculation Using the SOC Submodule

Zero-Field Splitting

The spin-orbit coupling (SOC) and spin-spin coupling (SSC) contributions to the zero-field splitting (ZFS) can be calculated very accurately using a wavefunction obtained from a multiconfigurational calculation of a multi-reference type such as CASSCF, MRCI, or MRPT as is described in QDPT Magnetic Properties Section *Magnetic properties through Quasi Degenerate Perturbation Theory*.

```
# In case that you want to run QDPT-SOC calculation with manually
#adjusted diagonal energies you can copy the following part into
#the %mrci soc block
#and modify it as needed(energies are given in
#wavenumbers relative to the lowest state)
# NOTE: It is YOUR responsibility to make sure that the CAS-CI state
#that you may want to dress with these energies correlate properly
#with the energies printed here. The order of states or even the
#identity of states may change with and without inclusion of
#dynamic correlation In the case that dynamic correlation strongly
#mixes different CAS-CI states there may not even be a proper
#correlation!
#
     EDiag[  0]        0.00  # root   0 of block 0
     EDiag[  1]    48328.40  # root   1 of block 0
     EDiag[  2]    48328.40  # root   2 of block 0
     EDiag[  3]    49334.96  # root   3 of block 0
     EDiag[  4]     7763.59  # root   0 of block 1
     EDiag[  5]     7763.59  # root   1 of block 1
     EDiag[  6]    11898.46  # root   2 of block 1
     EDiag[  7]    46754.23  # root   3 of block 1
```

Those transition energies can be substituted by a more accurate energies provided in the input file as follows:

```
%soc
dosoc true
dossc true
     EDiag[  0]        0.00  # root   0 of block 0
     EDiag[  1]    48328.40  # root   1 of block 0
     EDiag[  2]    48328.40  # root   2 of block 0
     EDiag[  3]    49334.96  # root   3 of block 0
     EDiag[  4]     7763.59  # root   0 of block 1
     EDiag[  5]     7763.59  # root   1 of block 1
     EDiag[  6]    11898.46  # root   2 of block 1
     EDiag[  7]    46754.23  # root   3 of block 1
end
```

Accurate diagonal energies generally improve the accuracy of the SOC and SSC splittings.

Local Zero-Field Splitting

The submodule can also be used to calculate the local ZFS splitting parameters of atomic centers. The method, referred to as local complete active space configuration interaction (L-CASCI), can be used to separate into atomic contributions the SOC part of the total ZFS tensor. The rational behind it and additional details are described in the original publication [716]; below are listed only the steps required to reproduce the calculation for the dimer complex presented there.

1. The first step consists in obtaining the molecular orbitals that are going to be used in the configuration interaction (CI) procedure. A good set of orbitals can be obtained from a restricted open-shell spin-averaged Hartree-Fock (SAHF) calculation. The relevant part of the input is listed below:

```
! def2-tzvp keepfock

% scf
    hftyp rohf
    rohf_case sahf
    rohf_numop 2
    rohf_nel[1] 9
    rohf_norb[1] 10
    end
```

For the present Mn(II)Mn(III) dimer there are a total of 9 electrons distributed into 10 d-orbitals.

2. Next, the molecular orbitals are localized using one of the implemented localization schemes. Below is the `orca_loc` input used in this case:

```
sahf.gbw
sahf.loc
0
200 # first of the 10 d-orbitals
209 # last of the 10 d-orbitals
128
0.000001
0.75
0.65
2
```

3. Following this, the localized orbitals are made locally canonical by block diagonalizing the Fock matrix using the `orca_blockf` utility.

```
orca_blockf sahf.fsv sahf.loc 200 204 205 209
```

The first two numbers define the range of molecular orbitals localized on one center; the last two are for the second center.

4. The recanonicalized orbitals stored in the `sahf.loc` file can be then used to calculate the SOC contribution to the local ZFS of the Mn(III) center using the following MRCI input:

```
! zora-def2-tzvp def2-tzvp/c zora
! nomulliken noloewdin
! moread noiter allowrhf
! moread

% mrci
    citype mrci
    tsel 0
    tpre 0
    intmode ritrafo
    solver diis
    soc
        intmode ritrafo
```

(continues on next page)

(continued from previous page)

```
    dosoc true
    end
  newblock 10 *
    nroots 5
    excitations none
    refs
      #  Mn(II)    Mn(III)
      {1 1 1 1 1  1 1 1 1 0}
      {1 1 1 1 1  1 1 1 0 1}
      {1 1 1 1 1  1 1 0 1 1}
      {1 1 1 1 1  1 0 1 1 1}
      {1 1 1 1 1  0 1 1 1 1}
      end
    end
  newblock 8 *
    nroots 45
    excitations none
    refs
      #  Mn(II)    Mn(III)
      {1 1 1 1 1  2 1 1 0 0}
      {1 1 1 1 1  2 1 0 1 0}
      {1 1 1 1 1  2 1 0 0 1}
      {1 1 1 1 1  2 0 1 1 0}
      {1 1 1 1 1  2 0 1 0 1}
      {1 1 1 1 1  2 0 0 1 1}
      {1 1 1 1 1  1 2 1 0 0}
      {1 1 1 1 1  1 2 0 1 0}
      {1 1 1 1 1  1 2 0 0 1}
      {1 1 1 1 1  1 1 2 0 0}
      {1 1 1 1 1  1 1 1 1 0}
      {1 1 1 1 1  1 1 1 0 1}
      {1 1 1 1 1  1 1 0 2 0}
      {1 1 1 1 1  1 1 0 1 1}
      {1 1 1 1 1  1 1 0 0 2}
      {1 1 1 1 1  1 0 2 1 0}
      {1 1 1 1 1  1 0 2 0 1}
      {1 1 1 1 1  1 0 1 2 0}
      {1 1 1 1 1  1 0 1 1 1}
      {1 1 1 1 1  1 0 1 0 2}
      {1 1 1 1 1  1 0 0 2 1}
      {1 1 1 1 1  1 0 0 1 2}
      {1 1 1 1 1  0 2 1 1 0}
      {1 1 1 1 1  0 2 1 0 1}
      {1 1 1 1 1  0 2 0 1 1}
      {1 1 1 1 1  0 1 2 1 0}
      {1 1 1 1 1  0 1 2 0 1}
      {1 1 1 1 1  0 1 1 2 0}
      {1 1 1 1 1  0 1 1 1 1}
      {1 1 1 1 1  0 1 1 0 2}
      {1 1 1 1 1  0 1 0 2 1}
      {1 1 1 1 1  0 1 0 1 2}
      {1 1 1 1 1  0 0 2 1 1}
      {1 1 1 1 1  0 0 1 2 1}
      {1 1 1 1 1  0 0 1 1 2}
      end
    end
  end
```

5. The three second order ZFS components printed at the end of the calculation (`Second order D-tensor: component 0`, etc.) are scaled using the S value for the complex, which in this case is 4.5 (9 electrons $\times$ 0.5). In order to obtain the correct local value of the ZFS, the three matrices have to be rescaled using the S value for Mn(III), which is to 2. Note that the three matrices have different scaling prefactors, and the dependence on S is

not the same:

$\mathbf{D}^{SOC-(0)} \propto \frac{1}{S^2}$

$\mathbf{D}^{SOC-(-1)} \propto \frac{1}{S(2S-1)}$

$\mathbf{D}^{SOC-(+1)} \propto \frac{1}{(S+1)(2S+1)}$

These equations can be used to calculate the required prefactors. For example in the case of the *SOC*-(0) the prefactor is equal to:

$\mathbf{D}^{SOC-(0)}_{\mathrm{Mn(III)}} = \frac{4.5^2}{2^2} \cdot \mathbf{D}^{SOC-(0)}_{\mathrm{dimer}} = 5.0625 \cdot \mathbf{D}^{SOC-(0)}_{\mathrm{dimer}}$

The final step is to scale the two remaining matrices using the appropriate prefactors, sum all three of them up, diagonalize the resulting the matrix, and use its eigenvalues to calculate the *D* and *E* parameters. These represent the local ZFS parameters of the Mn(III) center.

Zero-Field Splitting from an excited Multiplet

For an excited state Multiplet the Calculationof ZFS can be requested by

```
Lowest eigenvalue of the SOC matrix:    -149.86223277 Eh
Energy stabilization:    -2.54512 cm-1
Eigenvalues:     cm-1         eV      Boltzmann populations at T =  300.000 K
   0:            0.00        0.0000       3.36e-01
   1:            2.37        0.0003       3.32e-01
   2:            2.37        0.0003       3.32e-01
   3:         7757.65        0.9618       2.33e-17
   4:         7757.66        0.9618       2.33e-17
   5:        11913.81        1.4771       5.15e-26
```

```
soc
 DTensor true
 IStates 3,4,5
end
```

```
****************************************
 EXCITED STATE ZERO-FIELD SPLITTING:
****************************************

-------------------------------------------
Computing Excited State D Tensors of
Excited State Multiplet Consisting of States :  3  4  5
-------------------------------------------
          0         4
          1         5
          2         0
          3         1
          4         0
          5         2


-------------------------------------------
           ZERO-FIELD SPLITTING
(2ND ORDER SPIN-ORBIT COUPLING CONTRIBUTION)
-------------------------------------------

D   =   -2.668445  cm-1
E/D =    0.000103


...
```

(continues on next page)

(continued from previous page)

```
------------------------------------------------------------
                  ZERO-FIELD SPLITTING
EFFECTIVE HAMILTONIAN SOC CONTRIBUTION
------------------------------------------------------------

D   =   -2.674495  cm-1
E/D =    0.009610

...
```

g-Tensor

The `orca_mrci` program contains an option to calculate g-tensors using MRCI wavefunctions. For a system with an odd number of electrons, the doubly degenerate eigenvalues obtained from the QDPT procedure represent Kramers pairs, which are used to build the matrix elements of the total spin operator and the total angular momentum operator from the Zeeman Hamiltonian. Denoting Ψ as a solution and $\bar{\Psi}$ as its Kramers partner and using matrix element notations

$$\Phi_{11}^k = \langle\Psi|\,\hat{L}_k + g_e\hat{S}_k\,|\Psi\rangle\,,\ \Phi_{12}^k = \langle\Psi|\,\hat{L}_k + g_e\hat{S}_k\,|\bar{\Psi}\rangle\,,\ k = x,y,z \tag{7.240}$$

The elements of g-matrix are obtained as:

$$g_{kz} = 2\Phi_{11}^k,\ g_{ky} = -2\Im\left(\Phi_{12}^k\right),\ g_{kx} = 2\Re\left(\Phi_{12}^k\right) \tag{7.241}$$

Then, the true tensor G is built from g-matrices:

$$G = gg^T \tag{7.242}$$

G is subjected further to diagonalization yielding positive eigenvalues, the square roots of which give the principal values of g-matrix.

$$g_{xx} = \sqrt{G_{xx}},\ g_{yy} = \sqrt{G_{yy}},\ g_{zz} = \sqrt{G_{zz}} \tag{7.243}$$

A typical mrci block of the input file for a g-tensor calculation should (e.g. for a S=3/2 problem) look as the following:

```
%mrci   ewin -4,1000
        citype mrci
        cimode direct2
        intmode fulltrafo
        solver diis
        etol 1e-8
        rtol 1e-8
        tsel 1e-6
        tpre 1e-5
        soc
          PrintLevel 2
          GTensor true      # make g-tensor calculations
          NDoubGTensor 2    # number of Kramers doublets to account
                            # for every pair a separate
                            # calculation is performed
        end
        newblock 4 *
          excitations cisd
          nroots 10
          refs  cas(7,5) end
        end
end
```

The result for the first Kramers pair is printed as follows:

```
--------------
KRAMERS PAIR 1
--------------

Matrix elements Re<1|S|1>   -0.072128   0.024511  -2.998843
Matrix elements Re<1|S|2>   -0.001088   0.000366  -0.002010
Matrix elements Im<1|S|2>   -0.000354  -0.001037  -0.000173
Matrix elements Re<1|L|1>   -0.027067   0.009209  -1.123531
Matrix elements Re<1|L|2>   -0.000031   0.000010  -0.000763
Matrix elements Im<1|L|2>   -0.000006  -0.000011  -0.000065

-------------------
ELECTRONIC G-MATRIX
-------------------

g-matrix:
-0.002240     0.000754    -0.005551
0.000720     0.002100     0.000477
-0.198556     0.067498    -8.251703

g-factors:
0.002220     0.002222     8.254370 iso =     2.752937

g-shifts:
-2.000100    -2.000098     6.252051 iso =     0.750618

Eigenvectors:
0.057426     0.998060     0.024055
0.998327    -0.057244    -0.008177
0.006784    -0.024484     0.999677
```

Here for the L and S matrix elements indices 1 and 2 are assumed to denote Kramers partners, and three numbers in the first row stand for x, y, z contributions.

In addition the g-tensor is calculated within the Effective Hamiltonian formalism.

```
-----------------------------------------------
        ELECTRONIC G-MATRIX FROM EFFECTIVE HAMILTONIAN
        -----------------------------------------------

        g-matrix:
        1.978874    -0.000345     0.018908
        -0.000345     1.977899    -0.006433
        0.018879    -0.006418     2.763402

        g-factors:
        1.977789     1.978477     2.763909 iso =     2.240058

        g-shifts:
        -0.024530    -0.023843     0.761590 iso =     0.237739

        Eigenvectors:
        0.288884     0.957062     0.024060
        0.957364    -0.288770    -0.008181
        0.000882    -0.025397     0.999677

        # The g-factors are square roots of the eigenvalues of gT*g
        # Orientations are the eigenvectors of gT*g
```

Finally and only within the MRCI module the g-tensor is evaluated by using the Sum Over States formalism[609]:

```
-------------------------------------------------------------------------------
        SUM OVER STATES CALCULATION OF THE SPIN HAMILTONIAN (for g and HFC tensors)
        -----------------------------------------------------------------------


        Ground state index        =     0
        Ground state multiplicity =     4
        Ground state spin density = P[  1]
        State =   1  <0|P|I>=   2  <0|Q|I>=  19
        State =   2  <0|P|I>=   3  <0|Q|I>=  27
        State =   3  <0|P|I>=   4  <0|Q|I>=  34
        State =   4  <0|P|I>=   5  <0|Q|I>=  40
        State =   5  <0|P|I>=   6  <0|Q|I>=  45
        State =   6  <0|P|I>=   7  <0|Q|I>=  49
        State =   7  <0|P|I>=   8  <0|Q|I>=  52
        State =   8  <0|P|I>=   9  <0|Q|I>=  54
        State =   9  <0|P|I>=  10  <0|Q|I>=  55

        Origin for angular momentum        ... ( -0.0006, -0.0010,  0.0021)
        Kinetic Energy                     ... done
        Relativistic mass correction       ... done
        Gauge correction                   ... done
        Angular momentum integrals         ... done
        Reading Spin-Orbit Integrals       ... done
        ------------------------
        MATRIX ELEMENT PRINTING
        ------------------------

        Energy differences (DE=EI-E0) and spin-orbit matrix elements (SO=<I|HSO|0>) are
        printed in cm**-1. Orbital Zeeman matrix elements (L=<I|L|0>) are printed in au.

        State     DE         LX        LY        LZ           SOX         SOY         SOZ

        1    1349.3     0.0464   -0.0158    1.9264     -23.432       7.965    -974.312
        2   13026.2    -0.6596    0.6888    0.0214     337.028    -351.116     -10.966
        3   13615.1    -0.6961   -0.6514    0.0113     354.225     332.219      -5.736
        4   56686.3    -0.0053    0.0077    0.0971       1.794      -1.696     -36.786
        5   56954.2    -0.0516   -0.0048   -0.0042      28.211       5.821       1.459
        6   56994.0    -0.0418    0.0233   -0.0025      15.185      -2.144       1.145
        7   63371.5    -0.0211    0.0226    0.0078       3.833      -2.948      -2.724
        8   64176.0    -0.0652    0.0032   -0.0002      32.779       6.146       0.063
        9   74309.9    -0.0007    0.0032    0.0380       0.183      -1.058     -13.517

        -------------------
        ELECTRONIC G-MATRIX
        -------------------

        raw-matrix :
        2.025533    -0.000738     0.021755
        -0.000738     2.024537    -0.007389
        0.021755    -0.007389     2.928943

        g-factors:
        2.024122     2.025363     2.929527 iso =     2.326338

        g-shifts:
        0.021803     0.023044     0.927208 iso =     0.324018

        Eigenvectors:
        0.533896    -0.845208     0.024064
        0.845530     0.533866    -0.008182
        -0.005932     0.024715     0.999677
```

(continues on next page)

(continued from previous page)

```
Euler angles w.r.t. molecular frame (degrees):
-76.5038       1.4564    -161.2223

------------------------------
CONTRIBUTIONS TO THE G-MATRIX
------------------------------

Term                                g1           g2           g3
-----------------------------------------------------------------------
Relativistic mass correction:   -0.0008220   -0.0008220   -0.0008220
Gauge correction            :    0.0000000    0.0000000    0.0000000
g(OZ/SOC)                   :    0.0226250    0.0238662    0.9280297

State   1                           :    0.0000000   -0.0000000    0.9279829
State   2                           :    0.0013767    0.0223913    0.0000000
State   3                           :    0.0212332    0.0014408    0.0000000
State   4                           :    0.0000000    0.0000004    0.0000418
State   5                           :    0.0000074    0.0000099    0.0000001
State   6                           :    0.0000002    0.0000078    0.0000001
State   7                           :    0.0000000    0.0000015    0.0000002
State   8                           :    0.0000076    0.0000144    0.0000000
State   9                           :    0.0000000    0.0000000    0.0000046
----------------------------------------
Total g-shifts              :    0.0218030    0.0230442    0.9272077


# The g-factors are square roots of the eigenvalues of gT*g
# Orientations are the eigenvectors of gT*g
```

Note that within the SOS formalism in addition to the second order (SOC) contributions the bilinear to the field terms: Relativistic mass correction and diamagnetic spin-orbit term (Gauge) are evaluated. As can be seen these corrections are rather negligible in comparison to the second order SOC contributions and most of the time can be safely omitted. Moreover further insight is obtained by printing the individual contribution of each excited state to the g-tensor. In the example above the first excited state contributes to the g_z component while the next two to both the g_x and g_y components, respectively.

So to summarize the g-tensor calculations in the framework of wavefunction based methods like MRCI and/or CASSCF can be evaluated:

- via the QDPT approach within an individual Kramers doublet. This is valid analysis only for non-integer spin cases. In particular for systems with well isolated Kramers doublets where the EPR spectrum originates only from one Kramers doublet defined within the pseudo spin 1/2 formalism. This analysis has been proven useful in determining the sign of the ZFS and the electronic structure of the system under investigation.[547]
- within the effective Hamiltonian approach. This is a valid analysis for all spin cases as it provides the principal g-values of the system under investigation evaluated in the molecular axis frame. These g-values can be directly compared with the experimentally determined ones.[421]
- within the sum over states formalism (SOS). As above this analysis is valid for all spin cases and is only available via the MRCI module.

Magnetization and Magnetic Susceptibility

The MRCI and CASSCF modules of ORCA allow for the calculation of magnetization and magnetic susceptibility curves at different fields and temperatures by differentiation of the QDPT Hamiltonian with respect to the magnetic field. For magnetic susceptibility, calculations are performed in two ways when a static field different from zero is defined: (i) as the second derivative of energy with respect to the magnetic field and (ii) as the magnetization divided by the magnetic field. Although the first method corresponds to the definition of magnetic susceptibility, the second approach is widely used in the experimental determination of $\chi * T$ curves. If the static field is low, both formulas tend to provide similar values.

The full list of keywords is presented below.

```
%mrci
      citype mrci
      newblock 3 *
      excitations none
      refs cas(2,7) end
      end
      soc
        dosoc true
        domagnetization true       # Calculate magnetization (def: false)
        dosusceptibility true      # Calculate susceptiblity (def: false)
        LebedevPrec            5   # Precision of the grid for different field
                                   # directions (meaningful values range from 1
                                   # (smallest) to 10 (largest))
        nPointsFStep           5   # number of steps for numerical differentiation
                                   # (def: 5, meaningful values are 3, 5 7 and 9)
        MAGFieldStep       100.0   # Size of field step for numerical differentiation
                                   # (def: 100 Gauss)
        MAGTemperatureMIN    4.0   # minimum temperature (K) for magnetization
        MAGTemperatureMAX    4.0   # maximum temperature (K) for magnetization
        MAGTemperatureNPoints  1   # number of temperature points for magnetization
        MAGFieldMIN          0.0   # minimum field (Gauss) for magnetization
        MAGFieldMAX      70000.0   # maximum field (Gauss) for magnetization
        MAGNpoints            15   # number of field points for magnetization
        SUSTempMIN           1.0   # minimum temperature (K) for susceptibility
        SUSTempMAX         300.0   # maximum temperature (K) for susceptibility
        SUSNPoints           300   # number of temperature points for susceptibility
        SUSStatFieldMIN      0.0   # minimum static field (Gauss) for susceptibility
        SUSStatFieldMAX      0.0   # maximum static field (Gauss) for susceptibility
        SUSStatFieldNPoints    1   # number of static fields for susceptibility
      end
end
```

The same keywords apply for CASSCF calculations in rel block (instead of soc in MRCI). Although different aspects of integration and grid precision can be modified through keywords, default values should provide an accurate description of both properties. Calculated magnetization and susceptibility are printed in .sus and .mag files, respectively and also in the output file.

```
-------------------------------------------------------------------------------
FIELD DEPENDENT MAGNETIZATION AND MEAN SUSCEPTIBILITY (chi=M/B)
-------------------------------------------------------------------------------
TEMPERATURE (K)   M. FIELD (Gauss)   MAGNETIZATION (B.M.)   chi*T (cm3*K/mol)
-------------------------------------------------------------------------------

4.00               0.00               0.000000                inf
4.00            5000.00               0.350759            1.567189
4.00           10000.00               0.688804            1.538788
4.00           15000.00               1.003466            1.494496
4.00           20000.00               1.287480            1.438115
4.00           25000.00               1.537346            1.373773
4.00           30000.00               1.752841            1.305282
```

(continues on next page)

(continued from previous page)

```
4.00          35000.00              1.936067            1.235764
4.00          40000.00              2.090450            1.167516
4.00          45000.00              2.219920            1.102067
4.00          50000.00              2.328368            1.040315
4.00          55000.00              2.419335            0.982690
4.00          60000.00              2.495883            0.929301
4.00          65000.00              2.560582            0.880052
4.00          70000.00              2.615538            0.834730
-----------------------------------------------------------
```

```
-----------------------------------------------------------
        TEMPERATURE DEPENDENT MAGNETIC SUSCEPTIBILITY
        -------------------------------------------------------
        STATIC  FIELD    TEMPERATURE       chi*T (cm3*K/mol)
        (Gauss)            (K)          M/B          d2E/dB2
        -------------------------------------------------------

        0.00       1.00           ----       1.576836
        0.00       2.00           ----       1.576910
        0.00       3.00           ----       1.576951
        0.00       4.00           ----       1.576988
        0.00       5.00           ----       1.577023
        0.00       6.00           ----       1.577057
        0.00       7.00           ----       1.577091
        0.00       8.00           ----       1.577125
        0.00       9.00           ----       1.577159
        0.00      10.00           ----       1.577193
        0.00      11.00           ----       1.577227
        .....
        0.00     300.00           ----       1.586942
        1000.00       1.00       1.570517       1.558042
        1000.00       2.00       1.575324       1.572178
        1000.00       3.00       1.576246       1.574845
        1000.00       4.00       1.576590       1.575802
        1000.00       5.00       1.576768       1.576264
        1000.00       6.00       1.576880       1.576530
        1000.00       7.00       1.576961       1.576704
        1000.00       8.00       1.577026       1.576829
        .....
```

Note that the CASSCF module also supports the calculation of susceptibility tensors at non-zero user-defined magnetic fields. This is not yet possible with the MRCI module.

MCD and Absorption Spectra

The MRCI module of the ORCA program allows calculating MCD spectra and the SOC effects on absorption spectra. The formalism is described in detail by Ganyushin and Neese[283]. The approach is based on the direct calculation of the transition energies and transition probabilities between the magnetic levels. Namely, the differential absorption of LCP- and RCP photons for transitions from a manifold of initial states A to a manifold of final states J. Using Fermi's golden rule, the Franck-Condon approximation, assuming a pure electronic dipole mechanism and accounting for the Boltzmann populations of the energy levels, the basic equation of MCD spectroscopy may be written as (atomic units are used throughout):

$$\frac{\Delta\varepsilon}{E} = \gamma \sum_{a,j} \left(N_a - N_j\right) \left(\left|\langle \Psi_a \left| m_{\mathrm{LCP}} \right| \Psi_j \rangle\right|^2 - \left|\langle \Psi_a \left| m_{\mathrm{RCP}} \right| \Psi_j \rangle\right|^2 \right) f\left(E\right) \tag{7.244}$$

Here a and j label members of the initial and state manifold probed in the experiments.

$$N_a\left(B,T\right) = \frac{\exp\left(-E_a/kT\right)}{\sum\limits_i \exp\left(-E_i/kT\right)} \tag{7.245}$$

denotes the Boltzmann population and if the a-th ground state sublevel at energy E_a, $f(E)$ stands for a line shape function, and γ denotes a collection of constants. The electric dipole operators are given by:

$$m_{\text{LCP}} \equiv m_x - im_y \tag{7.246}$$

$$m_{\text{RCP}} \equiv m_x + im_y \tag{7.247}$$

They represent linear combinations of the dipole moment operator:

$$\vec{m} = \sum_N Z_N \vec{R}_N - \sum_i \vec{r}_i \tag{7.248}$$

where N and i denotes summations of nuclei (at positions $\vec{R}_N$ with charges Z_N) and electrons (at positions $\vec{r}_i$) respectively. The calculated transition dipole moment are subjected to the space averaging over the Euler angles which is performed by a simple summation over three angular grids.

$$\left(\frac{\Delta\varepsilon}{E}\right)_{ev} = \frac{1}{8\pi^2} \int_{\psi=0}^{2\pi} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \left(\frac{\Delta\varepsilon}{E}\right) \sin\theta d\theta d\phi d\psi \approx \sum_{\mu\eta\tau} \left(\frac{\Delta\varepsilon}{E}\right)_{\mu\eta\tau} \sin\theta_\tau \tag{7.249}$$

Finally, every transition is approximated by a Gaussian curve with a definite Gaussian shape width parameter. Hence, the final calculated MCD spectrum arises from the superposition of these curves.

As an illustration, consider calculation of a classical example of MCD spectrum of $[Fe(CN)_6]^{3-}$. The mrci block of the input file is presented below.

```
%mrci   ewin -4,10000
        citype mrddci2
        intmode ritrafo
        Tsel  1e-6
        Tpre  1e-5
        etol 1e-8
        rtol 1e-8
        cimode direct2
        maxmemint 300
        solver diis
        davidsonopt 0
        nguessmat 150
        MaxIter 50
        LevelShift 0.5
        PrintLevel 3
        soc
        printlevel 3
        Domcd true        # perform the MCD calculation
        NInitStates 24    # number of SOC and SSC state to account
                          # Starts from the lowest state
        NPointsTheta 10   # number of integration point for
        NPointsPhi 10     # Euler angles
        NPointsPsi 10     #
        B 43500           # experimental magnetic field strength
                          # in Gauss
        Temperature 299.0 # experimental temperature (in K)
        end
        newblock 2 *
         nroots 12
         excitations cisd
         refs cas(23,12) end
         end
        end
```

The parameters B and Temperature can be assigned in pairs, i.e. B = 1000, 2000, 3000..., Temperature = 4, 10, 300.... The program calculates the MCD and absorption spectra for every pair. Now for every point of the integration grid the program prints out the Euler angles, the orientation of the magnetic field in the coordinate system of a molecule, and the energy levels.

```
Psi = 36.000 Phi = 72.000 Theta = 20.000

Bx = 8745.0 By = 12036.5 Bz = 40876.6

Energy levels (cm-1,eV):Boltzmann populations for T = 299.000   K
   0 :         0.000     0.0000         4.53e-01
   1 :         3.943     0.0005         4.45e-01
   2 :       454.228     0.0563         5.09e-02
   3 :       454.745     0.0564         5.08e-02
   4 :      1592.142     0.1974         2.13e-04
   5 :      1595.272     0.1978         2.10e-04
   6 :     25956.363     3.2182         2.59e-55
   7 :     25958.427     3.2184         2.56e-55
   8 :     25985.656     3.2218         2.25e-55
   9 :     25987.277     3.2220         2.23e-55
  10 :     26070.268     3.2323         1.49e-55
  11 :     26071.484     3.2325         1.49e-55
  12 :     31976.645     3.9646         6.78e-68
  13 :     31979.948     3.9650         6.67e-68
  14 :     32018.008     3.9697         5.56e-68
  15 :     32021.074     3.9701         5.48e-68
  16 :     32153.427     3.9865         2.90e-68
  17 :     32157.233     3.9870         2.84e-68
  18 :     42299.325     5.2444         1.81e-89
  19 :     42303.461     5.2450         1.78e-89
  20 :     42346.521     5.2503         1.45e-89
  21 :     42348.023     5.2505         1.44e-89
  22 :     42456.119     5.2639         8.53e-90
  23 :     42456.642     5.2640         8.51e-90
```

In the next lines, ORCA calculates the strength of LCP and RCP transitions and prints the transition energies, the difference between LCP and RCP transitions (denoted as C), and sum of LCP and RCP transitions (denoted as D), and C by D ratio.

```
dE     Na         C           D        C/D

  0 -> 1        3.943 4.53e-01  1.14e-13  8.13e-13   0.00e+00
  0 -> 2      454.228 4.53e-01  5.01e-09  9.90e-09   5.06e-01
  0 -> 3      454.745 4.53e-01 -4.65e-09  7.00e-09  -6.65e-01
  0 -> 4     1592.142 4.53e-01 -8.80e-08  1.02e-07  -8.67e-01
  0 -> 5     1595.272 4.53e-01 -2.29e-08  2.97e-08  -7.71e-01
  0 -> 6    25956.363 4.53e-01  1.22e+01  9.60e+01   1.27e-01
  0 -> 7    25958.427 4.53e-01  3.44e+01  3.52e+01   9.77e-01
  0 -> 8    25985.656 4.53e-01  3.83e+01  1.70e+02   2.25e-01
  0 -> 9    25987.277 4.53e-01 -7.73e+00  6.03e+01  -1.28e-01
  0 ->10    26070.268 4.53e-01 -6.11e+00  2.85e+01  -2.14e-01
  0 ->11    26071.484 4.53e-01  6.17e+00  9.21e+00   6.70e-01
  0 ->12    31976.645 4.53e-01  2.45e+01  6.21e+01   3.95e-01
  0 ->13    31979.948 4.53e-01 -6.58e+01  6.93e+01  -9.50e-01
  0 ->14    32018.008 4.53e-01  3.42e-01  1.07e+02   3.21e-03
  0 ->15    32021.074 4.53e-01 -6.16e+00  3.24e+01  -1.90e-01
  0 ->16    32153.427 4.53e-01 -4.73e+01  1.37e+02  -3.46e-01
  0 ->17    32157.233 4.53e-01 -1.02e+00  5.97e+01  -1.71e-02
  0 ->18    42299.325 4.53e-01  6.47e+00  2.11e+01   3.07e-01
  0 ->19    42303.461 4.53e-01 -2.59e+00  7.61e+00  -3.40e-01
  0 ->20    42346.521 4.53e-01  1.90e+01  8.99e+01   2.11e-01
  0 ->21    42348.023 4.53e-01  3.36e+00  3.55e+00   9.48e-01
  0 ->22    42456.119 4.53e-01  2.52e-01  4.86e-01   5.20e-01
  0 ->23    42456.642 4.53e-01 -2.01e+00  2.91e+00  -6.91e-01
  1 -> 2      450.285 4.45e-01  4.59e-09  6.87e-09   6.69e-01
  1 -> 3      450.802 4.45e-01 -4.96e-09  9.73e-09  -5.09e-01
```

All C and D values are copied additionally into the text files input.1.mcd, input.2.mcd..., for every pair of Temperature and B parameters. These files contain the energies and C and D values for every calculated transition. These files are used by the program orca_mapspc to calculate the spectra lines. The orca_mapspc program generates from the raw transitions data into spectra lines. The main parameters of the orca_mapspc program are described in section 7.18.1. A typical usage of the orca_mapspc program for MCD spectra calculation for the current example may look as the following:

```
orca_mapspc input.1.mcd MCD -x020000 -x150000 -w2000
```

Here the interval for the spectra generation is set from 20000 cm^{-1} to 50000 cm^{-1}, and the line shape parameter is set to 2000 cm^{-1}.

Very often, it is desirable to assign different line width parameters to different peaks of the spectra to obtain a better fitting to experiment. orca_mapspc can read the line shape parameters from a simple text file named as input.1.mcd.inp. This file should contain the energy intervals (in cm^{-1}) and the line shape parameters for this energy interval in the form of:

```
20000 35000 1000
35000 40000 2000
40000 50000 1000
```

This file should not be specified in the executing command; orca_mapspc checks for its presence automatically:

```
orca_mapspc input.1.mcd MCD -x020000 -x150000
Mode is MCD
Number of peaks            ... 276001
Start wavenumber [cm-1]    ...  20000.0
Stop wavenumber [cm-1]     ...  50000.0
Line width parameters are taken from the file:input.1.mcd.inp
Number of points           ... 1024
```

Finally, the orca_mapspc program generates the output text file input.1.mcd.dat which contains seven columns of numbers: transition energies, intensities of MCD transitions (the MCD spectrum), intensities of absorption transitions (the absorption spectrum), the ratio between the MCD and absorption intensities, and the last three columns represent the "sticks" of the corresponding transitions.

```
Energy      C         D         C/D       C         D       E/D
24310.8   0.6673    980.2678   0.0006   0.0000   0.0000   0.0000
24340.1   0.8471   1174.3637   0.0007  -0.0001   0.0129  -0.0112
24369.5   1.0664   1408.5788   0.0007   0.0001   0.0281   0.0033
24398.8   1.3325   1690.5275   0.0007   0.0000   0.0000   0.0000
24428.1   1.6542   2029.0152   0.0008   0.0000   0.0000   0.0000
24457.4   2.0416   2434.1699   0.0008   0.0000   0.0332   0.0003
```

Now the MCD and the absorption spectra can be plotted with a suitable graphical program, for instance with the Origin program.

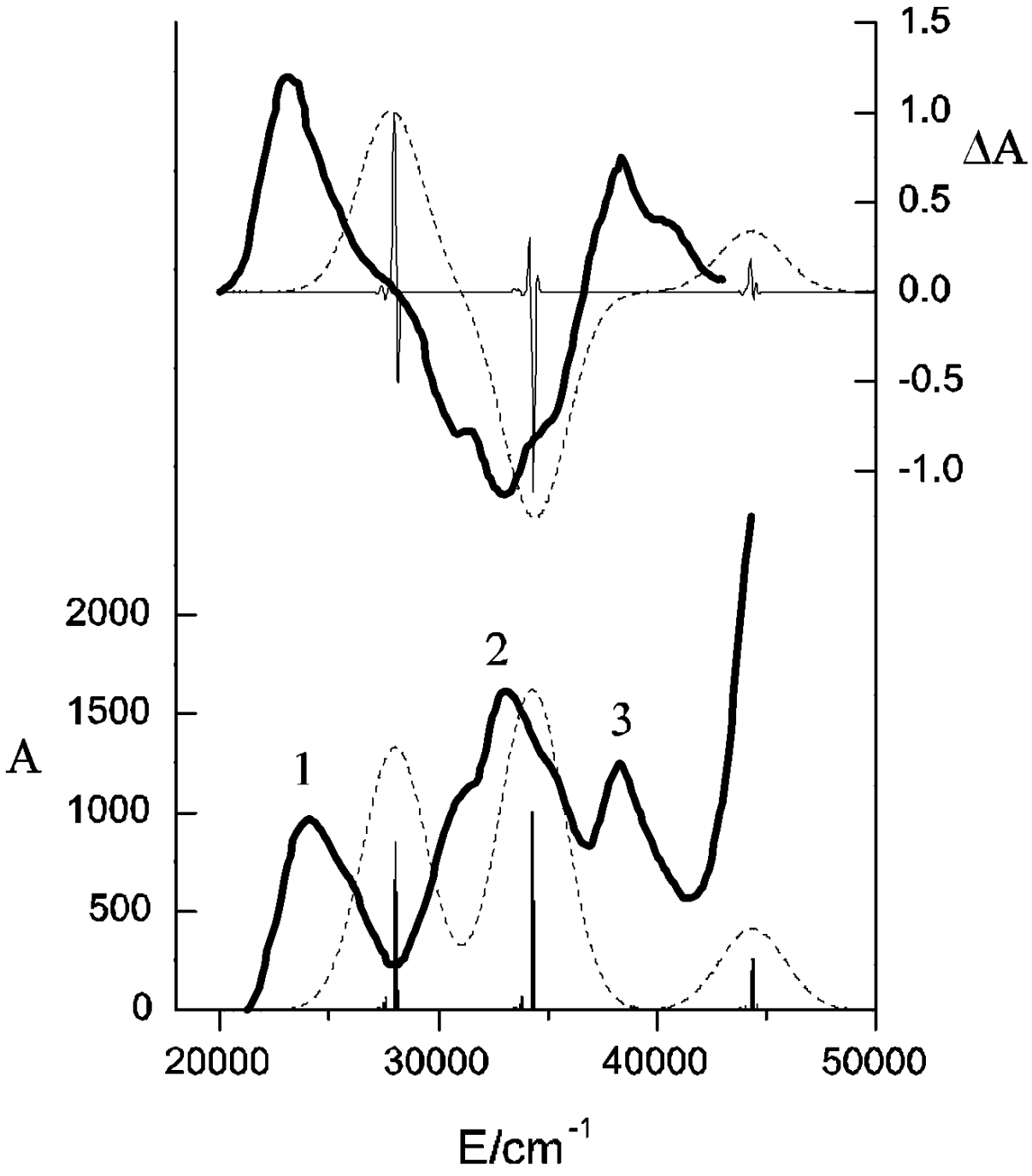


Fig. 7.41: Calculated MCD and absorption spectra of $[Fe(CN)_6]^{3-}$ (dash lines) compared to experimental spectra (solid lines).

Addition of Magnetic Fields

The inclusion of the Zeeman contribution into the QDPT procedure allows to obtain the splittings of the magnetic levels in an external magnetic field. The switch for this calculation and the magnetic field strength are defined in the soc subblock of the mrci block. Optionally the wave function decomposition can be printed for `MagneticField_PrintLevel` larger 0. The latter employs the thresh `TPrint` to omit small contributions from the printing:

```
%mrci
  soc
```

(continues on next page)

(continued from previous page)

```
    DoSOC true          #
    DoSSC true          #
    MagneticField true  # default false
    B 1,10,100,1000     # Strengh of the magnetic field in Gauss.
                        # 4000 is the default value

    # Optional printing of the wave function for each
    # magnetic field settings
    MagneticField_PrintLevel 0 # default (disabled)
    TPrint                   1e-3
  end
end
```

Then, the output contains three sets of data of splittings of the magnetic levels with the magnetic field applied parallel to x, y, and z directions:

```
End B (Gauss)  Energy levels (cm-1) and populations for B || x

     1.0   -0.030   0.333   0.012   0.333   0.018   0.333
    10.0   -0.030   0.333   0.012   0.333   0.018   0.333
   100.0   -0.031   0.333   0.012   0.333   0.020   0.333
  1000.0   -0.102   0.333   0.012   0.333   0.091   0.333

B (Gauss)  Energy levels (cm-1) and populations for B || y

     1.0   -0.030   0.333   0.012   0.333   0.018   0.333
    10.0   -0.030   0.333   0.012   0.333   0.018   0.333
   100.0   -0.032   0.333   0.014   0.333   0.018   0.333
  1000.0   -0.105   0.334   0.018   0.333   0.087   0.333

B (Gauss)  Energy levels (cm-1) and populations for B || z

     1.0   -0.030   0.333   0.012   0.333   0.018   0.333
    10.0   -0.030   0.333   0.011   0.333   0.018   0.333
   100.0   -0.030   0.333   0.005   0.333   0.025   0.333
  1000.0   -0.079   0.333  -0.030   0.333   0.108   0.333
```

Here the number in a row represents the strength of the magnetic field (in Gauss), and the following pairs of numbers denote the energy of the magnetic level (in cm^{-1}) with its occupation number. This table can be readily plotted with any suitable graphical program.

Relativistic Picture Change in Douglas-Kroll-Hess SOC and Zeeman Operators

The DKH correction to the SOC operator is implemented in ORCA as a correction to the one-electron part of the SOMF operator. The DKH transformation is performed up to the second order, and the two-electron part in our implementation is left untransformed. However, the electronic density employed for evaluating the SOMF matrix elements is obtained from a scalar relativistic calculation. The inclusion of the DKH correction is controlled by the picturechange key in the rel block:

```
%rel method DKH              # relativistic method
     picturechange 2         # include the DKH correction to SOC
     end
```

The "picturechange" key can be set to 0, 1, and 2 for no picture change, the first order, and the second order DKH transformations of the SOC operator.

With "picturechange" set to 1 or 2 the DKH correction are applied in the first order to the Zeeman operator. This correction has a visible effect on calculated g-tensors for molecules containing third-row and heavier atoms.

X-ray Spectroscopy

Likewise to the CASCI/NEVPT2 computational protocol presented in section *Core excited states with (C/R)ASCI/NEVPT2* starting from ORCA 4.2 the MRCI module can be used to compute core excited spectra, namely X-ray absorption (XAS) and resonant inelastic scattering (RIXS) spectra.

As discussed in the case of CASCI/NEVPT2 protocol *Core excited states with (C/R)ASCI/NEVPT2* a similar strategy is followed to compute XAS/RIXS spectra within the MRCI module. In principle the XAS/RIXS spectra calculations require two steps:

- In a first step one needs to optimize the valence active space orbitals in the framework of SA-CASSCF calculations, e.g. including valence excited states in the range between 6 to 15 eV.
- In a second step the relevant core orbitals are rotated into the active space and the MRCI problem is solved by saturating the excitation space with singly core-excited electronic configurations using the previously optimized sets of orbitals.
- The core orbitals are also included in the XASMOs definition. The use of this keyword is two fold. At first it effecteively reduces the number of the generated configuration state functions (CSFs) to those that exclusively contain contributions from the defined core orbitals. In the case of RIXS also XES (see below) the specified XASMOs are used to define intermediate or core ionized states.

A representative input for the case of $Fe(Cl)_4$ is provided bellow:

- In the first step one performs a SA-CASSCF calculation for the 5 and 15 quintet and triplet states (FeIICl4.casscf.inp).

```
!CC-PWCVTZ-DK cc-pVTZ/C  RIJCOSX SARC/J TightSCF DKH2

%rel
  FiniteNuc true
end

%basis
  newgto Cl "cc-pVTZ-DK" end
  newauxgto Cl "cc-pVTZ/C" end
end

%method FrozenCore FC_NONE
end

%casscf nel   6
  norb   5
  mult   5,3
  nroots 5,15
  switchstep nr
end

* xyz -2 5
Fe  -17.84299991694815     -0.53096694321123      6.09104775508499
Cl  -19.84288422845700      0.31089495619796      7.04101319789001
Cl  -17.84298666758073      0.11868125024595      3.81067954087770
Cl  -17.84301352218429     -2.87052442818457      6.45826391412877
Cl  -15.84311566482982      0.31091516495189      7.04099559201853
*
```

- In a second step the core orbitals are rotated in the active space and the MRCI problem is solved by saturating the excitation space with all the quintet and triplet states that involve single excitations from the core orbitals (FeIICl4-mrci.inp)

```
!MORead CC-PWCVTZ-DK cc-pVTZ/C  RIJCOSX SARC/J TightSCF DKH2

%moinp "FeIICl4-casscf.gbw"
```

(continues on next page)

(continued from previous page)

```
%rel
  FiniteNuc true
end

%method FrozenCore FC_NONE
end

%scf
  rotate { 6,42,90} { 7,43,90} { 8,44,90}  end
end

%basis
  newgto Cl "cc-pVTZ-DK" end
  newauxgto Cl "cc-pVTZ/C" end
end


%casscf
  nel    12
  norb   8
  mult   5,3
  nroots 34,195
  maxiter 1
  switchstep nr
end


%mrci
  CIType MRCI
  intmode fulltrafo
  XASMOs 42, 43, 44
  newblock 5 *
  nroots 34
  excitations cisd
  refs CAS(12,8)
  end
  end
  newblock 3 *
  nroots 195
  excitations cisd
  refs CAS(12,8)
  end
  end
  maxiter 100
  soc
  printlevel 3
  dosoc true
  end
end


* xyz -2 5
Fe  -17.84299991694815     -0.53096694321123      6.09104775508499
Cl  -19.84288422845700      0.31089495619796      7.04101319789001
Cl  -17.84298666758073      0.11868125024595      3.81067954087770
Cl  -17.84301352218429     -2.87052442818457      6.45826391412877
Cl  -15.84311566482982      0.31091516495189      7.04099559201853
*
```

In a similar fashion Multireference Equation of Motopn Couple Cluster MR-EOM-CC (see next section) can also be used to compute X-ray spectra. Further information can be found in reference[549]

As it is explicitly described in the respective ROCIS section RIXS spectra can be requested by the following keywords:

```
RIXS    true    # Request RIXS calculation (NoSOC)
RIXSSOC true    # Request RIXS calculation (with SOC)
Elastic true    # Request RIXS calculation (Elastic)
```

Please consult section *Resonant Inelastic Scattering Spectroscopy* for processing and analyzing the generated spectra

Likewise to TDDFT (*Use of TD-DFT for the Calculation of X-ray Absorption Spectra*) ROCIS (*General Use*) and CASSCF (*Core excited states with (C/R)ASCI/NEVPT2*) the computed transition densities also in the presence of SOC can be taken beyond the dipole approximation by using the *OPS tool* for details.

1. by performing a multiple expantion up to second order
2. by computing the exact transition moments

The whole set of spectroscopy tables can be requested with the following commands:

```
%mrci
 DoDipoleLength        true
 DoDipoleVelocity      true
 DoHigherMoments       true
 DecomposeFoscLength   true
 DecomposeFoscVelocity true
 DoFullSemiclassical   true
end
```

More details can be found in TDDFT (*Use of TD-DFT for the Calculation of X-ray Absorption Spectra*) ROCIS (*General Use*) and CASSCF (*Core excited states with (C/R)ASCI/NEVPT2*) sections.

Starting from ORCA 4.2 the previously reported RASCI-XES protocol reference[690], which can compute K_β Mainline XES spectra, can be processed entirely within the ORCA modules. In ORCA 5.0 a similar protocol (CASCI-XES) exist in the CASSCF module (*Core excited states with (C/R)ASCI/NEVPT2*)

- Like above or in the CASCI/NEVPT2 case in a first step one needs to optimize the valence active space orbitals in the framework of SA-CASSCF calculations, e.g. including valence excited states in the range between 6 to 15 eV for the N electron system.
- In a second step the metal 1s and 3p orbitals are rotated in the active space and the 1s MO is defined in the XASMOs list
- Computes the XES spectrum in the RASCI framework for the N-1 electron system in the presence of SOC if the XESSOC keyword for all the states that are dominated by 3p-1s electron decays.

A representative input sequence for the case of $Fe(Cl)_6$ is provided bellow:

As described in reference[690] at first for a CAS(5,5) the excitation space is saturated by the sextet as well as the 24 quartet and the 75 doublet states which are optimized in the SA-CASSCF fashion.

```
!ZORA  def2-TZVP  def2-TZVP/C

%cpcm
  epsilon 80
  refrac 1.33
  surfacetype gepol_ses
end

%scf
  MaxDisk 40000
end

%casscf
  nel 5
```

(continues on next page)

(continued from previous page)

```
  norb 5
  mult 6,4,2
  nroots 1,24,75
  shiftup 0.5
  shiftdn 0.5
  trafostep RI
  maxiter 150
end

*xyzfile -3  6
Fe      0.0000       0.0000     0.000000
Cl      2.478        0.0000     0.000
Cl     -2.478        0.0000     0.000
Cl      0.000005     2.478      0.00000
Cl      0.000005    -2.478     -0.0000
Cl     -0.000       -0.000      2.478
Cl      0.000       -0.0000    -2.478
*
```

In following the 1s and 3p Fe based MOs are rotated in the active space and the XES spectra are computed for the $[Fe(Cl)_6]^+$ system for the 4 septet and 81 quintet states.

```
! ZORA  def2-TZVP  def2-TZVP/C noiter moread AllowRHF

%moinp "fecl6_casscf.gbw"

%cpcm
  epsilon 80
  refrac 1.33
  surfacetype gepol_ses
end

%scf
  MaxDisk 40000
end

%scf
  rotate {0,59,90} {36, 60, 90} {37,61,90} {38,62,90}  end
end

%mrci citype mrci
  UseIVOs false
  Etol 1e-5
  newblock 5 *
  excitations none
  nroots 81
  refs ras(12:4 1/5/ 0 0) end
  end
  newblock 7 *
  excitations none
  nroots 4
  refs ras(12:4 1/5/ 0 0) end
  end
  XASMOs 59
  soc
  dosoc true
  XESSOC true
  end
end

*xyzfile -2  7
```

(continues on next page)

(continued from previous page)

```
Fe      0.0000       0.0000     0.000000
Cl      2.478        0.0000     0.000
Cl     -2.478        0.0000     0.000
Cl      0.000005     2.478      0.00000
Cl      0.000005    -2.478     -0.0000
Cl     -0.000       -0.000      2.478
Cl      0.000       -0.0000    -2.478
*
```

As a result the X-ray emission spectrum is calculated and the intensities are computed on the basis of the transition electric dipole moments

```
Printing the XES spectrum ...

        ------------------------------------------------------------------------------
        SPIN-ORBIT X-RAY EMISSION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
        ------------------------------------------------------------------------------
        Transition          Energy           INT             TX        TY        TZ
        1  421 ->    5      7228.632     0.000000000000     0.00000   0.00000   0.00000
        2  422 ->    5      7228.632     0.000000000000     0.00000   0.00000   0.00000
        3  423 ->    5      7228.632     0.000000000000     0.00000   0.00000   0.00000
        4  424 ->    5      7228.632     0.000000000000     0.00000   0.00000   0.00000
        5  425 ->    5      7228.632     0.000000000000     0.00000   0.00000   0.00000
        ...
        242  422 ->   25      7177.286     0.000917305388     0.00025   0.00171   0.00149
        243  423 ->   25      7177.286     0.002043577370     0.00197   0.00211   0.00181
        244  424 ->   25      7177.286     0.000789769987     0.00114   0.00133   0.00119
        245  425 ->   25      7177.286     0.000026130790     0.00018   0.00034   0.00002
        246  426 ->   25      7177.286     0.000035191741     0.00034   0.00028   0.00003
        247  427 ->   25      7177.286     0.005143175830     0.00294   0.00345   0.00296
        248  428 ->   25      7177.341     0.000000000000     0.00000   0.00000   0.00000
        249  429 ->   25      7177.341     0.000000000001     0.00000   0.00000   0.00000
        250  430 ->   25      7177.341     0.000000000001     0.00000   0.00000   0.00000
        251  431 ->   25      7177.341     0.000000000000     0.00000   0.00000   0.00000
        252  432 ->   25      7177.341     0.000000000000     0.00000   0.00000   0.00000
        ...
        4991  431 ->  420      7153.111     0.000195885011     0.00106   0.00000   0.00000
        4992  432 ->  420      7153.111     0.002719228427     0.00256   0.00299   0.00002

        All Done
        ------------------------------------------------------------------------------
```

The resulted XES spectrum can be visualized by processing the above output file with the `orca_mapspc`

```
orca_mapspc fecl6_xes.out XESSOC -x07140 -x17190 -w4.0 -eV -n10000
```

This will result in Fig. 7.42.

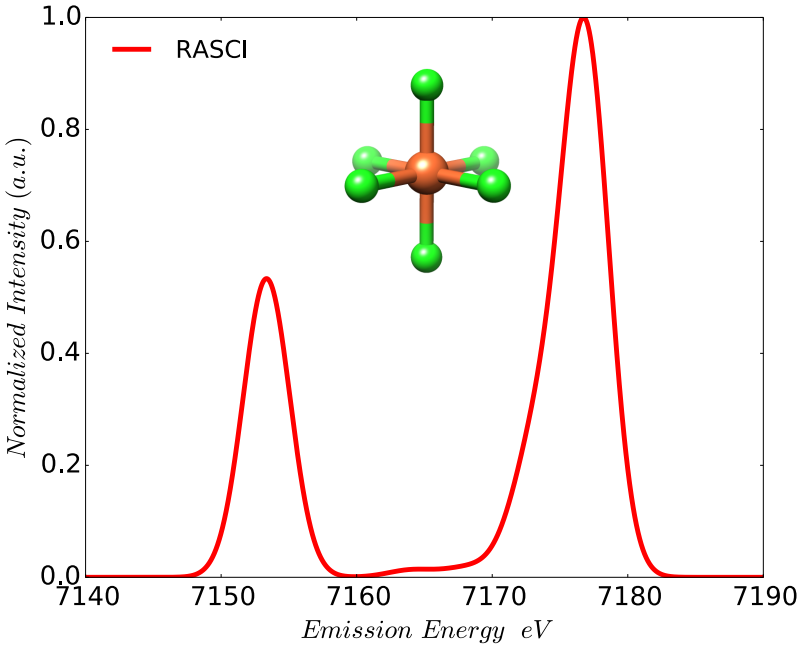


Fig. 7.42: Calculated RASCI K_β XES spectrum of $[Fe(Cl)_6]^+$.

7.40 Multireference Equation of Motion Coupled-Cluster (MR-EOM-CC) Theory

In analogy with single reference EOM-CC (see sections *Excited States with EOM-CCSD* and *Excited States via EOM-CCSD*) and STEOM-CC (see sections *Excited States with STEOM-CCSD* and *Excited States via STEOM-CCSD*), Multireference Equation of Motion Coupled-Cluster (MR-EOM-CC) theory [193, 194, 203, 406, 407, 639] can be viewed a transform and diagonalize approach to molecular electronic structure theory. An MR-EOM calculation involves a single state-averaged CASSCF calculation, incorporating a few low-lying states and the solution of a single set of cluster amplitudes, which define a sequence of similarity-transformed Hamiltonians. The ultimate goal of these many-body transformations is to effectively decouple the CAS configurations from important excited configurations (e.g., 2p2h, 2p1h, 1p1h, etc.) which comprise the first-order interacting space. Through the definition of suitable cluster operators, in each of the transformations, most of these excitations can be included in an internally contracted fashion. Hence, the resulting final transformed Hamiltonian can be diagonalized over a small subspace of the original first-order interacting space to gain access to many electronic states. As discussed in section *MR-EOM-CC: Multireference Equation of Motion Coupled-Cluster*, the MR-EOM implementation in ORCA therefore makes use of the CASSCF module (to obtain the state-averaged CASSCF reference), the MDCI module for the solution of the amplitude equations and the calculation of the elements of the various similarity transformed Hamiltonians and the MRCI module for the diagonalization of the final transformed Hamiltonian. Some desirable features of this methodology are:

- Many states can be obtained through the diagonalization of a similarity transformed Hamiltonian over a compact diagonalization manifold (e.g. the final diagonalization space in MR-EOM-T|$T^\dagger$|SXD|U only includes the CAS configurations and 1h and 1p configurations).
- Only a single state-averaged CASSCF calculation and the solution of a single set of amplitudes is required to define the final similarity transformed Hamiltonian and the results are typically quite insensitive to the precise definition of the CAS (only a few low-lying multiplets need to be included in the state-averaging)
- The MR-EOM approach is rigorously invariant to rotations of the orbitals in the inactive, active and virtual subspaces, and it preserves both spin and spatial symmetry.

Table 7.24: The details of the various MR-EOM transformations that are considered in the ORCA implementation of MR-EOM. The equations for the operator components and the residual equations which determine the corresponding amplitudes also appear in the Table. Note that we use the usual (Einstein) convention that repeated indices are summed over.

Name	Transformation	Operators	Operator Components	Residual Equation
T	$\widehat{\overline{H}} = e^{-\widehat{T}}\widehat{H}e^{\widehat{T}}$	$\widehat{T} = \widehat{T}_1 + \widehat{T}_2$	$\widehat{T}_1 = t_i^a \widehat{E}_i^a$	$R_i^a = \sum_{\mathfrak{m}} w_{\mathfrak{m}} \langle \Phi_{\mathfrak{m}} \vert \widehat{E}_a^i \widehat{\overline{H}} \vert \Phi_{\mathfrak{m}} \rangle$
	$= \overline{h}_0 + \overline{h}_q^p \{\widehat{E}_q^p\} + \overline{h}_{rs}^{pq} \{\widehat{E}_{rs}^{pq}\} + \ldots$		$\widehat{T}_2 = \frac{1}{2} t_{ij}^{ab} \widehat{E}_{ij}^{ab}$	$R_{ij}^{ab} = \overline{h}_{ij}^{ab}$
$T^\dagger$	$\widehat{\overline{H}}_2 e^{-\widehat{T}^\dagger}$	$\widehat{T}^\dagger = \widehat{T}_1^\dagger + \widehat{T}_2^\dagger$	$\widehat{T}_1^\dagger = t_a^i \widehat{E}_a^i$	None (*i.e.* set $t_a^i \approx t_i^a$)
	$= \widetilde{h}_0 + \widetilde{h}_q^p \{\widehat{E}_q^p\} + \widetilde{h}_{rs}^{pq} \{\widehat{E}_{rs}^{pq}\} + \ldots$		$\widehat{T}_2^\dagger = \frac{1}{2} t_{ab}^{ij} \widehat{E}_{ab}^{ij}$	None (*i.e.* set $t_{ab}^{ij} \approx t_{ij}^{ab}$)
SXD	$\widehat{\overline{F}} = \{e^{\widehat{S}+\widehat{X}+\widehat{D}}\}^{-1} \widehat{\mathcal{H}}_2 \{e^{\widehat{S}+\widehat{X}+\widehat{D}}\}$	$\widehat{S} = \widehat{S}_2$	$\widehat{S}_2 = s_{i'j'}^{aw} \widehat{E}_{i'j'}^{aw}$	$R_{i'j'}^{aw} = \overline{f}_{i'j'}^{aw}$
	$= \overline{f}_0 + \overline{f}_q^p \{\widehat{E}_q^p\} + \overline{f}_{rs}^{pq} \{\widehat{E}_{rs}^{pq}\} + \ldots$	$\widehat{X} = \widehat{X}_2$	$\widehat{X}_2 = x_{i'x}^{aw} \widehat{E}_{i'x}^{aw}$	$R_{i'x}^{aw} = \overline{f}_{,i'x}^{aw}$
		$\widehat{D} = \widehat{D}_2$	$\widehat{D}_2 = d_{xi'}^{aw} \widehat{E}_{xi'}^{aw}$	$R_{xi'}^{aw} = \overline{f}_{xi'}^{aw}$
U	$\widehat{G} = e^{-\widehat{U}} \widehat{\overline{F}}_2 e^{\widehat{U}}$	$\widehat{U} = \widehat{U}_2$	$\widehat{U}_2 = u_{i'j'}^{wx} \widehat{E}_{i'j'}^{wx}$	$R_{i'j'}^{wx} = g_{i'j'}^{wx}$
	$= g_0 + g_q^p \{\widehat{E}_q^p\} + g_{rs}^{pq} \{\widehat{E}_{rs}^{pq}\} + \ldots$			

As the details concerning the MR-EOM methodology are rather involved, we refer the interested reader to Refs. [193, 194, 203, 406, 407, 639] for a more detailed discussion. Note that the details concerning the implementation of MR-EOM in ORCA can be found in Refs. [406] and [407]. In the following discussion, we note that general spatial orbitals p, q, r, s, which comprise the molecular orbital basis, are partitioned into (doubly occupied) inactive core orbitals i', j', k', l', occupied orbitals i, j, k, l (i.e. the union of the inactive core and active orbital subspaces), active orbitals w, x, y, z and virtual orbitals a, b, c, d. In general, the many-body similarity transformations assume the general form

$$\widehat{G} = \{e^{\widehat{Y}}\}^{-1} \widehat{\overline{H}}_2 \{e^{\widehat{Y}}\} = g_0 + \sum_{p,q} g_q^p \{\widehat{E}_q^p\} + \sum_{p,q,r,s} g_{rs}^{pq} \{\widehat{E}_{rs}^{pq}\} + \ldots,$$

in which $\widehat{Y}$ is a cluster operator and $\widehat{\overline{H}}_2$ is the bare Hamiltonian or a similarity transformed Hamiltonian truncated up to two-body operators. The braces indicate Kutzelnigg-Mukherjee normal ordering [485, 596], which is used extensively in the definition of the MR-EOM formalism. The various transformations which need to be considered in the ORCA implementation of MR-EOM are summarized in Table Table 7.24. The table also includes the expressions for the operator components of the various internally contracted cluster operators and the residual equations that must be solved for the various amplitudes. Note that the residual equations are typically of the many-body type (i.e. obtained by setting the corresponding elements of the similarity transformed Hamiltonian to zero). The only exception is the residual equation which defines the t_i^a amplitudes, which is a projected equation of the form

$$R_i^a = \sum_{\mathfrak{m}} w_{\mathfrak{m}} \langle \Phi_{\mathfrak{m}} | \widehat{E}_a^i \widehat{H} | \Phi_{\mathfrak{m}} \rangle.$$

Here, $|\Phi_{\mathfrak{m}}\rangle$ is the $\mathfrak{m}^{\text{th}}$ state included in the state averaged CAS, with weight $w_{\mathfrak{m}}$. The reason the equation for the singles is of the projected form is that it satisfies the Brillouin theorem (i.e. the first order singles vanish for all i and a), whereas the corresponding many-body equation ($\bar{h}_i^a = 0$) does not.

Table 7.25: Details of the three MR-EOM approaches implemented in ORCA

Method	Input Keyword	Operators	Diagonalization Manifold
MR-EOM-T|T†-h-v	`MR-EOM-T|Td`	$\widehat{T}$; $\widehat{T}^\dagger$	CAS, 2h1p, 1h1p, 2h, 1h, 1p
MR-EOM-T|T†|SXD-h-v	`MR-EOM-T|Td|SXD`	$\widehat{T}$; $\widehat{T}^\dagger$; $\widehat{S}+\widehat{X}+\widehat{D}$	CAS, 2h, 1h, 1p
MR-EOM-T|T†|SXD|U-h-v	`MR-EOM`	$\widehat{T}$; $\widehat{T}^\dagger$; $\widehat{S}+\widehat{X}+\widehat{D}$; $\widehat{U}$	CAS, 1h, 1p

Note that there are three different MR-EOM approaches which have been implemented in ORCA. Namely, the current implementation allows for MR-EOM-T|T†-h-v, MR-EOM-T|T†|SXD-h-v and MR-EOM-T|T†|SXD|U-h-v calculations. At this point it is useful to discuss the naming convention used for these approaches. We use a vertical line to separate each transformation involved in the sequence of transformations defining the given MR-EOM approach. For example, T|T†|SXD indicates that a T transformation, is followed by a T† transformation, which is then followed by an SXD transformation. The h-v indicates that the elements of the transformed Hamiltonian have been hermitized (h) and vertex symmetrized (v) before entering the MRCI diagonalization (see Ref. [407] for more information). Essentially, this means that the full eightfold symmetry of the two-electron integrals (and hermiticity of the one-body elements) have been enforced upon the elements of the transformed Hamiltonian. The details of the three MR-EOM approaches are summarized in Table Table 7.25. This table includes the keyword (in the first line of input) used to initiate the calculation in ORCA, the various operators involved, and the configurations included in the final diagonalization manifold. One can clearly see that the MR-EOM-T|T†|SXD|U-h-v approach is the most cost effective, as it only includes the 1h and 1p configurations, beyond the CAS, in the final diagonalization manifold.

The various `%mdci` keywords, which are important for controlling MR-EOM calculations are (i.e. default values are given here):

```
%mdci
  STol              1e-7        #Convergence Tolerance on Residual Equations
  MaxIter           100         #Maximum Number of Iterations
  DoSingularPT      false       #Activate the Singular PT/Projection Procedure
  SingularPTThresh  0.01        #Threshold for the Singular PT/Projection
                                #Procedure
  PrintOrbSelect    false       #Print the Eigenvalues of the Orbital Selection
                                #Densities (and R_core and R_virt values)
                                #and Terminate the Calculation
  CoreThresh        0.0         #Core Orbital Selection Threshold
  VirtualThresh     1.0         #Virtual Orbital Selection Threshold
end
```

As discussed below, the orbital selection scheme is activated by adding `!OrbitalSelection` to the simple input line. Keywords that are specific to the CASSCF and MRCI modules are discussed in sections *The Complete Active Space Self-Consistent Field (CASSCF) Module* and *The Multireference Correlation Module*, respectively. We note that in MR-EOM-T|T†-h-v and MR-EOM-T|T†|SXD-h-v calculations, it is possible to override the default excitation classes in the final MRCI diagonalization. This is done by specifying `excitations none` and then explicitly setting the excitation flags within a given multiplicity block. For example, if we wanted to have 1h, 1p, 1h1p, 2h and 2h1p excitations in the final diagonalization manifold, we would specify (i.e. here we have requested 6 singlets and have a CAS(6,4) reference):

```
%mrci
    newblock 1 *
        excitations none
        Flags[is]   1
        Flags[sa]   1
        Flags[ia]   1
        Flags[ijss] 1
        Flags[ijsa] 1
        nroots      6
        refs
            cas(6,4)
        end
    end
end
```

7.40.1 The Steps Required to Run an MR-EOM Calculation

To illustrate the various steps required in a typical MR-EOM calculation, we will consider the calculation of the excitation energies of the neutral Fe atom at the MR-EOM-T|T†|SXD|U-h-v level of theory.

State-Averaged CASSCF Calculation

Evidently, the first step is to determine a suitable state-averaged CASSCF reference for the subsequent MR-EOM calculation. In choosing the state-averaged CAS for an MR-EOM calculation, we typically include a few of the low-lying multiplets that have the same character as the (much larger number of) states that we wish to compute in the final MR-EOM calculation. For the neutral Fe atom, we typically have electronic states which have either $4s^2 3d^6$ character or $4s^1 3d^7$ character. From the NIST atomic spectra database [1, 605], we find that the lowest lying a^5D multiplet is of $4s^2 3d^6$ character and the higher lying a^5F multiplet is of $4s^1 3d^7$ character. Hence, we can set up a state-averaged CASSCF(8,6) calculation (i.e. 8 electrons in 6 orbitals (4s and 3d)) which includes the ^5D and ^5F states and choose the weights such that the average occupation of the 4s orbital is 1.5. As discussed in Ref. [529], this is done to avoid a preference toward either of the 4s configurations in the state-averaging. We will run the state-averaged CASSCF calculation, making use of the second order DKH (see *The Douglas-Kroll-Hess Method*) method for the inclusion of relativistic effects in a Def2-TZVPP basis (i.e. the DKH-Def2-TZVPP relativistically recontracted basis, listed in section *Choice of Basis Set*). The input file for the state-averaged CASSCF(8,6) calculation takes the form:

```
!CASSCF DKH-Def2-TZVPP VeryTightSCF DKH

%casscf
  nel 8
  norb 6
  mult 5
  nroots 12
  weights[0] = 0.7, 0.7, 0.7, 0.7, 0.7, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5
end

* xyz 0 5
  Fe 0.000000 0.000000 0.000000
end
```

Here, we have requested 12 quintet states (the lowest lying ^5D and ^5F multiplets) and we have chosen the weights to be 0.7 for the five ^5D states and 0.5 for the seven ^5F states, such that the overall occupation of the 4s orbital will be 1.5. Once the calculation has converged, it is important to inspect the results printed in the final macro-iteration of the CASSCF calculation (macro-iteration 8 in this case). In this case, we have:

```
MACRO-ITERATION   8:
   --- Inactive Energy E0 = -1249.82392764 Eh
   --- All densities will be recomputed
CI-ITERATION   0:
 -1271.258899450   0.000000000001 (    0.00)
 -1271.258899450   0.000000000001
 -1271.258899450   0.000000000001
 -1271.258899450   0.000000000001
 -1271.258899450   0.000000000001
 -1271.186288591   0.000000000001
 -1271.186288591   0.000000000001
 -1271.186288591   0.000000000002
 -1271.186288591   0.000000000001
 -1271.186288591   0.000000000001
 -1271.186288591   0.000000000001
 -1271.186288591   0.000000000001
   CI-PROBLEM SOLVED
   DENSITIES MADE
   E(CAS)= -1271.222594021 Eh DE=   -1.591616e-12
   --- Energy gap subspaces: Ext-Act = 0.276   Act-Int = 2.469
```

(continues on next page)

(continued from previous page)

```
   N(occ)=  1.50000 1.30000 1.30000 1.30000 1.30000 1.30000
   ||g|| =     4.669702e-05 Max(G)=    1.104493e-05 Rot=68,1
```

Directly below `CI-ITERATION 0`, the final CAS-CI energies are printed, and one observes that they follow the correct degeneracy pattern (i.e. 5 states with energy `-1271.258899450` and 7 states with energy `-1271.186288591`). Furthermore, the final state-averaged CASSCF energy (`E(CAS)= -1271.222594021`) and occupation numbers (`N(occ)= 1.50000 1.30000 1.30000 1.30000 1.30000 1.30000`) are also printed. As expected, the occupation number of the 4s orbital is indeed 1.5, while the 3d orbitals each have an occupation of 1.3.

Selection of the States to Include in the MR-EOM Calculation

Once a satisfactory CASSCF reference has been obtained, the next step is to determine the number of states to include in the MR-EOM calculation. From the NIST atomic spectra database, one finds that the higher lying states of $4s^2 3d^6$ and $4s^1 3d^7$ character are either singlets, triplets, or quintets. To figure out how many states should be included in each multiplicity block, one can perform an inexpensive CAS-CI calculation. This is done by reading in the orbitals from the previous CASSCF calculation (here they are stored in the file `CAS.gbw`) and requesting a single iteration (i.e. using the `NoIter` keyword) of a state-averaged CASSCF calculation:

```
!CASSCF DKH-Def2-TZVPP ExtremeSCF DKH NoIter

!MOREAD
%moinp "CAS.gbw"

%casscf
  nel 8
  norb 6
  mult 5,3,1
  nroots 15,90,55
end

* xyz 0 5
  Fe 0.000000 0.000000 0.000000
end
```

Here, after some experimentation, we have chosen 15 quintets, 90 triplets and 55 singlets. It is important that we calculate states up to sufficiently high energy (i.e. all the states that are of interest) and it is imperative that we have complete multiplets. Hence, several iterations of this procedure might be required to choose the proper number of states for each multiplet. The relevant section of the output file which should be analyzed is the `SA-CASSCF TRANSITION ENERGIES`. For the above calculation, we obtain (i.e. only the CAS-CI energies for the first 33 roots are shown here):

```
------------------------------
SA-CASSCF TRANSITION ENERGIES
------------------------------

LOWEST ROOT (ROOT 0 ,MULT 5) =  -1271.258899450 Eh -34592.713 eV

STATE   ROOT MULT  DE/a.u.    DE/eV    DE/cm**-1
   1:    1    5   0.000000    0.000       0.0
   2:    2    5   0.000000    0.000       0.0
   3:    3    5   0.000000    0.000       0.0
   4:    4    5   0.000000    0.000       0.0
   5:    5    5   0.072611    1.976   15936.2
   6:    6    5   0.072611    1.976   15936.2
   7:    7    5   0.072611    1.976   15936.2
   8:    8    5   0.072611    1.976   15936.2
   9:    9    5   0.072611    1.976   15936.2
  10:   10    5   0.072611    1.976   15936.2
  11:   11    5   0.072611    1.976   15936.2
```

(continues on next page)

(continued from previous page)

```
  12:    0    3   0.092859     2.527  20380.2
  13:    1    3   0.092859     2.527  20380.2
  14:    2    3   0.092859     2.527  20380.2
  15:    3    3   0.092859     2.527  20380.2
  16:    4    3   0.092859     2.527  20380.2
  17:    5    3   0.092859     2.527  20380.2
  18:    6    3   0.092859     2.527  20380.2
  19:    7    3   0.092859     2.527  20380.2
  20:    8    3   0.092859     2.527  20380.2
  21:    9    3   0.092859     2.527  20380.2
  22:   10    3   0.092859     2.527  20380.2
  23:   11    3   0.101848     2.771  22353.1
  24:   12    3   0.101848     2.771  22353.1
  25:   13    3   0.101848     2.771  22353.1
  26:   14    3   0.101848     2.771  22353.1
  27:   15    3   0.101848     2.771  22353.1
  28:   16    3   0.101848     2.771  22353.1
  29:   17    3   0.101848     2.771  22353.1
  30:   18    3   0.102559     2.791  22509.1
  31:   19    3   0.102559     2.791  22509.1
  32:   20    3   0.102559     2.791  22509.1
```

Running the MR-EOM Calculation

Now that we have chosen a suitable CASSCF reference and the states that we wish to calculate, we can finally proceed with the MR-EOM calculation. The following input file runs an MR-EOM-T|T†|SXD|U-h-v calculation for 15 quintet, 90 triplet and 55 singlet states of the neutral Fe atom (i.e. the CASSCF orbitals are read from `CAS.gbw`):

```
!MR-EOM DKH-Def2-TZVPP VeryTightSCF DKH

!MOREAD
%moinp "CAS.gbw"

%method frozencore fc_ewin end

%casscf
  nel 8
  norb 6
  mult 5
  nroots 12
  weights[0] = 0.7, 0.7, 0.7, 0.7, 0.7, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5
end

%mdci
  ewin -6, 10000
  STol 1e-7
end

%mrci
  ewin -6, 10000
  MaxIter 200
  newblock 5 *
    nroots 15
    refs cas(8,6) end
  end
  newblock 3 *
    nroots 90
    refs cas(8,6) end
```

(continues on next page)

(continued from previous page)

```
  end
  newblock 1 *
    nroots 55
    refs cas(8,6) end
  end
end

* xyz 0 5
  Fe 0.000000 0.000000 0.000000
end
```

Note that since the default frozen core settings exclude the 3p orbitals from the correlation treatment, we have used an energy window (i.e. the line `ewin -6, 10000` in both the `%mdci` and `%mrci` blocks) such that they are included in the current calculation. We note that a detailed discussion of the input and output of an MR-EOM calculation has already been given in section *MR-EOM-CC: Multireference Equation of Motion Coupled-Cluster* and thus, we do not repeat it here. It is important to reiterate that one should always inspect the values of the largest (T, S and U) amplitudes. Ideally, the largest amplitudes should be smaller than 0.1 and should not exceed 0.15. If some amplitudes are larger than 0.15, it might be necessary to revisit the definition of the CAS and the weights used. For the T amplitudes, an alternative solution is to use the projection/singular PT scheme discussed in section *A Projection/Singular PT Scheme to Overcome Convergence Issues in the T Amplitude Iterations* below.

As discussed in section *MR-EOM-CC: Multireference Equation of Motion Coupled-Cluster*, the excitation energies are printed under the heading `TRANSITION ENERGIES`. For the current calculation, we obtain the following results (only the results for 33 states are shown here):

```
-------------------
TRANSITION ENERGIES
-------------------

The lowest energy is  -1271.833871822 Eh

State Mult Irrep Root Block    mEh          eV      1/cm
  0    5    -1    0    0      0.000      0.000       0.0
  1    5    -1    1    0      0.000      0.000       0.0
  2    5    -1    2    0      0.000      0.000       0.0
  3    5    -1    3    0      0.000      0.000       0.0
  4    5    -1    4    0      0.000      0.000       0.0
  5    5    -1    5    0     33.901      0.922    7440.4
  6    5    -1    6    0     33.901      0.922    7440.4
  7    5    -1    7    0     33.901      0.922    7440.4
  8    5    -1    8    0     33.901      0.922    7440.4
  9    5    -1    9    0     33.901      0.922    7440.4
 10    5    -1   10    0     33.901      0.922    7440.4
 11    5    -1   11    0     33.901      0.922    7440.4
 12    3    -1    0    1     54.743      1.490   12014.8
 13    3    -1    1    1     54.743      1.490   12014.8
 14    3    -1    2    1     54.743      1.490   12014.8
 15    3    -1    3    1     54.743      1.490   12014.8
 16    3    -1    4    1     54.743      1.490   12014.8
 17    3    -1    5    1     54.743      1.490   12014.8
 18    3    -1    6    1     54.743      1.490   12014.8
 19    5    -1   12    0     78.790      2.144   17292.5
 20    5    -1   13    0     78.790      2.144   17292.5
 21    5    -1   14    0     78.790      2.144   17292.5
 22    3    -1    7    1     95.413      2.596   20940.8
 23    3    -1    8    1     95.413      2.596   20940.8
 24    3    -1    9    1     95.413      2.596   20940.8
 25    3    -1   10    1     95.413      2.596   20940.8
 26    3    -1   11    1     95.413      2.596   20940.8
 27    3    -1   12    1     95.413      2.596   20940.8
 28    3    -1   13    1     95.413      2.596   20940.8
```

(continues on next page)

(continued from previous page)

```
 29    3    -1   14    1     95.413      2.596   20940.8
 30    3    -1   15    1     95.413      2.596   20940.8
 31    3    -1   16    1     95.413      2.596   20940.8
 32    3    -1   17    1     95.413      2.596   20940.8
```

It is also important to recall that one should always inspect the reference weights for each state, as only states which are dominated by reference space configurations can be treated accurately at the MR-EOM level of theory. Generally, the reference weights should be larger than (or close to) 0.9. In each multiplicity block, the individual state energies and reference weights can be found following convergence of the MRCI procedure, under the heading `CI-RESULTS` (see section *MR-EOM-CC: Multireference Equation of Motion Coupled-Cluster* for a more detailed discussion).

7.40.2 Approximate Inclusion of Spin-Orbit Coupling Effects in MR-EOM Calculations

The effects of spin-orbit coupling can approximately be included in MR-EOM calculations using the SOC submodule of the MRCI module, as outlined in section *Properties Calculation Using the SOC Submodule*. This can be viewed as a first order approximation to the inclusion of spin-orbit coupling effects in MR-EOM. In a more rigorous formulation, one would have to consider the various similarity transformations of the spin-orbit coupling operator. The details of the SOC submodule of the MRCI module have already been discussed in detail in Sec. *Properties Calculation Using the SOC Submodule* and its usage within the MR-EOM formalism is identical to that discussed therein. Let us consider the calculation of spin-orbit coupling effects in the excitation spectrum of the neutral Fe atom considered in the previous section. The input file for this calculation is:

```
!MR-EOM DKH-Def2-TZVPP ExtremeSCF DKH

%method frozencore fc_ewin end

%casscf
  nel 8
  norb 6
  mult 5
  nroots 12
  weights[0] = 0.7, 0.7, 0.7, 0.7, 0.7, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5
  etol 1e-12
  gtol 1e-12
end

%mdci
  ewin -6, 10000
  MaxIter 300
  STol 1e-12
end

%mrci
  ewin -6, 10000
  MaxIter 200
  newblock 5 *
    nroots 15
    refs cas(8,6) end
  end
  newblock 3 *
    nroots 90
    refs cas(8,6) end
  end
  newblock 1 *
    nroots 55
    refs cas(8,6) end
  end
```

(continues on next page)

(continued from previous page)

```
  soc
    dosoc true  #include spin-orbit coupling effects
  end
end

* xyz 0 5
  Fe 0.000000 0.000000 0.000000
end
```

In contrast with the calculation performed in section *The Steps Required to Run an MR-EOM Calculation*, the convergence thresholds have been tightened in all aspects of the calculation (i.e. the use of the `ExtremeSCF` keyword, `etol` and `gtol` (CASSCF energy and orbital gradient convergence tolerance) are set to 1×10^{-12} and the convergence tolerance for the residuals in the MR-EOM amplitude iterations have been set to 1×10^{-12}). We note that with the use of the `ExtremeSCF` keyword, the convergence tolerance on the energy (`Etol`) and residual (`Rtol`) in the MRCI portion of the calculation are also set to 1×10^{-12}. Although it is not absolutely necessary, we have used very strict convergence thresholds to preserve the degeneracies of the various multiplets as much as possible. The output of spin-orbit corrected MR-EOM spectrum appears under the heading `SOC CORRECTED ABSORPTION SPECTRUM VIA TRANSITION DIPOLE MOMENTS`:

```
-------------------------------------------------------------------------------------------------
↪---------
      SOC CORRECTED ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
↪---------
      Transition         Energy     Energy  Wavelength fosc(D2)      D2        |DX|      |DY|  ␣
↪   |DZ|
                          (eV)      (cm-1)    (nm)  (*population)  (au**2)    (au)      (au)  ␣
↪   (au)
-------------------------------------------------------------------------------------------------
↪---------
```

It is possible to obtain more accurate results by performing an MR-EOM-T|T†|SXD-h-v calculation and including the 1h1p excitations. It is important to note that these calculations are significantly more expensive. As discussed above, to run an MR-EOM-T|T†|SXD-h-v calculation, the keyword `MR-EOM-T|Td|SXD` must appear in the first line of input and, in order to activate the 1h1p excitations in each multiplicity block of the MRCI calculation, the `%mrci` block takes the form:

```
%mrci
  ewin -6, 10000
  MaxIter 200
  newblock 5 *
    nroots 15
    excitations none
    Flags[is] 1
    Flags[sa] 1
    Flags[ia] 1
    Flags[ijss] 1
    refs cas(8,6) end
  end
  newblock 3 *
    nroots 90
    excitations none
    Flags[is] 1
    Flags[sa] 1
    Flags[ia] 1
    Flags[ijss] 1
    refs cas(8,6) end
  end
  newblock 1 *
    nroots 55
```

(continues on next page)

(continued from previous page)

```
    excitations none
    Flags[is] 1
    Flags[sa] 1
    Flags[ia] 1
    Flags[ijss] 1
    refs cas(8,6) end
  end
  soc
    dosoc true
  end
end
```

We use `excitations none` to set the default excitation flags to false and then manually set the 1h (`Flags[is]`), 1p (`Flags[sa]`), 1h1p (`Flags[ia]`) and 2h (`Flags[ijss]`) excitation flags to true.

WARNINGS

- Currently, MR-EOM-T|T†|SXD|U-h-v calculations can only be run with the default excitation classes in the final MRCI (i.e. 1h and 1p). Any other input options for the excitation flags will automatically be overwritten and set to the default values.
- Only the inclusion of spin-orbit coupling effects has been tested for MR-EOM calculations. Other features that are available in the MRCI module (e.g. spin-spin coupling, magnetic property calculations, etc.) have not been tested for use within MR-EOM calculations.

7.40.3 A Projection/Singular PT Scheme to Overcome Convergence Issues in the T Amplitude Iterations

In certain cases, there may be nearly singular T_2 amplitudes (often, but not always large in magnitude), which can cause convergence issues in the solution of the T amplitude equations. Hence, it is sometimes necessary to discard some of the amplitudes to remedy these convergence problems. The nearly singular T_2 amplitudes are of the form t^{ab}_{wx}, where (w, x) is a pair of active orbitals which corresponds to a small eigenvalue (pair occupation number n_{wx}) of the two-body reduced density matrix (RDM). When nearly singular amplitudes are present, it is possible to employ a singular PT/projection scheme (i.e. Scheme I described in Ref. [193]), using the two-body RDM as the metric matrix, to discard these nearly singular amplitudes and replace them with suitable perturbative estimates. As a first example, let us consider the following calculation on the cyclopentadiene molecule:

```
!MR-EOM def2-SVP VeryTightSCF

%casscf
  nel 4
  norb 4
  nroots 2
  mult 3
end

%mdci
  STol 1e-7
  MaxIter 60
end

%mrci
  newblock 1 *
    nroots 3
    refs cas(4,4) end
  end
  newblock 3 *
    nroots 3
    refs cas(4,4) end
  end
```

(continues on next page)

(continued from previous page)

```
end

* xyz 0 3
  H       -0.879859        0.000000        1.874608
  H        0.879859        0.000000        1.874608
  H        0.000000        2.211693        0.612518
  H        0.000000       -2.211693        0.612518
  H        0.000000        1.349811       -1.886050
  H        0.000000       -1.349811       -1.886050
  C        0.000000        0.000000        1.215652
  C        0.000000       -1.177731        0.285415
  C        0.000000        1.177731        0.285415
  C        0.000000       -0.732372       -0.993420
  C        0.000000        0.732372       -0.993420
*
```

The T amplitude iterations do not converge after 60 iterations and show no signs of convergence (i.e. final largest residual of 0.000458135 and oscillatory behavior over a significant portion of the iterations). If we inspect the largest T amplitudes,

```
--------------------
LARGEST T AMPLITUDES
--------------------
  19-> 24  19-> 24       0.043103
  19-> 23  19-> 23       0.031162
  11-> 25  11-> 25       0.028458
  19-> 41  19-> 41       0.027950
  11-> 47  11-> 47       0.027026
  19-> 22  19-> 22       0.025163
  19-> 21  19-> 21       0.022167
  15-> 26  15-> 26       0.022084
  11-> 47  11-> 25       0.022033
  11-> 25  11-> 47       0.022033
  19-> 24  19-> 29       0.021769
  19-> 29  19-> 24       0.021769
  19-> 36  19-> 36       0.020986
  17-> 38  17-> 38       0.019743
  19-> 41  16-> 36       0.019107
  18-> 40  18-> 40       0.017949
```

one can see that there are no unusually large amplitudes. If we turn on the singular PT/projection scheme by adding the line `DoSingularPT true` to the `%mdci` block:

```
%mdci
  STol 1e-7
  MaxIter 60
  DoSingularPT true
end
```

and rerun the calculation, we find that the T amplitude iterations now successfully converge in 22 iterations. If we look at the largest T amplitudes:

```
--------------------
LARGEST T AMPLITUDES
--------------------
  11-> 25  11-> 25       0.028440
  11-> 47  11-> 47       0.027027
  15-> 26  15-> 26       0.022069
  11-> 47  11-> 25       0.022031
  11-> 25  11-> 47       0.022031
  19-> 41  19-> 41       0.020463
```

(continues on next page)

(continued from previous page)

```
  17-> 38  17-> 38       0.018288
  11-> 43  11-> 43       0.017250
  11-> 39  11-> 39       0.016838
  15-> 27  15-> 27       0.016001
  13-> 26  13-> 26       0.015985
  16-> 36  16-> 36       0.015759
  19-> 41  16-> 36       0.015697
  18-> 40  18-> 40       0.015376
  17-> 31  17-> 31       0.015074
  18-> 40  17-> 38       0.014470
```

most of the amplitudes corresponding to the active pair $(w, x) = (19, 19)$ no longer appear in the list (i.e. they are nearly singular amplitudes which have been projected out). The only one that does appear in the list, corresponds to a projected perturbative estimate (e.g. `19-> 41 19-> 41 0.020463`).

By default, when the singular PT/projection scheme is active, the amplitudes t_{wx}^{ab} for which the pair occupation numbers satisfy $n_{wx} < 0.01$ (i.e. `SingularPTThresh = 0.01`), are replaced by perturbative amplitudes in the procedure. However, in some cases, it might be necessary to increase the `SingularPTThresh` threshold beyond the default value to achieve convergence. One such example is the ferrocene molecule. Consider the following calculation:

```
!MR-EOM def2-SVP

%casscf
  nel 6
  norb 5
  mult  1,3
  nroots 5,6
end

%mdci
  DoSingularPT true
  MaxIter 50
end

%mrci
  newblock 1 *
    nroots 18
    refs cas(6,5) end
  end
  newblock 3 *
    nroots 10
    refs cas(6,5) end
  end
end

* xyz 0 1
  Fe 0.000000  0.000000  0.000000
  C  0.000000  1.220080  1.650626
  C -1.160365  0.377025  1.650626
  C -0.717145 -0.987065  1.650626
  C  0.717145 -0.987065  1.650626
  C  1.160365  0.377025  1.650626
  C  0.000000  1.220080 -1.650626
  C  1.160365  0.377025 -1.650626
  C  0.717145 -0.987065 -1.650626
  C -0.717145 -0.987065 -1.650626
  C -1.160365  0.377025 -1.650626
  H  0.000000  2.306051  1.635648
  H -2.193184  0.712609  1.635648
  H -1.355463 -1.865634  1.635648
```

(continues on next page)

(continued from previous page)

```
  H  1.355463 -1.865634  1.635648
  H  2.193184  0.712609  1.635648
  H  0.000000  2.306051 -1.635648
  H  2.193184  0.712609 -1.635648
  H  1.355463 -1.865634 -1.635648
  H -1.355463 -1.865634 -1.635648
  H -2.193184  0.712609 -1.635648
end
```

The T amplitude iterations do not converge after 50 iterations, even though the singular PT/projection scheme is activated. If we increase `SingularPTThresh` to 0.05 by adding `SingularPTThresh 0.05` to the `%mdci` block:

```
%mdci
  DoSingularPT true
  SingularPTThresh 0.05
  MaxIter 50
end
```

the T amplitude iterations successfully converge in 25 iterations.

In conclusion, it occasionally happens that the T amplitude iterations do not converge. In these cases, the singular PT/projection scheme can be activated (`DoSingularPT true`) to overcome these convergence difficulties. Sometimes, like in the case of ferrocene, it is necessary to adjust the threshold for the singular PT/projection procedure (`SingularPTThresh`) to achieve convergence. If the procedure still fails with larger values of the threshold, then it might be necessary to revisit the definition of the state-averaged CAS.

7.40.4 An Orbital Selection Scheme for More Efficient Calculations of Excitation Spectra with MR-EOM

As described in Ref. [406], the MR-EOM implementation in ORCA can make use of a sophisticated scheme to discard inactive and virtual orbitals, which are not important for the description of the excited states of interest. The selection of inactive core orbitals is based on the eigenvalues of the core orbital selection density

$$D_{i'j'} = D^t_{i'j'} + \frac{\mathrm{Tr}\,(D^t)}{\mathrm{Tr}\,(D^s) + \mathrm{Tr}\,(D^u)} \left(D^s_{i'j'} + D^u_{i'j'}\right), \tag{7.250}$$

in which

$$\begin{aligned}
D^t_{i'j'} &= \sum_{w,a,b} t^{ab\,(1)}_{i'w} \left(2t^{ab\,(1)}_{j'w} - t^{ba\,(1)}_{j'w}\right), \\
D^s_{i'j'} &= \sum_{k,w,a} \left[s^{aw\,(1)}_{i'k} \left(2^{aw\,(1)}_{j'k} - s^{aw\,(1)}_{kj'}\right) + s^{aw\,(1)}_{ki'} \left(2s^{aw\,(1)}_{kj'} - s^{aw\,(1)}_{j'k}\right)\right], \\
D^u_{i'j'} &= \sum_{k',w,x} u^{wx\,(1)}_{i'k'} \left(2u^{wx\,(1)}_{j'k'} - u^{wx\,(1)}_{k'j'}\right),
\end{aligned}$$

are respectively, the contributions from the first order $t^{ab\,(1)}_{i'w}$, $s^{aw\,(1)}_{i'k}$ and $u^{wx\,(1)}_{i'k'}$ amplitudes (i.e. note that all amplitudes have at least one active label). Similarly, the selection of virtual orbitals is based upon the eigenvalues of the virtual orbital selection density

$$\rho_{ab} = \rho^t_{ab} + \frac{\mathrm{Tr}\,(\rho^t)}{\mathrm{Tr}\,(\rho^s)} \rho^s_{ab}, \tag{7.251}$$

in which, the contribution ρ^t, from the first order T_2 amplitudes, is given by

$$\rho^t_{ab} = \sum_{k,w,c} t^{ac\,(1)}_{wk} \left(2t^{bc\,(1)}_{wk} - t^{cb\,(1)}_{wk}\right) + \sum_{i',w,c} t^{ac\,(1)}_{i'w} \left(2t^{bc\,(1)}_{i'w} - t^{cb\,(1)}_{i'w}\right), \tag{7.252}$$

and the contribution ρ^s, from the first order S_2 amplitudes, is given by

$$\rho^s_{ab} = \sum_{i',k,w} s^{aw\,(1)}_{i'k} \left(2s^{bw\,(1)}_{i'k} - s^{bw\,(1)}_{ki'}\right) + \sum_{i',w,x} s^{aw\,(1)}_{xi'} \left(2s^{bw\,(1)}_{xi'} - s^{bw\,(1)}_{i'x}\right). \tag{7.253}$$

Diagonalization of the core orbital selection density $D_{i'j'}$ and the virtual orbital selection density ρ_{ab} then yields two respective sets of eigenvalues $\{\lambda_{i'}\}$ and $\{\lambda_a\}$. We have found it useful to compute the ratios,

$$\mathcal{R}_{\text{core}} = \frac{\sum_{i'=0}^{n_{\text{core}}^{\text{excluded}}} \lambda_{i'}}{\sum_{i'=0}^{n_{\text{core}}} \lambda_{i'}} \times 100\%, \tag{7.254}$$

$$\mathcal{R}_{\text{virt}} = \frac{\sum_{a=0}^{n_{\text{virt}}^{\text{excluded}}} \lambda_{a}}{\sum_{a=0}^{n_{\text{virt}}} \lambda_{a}} \times 100\%, \tag{7.255}$$

of the sum of the excluded eigenvalues to the sum over all eigenvalues. The orbital selection in the core and virtual subspaces is then based upon the values of these ratios, as will be discussed below.

The orbital selection procedure is activated by adding the keyword `OrbitalSelection` to the first line of input, e.g.

```
! MR-EOM def2-TZVPP VeryTightSCF OrbitalSelection
```

There are two threshold parameters `CoreThresh` and `VirtualThresh`, which are used to determine which inactive core and virtual orbitals are to be discarded in the orbital selection procedure, respectively. Namely, all inactive core orbitals for which $\mathcal{R}_{\text{core}} <$ `CoreThresh` (i.e. $\mathcal{R}_{\text{core}}$ as defined in Eq. (7.254)) are discarded and all virtual orbitals satisfying the condition $\mathcal{R}_{\text{virt}} <$ `VirtualThresh` (i.e. $\mathcal{R}_{\text{virt}}$ as defined in Eq. (7.255)) are discarded. The default values of these thresholds are `CoreThresh = 0.0` (no core orbital selection) and `VirtualThresh = 1.0`. However, the values of these parameters can easily be changed by redefining them in the `%mdci` block:

```
%mdci
  CoreThresh 1.0
  VirtualThresh 1.0
end
```

Let us consider the calculation of the previous section (*A Projection/Singular PT Scheme to Overcome Convergence Issues in the T Amplitude Iterations*) on ferrocene, with the orbital selection procedure activated (using the default thresholds):

```
!MR-EOM def2-SVP OrbitalSelection

%casscf
  nel 6
  norb 5
  mult   1,3
  nroots 5,6
end

%mdci
  DoSingularPT true
  SingularPTThresh 0.05
  MaxIter 50
end

%mrci
  newblock 1 *
    nroots 18
    refs cas(6,5) end
  end
  newblock 3 *
    nroots 10
    refs cas(6,5) end
  end
end

* xyz 0 1
  Fe 0.000000  0.000000  0.000000
  C  0.000000  1.220080  1.650626
```

(continues on next page)

(continued from previous page)

```
  C -1.160365  0.377025  1.650626
  C -0.717145 -0.987065  1.650626
  C  0.717145 -0.987065  1.650626
  C  1.160365  0.377025  1.650626
  C  0.000000  1.220080 -1.650626
  C  1.160365  0.377025 -1.650626
  C  0.717145 -0.987065 -1.650626
  C -0.717145 -0.987065 -1.650626
  C -1.160365  0.377025 -1.650626
  H  0.000000  2.306051  1.635648
  H -2.193184  0.712609  1.635648
  H -1.355463 -1.865634  1.635648
  H  1.355463 -1.865634  1.635648
  H  2.193184  0.712609  1.635648
  H  0.000000  2.306051 -1.635648
  H  2.193184  0.712609 -1.635648
  H  1.355463 -1.865634 -1.635648
  H -1.355463 -1.865634 -1.635648
  H -2.193184  0.712609 -1.635648
end
```

The details of the orbital selection procedure are printed under the heading `ORBITAL SELECTION`:

```
-----------------------------------------------
ORBITAL SELECTION
-----------------------------------------------

T1 is NOT used in the construction of the orbital selection densities
Factor (in percent) for inactive (core) orbital selection                    ..
↪.      0.000000000
Factor (in percent) for virtual orbital selection                            ..
↪.      1.000000000
Inactive orbitals before selection:    15 ...   44 ( 30 MO's/ 60 electrons)
Virtual orbitals before selection:     50 ...  220 (171 MO's              )
Inactive orbitals after selection:     15 ...   44 ( 30 MO's/ 60 electrons)
Virtual orbitals after selection:      50 ...  126 ( 77 MO's              )

-----------------------------------------
TIMINGS FOR THE ORBITAL SELECTION PROCEDURE
-----------------------------------------

Total Time for Orbital Selection            ...       61.470 sec

First Half Transformation                   ...       58.041 sec ( 94.4%)
Second Half Transformation                  ...        1.789 sec (  2.9%)
Formation of Orbital Selection Densities    ...        1.629 sec (  2.7%)
Core Orbital Selection                      ...        0.000 sec (  0.0%)
Virtual Orbital Selection                   ...        0.003 sec (  0.0%)
Finalization of Orbitals                    ...        0.006 sec (  0.0%)
```

Comparing the number of virtual orbitals before the orbital selection procedure (171) with the number that are left after orbital selection (77), we see that more than half have been discarded (94). The canonical calculation (without orbital selection) takes 149373 seconds to run and yields the following excitation energies:

```
-------------------
TRANSITION ENERGIES
-------------------

The lowest energy is  -1648.190045042 Eh

State Mult Irrep Root Block    mEh          eV      1/cm
```

(continues on next page)

(continued from previous page)

```
  0    1    -1    0    0      0.000      0.000       0.0
  1    3    -1    0    1     65.110      1.772   14289.9
  2    3    -1    1    1     65.110      1.772   14289.9
  3    3    -1    2    1     70.413      1.916   15454.0
  4    3    -1    3    1     70.413      1.916   15454.0
  5    3    -1    4    1     95.979      2.612   21065.0
  6    3    -1    5    1     95.979      2.612   21065.0
  7    1    -1    1    0    105.302      2.865   23111.1
  8    1    -1    2    0    105.302      2.865   23111.1
  9    1    -1    3    0    107.034      2.913   23491.4
 10    1    -1    4    0    107.034      2.913   23491.4
 11    1    -1    5    0    160.595      4.370   35246.6
 12    1    -1    6    0    160.596      4.370   35246.6
 13    3    -1    6    1    164.694      4.482   36146.1
 14    3    -1    7    1    165.379      4.500   36296.6
 15    3    -1    8    1    165.379      4.500   36296.6
 16    3    -1    9    1    171.464      4.666   37632.1
 17    1    -1    7    0    208.587      5.676   45779.6
 18    1    -1    8    0    208.587      5.676   45779.6
 19    1    -1    9    0    213.093      5.799   46768.6
 20    1    -1   10    0    213.093      5.799   46768.6
 21    1    -1   11    0    216.225      5.884   47456.0
 22    1    -1   12    0    220.230      5.993   48334.9
 23    1    -1   13    0    220.230      5.993   48334.9
 24    1    -1   14    0    224.583      6.111   49290.3
 25    1    -1   15    0    224.583      6.111   49290.3
 26    1    -1   16    0    237.914      6.474   52216.0
 27    1    -1   17    0    237.914      6.474   52216.0
```

In contrast, the calculation with the orbital selection procedure activated runs in 28977 seconds (a factor of 5 speedup) and produces the following excitation energies:

```
-------------------
TRANSITION ENERGIES
-------------------

The lowest energy is  -1647.788478559 Eh

State Mult Irrep Root Block    mEh          eV      1/cm
  0    1    -1    0    0      0.000      0.000       0.0
  1    3    -1    0    1     65.112      1.772   14290.4
  2    3    -1    1    1     65.134      1.772   14295.3
  3    3    -1    2    1     70.520      1.919   15477.3
  4    3    -1    3    1     70.520      1.919   15477.3
  5    3    -1    4    1     96.105      2.615   21092.7
  6    3    -1    5    1     96.134      2.616   21099.0
  7    1    -1    1    0    105.415      2.868   23136.0
  8    1    -1    2    0    105.450      2.869   23143.5
  9    1    -1    3    0    107.294      2.920   23548.3
 10    1    -1    4    0    107.294      2.920   23548.3
 11    1    -1    5    0    161.082      4.383   35353.4
 12    1    -1    6    0    161.094      4.384   35356.0
 13    3    -1    6    1    164.786      4.484   36166.4
 14    3    -1    7    1    165.465      4.503   36315.4
 15    3    -1    8    1    165.465      4.503   36315.5
 16    3    -1    9    1    171.542      4.668   37649.1
 17    1    -1    7    0    208.853      5.683   45838.0
 18    1    -1    8    0    208.853      5.683   45838.0
 19    1    -1    9    0    213.419      5.807   46840.1
 20    1    -1   10    0    213.419      5.807   46840.1
 21    1    -1   11    0    216.526      5.892   47521.9
```

(continues on next page)

(continued from previous page)

```
 22    1    -1   12    0    220.611      6.003   48418.4
 23    1    -1   13    0    220.611      6.003   48418.5
 24    1    -1   14    0    225.135      6.126   49411.5
 25    1    -1   15    0    225.136      6.126   49411.5
 26    1    -1   16    0    238.388      6.487   52320.1
 27    1    -1   17    0    238.388      6.487   52320.1
```

We note that the excitation energies in the orbital selection procedure agree very nicely with those of the canonical calculation. However, the total energies are significantly different, as we currently have not implemented a procedure to correct them. Hence, the following warning is particularly important.

WARNING

- The orbital selection procedure should only be used for the calculation of excitation energies. Total energies computed with the orbital selection procedure have not been corrected and can differ greatly from the canonical results.

Before leaving the discussion of the orbital selection procedure, we note that there is also a keyword `PrintOrbSelect`, which can be added to the `%mdci` block to print the eigenvalues of the inactive core orbital selection and virtual orbital selection densities and the corresponding values of $\mathcal{R}_{\text{core}}$ and $\mathcal{R}_{\text{virt}}$ defined in Eqs. (7.254) and (7.255), respectively. This is useful if one wants to manually select the orbitals to discard in the orbital selection procedure by adjusting the values of `CoreThresh` and `VirtualThresh`. We note that the program terminates after printing. In the case of the calculation on ferrocene, if we modify the `%mdci` block to read

```
%mdci
  DoSingularPT true
  SingularPTThresh 0.05
  MaxIter 50
  PrintOrbSelect True
end
```

we find the following information in the `ORBITAL SELECTION` section of the output (only the first 50 values for the virtual orbital selection density are shown here):

```
Eigenvalues and corresponding R_core values for the core orbital selection density

Orbital        Eigenvalue        R_core
   0           0.00026936        0.419318
   1           0.00027080        0.840883
   2           0.00038739        1.443947
   3           0.00038739        2.047011
   4           0.00040299        2.674355
   5           0.00040299        3.301700
   6           0.00077636        4.510285
   7           0.00086085        5.850394
   8           0.00086085        7.190503
   9           0.00091850        8.620358
  10           0.00091850       10.050213
  11           0.00112826       11.806598
  12           0.00115561       13.605563
  13           0.00137961       15.753236
  14           0.00137961       17.900908
  15           0.00139093       20.066210
  16           0.00139093       22.231512
  17           0.00143349       24.463072
  18           0.00143350       26.694633
  19           0.00148539       29.006985
  20           0.00148539       31.319338
  21           0.00173415       34.018940
  22           0.00224131       37.508054
  23           0.00224132       40.997171
```

(continues on next page)

(continued from previous page)

```
  24           0.00533017       49.294785
  25           0.00533019       57.592429
  26           0.00658679       67.846267
  27           0.00662033       78.152314
  28           0.00701718       89.076149
  29           0.00701719      100.000000

 Eigenvalues and corresponding R_virt values for the virtual orbital selection density

Orbital        Eigenvalue        R_virt
   0           0.00000119        0.000450
   1           0.00000119        0.000899
   2           0.00000134        0.001404
   3           0.00000134        0.001909
   4           0.00000136        0.002423
   5           0.00000177        0.003091
   6           0.00000178        0.003764
   7           0.00000178        0.004437
   8           0.00000215        0.005248
   9           0.00000224        0.006096
  10           0.00000224        0.006944
  11           0.00000238        0.007844
  12           0.00000347        0.009154
  13           0.00000347        0.010465
  14           0.00000364        0.011841
  15           0.00000386        0.013299
  16           0.00000396        0.014793
  17           0.00000396        0.016287
  18           0.00000437        0.017937
  19           0.00000437        0.019587
  20           0.00000499        0.021472
  21           0.00000499        0.023357
  22           0.00000794        0.026354
  23           0.00000794        0.029352
  24           0.00000819        0.032447
  25           0.00000819        0.035543
  26           0.00000927        0.039044
  27           0.00000927        0.042546
  28           0.00001002        0.046332
  29           0.00001002        0.050119
  30           0.00001137        0.054415
  31           0.00001137        0.058711
  32           0.00001158        0.063086
  33           0.00001158        0.067461
  34           0.00001381        0.072678
  35           0.00001381        0.077894
  36           0.00001417        0.083249
  37           0.00001417        0.088604
  38           0.00001465        0.094137
  39           0.00001495        0.099785
  40           0.00001495        0.105432
  41           0.00001554        0.111302
  42           0.00001554        0.117172
  43           0.00001623        0.123303
  44           0.00001689        0.129685
  45           0.00001754        0.136310
  46           0.00001754        0.142934
  47           0.00001805        0.149752
  48           0.00001805        0.156570
  49           0.00002111        0.164546
```

(continues on next page)

(continued from previous page)

```
.
.
.
```

In conclusion, the orbital selection scheme provides a more efficient way to calculate accurate excitation spectra within the framework of MR-EOM. It can be used to extend the applicability of this approach to larger systems and we expect it to be much more effective in larger systems where the chromophore is localized to a small part of the molecule. We reiterate that it is currently limited to the calculation of excitation energies and should not be used if one is interested in total energies.

7.40.5 Nearly Size Consistent Results with MR-EOM by Employing an MR-CEPA(0) Shift in the Final Diagonalization Procedure

One drawback of the MR-EOM methodology is that it is not size-extensive (or size-consistent). The size-extensivity errors arise due to the final uncontracted MR-CI diagonalization step. Namely, they result from the components of the eigenvectors of the transformed Hamiltonian, which lie outside of the CASSCF reference space (e.g. 1h, 1p, etc. configurations). As more of the excitation classes are included through the successive similarity transformations of the Hamiltonian, the size of the final diagonalization manifold is greatly decreased resulting in much smaller size-extensivity errors upon going from MR-EOM-T|T†-h-v to MR-EOM-T|T†|SXD|U-h-v. To illustrate this, let us consider the O_2—O_2 dimer where the O_2 molecules are separated by a large distance. For the O_2 monomer, we employ a minimal active space consisting of 2 electrons distributed amongst the two π^* orbitals and we only consider the ground $^3\Sigma_g^-$ state (no state-averaging). In the MR-EOM calculations, we also calculate the higher lying $^1\Delta_g$ and $^1\Sigma_g^+$ singlet states. For example, the input file for the MR-EOM-T|T†|SXD|U-h-v calculation is given by:

```
!MR-EOM AUG-CC-PVTZ EXTREMESCF

%casscf
  nel 2
  norb 2
  nroots 1
  mult 3
end

%mdci
  MaxIter 300
  STol 1e-12
end

%mrci
  newblock 1 *
    nroots 3
    refs cas(2,2) end
  end
  newblock 3 *
    nroots 1
    refs cas(2,2) end
  end
end

* xyz 0 3
  O  0.00000000    -0.00000000    -0.60500000
  O -0.00000000     0.00000000     0.60500000
*
```

In the case of the dimer, we take the reference state as the coupled quintet state which is formed as the product $^3\Sigma_g^+ \otimes {}^3\Sigma_g^+$ of the monomer states. We note that at large separation, in the non-interacting limit, the dimer state energies can be decomposed as the sum of monomer state energies. There are various possibilities, taking into account the degeneracies of the various states:

1. a singlet, a triplet, and a quintet with energy $E(^3\Sigma_g^- + {}^3\Sigma_g^-)$,
2. four triplets with energy $E(^3\Sigma_g^- + {}^1\Delta_g)$,
3. two triplets with energy $E(^3\Sigma_g^- + {}^1\Sigma_g^+)$,
4. four singlets with energy $E(^1\Delta_g + {}^1\Delta_g)$,
5. four singlets with energy $E(^1\Delta_g + {}^1\Sigma_g^+)$,
6. a singlet with energy $E(^1\Sigma_g^+ + {}^1\Sigma_g^+)$.

Hence, in the final diagonalization step of the MR-EOM calculation, we must ask for 10 singlets, 7 triplets and 1 quintet. The input file for the MR-EOM-T|T†|SXD|U-h-v calculation on the dimer is given by:

```
!MR-EOM AUG-CC-PVTZ EXTREMESCF

%casscf
  nel 4
  norb 4
  nroots 1
  mult 5
  etol 1e-13
  gtol 1e-13
end

%mdci
  MaxIter 300
  STol 1e-12
end

%mrci
  newblock 1 *
    nroots 10
    refs cas(4,4) end
  end
  newblock 3 *
    nroots 7
    refs cas(4,4) end
  end
  newblock 5 *
    nroots 1
    refs cas(4,4) end
  end
end

* xyz 0 5
  O                  0.00000000     0.00000000  -500.60500000
  O                  0.00000000    -0.00000000  -499.39500000
  O                 -0.60500000     0.00000000   500.00000000
  O                  0.60500000    -0.00000000   500.00000000
*
```

In Table *Test for size consistency in MR-EOM: Differences in energy (in mE_\text{h}) between the _2—O_2 dimer energies (at large separation) and the sum of the monomer energies for the ground state and various excited states. The results were obtained in an aug-cc-pVTZ basis using minimal active spaces.*, we have compiled the results of the size consistency test, taking the difference of the dimer state energies (at large separation) and the sum of the monomer state energies (in mE_h). It is evident that as more excitation classes are included in the similarity transformed Hamiltonian and the size of the final diagonalization manifold is decreased, the size-consistency errors decrease. Of particular note are the results for the MR-EOM-T|T†|SXD|U-h-v approach (only includes 1h and 1p configurations in the final diagonalization manifold), for which the largest deviation is 1.25×10^{-2} mE_h. The much larger deviations for the MR-EOM-T|T†|SXD-h-v approach clearly demonstrate the large effect that the 2h excitations have on the size-consistency errors.

Table 7.26: Test for size consistency in MR-EOM: Differences in energy (in mE_h) between the $_2$—O_2 dimer energies (at large separation) and the sum of the monomer energies for the ground state and various excited states. The results were obtained in an aug-cc-pVTZ basis using minimal active spaces.

	T|T†-h-v	T|T†|SXD-h-v (with 1h1p)	T|T†|SXD-h-v	T|T†|SXD|U-h-v
$\Delta E(^3\Sigma_g^- + {}^3\Sigma_g^-)$	12.74	2.77	1.11	1.00×10^{-5}
$\Delta E(^3\Sigma_g^- + {}^1\Delta_g)$	14.20	3.84	1.85	1.54×10^{-4}
$\Delta E(^3\Sigma_g^- + {}^1\Sigma_g^+)$	17.21	5.52	2.83	4.13×10^{-4}
$\Delta E(^1\Delta_g + {}^1\Delta_g)$	15.69	3.10	2.31	5.26×10^{-3}
$\Delta E(^1\Delta_g + {}^1\Sigma_g^+)$	18.83	7.52	4.76	5.89×10^{-3}
$\Delta E(^1\Sigma_g^+ + {}^1\Sigma_g^+)$	22.34	10.75	7.31	1.25×10^{-2}

To reduce the size-consistency errors, one can make use of the MR-CEPA(0) shift in the final diagonalization step. This MR-CEPA(0) shift can easily be activated by adding the line

```
citype mrcepa_0
```

to the beginning of the `%mrci` block. The results of the size-consistency test with the use of the MR-CEPA(0) shift are tabulated in Table *Test for size consistency in MR-EOM, using the MR-CEPA(0) shift: Differences in energy (in mE_\text{h}) between the O_2—O_2 dimer energies (at large separation) and the sum of the monomer energies for the ground state and various excited states. The results were obtained in an aug-cc-pVTZ basis using minimal active spaces and the MR-CEPA(0) shift was applied in the final diagonalization in each case.*. For each of the methods, we see a marked improvement over the results of Table *Test for size consistency in MR-EOM: Differences in energy (in mE_\text{h}) between the _2—O_2 dimer energies (at large separation) and the sum of the monomer energies for the ground state and various excited states. The results were obtained in an aug-cc-pVTZ basis using minimal active spaces.*, which do not make use of the MR-CEPA(0) shift. The greatest improvement occurs in the MR-EOM-T|T†|SXD-h-v and the MR-EOM-T|T†|SXD|U-h-v results. Namely, the errors in the former case are on the order of nano Hartrees, while the errors in the MR-EOM-T|T†|SXD|U-h-v results are not detectable (sub-nano Hartree), as the energy is only printed with nine decimal places. It is interesting to note that upon adding the 1h1p configurations to the diagonalization manifold in the MR-EOM-T|T†|SXD-h-v calculations (i.e. with 1h1p), the size-consistency errors increase greatly. Hence, it appears that the use of the MR-CEPA(0) shift is most effective at reducing the size-consistency errors resulting from the presence of the 1h, 1p and 2h configurations in the final diagonalization manifold. In any case, one can easily take advantage of this approach to obtain nearly size-consistent results with both the MR-EOM-T|T†|SXD-h-v and MR-EOM-T|T†|SXD|U-h-v methods.

Table 7.27: Test for size consistency in MR-EOM, using the MR-CEPA(0) shift: Differences in energy (in mE_h) between the O_2—O_2 dimer energies (at large separation) and the sum of the monomer energies for the ground state and various excited states. The results were obtained in an aug-cc-pVTZ basis using minimal active spaces and the MR-CEPA(0) shift was applied in the final diagonalization in each case.

	T|T†-h-v	T|T†|SXD-h-v (with 1h1p)	T|T†|SXD-h-v	T|T†|SXD|U-h-v
$\Delta E(^3\Sigma_g^- + {}^3\Sigma_g^-)$	2.75×10^{-3}	0.01	2.00×10^{-6}	0.00
$\Delta E(^3\Sigma_g^- + {}^1\Delta_g)$	0.06	0.07	0.00	0.00
$\Delta E(^3\Sigma_g^- + {}^1\Sigma_g^+)$	0.14	0.15	4.00×10^{-6}	0.00
$\Delta E(^1\Delta_g + {}^1\Delta_g)$	0.21	0.22	1.00×10^{-6}	0.00
$\Delta E(^1\Delta_g + {}^1\Sigma_g^+)$	0.42	0.44	5.00×10^{-6}	0.00
$\Delta E(^1\Sigma_g^+ + {}^1\Sigma_g^+)$	0.82	0.87	9.00×10^{-6}	0.00

7.40.6 Perturbative MR-EOM-PT

The MR-EOMPT approach was developed for situations where the full accuracy of the iterative MR-EOMCC method is not required. It performs on par with other multireference perturbation theories such as fic-NEVPT2 and does not have the convergence difficulties with the $\hat{T}$ and $\hat{S}, \hat{X}, \hat{D}$ amplitudes like its iterative parent method as these amplitudes are computed in a non-iterative fashion. The only iterative part of the MR-EOMPT method is the calculation of the $\hat{U}$ amplitudes since they are quick to converge anyways [501].

The setup procedure for the MR-EOMPT method is the same as for the MR-EOMCC method, and the foregoing also applies to the perturbative variant. Please note that the orbital selection scheme has not been tested with this variant and should be unnecessary anyways since calculations are much faster than with the iterative MR-EOMCC method.

To invoke the new variant, set up the calculation as you would for an MR-EOMCC calculation and then add the keyword `DoMREOM_MRPT True` to the `%mdci` block.

The results are interpreted just like results for the iterative MR-EOMCC method. After transforming the Hamiltonian with the perturbatively estimated amplitudes and the final MRCIS diagonalization step, the final state energies are printed along with their reference weights. For reliable results, we again recommend that the reference weight be >90%.

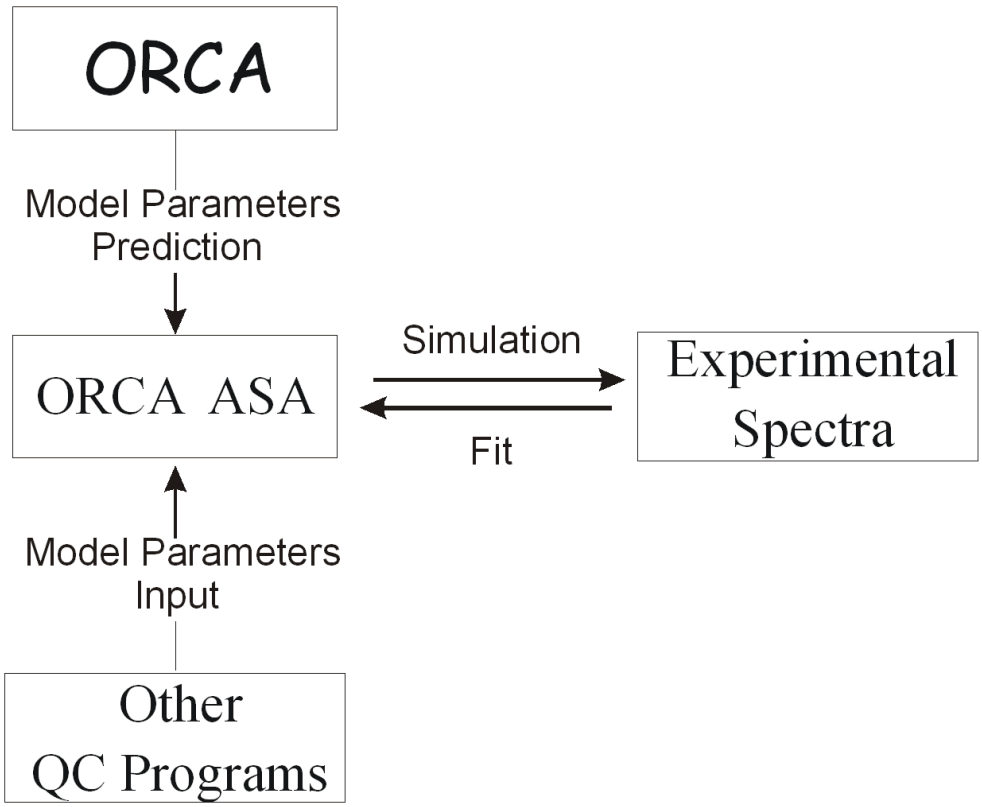
7.41 Simulation and Fit of Vibronic Structure in Electronic Spectra, Resonance Raman Excitation Profiles and Spectra with the orca_asa Program

Deprecated since ORCA 6.0.0

- The `orca_asa` program is no longer supported. It is still included in the 6.0.0 release and the documentation is preserved below. However, it may not function correctly and will be removed in a future ORCA version!
- The `!NMScan` and `!NMGrad` keywords are still available but these calculations may fail or not generate valid input for `orca_asa`. Please use the *ESD module* instead, if applicable.

In this section various aspects of the simulation and fit of optical spectra, including absorption, fluorescence, and resonance Raman are considered. This part of the ORCA is fairly autonomous and can also be used in a data analysis context, not only in a "quantum chemistry" mode. The program is called `orca_asa`, where ASA stands for "Advanced Spectral Analysis". The program was entirely designed by Dr. Taras Petrenko.

The general philosophy is as follows: An ORCA run produces the necessary data to be fed into the `orca_asa` program and writes an initial input file. This input file may be used to directly run `orca_asa` in order to predict an absorption, fluorescence or resonance Raman spectrum. Alternatively, the input file may be edited to change the parameters used in the simulations. Last – but certainly not least – the `orca_asa` program can be used to perform a fit of the model parameters relative to experimental data.



All examples below are taken from [678], which must be cited if you perform any work with the `orca_asa` program!

7.41.1 General Description of the Program

The program input comprises the following information: (1) model and specification of the model parameters characterizing the electronic structure of a molecule, as well as lineshape factors; (2) spectral ranges and resolution for simulations; (3) specification of vibrational transitions for rR excitation profile and spectra generation; (4) certain algorithm-selecting options depending on the model; (5) fitting options.

All optional parameters (1)-(3) are given in the `%sim` block, and fitting options are in the `%fit` block. The model parameters are specified within various blocks that will be described below. The program `orca_asa` is interfaced to ORCA and inherits its input style. The input for `orca_asa` run can be also generated upon ORCA run.

The current implementation features so called "simple", "independent mode, displaced harmonic oscillator" (IMDHO), and "independent mode, displaced harmonic oscillator with frequency alteration" (IMDHOFA) models.

7.41.2 Spectral Simulation Procedures: Input Structure and Model Parameters

Example: Simple Mode

This model represents the simplest approach which is conventionally used in analysis of absorption spectra. It neglects vibrational structure of electronic transitions and approximates each individual electronic band by a standard lineshape, typically a Gaussian, Lorentzian or mixed (Voigt) function. This model can only make sense if vibrational progressions are not resolved in electronic spectra. Upon this approximation the intensity of absorption spectrum depends on the energy of the incident photon (E_L), the electronic transition energy (E_T), the transition electric dipole moment ($\mathbf{M}$, evaluated at the ground-state equilibrium geometry). Lineshape factors are specified by homogeneous linewidth Γ and standard deviation parameter σ corresponding to Gaussian distribution of transition energies. The following example illustrates a simple input for simulation of absorption bandshapes using various intensity and lineshape parameters.

```
# example001.inp
#
# Input file to generate absorption spectrum consisting
# of 3 bands   with different lineshape factors:
#
# 1. Lorentzian centered at 18000cm**-1 (damping factor Gamma= 100 cm**-1)
# 2. Gaussian   centered at 20000cm**-1
#    (standard deviation Sigma= 100 cm**-1)
# 3. Mixed Gaussian-Lorentzian band representing Voigt profile
#    centered at 21000 cm**-1

%sim
         Model Simple

         # Spectral range for absorption simulation:
         AbsRange     17000.0,   23000.0

         # Number of points to simulate absorption spectrum:
         NAbsPoints  2000

end


#------------------------------------------------------------------------
#  Transition        Gamma    Sigma  Transition Dipole Moment (atomic unit)
#  Energy (cm**-1)  (cm**-1) (cm**-1)      Mx          My        Mz
#------------------------------------------------------------------------

$el_states
3     # number of electronic states
   1  18000.0        100.00      0.0        1.0         0.0       0.0
   2  20000.0          0.00    100.0        1.0         0.0       0.0
   3  22000.0         50.00     50.0        1.0         0.0       0.0
```

The parameters of of the final electronic states reached by the respective transitions are specified in the `$el_states` block. The spectral range and resolution used in the calculation are defined by the `AbsRange` and `NAbsPoints` keywords in `%sim` block. The calculation of the absorption spectrum is automatically invoked if `NAbsPoints`>1. After the `orca_asa` run you will find in your directory file `example001.abs.dat` containing absorption spectrum in simple two-column ASCII format suitable to be plotted with any spreadsheet program. Absorption spectra corresponding to individual electronic transitions are stored in file `example001.abs.as.dat` (the suffix "as" stands for "All States").

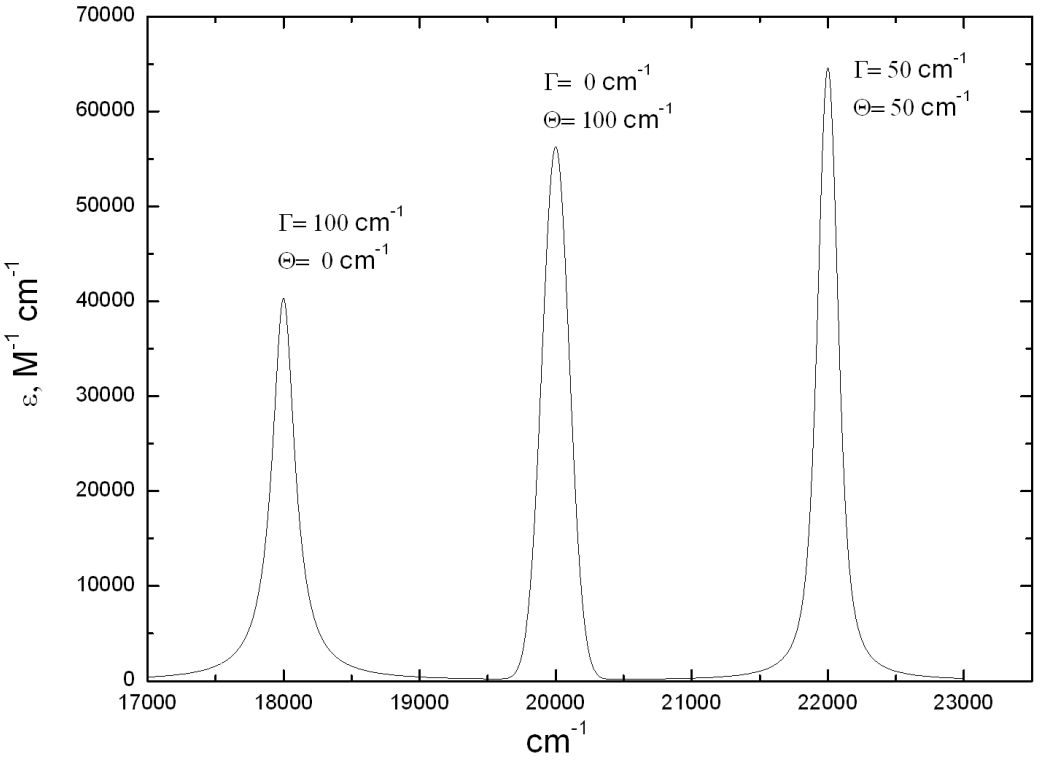


Fig. 7.43: Absorption spectrum generated after orca_asa run on file `example001.inp`. Three bands have different lineshape pararameters. Note that although all transitions are characterized by the same transition electric dipole moment their intensities are scaled proportionally to the transition energies.

The output of the program run also contains information about oscillator strengths and full-width-half-maximum (FWHM) parameters corresponding to each electronic band:

```
-----------------------------------------
   State       EV         fosc    Stokes shift
             (cm**-1)                (cm**-1)
-----------------------------------------
     1:      18000.00   0.054676      0.00
     2:      20000.00   0.060751      0.00
     3:      22000.00   0.066826      0.00


---------------------------------------
   BROADENING PARAMETETRS (cm**-1)
---------------------------------------
   State     Gamma     Sigma      FWHM
---------------------------------------
     1:        100.00      0.00    200.00
     2:          0.00    100.00    235.48
     3:         50.00     50.00    180.07
```

Note that although all three types of lineshape functions are symmetric this is not true for the overall shapes of individual absorption bands since the extinction coefficient (absorption cross-section) is also proportional to the incident photon energy. Therefore, if the linewidth is larger than 10% of the peak energy the asymmetry of the electronic band can be quite noticeable.

Example: Modelling of Absorption and Fluorescence Spectra within the IMDHO Model

The IMDHO model is the simplest approach that successfully allows for the prediction of vibrational structure in electronic spectra as well as rR intensities for a large variety of real systems. This model assumes:

1. harmonic ground- and excited-state potential energy surfaces;
2. origin shift of the excited-state potential energy surface relative to the ground-state one;
3. no vibrational frequency alteration or normal mode rotation occurs in the excited state;
4. no coordinate dependence of the electronic transition dipole moment.

In addition to the parameters that enter the "Simple model" defined above it requires some information about the vibrational degrees of freedom. The required information consists of the ground-state vibrational frequencies $\{\omega_{gm}\}$ and (dimensionless) origin shifts $\{\Delta_{mi}\}$, where i and m refer to electronic states and normal modes respectively. Δ is expressed in terms of dimensionless normal coordinates. Accordingly, for the IMDHO model one has to specify the following blocks

- The `$el_states` block contains the parameters $E_T, \Gamma, \sigma, \mathbf{M}$ for each electronic state. By default E_T is assumed to be adiabatic minima separation energy. Alternatively, it can be redefined to denote for the vertical transition energy.This is achieved by specifiying the keyword `EnInput=EV` in the `%sim` block.
- A `$vib_freq_gs` block specifies ground-state vibrational frequencies.
- A `$sdnc` block contains parameters $\{\Delta_{mi}\}$ in matrix form such that the i-th column represents the dimensionless displacements along all normal modes for the i-th excited-state PES.

The file `example002.inp` provides the input for simulation of absorption and fluorescence spectra of a system characterized by significant displacements of the excited-state origin along 5 normal coordinates.

```
# example002.inp
#
# Input file for simulation  of vibrational structure
# in absorption and fluorescence spectra assuming
# origin shift of excited PES along 5 normal coordinates.
# The simulated spectra closely reproduce the experimental
# optical bandshapes for the tetracene molecule.
#
%sim
     Model IMDHO

     # spectral range for absorption simulation (cm**-1)
     AbsRange    20000.0, 27000.0
     NAbsPoints  2000             # number of points in absorption spectrum

     # spectral range for simulation of fluorescence (cm**-1)
     FlRange     22000.0, 16000.0
     NFlPoints   2000      # number of points in fluorescence spectrum

     # the following options require the spectra to be normalized
     # so that their maxima are equal to 1.0
     AbsScaleMode Rel
     FlScaleMode  Rel   # default  for fluorescence

     # for absorption spectrum the default option is AbsScaleMode= Ext
     # which stands for extinction coefficient

end

#-----------------------------------------------------------------------------
#   Transition        Gamma    Sigma  Transition Dipole Moment (atomic unit)
#   Energy (cm**-1)  (cm**-1) (cm**-1)      Mx          My          Mz
#-----------------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
$el_states
1
   1  21140.0         50.00    100.0        1.0         0.0         0.0


# Block specifying Stokes Shift parameter for each electronic state
# This information is optional
$ss
1   # number of excited states
  1   300.0      # the Stokes shift for the 1st electronic transition

# Block providing the values of VIBrational FREQuencies
# for 5 Ground-State normal modes.
# Obligatory for IMDHO and IMDHOFA models.
$vib_freq_gs
5
    1      310.0
    2     1193.0
    3     1386.0
    4     1500.0
    5     1530.0


# Block specifying origin Shift of the excite-state PES
# along each normal mode in terms of the ground-state
# Dimensionless Normal Coordinates
# Obligatory for IMDHO and IMDHOFA models.
$sdnc
5 1
               1
     1       0.698
     2      -0.574
     3       0.932
     4      -0.692
     5       0.561
```

The calculation of absorption and fluorescence spectra is automatically invoked if the parameters `NAbsPoints>1` and `NFlPoints> 1`. The input file also contains the optional block `$ss` which specifies the Stokes shift λ for each electronic transition. This parameter is equal to the energy separation between the 0-0 vibrational peaks in the absorption and fluorescence spectra as shown in Fig. 7.44 . In general λ accounts for solvent induced effects as well as unresolved vibrational structure corresponding to low-frequency modes that are not specified in the input. Note that we have specified parameters `AbsScaleMode=Rel` and `FlScaleMode=Rel` in `%sim` block in order to ensure that the simulated spectra are normalized to unity. The calculated absorption and fluorescence spectra are stored in `example002.abs.dat` and `example002.fl.dat` files, respectively.

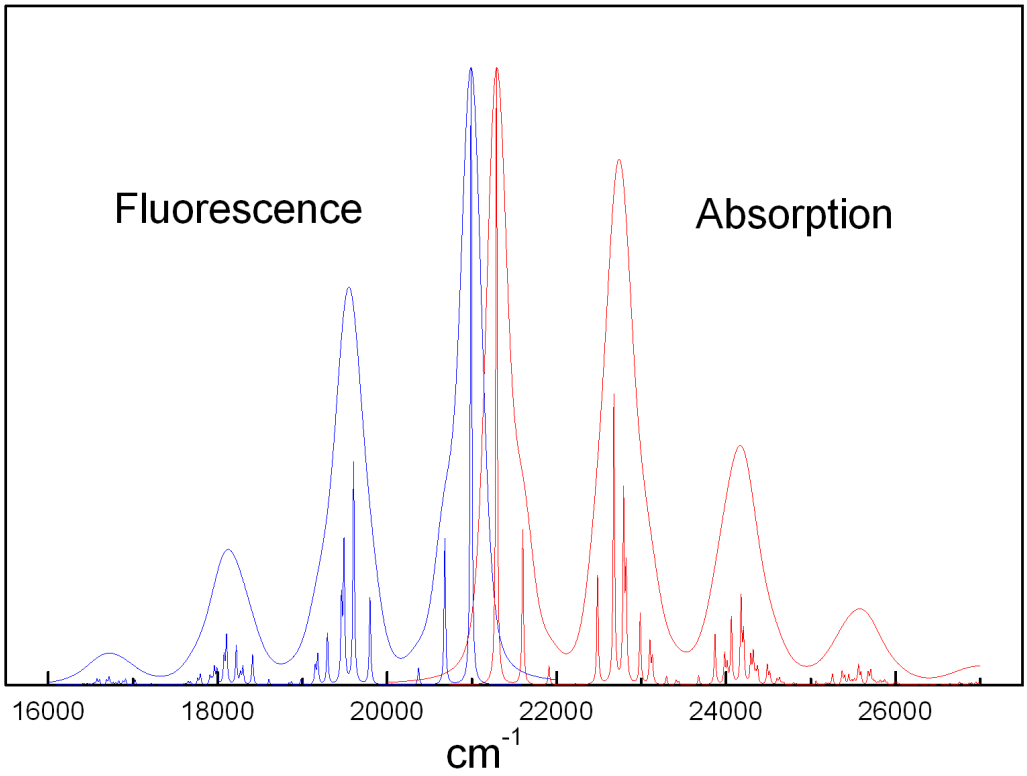


Fig. 7.44: Absorption and fluorescence spectra generated after `orca_asa` run on the file `example002.inp`. If the homogeneous broadening is set to be $\Gamma = 10\,\mathrm{cm}^{-1}$ one can resolve underlying vibrational structure and identify various fundamental and combination transitions.

Example: Modelling of Absorption and Fluorescence Spectra within the IMDHOFA Model

IMDHOFA (Independent Mode Displaced Harmonic Oscillators with Frequency Alteration) is based on the same assumptions as the IMDHO model except for vibrational frequency alteration in excited state can take place. The file `example003.inp` features almost the same input parameters as `example002.inp`. The IMDHOFA model is invoked by the keyword `Model=IMDHOFA` in the `%sim` block. Additionally, one has to provide the obligatory block `$vib_freq_es`. It contains the excited-state vibrational frequencies $\{\omega_{emi}\}$ in matrix form such that the i-th column represents the vibrational frequencies of all normal modes for the i-th excited-state PES.

```
# Block providing the values of VIBrational FREQuencies
# for 5 Excited-State normal modes.
# Obligatory for IMDHOFA model.

$vib_freq_es
5  1         # number of modes and number of excited states
             1
    1      410.0
    2     1293.0
    3     1400.0
    4     1600.0
    5     1730.0
```

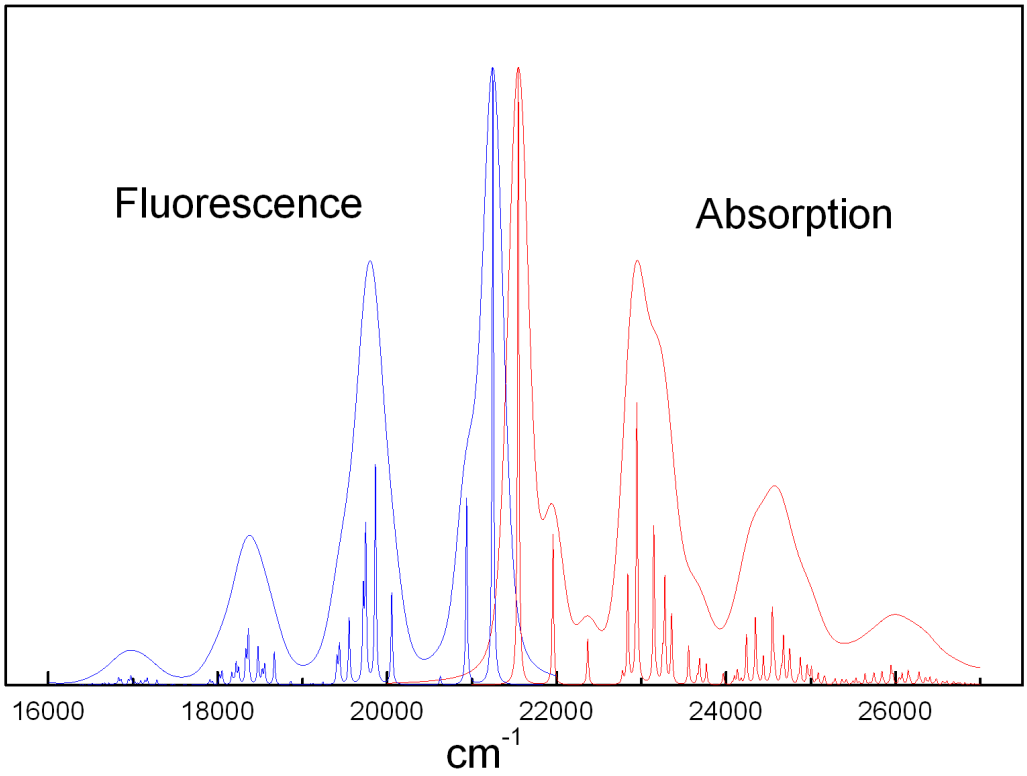


Fig. 7.45: Absorption and fluorescence spectra generated after `orca_asa` run on the file `example003.inp`. Also, the high-resolution spectra corresponding to homogeneous broadening $\Gamma = 10\,\mathrm{cm}^{-1}$ are shown.

Example: Modelling of Effective Broadening, Effective Stokes Shift and Temperature Effects in Absorption and Fluorescence Spectra within the IMDHO Model

For the IMDHO model the `orca_asa` is capable to model absorption and emission spectra in the finite-temperature approximation. While the keyword `Model=IMDHO` assumes the zero-temperature approximation, the value of `Model=IMDHOT` invokes the calculation of the spectra for the finite temperature which is specified by the paramter `TK` in the block `%sim`:

```
# example004.inp
#
#
%sim
     Model IMDHOT
     TK 300 # temperature (in Kelvin)

     # spectral range for absorption simulation (cm**-1)
     AbsRange    18000.0, 35000.0
     NAbsPoints  5000             # number of points in absorption spectrum

     # spectral range for simulation of fluorescence (cm**-1)
     FlRange     22000.0, 10000.0
     NFlPoints   5000      # number of points in fluorescence spectrum

     # the following options require the spectra to be normalized
```

(continues on next page)

(continued from previous page)

```
     # so that their maxima are equal to 1.0
     AbsScaleMode Rel
     FlScaleMode  Rel   # default  for fluorescence
end

#------------------------------------------------------------------------
#   Transition        Gamma    Sigma  Transition Dipole Moment (atomic unit)
#   Energy (cm**-1)  (cm**-1) (cm**-1)      Mx          My          Mz
#------------------------------------------------------------------------

$el_states
1
   1  21140.0         50.00    100.0        1.0         0.0         0.0


# Block specifying Stokes Shift parameter for each electronic state
$ss
1   # number of excited states
  1   300.0      # the Stokes shift for the 1st electronic transition

# Block providing the values of VIBrational FREQuencies
# for 10 Ground-State normal modes.
$vib_freq_gs
10
    1       30.0
    2       80.0
    3       100.0
    4       120.0
    5       130.0
    6      140.0
    7      160.0
    8      200.0
    9      310.0
   10     1300.0

# Block specifying origin Shift of the excite-state PES
# along each normal mode in terms of the ground-state
# Dimensionless Normal Coordinates
$sdnc
 10  1
               1
     1       2.5
     2       2.0
     3       1.8
     4       1.9
     5       1.5
     6       1.9
     7       2.4
     8       1.9
     9       2.5
    10       0.9
```

This example illustrates a typical situation in large molecules which feature a number of low frequency modes with significant values of dimensionless displacements for a given excited-state PES. In the case of high density of vibrational states with frequencies below or comparable to the intrincic value of FWHM (determined by Γ and σ) the vibrational progression is unresolved, whereby the spectra become very diffuse and show large separation between the maxima of absorption and emission spectra (Fig. 7.45). Besides, upon the condition $h\nu_i \leqslant kT$ the effective bandwidths and positions of maxima in the spectra can be strongly subject to temperature effects.

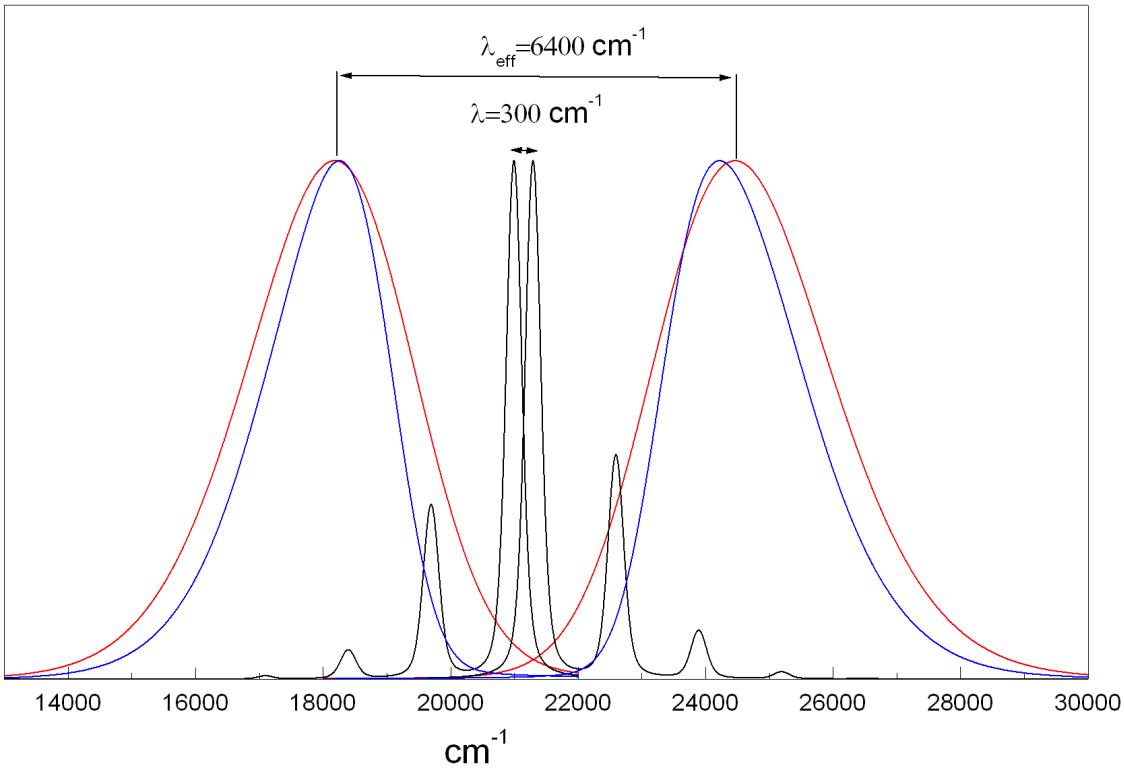


Fig. 7.46: Absorption and fluorescence spectra for T=0 K (blue) and T=300 K (red) generated after `orca_asa` run on the file `example004.inp`. Black lines show spectra corresponding to the case where all low-frequency modes were excluded from the calculation.

The effective Stokes shift and linewidth parameters which are evaluated in the simple self-consistent procedure are given in the output of the `orca_asa` run:

```
-------------------------------------------------------------------------
State       E0           EV        fosc     Stokes shift  Effective Stokes shift
          (cm**-1)     (cm**-1)                 (cm**-1)           (cm**-1)
-------------------------------------------------------------------------
    1:     21140.00   24535.85   0.074529    300.00             7091.70


-------------------------------------------------------------------------------------------------
                                BROADENING PARAMETETRS (cm**-1)
-------------------------------------------------------------------------------------------------
                 Intrinsic                                  Effective
 State   --------------------------  ------------------------------------------------------------
                                                 Sigma                          FWHM
         Gamma     Sigma     FWHM      --------------------------  --------------------------
                                             0K      298.15K   300.00K      0K      298.15K   300.00K
-------------------------------------------------------------------------------------------------
   1:     50.00    100.00    293.50     1125.34    1411.13   1413.57   2703.84   3376.75   3382.48
```

Note that the evaluation of the effective parameters is rather approximate and these values can noticeable deviate from those which can be directly deduced from the calculated spectra. However, such an information usually provides the proper order of magnitude of the effective vibronic broadening and Stokes shift. As indicated in the program output above, the effective bandshape has predominantly a Gaussian character which varies with the temperature so that $\sigma = 1125\,\mathrm{cm}^{-1}$ ($T = 0$ K) and $\sigma = 1414\,\mathrm{cm}^{-1}$ ($T = 300$ K). Indeed, as shown in Fig. 7.47 the absorption spectrum at $T = 300$ K can be well fitted using Gaussian lineshape with $\sigma = 1388\,\mathrm{cm}^{-1}$ (FWHM$= 3270\,\mathrm{cm}^{-1}$). One can see that at higher temperatures the deviation between the spectrum and its Gauss

fit becomes even smaller.

In molecules the normal distribution of the electronic transition energies in the ensemble would give rise to a Gaussian bandshape of the absorption band. However, the corresponding standard deviation is expected to be of the order of 100 cm^{-1}, whereby a typical Gaussian bandwidth of the order of 1000 cm^{-1} appears to result from unresolved vibronic progression. In general, this statement is supported by quantum chemical calculation of the model parameters. In principle the effective bandwidth parameters can also be used for characterization and assignement of individual electronic bands.

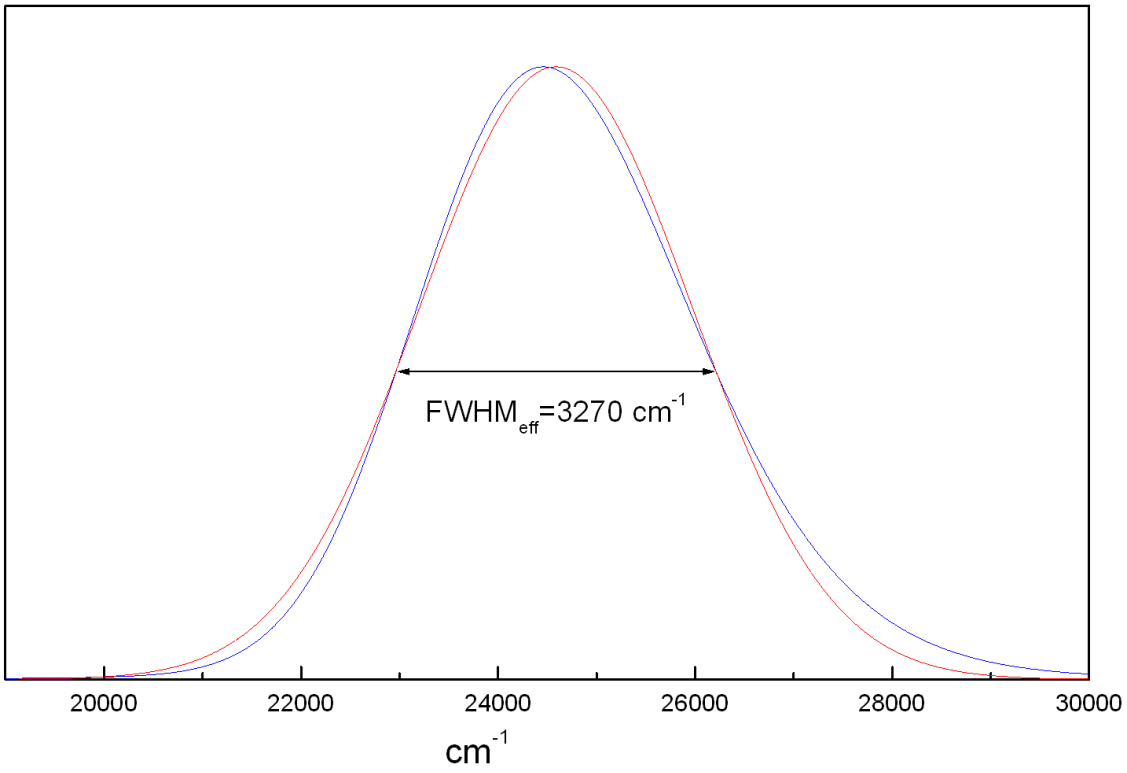


Fig. 7.47: Absorption spectrum (blue) for $T = 300$ K generated after `orca_asa` run on the file `example004.inp`. Red line represents the Gauss-fit of the calculated spectrum.

Example: Modelling of Absorption and Resonance Raman Spectra for the 1-$^1A_g \rightarrow$ 1-^1B_u Transition in *trans*-1,3,5-Hexatriene

The hexatriene molecule is characterized by 9 totally-symmetric normal modes which dominate vibrational structure in absorption and are active in rR spectra corresponding to the strongly dipole-allowed $1\,{}^1A_g \rightarrow 1\,{}^1B_u$ transition around 40000 cm^{-1} . Except for some peculiarities related to the neglect of normal mode rotations in the excited state the optical spectra are quite satisfactorily described by the IMDHO model.

The following input exemplifies simulation of absorption spectrum and rR spectra for an arbitrary predefined number of excitation energies.

```
#
# example005.inp
#
# input for simulation of absorption and resonance Raman spectra
# using experimental values of transition energy and displacement
# parameters corresponding to the strongly allowed 1-1Ag 1-1Bu transition
# in trans-1,3,5-hexatriene
#
```

(continues on next page)

(continued from previous page)

```
%sim
   Model IMDHO

   AbsRange     38000.0,   48000.0
   NAbsPoints   2000
   AbsScaleMode Rel

   # resonance Raman intensities will be calculated
   # for all vibrational states with excitation number
   # up to RamanOrder:
   RamanOrder  4

   # excitation energies (cm**-1) for which rR spectra will be calculated:
   RRSE 39500, 39800, 41400

   # full width half maximum of Raman bands in rR spectra (cm**-1):
   RRS_FWHM  10

   RSRange   0, 5000  # spectral range for simulation of rR spectra (cm**-1)
   NRRSPoints  5000   # number of points to simulate rR spectra (cm**-1)

end


$el_states
1
   1  39800.0      150.00    0.0     1.0   0.0   0.0


$vib_freq_gs
9
    1      354.0
    2      444.0
    3      934.0
    4     1192.0
    5     1290.0
    6     1305.0
    7     1403.0
    8     1581.0
    9     1635.0

$sdnc
9 1
                1
      1       0.55
      2       0.23
      3       0.23
      4       0.82
      5       0.485
      6       0.00
      7       0.085
      8       0.38
      9       1.32
```

After the `orca_asa` run the following files will be created:

- `example005abs.dat` contains the simulated absorption spectrum. It is shown in Fig. 7.48.
- `example005.o4.rrs.39500.dat`, `example005.o4.rrs.39800.dat` and `example005.o4.rrs.41400.dat` contain the simulated rR spectra for excitation energies at 39500, 39800 and 41400 cm^{-1}, respectively. The suffix "o4" stands for the order of Raman scattering specified in the input by keyword `RamanOrder=4`. The rR specta are shown in Fig. 7.49.

- `example005.o4.rrs.39500.stk`, `example005.o4.rrs.39800.stk` and `example005.o4.rrs.41400.stk` provide Raman shifts and intensities for each vibrational transition. Corresponding vibrational states are specified by the quantum numbers of excited modes.

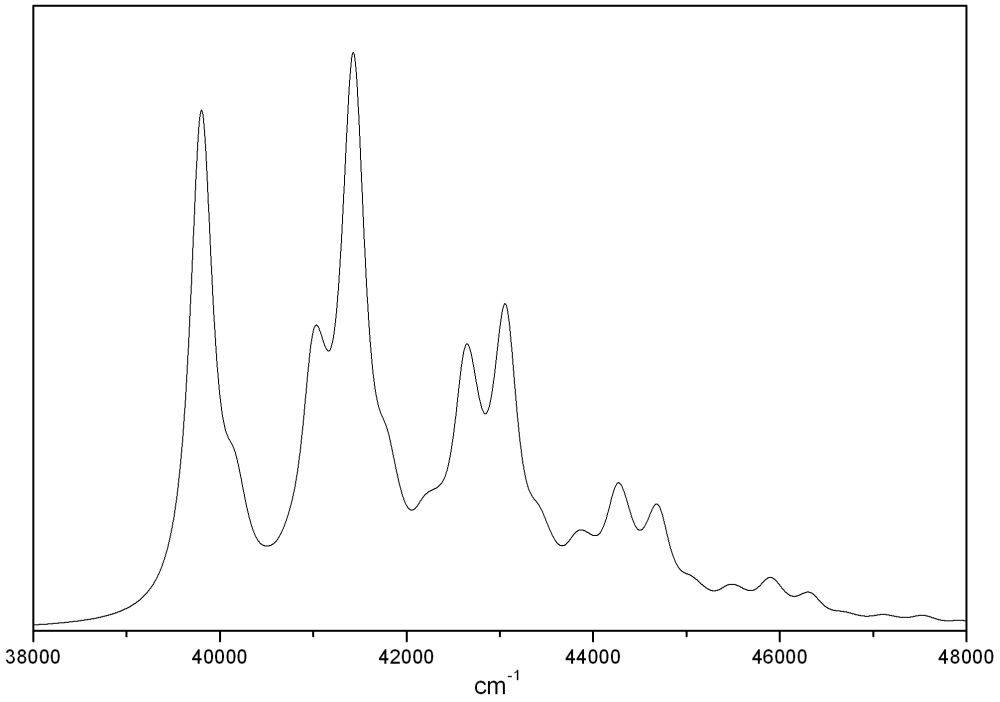


Fig. 7.48: Absorption spectrum corresponding to $1-^1A_g \rightarrow 1-^1B_u$ transition in *trans*-1,3,5-hexatriene generated after `orca_asa` run on the file `example005.inp`.

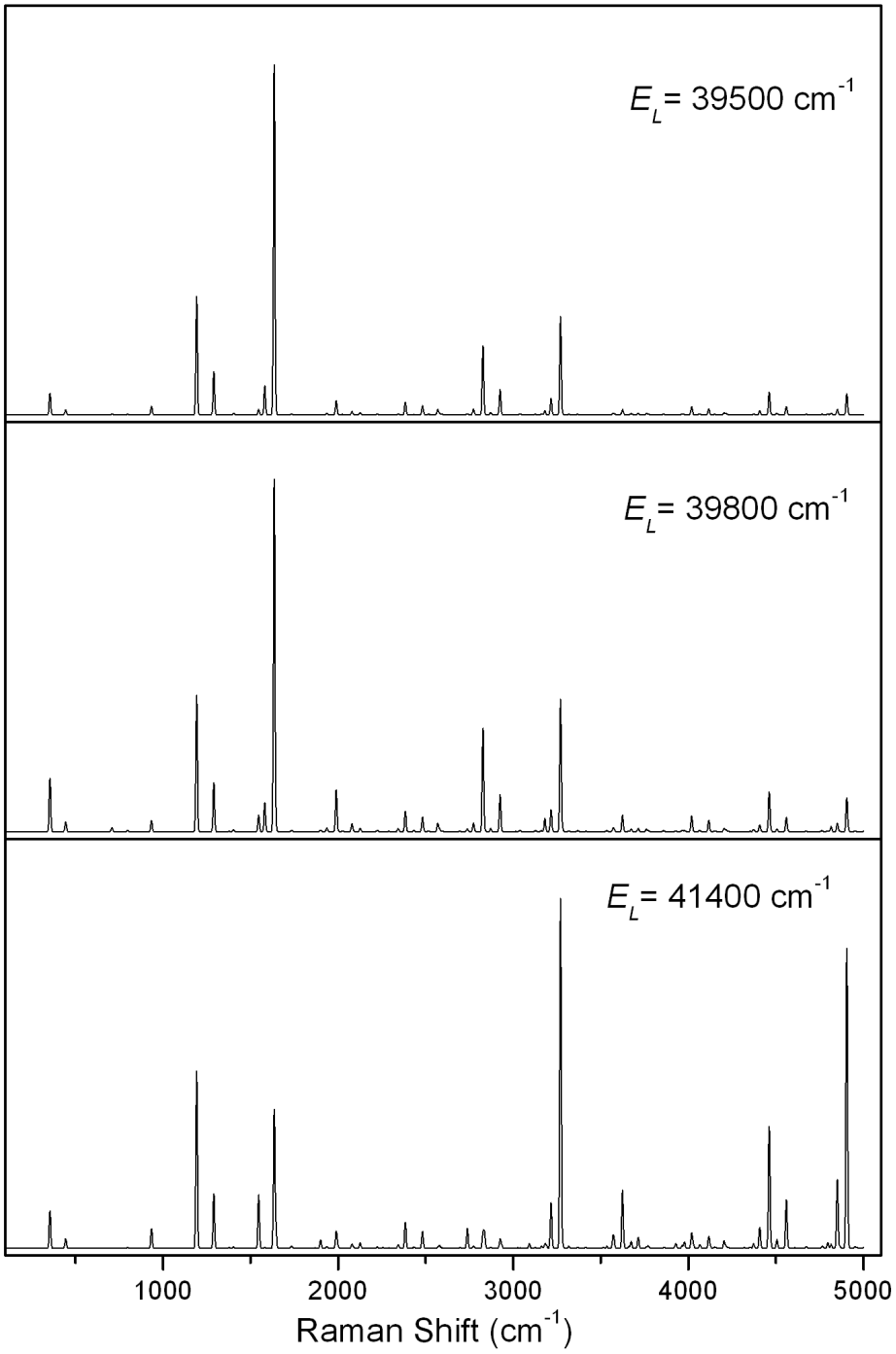


Fig. 7.49: Resonance Raman spectra for 3 different excitation energies which fall in resonance with $1\,^{-1}A_g \rightarrow 1\,^{-1}B_u$ transition in *trans*-1,3,5-hexatriene.

NOTE

- By default the program provides rR spectra on an arbitrary scale since only relative rR intensities within a single rR spectrum are of major concern in most practical cases. However, one can put rR spectra corresponding to different excitation energies on the same intensity scale by providing the keyword `RSISM=ASR`

in %sim block (RSISM – Raman Spectra Intensity Scaling Mode; ASR – All Spectra Relative). By default RSISM=SSR (SSR – Single Spectrum Relative) for which each rR spectrum is normalized so that the most intense band in it has intensity 1.0. The relative intensities of bands in rR spectra measured for different excitation energies can be compared if they are appropriately normalized relative to the intensity of a reference signal (e.g. Raman band of the solvent). We also keep in mind the possibility to extend our methodology in order to provide the absolute measure of rR intensities in terms of the full or differential cross-sections.

- Within the harmonic model, for a single electronic state neither relative rR intensities nor absorption bandshapes in the case of AbsScaleMode=Rel do depend on the values of the electronic transition dipole moment (unless it is precisely zero).

In the example above resonance Raman spectra have been generated for all vibrational transitions with total excitation number up to the value specified by the parameter RamanOrder. Its is also possible to make explicit specification of vibrational states corresponding to various fundamental, overtone and combination bands via the $rr_vib_states block. In such a case rR spectra involving only these vibrational transitions will be generated separately.

```
$rr_vib_states 5 # total number of vibrational transitions
 1
    modes 1
    quanta 1; # final vibrational state for the fundamental band corresponding to mode 1
 2
    modes 9
    quanta 1; # final vibrational state for the fundamental band corresponding to mode 9
 3
    modes 3, 4
    quanta 1, 1; # final vibrational state for the combination band involving single
                 # excitations in modes 3 and 4
 4
    modes 5
    quanta 3; # final vibrational state for the second overtone band corresponding to
              # mode 5
 5
     modes 1, 5,9
     quanta 1,2, 1; # final vibrational state for the combination band involving single
                    # excitations in modes 1 and 2, and double excitation in mode 5
```

Each vibrational transition is specified via the subblock which has the following structure:

```
k
          modes m1,m2,...mn
          quanta q1,q2,...qn;
```

This means that the k-th transition is characteriezed by excitation numbers q_i for modes m_i so that corresponding Raman shift is equal to $\nu = \sum q_i \nu_i$, where ν_i is vibrational frequency of the mode m_i.

After the orca_asa run the following files will be created in addition:

- example005.us.rrs.39500.dat, example005.us.rrs.39800.dat and example005.us.rrs.41400.dat contain the simulated rR spectra involving only vibrational transitions specified in the $rr_vib_states block, for excitations energies at 39500, 39800 and 41400 cm^{-1}, respectively. The suffix "us" stands for "User specified vibrational States".
- example005.us.rrs.39500.stk, example005.us.rrs.39800.stk and example005.us.rrs.41400.stk provide Raman shifts and intensities for each vibrational transition specified in the $rr_vib_states block.

Example: Modelling of Absorption Spectrum and Resonance Raman Profiles for the 1-^1A_g → 1-^1B_u Transition in *trans*-1,3,5-Hexatriene

The following example illustrates an input for simulation of absorption bandshape and resonance Raman profiles (RRP):

```
#
# example006.inp
#
# input for simulation of absorption and resonance Raman profiles
# using experimental values of transition energy and displacement
# parameters corresponding to the strongly allowed 1-1Ag 1-1Bu transition
# in trans-1,3,5-hexatriene
#

%sim
     Model IMDHO

     AbsRange     38000.0,   48000.0
     NAbsPoints   2000
     AbsScaleMode Rel

     RRPRange   38000.0, 48000.0 # spectral range for simulation of
                                 # rR profiles (cm**-1)

     NRRPPoints   2000 # number of points for simulation of rR profiles
     CAR 0.8

     RamanOrder 2
end

$el_states
1
   1  39800.0      150.00    0.0     1.0   0.0   0.0

$vib_freq_gs
9
    1      354.000000
    2      444.000000
    3      934.000000
    4     1192.000000
    5     1290.000000
    6     1305.000000
    7     1403.000000
    8     1581.000000
    9     1635.000000

$sdnc
9 1
                1
      1       0.55
      2       0.23
      3       0.23
      4       0.82
      5       0.485
      6       0.00
      7       0.085
      8       0.38
      9       1.32

$rr_vib_states 5 # total number of vibrational transitions
 1
```

(continues on next page)

(continued from previous page)

```
   modes 1
   quanta 1;
 2
   modes 9
   quanta 1;
 3
   modes 3, 4
   quanta 1, 1;
 4
   modes 5
   quanta 3;
5
    modes 1, 5,9
    quanta 1,2, 1;
```

The keyword `RamanOrder=2` will invoke generation of rR profiles for all vibrational transitions with total excitation number up to 2 in the range of excitation energies specified by the keywords `RRPRange` and `NRRPPoints`. Likewise, rR profiles for the vibrational states given in the `$rr_vib_states` block will be generated separately. Since in most cases only relative rR intensities are important, and one would be interested to compare absorption bandshape and shapes of individual rR profiles, the keyword `CAR = 0.8` is used to scale rR profiles for all vibrational transitions by a common factor in such a way that the ratio of the maximum of all rR intensities and the maximum of absorption band is equal to 0.8.

After the `orca_asa` run the following files will be created:

- `example006.abs.dat` contains the simulated absorption spectrum (Fig. 7.50).
- `example006.o1.rrp.dat` and `example006.o2.rrp.dat` contain rR profiles for vibrational transitions with total excitation numbers 1 and 2, respectively. RR profiles for all fundamental bands (from the file `example006.o1.rrp.dat`) are shown in Fig. 7.50.
- `example006.o1.info` and `example006.o1.info` contain specification of vibrational transitions with total excitation numbers 1 and 2, respectively, as well as corresponding Raman shifts.
- `example006.us.rrp.1.dat-example006.us.rrp.5.dat` contain rR profiles for vibrational transitions 1–5 specified in the `$rr_vib_states` block.

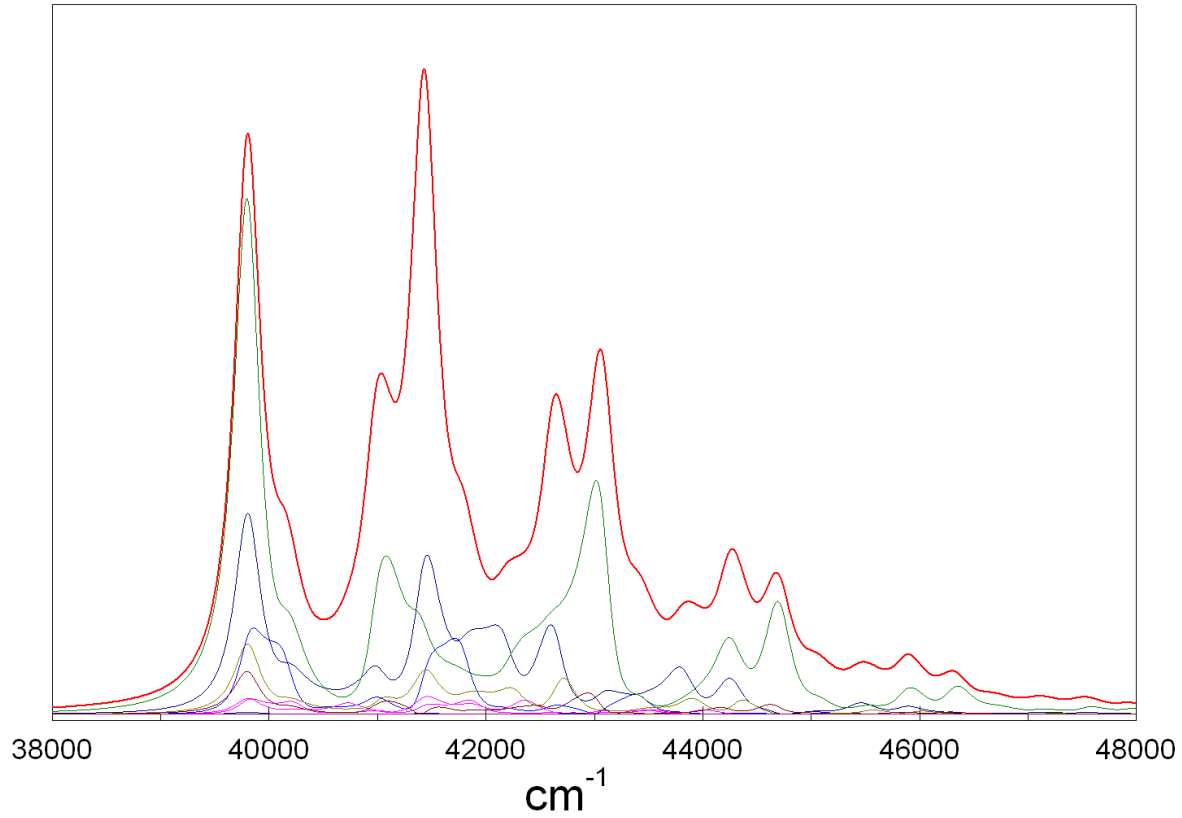


Fig. 7.50: Absorption spectrum and resonance Raman profiles of fundamental bands corresponding to $1\,^{-1}A_g \rightarrow 1\,^{-1}B_u$ transition in *trans*-1,3,5-hexatriene.

7.41.3 Fitting of Experimental Spectra

Example: Gauss-Fit of Absorption Spectrum

An absorption spectrum basically consists of a number of absorption bands. Each absorption band corresponds to a transition of the ground electronic state to an excited electronic state. In molecules such transitions are usually considerably broadened. In many cases there will be overlapping bands and one would need to deconvolute the broad absorption envelope into contributions from individual transitions. Within the "Simple model" the `orca_asa` program enables fit of an absorption spectrum with a sum of standard lineshape functions (Gaussian, Lorentzian) or more general Voigt functions. In most cases, one simply performs a "Gauss-Fit". That is, it is assumed that the shape of each individual band is that of a Gaussian function. Then one applies as many (or as few) Gaussians as are necessary for an accurate representation of the absorption envelope. In order to explain the fitting procedures within the "Simple model" let us consider an experimental absorption spectrum in Fig. 7.51:

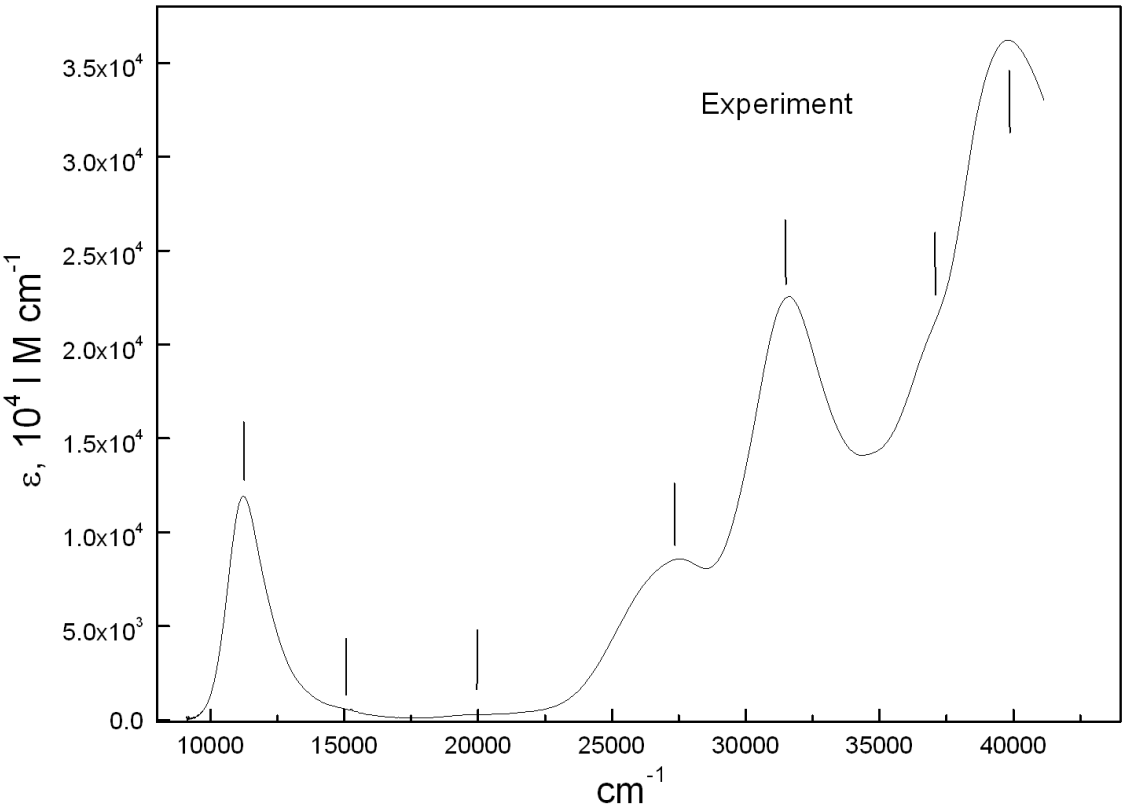


Fig. 7.51: Experimental absorption spectrum. Bars indicate transition energies which were used for the initial guess in the input for spectral fitting.

As shown in Fig. 7.51 one can identify roughly 7 electronic bands. The initial estimates of transition energies corresponding to the maxima and shoulders in the absorption spectrum (indicated by bars in Fig. 7.51) and rather approximate values of inhomogeneous broadening and transition dipole moment components are specified in the `$el_states` block of the input file for the spectral fitting:

```
# example007.inp
#
# Input file for fitting of experimental absorption spectrum
#

%sim
        model Simple
end

%fit
       Fit true               # Global flag to turn on the fit
       AbsFit true            # Flag to include absorption into the fit
       method Simplex
       WeightsAdjust true

       AbsRange 0.0, 100000.0 # absorption spectral range to be included in the fit;
                              # in the present case all experimental points
                              # will be included


       AbsName "absexp.dat"   # name of the file containing experimental
                              # absorption spectrum in a simple two-column
                              # ASCII format
```

(continues on next page)

(continued from previous page)

```
      ExpAbsScaleMode Ext    # This keyword indicates that the experimental
                             # absorption intensity is given in terms of
                             # the extinction coefficient. This is important
                             # for the proper fitting of transition dipole
                             # moments and oscillator strengths

      NMaxFunc 10000   # maximum number of function evaluations in simplex
                       # algorithm

      MWADRelTol 1e-5  # Relative Tolerance of the Mean Weighted Absolute
                       # Difference  (MWAD) function which specifies the
                       # convergence criterion


      E0Step 500.00    # initial step for the transition energies
                       # in the simplex fitting

      TMStep 0.5       # initial step for the transition dipole moments
                       # in the simplex fitting

      E0SDStep 500.0   # initial step for the inhomogeneous linewidth (Sigma)
                       # in the simplex fitting


end

# ! Parameters specified in the $el_states block
#   are used as initial guess in the fit

#-------------------------------------------------------------------------------
#  Transition        Gamma    Sigma  Transition Dipole Moment (atomic unit)
#  Energy (cm**-1)  (cm**-1) (cm**-1)      Mx          My        Mz
#-------------------------------------------------------------------------------

$el_states
7
    1 11270           0.0     1000.00    1.0000     0.0000     0.0000
    2 15100           0.0     1000.00    1.0000     0.0000     0.0000
    3 20230           0.0     1000.00    1.0000     0.0000     0.0000
    4 27500           0.0     1000.00    1.0000     0.0000     0.0000
    5 31550           0.0     1000.00    1.0000     0.0000     0.0000
    6 37070           0.0     1000.00    1.0000     0.0000     0.0000
    7 39800           0.0     1000.00    1.0000     0.0000     0.0000


# the integer values specified in $el_states_c block indicate parameters
# in the $el_states block to be varied
$el_states_c
7
    1    1             0          1          1         0         0
    2    2             0          2          2         0         0
    3    3             0          3          3         0         0
    4    4             0          4          4         0         0
    5    5             0          5          5         0         0
    6    6             0          6          6         0         0
    7    7             0          7          7         0         0
```

The functionality of the constraint block `$el_states_c` should be understood as follows: 1) 0 flag indicates that the corresponding parameter in the `$el_state` block will not be varied in the fitting; 2) if the number corresponding to a certain parameter coincides with the number of the corresponding electronic state this parameter will be varied independently. Thus, the block `$el_states_c` in the input indicates that all transition energies, inhomo-

geneous linewidths and x-components of the transition electric dipole moment will be varied independently, while homogeneous linewidths, y- and z-components of the transition dipole moment will be fixed to their initial values.

The following considerations are important:

- Since in conventional absorption spectroscopy one deals with the orientationally averaged absorption cross-section, the signal intensity is proportional to the square of the transition electric dipole moment $|\mathbf{M}|^2$. Thus, the intensities do not depend on the values of the individual components of $\mathbf{M}$ as long as $|\mathbf{M}|^2 = \text{const}$. Therefore, we have allowed to vary only M_x components. Otherwise there can be problems in convergence of the fitting algorithm.
- The sum of the weights of experimental points which enter the mean absolute difference function employed in the minimization is always kept equal to the number of experimental points. In the case of equidistant experimental photon energies all weights are assumed to be equal. However, in experimental electronic spectra the density of spectral points can increase significantly upon going from high- to low-energy spectral regions, which is due to the fact that experimental absorption spectra are initially acquired on the wavelength scale. In such a case the quality of the fit can be noticeably biased towards low-energy spectral region. Therefore, it is advisable to adjust relative weights of experimental points according to the their density which is controlled by the keyword `WeightsAdjust` in the `%fit` block. Although this parameter is not crucial for the present example, in general, it will provide a more balanced fit.
- The parameters `E0Step`, `TMStep`, `E0SDStep` in the `%fit` block specify the initial dimension of the simplex in the space of $E_T, \mathbf{M}, \sigma$ and should roughly correspond to the expected uncertainty of initial guess on these parameters in the `$el_states` block relative to their actual values. The quality of the fit can noticeably deteriorate if the parameters specifying initial steps are too low or too high.

The fit run of `orca_asa` on file `example007.inp` will converge upon approximately 3600 function evaluations (for `MWADRelTol=1e-5`). The results of the fit will be stored in file `example007.001.inp` which has the same structure as the input file `example007.inp`. Thus, if the fit is not satisfactory and/or it is not fully converged it can be refined in a subsequent `orca_asa` run upon which file `example007.002.inp` will be created, and so on. Some model parameters in intermediate files can be be additionally modified and/or some constraints can be lifted or imposed if so desired. The output file `example007.001.inp` will contain fitted model parameters stored in the `$el_states` block:

```
$el_states
7
   1 11368.24     0.00  732.50     1.6290   0.0000   0.0000
   2 15262.33     0.00  495.17    -0.2815   0.0000   0.0000
   3 19500.08     0.00  1023.39     0.2300   0.0000   0.0000
   4 26969.01     0.00  1832.30     1.4089   0.0000   0.0000
   5 31580.41     0.00  1440.87     1.8610   0.0000   0.0000
   6 35769.07     0.00  1804.02     1.5525   0.0000   0.0000
   7 39975.11     0.00  1909.38     2.4745   0.0000   0.0000
```

The overall quality of the fit is determined by the parameter MWAD which upon convergence reaches the value of $\approx$0.009 (MWAD stands for Mean Weighted Absolute Difference).

After the `orca_asa` run files `absexp.fit.dat` and `absexp.fit.as.dat` will be created. Both files contain the experimental and fitted spectra which are shown in Fig. 7.52 . In addition, the file `absexp.fit.as.dat` will contain individual contributions to the absorption spectrum corresponding to different excited states.

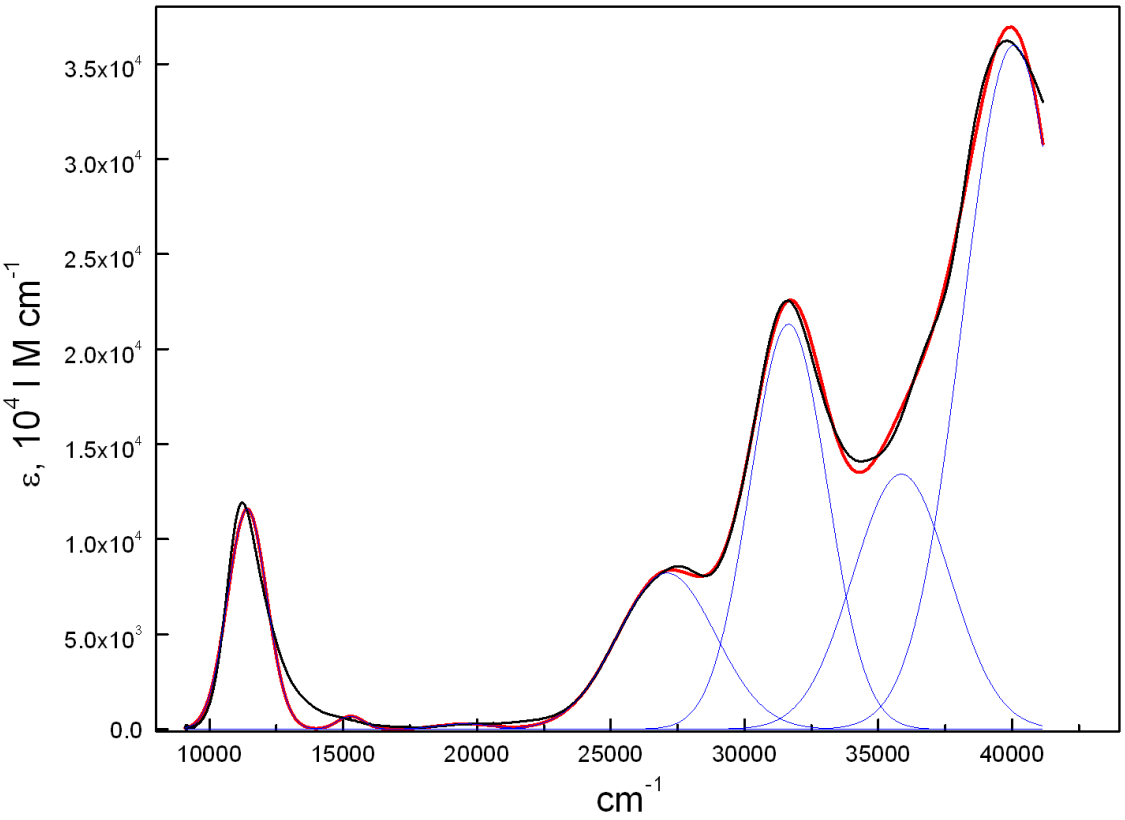


Fig. 7.52: Comparison of the experimental (black curve) and fitted (red) absorption spectra corresponding to the fit run of `orca_asa` on the file `example007.inp`. Blue curves represent individual contributions to the absorption spectrum from each state.

Since there is a noticeable discrepancy between the fitted and experimental spectra around 13000 cm^{-1} (Fig. 7.52) it is worthwhile to refine the fit after adding parameters for a new state in the file `example007.001.inp`:

```
$el_states
8
   1 11368.24     0.00  732.50     1.6290   0.0000   0.0000
... ... ...
   8 13280.00     0.00  1000.00     1.000   0.0000   0.0000


$el_states_c
8
   1     1         0        1          1        0        0
... ... ...
   8     8         0        8          8        0        0
```

Actually, the character of the discrepancy in the present case is very similar to that in Fig. 7.49 (section *Example: Modelling of Effective Broadening, Effective Stokes Shift and Temperature Effects in Absorption and Fluorescence Spectra within the IMDHO Model*) where a vibronically broadened absorption spectrum was fitted with a Gaussian lineshape. Thus, the poor fit in the region around 1300 cm^{-1} is most likely due to the essentially asymmetric character of the vibronic broadening rather than to the presence of another electronic band.

As shown in Fig. 7.53 the refined fit leads to much better agreement between the experimental and fitted absorption spectra (MWAD=0.0045).

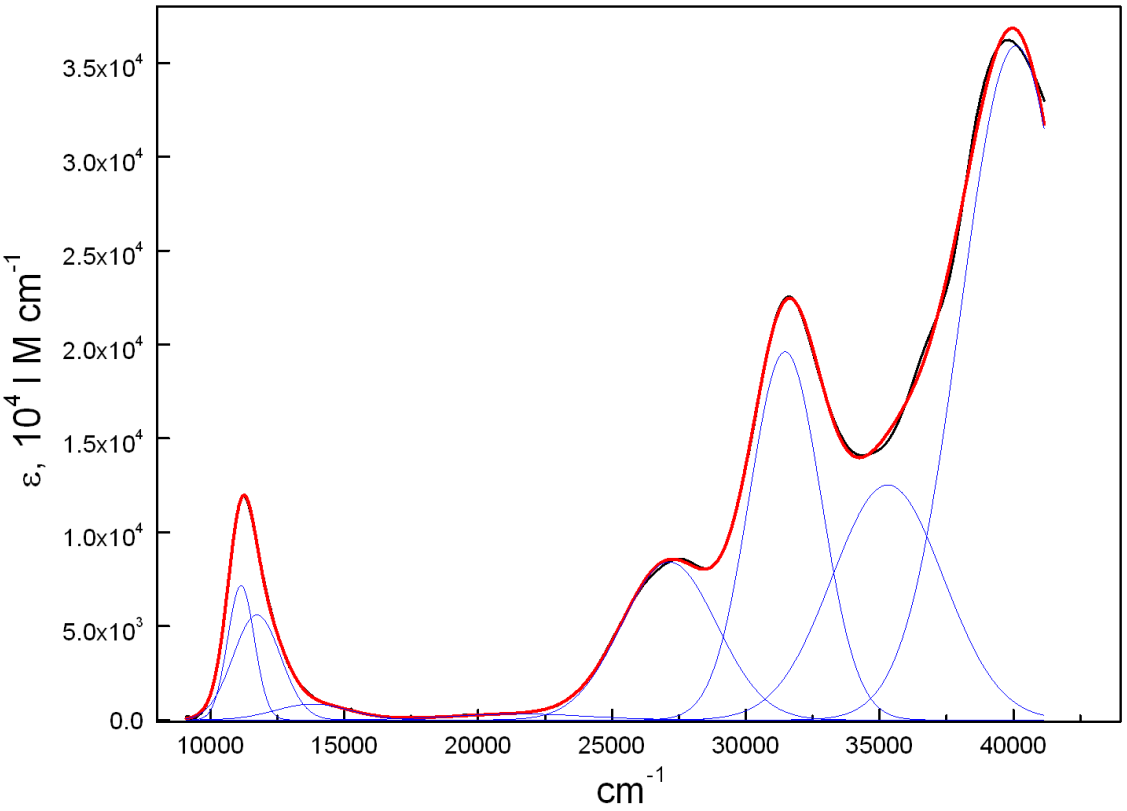


Fig. 7.53: Comparison of the experimental (black) and fitted (red) absorption spectra corresponding to the fit run of `orca_asa` on the file `example007.001.inp`. Blue curves represent individual contributions to the absorption spectrum from each state.

Due to some peculiarities of the simplex algorithm for function minimization, you can still refine the fit by rerunning `orca_asa` on the file `example007.002.inp`! This leads to an even lower value of the parameter MWAD= 0.0038, and therefore to better agreement of experimental and fitted spectra (even though the previous run has been claimed to be converged).

It is also possible to perform a fit using the same value of inhomogeneous linewidth for all electronic states. For this purpose one needs to choose as a guess the same linewidth parameters in the `$el_states` block:

```
$el_states
8
   1 11118.58     0.00  1000.0      1.0687   0.0000   0.0000
   2 13673.38     0.00  1000.0     -0.5530   0.0000   0.0000
   3 21267.40     0.00  1000.0      0.3675   0.0000   0.0000
   4 27024.71     0.00  1000.0      1.4041   0.0000   0.0000
   5 31414.74     0.00  1000.0      1.7279   0.0000   0.0000
   6 35180.77     0.00  1000.0      1.6246   0.0000   0.0000
   7 39985.52     0.00  1000.0      2.5708   0.0000   0.0000
   8 11665.01     0.00  1000.0      1.2332   0.0000   0.0000
```

In addition the constraint block should be modified as follows:

```
$el_states_c
8
   1     1        0        1        1        0        0
   2     2        0        1        2        0        0
   3     3        0        1        3        0        0
   4     4        0        1        4        0        0
   5     5        0        1        5        0        0
```

(continues on next page)

(continued from previous page)

```
  6     6        0        1        6        0        0
  7     7        0        1        7        0        0
  8     8        0        1        8        0        0
```

The constraint parameters for the inhomogeneous broadening were chosen to be 1, which means that formally σ_1 corresponding to the first state is varied independently while the linewidths $\{\sigma_i\}$ for other bands are varied in such a way that the ratios σ_i/σ_1 are kept fixed to their initial values, whereby the same linewidth parameter will be used for all states.

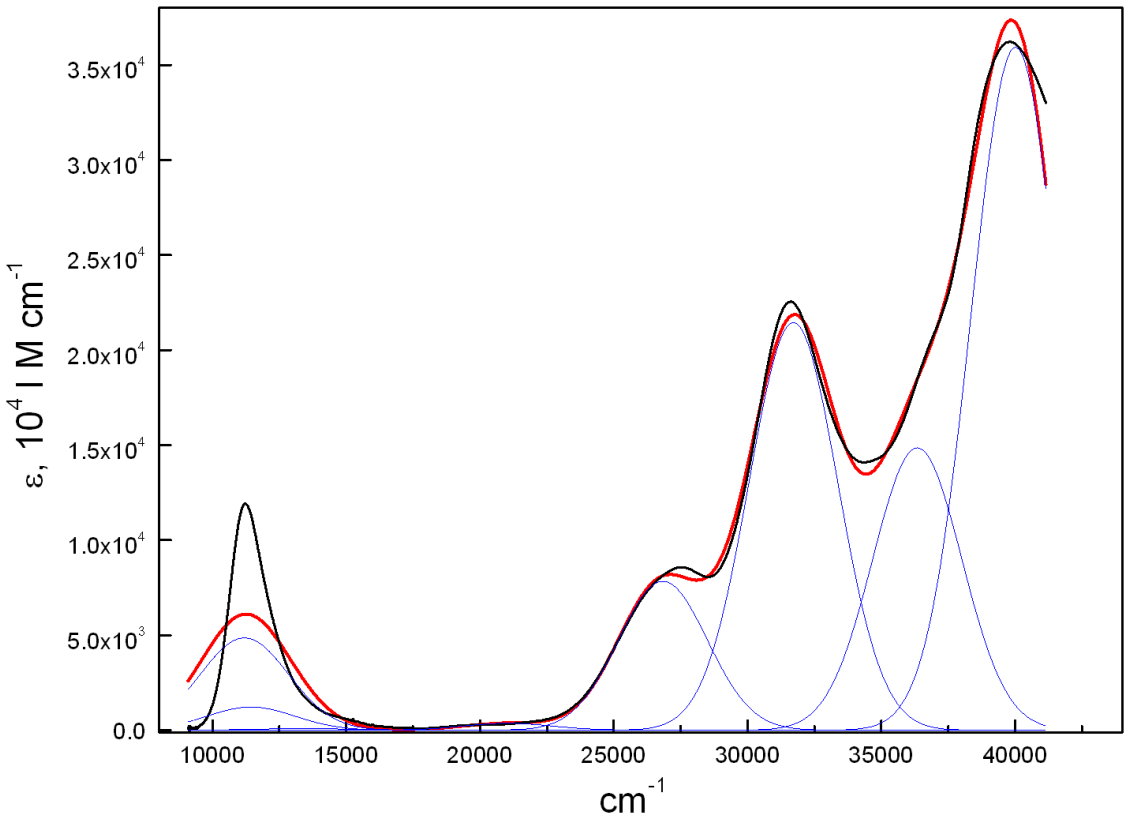


Fig. 7.54: Comparison of the experimental (black) and fitted (red) absorption spectra corresponding to the fit run of `orca_asa` on the file `example007.002.inp` in which equal broadening was assumed for all electronic bands. Blue curves represent individual contributions to the absorption spectrum from each state.

One can see (Fig. 7.54) that the assumption of equal linewidths for all electronic bands leads to a rather pronounced deterioration of the quality of the fit in the low-energy spectral range (MWAD=0.017). Apparently, this discrepancy can be fixed assuming more electronic states at higher energies.

NOTE

- The homogeneous linewidth parameters can also be included in the fit in a similar way. However, one can see that in most cases they appear to be much smaller than corresponding Gaussian linewidth parameters.
- Gauss-fit of absorption spectra is coventionally performed assuming the same linewidth parameters for all bands. However, since a large portion of Gaussian broadening is mainly due to the unresolved vibronic structure in the spectra which can significantly vary depending on the nature of transition, the assumption of unequal Gaussian bandwidths seems to be a physical one.

Example: Fit of Absorption and Resonance Raman Spectra for 1-$^1A_g \rightarrow$ 1-^1B_u Transition in *trans*-1,3,5-Hexatriene

Below we provide an example of the fit of the lineshape parameters and $\{\Delta_m\}$ corresponding to the strongly dipole-allowed 1-$^1A_g \rightarrow$1-^1B_u transition in hexatriene. It is known that the most intense bands in rR spectra correspond to the most vibronically active in absorption spectrum. For the IMDHO model this correlation is determined by the values of $\{\Delta_m\}$. Thus, the larger Δ, the larger is the rR intensity of a given mode and the more pronounced is the progression in the absorption spectrum corresponding to this mode. In principle, if all vibrational transitions in absorption are well resolved it is possible to determine $\{\Delta_m\}$ by a fit of the absorption spectrum alone. In practice this task is ambiguous due to the limited resolution of the experimental absorption spectra. The observation of a rR spectrum enables the identification of the vibrational modes that are responsible for the progression in the absorption spectrum, as well as a quantitative analysis in terms of $\{\Delta_m\}$. The file `example006.inp` provides a brute-force example on how to approach the fit employing the minimal possible experimental information: 1) An absorption spectrum; 2) relative rR intensities of fundamental bands for a given excitation energy. The rR spectrum upon the excitation in resonance with the 0-0 vibronic band at 39809 cm^{-1} is shown in Fig. 7.43.

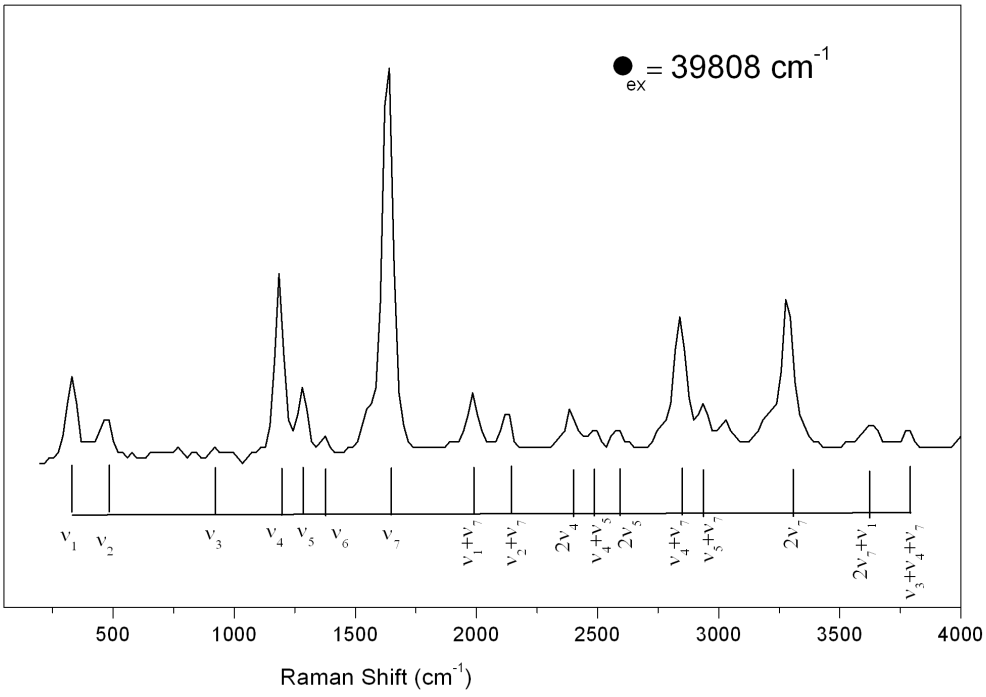


Fig. 7.55: Experimental Resonance Raman spectrum corresponding to 1-$^1A_g \rightarrow$ 1-^1B_u transition in *trans*-1,3,5-hexatriene.

The experimental rR spectrum has enabled the identification of seven vibrational modes that give rise to the most intense resonance Raman bands. Therefore, they are expected to have the largest excited-state displacements and the most pronounced effect on the vibrational structure of the absorption spectrum. Their vibrational frequencies have been entered as input for the fit as shown below:

```
#
# example008.inp
#
# Input for fit of absorption and resonance Raman spectra
# corresponding to the strongly allowed 1-1Ag 1-1Bu transition
# in 1,3,5 trans-hexatriene.
```

(continues on next page)

(continued from previous page)

```
#
# Parameters to be varied:
# 1) adiabatic minima transiton energy
# 2) homogeneous linewidth (Gamma)
# 3) dimensionless normal coordinate displacements of the
#    excited-state origin
#

%sim
      Model IMDHO
end


%fit

    Fit true     # boolean parameter to switch on the fit

    # boolean parameter to include experimental absorption
    # spectrum in the fit:
    AbsFit true

    # boolean parameter to include experimental rR spectra
    # specified in $rrs_exp block in the fit:
    RRSFit true

    AbsExpName "hex-abs.dat" # name of the file with experimental absorption
                             # spectrum

    # the following value of keyword ExpAbsScaleMode
    # indicates that only the shape of absorption band
    # but not its total intensity will be accounted in the fit:
    ExpAbsScaleMode Rel

    # the weight of absorption relative to the total weight of
    #  rR intensities in the difference function to be minimized:
    CWAR  5.0

    NMaxFunc  1000   # maximum number of function evaluations in simplex
                     # algorithm

    MWADRelTol 1e-4  # Relative Tolerance of the Mean Weighted Absolute
                     # Difference  (MWAD) function which specifies the
                     # convergence criterion
    SDNCStep  1.0

end


# The values specified in $el_states block serve as initial guess in the fit
$el_states
1
   1  40000.0      200.00    0.0     1.0   0.0   0.0

# the integer values specified in $el_states_c block indicate parameters
# in $el_states block to be varied
$el_states_c
1
   1     1            1       0       0     0     0

# 7 totally symmetric vibrations which give rise to the most
# intense bands in the rR spectra are included into analysis.
```

(continues on next page)

(continued from previous page)

```
# Experimental values of vibration frequencies are given:
$vib_freq_gs
7
    1      354.0
    2      444.0
    3      934.0
    4     1192.0
    5     1290.0
    6     1403.0
    7     1635.0

# Initial guess for the values of dimensionless normal
# coordinate displacements of the excited-state origin
$sdnc
7 1
                1
      1       0.0
      2       0.0
      3       0.0
      4       0.0
      5       0.0
      6       0.0
      7       0.0

# the integer values specified in $sdnc_c block indicate parameters
# in $sdnc block to be varied
$sdnc_c
7 1
              1
      1         1
      2         2
      3         3
      4         4
      5         5
      6         6
      7         7


# specification  of vibrational transitions and their intensities
# in experimental rR spectra:
$rrs_exp
1         # number of rR spectra
1 1       # start of the block specifying the 1st rR spectrum
    Ex 39809.0 # excitation energy for the first rR spectrum
    NTr  7     # number of vibrational transitions for which intensities are
               # provided
    1
       int   10.0  1.0
       modes  1
       quanta 1;
    2
       int   5.0  1.0
       modes  2
       quanta 1;
    3
       int   1.5  1.0
       modes  3
       quanta 1;
    4
       int   21.0  1.0
       modes  4
```

(continues on next page)

(continued from previous page)

```
      quanta 1;
   5
      int   7.5  1.0
      modes  5
      quanta 1;
   6
      int   2.0  1.0
      modes  6
      quanta 1;
   7
      int   46.0  1.0
      modes  7
      quanta 1;
```

The input of rR intensities for an arbitrary number of excitation energies follows the keyword `$rrs_exp` block:

```
$rrs_exp
1         # number of rR spectra
1 1
```

The first "1" in the last line denotes the number of the rR spectrum for which specification starts below. If the second number is the same as the number of the spectrum, then it means that only relative intensities for the first rR spectrum are meaningful in the fit. If several spectra are given in the input then the second number may have a different value, e.g.:

```
$rrs_exp
3         # number of rR spectra
1 2
...
```

This input is to be interpreted as indicating that 3 rR spectra are provided and the relative intensities for the first spectrum are given on the same scale as the second one that will be accounted for in the fit. The value of the excitation energies and the number of vibrational transitions specified are indispensable within the blocks specifying intensities for each rR spectrum.

Following the number of vibrational transitions given by the keyword `NTr` one has to specify each vibrational transition and its intensity. Thus, in the present case there are seven subblocks with the following structure:

```
k    int   I  W
          modes m1,m2,...mn
          quanta q1,q2,...qn;
```

This means that the k-th transition has intensity I and weight W in the mean absolute difference function that is used for the minimization (W is an optional parameter). The following 2 lines specify the vibrational transitions by providing excitation numbers q_i for modes m_i so that the corresponding Raman shift is equal to $\nu = \sum_{q_i \nu_i}$, where ν_i is vibrational frequency of the mode m_i.

The parameters that are to be varied are specified within the constraint blocks `$el_states_c` and `$sdnc_c`. Both blocks have the same structure and number of parameters as `$el_states` and `$sdnc`, respectively. A parameter from the `$el_states` block is supposed to be independently varied if its counterpart from the `$el_states_c` block is equal to the number of the electronic state. Likewise, a parameter from the `$sdnc` block is supposed to be independently varied if its counterpart from the `$sdnc_c` block is equal to the number of the normal mode. Model parameters that are set to 0 in the corresponding constraint blocks are not varied in the fit. The values of the following parameters may be important for the quality of the fit:

- `CWAR` in the `%fit` block specifies the weight of absorption relative to the weight of rR intensities in the difference function to be minimized. If this parameter was not specified the fit would be almost insensitive to the rR intensities in the input, since typically the number of experimental absorption points is much larger than the number of rR transitions in the input. In most cases the value of `CWAR` in the range 1.0–5.0 is a good choice since the error in the measured experimental intensity is expected to be much smaller for absorption

than for resonance Raman.

- SDNCStep in the %fit block specifies the initial dimension of the simplex in the space of $\{\Delta_m\}$ and should roughly correspond to the expected uncertainty of initial guess on $\{\Delta_m\}$ in the $sdnc block compared to their actual values. You can notice in the present example that if this parameter is too large (>2.0) or too small (<0.4) the quality of the fit may significantly deteriorate
- Although the default initial dimensions of the simplex have reasonable values for different types of parameters it may turn out to be helpful in some cases to modify the default values:

```
FREQGStep 10.0  # ground-state vibrational frequencies
          FREQEStep 10.0  # excited-state vibrational frequencies
          E0Step    300.0 # transition energies
          SSStep     20.0 # Stokes shift
          TMStep      0.5 # electronic transition dipole moment
          GammaStep  50.0 # homogeneous linewidth
          E0SDStep   50.0 # inhomogeneous linewidth
          SDNCStep 1.0 # origin shift along dimensionless normal coordinate
```

The fit run of orca_asa on the file example008.inp will converge upon approximately 700 function evaluations (for MWADRelTol=1e-4). The results of the fit will be stored in file example008.001.inp which has the same structure as the input file example008.inp. Thus, if the fit is not satisfactory and/or it is not fully converged it can be refined in subsequent orca_asa run upon which file example008.002.inp will be created, and so on. Some model parameters in intermediate files can be be additionally modified and/or some constraints can be lifted if so desired. The output file example008.001.inp will contain fitted displacement parameters $\{\Delta_m\}$stored in the $sdnc block:

```
$sdnc
7 1
                    1
      1       0.675000
      2      -0.194484
      3      -0.217527
      4       0.811573
      5       0.529420
      6      -0.149991
      7       1.314915
```

In the present example, these parameters are actually in very close agreement with those published for the hexatriene molecule!

The overall quality of the fit is determined by the parameter MWAD which upon convergence reaches the value of ≈0.027. The fitted rR intensities are presented in the commented lines next to the experimental rR intensities in file example008.001.inp:

```
$rrs_exp
1
   1   1 3.495285e+001
   Ex    39809.00
   NT   7
      1
        Int 10.0 1.0  # simulated intensity: 1.000982e+001
        modes  1
        quanta 1;
      2
        Int 5.0 1.0   # simulated intensity: 8.976285e-001
        modes  2
        quanta 1;
      3
        Int 1.5 1.0   # simulated intensity: 1.255880e+000
        modes  3
        quanta 1;
```

(continues on next page)

(continued from previous page)

```
4
  Int 21.0 1.0  # simulated intensity: 1.761809e+001
  modes  4
  quanta 1;
5
  Int 7.5 1.0   # simulated intensity: 7.499749e+000
  modes  5
  quanta 1;
6
  Int 2.0 1.0   # simulated intensity: 6.014466e-001
  modes  6
  quanta 1;
7
  Int 46.0 1.0  # simulated intensity: 4.600071e+001
  modes  7
  quanta 1;
```

The file hex-abs.fit.dat will contain the experimental and fitted absorption spectra in ASCII format which can be plotted in order to visualize the quality of absorption fit (Fig. 7.56).

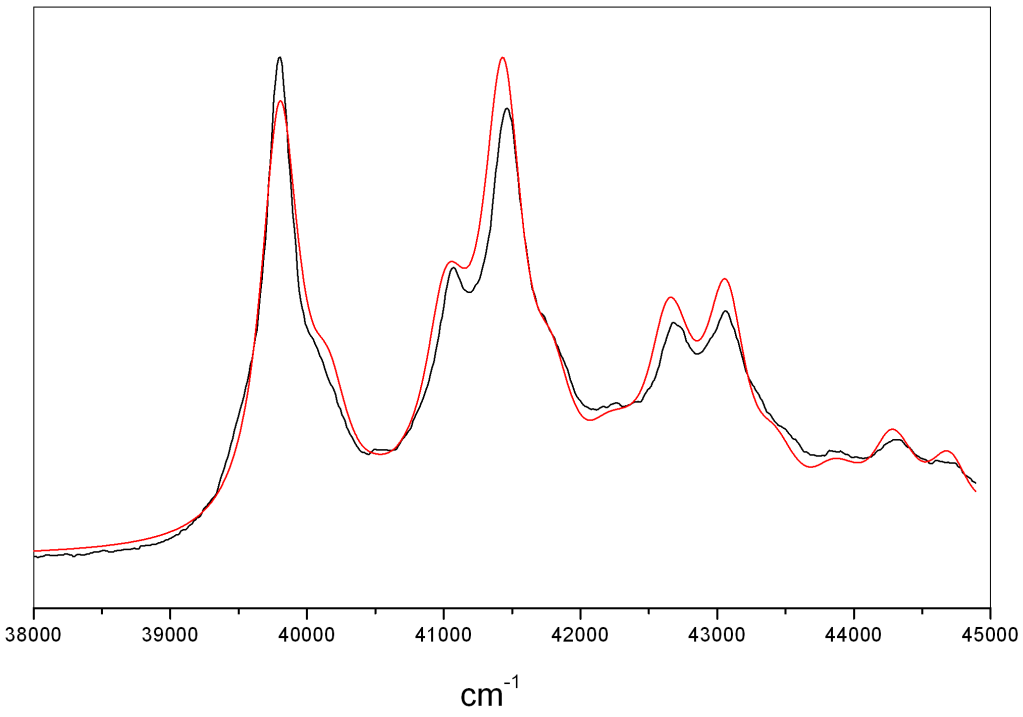


Fig. 7.56: Experimental (black) and fitted (red) absorption spectrum corresponding to 1-$^1A_g \rightarrow$ 1-^1B_u transition in 1,3,5 *trans*-hexatriene.

NOTE

- The more experimental rR intensities are included in the analysis the more reliable is the fit. In principle it is possible to obtain fully consistent results even if only a limited number of vibrational transitions is provided. However, in such a case it is desirable to include into analysis at least a single Raman transition involving the mode for which Δ is to be determined.

- The quality of the fit can be improved if the IMDHOFA model is invoked and excited-state vibrational frequencies are allowed to vary.
- Due to the initial guess and dimension of the simplex, as well as some peculiarities of the simplex algorithm for function minimization, you can still refine the fit by rerunning `orca_asa` on file `example008.001.inp` that may lead to an even lower value of the parameter `MWAD = 0.021`, and therefore to better agreement of experimental and fitted spectra (even though the previous run has been claimed to be converged).
- In this respect it appears to be wise to perform the fit in 3 steps:
  1. Fit the preresonance region below the 0-0 vibronic band with a single Lorentzian band, from which the adiabatic transition energy E_0, and homogeneous linewidth Γ are obtained. The range for fit of the absorption spectrum can be specified by the `AbsRange` keyword in the `%fit` block.
  2. Fix E_0 and Γ, and optimize $\{\Delta_m\}$ fitting the entire spectral range and rR intensities.
  3. Lift constraints on E_0 and Γ, and reoptimize simultaneously all parameters.

Example: Single-Mode Fit of Absorption and Fluorescence Spectra for 1-$^1A_g \rightarrow$ 1-$^1B_{2u}$ Transition in Tetracene

In this section we provide an example and discuss the most important aspects of joint fit of fluorescence and absorption spectra. Fig. 7.57 displays the experimental emission and absorption spectra corresponding to 1-$^1A_g \rightarrow$ 1-$^1B_{2u}$ transition in tetracene.

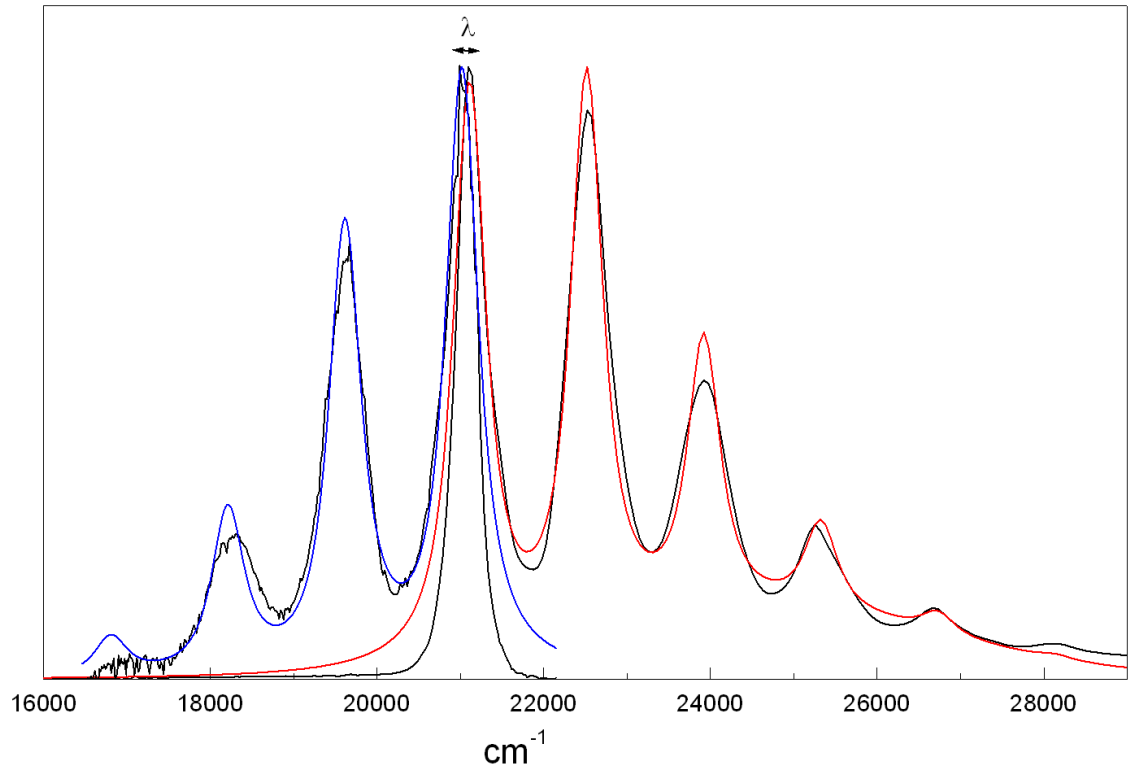


Fig. 7.57: Deconvoluted absorption (red) and fluorescence (blue) spectra of tetracene in cyclohexane upon the assumption of a single vibronically active mode. The black solid lines represent experimental spectra.

Both spectra show pronounced effective vibrational progressions that are dominated by 3 and 5 peaks, respectively. As can be shown on the basis of quantum chemical calculations this progression has essentially multimode character. However, the experimental spectra can be well fitted under the assumption of a single vibronically active mode. The input has the following structure:

```
#
# example009.inp
#
# Parameters to be varied:
# 1) adiabatic minima transition energy
# 2) homogeneous and inhomogeneous linewidths
# 3) normal mode frequency and corresponding dimensionless displacement of the
#    excited-state origin
#

%sim
      Model IMDHO
      EnInput E0 # we assume adiabatic minima separation energies
end

%fit
    Fit true       # global flag to turn on the fit
    AbsFit true    # flag to include absorption spectrum into the fit
    FlFit true     # flag to include fluorescence spectrum into the fit

    WeightsAdjust true

    AbsRange 19000.0, 28000.0 # spectral range for absorption
                              # which will be included into the fit

    FlRange  17800.0, 22300.0 # spectral range for absorption
                              # which will be considered in the fit

    AbsName "absexp.dat" # name of the file containing experimental
                         # absorption spectrum in a simple two-column
                         # ASCII format
    FlName  "flexp.dat"  # name of the file with experimental fluorescence spectrum

    ExpAbsScaleMode Rel  # flags indicating that only relative shapes of the
    ExpFlScaleMode Rel   # absorption and fluorescence bands will be fitted.


    CWAF   1.000 # important parameter to have a balanced relative quality of fit
                 # of fluorescence and absorption

    NMaxFunc 10000  # maximum number of function evaluations in simplex
                    # algorithm

    MWADRelTol= 0.0001  # Relative Tolerance of the Mean Weighted Absolute
                        # Difference  (MWAD) function which specifies the
                        # convergence criterion

    TMStep 0.5       # initial step for the transition dipole moments
                     # in the simplex fitting

    E0SDStep 500.0   # initial step for the inhomogeneous linewidth (Sigma)

    FREQGStep 100.00 # initial step for the vibrational frequencies

    E0Step 1000.0    # initial step for the transition energies

    SSStep 10.0      # initial step for the Stokes shift

    GammaStep 100    # initial step for the homogeneous linewidth

    SDNCStep 0.5    # initial step for the displacement parameter
```

(continues on next page)

(continued from previous page)

```
end

$el_states
2
   1 21100.00    100.00   100.00     1.0000   0.0000   0.0000
   2 24000.00    100.00   1000.00    1.0000   0.0000   0.0000

$el_states_c
2
   1     1         1        1          0        0        0
   2     2         2        2          2        0        0

$abs_bool
2
  1  1
  2  1

$fl_bool
2
  1  1
  2  0

$ss
2
    1      100.000000
    2        0.000000

$ss_c
2
    1            1
    2            0


$vib_freq_gs
1
    1      1500.0


$vib_freq_gs_c
1
    1            1

$sdnc
1 2
                   1          2
      1       2.0000000   0.000000

$sdnc_c
1 2
                   1          2
      1         1          0
```

The parameter `CWAF=1.0` in the `%fit` block specifies the weight of absorption relative to the weight of fluorescence in the difference function to be minimized. If this parameter was not specified the quality of the fit would be biased towards the spectrum with a larger number of experimental points. In some typical situations where the error in the measured experimental intensity is expected to be smaller for absorption than for emission it is desirable to choose the value of `CWAF` to be more than 1.0.

In order to account for a broad featureless background signal in the absorption spectrum above 24000 cm^{-1}, the second band was included into the analysis and approximated with a Voigt lineshape which means also that the corresponding frequency in the `$vib_freq_gs` block and displacement parameter in the `$sdnc` block are fixed to zero in the fit. Thus, the `$el_states` block contains an initial guess on the transition energies, transition electric

dipole moments and linewidth parameters for 2 states:

```
$el_states
2
   1 21100.00    100.00   100.00     1.0000   0.0000   0.0000
   2 24000.00    100.00   1000.00    1.0000   0.0000   0.0000
```

The initial value of the adiabatic minima separation energy for the first state was approximated by the energy corresponding to the first vibronic peak in the absorption spectrum (21100 cm^{-1}). The transition energies and linewidth parameters are varied independently as indicated in the `$el_states_c` block. Since we allow to fit only bandshapes, but not the overall intensities of the spectra, only relative absolute values of the transition electric dipole moments of two bands are important. Therefore it is reasonable to fix all components of the transition moment for the first state and vary only M_x component for the second one:

```
$el_states_c
2
   1     1         1        1          0        0        0
   2     2         2        2          2        0        0
```

Since we assume the absorption by both states and emission only from the first one, it is necessary to include Boolean arrays $abs_bool and $fl_bool which specify states which will be included in the treatment of the absorption and fluorescence spectra, respectively:

```
$abs_bool
2
  1  1 # 1 indicates that the corresponding state will be included in the calculation of
  2  1 # absorption


$fl_bool
2
  1  1
  2  0 # 0 indicates that the corresponding state will be excluded from the calculation
       # of emission spectrum
```

We need also to vary the value of vibrational frequency of the mode which determines separation of vibrational peaks in the spectra. This is done via the constraint block `$vib_freq_gs_c`:

```
$vib_freq_gs_c
1
    1            1
```

Note that it is meaningless to include into the treatment the Stokes shift for the second state which give rise to the background signal in the absorption since the corresponding emission is not present. Therefore λ for the second state is fixed to zero as indicated in the `$ss` block and its constraint counterpart `$ss_c`:

```
$ss
2
    1      100.000000 # initialization of the Stokes shift for the 1st electronic state
    2        0.000000

$ss_c
2
    1         1 # the Stokes shift for the 1st electronic state will be varied in the fit
    2         0 # the Stokes shift for the 2nd electronic state will be fixed in the fit
```

The fit run of `orca_asa` on file `example009.inp` will converge upon approximately 700 function evaluations (for `MWADRelTol=1e-4`). The file `example009.001.inp` will contain the fitted effective values of the vibrational frequency and dimensionless displacement: $\omega = 1404\,cm^{-1}$, $\Delta = 1.35$. One can notice that the fit is rather poor in the low- and high-energy edges of the absorption and fluorescence spectra, respectively (Fig. 7.57). The source of this discrepancy is the single-mode approximation which was employed here. The quality of the fit can be significantly improved assuming several modes with non-zero displacement parameters. Note that in such a case

the proper guess on the number of active modes and corresponding dimensionless displacements can be deduced from quantum chemical calculations.

7.41.4 Quantum-Chemically Assisted Simulations and Fits of Optical Bandshapes and Resonance Raman Intensities

In this section we finally connect the spectra simulation algorithms to actual quantum chemical calculations and outline a detailed approach for the analysis of absorption, fluorescence and resonance Raman spectra within the IMDHO model. Our procedure becomes highly efficient and nearly automatic if analytical excited state derivatives with respect to nuclear displacements are available. However, this availability is not mandatory and hence, spectral predictions may as well be achieved by means of normal mode scan calculations for high-level electronic structure methods for which analytic gradients have not been implemented.

Example: Quantum-Chemically Assisted Analysis and Fit of the Absorption and Resonance Raman Spectra for 1-$^1A_g \rightarrow$ 1-^1B_u Transition in *trans*-1,3,5-Hexatriene

The following input file for an ORCA run invokes the calculation of the excited-state origin displacements along all normal modes by means of energy and excited state gradient calculations at the ground-state equilibrium geometry. The method is valid for the IMDHO model for which the excited-state energy gradient along a given normal mode and corresponding origin shift are related in a very simple way.

```
#
# example010.inp
#
# TDDFT BHLYP Normal Mode Gradient Calculation
#
#   The keyword NMGrad invokes the normal mode gradient calculation
#
! RKS BHandHLYP TightSCF SV(P) NMGrad

%cis    NRoots 1
        triplets false
end

%rr
    # the nuclear Hessian must have been calculated before - for example by a
    # DFT calculation.
    HessName= "hexatriene.hess"

    states 1    # Perform energy-gradient calculations for the 1st
                # excited state.
    Tdnc 0.005  # Threshold for dimensionless displacements to be
                # included in the input file for spectral simulations
                # generated at the end of the program run.
                # By default Tdnc= 0.005

    ASAInput true # Generate the input file for spectra simulations
 end

* xyz   0 1
  C    -0.003374    0.678229   0.00000
  H    -0.969173    1.203538   0.00000
  C     1.190547    1.505313   0.00000
  H     2.151896    0.972469   0.00000
  C     1.189404    2.852603   0.00000
  H     0.251463    3.423183   0.00000
  H     2.122793    3.426578   0.00000
  C     0.003374   -0.678229   0.00000
  H     0.969172   -1.203538   0.00000
```

(continues on next page)

(continued from previous page)

```
  C    -1.190547   -1.505313   0.00000
  H    -2.151897   -0.972469   0.00000
  C    -1.189404   -2.852603   0.00000
  H    -0.251463   -3.423183    0.00000
  H    -2.122793   -3.426578    0.00000
*
```

In the ORCA run the TDDFT excited state gradient calculations are performed on top of a TDDFT calculation. Note, that the numbers of the excited-states which have to be included into analysis and input file for spectral simulations must be specified after the `States` keyword in the `%rr` block. They should also be consistent with the required number of roots in the `%tddft` block. The 1-^1B_u excited state appears to be the first root in the TDDFT calculation. Therefore, `NRoots=1` in the `%tddft` block, and `States=1` in the `%rr` block. One should also provide the name of the file containing the nuclear Hessian matrix via the `HessName` keyword in the `%rr` block. Here we used the `.hess` file obtained in a frequency calculation at the BHLYP/SV(P) level of theory.

After the ORCA calculation you will find in your directory a file called `example010.asa.inp` that is appropriate to be used together with the `orca_asa program` as defined in the preceding sections.

```
#
# example010.asa.inp
#
# ASA input
#
%sim
         model IMDHO
         method Heller
         AbsRange     5000.0,  100000.0
         NAbsPoints    0

         FlRange      5000.0,  100000.0
         NFlPoints     0

         RRPRange     5000.0,  100000.0
         NRRPPoints    0

         RRSRange        0.0,    4000.0
         NRRSPoints 4000

         RRS_FWHM 10.0


         AbsScaleMode Ext
         FlScaleMode Rel
         RamanOrder  0

         EnInput E0

         CAR   0.800

end

%fit
         Fit false
         AbsFit false
         FlFit false
         RRPFit fsalse
         RRSFit false
         method Simplex
         WeightsAdjust true

         AbsRange        0.0,  10000000.0
```

(continues on next page)

(continued from previous page)

```
        FlRange         0.0,  10000000.0
        RRPRange        0.0,  10000000.0
        RRSRange        0.0,  10000000.0
        AbsName ""
        FlName  ""

        ExpFlScaleMode  Rel
        ExpAbsScaleMode Rel

        CWAR  -1.000
        CWAF  -1.000

        NMaxFunc 100
        MWADRelTol= 1.000000e-004

        SFRRPSimStep= 1.000000e+002
        SFRRSSimStep 1.000000e+002
        FREQGStep 1.000000e+001
        FREQEStep 1.000000e+001
        E0Step 3.000000e+002
        SSStep 2.000000e+001
        TMStep 5.000000e-001
        GammaStep 5.000000e+001
        E0SDStep 5.000000e+001
        SDNCStep 4.000000e-001
end

$el_states
1
   1 42671.71   100.00    0.00     1.0725   3.3770  -0.0000

$vib_freq_gs
12
    1      359.709864
    2      456.925612
    3      974.521651
    4     1259.779018
    5     1356.134238
    6     1370.721341
    7     1476.878592
    8     1724.259894
    9     1804.572974
   10     3236.588264
   11     3244.034359
   12     3323.831066


$sdnc
12 1
                  1
      1     -0.594359
      2      0.369227
      3     -0.132430
      4     -0.727616
      5      0.406841
      6     -0.105324
      7      0.177617
      8     -0.090105
      9     -1.412258
     10      0.048788
     11      0.021438
```

(continues on next page)

(continued from previous page)

```
    12       0.008887
```

This input file can be used to construct theoretical absorption and rR spectra. **In order to compare experimental and theoretical rR spectra, it is necessary to use in both cases excitation energies that are approximately in resonance with the same vibrational transitions in the absorption spectrum. Therefore, in the case of the absorption spectrum with resolved or partially resolved vibrational structure it is necessary to modify the transition energies in the `%el_states` such that they coincide with the experimentally observed 0-0 vibrational peaks.** It is also desirable to roughly adjust homogeneous and, possibly, inhomogeneous linewidth parameters such that the experimental and calculated absorption spectra show similar slopes in the preresonance region (below the 0-0 transition). Then the assignment of experimental rR spectra can be done on the basis of comparison with the theoretical rR spectra calculated for the corresponding experimental excitation energies. For the sake of consistency and simplicity it is better to use those excitation energies which fall into the preresonace region and/or are in resonance with the 0-0 transition. **In the case of diffuse absorption spectra (i.e. those not showing resolved vibrational structure) it is also necessary to adjust the theoretical transition energies and linewidth parameters such that experimental and calculated positions of absorption maxima roughly coincide, and corresponding slopes below the maxima have a similar behavior.** According to above mentioned considerations one needs to modify the `%el_states` block in the file `example010.asa.inp`:

```
$el_states
1
   1 39808.0   150.00    0.00  1.0725   3.3770  -0.0000
```

The calculated absorption spectrum obtained by providing `AbsScaleMode= Rel`, `AbsRange= 39000, 49000` and `NAbsPoints= 2000` is shown in Fig. 7.58. Upon comparison with the experimental spectrum one can notice that the BHLYP functional gives relatively small discrepancies with somewhat lower intensity in the low-frequency edge and larger intensity on the high-energy side of the spectrum. Besides, there is a noticeable mismatch in the separation between individual vibronic peaks which is due to overestimation of vibrational frequencies by the BHLYP functional (typically by $\approx 10\%$).

You can arbitrarily vary various normal coordinate displacements in `%sdnc` block within 10–30% of their values in order to observe modifications of the calculated spectrum. This will tell you how these parameters influence the spectrum and probably it will be possible to obtain better initial guesses for the fit. In the present example you will find that reduction of the absolute value of the displacement parameter corresponding to the ninth mode by $\approx 10\%$, and reduction of vibrational frequencies by $\approx 10\%$ can noticeably improve the spectral envelope. Such a quick analysis suggests that experimentally observed peaks in the absorption spectrum represent different vibrational transitions corresponding to a single electronically excited state rather than to different electronic excitations. This conclusion will be confirmed upon establishing the fact that the absorption and rR spectra can be successfully fitted based on the assumption of a single electronic transition.

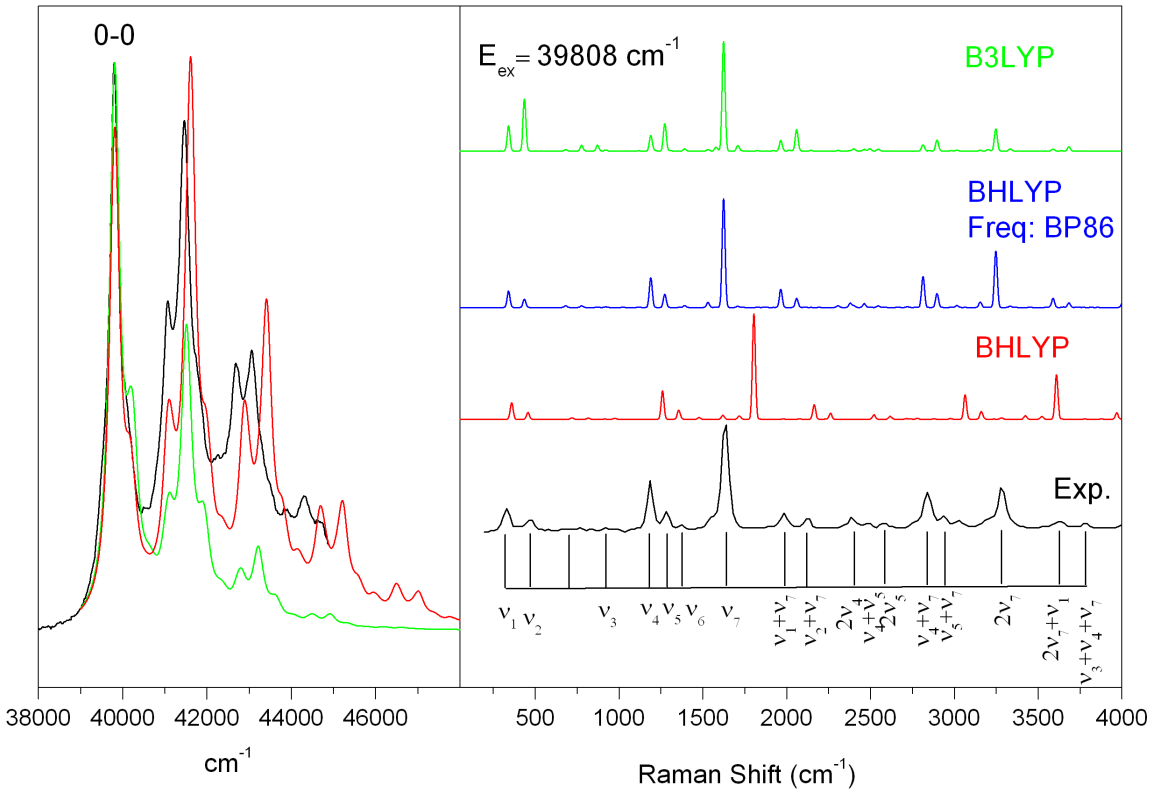


Fig. 7.58: Experimental and calculated at the BHLYP/SV(P) and B3LYP/SV(P) levels of theory absorption (left panel) and rR spectra (right panel) corresponding to 1-$^1A_g \rightarrow$ 1-^1B_u transition in *trans*-1,3,5-hexatriene.

In order to calculate the rR spectrum for experimental excitation energies you need to specify its value through `RRSE` keyword in `%sim` block as well as possibly to modify the parameters related to the spectral range and linewidth of rR bands which are suitable for comparison with the experimental rR spectrum:

```
# excitation energies (cm**-1) for which rR spectra will be calculated:
   RRSE 39808

   # full width half maximum of Raman bands in rR spectra (cm**-1):
   RRS_FWHM  20

   RSRange   0, 4000  # spectral range for simulation of rR spectra (cm**-1)
   NRRSPoints  4000   # number of points to simulate rR spectra (cm**-1)

   # resonance Raman intensities will be calculated
   # for all vibrational states with excitation number
   # up to RamanOrder:
   RamanOrder 3
```

The calculated rR spectrum is shown in Fig. 7.58. In the input we have invoked the calculation of rR intensities for the transitions with up to 3 vibrational quanta in the final vibrational state (`RamanOrder = 3`). Make sure that the rR intensity pattern in the given spectral range does not change noticeably upon further increase of this parameter. Typically, the larger are the normal coordinate displacements the greater order of Raman scattering is required in the calculation to account for all the most intense transitions in the rR spectrum. The inclusion of vibrational transitions beyond the fundamentals is a particular feature of the `orca_asa` program.

Comparison of the calculated and experimental rR spectra (Fig. 7.58) mainly shows discrepancies in the values of the Raman shifts that are mainly related to the low accuracy of the vibrational frequencies obtained at the BHLYP level (typically overestimated by $\approx 10\%$). However, the intensity patterns of the calculated and experimental rR spectra show very nice agreement with experiment that is already sufficient to assign the experimental peaks to

individual vibrational transitions. This can be done upon examination of file `example010.asa.o3.rrs.39808.stk` which provides intensity, Raman shift, and specification for each vibrational transition. It is actually one of the most consistent procedures that enables one to identify different fundamental, overtone and combination bands in the experimentally observed rR spectrum. Such an assignment is a necessary prerequisite for the fit. The current example is relatively straightforward since the spectral region 1–1700 cm^{-1} is actually dominated by fundamental bands while the most intense overtone and combination transitions occur at higher frequencies. However, in many cases even the low-frequency spectral range is characterized by significant contributions from overtone and combination bands that sometimes are even more intense than fundamental transitions! Thus, quantum chemical calculations can greatly facilitate the assignment of experimental rR bands.

After having performed the assignment it is advisable to discard those modes from the analysis that are not involved in any of the experimentally observed fundamental, overtone, or combination rR bands with noticeable intensities. In the present example these are the modes 6, 8, 10–12 from the input file given above. For these modes it is implied that the fitted displacement parameters are zero. You will find that the calculated displacement values are rather small indeed. Also it is advisable to change the ground-state vibrational frequencies in the `$vib_freq_gs` block to their experimental values.

Below is the modified input file for the fit run:

```
#
# example010-01.asa.inp
#
# ASA input
#
%sim
         model IMDHO
         method Heller
end


%fit

    Fit true
    AbsFit true
    RRSFit true
    AbsExpName "hex-abs.dat"
    ExpAbsScaleMode Rel
    CWAR  5.0

    NMaxFunc  1000

    SDNCStep  0.5

end

$el_states
1
   1 39808.0   150.00    0.00    -0.8533  -3.3690  -0.0000

$el_states_c
1
   1     1        1       0           0        0       0


$vib_freq_gs
7
    1      354.0
    2      444.0
    3      934.0
    4     1192.0
    5     1290.0
    6     1403.0
```

(continues on next page)

(continued from previous page)

```
    7     1635.0



$sdnc
7 1
                   1
      1      -0.594359
      2       0.369227
      3      -0.132430
      4      -0.727616
      5       0.406841
      6       0.177617
      7      -1.412258

$sdnc_c
7 1
              1
      1         1
      2         2
      3         3
      4         4
      5         5
      6         6
      7         7

$rrs_exp
1
1 1
    Ex 39809.0
    NTr  11
    1
       int   10.0  1.0
       modes  1
       quanta 1;
    2
       int   5.0  1.0
       modes  2
       quanta 1;
    3
       int   1.5  1.0
       modes  3
       quanta 1;
    4
       int   21.0  1.0
       modes  4
       quanta 1;
    5
       int   7.5  1.0
       modes  5
       quanta 1;
    6
       int   2.0  1.0
       modes  6
       quanta 1;
    7
       int   46.0  1.0
       modes  7
       quanta 1;

    8
```

(continues on next page)

(continued from previous page)

```
      int  6.8 1.0
      modes  1, 7
      quanta 1, 1;

   9
      int  4.0 1.0
      modes  2, 7
      quanta 1, 1;

  10
      int  2.0 1.0
      modes  3, 7
      quanta 1, 1;

  11
      int 17.0 1.0
      modes  7
      quanta 2;
```

In addition to the experimental intensities of fundamental bands the input file also contains the information about some overtone and combination transitions. Note that it is not really necessary to include all of them them into the fit, in particular if some of the rR bands are strongly overlapping with each other.

Fitted normal coordinate displacements of the excited-state origin show nice agreement with the published values:

```
$sdnc
7 1
                    1
      1      -0.638244
      2       0.455355
      3      -0.229126
      4      -0.854357
      5       0.501219
      6       0.197679
      7      -1.292997
```

NOTE

- It is not really important to employ the BHLYP/SV(P) method in the frequency calculations in order to obtain the `.hess` file (this was merely done to be consistent with the TDDFT/BHLYP/SV(P) method for the excited-state model parameters calculation). The frequency calculations can for example be carried out at the BP86/TZVP or RI-SCS-MP2/TZVP level of theory. This will provide displacements pattern very similar to that of the BHLYP/SV(P) method, but much more accurate vibrational frequencies which will further facilitate the assignment of rR spectra (Fig. 7.58). However, such a procedure can be inconsistent if the two methods give noticeably different normal mode compositions and/or vibrational frequencies. From our experience it can lead to significant overestimation of the excited-state displacements for some low-frequency modes.
- It is known that predicted dimensionless normal coordinate displacements critically depend on the fraction of the "exact" Hartree-Fock exchange (EEX) included in hybrid functionals. In general no universal amount of EEX exists that provides a uniformly good description for all systems and states. Typically, for a given molecule either the BHLYP/TZVP (50% of EEX) or B3LYP/TZVP (20% of EEX) methods yields simulated spectra that compare very well with those from experiment if vibrational frequencies are appropriately scaled.

Important Notes about Proper Comparison of Experimental and Quantum Chemically Calculated Resonance Raman Spectra

In order to compare experimental and theoretical rR spectra, **it is necessary to use in both cases excitation energies that are approximately in resonance with the same vibrational transitions in the absorption spectrum**. Therefore, in the case of diffuse absorption spectra (i.e. those not showing resolved vibrational structure) one needs to adjust the transition energies and linewidth paramters in the `%el_states` block such that the envelopes of the experimental and theoretical spectra rouhgly coincide, and then to employ experimental values of excitation energies to construct theoretical rR spectra. Typically in the case of diffuse absorption spectra rR profiles are rather smooth. Therefore, even though excitation energies are not in resonance with the same vibrational transition in the absorption spectrum, the rR spectra are not expected to vary significantly in the case of such mismatch.

In the case of the absorption spectrum with resolved or partially resolved vibrational structure it is necessary to modify the transition energies in the `%el_states` block such that the calculated and experimentally observed 0-0 vibrational peaks coincide, and modify linewidth parameters so that the low-energy slopes in the calculated and experimental spectra have a similar behavior.

Consider a single-mode model system for which "experimental" and calculated absorption spectra are shown in Fig. 7.59.

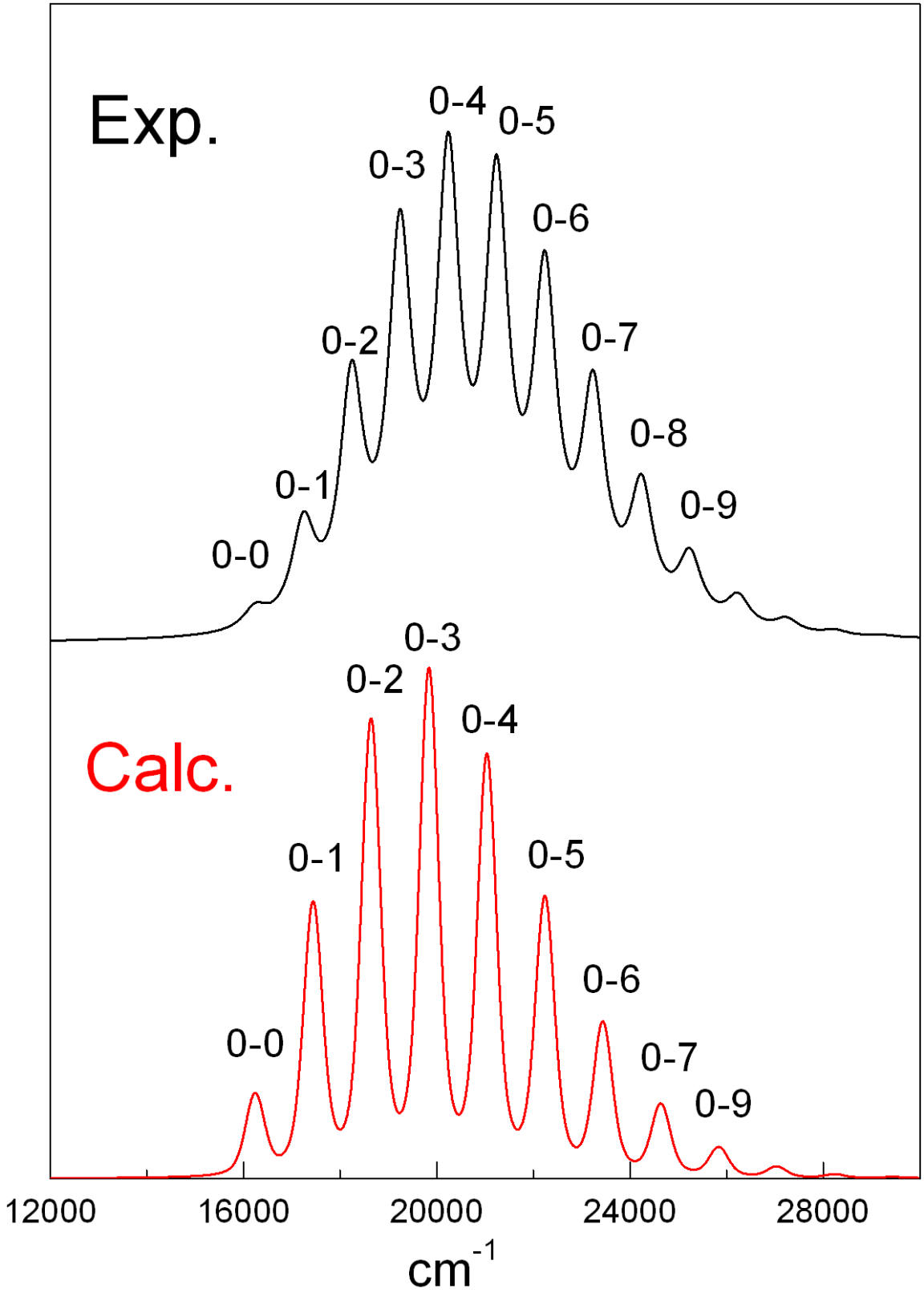


Fig. 7.59: Experimental and theoretical absorption spectra for a single-mode model system. The calculated spectrum is adjusted such that the position of 0-0 peak coincide with the experimental one.

Comparison of the calculated and experimental spectra shows that some adjustment of the linewidth parameters is neceassy before construction of theoretical rR spectra. One can directly compare calculated and experimental rR spectra upon the excitation at 16200 cm^{-1} which is in resonance with the 0-0 vibronic band. However, it is not consistent to use experimental values of the excitation energy in the calculation of rR spectrum which is in resonance with one of the other vibronic bands since the separation between vibrational peaks in the experimental and calculated spectra is different whereby positions of the peaks in both spectra do not coincide. Instead **one should use the excitaition energy which corresponds to the same vibronic peak in the calculated absorption spectrum as in the experimental one**. Alternatively, one can adjust theoretical value of vibrational frequency such that positions of corresponding vibronic peaks in the spectra coincide, and then use experimental values of excitation energies for the calculation of rR spectra.

Example: Normal Mode Scan Calculations of Model Parameters for 1-$^1A_g \rightarrow$ 1-^1B_u Transition in *trans*-1,3,5-Hexatriene

If excited state gradients are not available (which is the case for many of the electronic structure methods supported by ORCA), you have to resort to a more laborious procedure – single point calculations at geometries that are displaced along the various normal modes of the system. This roughly corresponds to taking numerical derivatives – however, once this extra effort is invested more information can be obtained from the calculation than what would be possible from an analytic derivative calculation.

The present example illustrates the application of normal mode scan calculations for the evaluation of excited state harmonic parameters that are necessary to simulate optical spectra within the IMDHO model. This method can be applied with any method like CIS, CASSCF, MRCI or TD-DFT.

The reference wavefunctions for the multireference calculations reported below are of the state-averaged CASSCF (SA-CASSCF) type. The complete active space CAS(6,6) includes all 6 valence shell π-orbitals. The average is taken over the first four states which was found necessary in order to include the ground state and the strongly allowed 1-^1B_ustate.

```
#
# example011.inp
#
# CASSCF normal mode scan calculations
#

# first do single point RHF calculation
! RHF TZVP TightSCF

* xyz 0 1
 C       -0.002759    0.680006    0.000000
 H       -0.966741    1.204366    0.000000
 C        1.187413    1.500920    0.000000
 H        2.146702    0.969304    0.000000
 C        1.187413    2.850514    0.000000
 H        0.254386    3.420500    0.000000
 H        2.116263    3.422544    0.000000
 C        0.002759   -0.680006    0.000000
 H        0.966741   -1.204366    0.000000
 C       -1.187413   -1.500920    0.000000
 H       -2.146702   -0.969304    0.000000
 C       -1.187413   -2.850514    0.000000
 H       -0.254386   -3.420500    0.000000
 H       -2.116263   -3.422544    0.000000
*


# perform SA-CASSCF calculation upon appropriate rotation of MOs
$new_job
! TZVP TightSCF
```

(continues on next page)

(continued from previous page)

```
%scf
    rotate {23,27} end
end

%casscf
        nel 6
        norb  6
        mult  1
        nroots  4
end

* xyz 0 1
 C       -0.002759    0.680006    0.000000
 H       -0.966741    1.204366    0.000000
 C        1.187413    1.500920    0.000000
 H        2.146702    0.969304    0.000000
 C        1.187413    2.850514    0.000000
 H        0.254386    3.420500    0.000000
 H        2.116263    3.422544    0.000000
 C        0.002759   -0.680006    0.000000
 H        0.966741   -1.204366    0.000000
 C       -1.187413   -1.500920    0.000000
 H       -2.146702   -0.969304    0.000000
 C       -1.187413   -2.850514    0.000000
 H       -0.254386   -3.420500    0.000000
 H       -2.116263   -3.422544    0.000000
*


# do normal mode scan calculations
# to map CASSCF ground and excited-state PESs
$new_job
! TZVP TightSCF  NMScan

%casscf
        nel 6
        norb  6
        mult  1
        nroots  4
end

%rr
    HessName "hexatriene_bp86.hess"
    NMList 10,11,18,24,26,28,29,31,32
    NSteps 6
    FreqAlter true
    EnStep 0.0001
    State 3
end

* xyz 0 1
 C       -0.002759    0.680006    0.000000
 H       -0.966741    1.204366    0.000000
 C        1.187413    1.500920    0.000000
 H        2.146702    0.969304    0.000000
 C        1.187413    2.850514    0.000000
 H        0.254386    3.420500    0.000000
 H        2.116263    3.422544    0.000000
 C        0.002759   -0.680006    0.000000
 H        0.966741   -1.204366    0.000000
 C       -1.187413   -1.500920    0.000000
```

(continues on next page)

(continued from previous page)

```
 H       -2.146702   -0.969304    0.000000
 C       -1.187413   -2.850514    0.000000
 H       -0.254386   -3.420500    0.000000
 H       -2.116263   -3.422544    0.000000
*
```

The file containing the Hessian matrix (`"hexatriene_bp86.hess"`) was obtained from the BP86/TZVP frequency calculations. The keyword `NMList` provides the list of the normal modes to be scanned. These should be only the totally symmetric vibrations, since only they can be significant for absorption and resonance Raman spectra within the constraints of the IMDHO model. The `FreqAlter` flag indicates whether frequency alterations are assumed in the post-scan potential surface fit. The Parameter `EnStep` is used to select the appropriate step during the scan calculations. The value is chosen such that the average energy change (in Eh) in both directions is not less than this parameter.

7.42 One Photon Spectroscopy

7.42.1 General Description

Introduced in Orca 6.0, the One Photon Spectroscopy (OPS) tool now takes charge of computing one-photon absorption (OPA), emission (OPE), and natural electric circular dichroism (ECD) intensities. In each of these processes, the intensity of the spectrum ($I(\omega)$) resulting from a transition between an initial state I and a final state J is determined by the square modulus of the transition moment $T_{IJ} = \langle \Psi_I | \hat{H}_1 | \Psi_J \rangle$, weighted by the populations N_I and N_J of states I and J, respectively.

$$I(\omega) = \alpha\omega^{-1} \sum_{IJ} (N_I - N_J) |\langle \Psi_I | \hat{H}_1 | \Psi_J \rangle|^2 \delta(E_{JI} \pm \hbar\omega) \tag{7.256}$$

In this context, α represents a positive constant, ω stands for photon energy, and the expression for $\hat{H}_1$ hinges on the specific modeling of photon-matter interaction.

7.42.2 Light-matter interaction approaches

Expressions for $\hat{H}_1$ can be derived from different theoretical perspectives. For instructional purposes, a classical electrodynamic approach is adopted here. By representing light classically through a vector potential ($\mathbf{A}$) and a scalar potential ($\phi = 0$), the radiation can be integrated into the Hamiltonian that models the molecular system.

$$\hat{H} = \sum_{i=1}^{N} \frac{1}{2m_e} \left[\hat{\mathbf{p}}_i - \frac{e}{c} \mathbf{A}(\mathbf{r}_i, t) \right]^2 - \frac{ge}{2m_e c} \sum_{i=1}^{N} \mathbf{B}(\mathbf{r}_i, t) \cdot \hat{s}_i + V(\mathbf{r}_1, ..., \mathbf{r}_N)$$

By disregarding the coupling between the magnetic field of light and the spin, as well as the $|\mathbf{A}|^2$ term, and applying Fermi's golden rule, an expression for $\langle \Psi_I | \hat{H}_1 | \Psi_J \rangle$ under the Full-Semiclassical Light-Matter interaction can be derived.

$$T_{IJ}^{FFMIO} = \frac{e}{m_e} \sum_{i=1}^{N} \langle I | \, \mathcal{E} \cdot \left[e^{i\mathbf{k}\cdot\mathbf{r}_i} \hat{\mathbf{p}}_i \right] | J \rangle \tag{7.257}$$

Where T_{IJ}^{FFMIO} is the transition moment for the **F**ull (semiclassical) **F**ield-**M**atter **I**nteraction **O**perator, while k and $\mathcal{E}$ denote the wave and polarization vectors describing the light, respectively. $\mathbf{r}_i$ and $\mathbf{p}_i$ represent the position and linear momentum operators for the i-th electron.

Proceeding from eq. (7.257) and approximating the exponential term of T_{IJ}^{FFMIO} with a Taylor expansion yields various orders of interaction. The zeroth order ($T_{IJ}^{[0]}$) results in the electric dipolar velocity formulation ($ED\ vel$), which, upon molecular orientational averaging, yields:

$$\epsilon_{ED\ vel}(\omega) = \sum_{IJ} (N_I - N_J) \frac{2}{3E_{JI}} |\langle \Psi_I | \hat{p} | \Psi_J \rangle|^2 \delta(E_{JI} \pm \hbar\omega) \tag{7.258}$$

For exact solutions of the Hamiltonian, or in theories that satisfy the Hypervirial theorem by construction, equation (7.258) can be transformed into a length formulation (7.259) ($ED\ len$).

$$\epsilon_{ED\ len}(\omega) = \sum_{IJ}(N_I - N_J)\frac{2E_{JI}}{3}|\langle\Psi_I|\hat{\mu}|\Psi_J\rangle|^2\delta(E_{JI} \pm \hbar\omega) \tag{7.259}$$

The absorption spectrum under $ED\ len$ and $ED\ vel$ formulations can be obtained in all Orca modules utilizing OPS by setting **DoDipoleLength** and **DoDipoleVelocity** to true, respectively, in the corresponding module's block. The results are then presented in the following tables:

```
-------------------------------------------------------------------------------------------------
-----
                    ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
-----
     Transition      Energy     Energy  Wavelength fosc(D2)      D2        DX        DY        ␣
↪DZ
                      (eV)       (cm-1)    (nm)               (au**2)    (au)      (au)      ␣
↪(au)
-------------------------------------------------------------------------------------------------
-----
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5   0.000000000   0.00000  -0.00000   0.00000   0.
↪00000
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7   0.000000000   0.00000  -0.00000   0.00000  -0.
↪00000
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5   0.000000000   0.00000   0.00000   0.00000  -0.
↪00000
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3   0.301976959   2.47833   1.57427  -0.00000  -0.
↪00025
```

```
-------------------------------------------------------------------------------------------------
-----
                    ABSORPTION SPECTRUM VIA TRANSITION VELOCITY DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
-----
     Transition      Energy     Energy  Wavelength fosc(P2)      P2        PX        PY        ␣
↪PZ
                      (eV)       (cm-1)    (nm)               (au**2)    (au)      (au)      ␣
↪(au)
-------------------------------------------------------------------------------------------------
-----
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5   0.000000000   0.00000   0.00000   0.00000  -0.
↪00000
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7   0.000000000   0.00000   0.00000   0.00000  -0.
↪00000
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5   0.000000001   0.00000  -0.00000  -0.00000  -0.
↪00001
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3   0.292852326   0.08029  -0.28335   0.00000  -0.
↪00002
```

Here, transitions are denoted using spectroscopic notation, such as 1-3B3g representing 1^3B_{3g}. If symmetry is not specified in the calculation (or is unavailable in the selected method), the system defaults to C_1 point group symmetry. The "fosc(D2)" and "fosc(P2)" columns indicate the computed oscillator strengths in lenght and velocity formulations respectively. Additionally, "D2", "P2", "DX", "DY", "DZ", "PX", "PY", and "PZ" represent the square modulus of the electric transition dipole moment, square modulus of the transition linear momentum, and its Cartesian components, respectively (with the imaginary unit in the linear momentum being implicit).

The first-order term ($T_{IJ}^{[1]}$) in the Taylor expansion of T_{IJ}^{FFMIO} gives rise to the electric quadrupole velocity formulation and the magnetic dipole contributions.

$$\langle\Psi_I|(T_{EQvel})_{\alpha\beta}|\Psi_J\rangle = \frac{ie}{2m_e}k_\alpha\epsilon_\beta\langle\Psi_I|r_\alpha p_\beta + p_\alpha r_\beta|\Psi_J\rangle \tag{7.260}$$

$$\langle \Psi_I | T_{MD} | \Psi_J \rangle = \frac{ie}{2m_e} \langle \Psi_I | (k \times \epsilon)(r \times p) | \Psi_J \rangle \tag{7.261}$$

Under the same conditions applied to the electric dipole transition moment, the quadrupole contribution can be reformulated in terms of a length representation.

$$\langle \Psi_I | (T_{EQlen})_{\alpha\beta} | \Psi_J \rangle = \frac{eE_{JI}}{2m_e} \sum_i k_\alpha \epsilon_\beta \langle \Psi_I | r_\alpha r_\beta | \Psi_J \rangle \tag{7.262}$$

By squaring the modulus of $(T_{IJ}^{[0]} + T_{IJ}^{[1]})$, six terms emerge in the oscillation strength intensity: the dipole square contribution as listed previously in the dipole-approximation tables, a magnetic dipole square, an electric quadrupole square, and three cross-product terms: electric dipole-electric quadrupole, electric dipole-magnetic dipole, and magnetic dipole-electric quadrupole. Enabling **DoHigherMoments** to true allows for the estimation of spectrum intensity by including all three squared terms in the calculation of the intensity. Similarly to the $ED\ len$ and $ED\ vel$ tables, the **DoDipoleLength** and **DoDipoleVelocity** keywords control the length and velocity representations of the electric operators.

```
-------------------------------------------------------------------------------------------------
-----------------------------------
                     ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM
-------------------------------------------------------------------------------------------------
-----------------------------------
     Transition      Energy     Energy  Wavelength fosc(D2)  fosc(M2)  fosc(Q2)  ␣
↪fosc(D2+M2+Q2)     D2/TOT    M2/TOT    Q2/TOT
                      (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-------------------------------------------------------------------------------------------------
-----------------------------------
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5     0.00000   0.49891   0.00010   0.
↪00000049901544   0.00000   0.99979   0.00021
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7     0.00000   0.00000   0.00000   0.
↪00000000000000   0.00000   0.00000   0.00000
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5     0.00000   1.12904   0.03298   0.
↪00000116202090   0.00000   0.97162   0.02838
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3     0.30198   0.00000   0.00000   0.
↪30197695902797   1.00000   0.00000   0.00000
```

```
-------------------------------------------------------------------------------------------------
-----------------------------------
                     ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM (Velocity)
-------------------------------------------------------------------------------------------------
-----------------------------------
     Transition      Energy     Energy  Wavelength fosc(P2)  fosc(M2)  fosc(Q2)  ␣
↪fosc(P2+M2+Q2)     P2/TOT    M2/TOT    Q2/TOT
                      (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-------------------------------------------------------------------------------------------------
-----------------------------------
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5     0.00000   0.49891   0.00010   0.
↪00000049901234   0.00000   0.99980   0.00020
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7     0.00000   0.00000   0.00000   0.
↪00000000000006   0.99964   0.00012   0.00024
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5     0.00000   1.12904   0.03441   0.
↪00000116403924   0.00051   0.96993   0.02956
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3     0.29285   0.00000   0.00000   0.
↪29285232633663   1.00000   0.00000   0.00000
```

Here, the column "fosc(D2+M2+Q2)" and "fosc(P2+M2+Q2)" refer to the total oscillator strengths obtained in length and velocity formulations, respectively, while the columns "fosc(D2)"/"fosc(P2)", "fosc(M2)", and "fosc(Q2)" refer to the individual contributions of each term.

In scenarios where the dipolar contribution to the intensity is non-zero, dividing the total intensity into D2/P2, M2, and Q2 contributions becomes challenging due to the dependence of M2 and Q2 terms on the chosen origin. To

address this origin-dependence issue, OPS offers additional formulations to the multipolar expansion. One possible approach devised by our group describes each transition from an origin that minimizes the contributions of M2 and Q2, thereby redistributing the intensity from these terms to all other components in the expansion of $e^{i\mathbf{k}\cdot\mathbf{r}_i}$. This refined formulation is readily accessible in an "origin-adjusted" table, provided that **DoHigherMoments** is enabled and additionally **DecomposeFoscLength** is set to true. [201], [200]

```
-----------------------------------------------------------------------------------------------
-----------------------------------
                     ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM (origin adjusted)
-----------------------------------------------------------------------------------------------
-----------------------------------
     Transition      Energy     Energy  Wavelength fosc(D2)  fosc(M2)  fosc(Q2)  ␣
↪fosc(D2+M2+Q2)     D2/TOT    M2/TOT    Q2/TOT
                      (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-----------------------------------------------------------------------------------------------
-----------------------------------
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5     0.00000   0.49891   0.00010   0.
↪00000049901544   0.00000   0.99979   0.00021
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7     0.00000   0.00000   0.00000   0.
↪00000000000000   0.00000   0.00000   0.00000
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5     0.00000   1.12904   0.03298   0.
↪00000116202090   0.00000   0.97162   0.02838
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3     0.30198   0.00000   0.00000   0.
↪30197695902797   1.00000   0.00000   0.00000
```

Alternatively, the intensity described through a truncated expansion of $e^{i\mathbf{k}\cdot\mathbf{r}_i}$ be achieved in an origin-independent manner. In this scenario, the cross-terms between the electric dipole moment and the electric octupole, as well as between the electric dipole moment and the magnetic quadrupole moments, arise from the second-order expansion ($T^{[2]}_{IJ}$), resolving the origin dependence on the D2, M2, and Q2 contributions. This formulation is provided in both length and velocity representations for the electric operators by setting the keywords **DecomposeFoscLength** and **DecomposeFoscVelocity** to true, and the results are presented in the tables "Origin Independent, Length" and "Origin Independent, Velocity," respectively."

```
-----------------------------------------------------------------------------------------------
-----------------------------------
                     ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM (Origin Independent, Length)
-----------------------------------------------------------------------------------------------
-----------------------------------
     Transition      Energy     Energy  Wavelength fosc(D2)  fosc(M2)  fosc(Q2)␣
↪fosc(D2+M2+Q2+DM+DO) D2/TOT    M2/TOT    Q2/TOT
                      (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-----------------------------------------------------------------------------------------------
-----------------------------------
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5     0.00000   0.49891   0.00010   0.
↪00000049901544   0.00000   0.99979   0.00021
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7     0.00000   0.00000   0.00000   0.
↪00000000000000   0.00000   0.00000   0.00000
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5     0.00000   1.12904   0.03298   0.
↪00000116202090   0.00000   0.97162   0.02838
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3     0.30198   0.00000   0.00000   0.
↪30197506344159   1.00001   0.00000   0.00000
```

```
-----------------------------------------------------------------------------------------------
-----------------------------------
                     ABSORPTION SPECTRUM COMBINED ELECTRIC DIPOLE + MAGNETIC DIPOLE + ELECTRIC␣
↪QUADRUPOLE SPECTRUM (Origin Independent, Velocity)
-----------------------------------------------------------------------------------------------
-----------------------------------
     Transition      Energy     Energy  Wavelength fosc(P2)  fosc(M2)  fosc(Q2)␣
```

(continues on next page)

(continued from previous page)

```
↪fosc(P2+M2+Q2+PM+PO) P2/TOT    M2/TOT    Q2/TOT
                      (eV)      (cm-1)    (nm)       (au)    (au*1e6)  (au*1e6)
-----------------------------------------------------------------------------------------------
↪-----------------------------------
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5     0.00000   0.49891   0.00010   0.
↪00000049901234   0.00000   0.99980   0.00020
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7     0.00000   0.00000   0.00000   0.
↪00000000000006   0.99962   0.00012   0.00024
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5     0.00000   1.12904   0.03441   0.
↪00000116403924   0.00051   0.96993   0.02956
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3     0.29285   0.00000   0.00000   0.
↪29285049443551   1.00001   0.00000   0.00000
```

Finally, intensity computation directly using T_{IJ}^{FFMIO} is achievable in OPS through the utilization of the keyword **DoFullSemiclassical** set to true. In this scenario, the orientational average is computed semi-numerically due to the absence of available analytical expressions.

```
-------------------------------------------------------------------------
                    ABSORPTION SPECTRUM VIA FULL SEMI-CLASSICAL FORMULATION
-------------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength   fosc(FFMIO)
                      (eV)      (cm-1)    (nm)
-------------------------------------------------------------------------
  0-1Ag ->  1-1B3g  2.507275   20222.5   494.5     0.00000049901134
  0-1Ag ->  2-1Au   2.726731   21992.6   454.7     0.00000000000260
  0-1Ag ->  3-1B2g  3.892633   31396.2   318.5     0.00000116403914
  0-1Ag ->  4-1B3u  4.973431   40113.4   249.3     0.29285049449335
```

The tables presented in this section can be generated using the following input example:

```
!B3LYP def2-TZVPP UseSym

%sym SymThresh 1.0e-2 end

%CIS
  Nroots 4
  Tda false
  DoDipoleLength true
  DoDipoleVelocity true
  DoHigherMoments true
  DoFullSemiclassical true
  DecomposeFoscLength true
  DecomposeFoscVelocity true
END

* xyz 0 1
  C  -0.03361342468929      0.00165373985356     -0.01949461738929
  C   0.00156939445126      0.00011912087577      1.46272959096597
  C   1.15884465742251      0.00010247651897      2.13104196345420
  C   2.45992795249087      0.00152015393852      1.42032876326693
  C   2.42474592208737      0.00010359895701     -0.06189678799337
  C   1.26746963770198      0.00012121580349     -0.73020920027406
  H  -0.96132586605922     -0.00069823088896      1.95925560299438
  H   1.20967888389840     -0.00071666462249      3.21322810501222
  O   3.51487697965937     -0.00034474583300      2.02878006087658
  H   3.38764113128475     -0.00071842966522     -0.55842278191419
  H   1.21663501697728     -0.00069837665390     -1.81239534279079
  O  -1.08856228522528     -0.00044385828375     -0.62794535620861
*
```

7.42.3 Natural electric circular dichroism

In the case of ECD modeling, intensity is determined by computing the difference between two absorption spectra: one acquired using left-circularly polarized light and the other utilizing right-circularly polarized light.

$$\Delta I(\omega) = I_{left}(\omega) - I_{rigth}(\omega)$$

To initiate these calculations, the keyword **DoCD** should be set to true. There are three available implementations, which involve utilizing the length or velocity representations for the electric dipole moment, and also using the FFMIO, selected by including the keywords **DoDipoleLength**, **DoDipoleVelocity**, and **DoFullSemiclassical** respectively.

```
-----------------------------------------------------------------------------
                    CD SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-----------------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength    R        MX        MY        MZ
                      (eV)      (cm-1)    (nm)   (1e40*cgs)   (au)      (au)      (au)
-----------------------------------------------------------------------------
  0-1A  ->  1-1A    5.858500   47252.0   211.6   -9.15001  -0.01149   0.10351  -0.06803
  0-1A  ->  2-1A    6.859736   55327.5   180.7  -10.82714   0.13087   0.17869   0.25077
  0-1A  ->  3-1A    7.025193   56662.0   176.5   10.05878  -0.00526  -0.09041  -0.13618
  0-1A  ->  4-1A    7.837041   63210.0   158.2   33.15402   0.02541  -0.23067   0.15096
```

```
-----------------------------------------------------------------------------
                    CD SPECTRUM VIA TRANSITION VELOCITY DIPOLE MOMENTS
-----------------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength    R        MX        MY        MZ
                      (eV)      (cm-1)    (nm)   (1e40*cgs)   (au)      (au)      (au)
-----------------------------------------------------------------------------
  0-1A  ->  1-1A    5.858500   47252.0   211.6  -10.85364  -0.01149   0.10351  -0.06803
  0-1A  ->  2-1A    6.859736   55327.5   180.7  -15.54410   0.13087   0.17869   0.25077
  0-1A  ->  3-1A    7.025193   56662.0   176.5   12.44178  -0.00526  -0.09041  -0.13618
  0-1A  ->  4-1A    7.837041   63210.0   158.2   38.90662   0.02541  -0.23067   0.15096
```

```
-------------------------------------------------------------------
                    CD SPECTRUM VIA FULL SEMI-CLASSICAL FORMULATION
-------------------------------------------------------------------
     Transition      Energy     Energy  Wavelength    R
                      (eV)      (cm-1)    (nm)   (1e40*cgs)
-------------------------------------------------------------------
  0-1A  ->  1-1A    5.858500   47252.0   211.6   -10.85394
  0-1A  ->  2-1A    6.859736   55327.5   180.7   -15.54405
  0-1A  ->  3-1A    7.025193   56662.0   176.5    12.44211
  0-1A  ->  4-1A    7.837041   63210.0   158.2    38.90630
```

The following input example may be used to generate the CD tables presented in this section:

```
!B3LYP def2-QZVPP

%CIS
  Nroots 4
  Tda false
  Docd true
  DoDipoleLength true
  DoDipoleVelocity true
  DoFullSemiclassical true
END

* xyz 0 1
  O   0.00292203187063     -0.00094788357265      0.00607763745451
  H  -0.01058222980812      0.00702286262037      0.97110447138368
```

(continues on next page)

(continued from previous page)

```
  O    1.43049571483128     -0.00638910838156     -0.24427444074129
  H    1.54056670040621     -0.87944353346616     -0.64068822569690
*
```

7.42.4 SOC and SSC corrected spectrum

All the OPS keywords and tables listed above are available for the cases in which spin-orbit coupling and spin-spin coupling effects are taken into account by a QDPT formulation (when the method is available in the selected ORCA module). In those cases, ORCA 6.0 does not provide symmetry after the relativistic correction; therefore, C_1 group symmetry transitions are reported. Additionally, the multiplicity of the relativistically corrected roots are not well defined anymore; therefore, the average value is reported.

For the last case example presented in this section, when **DoSOC** is set to true in the input, OPS reports:

```
-------------------------------------------------------------------------------------------------
      SOC CORRECTED ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
      Transition         Energy     Energy  Wavelength fosc(D2)      D2        |DX|      |DY|
    |DZ|
                          (eV)      (cm-1)    (nm)                (au**2)    (au)      (au)
    (au)
-------------------------------------------------------------------------------------------------
  0-1.0A  ->  1-3.0A    4.620455   37266.5   268.3   0.000000105   0.00000   0.00079   0.00024
    0.00049
  0-1.0A  ->  2-3.0A    4.620460   37266.5   268.3   0.000000009   0.00000   0.00003   0.00024
    0.00016
  0-1.0A  ->  3-3.0A    4.620491   37266.8   268.3   0.000000126   0.00000   0.00093   0.00034
    0.00035
  0-1.0A  ->  4-3.0A    5.626046   45377.1   220.4   0.000003140   0.00002   0.00043   0.00397
    0.00261
  0-1.0A  ->  5-3.0A    5.626322   45379.4   220.4   0.000000070   0.00000   0.00056   0.00020
    0.00040
  0-1.0A  ->  6-3.0A    5.626327   45379.4   220.4   0.000000003   0.00000   0.00002   0.00013
    0.00009
  0-1.0A  ->  7-1.0A    5.858722   47253.8   211.6   0.003490904   0.02432   0.01409   0.12986
    0.08520
  0-1.0A  ->  8-3.0A    6.537246   52726.4   189.7   0.000000006   0.00000   0.00017   0.00005
    0.00008
  0-1.0A  ->  9-3.0A    6.537246   52726.5   189.7   0.000000225   0.00000   0.00011   0.00099
    0.00065
  0-1.0A  -> 10-3.0A    6.537257   52726.5   189.7   0.000000032   0.00000   0.00004   0.00037
    0.00025
  0-1.0A  -> 11-1.0A    6.859693   55327.2   180.7   0.002900797   0.01726   0.12934   0.02129
    0.00891
  0-1.0A  -> 12-1.0A    7.025210   56662.1   176.5   0.011980549   0.06961   0.23240   0.08590
    0.09066
  0-1.0A  -> 13-3.0A    7.448215   60073.9   166.5   0.000000040   0.00000   0.00018   0.00022
    0.00037
  0-1.0A  -> 14-3.0A    7.448216   60073.9   166.5   0.000000003   0.00000   0.00002   0.00011
    0.00006
  0-1.0A  -> 15-3.0A    7.448231   60074.0   166.5   0.000000533   0.00000   0.00116   0.00078
    0.00098
  0-1.0A  -> 16-1.0A    7.837071   63210.2   158.2   0.012389093   0.06452   0.02426   0.21144
    0.13867
```

In this case, due to the complex nature of the relativistically-corrected wave functions, dipole moments are not necessarily real values, and their modulous are reported in the "|DX|", "|DY|", and "|DZ|" columns. Real and

imaginary components may be obtained in extended tables by increasing the selected print level. In this instance, selecting **Printlevel 4** in the %CIS block.

```
-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------
      SOC CORRECTED ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS (- Extended -)
-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------
      Transition         Energy     Energy  Wavelength     fosc(D2)          D2          
↪DX(re)         DX(im)         DY(re)         DY(im)         DZ(re)         DZ(im)
                          (eV)       (cm-1)    (nm)                          (au**2)       
↪(au)           (au)           (au)           (au)           (au)           (au)
-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------
  0-1.0A  ->  1-3.0A    4.620455   37266.5   268.3         1.05e-07       9.31e-07       4.10e-
↪04       6.80e-04       -1.22e-04      -2.02e-04      -2.55e-04      -4.24e-04
  0-1.0A  ->  2-3.0A    4.620460   37266.5   268.3         9.18e-09       8.11e-08      -2.25e-
↪05       1.30e-05        2.05e-04      -1.18e-04      -1.35e-04       7.78e-05
  0-1.0A  ->  3-3.0A    4.620491   37266.8   268.3         1.26e-07       1.11e-06       4.17e-
↪04      -8.35e-04        1.51e-04      -3.01e-04       1.58e-04      -3.17e-04
  0-1.0A  ->  4-3.0A    5.626046   45377.1   220.4         3.14e-06       2.28e-05      -4.31e-
↪04       5.49e-07        3.97e-03      -5.06e-06      -2.61e-03       3.32e-06
  0-1.0A  ->  5-3.0A    5.626322   45379.4   220.4         7.04e-08       5.11e-07       1.13e-
↪04       5.45e-04       -4.10e-05      -1.98e-04      -8.13e-05      -3.92e-04
  0-1.0A  ->  6-3.0A    5.626327   45379.4   220.4         3.42e-09       2.48e-08      -1.52e-
↪05      -1.15e-07        1.31e-04       9.97e-07      -8.58e-05      -6.51e-07
  0-1.0A  ->  7-1.0A    5.858722   47253.8   211.6         3.49e-03       2.43e-02      -1.41e-
↪02      -2.11e-05        1.30e-01       1.94e-04      -8.52e-02      -1.28e-04
  0-1.0A  ->  8-3.0A    6.537246   52726.4   189.7         6.35e-09       3.97e-08      -1.24e-
↪04       1.22e-04        3.61e-05      -3.56e-05       5.89e-05      -5.80e-05
  0-1.0A  ->  9-3.0A    6.537246   52726.5   189.7         2.25e-07       1.40e-06       1.10e-
↪04       1.07e-06       -9.87e-04      -9.61e-06       6.46e-04       6.29e-06
  0-1.0A  -> 10-3.0A    6.537257   52726.5   189.7         3.24e-08       2.02e-07      -4.25e-
↪05      -5.39e-08        3.73e-04       4.74e-07      -2.47e-04      -3.13e-07
  0-1.0A  -> 11-1.0A    6.859693   55327.2   180.7         2.90e-03       1.73e-02       4.49e-
↪02       1.21e-01        7.39e-03       2.00e-02       3.09e-03       8.35e-03
  0-1.0A  -> 12-1.0A    7.025210   56662.1   176.5         1.20e-02       6.96e-02      -2.18e-
↪01       8.17e-02       -8.04e-02       3.02e-02      -8.49e-02       3.19e-02
  0-1.0A  -> 13-3.0A    7.448215   60073.9   166.5         3.95e-08       2.17e-07       6.65e-
↪05      -1.67e-04       -8.19e-05       2.05e-04      -1.37e-04       3.42e-04
  0-1.0A  -> 14-3.0A    7.448216   60073.9   166.5         2.93e-09       1.61e-08      -2.09e-
↪05       6.86e-07        1.12e-04      -3.68e-06      -5.57e-05       1.83e-06
  0-1.0A  -> 15-3.0A    7.448231   60074.0   166.5         5.33e-07       2.92e-06      -6.54e-
↪04      -9.63e-04       -4.40e-04      -6.48e-04      -5.48e-04      -8.07e-04
  0-1.0A  -> 16-1.0A    7.837071   63210.2   158.2         1.24e-02       6.45e-02      -2.40e-
↪02       3.67e-03        2.09e-01      -3.20e-02      -1.37e-01       2.10e-02
```

7.42.5 OPS Full list of keywords

The folowing list of keywords may be included directly in the correcponding module block to trigger OPS compute the corresponding intensities.

```
**DoCD**                  (to request circular dichroism calculation)
**DoDipoleLength**        (to request the use of electric moments in a length formulation)
**DoDipoleVelocity**      (to request the use of electric moments in a velocity formulation)
**DoHigherMoments**       (to request the calculation of electric quadrupole and magnetic␣
↪dipole moments contributions)
**DoFullSemiclassical**   (to request the calculation of complete semiclassical multipolar␣
↪moments)
**DecomposeFoscLength**   (to request the decomposition of the oscillator strengths in a␣
↪multipolar expansion under a length formulation)
```

(continues on next page)

(continued from previous page)

```
**DecomposeFoscVelocity** (to request the decomposition of the oscillator strengths in a␣
↪multipolar expansion under a velocity formulation)
```

In Orca 6.0 the modules in which OPS is avaylable are: %cis/%tddft, %rocis, %casscf, %mrci, %mdci, %lft, %mcrpa and %xes.

7.42.6 Notes

1. All values in the OPS tables are expressed in atomic units unless otherwise specified.
2. Keyword corrections are done internally when necessary.
3. **IMPORTANT**: Input keywords have been standardized across all OPS-utilizing modules; Orca 5 keywords are no longer valid.
4. **IMPORTANT**: In CASSCF and MRCI, keywords controlling the absorption and ECD spectrum no longer belong in the "%rel" block.
5. Origin-adjusted formulation is approximated in SOC/SSC formulations.
6. OPS is not available for sTDDFT and sTDADFT.
7. The MDCI module reports spectra using left, right, and both solutions.
8. ADC2 and EOM-CCSD methods exclusively utilize ED length tables.

7.43 Magnetic properties through Quasi Degenerate Perturbation Theory

7.43.1 Quasi Degenerate Perturbation Theory (QDPT) in a nutshell

Quasi Degenerate Perturbation Theory offers a versatile and accurate approach to to a number of magnetic properties for basically every wavefunction based excited states method.

In a nutshell at the non relativistic limit for every excited state single or multireference wavefunction based method, bearing a CASSCF, MRCI, or a ROCIS type of zeroth order wavefunction one can set up an excitation problem that is a combination of the zeroth order wavefunction and excited spin-adapted configuration state functions (CSFs)$\left|\Phi_\mu^{SS}\right\rangle$.

That takes the form:

$$\left|\Psi_I^{SS}\right\rangle = \sum\nolimits_\mu C_{\mu I} \left|\Phi_\mu^{SS}\right\rangle \tag{7.263}$$

Here the upper indices SS stand for a wave function of the spin quantum number S and spin projection $M_S = S$. Since the BO Hamiltonian does not contain any complex-valued operator, the solutions $\left|\Psi_I^{SS}\right\rangle$ may be chosen to be real-valued.

By obtaining a solution to the above eigenvalue problem provides the coefficients with which the CSFs enter into the chosen wavefunction as well as the eigenstates of the spin-free operator. These eigenstates maube used to expand towards the respective relativistic eigenstates by setting up the relevant quasi-degenerate eigenvalue problem. In fact the spin-orbit coupling (SOC), the spin-spin coupling (SSC) effects along with the Zeeman interaction can be included by means of the quasi-degenerate perturbation theory (QDPT). In this approach the SOC, the SSC, and the Zeeman operators are calculated in the basis of pre-selected solutions of the BO Hamiltonian$\left\{\Psi_I^{SM}\right\}$.

$$\left\langle \Psi_I^{SM} \left| \hat{H}_{\mathrm{BO}} + \hat{H}_{\mathrm{SOC}} + \hat{H}_{\mathrm{SSC}} + \hat{H}_{\mathrm{Z}} \right| \Psi_J^{S'M'} \right\rangle = \delta_{IJ}\delta_{SS'}\delta_{MM'}E_I^{(S)} + \left\langle \Psi_I^{SM} \left| \hat{H}_{\mathrm{SOC}} + \hat{H}_{\mathrm{SSC}} + \hat{H}_{\mathrm{Z}} \right| \Psi_J^{S'M'} \right\rangle \tag{7.264}$$

Diagonalization of this matrix yields the energy levels and eigenvectors of the coupled states. These eigenvectors in fact represent linear combinations of the solutions of $\hat{H}_{\mathrm{BO}}$ with complex coefficients.

The effective one-electron SOC operator in second quantized form can be written as [610]:

$$\hat{H}_{\text{SOMF}} = \frac{1}{2}\sum_{pq} z_{pq}^{-}\hat{a}_p^{\uparrow}\hat{b}_q + z_{pq}^{+}\hat{b}_p^{\uparrow}\hat{a}_q + z_{pq}^{0}\left[\hat{a}_p^{\uparrow}\hat{a}_q - \hat{b}_p^{\uparrow}\hat{b}_q\right] \tag{7.265}$$

Here $\hat{a}_p^{\uparrow}$ and $\hat{b}_p^{\uparrow}$ stand for creation of α and β electrons respectively; $\hat{a}_p$ and $\hat{b}_p$ represent the corresponding annihilation operators. The matrix elements $z_{pq}^{-} = z_{pq}^{x} - iz_{pq}^{y}$, $z_{pq}^{+} = z_{pq}^{x} + iz_{pq}^{y}$, and $z_{pq}^{0} = z_{pq}^{z}$ (upper x, y, z indices denote the Cartesian components) are constructed from the matrix elements described in section *Zero-Field-Splitting*.

In this concept the SOC Hamiltonian reads:

$$\left\langle \Psi_I^{SM} \middle| \hat{H}_{\text{SOC}} \middle| \Psi_J^{S'M'} \right\rangle = \sum_{m=0,\pm1} (-1) \begin{pmatrix} S' & 1 \\ M' & m \end{pmatrix} \underbrace{\left\langle \Psi_I^{SS} || H_{-m}^{SOC} || \Psi_J^{SS} \right\rangle}_{Y_{II'}^{SS'}(m)} \tag{7.266}$$

where m represents the standard vector operator components. $\begin{pmatrix} S' & 1 \\ M' & m \end{pmatrix}$ is a Clebsch–Gordon coefficient that has a single numer- ical value that is tabulated. It satisfies certain selection rules and contains all of the M-dependence of the SOC matrix elements. The quantity $Y_{II'}^{SS'}(m)$ is a reduced matrix element. It only depends on the standard components of the two states involved. There are only three cases of non-zero $Y_{II'}^{SS'}(m)$ which arise from state pairs that either have the same total spin or differ by one unit.[621]

The SSC Hamiltonian reads:

$$\hat{H}_{\text{SSC}} = -\frac{3g_e^2\alpha^2}{8}\sum_{i\neq j}\sum_{m=0,\pm1,\pm2} \frac{(-1)^m}{r_{ij}^5} \left[\mathbf{r}_{ij} \times \mathbf{r}_{ij}\right]_{-m}^{(2)} \left[\mathbf{S}(i) \times \mathbf{S}(j)\right]_{m}^{(2)} \tag{7.267}$$

For matrix elements between states of the same multiplicity it can be simplified to

$$\begin{aligned} \left\langle aSM \middle| \hat{H}_{\text{SSC}} \middle| a'SM' \right\rangle = & \frac{\sqrt{(S+1)(2S+3)}}{\sqrt{S(2S-1)}} \\ & \times \sum_m (-1)^m \begin{pmatrix} S' & 2 \\ M' & m \end{pmatrix} \sum_{pqrs} D_{pqrs}^{(-m)} \left\langle aSS \middle| Q_{pqrs}^{0} \middle| a'SS \right\rangle \end{aligned} \tag{7.268}$$

Here

$$Q_{pqrs}^{(0)} = \frac{1}{4\sqrt{6}}\left\{E_{pq}\delta_{sr} - S_{ps}^{z}S_{rq}^{z} + \frac{1}{2}\left(S_{pq}^{z}S_{rs}^{z} - E_{pq}E_{rs}\right)\right\} \tag{7.269}$$

represents the two-electron quintet density. The operators $E_{pq} = \hat{a}_p^{\uparrow}\hat{a}_q + \hat{b}_p^{\uparrow}\hat{b}_q$ and $S_{pq}^{z} = \hat{a}_p^{\uparrow}\hat{a}_q - \hat{b}_p^{\uparrow}\hat{b}_q$ symbolize here the one-electron density operator and the spin density operator accordingly. The spatial part

$$D_{pqrs}^{(0)} = \frac{1}{\sqrt{6}}\iint \varphi_p(\mathbf{r}_1)\varphi_r(\mathbf{r}_2)\frac{3r_{1z}r_{2z} - \mathbf{r}_1\mathbf{r}_2}{r_{12}^5}\varphi_q(\mathbf{r}_1)\varphi_s(\mathbf{r}_2)\,d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.270}$$

denotes the two-electron field gradient integrals. These two-electron integrals can be evaluated using the RI approximation.

Finally, the Zeeman Hamiltonian is included in the form of:

$$\hat{H}_{\text{Z}} = \mu_B\left(\hat{\mathbf{L}} + g_e\hat{\mathbf{S}}\right)\mathbf{B} \tag{7.271}$$

with $\hat{\mathbf{L}}$ representing the total orbital momentum operator, and $\hat{\mathbf{S}}$ being the total spin operator.

In this concept solution of a selected relativistic Hamiltonian provide access to a numerous magnetic properties namely EPR properties *EPR and NMR properties* as well as Magnetization and Susceptibility properties *Magnetization and Magnetic Susceptibility* In addition monitoring the impact of an external Magnetic Field to the relativistic eigenstates and eigenvectors *Addition of Magnetic Fields* becomes straightforward.

Collectively within the QDPT framework the following magnetic properties become available

1) G-Tensor/Matrix

2) Zero Field Splitting

3) Hyperfine A-Tensor/Matrix

4) Electric Field Gradient

5) Magnetization

6) Susceptibility

7) Inclusion of Magnetic Fields

7.43.2 Magnetic properties through the Effective Hamiltonian

Since both the energies and the wavefunction of the low-lying spin-orbit states are available, the effective Hamiltonian theory can be used to extract EPR parameters such as the full G, Zero Field Splitting (ZFS) and hyperfine A tensors.

Provided that the ground state is non-degenerate. By applying this Hamiltonian on the basis of the model space, i.e. the $|S, M_S\rangle$ components of the ground state, the interaction matrix is constructed.

The construction of effective Hamiltonian relies on the information contained in both the energies and the wavefunctions of the low-lying spin-orbit states. Following des Cloizeaux formalism, the effective Hamiltonian reproduces the energy levels of the "exact" Hamiltonian E_k and the wavefunctions of the low-lying states projected onto the model space $\tilde{\Psi}$:

$$\hat{H}_{\text{eff}}|\tilde{\Psi}_k\rangle = E_k|\tilde{\Psi}_k\rangle$$

These projected vectors are then symmetrically orthonormalized resulting in an Hermitian effective Hamiltonian, which can be written as:

$$\hat{H}_{\text{eff}}|\tilde{\Psi}\rangle = \sum_k |S^{-\frac{1}{2}}\tilde{\Psi}_k\rangle E_k \langle S^{-\frac{1}{2}}\tilde{\Psi}_k|$$

The effective interaction matrix obtained by expanding this Hamiltonian into the basis of determinants belonging to the model space, is the compared to the matrix resulted from expanding the model Hamiltonian. Based on a singular value decomposition procedure, all 9 elements of the G, A and/or ZFS tensors may be extracted.

7.43.3 Organization of QDPT Magnetic Properties Computation

Starting from ORCA 6.0 the calculation of the magnetic properties through the Quasi Degenerate Perturbation Theory (QDPT) in all available correlation type modules is unified and simplified. Following the general architecture design of ORCA 6.0 the computation of all the involved magnetic properties are centrally performed by a driver data structure called the QDPT Driver. The Driver takes into account all the specific variables that are populated by the involved module and proceeds accordingly to calculated and represent the requested property in a uniform fashion. This presently involves the `casscf`, `mrci`, `rocis` and `lft` modules

In this way

1) Results analysis process from the user's perspective is simplified

2) Cross module correlation and comparisons are also easily accessible

The general keywords that activate the generation of QDPT properties are:

```
%method (casscf, mrci, lft, ...)
  relativistic block (soc, rel, ...)
    DoSOC true   # include the SOC contribution
    DoSSC true   # include the SSC contribution
  end
end
```

In a first step SOC contributions will be computed for any level of theory that is available

```
---------------------------------
QDPT WITH CASSCF/NEVPT2/MRCI... DIAGONAL ENERGIES
---------------------------------


************************************
COMPUTING QDPT HAMILTONIAN
************************************

************************************
Doing QDPT with ONLY SOC!
************************************
```

Upon request the non-zero SOC Matrix Elemnts will be printed

```
-----------------------------------
NONZERO SOC MATRIX ELEMENTS (cm**-1)
-----------------------------------

           Bra                         Ket
<Block Root  S    Ms  | HSOC |  Block Root  S    Ms>    =  Real-part     Imaginary part
-------------------------------------------------------------------------------------
   0    2  1.0  1.0              0     1  1.0  1.0            0.000            -71.172
   0    3  1.0  1.0              0     2  1.0  1.0            0.000              1.542
   0    4  1.0  1.0              0     0  1.0  1.0            0.000             50.048
   0    5  1.0  1.0              0     0  1.0  1.0            0.000            -48.827
   0    5  1.0  1.0              0     4  1.0  1.0            0.000            -40.119
   0    6  1.0  1.0              0     1  1.0  1.0            0.000             -0.197
   0    6  1.0  1.0              0     3  1.0  1.0            0.000              8.724
   0    7  1.0  1.0              0     1  1.0  1.0            0.000             -2.695
```

followed by printing of the SOC Hamiltonian

```
Note: In the following the full <I|HBO+SOC|J> are printed in the CI Basis.
      I,J are compound indices for |Block/Mult, Ms, Root>, where the states
      are ordered first by MultBlock, then Ms and finally Root.
-----------------
SOC MATRIX (A.U.)
-----------------
```

leading to the printout of the relativistically corrected eigenvalues and eigenvectors

```
Lowest eigenvalue of the SOC matrix:    -149.86223277 Eh
Energy stabilization:    -2.54512 cm-1
Eigenvalues:     cm-1         eV      Boltzmann populations at T =  300.000 K
   0:            0.00        0.0000       3.36e-01
   1:            2.37        0.0003       3.32e-01
   2:            2.37        0.0003       3.32e-01
   3:         7757.65        0.9618       2.33e-17
   4:         7757.66        0.9618       2.33e-17
   5:        11913.81        1.4771       5.15e-26

...

The threshold for printing is 0.0100
Eigenvectors:
                          Weight      Real          Image    : Block Root    Spin   Ms
 STATE   0:       0.0000
                          0.388265     0.410320    -0.468937 :     0    0    1    1
                          0.223270    -0.000000     0.472514 :     0    0    1    0
                          0.388265     0.410320     0.468937 :     0    0    1   -1
 STATE   1:       2.3703
```

(continues on next page)

(continued from previous page)

```
                        0.310686     0.534586     0.157809 :     0    0    1    1
                        0.378606     0.000017    -0.615309 :     0    0    1    0
                        0.310706     0.534623    -0.157747 :     0    0    1   -1
 STATE   2:       2.3703
                        0.300970    -0.214003    -0.505146 :     0    0    1    1
                        0.398078    -0.000007    -0.630934 :     0    0    1    0
                        0.300949    -0.214019     0.505119 :     0    0    1   -1

...
```

Then in following all the relevant QDPT properties will be printed:

```
************************************
COMPUTING QDPT PROPERTIES
************************************

1) for G-Tensor/Matrix
-------------------------------------------
ELECTRONIC G-MATRIX FROM EFFECTIVE HAMILTONIAN
-------------------------------------------

2) for ZFS, on the basis of the 2nd Order and Effective Hamiltonian approximations

-------------------------------------------
           ZERO-FIELD SPLITTING
2ND ORDER SOC CONTRIBUTION
-------------------------------------------
-----------------------------------------------------------
                  ZERO-FIELD SPLITTING
EFFECTIVE HAMILTONIAN SOC CONTRIBUTION
-----------------------------------------------------------

3) for A-Tensor/Matrix
-------------------------
QDPT HFC A-MATRICES
-------------------------

4) for Electric Field Gradient Tensor

---------------------------------------
EFG TENSOR
---------------------------------------

5) For Mahnetization/Susceptibility

-------------------------------------------------
SOC CORRECTED MAGNETIZATION AND/OR SUSCEPTIBILITY
-------------------------------------------------

6) For External Magnetic Fields Contributions

-------------------------------------------------------------------------------------------------------
↪-----------
              SOC TRANSITION MAGNETIC DIPOLE CONTRIBUTIONS IN EXTERNAL MAGNETIC FIELD
          Magnetic field Bx =       1.00 Gauss  By =       0.00 Gauss  Bz =       0.00 Gauss
-------------------------------------------------------------------------------------------------------
↪-----------
 States       Energy      Energy     Osh.Str           M2            MX             MY          ␣
↪    MZ
              (cm-1)       (eV)        (au)          (au**2)        (au)           (au)        ␣
↪   (au)
```

(continues on next page)

(continued from previous page)

```
-----------------------------------------------------------------------------------------------
↪-----------
 0  0         0.00       0.0000     0.00000000     0.00000145     0.00090159     0.00061858   ␣
↪ 0.00050413
 0  1        20.20       0.0025     0.00000000     0.00000126     0.00042808     0.00067630   ␣
↪ 0.00078545
 1  1         0.00       0.0000     0.00000000     0.00000000     0.00003430     0.00002347   ␣
↪ 0.00001909
-----------------------------------------------------------------------------------------------
↪-----------

Following this as Discussed in One Photon Spectroscopy Section {ref}`sec:ops.detailed` all␣
↪relativistically corrected optical spectra will also printed under the same  correction␣
↪scheme

-----------------------------------------------------------------------------------------------
↪---------
      SOC CORRECTED ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-----------------------------------------------------------------------------------------------
↪---------
      Transition         Energy     Energy  Wavelength fosc(D2)      D2       |DX|      |DY|  ␣
↪    |DZ|
                          (eV)      (cm-1)    (nm)  (*population)  (au**2)    (au)      (au)  ␣
↪    (au)
-----------------------------------------------------------------------------------------------
↪---------

...
```

If requested in a second step Spin-Spin Coupling contributions will be generated and wiil be added to the SOC Hamiltonian to generate SOC+SSC contributions

```
----------------------------------------
Calculating Spin-Spin Coupling Integrals
----------------------------------------
```

The the program will undergo the exact same analysis as above printing the SOC+SSC analysis

```
************************************************************
* DOING EVERYTHING A SECOND TIME: THIS TIME INCLUDING SSC *
************************************************************

************************************
COMPUTING QDPT HAMILTONIAN
************************************

************************************
Doing QDPT with SOC AND SSC!
************************************

-------------------------------------
NONZERO SOC and SSC MATRIX ELEMENTS (cm**-1)
-------------------------------------

          Bra                         Ket
<Block Root  S    Ms  | HSOC + HSSC |  Block Root  S    Ms>    =  Real-part     Imaginary part
-------------------------------------------------------------------------------------------
  0    2  1.0  1.0                      0     1  1.0  1.0           -0.000            -71.172
  0    3  1.0  1.0                      0     1  1.0  1.0           -0.001              0.000
  0    3  1.0  1.0                      0     2  1.0  1.0            0.020              1.542
  0    4  1.0  1.0                      0     0  1.0  1.0           -0.222             50.048
```

(continues on next page)

(continued from previous page)

```
  0    5  1.0  1.0                    0     0  1.0  1.0          -0.228          -48.827
  0    5  1.0  1.0                    0     4  1.0  1.0           0.332          -40.119
  0    6  1.0  1.0                    0     1  1.0  1.0          -0.030           -0.197
```

for both the Magnetic properties e.g. the ZFS

```
-----------------------------------------------------------
                   ZERO-FIELD SPLITTING
EFFECTIVE HAMILTONIAN SOC and SSC CONTRIBUTION
-----------------------------------------------------------
```

as well as the Optical properties

```
-------------------------------------------------------------------------------------------------
↪---------
  SOC+SSC CORRECTED ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------------------------
↪---------
     Transition         Energy     Energy  Wavelength fosc(D2)      D2       |DX|      |DY|  ␣
↪    |DZ|
                         (eV)      (cm-1)    (nm)  (*population)  (au**2)    (au)      (au)  ␣
↪    (au)
-------------------------------------------------------------------------------------------------
↪---------
```

7.44 Simulation of (Magnetic) Circular Dichroism and Absorption Spectra

7.44.1 General description of the program

ORCA can now simulate optical spectra that include spin-orbit coupling contributions at all levels of theory by using a common implementation. [269]

Following the energy-loss approach, the absorption cross section for a transition between states $\tilde{P}$ and $\tilde{Q}$ can be expressed as:

$$\sigma_{\tilde{P}\tilde{Q}} = \frac{4\pi^2}{c(E_{\tilde{Q}} - E_{\tilde{P}})}|T_{\tilde{P}\tilde{Q}}|^2,$$

where c is the speed of light; $E_{\tilde{P}}$ and $E_{\tilde{Q}}$ are the energy of the states $\tilde{P}$ and $\tilde{Q}$, respectively; $T_{\tilde{P}\tilde{Q}}$ is the transition moment between states $\tilde{P}$ and $\tilde{Q}$ and it can be computed with different expressions based on the applied approximation.

Under a dipolar approximation to the light-matter interaction, the transition moment takes the form:

$$T_{\tilde{P}\tilde{Q}} = \sum_{i=1}^{N} \langle \tilde{P}| \mathcal{E} \cdot \hat{\mathbf{p}}_i |\tilde{Q}\rangle,$$

where the sum runs over all electrons i; $\hat{\mathbf{p}}_i$ is the linear momentum operator; and $\mathcal{E}$ is the polarization vector of the incident light.

In order to take into account all the electric and magnetic mechanisms in the transition, it is necessary to use the full field-matter interaction operator (FFMIO). For transition moment, it leads to the equation (7.272).

$$T_{\tilde{P}\tilde{Q}} = \frac{e}{m_e} \sum_{i=1}^{N} \langle \tilde{P}| \mathcal{E} \cdot \left[e^{i\mathbf{k}\cdot\hat{\mathbf{r}}_i} \hat{\mathbf{p}}_i \right] |\tilde{Q}\rangle \tag{7.272}$$

where $\hat{r}_i$ is the position operator of i-th electron and $\mathbf{k}$ is the wave vector that points in the direction of the light propagation whose magnitude is related to the wavelength by $\lambda = 2\pi/|\mathbf{k}|$.

For free-rotating molecules, it is necessary to consider all possible orientations of the molecule with respect to the direction of the incident light. In some cases, such as the absorption of linear-polarized light under a dipolar approximation, the effect of the orientation can be averaged analytically. However, numerical integration over some selected molecular orientations, labeled as o in equation (7.273), is generally necessary.

$$< \sigma_{\tilde{P}\tilde{Q}} >= \sum_o w_o \frac{4\pi^2}{c(E_{\tilde{Q}}(o) - E_{\tilde{P}}(o))} |T_{\tilde{P}\tilde{Q}}(o)|^2 \tag{7.273}$$

where w_o is the weight of the orientation in the quadrature.

The implementation has been designed to compute the absorption of circularly-polarized light on systems under the effect of an additional external magnetic field, B, which modifies the states $\tilde{P}$ and $\tilde{Q}$ for each orientation o through a Zeeman perturbation. The computed results are presented as the difference in the absorption of the left (-) and right (+) circularly-polarized light (Δf_{osc}) and as the sum of the oscillator strength (f_{osc}), which corresponds to the linearly-polarized light absorption. The molecular orientations are constructed by using rotation matrices with three Euler angles: χ, θ, and ϕ. Herein, χ (the rotation angle on a plane perpendicular to the direction of external magnetic field/incident light) is integrated analytically whereas θ and ϕ are taken on a grid.

Finally, the states $\tilde{P}$ and $\tilde{Q}$ are obtained from QDPT by expanding the states over non-relativistic eigenstates of $\hat{H}_0$ ($\{I, J\}$) and the coefficients of the expansion ($d_{I\tilde{Q}}$) are obtained from the diagonalization of the complex matrix, which contains $\hat{H}_0$ as well as the SOC and Zeeman contributions (and SSC, if it is implemented in the selected electronic structure theory) expressed in $\{I, J\}$.

$$\langle \Psi_I^{SM} | \hat{H}_0 + \hat{H}_{SOC} + \hat{H}_{Zeeman} | \Psi_J^{S'M'} \rangle = \delta_{IJ}\delta_{SS'}\delta_{MM'}E_I^S + \langle \Psi_I^{SM} | \hat{H}_{SOC} + \hat{H}_{Zeeman} | \Psi_J^{S'M'} \rangle \tag{7.274}$$

7.44.2 Running and analyzing MCD calculations in TDDFT module

A minimum input to compute the Magnetic Circular Dichroism (MCD) requires setting the keyword DoMCD to true and to include an intensity for the external magnetic field B (in Gauss). An example input for the TD-DFT module is as follows:

```
! B3LYP def2-QZVPP

%tddft
    TDA False
    NROOTS 10
    DoSOC True
    DoMCD True
    B 50000.0
end

*xyz 0 1
 O     0.24287127056830      0.00033994295362      0.29479344369591
 C     0.13113617562040      0.00013705972874      1.65093880501262
 C     1.35780003491446     -0.00017874911331      2.22574232397512
 C     2.30247996517358     -0.00017814489451      1.14947275232337
 C     1.57300244613625      0.00013652073221      0.00793218434497
 H    -0.86951534119903      0.00025858475406      2.04160438459058
 H     1.56899490845286     -0.00038874637444      3.28047030533440
 H     3.37570289422870     -0.00038632501973      1.22171814100713
 H     1.83015664610447      0.00025985723335     -1.03506234028415
*
```

In the output file of this job, the estimated oscillator strengths (f_{osc}) and the difference between left and right circularly polarized light absorption (Δf_{osc}) are provided:

```
-------------------------------------------------------------------------------
           MCD Transitions via transition electric dipole moments
                B =  50000.00 Gauss  T =   300.000 K
-------------------------------------------------------------------------------

                   dfosc             fosc           dfosc/fosc

 0 -> 1       -0.0000000000       0.0000000003      -0.0663180503
 0 -> 2        0.0000000000       0.0000000005       0.0000209714
 0 -> 3        0.0000000000       0.0000000003       0.0661609417
 0 -> 4       -0.0000000001       0.0000000004      -0.2991916107
 0 -> 5       -0.0000000000       0.0000000006      -0.0009270335
 0 -> 6        0.0000000001       0.0000000004       0.2984154405
 0 -> 7        0.0000000003       0.0000000008       0.3721777693
 0 -> 8       -0.0000000000       0.0000000009      -0.0002960328
 0 -> 9       -0.0000000003       0.0000000008      -0.3689867758
 0 ->10       -0.0000050986       0.1741092913      -0.0000292840
```

These results may not be accurate when the energetic order of the states changes with respect to the relative orientation between the molecule and the external magnetic field. To obtain accurate results, it is necessary to perform a post-processing step for all orientations using the `orca_mapspc` program, which saves the results in a file that has the `.cis-el.dipole-length.1.mcd` extension:

```
orca_mapspc fur-mcd.cis-el.exact.1.mcd MCD -x050000 -x155000 -n10000 -w2000
```

In this example, we generate the spectrum (M^{-1}) between 50000 and 55000 cm^{-1} (-x050000 -x155000), using 10000 points (-n10000) and including a broadened normalized Gaussian function with a full width at half maximum of 2000cm^{-1} (-w2000).

Multiple MCD calculations can be performed in one run by setting multiple values for B. Transition moments can be also obtained through ED velocity formulation and FFMIO operator by setting the keywords `DoVelocity` and `DoQuadrupole` to true, respectively:

```
! B3LYP def2-QZVPP

%tddft
    TDA False
    nroots 10
    DoSOC True
    DoMCD True
    DoDipoleVelocity True
    DoFullSemiClassical True
    B 50000.0, 0.0
end
```

The results are printed separately in the output file for each setting:

```
-------------------------------------------------------------------------------
           MCD Transitions via transition electric dipole moments
                B =  50000.00 Gauss  T =   300.000 K
-------------------------------------------------------------------------------

                   dfosc             fosc           dfosc/fosc

 0 -> 1       -0.0000000000       0.0000000003      -0.0663180503
 .
 .
 .
 -------------------------------------------------------------------------------
           MCD Transitions via transition velocity dipole moments
                B =  50000.00 Gauss  T =   300.000 K
-------------------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
                  dfosc              fosc            dfosc/fosc

 0 -> 1       -0.0000000001       0.0000000013      -0.0477214281
.
.
.

-----------------------------------------------------------------------------
           MCD Transitions via full Semi-classical formulation
                B =  50000.00 Gauss  T =   300.000 K
-----------------------------------------------------------------------------

                  dfosc              fosc            dfosc/fosc

 0 -> 1       -0.0000000001       0.0000000016      -0.0731029604
.
.
.
```

Post-processing results are saved in the files having the `.cis-el.dipole-length.1.mcd`, `.cis-el.dipole-vel.1.mcd`, and `.cis-el.exact.1.mcd` extensions.

NOTE: It is worth enphasizing that the computed values of Δf_{osc} correspond to the difference in absorption between left and right circularly polarized light for the selected transition moments. In the case of both ED approximations, Δf_{osc} corresponds to the MCD signal. The sum of natural circular dichroism and magnetic-induced circular dichroism is obtained when the FFMIO is requested. To obtain only the MCD spectrum in an FFMIO scheme, it is necessary to subtract the natural circular dichroism by setting B to 0.0.

7.44.3 Running MCD calculations in other modules

The MCD implementation can also be used in other modules such as STEOM-CCSD, CAS, ROCIS, and MRCI (see the input file examples given below) by using the same keywords as those described for the TDDFT module. In the case of CAS, ROCIS, and MR-CI modules, it is necessary to include the keyword `NewMCD True`; otherwise, the previous MCD implementation will be called instead.

```
%mrci
       soc
         DoSOC True
         DoVelocity True
         DoQuadrupole True
         DoMCD True
         newMCD True
         B 50000.0
      end
end
```

```
%casscf
    DoDipoleVelocity True
    DoFullSemiClassical True
    rel
        DoSOC True
        DoMCD True
        B 50000.0
        Temperature 300.0
    end
end
```

```
%rocis
        SOC True
        DoMCD True
        B 50000.0
        DoDipoleVelocity True
        DoFullSemiClassical True
end
```

```
%mdci
        DoSOC True
        DoMCD True
        DoVelocity True
        DoQuadrupole True
        B 50000.0
        Temperature 300.0
end
```

It is important to keep in mind that the calculation of the MCD relies on the proper description of the transition moments and angular momentum for calculating the Zeeman perturbation. Therefore, the user is responsible for selecting the proper electronic structure method.

7.44.4 List of related keywords

```
%selected module
  DoMCD False         # Enables the use of the MCD module
  DoDipoleVelocity True # Use the electric dipolar velocity formulation
                        # for the light-matter interaction
  DoFullSemiClassical False  # Use the full semiclasical ligth-matter
                             # interaction transition moments
  MCDGridtype 1       # Grid for the molecular orientational average
                             # 1 = Lebedev grid (default)
                             # 2 = Regular grid
  MCDLebedev 14       # Number of points if Lebedev grid is selected
                      # 6, 12, 14(default), 26, 50, 110, 194, 302, 434, 590, 770
  NPOINTSPHI 10       # Number of points for phi angle if regular grid is selected
  NPOINTSTHETA 10     # Number of points for theta angle if regular grid is selected
  B 3000.0            # Magnetic field in Gauss
  Temperature 300.0   # Self-explanatory. One temperature must be defined
                      # for each B value
end
```

7.45 More on the Excited State Dynamics module

ORCA has now a module designed to calculate properties related to excited states named `ORCA_ESD`. It can be used to predict absorption/emission spectra, transition rates, resonant Raman, and MCD spectra, based on a path integral approach to the dynamic process [199]. It has some of the functionalities of `ORCA_ASA` and even more, as it will be discussed. What we do here is NOT a conventional dynamics with trajectories along time points, we rather solve the equation for the transition rates or intensities depending on the different cases considered.

This formulation works because there is an analytic solution to the path integral of the Multidimensional Harmonic Oscillator and the assumption of Harmonic nuclear movement is critical. In many cases that approximation does hold and the results are in very good agreement with the experiment. The general usage of the `ORCA_ESD` module and some examples are already presented on Sec. *Excited State Dynamics* and it is recommended to read that before going into the details here. We now will discuss the specifics and keywords related to of each part of the module. A complete keyword list can be found at the end of this section.

7.45.1 Absorption and Emission Rates and Spectrum

General Aspects of the Theory

The idea behind calculating the absorption or emission rates starts with the equation from the quantization of the electromagnetic field for the transition rates between and initial and a final state:

$$k(\omega)_{if} = \frac{4\omega^3 n^2}{3\hbar c^3} |\langle \Psi_i | \widehat{\mu} | \Psi_f \rangle|^2 \delta(E_i - E_f \pm \hbar\omega) \tag{7.275}$$

with $\hbar\omega$ being the energy of the photon, $\widehat{\mu}$ the dipole operator and n the refractive index of the solvent, as suggested by Strickler and Berg [830].

One way to obtain $k(\omega)$ is to compute it in the frequency domain, by calculating the Franck-Condon Factors between all initial and final states that satisfy the Dirac delta in Eq. *(7.275)*, considering the thermally accessible initial states with the appropriate weight,

$$k^{obs} = \int k(\omega) d\omega, \qquad k(\omega) = \sum_{if} P_i(T) k_{if}(\omega), \tag{7.276}$$

where $P_i(T) = e^{-\frac{\epsilon_i}{k_B T}}/Z$ is the Boltzmann population of a given initial state at temperature T, ϵ_i is the total vibrational energy of state i and Z is the vibrational partition function. However, this can lead to a very large number of states to be included, particularly if there are low frequency modes. In this work we will take the a different approach and switch to the time domain, by using the Fourier Transform representation of the Dirac delta,

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t} dt, \tag{7.277}$$

so that the equation to solve, in atomic units, is:

$$k(\omega) = \frac{2\omega^3}{3\pi c^3 Z} \sum_{if} e^{-\frac{\epsilon_i}{k_B T}} \langle \Theta_i | \vec{\mu}^e | \bar{\Theta}_f \rangle \langle \bar{\Theta}_f | \vec{\mu}^e | \Theta_i \rangle \int e^{i(E_i - E_f - \omega)t} dt,$$

with $\vec{\mu}^e$ being the "electronic transition dipole" and $|\Theta\rangle$ the vibrational wavefunction of the initial or final state.

After some extra steps, redefinition of the time variable and insertion of a resolution of identity, it can be shown that this equation is ultimately simplified to a Discrete Fourier Transform (DFT) of a correlation function $\chi(t)$ with a timestep Δt, multiplied by a prefactor α [199]:

$$\begin{aligned} k(\omega) &= \alpha \int Tr(\vec{\mu}^e e^{-i\widehat{H}\tau} \vec{\mu^e} e^{-i\widehat{\bar{H}}\bar{\tau}}) e^{i\Delta E t} e^{-i\omega t} dt \\ &= \alpha \int \chi(t) e^{-i\omega t} dt \\ &= 2\alpha \ \mathfrak{Re} \int_0^{\infty} \chi(t) e^{-i\omega t} dt \\ &\simeq 2\alpha\Delta t \ \mathfrak{Re} \ \ DFT\{\chi(t)\}, \end{aligned}$$

and this correlation function is then calculated using path integrals analytically at each time point t.

If one considers that the electronic part of the transition dipole varies with nuclear displacements and we allow for it to depend on the normal coordinates ($\mathbf{Q}$), such as:

$$\vec{\mu}^e(\mathbf{Q}) = \vec{\mu}_0^e + \sum_i \left. \frac{\partial \vec{\mu}^e}{\partial Q_i} \right|_{\mathbf{Q}=0} Q_i + \ldots, \tag{7.278}$$

we can even include vibronic coupling or the so-called Herzberg-Teller (HT) effect. The Frank-Condon (FC) approximation keeps only the coordinate-independent term. The correlation function for the HT cases can then be derived recursively from the FC one and the calculation is done quite efficiently. It is important to say the one must choose ONE set of coordinates in order to expand the transition dipole. In our formulation, it is always that of the FINAL state and that has some implications discussed below.

Other important aspect of this theory is that, in order to solve the path integrals, one has to work in one set of coordinates, the initial ($\mathbf{Q}$) or the final state ones ($\bar{\mathbf{Q}}$, with the bar indicating final coordinates). As we have a transition matrix element, one set of coordinates have to be transformed into the other and it is easy to show that they are related by

$$\mathbf{Q} = \mathbf{J}\bar{\mathbf{Q}} + \mathbf{K}. \tag{7.279}$$

That was first proposed by Duschinsky in the late 1930s[228] with the Duschinsky rotation matrix $\mathbf{J}$ and the displacement vector $\mathbf{K}$ defined as

$$\mathbf{J} = \mathbf{L}_x^{\mathrm{T}}\bar{\mathbf{L}}_x, \qquad \mathbf{K} = \mathbf{L}_x^{\mathrm{T}}(\bar{\mathbf{q}}_\mathbf{0} - \mathbf{q}_\mathbf{0}),$$

with $\mathbf{L}_x$ being the matrix containing the normal modes, here described in Cartesian coordinates (x), and $\mathbf{q_0}$ begin mass weighted coordinates ($q_i = \sqrt{m_i}x_i$).

The program runs by first reading and obtaining the initial and final state geometries and Hessians, then computes the Duschinsky rotation matrix and displacement vector, calculates the derivatives for the transition dipoles and computes the correlation function. After that, the DFT is done and the rates are obtained and printed when necessary. As the intensities observed experimentally are proportional to the rates, the spectrum is also calculated and printed on a BASENAME.spectrum file. If PRINTLEVEL HIGH is requested under %ESD, the correlation function is also printed on a BASENAME.corrfunc file.

OBS: The units for the Emission spectra are rather arbitrary, but for Absorption they are the experimental "molar absorptivity (ε)" in L mol^{-1}cm^{-1} [199]. Be aware that these are dependent on the line width of the curves.

Approximations to the excited state PES

As already mentioned in Sec. *Excited State Dynamics*, in order to predict the rates we need at least a ground state (GS) and an excited state (ES) geometry and Hessian. In ORCA, we have seven different ways to approximate this ES PES: AHAS, VH, VG, HHBS, HHAS, UFBS and UFAS. Those can can be choosen by setting the HESSFLAG under %ESD. If you actually optimize the ES geometry and input the Hessian, that will be called an Adiabatic Hessian (AH) method an no keyword must be given on the input.

OBS: In the present version, these approximations are only available for Absorption, Fluorescence and resonant Raman. They cannot be directly used for phosphorescence and ISC rate calculations. However, one can do the latter calculations by generating approximate ES Hessians from "fake" fluorescence or absorption runs, and then using the ES Hessians in AH calculations (vide infra).

The idea behind these approximations is to do a geometry update step ($\mathbf{\Delta S} = -\mathbf{gH^{-1}}$ for Quasi-Newton and $\mathbf{\Delta S} = -\mathbf{g}(\mathbf{H}+\mathbf{S})^{-1}$ for Augmented Hessian) to obtain the ES structure and somehow approximate the ES Hessian. The gradient ($\mathbf{g}$) and Hessian ($\mathbf{H}$) used on the step are on column Step of Table Table 7.28 below, with a description of the final ES Hessian:

Table 7.28: Methods used to estimate the ES PES

Method	Step	ES Hessian
AHAS	ES grad + GS Hessian	calculated on the ES geometry
VH	ES grad + ES Hessian at GS geometry	calculated on the GS geometry
VG (default)	ES grad + GS Hessian	equal to GS Hessian
VGFC	ES grad + GS Hessian	equal to GS Hessian (+ APPROXADEN TRUE)
HHBS	ES grad + Hybrid ES Hessian on GS geometry	Hybrid Hessian on GS geometry
HHAS	ES grad + GS Hessian	Hybrid Hessian on ES geometry
UFBS	ES grad + Updated frequencies ES Hessian on GS geometry	Updated frequencies ES Hessian on GS geometry
UFAS	ES grad + GS Hessian	Updated frequencies ES Hessian on ES geometry

OBS: Always use the GS geometry on the input file, equal to the one in the GSHESSIAN! If you asked for OPT FREQ at the input, a .xyz file is generated with the same geometry found on the .hess. If you picked the geometry from the .hess file, remember that it is in atomic units, so you have to use BOHRS on the main input.

After the calculation of the ES PES, a file named BASENAME.ES.hess is printed and can be used in future calculations. For example, one can use it in a separate AH calculation, as if the file contains the optimized ES geometry and the exact ES Hessian (of course, in reality both are approximate). This allows one to perform e.g. phosphorescence and ISC rate calculations with approximate ES PESs even though ESD does not support these calculations directly. For example, the phosphorescence rate from a triplet state to the S_0 state, with the VG approximation for the triplet PES and the ES gradient computed at the TDDFT level, can be computed as follows:

- Run an ESD fluorescence calculation using the VG approximation, but where "IROOTMULT TRIPLET" is added to the %TDDFT block. The resulting rate and spectrum are not meaningful, but the BASENAME.ES.hess file generated from this calculation is the correct VG approximation to the triplet PES.
- Run an ESD phosphorescence calculation using the AH method, with the aforementioned BASENAME.ES.hess file as the TSHESSIAN.

This way, the first ESD calculation only uses the information of the scalar triplet state in computing the VG step, while the second ESD calculation only uses the SOC-corrected states. The same ES Hessian file can therefore be used in the calculations of all three spin sublevels of the triplet state.

In addition to the BASENAME.ES.hess file, if there was any updates on the GS Hessian, like transition dipole derivatives, a BASENAME.GS.hess is also printed.

- The step can be controlled with the GEOMSTEP flag, with QN or AUGHESS options.
- Currently all ES Hessians are calculated numerically, if you want to change the parameters related to the frequency calculations, you can do that under %FREQ (Sec. *Vibrational Frequencies*). The numerical gradient settings are under %NUMGRAD (Sec. *Numerical Gradients*).
- The ES hybrid Hessian is calculated in the same way as described in Sec. *Hybrid Hessian*, except that the "model" Hessian is the GS one.
- The ES Hessian with updated frequencies is recalculated from the same GS normal modes, after an update on the frequencies, as $\mathbf{H}_{up} = \mathbf{L}\omega_{up}^2\mathbf{L}^T$. With $\mathbf{L}$ being the normal modes and ω_{up} the updated frequencies, with negative sign being kept after the square.
- The frequencies are updated depending on a calculation of the energy after a given step. If the ES modes are equal to the GS, then a step over a coordinate δq_i that would result in an energy difference δE is given by $\delta q_i = (-\frac{g_i}{\sqrt{\omega_i}} + \frac{\sqrt{g_i^2}}{\omega_i} + 2\delta E\omega_i)/\omega_i$. The default δE used is 10^{-4} Eh, in general above the error of the methods. If the error in energy after the step is larger than a threshold given by the UPDATEFREQERR flag (default 0.20 or 20%), the gradients are calculated and the frequency recomputed. If not not, that mode and frequency are assumed to be the same.
- The Updated Frequencies method can greatly accelerate the calculation of the Hessian, for much fewer gradient calculation are necessary, although you do not correct the modes. Also, the derivatives over the modes are already computed simultaneously.
- The expected energy error δE can be changed using the UF_DELE flag.
- The default method is the VG, but the AHAS is more trustworthy for unknown systems, although a lot heavier (Sec. *A better model, Adiabatic Hessian After a Step (AHAS)* and [199]).
- Always check the sum of K_i^2 printed on the output. If that number is too high (above 8 or so), it means that the geometries are too displaced and the theory might not work on these cases (check for different coordinate systems then, Sec. *Normal modes coordinate systems*).
- Also check for RMSD between the geometries after a step. If the difference is too big, there might be problems with the step.

Mixing methods

In principle, it is possible to use different methods to compute different parts needed for the ORCA_ESD module. You could, for instance use (TD)DFT analytical gradients for the ground/excited state geometries and Hessians and a more elaborate method such as STEOM-CCSD to get the energies and transition dipoles. If you want to do that, just input the Hessians and use the DELE flag for the energy difference between the states - at their own geometry! - and TDIP x,y,z to input the transition dipole. If there is SOC and the transition dipole is complex, use TDIP x.real, x.imag, y.real, y.imag, z.real, z.imag. The program will automatically detect each case. If you don't input these, they will be obtained by the module during the run, so you can set the excited method you want and let the program figure out DELE and TDIP.

OBS: It is not advisable to mix different levels of theory during a geometry step though. If you obtained a GS Hessian from DFT, doing a step based on a CASSCF ES numerical gradient might not lead to reasonable results. The same would be problematic even for different DFT functionals.

Removal of frequencies

If, after calculating an ES Hessian you end up with negative frequencies, the calculation of the correlation function might run into trouble. The default for the module is to turn all negative frequencies positive, printing a warning if any of them was lower than -300 cm^{-1}. If that is the case, you are probably on a saddle point and not even a minimum, so results should be taken with care,

You can also choose to completely remove the negative frequencies (and the corresponding from the GS), by setting IFREQFLAG REMOVE or leave them as negative with IFREQFLAG LEAVE under %ESD.

Sometimes, low frequencies have displacements that are just too large (check on the $\mathbf{K}$ vector), or the experimental low frequency modes are too anharmonic and you might want to remove them. It is also possible to do that setting the TCUTFREQ flag (in cm^{-1}), and all frequencies below the given threshold will be removed.

Normal modes coordinate systems

When working with systems with weak bonds, such as hydrogen bonds and π stacking, or with biphenyls and similar planar molecules it is common that there will be low frequency-high amplitude modes related to the angular bending. It has been shown that, in some cases, the normal modes transformed from Cartesian coordinates might be not sufficient to describe systems with these large amplitude motion [149]. In those, the definition of normal modes in terms of some curvilinear set of coordinates such as the internal ones are more suitable.

The first order transformation from Cartesian ($\mathbf{x}$) to internal ($\mathbf{s}$) coordinates is given by Wilson's $\mathbf{B}$ matrix[426] as

$$\mathbf{s} = \mathbf{B}(\mathbf{x} - \mathbf{x_0}), \tag{7.280}$$

and here we use Baker's[68] delocalized internal coordinates as default. First, a redudant set is build and an singular value decomposition of the $\mathbf{G} = \mathbf{B}\mathbf{B}^T$ matrix is performed to obtain the non-redundant set. The latter can be generated by $\mathbf{B}' = \mathbf{U}^T\mathbf{B}$, where $\mathbf{U}$ are the eigenvectors correponding to non-zero eigenvalues of $\mathbf{G}$. Then an orthogonal set is contructed from $\mathbf{B}'' = \mathbf{G}'^{-1/2}\mathbf{U}^T\mathbf{B}$. As the eigenvectors are not conitnuous functions of the coordinates, in order to avoid using an arbitrary selection, we will follow the work of Reimers[714] and set $\mathbf{G}^{-1/2}\mathbf{U}^T = \bar{\mathbf{G}}^{-1/2}\bar{\mathbf{U}}^T$, or use the same transformation for the initial an final coordinates. Please note that this may lead to numbers larger than 1 on the Duschinsky rotation matrix, for it is an approximation and the calculated rates may vary a little. The normal modes in internal coordinates ($\mathbf{L}_s$) are then obtained from those in Cartesian ones ($\mathbf{L}_x$) as

$$\mathbf{L}_s = \mathbf{B}''\mathbf{M}^{1/2}\mathbf{L}_x, \tag{7.281}$$

and the Duschinsky relation ((7.279)) still holds[149], with the displacement vector being simply

$$\mathbf{K}_s = \mathbf{L}_s^T(\bar{\mathbf{s}} - \mathbf{s}). \tag{7.282}$$

The options available for coordinate systems can be set under COORDSYS, and can be CARTESIAN, INTERNAL (for Baker delocalized - default), WINT (for weighted internals following Swart and Bickelhaupt [837]), FCWL (force constant weighted following Lindh [526]) and FCWS (same as before, but using Swart's model Hessian).

OBS: Calculating in internal coordinates is usually better but not necessarily. If something goes wrong, you may also want to try the Cartesian option.

Geometry rotation and Duschinsky matrices

The electronic transition should not to take place simultaneously with translations and rotations[746] of the molecular structure. Before further calculations take place, the initial and final state structures are superimposed to satisfy Eckart's conditions by obtaining a rotation matrix $\mathbb{B}$ that ensures $\sum m_i(\mathbb{B}\mathbf{R_i} \times \bar{\mathbf{R}}_\mathbf{i}) = 0$ [239], with $\mathbf{R}$ being Cartesian coordinates. As the initial geometry is rotated, so must be the corresponding normal modes $\mathbf{L}_x$ also. This can be turned of by setting the flag USEB FALSE.

By the default the Duschinsky rotation matrix is set to Identity, to take advantage of our much faster algorithm. To turn that on, just set USEJ TRUE. You can check the "diagonality" of this matrix by looking at the Diagonality Index (DI), here defined as $\sqrt{\sum_i \mathbf{J}_{ii}^2 / \sum_{ij} \mathbf{J}_{ij}^2}$. A DI=1 would be a perfectly diagonal matrix. The amount of mixing between modes is rpresented by the Mixing Index, with is here is given by $\langle |J_{max}| \rangle$, or the average value of the maximum J_i of each line. If DI=1, it means each normal coordinate from the initial state is equal to a mode of the final state. When USEJ=TRUE, the largest components of the $\mathbf{J}$ matrix are printed along with the $\mathbf{K}$ vector, so you can have a better idea of how the mixing is occuring.

Derivatives of the transition dipole

The derivatives of the transition dipoles with respect to the normal coordinates of the final state can be obtained directly from the derivatives with respect to the Cartesian coordinates as

$$\mathbf{U}(\bar{\mathbf{Q}}) = \bar{\mathbf{L}}_x^T \mathbf{M}^{-1/2} \mathbf{U}(\bar{\mathbf{R}}), \tag{7.283}$$

$\mathbf{U}$ being the matrix of the x,y and z components of the derivative, $\mathbf{M}$ a $3N \times 3N$ matrix with the atomic masses along the diagonal. Also, in case one already has the derivatives with respect to the initial state , those can be transformed into the derivatives with respect to the final state by using the Duschinsky relation, assuming that $\vec{\mu}_0^e(\bar{\mathbf{Q}}) + \sum_i \frac{\partial \vec{\mu}^e}{\partial \bar{Q}_i} \bar{Q}_i = \vec{\mu}_0^e(\mathbf{Q}) + \sum_i \frac{\partial \vec{\mu}^e}{\partial Q_i} Q_i$, so that

$$\frac{\partial \vec{\mu}^e}{\partial \bar{Q}_k} = \sum_j \mathbf{J}_{jk} \frac{\partial \vec{\mu}^e}{\partial Q_j}. \tag{7.284}$$

By default, this transformation is NOT done, since Eq. (7.284) is an approximation. If you want to turn it on, set CONVDER TRUE under %ESD.

OBS: Remember that, if you already have the Cartesian derivatives over the final state, like if you use AHAS for an absorption spectrum, the conversion should be exact (although there might be numerical issues, always use a larger GRAD for frequencies!).

Alternatively, these can be calculated numerically from displacements over each normal mode. In this case, it is convenient to use the dimensionless normal coordinates $q_i = \omega_i^{1/2} Q_i$ which represent proportional displacements on the PES [677]. We use $\Delta q = 0.01$ by default, but this can be changed setting the DER_DELQ flag.

- Again, DO NOT MIX different coordinates systems. If the derivatives were calculated over one coordinate set and you decide to change it, it has to be recalculated. You can manually delete them from the BASENAME.ES.hess file.
- For hybrid functionals, you can choose to use DFT for the gradient, energy and transition dipole, and the fast simplified TDA (Sec. *Simplified TDA and TD-DFT*) only for the derivatives by seting STDA TRUE under %ESD.
- A simple trick can be used to accelerate the computation of derivatives. If the first displacement gives a transition dipole that is too close to the reference, then the derivative can be assumed to be small and just the plus displacement may be taken to compute the derivative (with an usually small error). If it is large enough, then the minus displacement is also done and central differences is used. This is the default method and can be turned off by setting FASTDER to FALSE.
- The central differences option can be altogether turned off by setting CENTRALDIFF FALSE under %ESD.

- If you are having problems, set a larger PRINTLEVEL to check the individual calculation of the derivatives, you might be having some kind of root flipping during the displacement, or some other issue.

The Fourier Transform step

After the calculation of the correlation function, it is necessary to do a Fourier Transform (FT) step. To do that numerically, it is needed to correctly choose the grid in wich the time points will be computed, for that affects how the results will be obtained in the frequency domain. We have developed a method to generate an optimal set of parameters, depending on the final spectral resolution desired [199] and it will be used by default. Even so, you can choose your own grid by setting the NPOINTS and MAXTIME (in atomic units!) flags under %ESD. There are a few comments related to that:

- Because we solve the FT using a very efficient Cooley-Tukey algorithm, the NPOINTS should be always multiple of two. You can put any number on the input, but the next larger multiple of two will be calculated and set.
- The MAXTIME should be enough so that the correlation function goes to zero. If anything goes wrong, please check the BASENAME.corrfunc file for that.
- The finer the spectral resolution chosen with SPECRES, the largest MAXTIME must be.
- If you have a larger MAXTIME, you also must increase NPOINTS, otherwise the grid will be too sparse and many oscillations will be skipped.

Spectrum options

The final spectrum is saved in a BASENAME.spectrum file, with the total spectrum, the FC and HT parts discriminated, as explained in Sec. *Excited State Dynamics*. He are some detailed about it:

- The range for which the spectrum is saved is given by default, but it can be set using SPECRANGE flag under %ESD, as SPECRANGE 10000,70000.
- All of the INPUT units should always be in CM-1, but you can choose the OUTPUT units by setting the UNIT flag to CM-1, NM or EV.
- In order to better converge the correlation function and approximate experimental spectra, a lineshape function can be used instead of the delta. The default is to use a LOREZENTIAN lineshape, but LINES can be set to DELTA, LORENTZ, GAUSS or VOIGT.
- The DELTA lineshape might lead to a correlation function that oscillates forever, so please take care with that option.
- The default line widths are LINEW 50 and INLINEW 250.
- If you use a VOIGT lineshape, the Gaussian width can be controlled separately using the INLINEW flag. By default, it will be proportial to the Lorezntian to reach the same FWHM.
- The LINEW and INLINEW are NOT the full width half maximum ($FWHM$) of these curves. However they are related to them by: $FWHM_{lorentz} = 2 \times LINEW$ and $FWHM_{gauss} = 2.355 \times INLINEW$. For the VOIGT curve, it is a little more complicated but in terms of the other FWHMs, it can be aproximated as $FWHM_{voigt} = 0.5346 \times FHWM_{lorentz} + \sqrt{(0.2166 \times FWHM^2_{lorentz} + FWHM^2_{gauss})}$.
- The resolution of the spectrum can be modified with the SPECRES flag. By default it is a fraction of the LINEW. Please be aware that higher resolution (smaller SPECRES), mean a larger grid for the correlation function and more time points to calculate on.

General

- The temperature is accounted for exactly on Absorption and Emission [199] and can be set using the TEMP flag.
- PRINTLEVEL can be set to HIGH in order to print more details, including Huang-Rhys factors which are useful for rationalizing the contribution of different vibrational modes to the rate/spectrum.
- The frequencies read from the Hessian files can be scaled by any number by setting the SCALING flag under %ESD. The default is 1.0.
- If necessary, the transition dipole can also be scaled by setting the TDIPSCALING flag.
- If you just want to compute an ES PES and stop, set WRITEHESS to TRUE and the correlation function will be skipped.
- In order to make use of the fastest algorithm, set SAMEFREQ to TRUE and the DO method will be used, assuming equal Hessians between initial and final states and maximizing the efficiency when calculating the correlation function.
- If you want to calculate phosphorescence rates, you MUST input the adiabatic energy difference DELE manually (the energy difference between each state at its own geometry). And, of course, don't forget to set the SOC module to true.

7.45.2 Intersystem crossing rates

General Aspects of the Theory

Intersystem crossing (ISC) rates between a given initial state i and a final state f can be calculated from Fermi's Golden rule:

$$k(\omega)_{if} = \frac{2\pi}{\hbar}|\langle\Psi_f|\widehat{H}_{SO}|\Psi_i\rangle|^2\delta(E_i - E_f), \tag{7.285}$$

which is quite similar to the Eq. (7.275) for Fluorescence, except for the frequency term. The same trick used there can be applied here to swtich to the time domain. Then, we are left with a simple time integration, which is not anymore difficult to solve than the equations above.

One can use all of the infrastructure already presented to compute these ISC rates, including Duschisnky rotation, vibronic coupling effects, use of different coordinate systems and so on. Right now, its use is optimized for CIS/TDDFT, as explained in Section *ISC, TD-DFT and the HT effect*, but it can be applied in general by combining simpler methods to obtain the geometries and Hessians with more advanced methods to compute the SOC matrix elements, when needed.

Tips and Tricks

- The DELE must be given when using ESD(ISC), it is not automatically computed.That is the energy of the initial state minus the energy of the final state, each at its own geometry.
- A SOC matrix element calculated from any method can be given on the input using the SOCME Re, Im flag, where these are the real and imaginary parts of that number.
- The SOCMEs from TD-DFT are not bad, maybe except for those between the ground singlet and the triplets. In that case, a multireference calculation might be the preferred option.
- If the final state is higher in energy than then initial state, the DELE is a negative number. In that case, there is barrier to go up when doing the ISC and the rates becomes more sensitive to the temperature.
- In contrast to Fluorescence, the ISC rates depend strongly on the inclusion of Duschisnky rotations, please take care when using USEJ FALSE.
- The default LINES is GAUSS, and the default INLINEW of 250 cm^{-1} might be too large. One should always investigate it by varying the width a bit. Other LINES can increase the error, please take care when changing it.

7.45.3 Resonant Raman Spectrum

General Aspects of the Theory

Raman intensities can be obtained in many different ways, depending on the experimental set up. As discussed extensively by D. A. Long [197, 533], the part of it that is "set up independent" is the Scattering Factor (or Raman activity), given by:

$$S = 45a^2 + 7\gamma \quad \left(\frac{C^2m^2}{amuV^2}\right), \tag{7.286}$$

where the a is related to the isotropic part of the "transition polarizability" between an initial state and a final state with a different vibrational quantum number $\langle\Psi_f|\hat{\alpha}|\Psi_i\rangle = \alpha_{if}$:

$$a = \frac{1}{3}|(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})|$$

and γ is also related to its off-diagonal elements:

$$\gamma = \frac{1}{2}[|(\alpha_{xx} - \alpha_{yy})|^2 + |(\alpha_{yy} - \alpha_{zz})|^2 + |(\alpha_{zz} - \alpha_{xx})|^2 + 6(|\alpha_{xy}|^2 + |\alpha_{yz}|^2 + |\alpha_{xz}|^2)].$$

This transition polarizability itself can be computed using Kramers, Heisenberg and Dirac (KHD) formalism and it can be shown that each of its Cartesian components can be calculated as a sum-over-states:

$$\alpha^{if}_{\rho\sigma} = \frac{1}{\hbar}\sum_{n\neq i,f}\left(\frac{\langle\Psi_f|\mu_\rho|\Psi_n\rangle\langle\Psi_n|\mu_\sigma|\Psi_i\rangle}{\Delta E_{ni} - \omega_{laser} + i\gamma_{lt}} + \frac{\langle\Psi_f|\mu_\rho|\Psi_n\rangle\langle\Psi_n|\mu_\sigma|\Psi_i\rangle}{\Delta E_{ni} + \omega_{laser} + i\gamma_{lt}}\right) \tag{7.287}$$

In (7.287), the sum is over any number of electronic excited states n, ΔE_{in} is the energy difference between the initial state and the excited, ω_{laser} is the laser energy and γ_{lt} is the lifetime broadening to avoid a zero on the denominator. If we work with a laser for which the frequency is close to the excited state energy difference, the first term is much larger than the second and can approximate alpha by

$$\alpha^{if}_{\rho\sigma} \simeq \frac{1}{\hbar}\sum_{n\neq i,f}\left(\frac{\langle\Psi_f|\mu_\rho|\Psi_n\rangle\langle\Psi_n|\mu_\sigma|\Psi_i\rangle}{\Delta E_{ni} - \omega_{laser} + i\gamma_{lt}}\right). \tag{7.288}$$

This equation can be solved using a path integral approach by switching to the integral form of $1/x$:

$$\frac{1}{x} = \frac{i}{\hbar}\int_0^\infty e^{-ixt/\hbar}dt \tag{7.289}$$

So that the components of α_{if} can be given by:

$$\alpha^{if}_{\rho\sigma} \simeq \sum_{n\neq i,f}\frac{i}{\hbar^2}\int_0^\infty \langle\Psi_f|\mu_\rho|\Psi_n\rangle\langle\Psi_n|\mu_\sigma|\Psi_i\rangle e^{-it(\Delta E_{in} - \omega_{laser} - i\gamma_{lt})}dt \tag{7.290}$$

From here on, it is possible to show that the $\alpha^{if}_{\rho\sigma}$ can be calculated as an integral of a correlation function [197], which is similar to the one previously discussed. In order to calculate that, a path integral scheme is also used and a geometry and Hessian for the ES are needed. The `ORCA_ESD` module predicts the ES PES (if not inputed), computes the α_{if} and then prints the Scattering factor on a spectrum named BASENAME.spectrum.LASERE.

OBS.: The actual Raman Intensity collected with any polarization at 90 degrees, the I($\pi/2; \|^s + \perp^s, \perp^i$ [533]), can be obtained by setting RRINTES to TRUE under %ESD.

OBS2.: In the current implementation, if a multistate calculation is requested, Eq. (7.288) is solved for each state and all spectra are summed up afterwards.

Specific Keywords and Details

In order to solve Eq. (7.290), the same information as for Absorption/Emission is needed and to compute the ES PES all of the above approximations are also valid. The main difference here is that a laser energy, given by the LASERE flag, should be given. If it is not given, the default is to set it to the 0-0 energy difference between the ground and the excited state. As mentioned before, more information can be found on Sec. *Resonant Raman Spectrum*.

- You can choose more than one laser energy by setting LASERE 1,2,3,4 and etc. If so, each spectrum will be saved on a different BASENAME.spectrum.LASERE file.
- You can also choose more than one excited state to be accounted for with the flag STATES 1,2,3, etc. and the final spectrum will be the sum of all of Scattering Factors given by the α_{if}s. You can NOT choose several states and laser energies currently.
- The automatic selection for the integral grid is also done based on the same ideas as mentioned before.
- The default lineshape for resonant Raman is VOIGT.
- The lineshape of the RR spectra will be taken from the RRSLINEW flag. In this case, LINEW and INLINEW are used only in the calculation of the correlation function.
- Currently the only temperature for which this model works is at 0K.
- In terms of which vibrationally excited states to be considered, currently it goes up to Raman Order 2, which means fundamentals, first overtones and combination bands (up to a total quantum number of 2). You can reduced that using RORDER flag.
- It is also possible to include HT effect here for weak transitions, but be aware the calculation is much more costly. Due to technical reasons, the data is saved only on memory so, if you plan to go being 300 modes and do HT, there should be A LOT of memory available, about $8 \times NMODES^4$. Also, you should expect a VERY long time for the computation of the correlation function. We are currently working on ways to accelerate this particular case.
- As it is explained on the reference paper [197], the calculations using both Duschinsky rotation and HT effect can be greatly accelerated setting cutoffs for the derivatives and J matrix elements. The RRTCUTDER is a ratio with respect to the transition dipole moment below which the derivatives will be ignored and RRTCUTJ is a cutoff for J matrix elements. As the paper suggests, RRTCUTDER = 0.001 and RRTCUTJ = 0.01 are in general good numbers. We do recommend using these, but please be aware of the specific needs of your system.

7.45.4 Circular Polarized Spectroscopies

General Aspects of the Theory

Starting from ORCA 6, the ESD module has been expanded to include calculations of CP rates and spectra for chemical system. This enables the computation of the vibronic effect on ECD, CPF, and CPP spectra. The methodology is generally similar to the interaction of unpolarized light with matter, including absorption and photoluminescence. However, when a CP photon interacts with an optically active system, the electric field of the photon induces a linear displacement of the charge (transition electric dipole), while the magnetic field induces a circulation of the charge (transition magnetic dipole). These combined interactions cause an electron to be excited in a helical motion, involving both translation and rotation, along with their associated operators.

In a commonly use practice by defining a laboratory frame in which the z-axis defines the direction of the light trajectory, CP light interactions can be generated with the use of the complex vectors $\mathcal{E}\pm = \frac{1}{\sqrt{2}}(\hat{x} \pm i\hat{y})$

In this framework the FFMIO operator transforms as:

$$\mathrm{T}_{\mathrm{IF}}^{\pm} = 1/2 \sum_{j=1}^{N} \left\langle \mathrm{I} \middle| \mathrm{e}^{-\mathrm{ikr_j}} \, (\varepsilon \bullet \hat{\mathrm{p}}_{\mathrm{x}}) \middle| \mathrm{F} \right\rangle \pm \left\langle \mathrm{I} \middle| \mathrm{e}^{-\mathrm{ikr_j}} \, (\varepsilon \bullet \hat{\mathrm{p}}_{\mathrm{y}}) \middle| \mathrm{F} \right\rangle \tag{7.291}$$

In both ECD and CPL (CPF or CPP) spectroscopies the measured intensities are related to the difference of absorption or luminescence of the left and right polarized transition moments given by:

$$\Delta_{IF}^{L\pm R}(k,\epsilon) = \left|T_{IF}^{-}\right|^2 \pm \left|T_{IF}^{+}\right|^2 \tag{7.292}$$

which leads to the following expressions for the sum and the difference of the square moduli $\left|T_{IF}^{\pm}\right|^2$:

$$\Delta_{IF}^{L+R}(k,\epsilon) = 1/2 \left\langle I \left| \sum_{j=1}^{N} e^{-ikr_j} \left(\varepsilon \bullet \hat{p}_x\right) \right| F \right\rangle \left\langle I \left| \sum_{j=1}^{N} e^{-ikr_j} \left(\varepsilon \bullet \hat{p}_y\right) \right| F \right\rangle \tag{7.293}$$

$$\Delta_{IF}^{L-R}(k,\epsilon) = -\mathbf{Im}\left(\left\langle I \left| \sum_{j=1}^{N} e^{-ikr_j} \left(\varepsilon \bullet \hat{p}_x\right) \right| F \right\rangle \left\langle I \left| \sum_{j=1}^{N} e^{-ikr_j} \left(\varepsilon \bullet \hat{p}_y\right) \right| F \right\rangle \right) \tag{7.294}$$

Hence within the ED approximation, ECD and CPL radiative transition rates can be calculated through the orientational average of Equation (4), employing the Fermi's Golden rule:

$$k_{ECD}\left(\omega\right) = \frac{16\pi^2\omega_{ECD}}{3} \sum_{F} \mathbf{Im}\left(\left|\left\langle \Psi_I \left|\hat{\mu}\right| \Psi_F \right\rangle \left\langle \Psi_F \left|\hat{m}\right| \Psi_I \right\rangle\right|\right) \delta\left(E_{FI} \pm \omega_{ECD}\right) \tag{7.295}$$

$$k_{CPL}\left(\omega\right) = \frac{16{\omega_{CPL}}^3 n^2}{3\hbar c^3} \sum_{F} \mathbf{Im}\left(\left|\left\langle \Psi_I \left|\hat{\mu}\right| \Psi_F \right\rangle \left\langle \Psi_F \left|\hat{m}\right| \Psi_I \right\rangle\right|\right) \delta\left(E_{FI} \pm \omega_{CPL}\right) \tag{7.296}$$

where $\omega_{ECD/CPL}$ are the excitation and emission CP photon energies, respectively while ω_{FI} are the energies between the initial and final states reached in the absorption or the photoluminescent processes. Similarly E_{FI} is the transition energy and δ refers to the line-broadening mechanism arising from the lifetimes of the relevant final states and c is the speed of light. In the above expressions $\hat{\mu}$ defines electric dipole operator while $\hat{m}$ is the respective magnetic dipole operator $\hat{m} = \frac{1}{2m_e c}\sum_i r_i \times \widehat{p_i}$ and m_e is the electron mass. In the above expressions $\mathbf{Im}\left(\left|\langle\Psi_I|\hat{\mu}|\ \Psi_F\rangle\ \langle\Psi_F|\hat{m}|\ \Psi_I\rangle\right|\right)$ represents the rotatory strength (R_{IF}). As discussed above the transition rates including the vibronic coupling, on the Frank-Condon (FC) and Herzenberg-Teller (HT) limits, can be efficiently proceed through the path integral approach[199] in which it is possible to calculate $k_{ECD/CPL}^{obs}$ from the Fourier Transform (FT) of the respective correlation function $\chi_{(t)}$ that is computed from the path integral of the multidimensional harmonic oscillator according to[784]:

$$k_{ECD/CPL}^{obs}(\omega) = 2\alpha\mathcal{R}e\int_0^\infty \chi_{(t)} e^{\pm i\omega t} dt \tag{7.297}$$

with α being a collection of constants and for CP transition one-photon rates (ECD, CPL) considering electric dipole and magnetic dipole interactions in the expression of the rotatory strengths it takes the form:[784]

$$\chi_{(t)} = e^{\pm i\omega t}[\mathbf{Im}\left[\mu_e m_e^*\right]\rho^{FC}\left(t\right) + \sum_k \mathbf{Im}\left[\mu_e \frac{\partial m_e^*}{\partial \bar{Q}_k}\right]\rho_k^{\frac{HT}{FC}}\left(t\right) + \sum_k \mathbf{Im}\left[\frac{\partial \mu_e}{\partial \bar{Q}_k} m_e^*\right]\rho_k^{\frac{HT}{FC}}\left(t\right) + \sum_{kl} \mathbf{Im}\left[\frac{\partial \mu_e}{\partial \bar{Q}_k}\frac{\partial m_e^*}{\partial \bar{Q}_l}\right]\rho_k^{HT}\left(t\right)] \tag{7.298}$$

where μ_e and m_e represent the respective transition dipole $\langle\Psi_I|\hat{\mu}|\ \Psi_F\rangle$ and magnetic dipole $\langle\Psi_F|\hat{m}|\ \Psi_I\rangle$ moment integrals between initial and final states I, F while:

$$\rho^{FC} = Tr(e^{-i\hat{H}\tau}e^{-i\hat{H}\tau}) \tag{7.299}$$

$$\rho_k^{HT/FC} = Tr(\bar{Q}_k e^{-i\hat{H}\tau}e^{-i\hat{H}\tau}) \tag{7.300}$$

$$\rho_{kl}^{HT} = Tr(\bar{Q}_k e^{-i\hat{H}\tau}\bar{Q}_l e^{-i\hat{H}\tau}) \tag{7.301}$$

where, these traces can be evaluated following the approach discussed in Ref[199].

Finally, it is quite commonly that the ECD and CPL spectral intensities are represented against normalized absorption and photoluminescent intensities defining, similar expressions for, the dissymmetry factors g_{abs} and g_{lum}:

$$g_{abs/lum} = 2\frac{I_{LCP} - I_{RCP}}{I_{LCP} + I_{RCP}} \quad \left.\frac{4R}{D}\right|_{GS/ES\ Structure}; \quad -2 < g_{abs/lum} < 2 \tag{7.302}$$

Implying that the associated dissymmetry spectra can also be calculated, where D and R the square of the transition dipole and the rotatory strength, respectively.

7.45.5 Magnetic Circular Dichroism

General Aspects of the Theory

The formulation presented for the calculation of the absorption spectrum may be extended to the absorption of circularly polarized light (CPL) of a system under the effect of an external magnetic field in order to compute the MCD spectrum.[268] By assuming an electric dipolar approximation under a length formulation for the light-matter interaction, the molar absorptivity contribution by the transition between an initial state i and a final state f of one fixed oriented molecule may be expressed as:

$$\epsilon(\omega)_{if} = \alpha\omega|\mathcal{E}\cdot\langle\Psi_i|\hat{\mu}|\Psi_f\rangle|^2\delta(E_i - E_f \pm \hbar\omega) \tag{7.303}$$

where $\mathcal{E}$ is the polarization vector of the incident light, and α is a positive collection of constants.

By considering the case in which the light is propagating in the laboratory fixed $\hat{z}''$ direction, the circularly polarized light is described by $\mathcal{E} = \frac{1}{\sqrt{2}}(\hat{x}'' \pm i\hat{y}'')$ where the "$-$" sign corresponds to the left circularly polarized light and the "$+$" to the right circularly polarized light.

Similarly, as the MCD calculations presented in section *Simulation of (Magnetic) Circular Dichroism and Absorption Spectra*, the total absorptivity may be obtained by summing over all possible transitions and averaging the molecular orientations by using 3 rotation angles θ, ϕ, and χ.

$$\epsilon(\omega) = \alpha\omega\int_0^{2\pi}\int_0^{2\pi}\int_0^{\pi}\sum_{if}\frac{e^{-\frac{\epsilon_i}{k_BT}}}{Z}|\mathcal{E}\cdot\langle\Psi_i|\hat{\mu}|\Psi_f\rangle|^2 sin(\theta)\delta(E_i - E_f \pm \hbar\omega)d\theta d\phi d\chi \tag{7.304}$$

Considering χ the angle which rotates the molecule in the plane perpendicular to the magnetic field perturbation (which is colinear with the incident light propagation direction), the integrals of eq. (7.304) may be computed conveniently in a seminumerical scheme by using an intermediate r' frame.

$$\epsilon(\omega) = \alpha\omega\sum_p w_p\sum_{if}\frac{e^{-\frac{\epsilon_i}{k_BT}}}{Z}\sum_{\beta\gamma=x',y'}\overline{\overline{\mathcal{E}}}_{\beta\gamma}\langle\Psi_i|\hat{\mu}_\beta|\Psi_f\rangle\langle\Psi_f|\hat{\mu}_\gamma|\Psi_i\rangle\delta(E_i - E_f \pm \hbar\omega) \tag{7.305}$$

where θ and ϕ values are defined by a point p in a numerical grid, x', and y' are determined by θ and ϕ values according to eq. (7.306), and χ has been integrated analytically in the $\overline{\overline{\mathcal{E}}}_{\beta\gamma}$ values.

$$\begin{aligned}\hat{\mu}_{x'} &= cos(\theta)cos(\phi)\mu_{x''} + cos(\theta)sin(\phi)\mu_{y''} - sin(\theta)\mu_{z''}\\ \hat{\mu}_{y'} &= -sin(\phi)\mu_{x''} + cos(\phi)\mu_{y''}\end{aligned} \tag{7.306}$$

$$\overline{\overline{\mathcal{E}}}^{\pm}_{\beta\gamma} = \int_0^{2\pi}\mathcal{E}^{\pm}_{\beta}\mathcal{E}^{\pm *}_{\gamma}d\chi \tag{7.307}$$

Finally, by applying the Bohr-Oppenheimer approximation and writing the Dirac delta function in the time domain, we get:

$$\epsilon(\omega) = \alpha\omega\sum_p w_p\sum_{if}\frac{e^{-\frac{\epsilon_i}{k_BT}}}{Z}\sum_{\beta\gamma=x',y'}\overline{\overline{\mathcal{E}}}_{\beta\gamma}\langle\Theta_i|\hat{\mu}^e_\beta|\Theta_f\rangle\langle\Theta_f|\hat{\mu}^e_\gamma|\Theta_i\rangle\int e^{i(E_i-E_f-\omega)t}dt \tag{7.308}$$

and by taking the difference of absorbance between left and right CPL produce we reach an expression of the MCD intensities:

$$\Delta\epsilon(\omega) = \epsilon^-(\omega) - \epsilon^+(\omega) = -2\frac{\alpha}{Z}\omega\sum_{if}\sum_p w_p Im\Big[\int_{-\infty}^{\infty}\tilde{\chi}(t,x',y')e^{-i\hbar\omega t}dt\Big] \tag{7.309}$$

where $\tilde{\chi}$ under a first-order approximation of the transitions moments with respect to the nuclear displacement is:

$$\tilde{\chi}(t,\beta,\gamma) = e^{i\Delta Et}\Big[\hat{\mu}_{0\beta}\hat{\mu}^*_{0\gamma}\rho^{FC}(t) + \sum_k\frac{\partial\hat{\mu}_\beta}{\partial Q_k}\hat{\mu}^*_{0\gamma}\rho^{FC/HT}(t) + \sum_k\hat{\mu}_{0\beta}\frac{\partial\hat{\mu}^*_\gamma}{\partial Q_k}\rho^{FC/HT}(t) + \sum_{kl}\frac{\partial\hat{\mu}_\beta}{\partial Q_k}\frac{\partial\hat{\mu}^*_\gamma}{\partial Q_l}\rho^{HT}(t)\Big] \tag{7.310}$$

Similarly, as section *Simulation of (Magnetic) Circular Dichroism and Absorption Spectra*, the transition moments under the effect of an external magnetic field perturbation may be estimated by using a QDPT (eq. (7.274)), and the derivatives are approximated numerically in a similar way as the ESD-Absorption case.

Specific Keywords and recommendations

Once selected the ESD(MCD) calculation two variables need to be defined in the %esd block:

- The intensity of the magnetic field in Gauss "B"
- The grid to make the molecular orientational average "LEBEDEVINTEGRATIONPOINTS"

The MCD signals use to be much more sensitive than the corresponding absorption intensities. As result, the MCD calculations are much more sensitive to errors in the electronic structure. So it is highly recommended to first fully understand the electronic structure problem and verify there are no important problems. In this direction we suggest the following recommendations:

- Be sure the obtained electronic states are in the correct order by assigning point group symmetry labels and comparing them with better electronic structure methods. Due to the MCD intensities emerging as a result of the interaction of states by the magnetic field perturbation, a wrong-located state in energy may affect the MCD intensities, even if it is not in the energy range you are computing. We do not recommend using the method as a black box.
- It is recommended to do first an ESD-absorption calculation on the exact same level of theory, verify the intensities and also solve any problem related to the PES representation.
- The lineshape for the best agreement of the MCD intensities compared against the experimental measurements may differ for the best value for absorption intensities.

7.45.6 Complete Keyword List for the ESD Module

```
%ESD                            #The booleans are the defaults
  ESDFLAG    ABS   (default)     #Which calculation to make?
             ECD
             FLUOR (default)
             CPF/CPFLUOR
             PHOSP
             CPP/CPPHOSP
             MCD                #Only available with TDDFT
             RR
             ISC
             IC

  GSHESSIAN  "BASENAME.hess"     #The ground state Hessian
  ESHESSIAN  "BASENAME_S1.hess"  #The excited state Hessian
  TSHESSIAN  "BASENAME_T.hess"   #The triplet state Hessian

  HESSFLAG   AHAS                #How to obtained the ES PES?
             VH
             VG (default)
             VGFC                # VG + APPROXADEN TRUE
             HHBS
             HHAS
             UFBS
             UFAS

  STATES     1,2,3,4             #IROOTS to compute

  MODELIST   4,5,6               # only include the 4th, 5th and 6th vibrational modes
                                 # in the ESD calculation
  SINGLEMODE 1,5,10              # run three ESD calculations, each considering only
                                 # one of the 1st, 5th and 10th vibrational modes

  DOHT       FALSE               #Do HT effect?
  FASTDER    TRUE                #Use the fast derivatives algorithm?
  DELQ       0.01 (default)      #Dimensionless displacemente for derivatives
```

(continues on next page)

(continued from previous page)

```
STDA       FALSE               #Use sTDA during derivatives?

APPROXADEN FALSE               #Compute the DELE by extrapolating info from
                               #the Hessians, avoiding second single-point
                               #at the ES geometry

USEJ       FALSE               #Consider Duschinsky rotations?
USEB       TRUE                #Rotate the initial state?
SAMEFREQ   TRUE                #Use DO method and J=1.

DELE       12000               #Custom energy difference
TDIP       x,y,z               #Custom transition electric dipole
           x.re,x.im,y.re,y.im,z.re,z.im
TMDIP      x,y,z               #Custom transition magnetic dipole
           x.re,x.im,y.re,y.im,z.re,z.im
SOCME      x.re,x.im           #Custom spin-orbit coupling matrix elements, in a.u.
LASERE     34000               #The laser energy for RR

B          3000.0              #Magnetic field in Gauss for MCD
LEBEDEVINTEGRATIONPOINTS 14    #Lebedev Grid for MCD molecular
                               #orientational average

GEOMSTEP   AUGHESS (default)   #Geometry step?
           QN
STEPSCALING  1.0               #A number for scaling the steps
STEPCONSTR   0,2,5             #A list of atoms that will not be moved
COORDSYS   CART                #Coordinate system for the normal modes?
           INT (default)
           WINT
           FCWL
           FCWS

TCUTFREQ   100                 #Cutoff for frequencies
IFREQFLAG  POSITIVE (default)  #What to do with negative frequencies?
           LEAVE
           REMOVE

UF_DELE    1E-4                #Energy difference for updated freq.
UFFREQERR  0.2                 #Tolerated percentual error

TEMP       298.15 (default)    #Temperature to consider
UNIT       WN                  #wavenumbers  (Output units - input still in cm-1!)
           NM                  #nanometers
           EV                  #electron volts
CENTRALDIFF  TRUE              #Central differences?
CONVDER    FALSE               #Convert derivatives between state
SCALING    1.0 (default)       #Scaling for frequencies
TDIPSCALING  1.0 (default)     #Scaling for the transition dipole

LINES      DELTA               #The lineshape function
           LORENTZ (default)
           GAUSS
           VOIGT (default for RR)
LINEW      50 (default)
INLINEW    250 (default)
RRSLINEW   10 (default)

RORDER     2 (default)         #The Raman Order for RR
RRINTENS   false               #Calculate the intensities instead
RRTCUTDER  0 (default)         #A cutoff for derivatives
RRTCUTJ    0 (default)         #A cutoff for J matrix elements
```

(continues on next page)

(continued from previous page)

```
  WRITEHESS  FALSE               #Make ES PES and leave

  MAXTIME    12000               #Max time (atomic units!) for the FT
  NPOINTS    131072              #Total number of points

  SPECRANGE  10000,50000         #Spectral range
  SPECRES    1.0                 #Spectral resolution

  PRINTVIB   FALSE               #Save a .xyz file per each vibational mode.
                                 #(Requires DOHT true)
  PRINTLEVEL 0                   #If set to 1 (or 2, 3, ...), prints additional
                                 #details of the calculation
END
```

7.46 More details on the ORCA DOCKER

Note

The content here will assume you already have some basic knowledge from *The ORCA DOCKER: An Automated Docking Algorithm*, if not, please check that section first!

7.46.1 Underlying theory

The basic idea behind the DOCKER is to use a type of swarm intelligence [786] to find local minima for where the HOST would best be positioned, based on the energies and gradients of a given Potential Energy Surface (PES).

Swarm intelligence algorithms were invented trying to mimic the behavior of animals such as bees and ants, and it actually fits the problem here quite well (as shown below). The basic steps needed for the algorithm are:

1. Optimize HOST and GUEST.
2. Create a spacial grid around the HOST.
3. Initialize a possible set of random solutions.
4. Let the swarm intelligence find local minima.
5. Collect a fraction of the best solutions found on 4. and fully optimize them.
6. The structure with the lowest energy is considered the best solution.

To quickly demonstrate how the initial set of solutions "evolve" to find different local minima, take as an example the substituted biphenyl as HOST below. The figure below demonstrates how a water molecule (GUEST) is docked onto the initial HOST.

Each gray ball represents one possible placing of the water molecule (ignoring here its rotation angle). At the start, they are all distributed over the HOST without any bias. As the iterations go by, the algorithm starts to detect that both sides of the molecule have lower energy solutions than the aromatic rings, as expected, and the possible solutions start to converge to those points.

In the end, solutions containing water molecules binding to both sides are taken, the system is then fully optimized and one finds that the best biding occurs with the hydroxy group to the right.

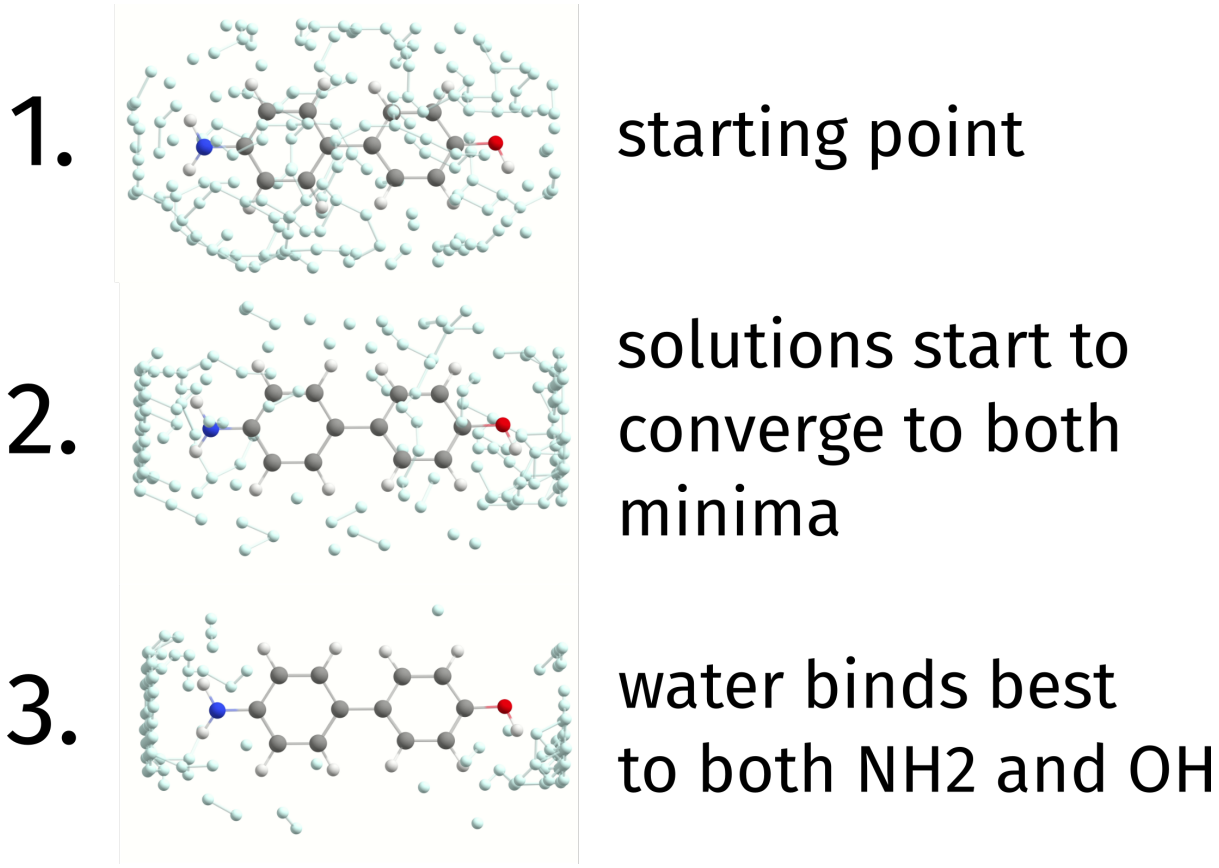


Fig. 7.60: The docking process of a water molecule on the substituted biphenyl.

Important

Right now the DOCKER can only be used together with the fast GFN-XTB methods, it will be generalized later. It is also not fully exploring the parallelization potential yet, and can be much faster for the next versions.

7.46.2 Looking Deeper into the Output

The details related to the grid and the swarm optimization can be found on the output, here is one example taken from the water dimer example from the earlier section:

```
Creating spatial grid
   Grid Max Dimension               5.50 Angs
   Angular Grid Step                32.73 degrees
   Cartesian Grid Step              0.50 Angs
   Points per Dimension             11 points
Initializing workers
   Population Density               0.50 worker/Ang^2
   Population Size                  57
Evolving structures
   Minimization Algorithm           mutant particle swarm
   Min, Max Iterations              (3 , 10)
```

The `Population Density` here defines how many solutions will be placed on the initial grid around the `HOST` and can be controlled by `%DOCKER POPDENSITY 0.50 END`, while the population size is a direct consequence of that, unless specified by `POPSIZE` under `%DOCKER`.

Another important piece of information is the `Min, Max Iterations`, which define the minimum number of

iterations before checking for convergence, and the maximum possible number of iterations. These can be changed by setting MINITER and MAXITER under the %DOCKER.

Then during the swarm search itself (here called the "evolution step"), the HOST and GUEST are partially optimized using the same criteria as from !SLOPPYOPT, and their energy at convergence is taken as a convergence criteria. The printing, still for the previous example, reads:

> **Note**
>
> All of the optimizations respect constraints and/or wall potentials given on the %GEOM block. The HOST can be easily frozen with %DOCKER FIXHOST TRUE END.

```
   Iter   Emin          avDE         stdDE          Time
          (Eh)       (kcal/mol)    (kcal/mol)      (min)
   ----------------------------------------------------

      1 -10.147462     2.756033     1.821981         0.03
      2 -10.147462     2.121389     1.610208         0.03
      3 -10.148583     2.313606     1.365227         0.03
      4 -10.148583     1.846998     1.188680         0.02
      5 -10.148583     1.587332     1.168207         0.02
No new minimum found after 3 (MinIter) steps.
```

Here we have the number of iterations, the energy of the best solution found during the process, the average energy difference from the lowest solutions and its standard deviation. The first number shows if the system is evolving towards a lower energy solutions and the later two give an idea of the "spread" of the energies found.

If no new minimum is found over MINITER iterations, the algorithm is considered to be converged and the best solutions are taken to the final optimization step.

7.46.3 The final steps

After the evolution step, a total of $max(sqrt(PopSize), 5)$ structures are taken from the set of solutions for a final full optimization. The output looks like:

```
Running final optimization
   Maximum number of structures      7
   Minimum energy difference         0.10 kcal/mol
   Maximum RMSD                      0.25 Angs
   Optimization strategy             regular
   Coordinate system                 redundant 2022
   Fixed host                        false
```

To avoid redundant solutions being optimized here, a check is made such that only structures with an energy difference of at least a Minimum energy difference and Maximum RMSD (obtained after an optimal rotation using the quaternion approach) are taken. The coordinate system might also be automatically set to Cartesian if the number of atoms is above 300 to speedup.

If some optimizations fail during the process, they will be flagged as optimization failed, as in the example below:

```
   Struc   Eopt       Interaction Energy       Time
           (Eh)               (kcal/mol)      (min)
   -------------------------------------------------

    1      -10.149006          -4.968378       0.01
    2      -10.149005          -4.967965       0.01
    3       optimization failed                0.01
    4      -10.149007          -4.968641       0.01
   (...)
```

That only means some of the solutions failed to fully optimize at the end, and on the printed file their energy is set to 1000 Eh to show the job was not completed. If you want to increase the number of iterations, or change the final convergence criteria, just change them using the %GEOM block as usual.

7.46.4 General Tips

1. In order to quick change the PES of the docking process, you can use !DOCK(GFN2-XTB), !DOCK(GFN1-XTB), !DOCK(GFN0-XTB) or !DOCK(GFN-FF).
2. For a quicker and less accurate docking the input !QUICKDOCK is also available.
3. There are four levels of "docking" procedure, from simple to more elaborate: SCREENING, QUICK, NORMAL (default) and COMPLETE, which can be given as %DOCK DOCKLEVEL COMPLETE END.
4. To increase the final number of structures optimized in the end, just change NOPT under %DOCKER.
5. If the guest topology is somehow changed during the docking process, that structure will be discarded. This can be turned off by setting %DOCK CHECKGUESTTOPO FALSE END.

Important

By the time of the ORCA6 release, this algorithm was still not published. Publications are under preparation.

7.46.5 Complete Keyword List

```
!QUICKDOCK       # simple keyord to set DOCKLEVEL QUICK
!NORMALDOCK      # simple keyord to set DOCKLEVEL NORMAL
!COMPLETEDOCK    # simple keyord to set DOCKLEVEL COMPLETE

!DOCK(GFN-FF)    # simple keyord to set EVPES GFNFF
!DOCK(GFN0-XTB)  # simple keyord to set EVPES GFN0XTB
!DOCK(GFN1-XTB)  # simple keyord to set EVPES GFN1XTB
!DOCK(GFN2-XTB)  # simple keyord to set EVPES GFN2XTB

%DOCKER

   #
   # general options
   #

   GUEST           "filename.xyz" # an .xyz file (can be multistructure), from where
                                  # the guest(s) will be read. can contain different
                                  # charges and multiplicities for each guests on the
                                  # comment line. will only be read if exactly two
                                  # integer numbers are given, otherwise ignored.

   DOCKLEVEL       SCREENING  # defines a general strategy for docking.
                   QUICK      # will alter things like that population density
                   NORMAL     # and final number of optimized structrures.
                   COMPLETE   # default is NORMAL.

   PRINTLEVEL      LOW
                   NORMAL     # default
                   HIGH       # will print many extra files!

   NREPEATGUEST    1          # number of times to repeat the content of the "GUEST" file
   CUMULATIVE      TRUE       # add the contents of the "GUEST" file one on
                              # top of each other?
                              # default is FALSE, meaning each will be done independently.
```

(continues on next page)

(continued from previous page)

```
GUESTCHARGE     0          # can be used to defined a charge for the guest (default 0)
GUESTMULT       3          # same for multiplicty (default 1)
                           # both can also be given via the comment line of the
                           # "filename.xyz", see above.

RAMDOMSEED      TRUE       # whether to allow for the process to be trully random and
                           # will give different results everytime (default FALSE)

FIXHOST         TRUE       # freeze coordinatef for the HOST during all steps?
                           # (default FALSE)

CHECKGUESTTOPO  FALSE      # make sure that the topolgy of the GUEST is always kept
                           # (default TRUE)

#
# evolution step
#

MAXITER         10         # maximum number of iterations
MINITER         3          # minimum number to check for convergence
                           # and minimum required of equal steps
                           # to signal convergence
POPDENSITY      0.5        # the population density based on the HOST grid
POPSIZE         150        # a fixed number for the swarm population size

EVOPTLEVEL      NORMALOPT  # use default optimization criteria during the evolution step
                           # instead of the !SLOPPYOPT? more accurate, but slower.

EVPES           GFNFF      # which PES to use **only** during the evolution step.
                GFN0XTB    # can be different from the final optimization.
                GFN1XTB
                GFN2XTB
#
# final optimization
#

NOPT            10         # a fixed number of structures to be optimized
NOOPT           FALSE      # do not optimize any structure at all? (default FALSE)
```

7.47 More on the ORCA SOLVATOR

Here we will present a few more technical details about the SOLVATOR that were too specific to be presented on the more general section. This section presumes that the section *ORCA SOLVATOR: Automatic Placement of Explicit Solvent Molecules* was already read. If not, please return here after that.

7.47.1 The Simpler Stochastic Mode

The `CLUSTERMODE STOCHASTIC`, as the name suggests, uses random trial and error to assign the placing of the solvents. Well, it is actually more complicated than that.

The algorithm actually uses information from self-consistent EEQ charges [132] as part of an extremely simplistic potential to guide the placing of polar molecules in a more reasonable way.

After trying to distribute the solvent molecule somewhere and checking for clashes, we first compute the electrostatic energy between solvent and solute:

$$E_{elec} = \sum_{A}^{solute} \sum_{B}^{solvent} \frac{Q_A Q_B}{R_{AB} 4\pi\epsilon_{solv}}, \tag{7.311}$$

and define our target function to be minimized (V) as a damped version of that, using the minimum distance R_{min} between the solute and solvent atoms:

$$\begin{aligned} V &= E_{elec} e^{-R_{min}^2} && \text{if } E_{elec} < 0 \\ &= R_{min} && \text{otherwise,} \end{aligned} \tag{7.312}$$

The `STOCHASTIC` mode then consists of finding the correct solvent placement that minimizes V for a given solute. The damping function is there to ensure that:

1. The electrostatic interaction decays strongly with distance,
2. Repulsive energies will be so unfavorable that only the distance will matter.

The result is such that solvent molecules are placed as close as possible to the solute and maximizing electrostatic interactions. This helps to create the solvent shell such that is does look like the actual result one would expect from a more elaborate calculation, but with essentially zero cost.

The best value for V found after each solvent was added is what is printed in the output as `Target function`:

```
     Iter     Target function   Time
                               (min)
     -------------------------------
        1     -4.342597e-07     0.00
        2     -3.166857e-07     0.00
        3     -4.814590e-08     0.00
```

When using the `DROPLET` mode, the R_{min} is defined as the distance to the centroid of the solute, instead of that of the closest atom pair instead.

If you don't want to include the electrostatic component for any reason, just set `%SOLVATOR USEEEQCHARGES FALSE END`. In the future other charge models will be available as well.

7.47.2 Adding Explicit Solvents with the Docker

The Wall Potential

If one uses the default approach using the DOCKER, a fictitious wall potential is added to guarantee that the solvents are added such that they fill most of the first solvation sphere around the solute before being placed further.

> **Note**
>
> This resembles to some extent what was published recently by the group of Prof. Stefan Grimme (called "quantum cluster growth"), but here only a single outer wall potential is used [812]. Otherwise, the present algorithm is independent and unrelated to it.

As one can see from the output of the Histine example of the before mentioned section, by default an ellipsoid potential is built:

```
Ellipsoid potential radii:                    .... 8.37, 6.28, 5.76 Angs
```

with dimensions such that it will enclose the solute plus at least one molecule of the solvent in all directions.

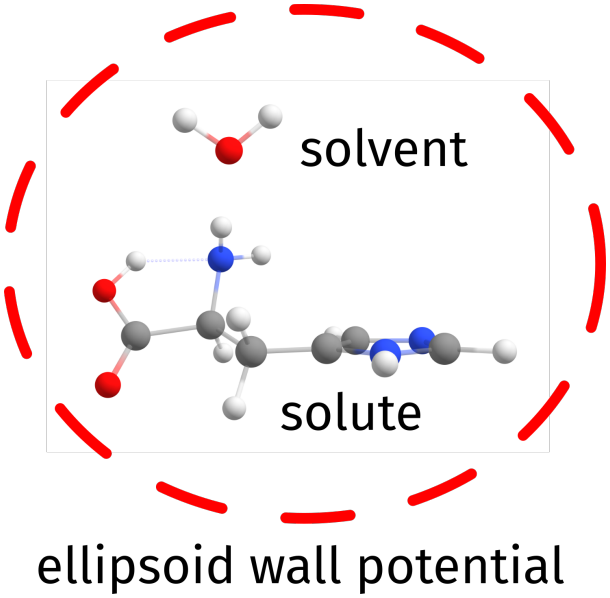


Fig. 7.61: A simple scheme to show how the wall potential is built to keep the solvent molecules close to the solute space.

A single parameter controlled by `SOLVWALLFAC` under the `%SOLVATOR` block defines how further this wall is built outside the solute. Its default value is 1.0, and increasing it to larger values will increase the default initial wall by about half the sum of the maximum dimensions of solute plus solvent for each unit.

The initial wall potential is by default not changed, unless a) the algorithm can not place a solvent, or b) the energy of the placed solvent is higher than before. Only then the walls will be updated to help allocating the next solvent molecule, and a message will be printed with the current scaling factor.

The DOCKER

All options related to the docking process can be controlled as usual via the `%DOCKER` block. Please check the section *The ORCA DOCKER: An Automated Docking Algorithm* for more info.

Also, all options given under the `%GEOM` block such as constraints and etc., will be respected during the docking of the solvent, as with the general docking, but these can only be given for the solute.

7.47.3 Controlling True Randomness

Both these algorithms are intrinsically dependent on random numbers, but ORCA sets a fixed random seed such that the same results are always obtained on the same machine if calculations are repeated.

In order to make both fully random, please set `%SOLVATOR RANDOMSOLVATION TRUE END`.

7.47.4 Vacuum Search

One option that might come in handy under certain conditions is to use the SOLVATOR to add the explicit solvent based on the implicit solvation information, but actually **not** use any implicit solvation while trying to place them.

That is useful for instance when trying to generate aggregates of solute plus solvents that can form in gas-phase only, where there will be no other solvent molecules around. That can be set with `%SOLVATOR VACUUMSEARCH TRUE END` and the implicit solvation method will be ignored to compute energies and gradients. In the `STOCHASTIC` method, the ϵ_{solv} will be set to 1.0.

Important

By the time of the ORCA6 release, this algorithm was still not published. Publications are under preparation.

7.47.5 Complete Keyword List

```
%SOLVATOR

   #
   # general options
   #

   NSOLV       10   # number of explicit solvent molecules to be added.

   SOLVENTFILE "solvent_file_name.xyz" # a file for custom solvents.
                                       # NOT needed for the regular
                                       # solvents available via
                                       # ALPB(solvent). charge and
                                       # multipl. can be give on the comment.

   CLUSTERMODE STOCHASTIC   # method for adding new solvent molecules.
               DOCKING      # default

   PRINTLEVEL  LOW
               NORMAL       # default
               HIGH

   RANDOMSOLV  TRUE         # make it completely random? default FALSE.

   FIXSOLUTE   FALSE        # keep the solute constrained? default TRUE.

   #
   # stochastic method
   #

   USEEEQCHARGES FALSE      # use EEQ charges during the STOCHASTIC mode?
                            # default is TRUE

   DROPLET       FALSE      # create a spherical droplet? default FALSE

   RADIUS        10         # a radius in Angstroem for the droplet.
                            # solvent molecules will be added until the
                            # target radius is reached.

   #
   # docking method
   #

   WALLFAC     1.0          # factor use to define the initial
```

(continues on next page)

(continued from previous page)

```
                          # size of the wall potential.

  # all other docking options are controlled
  # as usual via the %DOCKER block.
  # flags such as !NORMALDOCK and
  # !COMPLETEDOCK will apply here as well
```

7.48 *Ab initio* Molecular Dynamics Simulations

A few years ago, we included an *ab initio* molecular dynamics (AIMD) module into ORCA.[1] As a plethora of different electron structure methods with analytical gradients is already implemented, all these methods are now available also for MD simulations, offering a wide range of accuracy/performance trade-offs.

Despite the relatively short history inside of the ORCA package, the MD module has grown considerably over the last years. A few features found in other MD codes are still missing. In future releases, many more new features and methods will (hopefully) be added to this part of the program. We will always do our best to keep a strict backward compatibility, such that the sample inputs from this section will remain valid in all future releases.

For some more information as well as input examples for the ORCA MD module, please visit

https://brehm-research.de/orcamd

7.48.1 Changes in ORCA 5.0

- Added a `Metadynamics` module with many features and options:
  - Can perform one-dimensional and two-dimensional Metadynamics simulations [487] to explore free energy profiles along reaction coordinates, called collective variables ("Colvars").
  - `Colvars` can be distances *(including projections onto vectors and into planes)*, angles, dihedrals, and coordination numbers [408]. The latter allows, *e. g.*, to accurately compute pK_A values of weak acids [854, 855].
  - For all Colvars, groups of atoms *(e. g., centers of mass)* can be used instead of single atoms.
  - Metadynamics simulations can be easily restarted and split over multiple runs.
  - Ability to run well-tempered Metadynamics [73] for a smoothly converging free energy profile.
  - Ability to run extended Langrangian Metadynamics [408], where a virtual particle on the bias profile is coupled to the real system via a spring. The virtual particle can be thermostated.
- Added two modern and powerful `thermostats` *(both available as global and massive)*:
  - The widely used Nosé–Hoover chain thermostat (NHC) with high-order Yoshida integrator [562, 563]; allows for a very accurate sampling of the canonical ensemble.
  - The stochastical "Canonical Sampling through Velocity Rescaling" (CSVR) thermostat [129] which has become quite popular recently.
- Can define harmonic and Gaussian `restraints` for all Colvars *(distance, angle, dihedral, coordination number)*. This allows for umbrella sampling [430, 482], among other methods. Can also define one-sided restraints which act as lower or upper wall.
- Can now print the instantaneous and average force on `constraints` and `restraints`; this allows for thermodynamic integration [430].
- The target value for `constraints` and `restraints` can now be a ramp, so that it can linearly change during the simulation.

[1] Strictly speaking, these simulations are Born–Oppenheimer molecular dynamics simulations (BOMD), because they approximately solve the time-independent Schrödinger equation to compute gradients and then move the atoms according to these gradients.

- Can now keep the system's center of mass fixed during MD runs.
- Can now print population analyses, orbital energies, and .engrad files in every MD step if requested.

7.48.2 Changes in ORCA 4.2 (Aug 2019)

- Added a Cartesian minimization command to the MD module, based on L-BFGS and simulated annealing. Works for large systems ($> 10\,000$) atoms and also with constraints. Offers a flag to only optimize hydrogen atom positions *(for crystal structure refinement)*. See `Minimize` command.
- The MD module can now write trajectories in DCD file format *(in addition to the already implemented XYZ and PDB formats)*, see `Dump` command.
- The `thermostat` is now able to apply temperature ramps during simulation runs.
- Added more flexibility to region definition *(can now add/remove atoms to/from existing regions)*. Renamed the `Define_Region` command to `Manage_Region`.
- Added two new constraint types which keeps centers of mass fixed or keep complete groups of atoms rigid, see `Constraint` command.
- Ability to store the GBW file every n-th step during MD runs *(e.g. for plotting orbitals along the trajectory)*, see `Dump` command.
- Can now set limit for maximum displacement of any atom in a MD step, which can stabilize dynamics with poor initial structures. Runs can be cleanly aborted by "`touch EXIT`". See `Run` command.
- Better handling/reporting of non-converged SCF during MD runs.
- Fixed an issue which slowed down molecular dynamics after many steps.
- Stefan Grimme's xTB method can now be used in the MD module, allowing fast simulations of large systems.

7.48.3 Changes in ORCA 4.1 (Dec 2018)

- Molecular dynamics simulation can now use Cartesian, distance, angle, and dihedral angle constraints. These are managed with the `Constraint` command.
- The MD module now features cells of several geometries (cube, orthorhombic, parallelepiped, sphere, ellipsoid), which can help to keep the system inside of a well-defined volume. The cells have repulsive harmonic walls.
- The cells can be defined as elastic, such that their size adapts to the system. This enables to run simulations under constant pressure.
- Trajectories can now be written in XYZ and PDB file format.
- A restart file is written in each simulation step. With this file, simulations can be restarted to seamlessly continue (useful for batch runs or if the job crashed). Restart is handled via the `Restart` command; see below.
- Introduced regions (*i. e.*, subsets of atoms), which can be individually defined. Regions can be used to thermostat different parts of the system to different temperatures (*e. g.*, cold solute in hot solvent), or to write subset trajectories of selected atoms.
- The energy drift of the simulation is now displayed in every step (in units of Kelvin per atom). Large energy drift can be caused by poor SCF convergence, or by a time step length chosen too large.
- Physical units in the MD input are now connected to their numeric values via underscore, such as `350_pm`. A whitespace between value and unit is no longer acceptable. This slightly breaks backward compatibility – sorry.
- Fixed a bug in the time integration of the equations of motion, which compromised energy conservation.
- Fixed crashes for semiempirics and if ECPs were employed. You can now run MD simulations with methods such as PM3 and with ECPs.

7.48.4 Input Format

The molecular dynamics module is activated by specifying "MD" in the simple input line. The actual MD input which describes the simulation follows in the "%md" section at some later position in the input file. The contents of this section will subsequently be referred to as "MD input".

```
! MD BLYP D3 def2-SVP
%md
  Timestep 0.5_fs # This is a comment
  Initvel 350_K
  Thermostat NHC 350_K Timecon 10.0_fs
  Dump Position Stride 1 Filename "trajectory.xyz"
  Run 200
end
* xyz 0 1
O         -4.54021        0.78439         0.09307
H         -3.64059        0.38224        -0.01432
H         -4.63463        1.39665        -0.67880
*
```

Please note that the MD input is not processed by ORCA's main parser, but by a dedicated parser in the MD module. Therefore, the MD input is not required to obey the general ORCA syntax rules. The syntax will be described in the following.

In contrast to general ORCA input, the MD input is not based on keywords, but on *commands*, which are executed consecutively on a line-by-line basis starting at the top (like, *e. g.*, in a shell script). This means that identical commands with different arguments may be given, coming into effect when the interpreter reaches the corresponding line. This enables to perform multiple simulations (*e. g.*, pre-equilibration and production run) within a single input file:

```
%md
  Timestep 1.0_fs
  Run 200
  Timestep 2.0_fs
  Run 500
end
```

Work is already under way to add variable definitions, loops, and conditional branching to the MD input.[2] This will enable even larger flexibility (*e. g.*, to run a simulation until a certain quantity has converged). The MD input is written in the SANscript language ("Scientific Algorithm Notation Script"), which is under development. A first glimpse can be found at

https://brehm-research.de/sanscript

As in standard ORCA input, **comments** in the MD input are initiated by a "#" sign and span to the end of the current line. Commands can be started both at the beginning of a line and after a command. The only place where a "#" is **not** treated as start of a comment is inside of a string literal (*e. g.*, in file names).

```
%md
  # Comment
  Timestep 0.5_fs # Comment
  Dump Position Filename "trajectory#1.xyz"
end
```

Some more MD input syntax rules:

- The MD input is generally **not** case-sensitive. The only exception are file names on platforms with case-sensitive file systems (such as GNU Linux).
- Empty lines are allowed.

[2] Technically speaking, ORCA will then be a Turing-complete script interpreter, such that *any* computational problem can be solved with ORCA :-)

- Commands and options are separated by space or tabulator characters. Any combination of these characters may be used as separator.
- Both DOS and UNIX line break style is acceptable.

Commands

As already noted above, the central item of the MD input is a command. Each input line contains (at most) one command, and these commands are executed in the given order. A command typically takes one or more arguments, which are given behind the command name, separated by whitespaces, tabulator characters, or commas (optional). The order of the arguments for a command is fixed (see command list in section *Command List*. Commands may have optional arguments, which are always specified at the end of the argument list, after the last non-optional argument. If there exist multiple optional arguments for a command, not all of them need to be specified; however, they need to be specified in the correct order and without gaps:

```
%md
  Command Arg1 Arg2 Arg3                  # fine
  Command Arg1, Arg2, Arg3                # fine
  Command Arg1 Arg2 Arg3 Optarg1          # fine
  Command Arg1 Arg2 Arg3 Optarg1 Optarg2  # fine
  Command Arg1 Arg2 Arg3 Optarg2          # will not work
end
```

Apart from arguments and optional arguments, commands can also have *modifiers*. These can be considered as "sub-commands", which modify a given command, and may possess their own argument lists. Modifiers generally follow after all non-optional and optional arguments, and they may **not** possess optional arguments on their own. If a command has multiple modifiers, the order in which they are given is not important.

In the following input example, "Mod1" and "Mod2" are modifiers of "Command". "Mod1" takes one argument, "Mod2" does not take arguments:

```
%md
  Command Arg1                            # fine
  Command Arg1 Optarg1                    # fine
  Command Arg1 Mod1 Modarg1 Mod2          # fine
  Command Arg1 Mod2                       # fine
  Command Arg1 Mod2 Mod1 Modarg1          # fine
  Command Arg1 Optarg1 Mod1 Modarg1 Mod2  # fine
end
```

To make this abstract definition a little more illustrative, please consider again one line from the input sample at the beginning of this section:

```
%md
  Dump Position Stride 1 Filename "trajectory.xyz"
end
```

Here, "Dump" is the command, which takes one non-optional argument to specify which quantity shall be dumped – in this case, "Position". The "Dump" command has two modifiers, namely "Stride" and "Filename". The former takes one integer argument, the latter a string argument. Swapping the two modifiers (together with their respective arguments, of course) would not change the behavior.

Separating Arguments

As shown above, the arguments which are passed to a command do not need to be separated by commas. However, it is allowed (and recommended) to still use commas. First, it can increase the readability of the input file. Secondly, there exist a few ambiguous cases in which commas (or parentheses) should be used to clarify the intended meaning. One of these cases is the arithmetic minus operator. It can either be used as binary operator (subtracting one number from another), or as unary operator (returning the negative of a number). By default, the minus operator will be considered as binary operator (if possible).

Consider the case in which you want to pass two integer arguments "10" and "-10" to a command. Without commas (or parentheses), the minus is mistreated as binary operator, and only one argument will be passed to the command:

```
Command 10 -10    #  Pitfall: treated as "Command (10 - 10)", i.e., "Command 0"
Command 10, -10   #  Two arguments, as intended
Command 10 (-10)  #  Also works
```

Physical Units

In many cases, it is required to specify quantities which bear a physical unit in an input file (*e. g.*, temperature, time step lengths, ...). For many quantities, there are different units in widespread use, which always leads to some confusion (just consider the "kcal vs kJ" case). ORCA handles this problem by defining default units for each quantity and requiring that all quantities are given in their default unit. ORCA's default units are the atomic units, which are heavily used in the quantum chemistry community, but not so much in the molecular dynamics community. As an *ab initio* molecular dynamics module exists in the small overlap region of both communities, some "unit conflicts" might arise. To prevent those from the beginning, it is allowed to specify units of personal choice within ORCA's MD input.

Luckily, this is as simple and convenient as it sounds. The parser of the MD module checks if a unit is given after a numeric constant, and automatically converts the constant to the internal default unit. If no explicit unit is given, the default unit is assumed. Please note that the default units within the MD module are not necessarily atomic units (see table below). Units are connected to the preceding numerical value by an underscore:

```
%md
  Timestep  1.0_fs
  Timestep 41.3_au # identical
  Timestep  1.0    # identical, as default time unit in MD module is fs
end
```

In the following, all units which are currently known to the MD module's parser are listed, sorted by physical quantities. The default unit for each quantity is printed in **bold letters**. Additive constant and factor are applied to convert a unit into the default unit. The additive constant is applied before the factor. A "—" means that the constant/factor is not applied. More units will be probably added in the future.

Unit Symbol	Additive Constant	Factor
— Length Units —		
Angstrom	–	–
A	–	–
Bohr	–	0.5291
pm	–	0.01
nm	–	10.0
— Time Units —		
fs	–	–
ps	–	1000
au	–	0.02419
— Temperature Units —		
Kelvin	–	–
K	–	–
Celsius	273.15	–
C	273.15	–
— Angle Units —		
Deg	–	–
Rad	–	$180/\pi$

7.48.5 Discussion of Features

Restarting Simulations

Ab initio molecular dynamics simulation are computationally expensive, and will typically run for a long time even in the case of medium-sized systems. Often, it is desirable to perform such a simulation as a combination of multiple short runs (*e. g.*, if the queuing system of the cluster imposes a maximum job time). The ORCA MD module writes a restart file in each simulation step, which allows for the seamless continuation of simulations. This restart file has the name "`basename.mdrestart`", where `basename` is the project's base name. To load an existing restart file, use the *Restart* command *(see command list below)*.

In the first run of a planned sequence of runs, no restart file exists yet. for this case, the *Restart* command offers the `IfExists` modifier. The restart file is only loaded if it exists. If not, the restart is simply skipped, and no error is thrown. By using this modifier, you can have the *Restart* command already in place in the first run of a sequence (where no restart file exists in the beginning), and do not need to modify the input after the first run has finished.

Concerning the *Dump* command, it is good to know that trajectory files are appended (*not* overwritten) by default. If you ever want to overwrite an existing trajectory file by a *Dump* command, use the `Replace` modifier.

Please note that **only** the positions, velocities, thermostat internal state *(only for NHC)*, Metadynamics hills, and time step counters are restarted when executing a *Restart* command. All other properties (thermostats, regions, trajectory dumps, constraints, cells, etc.) are **not** restarted. They should all remain in the input file, as executed in the first run of a sequence. Just add the *Restart* command after all other relevant commands have been executed, directly before the *Run* command.

To conclude this discussion, a short example is given. If the MD input file

```
%md
  Timestep 0.5_fs
  Initvel 300_K
  Thermostat NHC 300_K Timecon 10.0_fs
  Dump Position Stride 1 Filename "trajectory.xyz"
```

(continues on next page)

(continued from previous page)

```
  Restart IfExists
  Run 100
end
```

is subsequently executed ten times (without any modification), the resulting trajectory file will be identical *(apart from numerical noise)* to that obtained if the following input is executed once:

```
%md
  Timestep 0.5_fs
  Initvel 300_K
  Thermostat NHC 300_K Timecon 10.0_fs
  Dump Position Stride 1 Filename "trajectory.xyz"
  Run 1000
end
```

Regions

In the ORCA MD module, **regions** can be defined. This concept does *not* refer to regions in space, but rather to subsets of atoms in the system. A region is nothing more than a list of atoms. Regions may overlap, *i. e.*, atoms can be part of more than one region at a time. The atoms which are part of a certain region remain the same until the region is manually re-defined, *i. e.*, regions are fixed and do not adapt to any changes in the system. There exist a few pre-defined regions which have a name. User-defined regions, in contrast, only carry an integer identifier. The following regions are pre-defined in any case:

- `all`: Contains all atoms of the system. This is the default if no region is specified in some commands, so by default, these commands will always act on the whole system.
- `active`: This region contains all movable ("non-frozen") atoms. By default, it is identical to the `all` region. Atoms inside of this region are updated by the time integration in a molecular dynamics run, displaced in a minimization, and are considered for computing the kinetic energy.
- `inactive`: This region contains all atoms which are *not* part of the `active` region. These atoms are "frozen"; they are ignored by the time integration / minimization, and also not considered for the computation of the kinetic energy. They simply remain on their initial positions. This is in principle identical to applying Cartesian constraints to the atoms; however, it is much faster. As constraints have to be solved iteratively *(see below)*, Cartesian constraints become quite computationally demanding if applied to thousands of atoms.

From these three pre-defined regions, only the `active` region can be manually modified. Changes in the composition of the `active` region automatically modify the `inactive` region. The `all` region obviously cannot be changed.

In case of a QM/MM simulation, the following four additional regions can be used:

- `qm`: This is the "quantum mechanics" region – it contains all atoms which are treated by the electron structure method.
- `mm`: This is the "molecular mechanics" region – it contains all atoms which are treated by a force-field approach. It exactly contains those atoms which are not part of the `qm` region.
- `active_qm`: Contains exactly those atoms which are part of both the `qm` and the `active` regions.
- `active_mm`: Contains exactly those atoms which are part of both the `mm` and the `active` regions.

These regions can **not** be modified in the MD input. The MD module just reads the region definitions from the QM/MM module, but is not able to make any changes here.

Regions can be useful for many purposes. For example, a "realistic" wall of atoms can be built around the system by defining the `active` region such that it only contains the non-wall atoms. The wall atoms will then be frozen. Apart from that, trajectories of regions can be written to disk, only containing the "interesting" part of a simulation. Furthermore, velocity initialization can be applied to regions, enabling to start a simulation in which different sets of atoms possess different initial temperatures. Thermostats can be attached to regions to keep different sets of

atoms at different temperatures during the whole simulation. This allows for sophisticated simulation setups (cold solute in hot solvent, temperature gradient through the system, etc).

Regions are defined or modified by the *Manage_Region* command. Many other commands take regions as optional arguments. Please see the command list below.

Metadynamics

Metadynamics is a powerful tool to analyze free energy profiles of reactions and other processes *(solvation, aggregation, conformer change, dissociation)* based on molecular dynamics simulations. It has been developed by Laio and Parrinello in 2002 [487]. In principle, the frequency of observing a certain process in MD simulations is directly related to the free energy barrier of the process. However, many interesting processes *(such as chemical reactions)* possess such a high free energy barrier that they will never occur on the time scales typical for AIMD simulations (100 ps). To increase the frequency at which such processes happen, so-called rare event sampling methods can be employed. Metadynamics is one among those. It works by building up a bias potential as a sum of Gaussian hills, so that free energy minima are slowly filled up and the system is gradually pushed away from its resting points.

Please note that there is also a method with the same name for exploring conformation space that has been published by Grimme in 2019 [330]. It is in principle based on the original "Parrinello" Metadynamics, but with several modifications and extensions. The ORCA MD module contains the original Parrinello variant of Metadynamics [487], together with several extensions such as well-tempered Metadynamics [73] and extended Lagrangian Metadynamics [408]. The Grimme method for conformer search will probably be implemented in the future.

In Metadynamics, one has to define one or more "collective variables" (Colvars) along which the free energy profile of the system will be sampled. A Colvar is in principle nothing more than a continuous function of all atom positions which returns a real number. A simple example of a Colvar is the distance between two atoms, which could be used to explore the free energy profile of a bond formation or cleavage. In the ORCA MD module, Colvars can be defined via the *Manage_Colvar* command. Available Colvar types are distances *(including projections onto lines or into planes)*, angles, dihedral angles, and coordination numbers [408]. The latter allows, *e. g.*, to accurately compute pK_A values of weak acids in solvent [854, 855]. For the distances, angles, and dihedral angles, atom groups instead of single atoms can be specified, so that, *e. g.*, the distance between the centers of mass of two molecules can be defined as a Colvar.

Based on one or two Colvars (ORCA supports one-dimensional and two-dimensional Metadynamics), a Metadynamics simulation can be set up. There are many parameters to choose, which are described in the section of the *Metadynamics* command. After all parameters have been set, the actual simulation is simply started via the *Run* command. It is also possible to restart Metadynamics simulations so that they can be split into multiple successive runs; see the *Restart* command. A full example for a two-dimensional well-tempered extended Lagrangian Metadynamics simulation can be found on *below*.

Note that Metadynamics simulations typically require very much computational time *(at least several 10 000 MD steps for a roughly converged result, depending on the Colvar choice)*. So this is by no means a method to "shortly try out". However, there are no cheaper methods for predicting free energy profiles *(apart from very simple approximations such as the harmonic oscillator)*, and the predictive power of computing free energy profiles comes at a price.

7.48.6 Command List

In the following, an alphabetical list of all commands currently known to the MD module is given. The description of each command starts with a small box which contains the command's name and a table of arguments and modifiers. The last-but-one column in the table specifies the type of each argument. Possible types are "Integer", "Real", "String", and "Keyword". In the latter case, the last column contains a list of allowed keyword values in { `braces` }. If the type is "Real" and is a physical quantity with unit, the quantity is given in the last column in [square brackets]. Each such box is followed by a textual description of the corresponding command.

7.48.7 Command Overview

Command	Description
Cell	Defines and modifies cells
Constraint	Manages constraints
Dump	Controls trajectory output
Initvel	Randomly initializes atom velocities
Manage_Colvar	Manages collective variables ("Colvars")
Manage_Region	Manages regions
Metadynamics	Sets parameters for Metadynamics runs
Minimize	Performs a Cartesian energy minimization
PrintLevel	Controls the output verbosity
Randomize	Sets the random seed
Restart	Restarts a simulation to seamlessly continue
Restraint	Manages restraints on Colvars
Run	Performs a molecular dynamics run
SCFLog	Controls the ORCA log file output
Screendump	Prints current MD state to screen
Thermostat	Manages thermostats
Timestep	Sets the integrator time step Δt

Cell

Manadory Arguments:		-		
Optional Arguments:		-		
Modifiers:	`Cube`	...	...	*see text*
	`Rect`	...	...	*see text*
	`Rhomb`	...	...	*see text*
	`Sphere`	...	...	*see text*
	`Ellipsoid`	...	...	*see text*
	`None`	—	—	—
	`Spring`	k	Real	*see text*
	`Elastic`	t_{avg}	Real	[time]
		c_{response}	Real	*see text*
	`Anisotropic`	—	—	—
	`Pressure`	...	...	*see text*
	`Fixed`	—	—	—

Defines a harmonic repulsive wall around the system *(the wall is "soft" with a spring constant and atoms can slightly penetrate; "hard" repulsive walls are not supported)*. This helps to keep the molecules inside of a well-defined volume, or to keep a constant pressure in the system. In the latter case, the cell can be defined as elastic, such that it exerts a well-defined pressure *(see below)*. Please note that ORCA does not feature periodic boundary conditions, and therefore, all cells are non-periodic (just repulsive walls). There are several cell geometries available *(only one type of cell can be active at a time)*:

- **`Cube`**: Defines a cubic cell. If two real values p_1 and p_2 are specified as coordinates, the cell ranges from (p_1, p_1, p_1) to (p_2, p_2, p_2). If only one real value p is supplied, the cell ranges from $(-\frac{p}{2}, -\frac{p}{2}, -\frac{p}{2})$ to $(\frac{p}{2}, \frac{p}{2}, \frac{p}{2})$, *i. e.* it is centered at the origin with edge length p.
- **`Rect`**: Defines an orthorhombic cell. Six real values x_1, y_1, z_1, x_2, y_2, and z_2 have to be specified as coordinates *(in this order)*. The cell will range from (x_1, y_1, z_1) to (x_2, y_2, z_2).
- **`Rhomb`**: Defines a parallelepiped-shaped cell (also termed as rhomboid sometimes). You have to specify twelve real values in total. The first three define one corner point p of the cell, and the remaining nine define three cell vectors v_1, v_2, and v_3, each given as Cartesian vector components. The cell is then defined as the

set of points $\{p + c_1 v_1 + c_2 v_2 + c_3 v_3 \mid 0 \leq c_1, c_2, c_3 \leq 1\}$ The vectors v_1, v_2, and v_3 do not need to be orthogonal to each other, but they may not all lie within one plane (cell volume would be zero).

- **Sphere**: Defines a spherical cell. You need to specify four real values c_x, c_y, c_z, and r. The cell will then be defined as a sphere around the central point (c_x, c_y, c_z) with radius r.
- **Ellipsoid**: Defines an ellipsoid-shaped cell. As first three arguments, you have to specify three real values c_x, c_y, c_z, which define the center of the ellipsoid to be (c_x, c_y, c_z). As fourth argument, a keyword has to follow, which may either be "XYZ" or "Vectors". In the "XYZ" case, three more real values r_x, r_y, and r_z have to be specified, which define the partial radii of the ellipsoid along the X, Y, and Z coordinate axes. If instead "Vectors" was given, nine more real values v_x^1, v_y^1, v_z^1, v_x^2, v_y^2, v_z^2, v_x^3, v_y^3, v_z^3 have to follow after the keyword. These values define three vectors $v^1 := (v_x^1, v_y^1, v_z^1)$, $v^2 := (v_x^2, v_y^2, v_z^2)$, and $v^3 := (v_x^3, v_y^3, v_z^3)$, which are the principal axes of the ellipsoid. These vectors have to be strictly orthogonal to each other. The length of each vector defines the partial radius of the ellipsoid along the corresponding principal axis.

All cell types define a harmonic potential $E_{\text{cell}}(r) := k \cdot r^2$ which acts on all atoms in the system **outside of the cell**, where r is the closest distance from the atom's center to the defined cell surface. Atoms whose center is inside of the cell or directly on the cell surface do not experience any repulsive force. Following from the definition, the force which acts on an atom outside of the cell is always parallel to the normal vector of the cell surface at the point which is closest to the atom center. This is trivial in case of cubic, rectangular, rhombic, and spherical cells, but not so trivial for ellipsoid-shaped cells.

The spring constant k in the above equation (*i. e.*, the "steepness" of the wall) can be specified by the "Spring" modifier, which expects one real value as argument. The spring constant has to be specified in the unit kJ mol^{-1} Å^{-2}, other units cannot be specified here. The default value is 10 kJ mol^{-1} Å^{-2}. Larger spring constants reduce the penetration depth of atoms into the wall, but may require shorter integration time steps to ensure energy conservation. If jumps in the total energy occur, try to use a smaller spring constant (*e. g.*, the default value).

The command "Cell None" disables any previously defined cell.

If you want to perform simulations under constant pressure, you can define an elastic cell. Then, ORCA accumulates the force which the cell exerts on the atoms in each time step, and divides this total force by the cell surface area to obtain a pressure. As this momentarily pressure heavily fluctuated, a running average is used to smooth this quantity. If the averaged pressure is larger than the external pressure which was specified, the cell will slightly grow; if it was smaller, the cell will slightly shrink. In the beginning of a simulation, the cell size will not vary until at least the running average history depth of steps have been performed.

An elastic cell is enabled by using the "Elastic" modifier after the cell geometry definition. Subsequently, two real values t_{avg} and c_{response} are required. While t_{avg} defines the length of the running average to smooth the pressure *(in units of physical time, **not** time steps)*, the c_{response} constant controls how fast the cell size will change at most. More specific, c_{response} is the fraction of the cell volume growth per time step if the ratio of averaged and external pressure would be infinite, and at the same time the fraction of the cell volume reduction per step if the aforementioned ratio is zero. Put into mathematical form, the cell volume change per time step is

$$V_{\text{new}} := \begin{cases} V_{\text{old}} \cdot \frac{c_{\text{response}} \cdot \frac{\langle p \rangle}{p_{\text{ext}}} + 1}{c_{\text{response}} + 1} & \text{if } \frac{\langle p \rangle}{p_{\text{ext}}} \leq 1, \\ V_{\text{old}} \cdot \frac{(c_{\text{response}} + 1) \cdot \frac{\langle p \rangle}{p_{\text{ext}}}}{\frac{\langle p \rangle}{p_{\text{ext}}} + c_{\text{response}}} & \text{if } \frac{\langle p \rangle}{p_{\text{ext}}} > 1, \end{cases} \tag{7.313}$$

where $\langle p \rangle$ represents the averaged pressure the system exerts on the walls, and p_{ext} is the specified external pressure. Good starting points are $t_{\text{avg}} = 100$ fs and $c_{\text{response}} = 0.001$. Please note that larger values of c_{response} or smaller values of t_{avg} may lead to uncontrolled fluctuations of the cell size. An already defined fixed cell can be switched to elastic by the command "Cell Elastic ..." (the dots represent the two real arguments).

By default, the size change of an elastic cell due to pressure is performed *isotropically*, *i. e.*, the cell is scaled as a whole, and exactly retains its aspect ratio. By specifying the "Anisotropic" modifier after switching on an elastic cell, the cell pressure is broken down into individual components, and the size of the cell is allowed to change independently in the individual directions. This, of course, only makes sense for the cell geometries Rect, Rhomb, and Ellipsoid. An already defined isotropic cell can be switched to anisotropic by simply executing "Cell Anisotropic".

In case of an elastic cell, the external pressure is defined by the modifier "Pressure", which expects either one or three real values as arguments. If one argument is given, this is the isotropic external pressure. If three arguments are supplied, these are the components of the pressure in X, Y, and Z direction *(in case of orthorhombic cells)*

or along the direction of the three specified vectors *(in case of parallelepiped-shaped and ellipsoid-shaped cells)*. This allows for anisotropic external pressure (probably only useful for solid state computations). Both the pressure and the pressure components have to specified in units of bar ($= 10^5$ N m^{-2}), other units cannot be used. If this modifier is not used, the default pressure will be set to 1.0 bar (isotropic) if an elastic cell is used. The external pressure of an already defined cell can be changed by the command "`Cell Pressure ...`" (the dots represent the real argument(s)).

As all cells are non-elastic by default, there is no keyword to explicitly request this at the time of cell definition. However, possible applications might require to use an elastic cell during equilibration period, and then "freeze" this cell at the final geometry for the production run. This can be achieved by using the "`Cell Fixed`" command (without any additional arguments).

If the cell is elastic, there is a volume work term which contributes to the total energy of the system. ORCA computes this term in every step and adds it to the potential energy. Without this contribution, the conserved quantity would drift excessively in elastic cell runs.

To completely switch off a previously defined cell, simply use "`Cell None`".

Please note that cells are **not** automatically restarted by using the *Restart* command.

Examples:

Cubic cell with edge length 10 Å centered around origin:
`Cell Cube 10`

Spherical cell with radius 5 Å centered around origin and 20 kJ mol^{-1} Å^{-2} wall steepness:

`Cell Sphere 0, 0, 0, 5 Spring 20`

Elastic orthorhombic cell from $(-2, -2, 0)$ to $(12, 12, 10)$, $t_{\text{avg}} = 100$ fs, $c_{\text{response}} = 0.001$:

`Cell Rect -2, -2, 0, 12, 12, 10 Elastic 100, 0.001`

Ellipsoid-shaped cell centered on origin with partial radii 5, 10, 15 Å along the X, Y, Z axes:

`Cell Ellipsoid 0, 0, 0 XYZ 5, 10, 15`

The commas are optional, but make sure to use them with negative numbers. By default, the minus operator will act as binary operator if possible *(see* **discussion above***)*.

Constraint

Mandatory Arguments:		*operation*	Keyword	{ `Add`, `Remove`, `List` }
		type	Keyword	{ `Cartesian`, `Distance`, `Angle`, `Dihedral`, `Center`, `Rigid` }
		atom(s)	Integer	—
Optional Arguments:		—		
Modifiers:	`Target`	*value(s)*	Real	—
	`Ramp`	*value(s)*	Real	—
	`Weights`	...	...	*see text*
	`All`	—		
	`Noprint`	—		

Manages constraints in the molecular dynamics simulation. Unlike *Restraints*, constraints are geometric relations which are strictly enforced at every time (*i. e.*, they do not fluctuate around their target value). All atoms involved in constraints have to be included in the active region. In principle, constraints also work in Cartesian geometry optimizations with the *Minimize* command, but the performance together with L-BFGS may be poor (except for Cartesian constraints, which work flawlessly in L-BFGS). In these cases, try to use the simulated annealing method instead.

The simplest possibility is to constrain the Cartesian position of an atom to some value. A zero-based atom index is required. The command `Constraint Add Cartesian 3` would fix the fourth atom in the simulation at its current

position in space. If the desired position shall be explicitly given, it can be specified via the `Target` modifier, *e. g.*, `Constraint Add Cartesian 3 Target 5.0 1.0 1.0`. To determine which dimensions to fix, one of the `XYZ`, `XY`, `XZ`, `YZ`, `X`, `Y`, or `Z` modifiers can be added. For example, `Constraint Add Cartesian 3 X Target 1.0` would constrain the X coordinate of atom 3 to the absolute value 1.0, but would not influence movement along the Y and Z coordinate at all.

By using the `Distance` keyword, distances between atoms can be fixed. The command `Constraint Add Distance 3 5` would fix the distance between atom 3 and 5 to its current value. You need to specify exactly two atom indices; multiple distance constraints are entered via multiple `Constraint` commands. Also here, a desired distance value can be given via the `Target` modifier, such as `Constraint Add Distance 3 5 Target 350_pm`.

Similarly, angles and dihedral angles between atoms can be fixed with the `Angle` and `Dihedral` keywords. Angles are defined by three atom indices, and dihedral angles by four atom indices. Also here, target values may be specified. Any combination of Cartesian, distance, angle, and dihedral constraints may be used simultaneously, and may even be applied to the same group of atoms. A molecule can be made completely rigid by constraining all its bonds, angles, and torsions. Please make sure that your constraints are not over-determined, and do not contradict each other. Otherwise, they can't be enforced and the simulation will print warnings or crash.

A different and powerful class of constraints can be defined with the `Center` argument. Directly after the keyword, a list of integer atom numbers is expected. This list can be a combination of numbers and ranges, *e. g.*, "`1, 3, 5..11, 14`". The weighted average position of this subset of atoms is then constrained to a fixed position in Cartesian space. By default, the weights are taken as the atom masses, such that the center of mass of the selected atoms is kept fixed. This allows, *e. g.*, to run a MD simulation of two molecules with fixed center of mass, such that their center of mass distance remains constant. Custom weights for the definition of the center can be entered by using the `Weights` modifier after the atom list. It expects exactly the same number of real arguments as the length of the specified atom list. The geometric center of a group of atoms can be held fixed by setting all weights to 1.0, for example "`Constraint Add Center 2, 5..7 Weights 1.0 1.0 1.0 1.0`". If desired, a `Target` for the center position can be given, which expects three real numbers for the X, Y, and Z coordinate after the keyword. If no target is specified, the current center position is held fixed.

With the `Rigid` type of constraints, complete groups of atoms can be kept rigid, *i. e.*, keep all their distances and angles relative to each other, but move as a whole. After the `Rigid` keyword, a list of atom numbers is expected. More than one group of atoms can be kept rigid at the same time – just call the `Constraint Add Rigid` command multiple times with different atom lists. Internally, the rigid constraint is realized by defining the correct number of distance constraints. Such a large number of distance constraints is hard to converge; therefore, warning messages that RATTLE did not converge will **not** be shown if a rigid constraint is active. Almost planar (or even linear) groups of atoms are hard to keep rigid by using only distance constraints. It might help do add a dummy atom outside of the plane and include this into the constraint.

ORCA supports constraints with linearly changing target value during the simulation. To define such a constraint, write "`Ramp`" directly after the "`Target`" modifier. After "`Ramp`", twice the number of real numbers that would have been required for "`Target`" follows (two instead of one for distances, angles, and dihedrals; six instead of three for "`Cartesian XYZ`", and so on). The first half of these arguments are the starting values, the second half are the final target values. For example, "`Constraint Add Distance 3 5 Target Ramp 300_pm 400_pm`" will define a distance constraint with a target value rising from 300 pm to 400 pm. The ramp will be performed once during the *Run* command which follows next after the constraint definition. Therefore, the number of steps specified in this *Run* command also specifies the rate at which the constraint target is modified. After the ramp has been completed once, the final (constant) target value(s) will be used for all subsequent *Run* commands.

If an already defined constraint is defined again, it is overwritten, *i. e.*, the old version of the constraint is automatically deleted first.

Constraints are removed with the `Remove` keyword. You can either remove single constraints, *e. g.*, `Constraint Remove Distance 3 5`, or groups of similar constraints. To remove all angle constraints, use `Constraint Remove Angle All`. To remove all restraints, enter `Constraint Remove All`.

The `List` argument prints all currently active constraints to the screen and log file. No additional arguments can be specified.

By default, the external force acting on each constraint is computed in every MD step and written to a file named "`basename-constraints.csv`" *(one column per constraint)*. This can be useful – the average force acting on a constraint can be, *e. g.*, used for thermodynamic integration [430]. If a large number of constraints is defined, this

might waste computer time if it is not required. In these cases, the constraints can be defined with the `Noprint` modifier. For such constraints, the acting forces are not computed and not written to the file. Note that constraints which have been pre-defined *(e. g., by the force field for rigid molecules such as TIP3P water)* automatically have this modifier.

Please note that each constraint decreases the number of the system's degrees of freedom (DoF) by one. This effect is included, *e. g.*, in the temperature computation, where the DoF count enters. From this consideration, it can also be understood that a constraint behaves significantly different from a restraint with very large spring constant: In the former case, the DoF is removed from the system; in the latter case, the DoF is still there, but can only move in a tiny interval.

It is computationally inefficient to define a large number of Cartesian constraints if a subset of atoms simply shall be fixed. A more efficient approach is to define an active region which only contains the atoms which shall be movable (see *Manage_Region* command). All atoms outside of the active region will not be subject to time integration and therefore keep their positions. However, please note that these atoms may not be involved in any other (distance, angle, dihedral) constraint.

Dump

Mandatory Arguments:		*quantity*	Keyword	{ `Position`, `Velocity`, `Force`, `GBW`, `EnGrad` }
Optional Arguments:	—			
Modifiers:	`Format`	*fmt*	Keyword	{ `XYZ`, `PDB`, `DCD` }
	`Stride`	*n*	Integer	—
	`Filename`	*fname*	String	—
	`Region`	*region*	…	…
	`Replace`	—		
	`None`	—		

Specifies how to write the output trajectory of the simulation. The *quantity* argument can be one of the keywords `Position`, `Velocity`, `Force`, `GBW`, and `EnGrad`. While the velocities are written in Angstrom/fs, the unit of the forces is Hartree/Angstrom. The following paragraphs only apply to the first three quantities. Dumping GBW and EnGrad files works differently, and is described at the very end of this section.

The `Stride` modifier specifies to write only every n-th time step to the output file (default is $n = 1$, *i. e.*, every step). A stride value of zero only writes one frame to the trajectory at the time when the `Dump` command is called – no further frames will be written during the run. This can be helpful, *e. g.*, to write an initial PDB snapshot for DCD trajectories, or to keep a single GBW file at some point.

The `Format` modifier sets the format of the output file. Currently, only the `XYZ`, `PDB`, and `DCD` formats are implemented. Please note that the `DCD` format is not well-defined, and different programs use different formats with this extension. Furthermore, `DCD` files do not store atom type information and are only valid together with a PDB snapshot of the system (a single PDB snapshot can be written via "`Dump Position Format PDB Stride 0`"). If not specified, ORCA tries to deduce the format from the file extension of the specified file name. If also no file name is given, trajectories will be written in XYZ format by default.

The `Filename` modifier sets the output file name. If not specified, the default file name will have the form "`proj-qty-rgn.ext`", where `proj` is the base name of the ORCA project, `qty` is one of `postrj`, `veltrj`, or `frctrj`, `rgn` specifies the name or number of the region for which the dump is active, and `ext` is the file extension selected by the `Format` modifier.

If the trajectory file already exists at the beginning of a *Run* command, new frames will be appended to its end by default. If you want to overwrite the existing file instead, use the `Replace` modifier. The old existing file is erased only once after a dump with this modifier has been specified. If multiple *Run* commands follow after the dump definition, the trajectory will **not** be replaced before each of these runs, only before the first one among them. To overwrite the file another time, simply re-define the dump with the `Replace` modifier. If the file does not yet exist at the beginning of a run, this modifier has no effect. Appending frames to DCD trajectories is not possible *(because they store the total frame count in the header)*. Therefore, `Replace` is automatically switched on if the format is DCD.

With the `Region` modifier, the trajectory output can be restricted to a specific region (*i. e.*, subset of atoms). This modifier expects one argument, which is either the name of a pre-defined region or the number of a user-defined region *(see above)*. If not specified, the trajectory of the whole system will be written. Multiple dump commands for multiple regions can be active at the same time, but each pair of region and quantity *(position/velocity/force)* can have only one attached dump command at a time (re-defining will overwrite the dump settings).

Use the `None` modifier to disable writing this quantity to an output file. The command "`Dump Position None`" will disable writing of all position trajectories for all regions. To disable only the dump for a specific region, use "`Dump Position Region r None`", where `r` is the name or number of the region.

The default is to write a position trajectory with `Stride 1` and `Format XYZ` to a file named "`proj-postrj-all.xyz`", where "proj" is the base name of the ORCA project. If you want to create no output trajectory at all, use "`Dump Position None`" as described above.

The `Dump GBW` command keeps a copy of the GBW file every n steps, which can be used for computing properties along the MD trajectory, *e. g.*, plotting orbitals. This does not yield a trajectory, as all the GBW files are stored individually. The value of n is controlled by the `Stride` modifier. The file names are formed by appending the step number *(six digits with leading zeros)* followed by ".gbw" to the `Filename` argument. Therefore, this argument should not contain the ".gbw" extension by itself. If the `Filename` modifier is not specified, the default will be "proj-step", where "proj" is the base name of the ORCA project. This will lead to files such as "`proj-step000001.gbw`", etc. The `Format` and `Region` modifiers can not be used for `Dump GBW`.

In a very similar way, `Dump EnGrad` stores an ORCA `.engrad` file *(energy and gradient)* every n steps. All the `.engrad` files are stored individually *(not as a continuous trajectory)*. The value of n is controlled by the `Stride` modifier. By default file names such as "`proj-step000001.engrad`" will be used.

Initvel

Mandatory Arguments:		*temp*	Real	[temperature]
Optional Arguments:		—		
Modifiers:	`Region`	*region*	...	...
	`No_Overwrite`	—		

Initializes the velocities of the atoms by random numbers based on a Maxwell–Boltzmann distribution, such that the initial temperature matches $temp$ (see also section *1.5.2*). Please note that this overwrites all velocities, so do not call this command when your system is already equilibrated (*e. g.*, to change temperature – use a thermostat instead).

The total linear momentum of the initial configuration is automatically removed, such that the system will not start to drift away when the simulation begins. This only concerns the initial configuration. Total linear momentum might build up during the simulation due to numeric effects.

With the `Region` modifier, the initialization of velocities can be performed for a specific region (*i. e.*, subset of atoms). This modifier expects one argument, which is either the name of a pre-defined region or the number of a user-defined region *(see above)*. If not specified, the command acts on the whole system.

The `No_Overwrite` modifier only initializes the velocities if no atom velocities have been defined/read before. This is useful in combination with the `Restart` command: After reading an existing restart file, the velocities are already known, and the initialization will be skipped if this modifier is used. The following combination of commands in a MD input would initialize the velocities only upon first execution, and restart the positions and velocities on all following executions of the same input:

```
Restart IfExists
Initvel 350_K No_Overwrite
```

If neither the `Initvel` command nor a *Restart* command is not invoked before a *Run* call, the atom velocities will be initialized to zero before starting the run.

Manage_Colvar

Mandatory Arguments:		*operation*	Keyword	{ Define }
		id	Integer	—
		type	Keyword	{ Distance, Angle, Dihedral, CoordNumber }
Optional Arguments:		—		
Modifiers:	Atom	*atom*	Integer	—
	Group	*atomlist*	Integers	—
	Weights	*weights*	Reals	—
	Cutoff	*cutoff*	Real	[length]
	Noprint	—		

Defines collective variables ("Colvars") which are used for Metadynamics or to impose Restraints on the system. In a general sense, a Colvar is simply a continuous function of all the atom positions which returns a real number. As Colvars don't have any effect on the simulation by themselves, they can currently only be defined or re-defined; there is no requirement for deleting them. The second argument of the Manage_Colvar command is the number of the Colvar. This number is used to address the Colvar later. Allowed numbers are within the range of $1 \ldots 10000$. If a Colvar number which had previously been defined is defined again, it is simply overwritten *(and all restraints based on the old Colvar are deleted!)*. The third mandatory argument is the type of the Colvar, which can be Distance, Angle, Dihedral, and CoordNumber. More Colvar types will probably be added in the future *(feel free to make suggestions in the forum!)*.

Distance Colvars are defined between two points in space. Each point can either be a single atom (expressed by "Atom") or the weighted average (center) of the positions of a group of atoms (expressed by "Group"). For example, the command "Manage_Colvar Define 1 Distance Atom 0 Atom 7" defines Colvar 1 to be the distance between atoms 0 and atom 7 *(as always, the atom count starts at zero)*. On the other hand, the command "Manage_Colvar Define 2 Distance Group 0 1 2 Group 3 4 5" sets Colvar 2 to be the distance between the centers of atoms 0, 1, 2 and atoms 3, 4, 5. If many atoms shall be selected, the range syntax "Group 0..2" can be used, including multiple such ranges if required, such as in "Group 0..2, 5, 7..11" *(see also discussion of the* Manage_Region *command)*. By default, the center of mass is used for groups. However, weights can be manually specified if required by using the "Weights" modifier directly after the atom list for the center is finished. "Weights" expects as many real numbers as the group possesses atoms, for example "Manage_Colvar Define 2 Distance Atom 0 Group 1 2 3 Weights 1.0 1.0 1.0". The "Atom" and "Group" syntax can be mixed, *e. g.*, to define the distance between a single atom and a center of mass. When defining distance Colvars, one of the modifiers X, Y, Z, XY, XZ, YZ, and XYZ may be specified directly after "Distance". The first three among them denote that the positions shall be projected onto the corresponding Cartesian vector before computing the distance. The following three modifiers require that the two positions are projected into the corresponding Cartesian plane prior to computing the distance. The last one is the default *(just measure the standard distance in 3D space)* and does not need to be specified explicitly.

In a very similar manner, angle Colvars can be defined. Instead of two points in space, an angle Colvar is defined via three points in space, each of which can either be an "Atom" or a "Group" *(see above)*. For example, the command "Manage_Colvar Define 3 Angle Group 0 1 2 Atom 3 Atom 4" defines Colvar 3 to be the angle spanned by the mass center of atoms 0, 1, 2, atom 3, and atom 4, respectively.

Dihedral Colvars are defined through four points in space, each of which can either be an "Atom" or a "Group" *(see above)*. For example, the command "Manage_Colvar Define 4 Dihedral Atom 0 Group 1..5 Atom 6 Atom 7" defines Colvar 4 to be the dihedral angle spanned by atom 0, the mass center of atoms 1, 2, 3, 4, 5, atom 6, and atom 7, respectively.

The Colvar type "CoordNumber" has been suggested in literature [408] to measure the coordination number of some atom species around some other atom. An example where this type of Colvar has been successfully applied is the calculation of pK_A values of weak acids in solvent via Metadynamics [854, 855]. The Colvar is defined by the following equation

$$C := \frac{1}{N_A} \sum_i^{N_A} \sum_j^{N_B} \frac{1 - \left(\frac{r_{ij}}{r_{\text{cut}}}\right)^6}{1 - \left(\frac{r_{ij}}{r_{\text{cut}}}\right)^{12}},$$

where N_A is the set of atoms which is coordinated *(typically only one atom)*, N_B is the set of coordinating atoms, r_{ij} is the distance between atoms i and j, and r_{cut} is a constant cutoff distance which specifies a threshold for coordination. After "`CoordNumber`", two atoms or groups of atoms must follow, which correspond to N_A and N_B, respectively. The distance cutoff is specified after the "`Cutoff`" modifier which should follow the two group definitions. For example, the command "`Manage_Colvar Define 5 CoordNumber Atom 0 Group 1..10 Cutoff 200_pm`" defines Colvar 5 as the coordination number of the group of atoms 1 to 10 around atom 0 with a distance cutoff of $r_{\mathrm{cut}} = 200$ pm.

For every defined Colvar, the temporal development of the position and the external force acting on the Colvar is written to a text file named "`basename-colvars.csv`" in every MD step by default. If a large number of Colvars are defined, this might be a waste of time and disk space. In these cases, the "`Noprint`" modifier can be specified when defining the Colvar. Colvars defined with this modifier will not appear in the text file, and the force acting on the Colvar will not be computed *(if not required otherwise, e. g., for restraints)*.

`Manage_Region`

Mandatory Arguments:		*identifier*	Keyword/Integer	...
		operation	Keyword	{ `Define`, `AddAtoms`, `RemoveAtoms` }
		atomlist	Integer(s)	—
Optional Arguments:		—		
Modifiers:	`Element`	*elem*	String	—

Defines or modifies regions. Regions are just subsets of atoms from the system – see [**Section *1.3***](#moldyn:sec_regions) above.

As described above, there exist several pre-defined regions which are identified by names. The only such pre-defined region which can be re-defined by the user is the `active` region. All atoms in this region are subject to time integration in molecular dynamics and displacement in minimization runs. All other atoms are simply ignored and remain on their initial positions. Please note that the `active` region may never be empty.

To re-define the `active` region, use the command "`Manage_Region active Define 1 5 7 ...`". The integer arguments after `active` are the numbers of the atoms to be contained in the region, in the order given in the ORCA input file. Atom numbers are generally zero-based in ORCA, *i. e.*, counting starts with 0.

Apart from that, user-defined regions are supported. These are identified with an integer number instead of a name. The integer numbers do not need to be sequential, *i. e.*, it is fine to define region 2 without defining region 1. To give an example, the command "`Manage_Region 1 Define 17 18 19`" defines region 1, and adds atoms 17, 18, and 19 to this newly defined region. Using `Define` without an atom list, such as in "`Manage_Region 1 Define`", deletes the user-defined region, as it will be empty then. Atoms can be added to or removed from previously defined regions (including the `active` region) with the `AddAtoms` and `RemoveAtoms` operations. The atom numbers specified after the operation name are added to or removed from the region. For example, "`Manage_Region active RemoveAtoms 15 16 17`" will remove atoms 15 to 17 from the `active` region (and add them to the `inactive` region instead).

If you want to specify a range of atoms, you can use the syntax "`a..b`" to include all atom numbers from `a` to `b`. If you want only, *e. g.*, every third atom in a range, you can use "`a..b..i`" to add the range from `a` to `b` with increment `i`. As an example, "`2..10..3`" will expand to the list `2, 5, 8`. You can mix atom numbers and ranges, as shown in the following two examples *(as always, the commas are optional)*:

```
Manage_Region active Define 1, 4, 5..11, 14, 17..30..2
Manage_Region active RemoveAtoms 4, 15..17
```

Instead of an atom list, the `Element` modifier can be used, followed by a string which represents an element label. This will have the same effect as specifying an atom list with all atoms of this element type instead. Don't forget the double quotes around the element label string. For example, `Manage_Region active RemoveAtoms Element "H"` removes all hydrogen atoms from the active region.

Metadynamics

Mandatory Arguments:		—		
Optional Arguments:		—		
Modifiers:	Off	—		
	Reset	—		
	Colvar	*colvar*	Integer	—
	Scale	*scale*	Real	...
	Wall	*side*	Keyword	{ Lower, Upper }
		target	Real	phys. unit
		k	Real	kJ mol^{-1} phys. $unit^{-2}$
	HillSpawn	*frequency*	Integer	—
		height	Real	kJ mol^{-1}
		sigma	Real	phys. unit
	Range	*from*	Real	phys. unit
		to	Real	phys. unit
		resolution	Integer	—
	Store	*store*	Integer	—
	Temperature	*temp*	Real	[temperature]
	WellTempered	*biastemp*	Real	[temperature]
	Lagrange	*mass*	Real	a.m.u.
		k	Real	kJ mol^{-1} $scaleunit^{-2}$
		target_temp	Real	[temperature]
		tau	Real	[time]

Sets the parameters for a *Metadynamics* simulation [487]. After all parameters have been set, the actual simulation can be started by a *Run* command. The parameters can either all be set in a single call to the Metadynamics command, or distributed over multiple such calls to avoid very long lines. In both cases, there are some rules for the order of parameter settings. All Colvars for the Metadynamics need to be specified before setting any other parameters. Modifiers which are related to Colvars (such as Scale, Wall, or Range) only apply to the Colvar that was specified last before them in the Metadynamics command.

The Colvar modifier specifies a Colvar to be used in the Metadynamics simulation. It expects one integer argument, which is the number of the Colvar, as defined before via the *Manage_Colvar* command. The ORCA MD module supports one- and two-dimensional Metadynamics, so either one or two Colvar modifiers can be given. The number of Colvar modifiers specified defines the dimensionality of the Metadynamics simulation. Please note that modifiers which are related to Colvars (such as Scale, Wall, or Range) only apply to the Colvar that was specified last before them, so after specifying the first Colvar, these should be set before specifying the second Colvar.

Colvars can have different physical units, such as Angstrom for distances and degree for angles. In a multi-dimensional Metadynamics run, the different numerical magnitude of the corresponding numbers can be an issue: Angles span over a range of 180 degree, while distances will often be within an interval of only 10 Angstrom. To bring all Colvars to a similar scale, the Metadynamics module internally divides every Colvar by a user-supplied constant. These internal values are dimensionless, they will be referred to as "scale units". For a previously specified Colvar, the scale constant can be set via the Scale modifier. It expects one real argument which has to be specified in physical units of the Colvar (length units for distance Colvars, angle units for angle and dihedral Colvars). Coordination number Colvars are dimensionless anyway. If not specified, reasonable default values for the scale are used, which are 1.0 Angstrom for distance Colvars, 20.0 degree for angle and dihedral Colvars, and 0.2 for coordination number Colvars. Note that the Scale modifier only applies to the Colvar given last before it in the Metadynamics command.

As an example, consider the following commands to set up a two-dimensional Metadynamics simulation:

```
Manage_Colvar Define 1 Distance Atom 0 Atom 1
Manage_Colvar Define 2 Angle Atom 0 Atom 1 Atom 2
Metadynamics Colvar 1 Scale 1.0_A Colvar 2 Scale 10.0_Deg
```

To keep Colvars within the region of interest during Metadynamics simulations, harmonic walls can be imposed on Colvars. This is achieved via the `Wall` modifier. As a first argument, it expects the direction of the wall, which can be `Lower` or `Upper`. The second argument is the position of this wall – given in physical units of the Colvar (*e. g.*, Angstrom for distance Colvars), **not** in scale units. As an optional third real argument, the spring constant of the harmonic wall can be specified in kJ mol^{-1} unit^{-2}, where unit is the default physical unit of the Colvar (Angstrom for distances, degree for angles). If omitted, a spring constant of 50 kJ mol^{-1} Angstrom^{-2} for distances, 0.5 kJ mol^{-1} degree^{-2} for angles, and 250.0 kJ mol^{-1} for coordination numbers is used *(a reasonable choice)*. Both lower and upper wall can be defined after one `Wall` modifier, such as in "`Metadynamics Colvar 1 Wall Lower 3.0_A 50.0 Upper 10.0_A 50.0`". Note that one-sided harmonic walls can also be imposed on Colvars via the *Restraint* command. For standard Metadynamics, this is redundant. However, for extended Lagrangian Metadynamics *(see below)*, it makes a difference: Restraints act on the Colvar and therefore on the real atomistic system, whereas the walls defined in the `Metadynamics` command act directly on the virtual particle. Again, note that the `Wall` modifier only applies to the Colvar given last before it in the `Metadynamics` command.

The last modifier which applies to Colvars is the `Range` modifier. It has no influence on the Metadynamics run itself, and only controls the output of the free energy profiles. The `Range` modifier expects three arguments. The first two have to be real numbers and define the lower and upper interval borders for which the free energy profile with respect to this Colvar shall be output. The third argument is of integer type and controls the number of grid points to produce for this interval. In two-dimensional Metadynamics, both Colvars can have associated `Range` modifiers, which then control the interval and resolution of the 2D grid for the free energy profile. In this case, the grid should not be much finer than 100×100; otherwise, the evaluation of the grid points will become quite slow. If no `Range` modifier is given, default values are used (range of $0 \ldots 20$ Angstrom for distance Colvars, $0 \ldots 180$ degree for angle Colvars, $-180 \ldots 180$ degree for dihedral Colvars, and $0 \ldots 3$ for coordination number Colvars). As above, note that the `Range` modifier only applies to the Colvar given last before it in the `Metadynamics` command.

Addition of new Gaussian hills to the bias potential is controlled via the `HillSpawn` modifier. It expects three arguments. The first argument has to be of integer type and defines the hill spawning frequency, *i. e.*, every how many MD steps a new hill is added *(typically every 10 – 50 fs)*. The second argument is a real number and specifies the height of each new hill in units of kJ mol^{-1} *(typically 0.1 – 1.0 kJ mol^{-1})*. The third argument sets the width of the Gaussian hills σ (which is the standard deviation, **not** the variance σ^2) in "scale units" – see the `Scale` modifier above. In two-dimensional Metadynamics simulations, the width applies for both dimensions at the same time, and the scales of the two Colvars need to be adjusted to obtain the correct "aspect ratio" of the hill width. Standard choices for σ are 0.1 – 1.0 scale units. The hill spawning parameters can be changed at any point during a Metadynamics simulations, not modifying the hills which are already present. If spawning of new hills shall be temporarily suspended during a Metadynamics simulations, "`Metadynamics HillSpawn Off`" can be specified. If you want do delete all hills, consider the `Reset` modifier described below.

The `Store` modifier controls how often the current intermediate free energy profile is saved to disk. It expects one integer argument which specifies the number of MD simulation time steps between two such stores. In case of one-dimensional Metadynamics, these free energy profiles have the file names "`basename-metadynamics_profile_###.csv`", where "`###`" indicates the step number after which the profile was written. In addition to that, a file "`basename-metadynamics_profile_history.csv`" is written, which contains all the previously computed free energy profiles as columns, so that they can easily be printed in one single plot. For two-dimensional Metadynamics, Gnuplot source files with file names "`basename-metadynamics_2d_profile_###.gp`" are written, which can be converted into contour plots with the freeware tool Gnuplot *(runs both on Windows and GNU Linux)*. The raw data for these contour plots can be found in corresponding files named "`basename-metadynamics_2d_profile_###.gp.csv`". Note that in both cases, the free energy scale origin is set to the deepest free energy well, so that all numbers are positive. If the `Store` modifier is not specified, the free energy profiles are stored every 1000 MD steps per default.

The `WellTempered` modifier switches on well-tempered Metadynamics [73]. In contrast to standard Metadynamics, the free energy profile converges towards a limit for long runs with this approach. In short terms, this approach scales down the hill size at positions where already many hills have been spawned before, so that the changes in the bias potential become smaller over time *(convergence)*. The `WellTempered` modifier expects one real argument, which is the so-called bias temperature, specified in temperature units. The bias temperature should be chosen in a way so that $\frac{1}{2} \cdot k_\mathrm{B} \cdot T_\mathrm{Bias}$ is around the same size as the largest barrier which the simulation shall overcome. For example, a bias temperature of 12 000 K is well-suited to overcome barriers of around 100 kJ mol^{-1}. Note that the Metadynamics module needs to know the simulation temperature in order to reconstruct the free energy profile in a well-tempered Metadynamics run. Typically, a thermostat should be active during a Metadynamics run, keeping the simulation temperature constant. In this case, the temperature is simply obtained from the thermostat.

However, if no thermostat for the region `all` is specified, the simulation temperature has to be specified manually for well-tempered Metadynamics. This can be achieved by the `temperature` modifier, which expects one real argument – the simulation temperature in temperature units.

The `Lagrange` modifier switches on extended Lagrangian Metadynamics [408]. In this variant, a virtual particle *(with mass and velocity)* moves in the space spanned by the Colvars, and the only connection between this particle and the real atomistic system is a harmonic spring. The bias potential *(the Gaussian hills)* only acts on the virtual particle. The first argument is the mass of the virtual particle in a.m.u. The second argument is the harmonic spring constant in units of kJ mol^{-1}, which is evaluated in scale units – see the `Scale` modifier above. Typical values depend on the system and Colvars, but might be 100 a.m.u. and 10 kJ mol^{-1}. Optionally, a third and fourth parameter can be given to switch on thermostating of the virtual particle. A simple Berendsen thermostat is applied here. The third argument is the target temperature of the virtual particle, and the fourth argument is the thermostat time constant τ. A good choice would be a target temperature of 100 K and $\tau = 10$ fs. Note that in contrast to the normal Berendsen thermostat, the virtual particle is only cooled, but never heated. In other words, the thermostat only becomes active if the instantaneous temperature of the virtual particle becomes larger than the target temperature. This is to ensure that the virtual particle can change its direction – otherwise, it might happen that it is driven in the same direction for very long time intervals.

The `Reset` modifier resets the bias profile, *i. e.*, it deletes all hills which had been spawned, so that the bias profile becomes flat again. All other parameters of the Metadynamics simulation are not modified. If, for example, hill spawning is still on, then new hills will be spawned in the next simulation run.

The `Off` modifier completely switches off Metadynamics. It deletes all hills and turns off the Metadynamics module. It also resets the choice of Colvars for Metadynamics, so you will need to use this first if you want to set up a second different Metadynamics run within the same input script. This modifier can only be given as first argument to the `Metadynamics` command, and no further arguments can follow.

A restart file for the Metadynamics module (file name "`basename.metarestart`") is written each time a new hill has been spawned. The *`Restart`* command detects this file and automatically restarts the Metadynamics run *(i. e., loads all hills and the positions and velocities of the extended Lagrangian virtual particle if active)*. However, this only happens when Metadynamics is active and set up at the time when the *`Restart`* command is invoked. The parameters for the Metadynamics simulation are **not** restarted. Therefore, leave all parameter settings via calls to the `Metadynamics` command in place in your input file, and simply call the *`Restart`* command after all those, directly before the *`Run`* command.

Please see also the discussion on Metadynamics in [**Section** *1.3*](#moldyn:sec_metadynamics).

This section is concluded with a full example for a two-dimensional well-tempered extended Lagrangian Metadynamics run with restart ability *(just run the same input again to continue the simulation where it ended last)*. The two Colvars are defined as distances between atoms. You need to adapt all parameters in blue to your question and system:

```
Timestep 0.5_fs
Initvel 350_K
Thermostat NHC 350_K Timecon 100.0_fs
Dump Position Stride 1 Filename "trajectory.xyz"
Manage_Colvar Define 1 Distance Atom 0 Atom 1
Manage_Colvar Define 2 Distance Atom 2 Atom 3
Metadynamics Colvar 1 Scale 1.0_A Wall Lower 3.0 50.0 Upper 10.0 50.0 Range 0.0 15.0 100
Metadynamics Colvar 2 Scale 1.0_A Wall Lower 1.0 50.0 Upper 8.0 50.0 Range 0.0 13.0 100
Metadynamics HillSpawn 40 0.5 0.5 Store 2000
Metadynamics WellTempered 6000_K
Metadynamics Lagrange 100.0 10.0 200.0_K 10.0_fs
Restart IfExists
Run 100000
```

Minimize

Mandatory Arguments:		—		
Optional Arguments:		*method*	Keyword	{ Combined, LBFGS, Anneal }
Modifiers:	Steps	*n*	Integer	—
	MaxGrad	*thres*	Real	[kJ mol^{-1} Å^{-1}]
	RMSGrad	*thres*	Real	[kJ mol^{-1} Å^{-1}]
	TempConv	*thres*	Real	[temperature]
	Accel	*value*	Real	—
	Damp	*value*	Real	—
	StepLimit	*value*	Real	[length]
	History	*n*	Integer	—
	Noise	*value*	Real	[length]
	OnlyH	—		

Performs a Cartesian energy minimization of the system. For molecules, this is less efficient than ORCA's built-in geometry optimization in internal coordinates *(i. e., requires more steps to converge)*. However, the algorithms employed here also work with large atom counts *(e. g., 50 000)* as sometimes encountered in QM/MM simulations, which is absolutely out of scope of ORCA's primary optimization module. Furthermore, the minimization also works under all types of constraints (which some limitations in the case of L-BFGS) that have been set with the *Constraint* command, and also includes the effect of the repulsive simulation cell if activated. Only atoms contained in the active region are displaced, while all other atoms are kept at their positions.

The simplest way of performing a minimization is simply calling the Minimize command without arguments. This defaults to the L-BFGS method, which is fairly robust and efficient. If the minimization seems unstable, try to reduce the History or StepLimit parameters. L-BFGS may sometimes show poor performance with constraints other than Cartesian type. Apart from that, there is also a simulated annealing method implemented, which can be selected by specifying Anneal as the first argument. In contrast to L-BFGS, the simulated annealing method works equally well with all types of constraints. There is also a Combined method, which is a combination of some L-BFGS steps in the beginning, followed by a simulated annealing run until the temperature falls below a threshold, and another final L-BFGS run until the convergence criteria are reached.

With the Steps modifier, the maximum number of minimization steps can be specified. If this number of steps has been performed, the minimization finishes, no matter if the convergence criteria are fulfilled or not. The default value is 500.

The MaxGrad and RMSGrad modifiers control the convergence thresholds for the largest gradient on some atom and the root mean square average of the gradients. The default values are currently set to 5.0 and 1.0 kJ mol^{-1} Å^{-1}, respectively, which is about the same criterion as the default setting in the primary ORCA geometry optimization.

If the TempConv modifier is given, a simulated annealing run finished after the temperature was monotonously decreasing within 5 successive steps, and dropped below the specified value. Note that the simulated annealing run will finish if either this condition is reached, or the gradient thresholds are observed. It is not required to fulfill both criteria.

The Accel modifier specifies the acceleration factor for simulated annealing runs *(has no effect on L-BFGS)*. As long as the angle between velocity vector and gradient vector of some atom is below 90 degrees, the gradient is multiplied by this factor and the velocity is multiplied by a fraction of this factor. This helps to enforce a faster movement in gradient direction. The default value is 4.0. If this feature is not desired, use `Accel 1.0` to switch it off (1.0 means "no artificial acceleration").

The Damp modifier is the damping factor for simulated annealing runs *(has no effect on L-BFGS)*. Atom velocities are multiplied by this factor in every integration step. The default value is 0.98. Smaller values will make the algorithm more stable and less prone to oscillations and overshoots, but will also require significantly more steps to converge. Don't use values ≥ 1, as then it won't be an "annealing" anymore :-)

The StepLimit modifier specifies the maximum displacement of any atom *(in length units)* that can happen in one step of a minimization run. This can help to avoid large, unreliable steps which could lead to abrupt jumps in geometry and very high potential energies. This modifier concerns both L-BFGS and simulated annealing runs.

Negative values disable the step limit. The step limit is disabled by default. If you need to switch it on, try something in the order of 0.1 Å.

The `History` modifier controls the depth of gradient and position vector history that is used in the L-BFGS method to approximate the inverse Hessian. The default value is 20. Smaller values can help to stabilize the algorithm.

With the `Noise` modifier, small random numbers can be added to the atom positions before the minimization starts. This can help to escape local maxima and saddle points in the minimization. For example, a minimization of an initially linear water molecule would not be able to leave this maximum – but with some random "noise", it will be possible. The modifier expects one real argument which specifies the maximum atom displacement in length units (something like 0.01 Å will be reasonable). This feature is switched off by default.

If the `OnlyH` modifier is given, all non-hydrogen atoms are removed from the active region before the minimization starts. After the minimization has finished, the original active region is restored. This is helpful if only hydrogen positions shall be optimized, *e. g.*, to refine experimental crystal structures.

PrintLevel

Mandatory Arguments:	*value*	Keyword	{ `Low`, `Medium`, `High`, `Debug` }
Optional Arguments:	—		
Modifiers: —			

Controls the amount of information which is printed to the screen during the simulation. `Debug` should be used only in rare cases, because it might slow the simulation down heavily.

The default value is `Medium`.

Randomize

Mandatory Arguments:	—		
Optional Arguments:	*seed*	Integer	—
Modifiers: —			

There are a few algorithms in the ORCA MD module which rely on random numbers, *e. g.*, the initialization of atom velocities with the *Initvel* command. These random numbers are so-called "pseudo-random numbers", produced by a deterministic generator. This generator has a *state*, which is simply an integer number. If initialized to the same state, the generator will always create the same sequence of "random" numbers. This sounds like a deficiency at first thought, but is a very important feature for scientific reproducibility and for debugging purposes. If you start the same MD input file with "random" velocity initialization a couple of times, the trajectory will be exactly identical in all runs.

However, there are cases in which this behavior is not desired, *e. g.*, if you want to average a property over multiple trajectories of the same system. In these cases, call the `Randomize` command in the beginning of the input. If no argument is given, the random number generator is initialized with the current system time as a seed. MD runs started at different times will have different random velocities in the beginning. If you want more control over this process, you can also specify a positive integer number as argument, which is used as initial random seed. Simulations started with the same seed argument will have identical initial random velocities (if all other system parameters such as atom count, atom types, ... remain identical).

Without a call to `Randomize`, a seed of 1 is always used.

Restart

Mandatory Arguments:		—		
Optional Arguments:		*fname*	String	—
Modifiers:	`IfExists`	—		

Reads a restart file to continue a previous molecular dynamics run. Such a restart file is written after every simulation step, such that a crashed simulation may easily be recovered. The file name of the restart file may be given via *fname*; otherwise, it is deduced from the project's base name as `<basename>.mdrestart`.

If the `IfExists` modifier is specified, a restart is only performed if the restart file exists. The error and abort that would normally occur in case of a non-existent restart file are suppressed by this flag. This is useful in the first of a series of batch runs, where the restart file does not yet exist in the beginning.

Please note that the following quantities are stored to/loaded from restart files:

- Atom Positions
- Atom Velocities
- Thermostat internal state *(only for NHC)*
- Metadynamics hills and extended Lagrangian internal state
- Simulation step number and elapsed physical time

All other quantities (timestep, regions, thermostat, constraints, cells, etc.) are **not** restarted and need to be set in the input file, typically **before** the `Restart` command. It is safe to just call the `Restart` command immediately before the *Run* command.

Please see also the discussion on restarting simulations in [**Section** *1.3*](#moldyn:sec_restart).

Restraint

Mandatory Arguments:		*operation*	Keyword	{ `Add`, `Reset` }
			Keyword	`Colvar`
		colvar	Integer	—
Optional Arguments:		—		
Modifiers:	`Harmonic`	—		
	`Gaussian`	—		
	`Spring`	—		
	`Sigma`	*Gaussian Sigma*	Real	…
	`Height`	*Gaussian Height*	Real	kJ mol^{-1}
	`Target`	*target*	Real	…
	`Lower`	*lower wall*	Real	…
	`Upper`	*upper wall*	Real	…
	`Ramp`	*initial target*	Real	…
		final target	Real	…
	`Noprint`	—		

This command imposes restraints on collective variables ("Colvars") defined before via the *Manage_Colvars* command. As a first argument, it expects the kind of operation to perform, which can be `Add` and `Reset`. The second argument needs to be the keyword "`Colvar`", and the third argument is an integer number specifying the Colvar on which the operation shall be performed.

If the first argument is "`Add`", a new restraint is added to the specified Colvar. Note that an arbitrary number of restraints of different types can be active on a Colvar at the same time. The next argument after the Colvar number needs to be the type of the restraint. Currently, `Harmonic` and `Gaussian` are allowed. When adding harmonic restraints, the `Spring` modifier can be given, specifying the harmonic spring constant of the restraint in

kJ mol^{-1} unit^{-2}, where unit is the default physical unit of the Colvar (Angstrom for distances, degree for angles). If not specified, a value of 50 kJ mol^{-1} unit^{-1} is used. When adding Gaussian restraints, the `Height` and `Sigma` modifiers are allowed. The former sets the height of the Gaussian hill in kJ mol^{-1}, while the latter sets the width of the Gaussian function in physical Colvar units *(σ is the standard deviation, **not** the variance σ^2)*. The height can be either positive or negative, allowing for both Gaussian hills and Gaussian wells. If not specified, the default values of -10 kJ mol^{-1} for the height *(i. e., Gaussian well)* and 10 Colvar units *(e. g., Angstrom or degree)* for sigma are used.

The position of the new restraint is controlled via the `Target` modifier, which expects one real argument in Colvar units. If the restraint shall be an one-sided wall, the modifiers `Lower` and `Upper` can be used instead of `Target`. It is also possible to specify both `Lower` and `Upper` in order to define a lower and an upper wall at different positions in one command. If `Ramp` is given directly after `Target`, `Lower`, or `Upper`, a restraint with linearly moving target position over time is defined. `Ramp` expects two arguments, which are the initial restraint position, and the final restraint position after the next subsequent *Run* command.

The following example shows how to assign a harmonic two-sided restraint with different lower and upper wall parameters to a distance Colvar with number 7 that has been previously defined via the *Manage_Colvars* command:

```
Restraint Add Colvar 7 Harmonic Lower 400_pm Spring 50.0
Restraint Add Colvar 7 Harmonic Upper 800_pm Spring 80.0
```

By default, some additional data *(current position, potential energy, external force, internal force)* for each restraint is printed to a file with the name "`basename-restraints.csv`" in every MD step. This data can be used, *e. g.*, for thermodynamic integration. If a large number of restraints is defined, this can waste time and disk space. To switch this off for a restraint, specify the `Noprint` modifier when defining it.

If the first argument was "`Reset`", all restraints imposed on the specified Colvar are deleted. No further arguments or modifiers *(apart from the three mandatory arguments described above)* can be given.

Run

Mandatory Arguments:		*n*	Integer	—
Optional Arguments:		—		
Modifiers:	`StepLimit`	*value*	Real	[length]
	`CenterCOM`	—		

Performs a molecular dynamics run over n time steps with the current settings, applying the velocity Verlet algorithm to solve the equations of motion (see section *1.5.1*). You might want to call commands like *Timestep*, *Initvel*, *Thermostat*, and *Dump* before. Please note that only atoms within the `active` region will be subject to time integration. All other atoms will be skipped, and will therefore retain their initial positions.

The `StepLimit` modifier can be used to limit the maximum displacement of any atom in a MD time integration step. In addition to the displacement, also the velocities will be limited to a maximum of *value*$\cdot\Delta t$. This can help to stabilize the dynamics if the initial geometry is poor and large forces are acting (close atoms, etc.). The keyword expects one real argument in distance units. A reasonable choice would be 0.1 Å.

If the `CenterCOM` modifier is given, the center of mass (CoM) of the total system is kept fixed. Normally, the CoM should not drift anyway, because the velocity initialization is performed in a way which gives the CoM a zero initial velocity, and the conservation of momentum should keep it like that. However, numerical errors and massive *Thermostats* *(among other factors)* can break this momentum conservation, leading to a drift of the CoM over time. If this shall be avoided, specify this modifier.

If no call to *Initvel* occurred before this command, the atom velocities are initialized to zero. If no call to *Timestep* occurred before this command, a default time step of 0.5 fs is set.

You can cleanly end a MD run by creating an empty file with the name "`EXIT`" *(note the all-uppercase letters on case-sensitive file systems)*. On Unix operating systems such as GNU Linux, this can easily be achieved by the command "`touch EXIT`". will detect the file, abort the MD run, and delete the file. You will still get the remaining output *(such as the timing statistics)*, and you don't have to delete all the remaining "`.tmp`" files, which both would not be the case if you would have killed the process instead.

SCFLog

Mandatory Arguments:	*value*	Keyword	{ Discard, Last, Append, Each }
Optional Arguments:	—		
Modifiers: —			

Controls how/if the detailed output from the electron structure calculation (*i. e.*, integrals, scf, gradient, ...) will be written to log files. `Discard` completely discards the output. `Last` only keeps the last output for each program call (useful to read error message if simulation aborts). `Append` redirects all the output into one single log file ("basename.scf.log", "basename.int.log", "basename.grad.log", ...), appending each step at the end of the file. `Each` writes the output for each step and each program to different log files, which have the step number in their file names.

The amount of information which is printed to the SCF log file can be controlled by the standard ORCA print flags, such as "`%output PrintLevel Maxi end`". Note that by default, ORCA reduces the print level after the first SCF. Due to this, properties such as orbital energies and population analyses will only be printed once by default. If you want to keep the print level constant for subsequent SCF runs, disable this feature via "`%method ReducePrint false end`" in the ORCA input.

The default value is `Append`. Note that this can lead to large log files in long runs.

Screendump

Mandatory Arguments:	—
Optional Arguments:	—
Modifiers: —	

Prints the current state of the MD module (atom positions, velocities, potential and kinetic energy, cell properties, etc.) to the screen and log file in a well-defined and "grepable" format. This is mostly useful for unit testing, *e. g.*, to verify if the system state after a MD run equals the state obtained from some other ORCA binary distribution.

Thermostat

Mandatory Arguments:		*type*	Keyword	{ Berendsen, CSVR, NHC, None }
Optional Arguments:		*temperature*	Real	[temperature]
Modifiers:	Timecon	*tau*	Real	[time]
	Ramp	*target_temp*	Real	[temperature]
	Chain	*chain_length*	Integer	—
	MTS	*mts*	Integer	—
	Yoshida	*yoshida*	Integer	—
	Region	*region*	...	...
	Massive	—		

Changes the atom thermostat settings for subsequent simulation runs. "*Type*" sets the thermostat type. Currently, three thermostat types are implemented: Berendsen [92], Nosé–Hoover chains (NHC) [562, 563], and "Canonical Sampling through Velocity Rescaling" (CSVR) [129]. The very basic and robust Berendsen thermostat should only be used for early pre-equilibration runs, as it does **not** sample the canonical ensemble and leads to problems such as the flying ice cube effect. Both the NHC and the CSVR thermostats are very sophisticated, and correctly sample the canonical ensemble. One of these two should be used in all standard NVT simulations. Use `None` as type to disable the thermostat.

The optional *temperature* argument sets the target temperature to which the system is thermostated. If this argument is omitted, the temperature from the last call to the *Initvel* command is used (if no such call was invoked before, the simulation is aborted).

The `Timecon` modifier sets the coupling strength of the thermostat (large time constants correspond to weak coupling). The default value is 10 fs, which is a relatively strong coupling. For a production run, 100 fs would be appropriate. Values in the range of 10 . . . 100 fs are reasonable (see also section *1.5.3*).

If the `Ramp` modifier is used, a temperature ramp can be applied during a MD run. The final temperature at the end of the ramp has to be specified directly after the modifier. The initial temperature at the beginning of the ramp is taken from the *temperature* argument *(or from the last* `Initvel` *command if this argument is missing)*. The temperature ramp is applied only to the *Run* command which first follows the ramp definition. The slope of the ramp is chosen such that the final temperature is reached at the end of the run. Any subsequent *Run* command will simply use the final temperature for thermostating. To apply another temperature ramp, you need to explicitly define it again.

The `Chain`, `MTS`, and `Yoshida` modifiers only apply to NHC thermostats. They specify the chain length of the Nosé–Hoover chain (default: 3), the number of multiple time steps in which the thermostat integration is performed (default: 2), and the order of the Yoshida integrator used (default: 3, allowed: 1, 3, 5, 7), respectively. Normally, there is little need to modify one of these parameters. For more information, refer to the original publications [562, 563].

The `Massive` modifier activates *massive thermostating*, which means that each degree of freedom is assigned to an independent thermostat. This is useful for pre-equilibration runs (helps to reach energy equipartition) and should not be used during production runs, as it might heavily distort the dynamics. Note that massive thermostats also break the conservation of momentum (both linear and angular), so better specify the `CenterCOM` modifier for the *run* command if this is an issue. Please also note that massive NHC thermostats of large systems can be quite slow, because each NHC thermostat is a dynamical system on its own which needs to be time integrated.

With the `Region` modifier, the thermostat can be attached to a specific region (*i. e.*, subset of atoms). This modifier expects one argument, which is either the name of a pre-defined region or the number of a user-defined region *(see above)*. If not specified, the thermostat acts on the whole system. Multiple thermostats for multiple regions can be active at the same time, but each region can have only one attached thermostat at a time (re-defining will overwrite the thermostat settings).

The command "`Thermostat None`" will remove all thermostats from all regions. If you want to disable a thermostat for a specific region only, use "`Thermostat None Region r`", where `r` is the name or number of the region.

Please note that all three implemented thermostat types will show no effect (or unexpected effects) if the system's temperature is close to 0 K, as they all work by multiplying the velocities with a *(more or less complicated)* factor.

Timestep

Mandatory Arguments:	*dt*	Real	[time]
Optional Arguments:	—		
Modifiers: —			

Sets the simulation time step Δt used to integrate the equations of motion for all following runs to dt. If your system contains hydrogen atoms, a time step not above 0.5 fs is recommended. If only heavier atoms are present, a larger time step may be chosen. A good estimate for a time step that still allows for an accurate simulation is $\Delta t = \sqrt{m} \cdot 0.5\text{fs}$, where m is the mass of the lightest atom in the system (in a.m.u.). This is one reason why some scientists perform simulations with fully deuterated compounds: It allows to increase the time step by a factor of ≈ 1.4 :-)

If this command is not invoked before a *Run* call, a default time step of 0.5 fs will be set before starting the run.

7.48.8 Scientific Background

In this section, some of the methods and algorithms used within ORCA's MD module are described in some more depth, with a focus on the scientific background.

Time Integration and Equations of Motion

The central concept of molecular dynamics simulations is to solve Newton's equations of motion (at least as long as the atom cores are treated classically). These read

$$\ddot{x}_i\left(t\right) = \frac{F_i\Big(\vec{x}\left(t\right)\Big)}{m_i}, \quad i = 1 \ldots N,$$

where $x_i\left(t\right)$ denotes the position of the i-th degree of freedom at time t, m the corresponding mass, and F_i the force acting upon this degree of freedom. As the force may depend on all positions, this is a coupled system of N ordinary differential equations (ODEs). In the general case, it is not possible to obtain an analytical solution of this system, and therefore numerical solution methods are applied. These are almost always based on discretizing the time variable and approximately solving the system by taking finite time steps.

Of all different methods to numerically solve coupled systems of ODEs, the *symplectic integration schemes* for Hamiltonian systems attained special attention in the field of molecular dynamics. They possess a very good conservation of energy. In contrast to many other methods, they show a reasonable behavior when investigating the long-term evolution of chaotic Hamiltonian systems (like, *e. g.*, MD simulations). Three popular such symplectic integration schemes are the *Leapfrog* algorithm, the *Verlet* method, and the *Velocity Verlet* integrator. Despite their different names, they are very similar. It can be easily seen that the Verlet and Velocity Verlet methods are algebraically equivalent (by eliminating the velocities from the Velocity Verlet algorithm), and it can be shown that, eventually, all three methods are identical.[3] All three methods are explicit integration methods with a global error of order 2, and therefore one order better than the semi-implicit Euler method, which is also a symplectic integration scheme. As the Velocity Verlet algorithm is the only of these three methods which yields velocities and positions at the same point in time, many popular molecular dynamics packages (CP2k, CPMD, LAMMPS) use this scheme. For the same reasons, the ORCA MD module uses the Velocity Verlet algorithm as time integration method.

The general equations of the Velocity Verlet scheme read

$$\begin{aligned}
\vec{x}\left(t + \Delta t\right) &= \vec{x}\left(t\right) + \vec{v}\left(t\right)\Delta t + \frac{1}{2}\vec{a}\left(t\right)\Delta t^2, \\
\vec{v}\left(t + \Delta t\right) &= \vec{v}\left(r\right) + \frac{\vec{a}\left(t\right) + \vec{a}\left(t + \Delta t\right)}{2}\Delta t.
\end{aligned}$$

By inserting

$$\vec{a}_i\left(t\right) = \frac{\vec{F}_i\left(t\right)}{m_i}, \quad i = 1 \ldots N,$$

one arrives at the two-step method

$$\begin{aligned}
\vec{x}_i\left(t + \Delta t\right) &= \vec{x}_i\left(t\right) + \vec{v}_i\left(t\right)\Delta t + \frac{\vec{F}_i\left(t\right)}{2m_i}\Delta t^2, \qquad i = 1 \ldots N, \\
\vec{v}_i\left(t + \Delta t\right) &= \vec{v}_i\left(r\right) + \frac{\vec{F}_i\left(t\right) + \vec{F}_i\left(t + \Delta t\right)}{2m_i}\Delta t, \qquad i = 1 \ldots N,
\end{aligned}$$

which is implemented in ORCA's MD module.

[3] Hairer, Lubich, Wanner, "Geometric Numerical Integration", Springer 2006.

Velocity Initialization

In the beginning of a MD simulation, it is often the case that only the initial positions of the atoms are known, but not the velocities. As MD simulations are performed at some finite temperature, it is a good idea to initialize the velocities in a way such that the desired simulation temperature is already present in the beginning. In statistical mechanics, it is often assumed that the velocity distribution of atoms is given by a Maxwell–Boltzmann distribution (which is strictly only the case in idealized gases). Therefore, it is a reasonable choice to initialize the atom's velocities according to the Maxwell–Boltzmann equation in the beginning of a MD simulation. The goal is to find an initial velocity distribution in which each degree of freedom possesses a similar amount of energy, such that the equipartition theorem is approximately fulfilled.

The scalar Maxwell–Boltzmann velocity distribution (leaving out the normalization factor) at temperature T is given by

$$f\left(v\right) = v^2 \exp\Big(-\frac{mv^2}{2k_B T}\Big).$$

To initialize the particle's velocities such that this distribution function is fulfilled, one starts with a series of normal-distributed random numbers with mean 0 and variance 1, denoted by $\mathcal{N}\left(0,1\right)$. The Cartesian velocity components for each atom are then computed by

$$v_{i,\alpha} := \sqrt{\frac{k_B T}{m_i}}\,\mathcal{N}\left(0,1\right), \quad \alpha \in \{x,y,z\}, \;\; i = 1 \ldots N.$$

As the C++98 standard does not offer a platform-independent way of obtaining normal-distributed random numbers, these are internally computed from uniformly distributed random numbers by applying the *Box–Muller transform* [115]: Assuming that u_1 and u_2 are two uniformly distributed random numbers from the interval $[0,1]$, the equations

$$\begin{aligned} z_1 &:= \sqrt{-2\log\left(u_1\right)}\cos\left(2\pi u_2\right), \\ z_2 &:= \sqrt{-2\log\left(u_1\right)}\sin\left(2\pi u_2\right) \end{aligned}$$

yield two new random numbers z_1 and z_2 which obey a normal distribution with mean 0 and variance 1.

After the velocities have been initialized, the total linear momentum of the system will probably have some finite value other than zero. As the linear momentum is (approximately) conserved within a molecular dynamics simulation, this would result in the system drifting away into one direction during the course of the simulation, which is probably not desired. Therefore, the total momentum is explicitly set to zero after the Maxwell–Boltzmann initialization:

$$\begin{aligned} \vec{P}_{\mathrm{tot}} &:= \sum_{i=1}^{N} m_i \vec{v}_{i,\mathrm{old}}, \\ \vec{v}_{i,\mathrm{new}} &:= \vec{v}_{i,\mathrm{old}} - \frac{\vec{P}_{\mathrm{tot}}}{m_i N}, \quad i = 1 \ldots N. \end{aligned}$$

This, of course, might change the initial temperature. Therefore, a final step is performed, in which all velocity vectors are multiplied with a factor that is determined such that the initial temperature exactly matches the target value.

Thermostats

After the initial velocities have been initialized to some finite temperature, it might be assumed that one can simply start the time integration of the dynamical system (equivalent to the *NVE ensemble*), and the starting temperature would be approximately preserved. In a real system, however, there are (at least) two reasons why the temperature will strongly deviate from the initial value already after a few steps. First, the initial velocity distribution only considers the kinetic energy of the particles, but some amount of energy will be exchanged with the potential energy contribution (*e. g.*, bond stretching) immediately, altering the temperature. Secondly, the numerical errors introduced due to the finite time step (and in case of *ab initio* MD, also due to the approximate forces) will lead to a drift in energy and therefore in temperature. To counter these effects, it is often desirable to have a temperature control during the course of the simulation (which then runs in the *NVT ensemble*), which is called a thermostat.

There exist many different kinds of thermostats, ranging from simple expressions up to highly complex dynamical systems on their own. But all of them share a common issue: If the thermostat is coupled only weakly to the system, the temperature will change anyway. However, if the thermostat is coupled more strongly to the system (*i. e.*, intervenes stronger), then the dynamics of the simulation will change, no longer resembling the undisturbed original dynamics which one wants to investigate. Therefore, it is always a tradeoff between temperature stability and disturbed dynamics to decide how strong a thermostat should be coupled to the system.

In ORCA, currently three thermostats are implemented: The Berendsen thermostat [92], the Nosé–Hoover chain thermostat (NHC) [562, 563], and the "Canonical Sampling through Velocity Rescaling" thermostat (CSVR) [129].

Berendsen Thermostat

The Berendsen thermostat [92] is similar to the simple velocity rescaling scheme, but enhanced by a time constant τ to control the coupling strength. Let T_0 be the desired target temperature and T the current temperature of the system. Then the temperature gradient caused by the thermostat can be expressed as

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}.$$

Considering the fact that discrete time steps Δt are used, the correction factor for the velocities in each time step is determined by

$$f := \sqrt{1 + \frac{\Delta t\,(T_0 - T)}{T\tau}}$$

The new velocities are then easily obtained as

$$\vec{v}_{i,\mathrm{new}} := f \cdot \vec{v}_{i,\mathrm{old}}, \quad i = 1 \ldots N.$$

Let's consider some special cases. If $\tau = \Delta t$, the whole temperature deviation from T_0 is corrected immediately, such that the temperature is always exactly kept at the target value. This is identical to simple velocity rescaling (without any time constant), which is known to work poorly for most systems (a single harmonic oscillator would, *e. g.*, simply explode). With a larger time constant $\tau > \Delta T$, the coupling strength is reduced, leading to reasonable results. Typically, a value of τ in the range of $20 \ldots 200 \cdot \Delta T$ will be applied. For $\tau \to \infty$, the coupling strength goes to zero, such that the thermostat is no longer active. Values of $\tau < \Delta T$ are not allowed.

From the formula, it becomes clear that a Berendsen thermostat will have no effect if the system has a temperature of 0 K (or in the "massive" case: if the considered degree of freedom has 0 K), because it is based on multiplying the velocities by a factor to modify the temperature. Therefore, this type of thermostat can't be used to heat a system up starting from 0 K.

Constraints

Unlike restraints, constraints are geometric relations which are strictly enforced at every time (*i. e.*, they do not fluctuate around their target value). Many molecular dynamics techniques make use of geometric constraints (*e. g.*, to keep water molecules rigid, or to fix some reaction coordinate). Standard BOMD describes the nuclei as point charges in space, such that the motion of the atoms is governed by the laws of classical mechanics. Systems in classical mechanics can be described by the Lagrange formalism, which contains a well established sub-formalism for holonomic constraints, namely the method of Lagrange multipliers.

However, molecular dynamics discretizes time to solve the equations of motions with finite time steps, often using a Verlet integrator. With discretized time, it is slightly more involved to enforce and keep exact constraints. Within the last decades, algorithms have been developed to do so. One famous among them is the SHAKE algorithm. However, it comes with the disadvantage of only enforcing the constraints in the positions, not in the velocities. This may lead to problems such as artificially high temperature values due to "hidden" velocities along the constrained directions. An extension of SHAKE which also enforces the constraints for the velocities is the RATTLE algorithm, which is implemented in the AIMD module of ORCA.

The RATTLE scheme is a generalization of the Velocity Verlet integrator to allow for constraints. This means that RATTLE is not applied in addition to the Velocity Verlet integrator, but replaces it. In case of no active constraints,

both methods are identical. A system of coupled constraints cannot be solved exactly in one step, and RATTLE uses an iterative approach to enforce all constraints simultaneously. This is often a matter of concern with respect to performance. However, in AIMD, the energy and gradient calculations typically take seconds or even minutes per step, such that the additional computation time for iteratively solving the constraints can be totally neglected.

As an iterative procedure, RATTLE is not able to give exact solutions, but only converged up to a given tolerance. In the ORCA MD module, the tolerance is currently set to 10^{-2} pm for distances, and 10^{-4} degree for angles and dihedral angles. This tolerance is typically reached within a few dozen iterations. In some cases, it might happen that the RATTLE iterations do not converge to the required tolerance. This is typically the case if the set of constraints is over-determined or contradictory.

The mathematical and technical details of RATTLE are not described here, they can be found in the literature. The general concept of RATTLE was suggested by Andersen [35]. The original article only covered distance constraints. A follow-up work describes how to handle any holonomic constraints, in particular how to constrain angles and dihedral angles [483]. The Wilson vectors (*i. e.*, derivatives of angles and dihedral angles with respect to Cartesian atom positions) are taken from Wilson's original work [889].

7.49 Fast Multipole Method

The Fast Multipole Method (FMM) algorithm was proposed in the 1980s[317], to reduce the computation time for two-centre interactions in large systems, by moving from a quadratic relationship (O(N^2)) to a quasi-linear relationship (O($Nlog(N)$)) of the computation time with the number N of particles (atoms, point charges, etc.). This algorithm is particularly useful for long-range interactions that are difficult to ignore, such as the Coulombic interaction (1/r) which is fundamental to any system but grows quadratically with the number of centers.

The FMM is used in ORCA for accelerating QMMM calculations in the scope of electrostatic embedding: FMM-QMMM (cf. *Embedding Types*).

7.49.1 The Octree hierarchy

There exists a lot of detailed and pedagogical literature on the subject (we recommend for instance [49] and [377]). In the following we will only describe the main parameters.

The idea of the algorithm is to divide space to handle differently short- and long-range interactions. The latter will be approximated (cf. *Approximation of the Far Field interactions*). To do so, the whole system is placed in a cubic box which is iteratively divided in 8 children boxes (forming levels L: $\mathrm{L} = 0 \rightarrow \mathrm{L} = \mathrm{L}_{\mathrm{max}}$) so as to form a structure called an Octree (see Fig. Fig. 7.62).

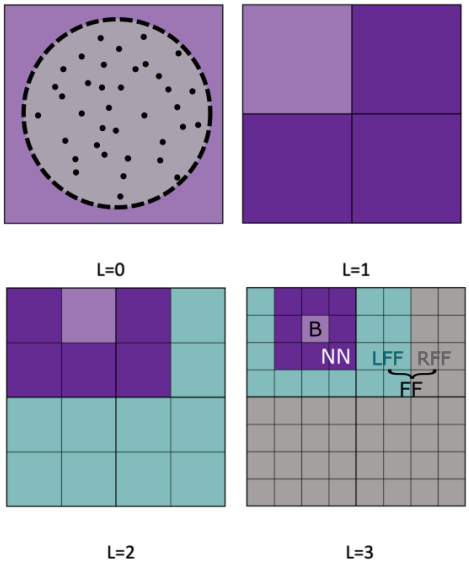


Fig. 7.62: Division of the system into boxes, schematic representation in 2D.

At the deepest level, $L = L_{max}$, for every box, we can define a layer of nearest neighbours boxes (NN) which will be responsible for the Near-Field (NF, i.e.short-range). The electrostatic potential due to the rest of the boxes is define as the Far-Field (FF, i.e. long-range).

$$\forall B \ , \ \mathbf{V}_B = \mathbf{V}_B^{\mathbf{NF}} + \mathbf{V}_B^{\mathbf{FF}}$$

As visible on Figure Fig. 7.62, one can divide the FF area between a Local Far Field (LFF, green) area and a Remote Far Field area (RFF, grey), so that a recursive scheme between levels appears:

$$\begin{aligned} \forall B \ \ \mathbf{V}_B^{\mathbf{FF}} &= \mathbf{V}_B^{\mathbf{LFF}} + \mathbf{V}_B^{\mathbf{RFF}} \\ &= \mathbf{V}_B^{\mathbf{LFF}} + \mathbf{W}^T_{Parent(B)\to B}\mathbf{V}^{\mathbf{FF}}_{Parent(B)} \end{aligned}$$

The LFF boxes are the boxes in the NF of the parent but not in the NF of the children, that represents a maximum of 189 boxes. The RFF potential is due to boxes which actually represent the FF area of the parent box of B. This means that one needs to calculate only the LFF term at every level, the rest of the potential will be inherited from the parent box.

The number of levels, L_{max}, can be setup in the input directly (FMMQMMM_Levels), or will be deduced from the provided box dimension (FMMQMMM_BoxDimInp, dimension of the box at L_{max}). The second option (default) is recommended, with a box dimension around 9.0 Bohr (FMMQMMMM_BoxDimInp 9.0). If FMMQMMM_DoBoxDimOpt option is turned to TRUE (default), the box dimension will be reduced as much as possible while keeping the same value of L_{max}, this to ensure the algorithm works optimally regarding both accuracy and efficiency.

```
%method
  DOFMMQMMM             True  #turn ON the use of the FMM
  FMMQMMM_BoxDimInp      9.0  #Higher boundary to set up the dimension (in Bohr!)
  FMMQMMM_DoBoxDimOpt   True  #Optimize the box dimension
end
```

The higher L_{max}, the bigger the number of boxes: at $L = 2$ one has 64 boxes only but for $L = 6$ the code generates more than 200,000 boxes (8^6). For $L_{max} \geq 7$, more than 2 millions boxes are generated, this can start leading to memory issues, so that the user should ensure enough memory is available (cf. *Global memory use*).

7.49.2 Approximation of the Far Field interactions

In a space of origin O(0,0,0), let P ($\mathbf{r}_P$) be the center of a charge distribution (e.g. a Gaussian overlap or a distribution of point charges). Let's consider a point A in the vicinity of P, such that A = (q_A; $\mathbf{r}_A$), with q_A the charge of a point charge or $q_A = -1$ if we consider a Gaussian overlap. In that case, the charge would indeed be the one of the electron situated at a position $\mathbf{r}_{AP} = \mathbf{r}_A - \mathbf{r}_P$ from the center of the Gaussian overlap. Similarly, let Q ($\mathbf{r}_Q$) be the center of a charge distribution and B, a point in the vicinity of Q, such that B = (q_B; $\mathbf{r}_B$). We note $\mathbf{r}_{QP} = \mathbf{r}_Q - \mathbf{r}_P$.

The Coulomb interaction between A and B can be expressed as:

$$U_{AB} = f\Big(\frac{1}{|\mathbf{r}_a - \mathbf{r}_b|}\Big),$$

The name of the algorithm comes from the way the FF interactions are evaluated. It is done through a mutlipole expansion of the $\frac{1}{|\mathbf{r}_a - \mathbf{r}_b|}$ term, leading to the following expression:

$$\frac{1}{|\mathbf{r}_A - \mathbf{r}_B|} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sum_{j=0}^{\infty}\sum_{k=-j}^{j}(-1)^j R_{l,m}(\mathbf{r}_{AP}) I_{l+j,m+k}(\mathbf{r}_{QP}) R_{j,k}(\mathbf{r}_{BQ}),$$

This series converges for well separated centers only, that is why a NF layer is required. $R_{l,m}$ and $I_{l,m}$ are regular and irregular solid scaled harmonics [706]:

$$R_{l,m}(\mathbf{r}) = \frac{1}{\sqrt{(l-m)!(l+m)!}}\mathbf{r}^l\sqrt{\frac{4\pi}{2l+1}}Y_{l,m}(\mathbf{r})$$

$$I_{l,m}(\mathbf{r}) = \sqrt{(l-m)!(l+m)!}\frac{1}{\mathbf{r}^{l+1}}\sqrt{\frac{4\pi}{2l+1}}Y_{l,m}(\mathbf{r})$$

$Y_{l,m}$ are spherical harmonics of degree l and order m. For details on the derivation of this equation see [49].

We can now introduce the multipole expansion centered in P ($\mathbf{r}_P$), of a charge distribution $\mathbf{Q}(\mathbf{r}, P)$, as:

$$\mathbf{Q}(\mathbf{r}, P) = f\Big(\sum_{l=0}^{\infty}\sum_{m=-l}^{l} R_{l,m}(\mathbf{r} - \mathbf{r}_P)\Big)$$

In practise:

$$\mathbf{Q}(\mathbf{r}, P) \approx f\Big(\sum_{l=0}^{MAM}\sum_{m=-l}^{l} R_{l,m}(\mathbf{r} - \mathbf{r}_P)\Big)$$

Q is a vector, with elements $Q_{l,m}$ defined by the pair of values (l,m). In the case of a point charge A (q_A,**r**) we have:

$$Q^A_{l,m}(\mathbf{r}, P) = q_A \times R_{l,m}(\mathbf{r} - \mathbf{r}_P),$$

while for a Gaussian overlap distribution $\langle\mu|\nu\rangle$ it becomes:

$$\begin{aligned} Q^{\mu\nu}_{l,m}(\mathbf{r}, P) &= \langle\mu|\, R_{l,m}(\mathbf{r} - \mathbf{r}_P)\,|\nu\rangle \\ &= \int \mu(\mathbf{r'}) R_{l,m}(\mathbf{r} - \mathbf{r}_P)\nu(\mathbf{r'})d\mathbf{r'} \end{aligned}$$

So that we can approximate the electrostatic interaction between a Gaussian overlap $\langle\mu|\nu\rangle$ with center in $\mathbf{r}_B$ and a point charge A (q_A,$\mathbf{r}_A$) in the FF, through the following expansion:

$$\frac{\int q_A\mu(\mathbf{r})\nu(\mathbf{r})d\mathbf{r}}{|\mathbf{r}_A - \mathbf{r}_B|} \approx \mathbf{Q}^{\mu\nu}(\mathbf{r}_b, P')\mathbf{I}(\mathbf{r}_{PP'})\mathbf{Q}^A(\mathbf{r}_A, P),$$

with P and P' the respective center of the multipole expansions, and I the interaction matrix.

The truncation parameter of the expansion, MAM, is the Maximum Angular Momentum of the underlying solid scaled harmonics. In our setup, it is the same value for the expansion of the point charges in the embedding or the Gaussian overlaps in the QM part. It can be setup through the keyword FMMQMMM_MAM. We recommend FMMQMMM_MAM = 20 to ensure a good accuracy, in the order of 0.0001-0,0001 Ha. With a lower value of MAM=15, one can usually expect a mHa (0.63 kcal.mol^{-1}) precision. However it is system dependent, so that we recommend you to start with MAM=20 and see how it behaves if you decrease it to a value of 15. The available values for MAM are in the range 1-25.

```
%method
  DOFMMQMMM             True  #turn ON the use of the FMM
  FMMQMMM_MAM             20   #Default Maximum Angular Momentum
end
```

7.49.3 The "Very" Fast Multipole Method (VFMM)

Due to the recursive scheme between levels, the FF felt in the center of a box is built by only calculating interactions with the LFF boxes at every level. The evaluation of V^{LFF} becomes the bottleneck of the algorithm. To accelerate it, an option has been implemented in which different MAM truncation parameters will be used for all the (>189) boxes in the LFF area. This option FMMQMMMM_DoVFMM is turned ON by default, but should be switched OFF whenever the MAM is smaller than 15, that would impact to much the accuracy without improving much the efficiency.

```
%method
  DOFMMQMMM             True  #turn ON the use of the FMM
  FMMQMMM_DoVFMM        True  #Use the Very Fast Multipole Method
end
```

7.49.4 Recommended input

Whenever there is an electrostatic embedding, for systems containing more than 10,000 point charges (ECM) or MM atoms (QMMM), it is recommended to turn on the FMM in order to accelerate the calculation of the electrostatic potential. However, when using heavy parallelization with more than 24 processors, the impact of enabling the Fast Multipole Method (FMM) may be negligible for small embedding size.

Caution

FMM-QMMM does not replace the QMMM keyword.

The two elements one can play on are the truncation parameter of the expansion, MAM, and the box dimension at the deepest level (L_{max}). Recommended/default parameters can be called in the keyword line directly using FMM-QMMM or by setting the parameters accordingly in the %method block:

```
#keyword line
! FMM-QMMM

#OR through method block

%method
  DOFMMQMMM             True  #turn ON the use of the FMM
  USEFMMQMMM            True  #turn ON the use of the FMM
  ###parameters:
  FMMQMMM_DoVFMM        True  #Use the Very Fast Multipole Method
  FMMQMMM_MAM             20  #Maximum Angular Momentum
  FMMQMMM_Levels           0  #Number of levels set to 0, will be set automatically by box dim
  FMMQMMM_BoxDimInp      9.0  #Higher boundary to set up the dimension (in Bohr!)
  FMMQMMM_DoBoxDimOpt   True  #Optimize the box dimension
end
```

The parameters used by the algorithm are printed in the output by default, we recomend you to check them in your first calculations:

```
----------------------
SHARK INTEGRAL PACKAGE
----------------------
```

(continues on next page)

(continued from previous page)

```
[...]

Use FMM for one center integrals with PCs

- - - - - - - FMM PARAMETERS- - - - - - -
num PCs              75894
VFMM                  USED
MAM                   20
Input box dim        9.000
Refined box dim      5.197
Tree depth/levels      6
- - - - - - - - - - - - - - - - - - - - -
```

This can be turned off by modifying the relative printing options.

```
%method
 FMMQMMM_Printing 1 # remove FMM PARMETERS printing, default value is  2
end
```

7.49.5 Some examples

Call the recommanded FMM parameters through the keyword line:

```
!QMMM wB97X-D3BJ RIJCOSX FMM-QMMM

%qmmm
 ORCAFFFilename "file.prms"
 Embedding Electrostatic
 ChargeAlteration CS
 QMAtoms {2016:2026} end
end


*pdbfile 0 1 file.pdb
```

Change the MAM from 20 (value when one uses the keyword line) to 15, and use 5 levels:

```
!QMMM wB97X-D3BJ RIJCOSX

%method
  DOFMMQMMM             True  #turn ON the use of the FMM
  USEFMMQMMM            True  #turn ON the use of the FMM
  ###parameters:
  FMMQMMM_DoVFMM        True  #Use the Very Fast Multipole Method
  FMMQMMM_MAM             15  #Maximum Angular Momentum
  FMMQMMM_Levels           5  #number of levels
end

%qmmm
 ORCAFFFilename "file.prms"
 Embedding Electrostatic
 ChargeAlteration CS
 QMAtoms {2016:2026} end
end


*pdbfile 0 1 file.pdb
```

Specify the box dimension without optimizing it:

```
!QMMM wB97X-D3BJ RIJCOSX

%method
  DOFMMQMMM             True  #turn ON the use of the FMM
  USEFMMQMMM            True  #turn ON the use of the FMM
  ###parameters:
  FMMQMMM_DoVFMM        True  #Use the Very Fast Multipole Method
  FMMQMMM_MAM             15  #Maximum Angular Momentum
  FMMQMMM_Levels           0  #number of levels
  FMMQMMM_BoxDimInp      9.0  #Higher boundary to set up the dimension (in Bohr!)
  FMMQMMM_DoBoxDimOpt   False #Does not optimize the box dimension
end

%qmmm
 ORCAFFFilename "file.prms"
 Embedding Electrostatic
 ChargeAlteration CS
 QMAtoms {2016:2026} end
end


*pdbfile 0 1 file.pdb
```

Specifying the box dimension and ask to optimize it:

```
!QMMM wB97X-D3BJ RIJCOSX

%method
  DOFMMQMMM             True  #turn ON the use of the FMM
  USEFMMQMMM            True  #turn ON the use of the FMM
  ###parameters:
  FMMQMMM_DoVFMM        True  #Use the Very Fast Multipole Method
  FMMQMMM_MAM             15  #Maximum Angular Momentum
  FMMQMMM_Levels           0  #number of levels
  FMMQMMM_BoxDimInp      9.0  #Higher boundary to set up the dimension (in Bohr!)
  FMMQMMM_DoBoxDimOpt   True  #Optimize the box dimension
end

%qmmm
 ORCAFFFilename "file.prms"
 Embedding Electrostatic
 ChargeAlteration CS
 QMAtoms {2016:2026} end
end


*pdbfile 0 1 file.pdb
```

7.50 Implicit Solvation Models

Implicit solvation models play an important role in quantum chemistry. Without resorting to placing multiple solvation shells of solvent molecules implicit solvent models are able to mimic the effect of a specific solvent on the solute.

The implicit solvent models available in ORCA are

1. C-PCM[76] : The Conductor-like Continuum Polarization Model
2. SMD[558] : The Solvation Model based on Density
3. OpenCOSMO-RS[293] : Interface to the open source implementation of the COSMO-RS model

4. ALPB/ddCOSMO/CPCM-X : The solvation models available in XTB

Regarding C-PCM and SMD, they are natively implemented in ORCA. Points 1, 2 and 3 are covered in this section. For more information on point 4, check Section *ONIOM Methods*.

7.50.1 The Conductor-like Polarizable Continuum Model (C-PCM)

The conductor-like polarizable continuum model (C-PCM) is an implementation of the conductor-like apparent surface charge methods. In these models the solute is placed in a cavity of roughly molecular shape. The solvent reaction field is described by apparent polarization charges on the cavity surface, which are in turn determined by the solute. These charges can be treated as punctual (point charges) or be modelled as spherical Gaussians [904]. The cavity in ORCA is constructed differently depending on how the charges are treated. In the case of using point charges, the cavity is generated through the GEPOL[649, 650, 651] algorithm, either as solvent-excluding surface (SES), or solvent-accessible surface (SAS). When Gaussian charges are considered, the user can choose between a scaled vdW surface or the GEPOL SES, and the charge positions are determined following a Lebedev quadrature approach. This scheme is known as Gaussian Charge Scheme[287] and more details on how to use it are given in Section *Use of the Gaussian Charge Scheme*.

The ORCA C-PCM implementation closely follows the C-PCM[76] paper. The molecular Hamiltonian of the isolated system is perturbed by the solvent:

$$\hat{H} = \hat{H}^0 + \hat{V}$$

where $\hat{H}^0$ is the Hamiltonian of the isolated molecule, whereas $\hat{V}$ describes the solute – solvent interactions. The SCF procedure leads to the variational minimization of the free energy of the solute, G:

$$G = \left\langle \Psi \left| \hat{H}^0 \right| \Psi \right\rangle + \frac{1}{2} \left\langle \Psi \left| \hat{V} \right| \Psi \right\rangle$$

Using the conductor-like boundary condition the electrostatic potential can be determined by

$$V(\vec{r}) + \sum_{i}^{N_q} V_{q_i}(\vec{r}) = 0$$

where V and V_{q_i} are the electrostatic potential due to the solute and to the polarization charges, $\vec{r}$ is a point on the cavity surface, and N_q is the total number of solvation charges. The vector of polarization charge can then be determined by

$$\mathbf{AQ} = -\mathbf{V} \tag{7.314}$$

where the vector $\mathbf{V}$ contains the electrostatic potential due to the solute at the position of the charges. The elements of the matrix $\mathbf{A}$ have a different functional form depending on how the charges are treated. If we use point charges:

$$A_{ii} = 1.07\sqrt{\frac{4\pi}{S_i}}$$

$$A_{ij} = \frac{1}{r_{ij}} \tag{7.315}$$

in which S_i is the area of the surface element i, and $r_{ij} = |\vec{r}_i - \vec{r}_j|$. When Gaussian charges are considered:

$$A_{ii} = \frac{\zeta_i \sqrt{2/\pi}}{F_i}$$

$$A_{ij} = \frac{\mathrm{erf}\left(\zeta_{ij} r_{ij}\right)}{r_{ij}}$$

Here, ζ_i is the exponent of the Gaussian charge i (i belongs to sphere I). This quantity is calculated as $\zeta_i = \zeta/(R_I\sqrt{w_i})$, where R_I is the radius of sphere I, w_i is the weight of the Lebedev point i, and ζ is a width parameter optimized for each particular Lebedev grid [904]. On the other hand, $\zeta_{ij} = \zeta_i\zeta_j/\sqrt{\zeta_i^2 + \zeta_j^2}$. The function F_i,

known as switching function, measures the contribution of the Gaussian charge i to the solvation energy. This function is calculated as

$$F_i = \prod_{J,i\notin J}^{\text{atoms}} g(\vec{r}_i, \vec{R}_J)$$

where $g(\vec{r}_i, \vec{R}_J)$ is the elementary switching function. In ORCA we use the improved Switching/Gaussian (ISWIG) function for $g(\vec{r}_i, \vec{R}_J)$ proposed in ref. [494]:

$$g(\vec{r}_i, \vec{R}_J) = 1 - \frac{1}{2}\left\{\text{erf}\left[\zeta_i\left(R_J - r_{iJ}\right)\right] + \text{erf}\left[\zeta_i\left(R_J + r_{iJ}\right)\right]\right\}$$

If $g(\vec{r}_i, \vec{R}_J) < 10^{-7}$ the value of g is set equal to 0.

If we consider a solvent with a dielectric constant ε, eq. (7.314) reads as

$$\mathbf{AQ} = -f(\varepsilon)\mathbf{V} \tag{7.316}$$

where $f(\varepsilon) = (\varepsilon - 1)/(\varepsilon + x)$ is a scaling function, and x is in the range 0-2. In C-PCM x is equal to 0.

The C-PCM model can be used via

```
! CPCM(solvent)
```

where `solvent` is one of the available solvents in Table 7.29

Table 7.29: List of available solvents for the different implicit solvation methods in ORCA. The data for the dielectric constant used within C-PCM is that at 293.15,[368] except for ammonia, which has a boiling point of 239.81 K. For the rest of solvation models, see the corresponding sources.[558][814]

Solvent	C-PCM	SMD	COSMO-RS	ALPB	ddCOSMO	CPCM-X
1,1,1-trichloroethane	X	X				
1,1,2-trichloroethane	X	X				
1,2,4-trimethylbenzene	X	X				X
1,2-dibromoethane	X	X	X			
1,2-dichloroethane	X	X	X			X
1,2-ethanediol	X	X				
1,4-dioxane / dioxane	X	X	X	X	X	
1-bromo-2-methylpropane	X	X				
1-bromooctane / bromooctane	X	X	X			
1-bromopentane	X	X				
1-bromopropane	X	X				
1-butanol / butanol	X	X	X			X
1-chlorohexane / chlorohexane	X	X	X			X
1-chloropentane	X	X				
1-chloropropane	X	X				
1-decanol / decanol	X	X	X			X
1-fluorooctane	X	X	X			X
1-heptanol / heptanol	X	X	X			X
1-hexanol / hexanol	X	X	X			X
1-hexene	X	X				
1-hexyne	X	X				
1-iodobutane	X	X				
1-iodohexadecane / hexadecyliodide	X	X	X			X
1-iodopentane	X	X				
1-iodopropane	X	X				
1-nitropropane	X	X				
1-nonanol / nonanol	X	X	X			X

continues on next page

Table 7.29 – continued from previous page

Solvent	C-PCM	SMD	COSMO-RS	ALPB	ddCOSMO	CPCM-X
1-octanol / octanol	X	X	X	X	X	X
1-pentanol / pentanol	X	X	X			X
1-pentene	X	X				
1-propanol / propanol	X	X	X			X
2,2,2-trifluoroethanol	X	X				
2,2,4-trimethylpentane / isooctane	X	X	X			X
2,4-dimethylpentane	X	X				
2,4-dimethylpyridine	X	X				
2,6-dimethylpyridine	X	X	X			X
2-bromopropane	X	X				
2-butanol / secbutanol	X	X	X			X
2-chlorobutane	X	X				
2-heptanone	X	X				
2-hexanone	X	X				
2-methoxyethanol / methoxyethanol	X	X	X			X
2-methyl-1-propanol / isobutanol	X	X	X			X
2-methyl-2-propanol	X	X				
2-methylpentane	X	X				
2-methylpyridine / 2methylpyridine	X	X	X			X
2-nitropropane	X	X				
2-octanone	X	X				
2-pentanone	X	X				
2-propanol / isopropanol	X	X	X			X
2-propen-1-ol	X	X				
e-2-pentene	X	X				
3-methylpyridine	X	X				
3-pentanone	X	X				
4-heptanone	X	X				
4-methyl-2-pentanone / 4methyl2pentanone	X	X	X			X
4-methylpyridine	X	X				
5-nonanone	X	X				
acetic acid / aceticacid	X	X	X			X
acetone	X	X	X	X	X	
acetonitrile / mecn / ch3cn	X	X	X	X	X	X
acetophenone	X	X	X			X
ammonia	X		X			
aniline	X	X	X	X	X	X
anisole	X	X	X			X
benzaldehyde	X	X	X	X	X	
benzene	X	X	X	X	X	X
benzonitrile	X	X	X			X
benzyl alcohol / benzylalcohol	X	X	X			X
bromobenzene	X	X	X			X
bromoethane	X	X	X			X
bromoform	X	X	X			X
butanal	X	X				
butanoic acid	X	X				
butanone	X	X	X			X
butanonitrile	X	X				
butyl ethanoate / butyl acetate / butylacetate	X	X	X			X
butylamine	X	X				
n-butylbenzene / butylbenzene	X	X	X			X
sec-butylbenzene / secbutylbenzene	X	X	X			X
tert-butylbenzene / tbutylbenzene	X	X	X			X
carbon disulfide / carbondisulfide / cs2	X	X	X	X	X	X

continues on next page

Table 7.29 – continued from previous page

Solvent	C-PCM	SMD	COSMO-RS	ALPB	ddCOSMO	CPCM-X
carbon tetrachloride / ccl4	X	X	X			X
chlorobenzene	X	X	X			X
chloroform / chcl3	X	X	X	X	X	X
a-chlorotoluene	X	X				
o-chlorotoluene	X	X				
conductor	X				X	
m-cresol / mcresol	X	X	X			X
o-cresol	X	X				
cyclohexane	X	X	X			X
cyclohexanone	X	X	X			X
cyclopentane	X	X				
cyclopentanol	X	X				
cyclopentanone	X	X				
decalin	X	X	X			X
cis-decalin	X	X				
n-decane / decane	X	X	X			X
dibromomethane	X	X				X
dibutylether	X	X	X			X
o-dichlorobenzene / odichlorobenzene	X	X	X			X
e-1,2-dichloroethene	X	X				
z-1,2-dichloroethene	X	X				
dichloromethane / ch2cl2 / dcm	X	X	X	X	X	X
diethyl ether / diethylether	X	X	X	X	X	X
diethyl sulfide	X	X				
diethylamine	X	X				
diiodomethane	X	X				
diisopropyl ether / diisopropylether	X	X	X			X
cis-1,2-dimethylcyclohexane	X	X				
dimethyl disulfide	X	X				
n,n-dimethylacetamide / dimethylacetamide	X	X	X			X
n,n-dimethylformamide / dimethylformamide / dmf	X	X	X	X	X	X
dimethylsulfoxide / dmso	X	X	X	X	X	X
diphenylether	X	X	X			X
dipropylamine	X	X				
n-dodecane / dodecane	X	X	X			X
ethanethiol	X	X				
ethanol	X	X	X	X	X	X
ethyl acetate / ethylacetate / ethanoate	X	X	X	X	X	X
ethyl methanoate	X	X				
ethyl phenyl ether / ethoxybenzene	X	X	X			X
ethylbenzene	X	X	X			X
fluorobenzene	X	X	X			X
formamide	X	X				
formic acid	X	X				
furan / furane			X	X	X	
n-heptane / heptane	X	X	X			X
n-hexadecane / hexadecane	X	X	X	X	X	X
n-hexane / hexane	X	X	X	X	X	X
hexanoic acid	X	X				
iodobenzene	X	X	X			X
iodoethane	X	X				
iodomethane	X	X				
isopropylbenzene	X	X	X			X
p-isopropyltoluene / isopropyltoluene	X	X				X
mesitylene	X	X	X			X

continues on next page

Table 7.29 – continued from previous page

Solvent	C-PCM	SMD	COSMO-RS	ALPB	ddCOSMO	CPCM-X
methanol	X	X	X	X	X	X
methyl benzoate	X	X				
methyl butanoate	X	X				
methyl ethanoate	X	X				
methyl methanoate	X	X				
methyl propanoate	X	X				
n-methylaniline	X	X				
methylcyclohexane	X	X				
n-methylformamide / methylformamide	X	X	X			X
nitrobenzene / phno2	X	X	X			X
nitroethane	X	X	X			X
nitromethane / meno2	X	X	X	X	X	X
o-nitrotoluene / onitrotoluene	X	X				X
n-nonane / nonane	X	X	X			X
n-octane / octane	X	X	X			X
n-pentadecane / pentadecane	X	X	X			X
octanol(wet) / wetoctanol / woctanol				X	X	
pentanal	X	X				
n-pentane / pentane	X	X	X			X
pentanoic acid	X	X				
pentyl ethanoate	X	X				
pentylamine	X	X				
perfluorobenzene / hexafluorobenzene	X	X	X			X
phenol	X		X	X	X	
propanal	X	X				
propanoic acid	X	X				
propanonitrile	X	X				
propyl ethanoate	X	X				
propylamine	X	X				
pyridine	X	X	X			X
tetrachloroethene / c2cl4	X	X	X			X
tetrahydrofuran / thf	X	X	X	X	X	X
tetrahydrothiophene-s,s-dioxide /	X	X				X
/ tetrahydrothiophenedioxide / sulfolane	X	X				X
tetralin	X	X	X			X
thiophene	X	X				
thiophenol	X	X				
toluene	X	X	X	X	X	X
trans-decalin	X	X				
tributylphosphate	X	X	X			X
trichloroethene	X	X				
triethylamine	X	X	X			X
n-undecane / undecane	X	X	X			X
water / h2o	X	X	X	X	X	X
xylene	X	X				X
m-xylene	X	X				
o-xylene	X	X				
p-xylene	X	X				

The parameters can be more accurately defined using the `%cpcm` block input. The available options are as follows

```
%cpcm  epsilon             80.0    # Dielectric constant
       refrac               1.0    # Refractive index
       rsolv                1.3    # Solvent probe radius
       rmin                 0.5    # Minimal GEPOL sphere radius
```

(continues on next page)

(continued from previous page)

```
      pmin                 0.1    # Minimal distance between two surface points
      fepstype            cpcm    # Epsilon function type: cpcm, cosmo
      xfeps                0.0    # X parameter for the feps scaling function
      surfacetype vdw_gaussian    # Cavity surface: gepol_ses, gepol_sas
                                                    vdw_gaussian, gepol_ses_gaussian
      ndiv                   5    # Maximum depth for recursive sphere generation
      num_leb              302    # Lebedev points for the Gaussian charge scheme
      radius[N]            1.3    # Atomic radius for atomic number N in Angstrom
      AtomRadii(N,1.4)            # Atomic radius for the Nth atom in Angstrom
      scale_gauss          1.2    # Scaling factor for the atomic radii in the
                                    Gaussian charge scheme
      cut_area             0.0    # Cutoff for the area of a surface segment in a.u.
                                    Only valid for the Gaussian charge scheme
      cut_swf             1e-7    # Cutoff for the switching function
                                    Only valid for the Gaussian charge scheme
      thresh_h             5.0    # Threshold for the charge density on a hydrogen
                                    atom in charges/Å^2 (isodensity scheme)
      thresh_noth          5.0    # Threshold for the charge density on non-hydrogen
                                    atoms in charges/Å^2 (isodensity scheme)
      CPCMccm                0    # Coupled-cluster/C-PCM scheme
      cds_cpcm               0    # Use of the GVDW_nel or GSES_nel scheme
      end
```

Regarding the parameters shown above, some of them can be just used within a given type of cavity surface and charge scheme, as it is mentioned in subsections *Use of the Gaussian Charge Scheme*, and *Use of the Point Charge Scheme*. ORCA supports two types of solvation charge schemes: (i) the point charge scheme, (ii) the Gaussian charge scheme. The default solvation scheme from ORCA 5.0 on is the Gaussian charge scheme with a vdW-type cavity (”`surfacetype vdw_gaussian`”). Older ORCA versions considered “`surfacetype gepol_ses`” and the point charge scheme as defaults. So, if one wants to reproduce the results obtained with ORCA versions older than ORCA 5.0, please use the tag “`surfacetype gepol_ses`” in the “`%cpcm`” block.

The different charge schemes are described in the subsections *Use of the Gaussian Charge Scheme*, and *Use of the Point Charge Scheme*. The availability of the analytical gradient and Hessian for the different combinations of charge scheme and surface is shown in Table Table 7.30.

Table 7.30: Available type of gradients and Hessians within the C-PCM in ORCA. The tag “YES” means that the feature is implemented, and “NO” that it isn’t. For clarity, we denote by “*” the default scheme in ORCA (`surfacetype vdw_gaussian`).

`surfacetype`	Charge type	Gradient		Hessian	
		Analytical	Numerical	Analytical	Numerical
`gepol_ses`	Point	Yes	Yes	Yes	Yes
`gepol_sas`	Point	Yes	Yes	Yes	Yes
`vdw_gaussian^*`	Gaussian	Yes	Yes	Yes	Yes
`gepol_ses_gaussian`	Gaussian	No	Yes	No	Yes

Note: If the user wants to turn off the C-PCM, one has to write the following tag in the simple input:

```
! NOCPCM
```

This is needed, for instance, in the context of concatenated calculations using the `$new_job` feature, where the previous calculation involved solvation, but one wants to turn it off for the next one.

Use of the Gaussian Charge Scheme

The Gaussian charge scheme avoids the Coulomb singularity present in conventional point charge surface element models. This approach, when applied together with a switching function, results in a smooth solvation potential and, more importantly, on smooth derivatives of this quantity with respect to external perturbations. Then, it is highly recommended to adopt this approach within the C-PCM. The Gaussian charge scheme can be used with two types of solute cavity surfaces: (1) a scaled vdW surface, (2) a solvent-excluded surface (SES). To assign the radii for the different atoms we follow the scheme proposed in ref. [494]. That is, we use Bondi radii [114] for all elements, except for hydrogen where we adopt 1.1 Å. For 16 of the main-group elements in the periodic table, where Bondi's radii are not defined, we adopt the radii proposed in ref. [552] by Mantina et al. This is the case for elements: Be, B, Al, Ca, Ge, Rb, Sr, Sb, Cs, Ba, Bi, Po, At, Rn, Fr, Ra. For the elements that are not covered neither by Bondi nor by Mantina, we consider a radius of 2 Å.

- **Scaled vdW cavity**

The Gaussian charge scheme with a scaled vdW-type cavity is now the default in ORCA, so one just needs to add the C-PCM tag in the `%cpcm` block in the input file. That would correspond to (although the user does not need to write that, as it is internally processed by ORCA):

```
%cpcm
surfacetype vdw_gaussian
end
```

In this case, the radius R_I of atom I for the scaled vdW cavity is calculated as

$$R_I = f_{scal} R_I^{\mathrm{vdW}}$$

where R_I^{vdW} is the vdW radius of atom I and f_{scal} is a scaling factor. This parameter is by default equal to 1.2, as suggested in ref. [494]. However, the user can modify its value through the `scale_gauss` tag in the `%cpcm` block in the input file. The number of C-PCM charges per atom is such that we have an approximate density of 5.0 charges/Å^2 on the surface of the cavity. This number corresponds approximately to 110 points on a hydrogen atom ($110/(4\pi \times 1.32^2) \approx 5.0\text{Å}^{-2}$). ORCA will choose then different levels of discretization (Lebedev grids) depending on the radii of the atoms. This scheme is called isodensity scheme. The threshold for the number of charges per unit of area on hydrogens and non-hydrogen atoms can be specified by the user via `thresh_h` and `thresh_noth`, respectively. Alternatively, the user can request a fixed number of Lebedev points per sphere in the cavity (independent of the radius of the sphere). This can be done via the `num_leb` tag. This parameter can adopt the following values: 50, 110, 194, 302, 434, 590, 770, 974, and 1202. ORCA versions older than ORCA 6.0 use this last scheme.

The analytical gradient, as well as the analytical Hessian are available for this solvation method.

- **Solvent-excluded surface**

The GEPOL-generated SES can be used together with the Gaussian charge scheme. In this case, the Gaussian charges are not only placed on the surface of the atomic spheres but also on the surface of the new spheres generated through the GEPOL algorithm. To use this approach we should modify the ORCA input file as follows

```
%cpcm
surfacetype gepol_ses_gaussian
end
```

The SES is in general recommended when the solute is explicitly solvated by few solvation layers. In this case, the additional GEPOL spheres prevent the solvent to fill the space between explicit solvent molecules or between those molecules and the solute. The radius of the solvent sphere that rolls over the solute vdW-type surface to generate the SES is controlled by the parameter "`rsolv`". This radius has a default value of 1.30 Å, but the user can change it by specifying another value for "`rsolv`" in the `%cpcm` block. The minimum radius for an added GEPOL sphere is controlled by "`rmin`".

Neither the analytical gradient nor the analytical Hessian are available for this strategy. They should be computed numerically. Due to the interdependency between GEPOL spheres and the atoms present in our system, the analytical gradient computed using the SES does not converge as smooth as compared to that using the vdW-type cavity

(see ref. [287]) and can lead to wrong minima. The Hessian is affected in the same way. Then, ORCA 6.0 does not support anymore the analytical gradient/Hessian for such surface to prevent inaccurate results.

Use of the Point Charge Scheme

Within this scheme, the solvation charges are treated as punctual and no switching function is used to accept/discard charges in the intersection between the spheres that form the solute cavity. These two facts lead to a discontinuous potential energy surface (numerical instabilities in the SCF energy and its derivatives). Then, the point charge scheme is not recommended within the C-PCM. If the user still wants to use this charge scheme it can be used together with two different type of surfaces for the solute cavity: (1) the SES, and (2) the SAS.

- **Solvent-excluded surface**

In the same way as for the Gaussian charge scheme with the SES, this surface is generated for the point charge scheme through the GEPOL algorithm. However, while for the case of `surfacetype gepol_ses_gaussian`, the surface is discretized using Lebedev-type grids, in the point charge scheme, the surface is divided in spherical triangles called "tesserae". The level of tesselation is controlled by the tag `ndiv`. The radii used for the solute atoms in order to construct the GEPOL cavity are those optimized for the COSMO-RS model[445] for H, C, N, O, F, S, Cl, Br, and I, while for the rest of elements scaled Bondi radii are used.

In order to use the point charge scheme with the SES, we should add the following tag in the `%cpcm` block:

```
%cpcm
surfacetype gepol_ses
end
```

- **Solvent-accessible surface**

Here, the surface of the solute cavity is generated by following the center of a sphere with radius "`rsolv`" (representing the solvent molecule) rolling over the surface of the vdW surface of the solute. The discretization of the resulting surface is done via tesserae, as done for the SES. In order to use the point charge scheme together with the SAS, one should add the following tag in the `%cpcm` block:

```
%cpcm
surfacetype gepol_sas
end
```

- **How to circumvent the numerical instabilities in the point charge scheme**

Numerical instabilities are implicit in the point charge scheme due to the "punctual" nature of the charges and to the fact that no switching function is considered in the intersection between the spheres that form the solute cavity. At the same time, there is no way to predict, in advance if there will be discontinuities in the SCF energy and/or in its derivatives. However, if the user still wants to use this charge scheme, no matter with which type of surface (SES or SAS), there are two things that one can try to minimize the aforementioned problems:

- **Change `ndiv`** : The parameter "`ndiv`" controls the number of triangles per sphere in the solute cavity. By changing the number of triangles, it also changes the number of triangles in the intersection between spheres and, then, this can solve those situations where two point charges get too close making the elements A_{ij} in eq (7.315) to diverge (and the solvation charges to suddenly have very large values). However, this strategy can be a solution for a particular case, but will never ensure that for the same system with another geometry the discontinuities in the SCF energy do not show (as it does not prevent point charges to be too close from each other).
- **Increase `pmin`** : This parameter removes those charges that are at a distance lower than `pmin` from each other. The default value for `pmin` is of 0.1 a.u. (≈ 0.0529 Å). Then, by incrasing `pmin` one removes all pairs of charges that are too close from each other in the intersection between the spheres. **Note:** Be careful when increasing `pmin`. Although this prevents sudden jumps in the SCF energy, it can lead to "biasing" the solute-solvent interaction, as one is removing a significant number of charges that represent the effect of the solvent.

Calculation of the free energy of solvation within the C-PCM

The solvation free energy, ΔG_{solv}, is defined as the free energy of transfer of a solute from the gas phase to the condensed phase. This quantity can be written as (see eq 27 in ref. [287])

$$\Delta G_{solv} = E_{solv}(\vec{R}_l, \vec{R}_v) + \Delta G_{el} + \Delta G_{cav} + \Delta G_{disp} \tag{7.317}$$

The first term, $E_{solv}(\vec{R}_v, \vec{R}_l)$, corresponds to the difference between the liquid-phase expectation value of the gas-phase Hamiltonian, $E(\vec{R}_l)$, and the gas-phase potential energy surface $E(\vec{R}_v)$

$$E_{solv}(\vec{R}_v, \vec{R}_l) = E(\vec{R}_l) - E(\vec{R}_v)$$

The second term in eq (7.317) accounts for the electronic-polarization contribution to ΔG_{solv} and is calculated from the charges spread over the surface of the solute cavity (vdW-type or the SES). Finally, ΔG_{cav} is the cavitation energy, that is, the reversible work required to create a cavity inside the bulk of the solvent in order to accommodate the solute, while ΔG_{disp} accounts for the changes in the dispersion energy occuring when solvating the solute. The sum of these last two terms correspond to the non-electrostatic contribution, ΔG_{nel}, to ΔG_{solv}

$$\Delta G_{nel} = \Delta G_{cav} + \Delta G_{disp}$$

In ORCA, if a system is solvated within the C-PCM (defining `!CPCM(solvent name)` in the input file, within either the point charge scheme or the Gaussian charge scheme) there is no "non-electrostatic solvation contribution" neither to the SCF energy, nor to its derivatives. In order to have a rough idea of the value of ΔG_{nel}, one can estimate this quantity through empirical equations available in the literature obtained from experimental data. For instance, ΔG_{nel} can be estimated from the free energy of hydration for linear-chain alkanes[852]

$$\Delta G_{\text{nel}} = 1.321 + 0.0067639 \times \text{SASA} \tag{7.318}$$

where SASA is the solvent-accessible surface area. However, although eq (7.318) may give a good estimation of ΔG_{nel} for organic molecules in water, it is a bad approximation for non-electrostatic solvation effects when considering solvents different than water (as well as for many solutes in water). There are two alternatives to this approximation that one can adopt in order to include non-electrostatic solvation effects in ORCA calculations:

1. The use of the SMD model in combination with the C-PCM (see section *The SMD Solvation Model*).
2. The use of Gaussian charges to compute the electrostatic contribution to ΔG_{solv} together with an equation for ΔG_{nel} calculated from the contribution of the solute atoms to the SASA.[287]

Focusing on the second strategy, the cavitation energy, ΔG_{cav} is calculated through the equation proposed by Pierotti and Claverie.[180, 680] In this case, ΔG_{cav} is expanded in powers of the ratio $\overline{\mathbf{R}}$ between the radius of the solute sphere R_I and that of the solvent spheres R_S. Then, the cavitation energy for the solute sphere centered at atom I reads as,

$$\Delta G_{cav,I} = -\ln(1-y) + \left(\frac{3y}{1-y}\right)\overline{\mathbf{R}} + \left[\frac{3y}{1-y} + \frac{9}{2}\left(\frac{y}{1-y}\right)^2\right]\overline{\mathbf{R}}^2 + \frac{yP}{\rho_S k_B T}\overline{\mathbf{R}}^3$$

with $y = 8\pi\rho_S R_S^3/6$, being ρ_S the density of the solvent, $\overline{\mathbf{R}} = R_I/R_S$, T the temperature and P the pressure. If we consider the solvent-accessible surface (SAS), then, the total ΔG_{cav} is the sum of the $\Delta G_{cav,I}$ weighted by a factor that depends on the exposed SASA_I of sphere I

$$\Delta G_{cav} = \sum_{i=1}^{N_{\text{sph}}} \frac{\text{SASA}_I}{4\pi R_I^2}\Delta G_{cav,I} = \sum_{i=1}^{N_{el}} \frac{\text{SASA}_i}{4\pi R_I^2}\Delta G_{cav,i} \tag{7.319}$$

Here, the summation over the total number of spheres (N_{sph}) that conform the SAS is replaced by a summation over the total number of surface elements (N_{el}) in which the SAS is divided into.

With respect to ΔG_{disp}, this quantity is assumed to depend linearly on the contribution of each surface element to the SASA

$$\Delta G_{disp} = \sum_{i=1}^{N_{\text{el}}} \sigma_I \text{SASA}_i \tag{7.320}$$

Here, the factor σ_I is the atomic surface tension of the sphere I to which the surface element i belongs to.

This strategy, that considers eqs (7.319) and (7.320) to compute ΔG_{nel} assumes the use of Gaussian charges for the calculation of ΔG_{el} in eq (7.317). When a vdW-type cavity is used to account for electrostatic solvation effects, then the solvation approach is called GVDW_nel, while when a SES-type cavity is used we refer to the solvation approach as GSES_nel ("G" for Gaussian charges and "nel" because of the inclusion of non-electrostatic solvation effects). Details on the GVDW_nel and GSES_nel models can be found in ref. [287].

The GVDW_nel and GSES_nel models have been parametrized for solutes containing H, C, N and O atoms in the following solvents: benzene, chloroform, cyclohexane, octanol, toluene and water. The parameters for the σ_I atomic surface tensions in eq (7.320) together with the radius of the solvent molecules R_S used to generate the SAS are provided in ref. [287]. The use of the GVDW_nel and GSES_nel is controlled by the string `cds_cpcm` to add in the `%cpcm` block. This flag should be equal to 2 within these two models (by default, when we use C-PCM together with the Gaussian Charge Scheme without non-electrostatic solvation effects, `cds_cpcm` is equal to 0). In order to calculate ΔG_{solv} for the GVDW_nel scheme one has to do the following steps:

1. Optimize the geometry of the solute (anisole in this case) in vacuum. For a DFT calculation at the B3LYP/def2-TZVP level of theory (with RIJCOSX), our input file would read like:

```
! B3LYP def2-TZVP D3BJ tightopt

%maxcore 4000

* xyz 0 1
C   0.66588156410505    1.19833878437922    -0.02647672759883
C   1.95470174000504    0.69483871574862    -0.01927812843331
C   2.17893173780041   -0.68097630146538     0.00933914777704
C   1.09270919891774   -1.54252176637089     0.03141990372830
C  -0.21073409652496   -1.05179390481975     0.02471602253843
C  -0.42517201252444    0.32533458015547    -0.00499572635290
H   0.47565631823749    2.26336213776820    -0.05012843684387
H   2.79210971512409    1.38085300778245    -0.03749594945361
H   3.18771596563920   -1.07149825338966     0.01433233618265
H   1.25061585302875   -2.61357041387417     0.05495238973649
H  -1.03877793934640   -1.74453554370491     0.04346162651229
O  -1.65366310927299    0.90979347913477    -0.01693978807213
C  -2.79791686738707    0.07370821312998     0.01117694976431
H  -2.83142307710305   -0.58944196707467    -0.85867919998175
H  -3.65619850989594    0.74176224021803    -0.00958531946332
H  -2.82883648080292   -0.53125300761732     0.92288089996024
*
```

2. Use the resulting structure (`anisole.xyz`) as input structure for the subsequent geometry optimization in solution. In this case, we consider water as the solvent. The input file looks like:

```
! B3LYP def2-TZVP D3BJ tightopt CPCM(Water)

%cpcm
cds_cpcm 2
end

%maxcore 4000

* xyzfile 0 1 anisole.xyz
```

3. Search for the "`FINAL SINGLE POINT ENERGY`" in the output file for the solvated system (we call it E_s):

```
-------------------------   --------------------
FINAL SINGLE POINT ENERGY      -346.723542928530
-------------------------   --------------------
```

4. Search for the "`FINAL SINGLE POINT ENERGY`" in the output file for the system in vacuum (we call it E_0):

```
-------------------------   --------------------
FINAL SINGLE POINT ENERGY      -346.719657762242
-------------------------   --------------------
```

5. Calculate $\Delta G_{solv} = E_s - E_0 = -0.003885$ a.u. If we convert it to kcal/mol and round it to two decimal digits, we have $\Delta G_{solv} = -2.44$ kcal/mol. This quantity is in very good agreement with its experimental counterpart (-2.45 kcal/mol[558]).

6. In case the user is interested in the value of ΔG_{nel}, this quantity is printed in the "`CPCM Solvation Model Properties`" block:

```
CPCM Solvation Model Properties:
Surface-charge          :  -0.02580685887229
Corrected charge        :   0.00000000000000
Outlying charge corr.   :   0.00005046198558 Eh    0.00137 eV
Free-energy (cav+disp)  :   0.00272597237192 Eh    0.07418 eV
```

Here, ΔG_{nel} corresponds to "`Free-energy (cav+disp)` " and is calculated through eqs (7.319) and (7.320). The term called "`Surface-charge`" corresponds to the solute net charge that lies outside the C-PCM cavity. It is basically the sum of the C-PCM charges (without the scaling by $f(\varepsilon)$). The effect of this excess of charge on the energy and the C-PCM charges themselves can be corrected by the so-called outlying charge correction. In ORCA, this correction is calculated through a Lagrangian-based algorithm,[705] but only printed for information purposes. That is, the outlying charge effect is neither added to the SCF energy nor to its derivatives. In any case, the corrected total charge is printed in "`Corrected charge`", while the correction term for the energy is printed in "`Outlying charge corr.`". To further help the user, ORCA also prints a file with extension `.cpcm_corr`, where the corrected C-PCM charges are provided.

If we want to use, instead, the GSES_nel model, one just needs to add `surfacetype gepol_ses_gaussian` in the `%cpcm` block.

Note: The analytical Hessian is not available for the GVDW_nel and GSES_nel models as the second derivative of ΔG_{nel} with respect to nuclear displacements is not implemented.

7.50.2 The Conductor-like Screening Solvation Model (COSMO)

Note

The COSMO solvation model has been removed from ORCA v4.0.0 !!!

Please use the C-PCM solvation model in combination with the COSMO epsilon function if required!
As a short form to use the C-PCM model with the COSMO epsilon function, you can specify the solvent via `!CPCMC(solvent)` .

7.50.3 The SMD Solvation Model

The SMD solvation model has been proposed by the Cramer and Truhlar groups,[558] and is based on the quantum mechanical charge density of a solute molecule interacting with a continuum description of the solvent. In the model the full solute electron density is used without defining partial atomic charges and the solvent is not represented explicitly but rather as a dielectric medium with the surface tension at the solute–solvent boundary. SMD is a universal solvation model, in the sense that it is applicable to any charged or uncharged solute in any solvent or liquid medium for which a few key descriptors are known. In particular, these descriptors are the dielectric constant, refractive index, bulk surface tension, and acidity and basicity parameters. Neglecting the concentration contribution, the model separates the observable solvation free energy into two main components,

$$\Delta G_{\mathrm{S}} = \Delta G_{\mathrm{ENP}} + \Delta G_{\mathrm{CDS}}. \tag{7.321}$$

In ORCA, the first component is the bulk electrostatic contribution arising from a self-consistent reaction field treatment that involves the electrostatic interaction using the Conductor-like Polarizable Continuum Model (C-PCM). However, the radii are set to "intrinsic atomic Coulomb radii". The second component, called the cavity-dispersion solvent-structure (CDS) term, is the contribution resulting from short-range interactions between the solute and solvent molecules in the first solvation shell. This contribution is a sum of terms that are proportional (with geometry-dependent proportionality constants called atomic surface tensions) to the solvent-accessible surface areas of the individual atoms of the solute. The CDS contribution to the free energy of solvation is given by

$$\Delta G_{\mathrm{CDS}} = \sum_{k}^{\mathrm{atoms}} \sigma_k A_k(\mathbf{R}, R_{Z_k} + r_s) + \sigma^{[\mathrm{M}]} \sum_{k}^{\mathrm{atoms}} A_k(\mathbf{R}, R_{Z_k} + r_s),$$

where σ_k and $\sigma^{[\mathrm{M}]}$ are the atomic surface tension of atom k and the molecular surface tension, respectively, and A_k is the solvent-accessible surface area (SASA). The SASA depends on the geometry $\mathbf{R}$, the set R_{Z_k} of all atomic van der Waals radii, and the solvent radius r_s, which is added to each of the atomic van der Waals radii. In the program Bondi radii are used for CDS contribution.

More details can be found in the original paper of Marenich et al. [558], which should be cited in publications using results of SMD calculations.

SMD can be employed in single point calculations and geometry optimizations, using single-determinant SCF (HF and DFT) and CASSCF methods. In post-SCF methods the result is corrected in the reference wave function. The SMD solvation model is invoked in the input file via

```
! SMD(solvent)
```

where `solvent` is one of the 179 solvents in the SMD library (see Table Table 7.29). Alternatively, one can request the SMD model via the `%cpcm` block by writing:

```
%cpcm   smd true              # turn on SMD
        SMDsolvent "solvent" # specify the name of solvent from the list
 end
```

Independently on the way the user invokes the SMD model, ORCA automatically sets a number of default SMD parameters for the chosen solvent. If required, the user can also manually specify the solvent descriptors used in an SMD calculation in the `%cpcm` block.

```
%cpcm   soln               # index of refraction at optical frequencies at 293 K
        soln25             # index of refraction at optical frequencies at 298 K
        sola               # Abraham's hydrogen bond acidity
        solb               # Abraham's hydrogen bond basicity
        solg               # relative macroscopic surface tension
        solc               # aromaticity, fraction of non-hydrogenic solvent atoms
                           # that are aromatic carbon atoms
        solh               # electronegative halogenicity, fraction of non-hydrogenic
                           # solvent atoms that are F, Cl, or Br
 end
```

Let's consider the following input for a water molecule solvated by water,

```
! B3LYP def2-SVP tightscf SMD(water)

* xyz 0 1
O  -0.00000018976103    0.00606010894837    0.00000000004527
H   0.76098169249695   -0.58891312953082   -0.00000000000022
H  -0.76098151333900   -0.58891299029372   -0.00000000000022
*
```

Before the SCF part starts, the program prints the SMD information. This part reads as:

```
--------------------
CPCM SOLVATION MODEL
--------------------
CPCM parameters:
  Epsilon                                     ...      78.3550
  Refrac                                      ...       1.3328
  Rsolv                                       ...       1.3000
  Surface type                                ... GAUSSIAN VDW
  Discretization scheme                       ... Constant charge density
    Threshold for H atoms                     ...       5.0000 (charges/Ang^2)
    Threshold for non-H atoms                 ...       5.0000 (charges/Ang^2)
  Epsilon function type                       ...         CPCM
Solvent:                                      ... WATER
SMD-CDS solvent descriptors:
  Soln                                        ...       1.3328
  Soln25                                      ...       1.3323
  Sola                                        ...       0.0000
  Solb                                        ...       0.0000
  Solg                                        ...       0.0000
  Solc                                        ...       0.0000
  Solh                                        ...       0.0000
Radii:
  Scheme                                      ... Element-dependent radii
  Radius for O  used is    2.8724 Bohr (=   1.5200 Ang.)
  Radius for H  used is    2.2677 Bohr (=   1.2000 Ang.)
Calculating surface                           ...        done! (  0.0s)
Cavity surface points                         ...          244
Cavity Volume                                 ...     130.0616
Cavity Surface-area                           ...     129.3354
Calculating surface distance matrix           ...        done! (  0.0s)
Performing Cholesky decomposition & store     ...        done! (  0.0s)
Overall time for CPCM initialization          ...                0.0s
```

After the SCF is converged, the SMD contribution to the total energy is printed (this term is labelled as "SMD CDS" term).

```
               *****************************************************
               *                        SUCCESS                    *
               *           SCF CONVERGED AFTER  10 CYCLES          *
               *****************************************************

Recomputing exchange energy using gridx3       ... done (    0.056 sec)
Old exchange energy                            :     -1.791736873 Eh
New exchange energy                            :     -1.791694663 Eh
Exchange energy change after final integration :      0.000042211 Eh
Total energy after final integration           :    -76.335230273 Eh

SMD CDS free energy correction energy :                1.44927     Kcal/mol
Total Energy after SMD CDS correction =               -76.332920707 Eh
             **** ENERGY FILE WAS UPDATED (water.en.tmp) ****

----------------
TOTAL SCF ENERGY
----------------

Total Energy       :        -76.33292070695327 Eh           -2077.12437 eV

Components:
Nuclear Repulsion  :          9.11286213981824 Eh             247.97359 eV
Electronic Energy  :        -85.43200537714571 Eh           -2324.72305 eV
One Electron Energy:       -123.06209110061025 Eh           -3348.68974 eV
Two Electron Energy:         37.63008572346453 Eh            1023.96669 eV
```

(continues on next page)

(continued from previous page)

```
CPCM Dielectric    :         -0.01612924659271 Eh             -0.43890 eV
SMD CDS (Gcds)     :          0.00230956608299 Eh              0.06285 eV
```

Notes:

- If one is interested in the calculation of free energies of solvation using the SMD model, one just needs to compute ΔG_{S} according to eq (7.321). However, here one should take into account that the SMD model considers the same concentration (1 mol/L) in both the gaseous and solution phases. Then, if a gas-phase standard state of 1 atm is considered, and we want to compare the calculated ΔG_{S} with its experimental counterpart, a concentration term equal to 1.89 kcal/mol has to be added to the calculated ΔG_{S}.

7.50.4 Dynamic Radii Adjustment for Continuum Solvation (DRACO)

The DRACO scheme is an approach that improves the performance of implicit solvation models, in particular the accuracy of the calculated solvation free energies.[688] It is based on a dynamic scaling of the original static radii used to describe the atoms/spheres that define the cavity of the solute.

In this approach, the original radii r_i of an atom i is scaled according to

$$r_{i\mathrm{scale}} = f_{i\mathrm{scale}} r_i \tag{7.322}$$

with the scaling factor $f_{i\mathrm{scale}}$ determined from

$$f_{i\mathrm{scale}} = \mathrm{erf}(a_Z(q_{\mathrm{eff},i} - b_Z)) + 1 \tag{7.323}$$

Here, $q_{\mathrm{eff},i}$ corresponds to the effective partial charge and is defined as

$$q_{\mathrm{eff},i} = q_i + k_Z q_i \mathrm{CN}_i \tag{7.324}$$

where q_i is the atomic partial charge. The parameters a_Z, b_Z and k_Z are element-specific parameters, and CN_i is the fractional coordination number.

For oxygen atoms, the $f_{i\mathrm{scale}}$ is corrected via a term that depends on the Abraham's hydrogen bond acidity (α):

$$f^{\mathrm{O}}_{i\mathrm{scale}} = f_{i\mathrm{scale}} + c_{\mathrm{O}}(0.43 - \alpha) \tag{7.325}$$

with c_{O} a parameter that is different for pure C-PCM or SMD. The correction in eq (7.325) is only applied for solvents with $\alpha < 0.43$.

DRACO is parametrized for C-PCM and SMD for the following solvents: acetonitrile, DMSO, methanol, and water. The element-specific parameters in eqs (7.323),(7.324), and (7.325) are available in https://github.com/grimme-lab/DRACO.

The use of DRACO with C-PCM or SMD is triggered via the following tags in the simple input:

```
! CPCM(solvent) DRACO
```

or

```
! SMD(solvent) DRACO
```

Alternatively, it can also be requested in the `%cpcm` block:

```
%cpcm
DRACO true
end
```

In ORCA, the default scheme to calculate the partial charges in eq (7.324) is the electronegativity-equilibration (EEQ) charge model (D4 case),[132]. However, one can also request the Charge Extended Hückel (CEH) model.[541] The charge scheme within DRACO is controlled by the following tag in the `%cpcm` block:

```
%cpcm
draco_charges ceh  # default = eeq
end
```

If CEH charges are requested, ORCA needs to run an XTB calculation to generate them. Here, one should have at least the version 6.7.1 of XTB (older versions will not generate the CEH charges, see section *Program Components*). Regarding the fractional coordination number in eq (7.324), it is calculated as described in GFN2-XTB calculations.[70]

Note: DRACO can only be used, for the moment, in single-point energy calculations.

7.50.5 OpenCOSMO-RS

ORCA is interfaced to openCOSMO-RS,[293] an open source implementation of the COSMO-RS model.[447, 448] This model is widely used in both academia and industry to predict fluid phase thermodynamics.

The main idea behind COSMO-RS is that the interaction between molecules in the liquid phase can de depicted as an ensemble of interacting surface segments. The properties on the surface segments are calculated via QM calculations of the solute and the solvent in a perfect conductor ($\epsilon = \infty$), and the interaction free energies between segments are functions of a set of descriptors. Among them, the most relevant one is the screening charge density (σ).

A complete description of the COSMO-RS model used in ORCA is provided in ref [293], and the code is available in https://github.com/TUHH-TVT. Since openCOSMO-RS has been parametrized for ORCA 6.0, the executable is shipped together with the ORCA 6.0 binaries and called `openCOSMORS`. ORCA will then call internally this executable whenever it is needed.

COSMO-RS calculations within ORCA are a special type of calculation. By requesting COSMO-RS, ORCA runs a set of single-point-energy calculations for both the solute and the solvent and then calls the openCOSMO-RS executable, which expects an input geometry for the solute and the solvent and calculates the free energy of solvation of the solute in the solvent. A workflow of a calculation requesting COSMO-RS in ORCA is the following:

1. Single-point-energy calculation of the solute in the gas-phase
2. Single-point-energy calculation of the solute in a conductor ($\epsilon = \infty$)
3. Single-point-energy calculation of the solvent in a conductor ($\epsilon = \infty$)
4. Do COSMO-RS via the openCOSMO-RS executable and compute solvation properties

Points 1 to 3 correspond to DFT calculations at the BP86/def2-TZVPD level. This is the level of theory used to parametrize COSMO-RS for ORCA 6.0. As it is shown later, the user can change the functional and basis set, but this is not recommended!

Regarding calculations 2 and 3, to produce smooth σ profiles, ORCA discards the surface segments with an area < 0.01 $Å^2$. For these two calculations, the radii of several elements used to construct the cavity of solute and solvent are different than the defaults employed within non-COSMO-RS calculations (using C-PCM). This is due to the fact that the provided openCOSMO-RS binaries involve a special parametrization of COSMO-RS for ORCA 6.0, and this does not only affect the COSMO-RS parameters, but also the radii of several elements used in the C-PCM part.

After point 4 is done, the free energy of solvation (ΔG_S) of the solute in the solvent is printed in the ORCA output file. In ORCA 6.0 the use of openCOSMO-RS is restricted to the calculation of ΔG_S.

How to Run a ORCA/COSMO-RS Calculation

ORCA/COSMO-RS calculations are controlled through the `%cosmors` block. These type of calculations require two input structures, one for the solute and one for the solvent. The solute coordinates are provided in the input file as done for any other type of calculation. However, regarding the structure for the solvent, there are two options:

1. Retrieve the structure of the solvent from a database
2. Provide the structure in a separate file

To use strategy 1, we need to request the solvent via:

```
COSMORS(Solvent)
```

in the simple input or using the `%cosmors` block:

```
%cosmors
solvent "Solvent"
end
```

The list of internal solvents available in ORCA are shown in Table 7.29. For instance, for a water molecule solvated by acetonitrile, the input file looks like:

```
!COSMORS(Acetonitrile)
* xyz 0 1
  O    0.00000006589375       0.00157184228646       0.00000000004493
  H    0.77316868532439      -0.58666889665624      -0.00000000000005
  H   -0.77316876182122      -0.58666895650640      -0.00000000000005
*
```

If the user wants to provide a structure for the solvent (strategy 2), then a separate file with extension **.cosmorsxyz** should be available. The name of this file (without extension) is controlled by the tag `solventfilename` in the `%cosmors` block. For instance, if we want to calculate the free energy of solvation of acetone in water, the ORCA input file would look like this:

```
%cosmors
solventfilename "water"
end

* xyz 0 1
  H    1.99757808828569      0.25022586507917      0.72957579847856
  C    1.44666788654273      0.06088176074125     -0.18975506731892
  H    1.67115809398389      0.82690156694531     -0.93346420198265
  H    1.76410010378270     -0.89580894800398     -0.61702931492419
  C   -0.03218812888253     -0.00668250402989      0.08117960952503
  O   -0.47126229461002     -0.06383203315654      1.21590571064657
  C   -0.93671505604705     -0.00759565576779     -1.12174123739234
  H   -0.88761294036774      0.97611724740670     -1.59852429171257
  H   -0.58940208311204     -0.73299176093053     -1.85998395179737
  H   -1.96332366957567     -0.22151553828368     -0.83026305352213
*
```

The structure of the solvent (water) is the one in the `water.cosmorsxyz` file:

```
3
0 1
  O    0.00000006589375       0.00157184228646       0.00000000004493
  H    0.77316868532439      -0.58666889665624      -0.00000000000005
  H   -0.77316876182122      -0.58666895650640      -0.00000000000005
```

where the first line corresponds to the number of atoms, and in the second line the charge and multiplicity are provided.

The output for COSMO-RS is printed in the ORCA output file after the line that reads `OPENCOSMO-RS CALCULATION`. First of all, the information regarding the level of theory, the solute and the solvent is printed. For the example above (acetone in water), it reads as,

```
-------------------------------------------------------------------------------
                          OPENCOSMO-RS CALCULATION
-------------------------------------------------------------------------------

----------------------
GENERAL INFORMATION
----------------------
Calculation method                          ... DFT
Functional                                  ... BP86
Basis set                                   ... DEF2-TZVPD

----------------------
SOLUTE INFORMATION
----------------------
Number of atoms                             ...     10
Total charge                                ...      0
Multiplicity                                ...      1

CARTESIAN COORDINATES (ANGSTROEM)
  H      1.997578    0.250226    0.729576
  C      1.446668    0.060882   -0.189755
  H      1.671158    0.826902   -0.933464
  H      1.764100   -0.895809   -0.617029
  C     -0.032188   -0.006683    0.081180
  O     -0.471262   -0.063832    1.215906
  C     -0.936715   -0.007596   -1.121741
  H     -0.887613    0.976117   -1.598524
  H     -0.589402   -0.732992   -1.859984
  H     -1.963324   -0.221516   -0.830263

----------------------
SOLVENT INFORMATION
----------------------
Solvent name                                ... water
Number of atoms                             ...      3
Total charge                                ...      0
Multiplicity                                ...      1

CARTESIAN COORDINATES (ANGSTROEM)
  O      0.000000    0.001572    0.000000
  H      0.773169   -0.586669   -0.000000
  H     -0.773169   -0.586669   -0.000000
```

After these lines, ORCA prints the final single point energy for each of the QM calculations, together with the output file to which the ORCA output is redirected:

```
---------------------------------------------
Single Point Calculation (solute / gas-phase)
---------------------------------------------
Output single point calculation redirected to >test.solute_vac.lastout

FINAL SINGLE POINT ENERGY (Solute-gas-phase)     -193.233975714789

---------------------------------------------
Single Point Calculation (solute / CPCM)
---------------------------------------------
Output single point calculation redirected to >test.solute_cpcm.lastout
```

(continues on next page)

(continued from previous page)

```
FINAL SINGLE POINT ENERGY (Solute-CPCM)     -193.243987123912

-----------------------------------------------
Single Point Calculation (solvent / CPCM)
-----------------------------------------------
Output single point calculation redirected to >test.solvent_cpcm.lastout

FINAL SINGLE POINT ENERGY (Solvent-CPCM)      -76.479155022866
```

Once this information is printed, ORCA calls the openCOSMO-RS executable and ΔG_S is printed in the following block:

```
----------------------
SOLVATION DATA
----------------------
Reference temperature              :                 298.15 K
Free energy of solvation (dGsolv)  :       -0.006626234385 Eh             -4.158026 kcal/mol

Note: In order to calculate the free energy of the solvated solute (Gsolv), one should add
      the computed dGsolv to the "Final Gibbs free energy" (Gvac = H-T*S) of the solute in gas-
↪phase.
      That is: Gsolv = Gvac + dGsolv. Here, the Gvac has been calculated previously after a␣
↪frequency
      calculation of the solute at a certain level of theory and printed in the
↪"THERMOCHEMISTRY" block.
```

As pointed out in the last paragraph of the COSMO-RS output, to calculate the Gibbs free energy of the solute solvated by the given solvent, one should add the calculated ΔG_S to the `Final Gibbs free energy` of the solute in the gas-phase.

The parameters that can be defined in the `%cosmors` block in the ORCA input file are the following:

```
%cosmors
  aeff                5.92500 # Effective contact area between surface segments (Å^2)
  lnalpha             0.20200 # Logarithm of the misfit prefactor
  lnchb               0.16600 # Hydrogen bond (HB) strength parameter
  chbt                1.50    # Parameter for the temperature dependence of the HB
  sigmahb             9.61e-3 # HB threshold parameter (e/Å^2)
  rav                 0.50    # Radius to average ideal screening charges in Å
  fcorr               2.40    # Parameter adjusted from dielectric screening energies
  ravcorr             1.00    # Additional radius to calculate the misfit energy in Å
  astd                41.6240 # Standard surface area (normalization factor) in Å^2
  zcoord              10.0    # Coordination number
  dgsolv_eta         -4.44800 # Offset for the solv. energy calculation
  dgsolv_omegaring    0.26300 # Correction for solv. energy of molecules with rings
  temp                298.15  # Reference temperature in Kelvin
  dftfunc             "BP86"  # String for the DFT functional
  dftbas              "def2-TZVPD"  # String for the basis set
  solvent             "THF"   # Solvent from the internal database
  solventfilename     "water" # Name of the .cosmorsxyz solvent file
end
```

It is not recommended to change the defaults of the COSMO-RS parameters.

Note: The workflow explained above for ORCA/openCOSMO-RS calculations involves the same structure for the solute in the gas-phase and in solution. However, these structures may differ substantially depending on the type of solute and solvent. In ORCA, it is possible to optimize the structures for each of the three calculations needed in the ORCA/openCOSMO-RS workflow. That is, (1) the solute in gas-phase, (2) the solute in a conductor, and (3) the solvent in a conductor. To do that, one needs to add the level of optimization via the `dftfunc` tag, which is exclusive for the DFT functional, but as an exception, can be extended with the optimization tag:

```
%cosmors
  dftfunc  "BP86 tightopt"
end
```

7.50.6 Implicit Solvation in Coupled-Cluster Methods

The coupled-cluster Lagrangian, $\mathcal{L}$, for a system implicitly solvated reads as follows,[133, 142, 285]

$$\mathcal{L}(\Lambda,T) = \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}H_0\mathrm{e}^{T}|\psi_0\rangle + \frac{1}{2}\bar{\mathbf{Q}}(\Lambda,T)\cdot\bar{\mathbf{V}}(\Lambda,T) \tag{7.326}$$

where ψ_0 is the reference wave function, and H_0 is the Hamiltonian for the isolated molecule. The operator T for CCSD is defined in terms of single and double excitations ($T = T_1 + T_2$), and Λ is the de-excitation operator, defined in terms of the Lagrange multipliers:

$$T = T_1 + T_2 = \sum_{ia} t_a^i a_a^+ a_i + \sum_{ijab} t_{ab}^{ij} a_a^+ a_b^+ a_j a_i \tag{7.327}$$

$$\Lambda = \sum_{ia} \lambda_i^a a_i^+ a_a + \frac{1}{2}\sum_{ijab} \lambda_{ij}^{ab} a_i^+ a_a a_j^+ a_b \tag{7.328}$$

Here, t_a^i and t_{ab}^{ij} are the singles and doubles wave function amplitudes and a_i and a_a^+ are standard fermion annihilation and creation operators, respectively. Canonical occupied orbitals are denoted by the symbols $i, j, k, \ldots$, virtual orbitals by the symbols $a, b, c, \ldots$, and we use the symbols $p, q, r \ldots$ for general orbital indices.

The quantities $\bar{\mathbf{Q}}$ and $\bar{\mathbf{V}}$ are the CC expectation values of the C-PCM operators $\mathbf{Q}$ and $\mathbf{V}$, which are the solvation charges vector and solute potential vector defined at the position the charges, respectively.

$$\bar{\mathbf{Q}}(\Lambda,T) = \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}\mathbf{Q}\mathrm{e}^{T}|\psi_0\rangle \tag{7.329}$$

$$\bar{\mathbf{V}}(\Lambda,T) = \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}\mathbf{V}\mathrm{e}^{T}|\psi_0\rangle \tag{7.330}$$

Equation (7.326) can be rewritten by introducing the normal product form of an operator:

$$X_N = X - \langle\psi_0|X|\psi_0\rangle = X - X_0 \tag{7.331}$$

If one uses this result in eq (7.326), together with the fact that $\mathbf{Q}$ and $\mathbf{V}$ are related through eq (7.316), then eq (7.326) reads as,

$$\begin{aligned}\mathcal{L}(\Lambda,T) =& \langle\psi_0|H_0|\psi_0\rangle + \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}H_{0N}\mathrm{e}^{T}|\psi_0\rangle + \frac{1}{2}\mathbf{Q}_0\cdot\mathbf{V}_0 + \mathbf{Q}_0\cdot\bar{\mathbf{V}}_N(\Lambda,T) + \frac{1}{2}\bar{\mathbf{Q}}_N(\Lambda,T)\cdot\bar{\mathbf{V}}_N(\Lambda,T) = \\ =& E_0 + \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}H_{0N}\mathrm{e}^{T}|\psi_0\rangle + \mathbf{Q}_0\cdot\bar{\mathbf{V}}_N(\Lambda,T) + \frac{1}{2}\bar{\mathbf{Q}}_N(\Lambda,T)\cdot\bar{\mathbf{V}}_N(\Lambda,T)\end{aligned} \tag{7.332}$$

Here, $\mathbf{Q}_0$ and $\mathbf{V}_0$ are the $\mathbf{Q}$ and $\mathbf{V}$ vectors calculated with the ψ_0 wave function, and E_0 is the reference energy ($E_0 = \langle\psi_0|H_0|\psi_0\rangle + \frac{1}{2}\mathbf{Q}_0\cdot\mathbf{V}_0$). Different approximations can be adopted in eq (7.332) depending on how one calculates its last term $\frac{1}{2}\bar{\mathbf{Q}}_N(\Lambda,T)\cdot\bar{\mathbf{V}}_N(\Lambda,T)$.

In ORCA there are three different CCSD/CPCM approaches: (i) the PTE scheme, (ii) the PTE(S) scheme, and the (iii) the PTES scheme, being the last one the default. Here, the acronym PTE stands for "perturbation theory and energy" and "S" for singles. The choice of any of these approaches is controlled via the tag "`CPCMccm`". Information about which CCSD/C-PCM is used by ORCA in a calculation is printed in the "`ORCA-MATRIX DRIVEN CI`" block in the output file in the line starting by "`CPCM scheme`":

```
-------------------------------------------------------------------------------
                            ORCA-MATRIX DRIVEN CI
-------------------------------------------------------------------------------

--------------------------------
```

(continues on next page)

(continued from previous page)

```
AUTOMATIC CHOICE OF INCORE LEVEL
--------------------------------

Memory available                           ...   2000.00 MB
Memory needed for S+T                      ...      9.26 MB
Memory needed for J+K                      ...     18.57 MB
Memory needed for DIIS                     ...    129.61 MB
Memory needed for 3-ext                    ...     72.69 MB
Memory needed for 4-ext                    ...    486.14 MB
Memory needed for triples                  ...      0.00 MB
 -> Final InCoreLevel    ... 5

Wavefunction type
-----------------
Correlation treatment                      ...      CCSD
Single excitations                         ... ON
Orbital optimization                       ... OFF
Calculation of Lambda equations            ... ON
Calculation of Brueckner orbitals          ... OFF
Perturbative triple excitations            ... OFF
CPCM scheme                                ... PTE(S)
Calculation of F12 correction              ... OFF
```

In the following subsections, we describe the different CCSD/C-PCM approaches available in ORCA and how to use them.

PTE scheme

In the "perturbation theory energy" (PTE) scheme, the last term in eq (7.332) is equal to zero (this term does not depend on Λ and T),

$$\mathcal{L}(\Lambda,T) = E_0 + \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}H_{0N}\mathrm{e}^{T}|\psi_0\rangle + \mathbf{Q}_0\cdot\bar{\mathbf{V}}_N(\Lambda,T) \tag{7.333}$$

The potential $\bar{\mathbf{V}}_N$ can be written as follows:

$$\bar{\mathbf{V}}_N(\Lambda,T) = \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}\sum_{pq}\mathbf{v}_{pq}\{p^+q\}\mathrm{e}^{T}|\psi_0\rangle = \sum_{pq}\mathbf{v}_{pq}\langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}\{p^+q\}\mathrm{e}^{T}|\psi_0\rangle = \sum_{pq}\mathbf{v}_{pq}\Gamma_{pq} \tag{7.334}$$

where we have used that $\mathbf{V}_N = \sum_{pq}\mathbf{v}_{pq}\{p^+q\}$ (second-quantized form of a normal ordered operator), with $\mathbf{v}_{pq}$ the components of the solute potential in the MO basis. The matrix Γ is the CCSD relaxed one-electron density matrix. Then, the contribution to the equations for the T amplitudes comes from the derivative of $\bar{\mathbf{V}}_N(\Lambda,T)$ with respect to the Λ amplitudes ($\mathbf{Q}_0$ does not depend on the Lagrange multipliers). In this context, the Hamiltonian H_{0N} contains a term that depends on the elements of the Fock matrix ($\sum_{pq}f_{pq}\{p^+q\}$) and that has the same functional form as $\mathbf{V}_N$. As the Fock matrix is updated in the reference calculation with a C-PCM term that reads as (in AO basis) $F_{\mu\nu}^{\mathrm{CPCM}} = \mathbf{Q}_0\cdot\mathbf{V}_{\mu\nu}$, then the term $\mathbf{Q}_0\cdot\bar{\mathbf{V}}_N(\Lambda,T)$ is added implicitly to eq (7.333).

Once the T amplitudes are obtained, the total energy, E, is calculated as

$$E = E_0 + \langle\psi_0|\mathrm{e}^{-T}(H_{0N}+\mathbf{Q}_0\cdot\mathbf{V}_N)\mathrm{e}^{T}|\psi_0\rangle = E_0 + \sum_{ia}F_{ia}t_a^i + \frac{1}{4}\sum_{ijab}\langle ij||ab\rangle\left(t_{ab}^{ij}+2t_a^it_b^j\right) \tag{7.335}$$

Then, the C-PCM contribution to the CC energy within the PTE scheme occurs through the term $\frac{1}{2}\mathbf{Q}_0\cdot\mathbf{V}_0$ in E_0 and implicitly through the Fock matrix elements F_{ia} ($F_{ia} = F_{ia}^0 + \mathbf{Q}_0\cdot\mathbf{v}_{ia}$).

The PTE scheme corresponds to "`CPCMccm 0`", and is implemented for canonical-CCSD (RHF and UHF) and DLPNO-CCSD (RHF and UHF). For instance, for a DLPNO-CCSD calculation (closed-shell reference) of a system solvated by water using the PTE scheme, the input file looks like:

```
! DLPNO-CCSD cc-pVTZ cc-PVTZ/C TightSCF CPCM(Water)

%cpcm
CPCMccm 0
end

* xyzfile 0 1 water.xyz
```

PTE(S) scheme

In this scheme (where the "S" stands for singles), the last term in eq (7.332) depends on the T amplitudes, but not on the Λ amplitudes,

$$\mathcal{L}(\Lambda,T) = E_0 + \langle\psi_0|(1+\Lambda)\mathrm{e}^{-T}H_{0N}\mathrm{e}^{T}|\psi_0\rangle + \mathbf{Q}_0\cdot\bar{\mathbf{V}}_N(\Lambda,T) + \frac{1}{2}\bar{\mathbf{Q}}_N(T)\cdot\bar{\mathbf{V}}_N(T) \tag{7.336}$$

Again, in the same way as for the PTE scheme, the C-PCM contribution to the equations for the T amplitudes comes from the term $\mathbf{Q}_0\cdot\bar{\mathbf{V}}_N(\Lambda,T)$ in eq (7.336), which is implictly added to the Fock matrix elements in the MO basis. The last term in eq (7.336) does not depend on the Λ amplitudes and then does not contribute to the equations for the T amplitudes. However, this term depends on the T amplitudes through the elements γ_{ai} of the density matrix Γ,

$$\gamma_{ai} = t_a^i + \sum_{me}\left(t_{ae}^{im} - t_e^i t_a^m\right)\lambda_m^e - \frac{1}{2}\sum_{mnef}\lambda_{mn}^{ef}\left(t_{ef}^{in}t_a^m + t_{af}^{mn}t_e^i\right) \tag{7.337}$$

and then it contributes to the final energy E in the following way:

$$\frac{1}{2}\bar{\mathbf{Q}}_N(T)\cdot\bar{\mathbf{V}}_N(T) = \frac{1}{2}\left(\sum_{ai}t_a^i\mathbf{q}_{ai}\right)\cdot\left(\sum_{ai}t_a^i\mathbf{v}_{ai}\right) \tag{7.338}$$

That gives the final equation for the total energy of our system,

$$\begin{aligned} E &= E_0 + \langle\psi_0|\mathrm{e}^{-T}(H_{0N}+\mathbf{Q}_0\cdot\mathbf{V}_N)\mathrm{e}^{T}|\psi_0\rangle + \frac{1}{2}\left(\sum_{ai}t_a^i\mathbf{q}_{ai}\right)\cdot\left(\sum_{ai}t_a^i\mathbf{v}_{ai}\right) = \\ &= E_0 + \sum_{ia}F_{ia}t_a^i + \frac{1}{4}\sum_{ijab}\langle ij||ab\rangle\left(t_{ab}^{ij} + 2t_a^i t_b^j\right) + \frac{1}{2}\left(\sum_{ai}t_a^i\mathbf{q}_{ai}\right)\cdot\left(\sum_{ai}t_a^i\mathbf{v}_{ai}\right) \end{aligned} \tag{7.339}$$

Therefore, the CC energy for a solvated system within the PTE(S) scheme involves three C-PCM contributions: (1) the term $\frac{1}{2}\mathbf{Q}_0\cdot\mathbf{V}_0$ included in E_0, (2) the term $\sum_{ia}\mathbf{Q}_0\cdot\mathbf{v}_{ia}t_a^i$ that occurs implicitly through $\sum_{ia}F_{ia}t_a^i$ and (3) the term $\frac{1}{2}\left(\sum_{ai}t_a^i\mathbf{q}_{ai}\right)\cdot\left(\sum_{ai}t_a^i\mathbf{v}_{ai}\right)$.

This scheme is available in ORCA for canonical-CCSD (RHF and UHF) and DLPNO-CCSD (RHF and UHF). In order to use it, the user needs to add the tag "`CPCMccm 1`" in the `%cpcm` block.

```
! DLPNO-CCSD cc-pVTZ cc-PVTZ/C TightSCF CPCM(Water)

%cpcm
CPCMccm 1
end

* xyzfile 0 1 water.xyz
```

PTES scheme

In this scheme, both $\bar{\mathbf{Q}}_N$ and $\bar{\mathbf{V}}_N$ in the last term in eq (7.332) depend on the T amplitudes but just one of them depends on the Λ amplitudes,

$$\mathcal{L}(\Lambda,T) = E_0 + \langle\psi_0|(1+\Lambda)\mathrm{e}^T H_{0N}\mathrm{e}^T|\psi_0\rangle + \mathbf{Q}_0\cdot\bar{\mathbf{V}}_N(\Lambda,T) + \frac{1}{2}\bar{\mathbf{Q}}_N(T)\cdot\bar{\mathbf{V}}_N(\Lambda,T) \tag{7.340}$$

The C-PCM terms enter these equations on the one hand, implicitly, through the elements of the Fock matrix (like for the PTE and PTE(S) schemes), and on the other hand, explicitly through the derivatives of $\frac{1}{2}\bar{\mathbf{Q}}_N(T)\cdot\bar{\mathbf{V}}_N(\Lambda,T)$ with respect to Λ. If we call $\mathcal{L}_{\mathrm{CPCM}}$ the last C-PCM term in eq (7.340), then the contribution from this term to the T amplitudes equations read as:

$$\frac{\partial\mathcal{L}_{\mathrm{CPCM}}}{\partial\lambda_i^a} = \frac{1}{2}\left(\sum_{ai} t_a^i\mathbf{q}_{ai}\right)\cdot\left[-\sum_j t_a^j\mathbf{v}_{ji} + \sum_b t_b^i\mathbf{v}_{ab} + \mathbf{v}_{ia} + \sum_{bj}\left(t_{ba}^{ji} - t_a^j t_b^i\right)\mathbf{v}_{bj}\right] \tag{7.341}$$

$$\frac{\partial\mathcal{L}_{\mathrm{CPCM}}}{\partial\lambda_{ij}^{ab}} = \frac{1}{2}\left(\sum_{ai} t_a^i\mathbf{q}_{ai}\right)\cdot\left[-\frac{1}{2}\sum_k t_{ab}^{kj}\mathbf{v}_{ki} + \frac{1}{2}\sum_c t_{ac}^{ij}\mathbf{v}_{bc} - \frac{1}{2}\sum_{ck}\left(t_{ab}^{kj}t_c^i + t_{cb}^{ij}t_a^k\right)\mathbf{v}_{ck}\right] \tag{7.342}$$

The contribution to the energy is the same as that for the PTE(S) scheme, but with different values for the T amplitudes (as the equations to calculate them differ slightly from those for the PTE(S) scheme).

$$\begin{aligned} E &= E_0 + \langle\psi_0|\mathrm{e}^{-T}(H_{0N} + \mathbf{Q}_0\cdot\mathbf{V}_N)\mathrm{e}^T|\psi_0\rangle + \frac{1}{2}\left(\sum_{ai} t_a^i\mathbf{q}_{ai}\right)\cdot\left(\sum_{ai} t_a^i\mathbf{v}_{ai}\right) = \\ &= E_0 + \sum_{ia} F_{ia}t_a^i + \frac{1}{4}\sum_{ijab}\langle ij||ab\rangle\left(t_{ab}^{ij} + 2t_a^i t_b^j\right) + \frac{1}{2}\left(\sum_{ai} t_a^i\mathbf{q}_{ai}\right)\cdot\left(\sum_{ai} t_a^i\mathbf{v}_{ai}\right) \end{aligned} \tag{7.343}$$

This scheme is the default CCSD/C-PCM approach in ORCA and is available in ORCA for canonical-CCSD (RHF and UHF) and DLPNO-CCSD (RHF and UHF). In this case, the tag "`CPCMccm`" in the `%cpcm` block is equal to 2. However, as the PTES scheme is the default in ORCA, the user just needs to add the information about the solvent in the input file, in order to use this approach.

```
! DLPNO-CCSD cc-pVTZ cc-PVTZ/C TightSCF CPCM(Water)

* xyz 0 1
O  -0.00000018976103    0.00606010894837    0.00000000004527
H   0.76098169249695   -0.58891312953082   -0.00000000000022
H  -0.76098151333900   -0.58891299029372   -0.00000000000022
*
```

Notes regarding the use of the CCSD/C-PCM schemes:

- For calculations within the PTE(S) and the PTES schemes, the explicit C-PCM contribution to the total energy (see eqs. (7.339) and (7.343)) is printed in the "`COUPLED CLUSTER ENERGY`" block after the equations for the "T" amplitudes converge. In this case, this energy is labelled as "`C-PCM corr. term`" and is already included in the correlation energy, "`E(CORR)` ". For the input example from above, this information looks like:

```
----------------------
COUPLED CLUSTER ENERGY
----------------------

E(0)                                    ...    -76.066903687
E(CORR)(strong-pairs)                   ...     -0.267203798
E(CORR)(weak-pairs)                     ...     -0.000106106
E(CORR)(corrected)                      ...     -0.267309904
C-PCM corr. term (included in E(CORR))  ...     -0.000003158
E(TOT)                                  ...    -76.334213591
Singles Norm <S|S>**1/2                 ...      0.017784967
T1 diagnostic                           ...      0.006287935
```

This contribution does not represent the whole C-PCM contribution to the correlation energy, as this one also occurs, implicitly, through the "T" amplitudes.

- The C-PCM contribution to the Λ equations is implemented in ORCA for the PTE(S) and PTES schemes. Then, the user can request unrelaxed densities.

7.51 Calculation of Properties

7.51.1 Electric Properties

Calculation of first order (electric dipole and quadrupole moments) and second order (polarizabilities) electric properties can be requested in the `%elprop` input block. The second order properties can be calculated through the solution of the CP-SCF (see *CP-SCF Options*) or CP-CASSCF (see *CASSCF Linear Response*) equations. Details are shown below:

```
%elprop
  Dipole     true
  Quadrupole true
  Polar      true
  PolarVelocity true # polarizability w.r.t. velocity perturbations
  PolarDipQuad  true # dipole-quadrupole polarizability
  PolarQuadQuad true # quadrupole-quadrupole polarizability
  freq_r 0.00   # purely real frequency (default: static calculation)
  freq_i 0.00   # purely imaginary frequency (default: static calculation)
  Solver  CG    # CG(conjugate gradient)
                # other options: DIIS or POPLE(default)
  MaxDIIS  5    # max. dimension of DIIS method
  Shift   0.2   # level shift used in DIIS solver
  Tol     1e-3  # Convergence of the CP-SCF equations
                # (norm of the residual)
  MaxIter 64    # max. number of iterations in CPSCF
  PrintLevel 2
  Origin  CenterOfElCharge  # center of electronic charge
          CenterOfNucCharge # center of nuclear charge
          CenterOfSpinDens        # center of spin density
          CenterOfMass      # center of mass (default)
          N                 # position of atom N (starting at 0)
          X,Y,Z             # explicit position of the origin
                            # in coordinate input units (Angstrom by default)
  end
```

The most efficient and accurate way to calculate the polarizability analytically is to use the coupled-perturbed SCF method. The most time consuming and least accurate way is the numerical second derivative of the total energy. Note that the numerical differentiation requires: (a) tightly or even very tightly converged SCF calculations and (b) carefully chosen field increments. If the field increment is too large then the truncation error will be large and the values will be unreliable. On the other hand, if the field increment is too small the numerical error associated with the finite difference differentiation will get unacceptably large up to the point where the whole calculation becomes useless.

7.51.2 The Spin-Orbit Coupling Operator

Several variants of spin-orbit coupling operators can be used for property calculations [610]. They are based on effective potential and mean-field approaches, and have various parameters that can be selected via the `%rel` block. Note that the SOMF operator depends on the density matrix, so the operator itself can differ for example between a CASSCF and an MRCI calculation.

Note: The defaults have slightly changed in ORCA 5.0, see `SOCFlags` in the following.

```
%rel
     # ------------------------------------------------
     # SPIN ORBIT COUPLING OPERATORS
     # ------------------------------------------------
     SOCType 0  # none
             1  # effective nuclear charge
             2  # mean-field with atomic densities read from
                # disk; similar to SOCType=4
             3  # mean-field/effective potential (default)
             4  # mean-field with atomic densities generated
                # on the fly; see bellow
     # ------------------------------------------------
     # Flags for construction of potential; operative
     # only for SOCType 3.
     # ------------------------------------------------
     SOCFlags 1,4,3,0 # default if nothing is specified
                      # 1,3,3,0 default if RI is applied and AuxJ available
                      # e.g. when using !RIJCOSX (default for DFT) or !RIJONX
     # Flag 1 = 0 - do not include 1-electron terms
     #        = 1 - do include 1-electron terms
     # Flag 2 = 0 - do not include Coulomb terms
     #        = 1 - compute Coulomb terms fully numeric
     #        = 2 - compute Coulomb term seminumeric
     #        = 3 - compute Coulomb term with RI approx
     #        = 4 - compute Coulomb term exactly
     # Flag 3 = 0 - do not include exchange terms
     #        = 1 - do include local X-alpha exchange
     #              the X-Alpha parameter can be chosen via
     #              % rel Xalpha 0.7 (default)
     #        = 2 - same as 1 but with sign reversed
     #        = 3 - exchange via one-center exact
     #              integrals including the spin-other
     #              orbit interaction
     #        = 4 - all exchange terms full analytic
     #              (this is expensive)
     # Flag 4 = 0 - do not include DFT local correlation
     #              terms
     #        = 1 - do include local DFT correlation (here
     #              this is done with VWN5)
     #
     SOCMaxCenter 4 # max. number of centers to include in
                    # the integrals (not fully consistently
                    # implemented yet; better leave equal to 4)
     #  The following simple input equivalents can also be used:
     #  SOMF(1X)      = SOCType 3, SOCFlags 1,2,3,0 and SOCMaxCenter 4
     #  RI-SOMF(1X)   = SOCType 3, SOCFlags 1,3,3,0 and SOCMaxCenter 4
     #  VEFF-SOC      = SOCType 3, SOCFlags 1,3,1,1 and SOCMaxCenter 4
     #  VEFF(-2X)-SOC = SOCType 3, SOCFlags 1,3,2,1 and SOCMaxCenter 4
     #  AMFI          = SOCType 3, SOCFlags 1,4,3,0 and SOCMaxCenter 1
     #  AMFI-A        = SOCType 4, SOCFlags 1,4,3,0 and SOCMaxCenter 1
     #                  (AMFI-like approach that uses pre-calculated atomic densities;
     #                   this can be combined with the SOCOff option
     #                   to exclude contributions from specific atoms)
```

(continues on next page)

(continued from previous page)

```
#  NOTE: If you choose the RI option you need to specify an auxiliary basis set
#        even if the underlying SCF calculation does not make use of any form
#        of RI!
# ------------------------------------------------
# For the effective nuclear charge SOC operator
# the nuclear charges can be adjusted.
# ------------------------------------------------
Zeff[26] 0.0   # set the effective nuclear charge
               # of iron (Z = 26) to zero
# ------------------------------------------------
# Neglecting SOC contributions from particular
# atoms
# ------------------------------------------------
SOCOff 0,5     # turn off the SOC for atoms 0 and 5
               # this makes sense if the SOC operator
               # has only one center contributions
               # (e.g. effective nuclear charge)
```

Simple input equivalents are described in more detail in [610]. More details on the AMFI-A approach which uses pre-calculated atomic densities can be found in [284].

The Breit-Pauli spin-orbit coupling operator is given by:

$$\hat{H}_{\mathrm{SOC}} = \hat{H}^{(1)}_{\mathrm{SOC}} + \hat{H}^{(2)}_{\mathrm{SOC}}$$

with the one- and two-electron contributions

$$\hat{H}^{(1)}_{\mathrm{SOC}} = \frac{\alpha^2}{2} \sum_i \sum_A Z_A \frac{(\mathbf{r}_i - \mathbf{R}_A) \times \mathbf{p}_i}{|\mathbf{r}_i - \mathbf{R}_A|^3} \hat{s}_i \equiv \frac{\alpha^2}{2} \sum_i \sum_A Z_A r_{iA}^{-3} \hat{\mathbf{l}}_{iA} \hat{\mathbf{s}}_i \tag{7.344}$$

$$\hat{H}^{(2)}_{\mathrm{SOC}} = -\frac{\alpha^2}{2} \sum_i \sum_{j \neq i} \frac{(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{p}_i}{|\mathbf{r}_i - \mathbf{r}_j|^3} (\hat{\mathbf{s}}_i + 2\hat{\mathbf{s}}_j) \tag{7.345}$$

$$\equiv -\frac{\alpha^2}{2} \sum_i \sum_{j \neq i} \hat{\mathbf{l}}_{ij} r_{ij}^{-3} (\hat{\mathbf{s}}_i + 2\hat{\mathbf{s}}_j) \tag{7.346}$$

This operator would be hard to handle exactly; therefore it is common to introduce mean field and/or effective potential approaches in which the operator is written as an effective one-electron operator:

$$\hat{H}_{\mathrm{SOC}} \cong \sum_i \hat{\mathbf{h}}_i^{(\mathrm{eff})} \hat{\mathbf{s}}_i \tag{7.347}$$

The simplest approximation is to simply use the one-electron part and regard the nuclear charges as adjustable parameters. Reducing their values from the exact nuclear charge is supposed to account in an average way for the screening of the nuclear charge by the electrons. In our code we use the effective nuclear charges of Koseki et al. This approximation introduces errors which are usually smaller than 10% but sometimes are larger and may approach 20% in some cases. The approximation is best for first row main group elements and the first transition row (2p and 3d elements). For heavier elements it becomes unreliable.

A much better approximation is to take the two-electron terms into account precisely. Without going into details here – the situation is as in Hartree-Fock (or density functional) theory and one gets Coulomb, exchange and correlation terms. The correlation terms (evaluated in a local DFT fashion) are negligible and can be safely neglected. They are optionally included and are not expensive computationally. The Coulomb terms is (after the one-electron term) the second largest contribution and is expensive to evaluate exactly. The situation is such that in the Coulomb-part the spin-other orbit interaction (the second term in the two-electron part) does not contribute and one only has to deal with the spin-own-orbit contribution. The exact evaluation is usually too expensive to evaluate. The RI and seminumeric approximation are much more efficient and introduce only minimal errors (on the order of usually not more than 1 ppm in g-tensor calculations for example) and are therefore recommended. The RI approximation is computationally more efficient. Please note that you have to specify an auxiliary basis set to take advantage of the RI approximation, even if the preceding SCF calculation does not make use of any form of RI. The one-center

approximation to the Coulomb term introduces much larger errors. The fully numeric method is both slower and less accurate and is not recommended.

The exchange term has contributions from both the spin-own-orbit and spin-other-orbit interaction. These are taken both into account in the mean-field approximation which is accessed by `Flag 3 = 3`. Here a one-center approximation is much better than for the Coulomb term since both the integrals and the density matrix elements are short ranged. Together with the Coulomb term this gives a very accurate SOC operator which is recommended. The DFT-Veff operator suffers from not treating the spin-other-orbit part in the exchange which gives significant errors (also, local DFT underestimates the exchange contributions from the spin-same-orbit interaction by some 10% relative to HF but this is not a major source of error). However, it is interesting to observe that in the precise analytical evaluation of the SOMF operator, the spin-other-orbit interaction is exactly -2 times the spin-own-orbit interation. Thus, in the DFT framework one gets a *much* better SOC operator if the sign of the DFT exchange term is simply reversed! This is accessed by `Flag 3 = 2`.

Exclusion of Atomic Centers

In ORCA it is possible to change the spin-orbit coupling operator in order to exclude contributions from user-defined atoms. This approach can be useful, for example, in quantifying the contribution of the ligands to the zero-field splitting (ZFS); for an application of this method see Ref. [834].

This is illustrated for the calculation of the SOC contribution to the ZFS of the triplet oxygen molecule. Using the input below we start by a normal calculation of the ZFS, including both oxygen atoms. Note that we use here the effective nuclear charge operator. This is required as not all implemented SOC operators are compatible with the decomposition in terms of individual centers contributions.

```
! def2-TZVP def2-TZVP/c

%casscf nel 8
        norb 6
        mult 3,1
        nroots 1,3
        rel dosoc true
            end
        end

%rel
    SOCType 1
    end

*xyz 0 3
O 0 0 0
O 0 0 1.207
*
```

The calculated value of the D parameter is approximately 2.573 cm^{-1}. In a second calculation we exclude the contribution from the first oxygen atom. For this we change the `%rel` block to the one below.

```
%rel
    SOCType 1
    SOCOff 0
    end
```

Now the D parameter is calculated to be approximately 0.643 cm^{-1}, a result that deviates quite significantly from half of the value calculated previously, implying that non-additive effects are important. In addition to the effective nuclear charge operator, the AMFI-A operator described previously can be used. Given that this is based on pre-calculated atomic densities, it might be preferred for heavier elements where the effective nuclear charge operator becomes unreliable. The method is not limited to CASSCF calculations as described above, and can be used in DFT, MRCI and ROCIS calculations.

7.51.3 EPR and NMR properties

Calculation of EPR and NMR response properties can be requested in the %eprnmr input block. The individual flags are given below.

```
%eprnmr
  # Calculate the g-tensor using CP-KS theory
  gtensor true
  # Calculate and print one- and two-electron contributions to the g-tensor
  gtensor_1el2el true

  # Calculate the D-tensor
  DTensor so      # spin orbit part
          ss      # spin-spin part
          ssandso # both parts
  DSOC    qro     # quasi-restricted method; must be done with the keyword !uno
          pk      # Pederson-Khanna method.
                  # NOTE: both approximations are only valid for
                  # pure functionals (no HF exchange)
          cp      # coupled-perturbed method (default)
          cvw     # van W\"ullen method
  DSS     direct  # directly use the canonical orbitals for the spin density
          uno     # use spin density from UNOs

  PrintLevel n    # Amount of output (default 2)

  # whether to calculate and print the Euler angles via `orca_euler` if the
  # calculation of the g-tensor or the D-tensor is requested
  PrintEuler false

  # For the solution of the CP-SCF equations:
  Solver   Pople  # Pople solver (default)
           CG     # conjugate gradient
           DIIS   # DIIS type solver
  MaxIter  64     # maximum number of iterations
  MaxDIIS  10     # max. number of DIIS vectors (only DIIS)
  Tol      1e-3   # convergence tolerance
  LevelShift 0.05 # level shift for DIIS (ignored otherwise)

  Ori     CenterOfElCharge  # center of electronic charge
          CenterOfNucCharge # center of nuclear charge
          CenterOfSpinDens        # center of spin density
          CenterOfMass      # center of mass
          GIAO              # use the GIAO formalism (default)
          N                 # number of the atom to put the origin
          X,Y,Z             # explicit position of the origin
                            # in coordinate input units (Angstrom by default)

  # Calculate the NMR shielding tensor
  NMRShielding 1 # for chosen nuclei - specified with the Nuclei keyword
               2 # for all nuclei - equivalent to the 'NMR' simple input keyword

  # treatment of 1-electron integrals in the RHS of the CPSCF equations
  giao_1el = giao_1el_analytic   # analytical, default
             giao_1el_numeric    # numerical - for testing only

  # treatment of 2-electron integrals in the RHS of the CPSCF equations
  # various options for combination of approximations in Coulomb (J) and
  # exact (HF) exchange (K) part. The default is the same as used for the SCF.
  giao_2el = giao_2el_rijk        # RIJK approximation
             giao_2el_same_as_scf # use same scheme as in SCF
             giao_2el_analytic    # fully analytical, for testing only
```

(continues on next page)

(continued from previous page)

```
            giao_2el_rijonx      # RIJ approximation with analytical K
            giao_2el_cosjonx     # COSJ approximation with analytical K
            giao_2el_rijcosx     # RIJ approximation with COSX approximation
            giao_2el_cosjx       # COSJ approximation with COSX apprxomation
            giao_2el_exactjcosx  # exact J with COSX approximation
            giao_2el_exactjrik   # exact J with RIK approximation

# for g-tensor calculations using the SOMF-operator for the SOC
# treatment, the 2-electron contribution to the GIAO terms can be
# computed as well, but they take much more time and usually do not
# contribute significantly and therefore are disabled by default
do_giao_soc2el false

# treatment of tau in meta-GGA DFT - see below
Tau = Dobson # (default) Other options: 0, MS, GI

# use effective nuclear charges for the gauge correction to the A-tensor
# (this makes sense if an effective 1-electron SOC operator is used)
hfcgaugecorrection_zeff true

# calculate diamagnetic spin-orbit (DSO) integrals needed for the gauge
# correction to the A-tensor numerically (faster than the analytical way)
hfcgaugecorrection_numeric true

# Grid settings for the above: <0 means to use the DFT grid setting
hfcgaugecorrection_angulargrid -1
hfcgaugecorrection_intacc      -1
hfcgaugecorrection_prunegrid   -1
hfcgaugecorrection_bfcutoff    -1
hfcgaugecorrection_wcutoff     -1

Nuclei = all type { flags }
  # Calculate nuclear properties. Here the properties
  # for all nuclei of "type" are calculated ("type" is
  # an element name, for example Cu.)
  # Flags can be the following:
  #   aiso  - calculate the isotropic part of the HFC
  #   adip  - calculate the dipolar part of the HFC
  #   aorb  - 2nd order contribution to the HFC from SOC
  #   fgrad - calculate the electric field gradient
  #   rho   - calculate the electron density at the
  #           nucleus
  #   shift - NMR shielding tensor (orbital contribution)
  #   srot  - spin-rotation tensor
  #   ssdso - spin-spin coupling constants, diamagnetic spin orbit term
  #   sspso - spin-spin coupling constants, paramagnetic spin orbit term
  #   ssfc  - spin-spin coupling constants, Fermi contact term
  #   sssd  - spin-spin coupling constants, spin dipole term
  #   ssall - spin-spin coupling constants, calculate all above contributions

# In addition you can change several parameters
# e.g. for a different isotope.
Nuclei = all N { PPP=39.1, QQQ=0.5, III=1.0 };
    # PPP   - the HFC proportionality factor for this nucleus
    #         = ge*gN*betaE*betaN
    # QQQ   - the quadrupole moment for this nucleus
    # III   - the spin for this nucleus
    # ist   - isotope
    # ssgn  - nuclear g-factor for spin-spin coupling
    #         and spin-rotation constants (overrides ist)
```

(continues on next page)

(continued from previous page)

```
# For example:
# calculates the hyperfine coupling for all nitrogen atoms
Nuclei = all N { aiso, adip, fgrad, rho};

# calculates the spin-spin coupling constants for all carbon atoms
# assuming all carbon atoms are 13C
Nuclei = all C { ssall, ist = 13};

# You can also calculate properties for individual atoms
# as in the following example:
Nuclei = 1,5,8 { aiso, adip};

# NOTE: Counting starts with atom 1!
# WARNING: All the nuclei, mentioned in one line
# as above will be assigned the same parameters !

# For spin-spin coupling constants, a distance threshold is
# applied in the eprnmr block to restrict the number of couplings
# to be computed, given in Angstroms:
SpinSpinRThresh 5.0 # default

# Coupling can be restricted to certain element pairs
# (if they are added to the "Nuclei" list and are within RThresh).
# The syntax accepts multiple pairs of element symbols or
# atomic numbers or "*" as a wildcard
SpinSpinElemPairs {C C} {H *} {6 7} # default is {* *}, i.e. all

# Similarly, coupling can be restricted to certain atom pairs
# (if they are added to the "Nuclei" list and are within RThresh).
# The syntax accepts multiple pairs of indices (starting at 0!)
# or "*" as a wildcard
SpinSpinAtomPairs {1 0} {5 *} # default is {* *}, i.e. all

# whether to print reduced spin-spin coupling constants
PrintReducedCoupling false

end
```

Hyperfine and Quadrupole Couplings

The hyperfine coupling has four contributions:

(a) The isotropic Fermi contact term that arises from the finite spin density on the nucleus under investigation. It is calculated for nucleus N from:

$$a_{\text{iso}}(N) = \left(\frac{4}{3}\pi \langle S_z \rangle^{-1}\right) g_e g_N \beta_e \beta_N \rho\left(\vec{R}_N\right) \tag{7.348}$$

Here, $\langle S_z \rangle$ is the expectation value of the z-component of the total spin, g_e and g_N are the electron and nuclear g-factors and β_e and β_N are the electron and nuclear magnetons respectively. $\rho\left(\vec{R}_N\right)$ is the spin density at the nucleus. The proportionality factor $P_N = g_e g_N \beta_e \beta_N$ is commonly used and has the dimensions MHz bohr^3 in ORCA.

(b) The spin dipole part that arises from the magnetic dipole interaction of the magnetic nucleus with the magnetic moment of the electron. It is also calculated as an expectation value over the spin density as:

$$A^{\text{dip}}_{\mu\nu}(N) = P_N \sum kl \rho_{kl} \left\langle \phi_k \left| r_N^{-5} \left(3\vec{r}_{N\mu}\vec{r}_{N\nu} - \delta_{\mu\nu} r_N^2\right) \right| \phi_l \right\rangle \tag{7.349}$$

where ρ is the spin-density matrix and $\vec{r}_N$ is a vector of magnitude r_N that points from the nucleus in question to the electron ($\{\phi\}$ is the set of basis functions).

(c) The second order contribution that arises from spin-orbit coupling. Presently ORCA can calculate all these contributions. The first two are calculated as simple expectation values of the appropriate operators over the self-consistent spin density, but the second order contribution requires the solution of the coupled-perturbed SCF equations and is consequently computationally more demanding. The contribution can be written:

$$A_{\mu\nu}^{\text{orb}}(N) = -\frac{1}{2S} P_N \sum_{kl} \frac{\partial \rho_{kl}}{\partial I_\nu} \left\langle \phi_k \left| h_\mu^{\text{SOC}} \right| \phi_l \right\rangle \tag{7.350}$$

The derivative of the spin density is computed from solving the coupled-perturbed SCF equations with respect to the nucleus-orbit coupling as perturbation. The nucleus-orbit coupling is represented by the operator

$$h_\nu^{\text{NOC}}(A) = \sum_i r_{iA}^{-3} l_{i,\nu}^{(A)} \tag{7.351}$$

where the sum is over electrons and A is the nucleus in question.

(d) The gauge-correction contribution.[492] This term is often small. However, it is needed in order to get exactly gauge-invariant results. We recently showed that the gauge correction can become crucial in the long-distance limit between the nuclear spin and the electron spin. This is relevant for pseudocontact NMR chemical shifts (PCS).[492]

The **field gradient tensor** is closely related to the dipole contribution to the hyperfine coupling. The main differences are that the electron instead of the spin density enters its calculation and that it contains a nuclear contribution due to the surrounding nuclei. It is calculated from

$$\begin{aligned} V_{\mu\nu}(N) = &- \sum_{kl} P_{kl} \left\langle \phi_k \left| r_N^{-5} \left(3\vec{r}_{N\mu}\vec{r}_{N\nu} - \delta_{\mu\nu} r_N^2 \right) \right| \phi_l \right\rangle \\ &+ \sum_{A \neq N} Z_A \vec{R}_{AN}^{-5} \left(3\vec{R}_{AN\mu}\vec{R}_{AN\nu} - \delta_{\mu\nu} R_{AN}^2 \right) \end{aligned} \tag{7.352}$$

with Z_A as the nuclear charge of nucleus A and $\vec{R}_{AN}$ as a vector of magnitude R_{AN} that points from nucleus A to nucleus N. $\mathbf{P}$ is the first order density matrix.

NOTE:

- Hyperfine and quadrupole couplings are properties where the standard basis sets that have been designed for geometry optimization and the like may not be entirely satisfactory (especially for atoms heavier than Ne). You should probably look into tailoring the basis set according to your needs. While it is likely that a later release will provide one or two special basis sets for "core-property" calculations at this time you have to make sure yourself that the basis set has enough flexibility in the core region, for example by uncontracting core basis functions and adding s-primitives with large exponents (or using the "decontraction feature", section *Choice of Basis Set*). If you add these tight functions and use DFT make sure that the numerical integration is still satisfactory. Use the "SpecialGrid" feature to enlarge grids for individual atoms without increasing the computational effort too drastically.
- For heavy nuclei you may want to consider the possibility of relativistic effects. Scalar relativistic effects can be handled with several quasi-relativistic Hamiltonians in ORCA, an overview of the possibilities and some recommendations can be found in *Relativistic Calculations*. Note that relativistic calculations may have special requirements on basis sets, and in the context of property calculations, you should be especially aware of the importance of using picture change corrections (see *Relativistic Calculations* and *Picture-Change Effects*). In quasi-relativistic calculations with DFT, one should also be very cautious about accuracy of the numerical integration, especially for heavier (transition metal) nuclei.

Second order HFCs require the calculation of the spin-orbit coupling contributions which in turn requires the calculation of the coupled perturbed SCF equations. These effects can be quite significant for heavier nuclei and should definitely be included for transition metal complexes. The spin-orbit coupling treatment used is the same as described under *The Spin-Orbit Coupling Operator*.

The g-Tensor

The EPR g-tensor is a property that can be calculated as a second derivative of the energy and it is implemented as such in ORCA for the SCF methods—e.g. HF and DFT—, CASSCF, as well as all-electron MP2 (or RI-MP2) and double-hybrid DFT. The following four contributions arise for the g-tensor (SZ = spin Zeeman, RMC = relativistic mass correction, DSO = diamagnetic spin-orbit correction, PSO = orbital Zeeman/SOC term):

$$g_{\mu\nu}^{(\mathrm{SZ})} = \delta_{\mu\nu} g_e \tag{7.353}$$

$$g_{\mu\nu}^{(\mathrm{RMC})} = -\frac{\alpha^2 g_e}{2S} \sum_{k,l} P_{kl}^{\alpha-\beta} \left\langle \phi_k \left| \hat{T} \right| \phi_l \right\rangle \tag{7.354}$$

$$g_{\mu\nu}^{(\mathrm{DSO})} = \frac{\alpha^2 g_e}{4S} \sum_{k,l} P_{kl}^{\alpha-\beta} \left\langle \phi_k \left| \sum_A \xi\left(r_A\right) \left[\mathbf{r}_A \mathbf{r}_O - \mathbf{r}_{A,\mu} \mathbf{r}_{O,\nu}\right] \right| \phi_l \right\rangle \tag{7.355}$$

$$g_{\mu\nu}^{(\mathrm{PSO})} = -\frac{g_e}{2S} \sum_{k,l} \frac{\partial P_{kl}^{\alpha-\beta}}{\partial B_\mu} \left\langle \phi_k \left| h_\nu^{\mathrm{SOC}} \right| \phi_l \right\rangle \tag{7.356}$$

Here, g_e is the free-electron g-value (=2.002319…), S is the total spin, α the fine structure constant, $P^{\alpha-\beta}$ is the spin density matrix, $\phi$is the basis set, $\hat{T}$ is the kinetic energy operator, $\xi\left(r_A\right)$ an approximate radial operator, h^{SOC} the spatial part of an effective one-electron spin-orbit operator and B_μ is a component of the magnetic field. The calculation of the derivative of the spin-density depends on the chosen level of theory. For the SCF-level it is done based on the coupled-perturbed SCF theory with respect to a magnetic field perturbation.

Accuracy. g-tensor calculations at the SCF level are not highly demanding in terms of basis set size. Basis sets that give reliable SCF results (at least valence double-zeta plus polarization) usually also give reliable g-tensor results. For many molecules the Hartree-Fock approximation will give reasonable predictions. In a number of cases, however, it breaks down completely. DFT is more robust in this respect and the number of molecules where it fails is much smaller. Among the density functionals, the hybrid functionals seem to be the most accurate. In my hands PBE0 is perhaps the best although PWP1 and B3LYP are not much worse. The GGA functionals such as BP, PW91, BLYP or PBE are equally good for small radicals but are significantly inferior to their hybrid counterparts for transition metal complexes.

Gauge dependence. Unfortunately, the g-tensor is a gauge dependent property, i.e. the results depend on where the origin is chosen within the molecule. Unless fully invariant procedures (such as GIAOs) are used, this undesirable aspect is always present in the calculations. GIAOs are now available for calculations on the SCF-level in ORCA. However, if the choice of gauge origin is not outrageously poor, the gauge dependence is usually so small that it can be ignored for all practical purposes, especially if large basis sets are used. ORCA gives you considerable freedom in the choice of gauge origin. It can either be the center of mass, the center of nuclear charge, the center of electronic charge, GIAOs (recommended if available), a special atom or a user-defined point in space. It is wise to check the sensitivity of the results with respect to the choice of origin, especially when small g-shifts on the order of only a few hundred ppm are calculated.

Spin-orbit coupling operator. In previous versions of the code, the g-tensor module used the parameterization of Koseki *et al.* [462, 463, 464] for the spin-orbit operator. This is expected to be a reasonable approximation for the 2p and 3d elements and less satisfactory for heavier main group or transition metal containing systems. Thus, the main target molecules with the simple operators are radicals made of light atoms and first row transition metal complexes. More accurate SOC operators (at only moderately increased computational cost) have now been implemented and are described in section *The Spin-Orbit Coupling Operator*. With these operators there are fewer restrictions. However, for very heavy elements they will suffer from the shortcomings of the Breit-Pauli approximation and future releases will modify these operators to take into account the ZORA or DKH corrections to the SOC.

Zero-Field-Splitting

It is well known that the ZFS consists of a first order term arising from the direct spin-spin interaction[364]:

$$D_{KL}^{(\mathrm{SS})} = \frac{1}{2}\frac{\alpha^2}{S\left(2S-1\right)}\left\langle 0SS\left|\sum_i\sum_{j\neq i}\frac{r_{ij}^2\delta_{KL}-3\left(\mathbf{r}_{ij}\right)_K\left(\mathbf{r}_{ij}\right)_L}{r_{ij}^5}\left\{2\hat{s}_{zi}\hat{s}_{zj}-\hat{s}_{xi}\hat{s}_{xj}-\hat{s}_{yi}\hat{s}_{yj}\right\}\right|0SS\right\rangle \quad (7.357)$$

(K,L =x,y,z). Here α is the fine structure constant ($\approx$ 1/137 in atomic units), $\mathbf{r}_{ij}$ is the electronic distance vector with magnitude r_{ij} and $\hat{s}_i$ is the spin-vector operator for the i'th electron. $|0SS\rangle$ is the exact ground state eigenfunction of the Born-Oppenheimer Hamiltonian with total spin S and projection quantum number $M_S = S$. Since the spin-spin interaction is of first order, it presents no particular difficulties. The more complicated contribution to the **D**-tensor arises from the spin-orbit interaction, which gives a second order contribution. Under the assumption that the spin-orbit coupling (SOC) operator can to a good approximation be represented by an effective one-electron operator ($\hat{H}_{\mathrm{SOC}} = \sum_i \hat{\mathbf{h}}_i^{\mathrm{SOC}}\hat{\mathbf{s}}_i$), ref [621] has derived the following sum-over-states (SOS) equations for the SOC contribution to the ZFS tensor:

$$D_{KL}^{\mathrm{SOC}-(0)} = -\frac{1}{S^2}\sum_{b(S_b=S)}\Delta_b^{-1}\left\langle 0^{SS}\left|\sum_i\hat{h}_i^{K;\mathrm{SOC}}\hat{s}_{i,0}\right|b^{SS}\right\rangle\left\langle b^{SS}\left|\sum_i\hat{h}_i^{L;\mathrm{SOC}}\hat{s}_{i,0}\right|0^{SS}\right\rangle \quad (7.358)$$

$$D_{KL}^{\mathrm{SOC}-(-1)} = -\frac{1}{S\left(2S-1\right)}\sum_{b(S_b=S-1)}\Delta_b^{-1}\left\langle 0^{SS}\left|\sum_i\hat{h}_i^{K;\mathrm{SOC}}\hat{s}_{i,+1}\right|b^{S-1S-1}\right\rangle\left\langle b^{S-1S-1}\left|\sum_i\hat{h}_i^{L;\mathrm{SOC}}\hat{s}_{i,-1}\right|0^{SS}\right\rangle \quad (7.359)$$

$$D_{KL}^{\mathrm{SOC}-(+1)} = -\frac{1}{(S+1)(2S+1)}\cdot \sum_{b(S_b=S+1)}\Delta_b^{-1}\left\langle 0^{SS}\left|\sum_i\hat{h}_i^{K;\mathrm{SOC}}\hat{s}_{i,-1}\right|b^{S+1S+1}\right\rangle\left\langle b^{S+1S+1}\left|\sum_i\hat{h}_i^{L;\mathrm{SOC}}\hat{s}_{i,+1}\right|0^{SS}\right\rangle \quad (7.360)$$

Here the one-electron spin-operator for electron i has been written in terms of spherical vector operator components $s_{i,m}$ with $m = 0, \pm 1$ and $\Delta_b = E_b - E_0$ is the excitation energy to the excited state multiplet $\left|b^{SS}\right\rangle$ (all M_S components are degenerate at the level of the BO Hamiltonian).

One attractive possibility is to represent the SOC by the spin-orbit mean-field (SOMF) method developed by Hess et al.,[390] widely used in the AMFI program by Schimmelpfennig [761] and discussed in detail by Berning et al.[97] as well as in ref. [610]. In terms of an (orthonormal) one-electron basis, the matrix elements of the SOMF operator are:

$$\begin{aligned} h_{rs}^{K;\mathrm{SOC}} &= \left(p\left|\hat{h}_K^{1el-\mathrm{SOC}}\right|q\right) \\ &+\sum_{rs}P_{rs}\left[\left(pq\left|\hat{g}_K^{\mathrm{SOC}}\right|rs\right)-\tfrac{3}{2}\left(pr\left|\hat{g}_K^{\mathrm{SOC}}\right|sq\right)-\tfrac{3}{2}\left(sq\left|\hat{g}_K^{\mathrm{SOC}}\right|pr\right)\right] \end{aligned} \quad (7.361)$$

and:

$$\hat{h}_k^{1el-\mathrm{SOC}}\left(\mathbf{r}_i\right) = \frac{\alpha^2}{2}\sum_i\sum_A Z_A r_{iA}^{-3}\hat{\mathbf{l}}_{iA;k} \quad (7.362)$$

$$\hat{g}_k^{\mathrm{SOC}}\left(\mathbf{r}_i,\mathbf{r}_j\right) = -\frac{\alpha^2}{2}\hat{\mathbf{l}}_{ij;k}r_{ij}^{-3} \quad (7.363)$$

$\hat{\mathbf{l}}_{iA} = \left(\hat{\mathbf{r}}_i - \mathbf{R}_A\right)\times\hat{\mathbf{p}}_i$ is the angular momentum of the i'th electron relative to nucleus A. The vector $\hat{\mathbf{r}}_{iA} = \hat{\mathbf{r}}_i - \mathbf{R}_A$ of magnitude r_{iA} is the position of the i'th electron relative to atom A. Likewise, the vector $\hat{\mathbf{r}}_{ij} = \hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j$ of magnitude r_{ij} is the position of the ith electron relative to electron j and $\hat{\mathbf{l}}_{ij} = \left(\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j\right)\times\hat{\mathbf{p}}_i$ is its angular momentum relative to this electron. **P** is the charge density matrix of the electron ground state ($P_{pq} = \left\langle 0SS\left|E_q^p\right|0SS\right\rangle$with $E_q^p = a_{p\beta}^+a_{q\beta} + a_{p\alpha}^+a_{q\alpha}$ where $a_{p\sigma}^+$ and $a_{q\sigma}$ are the usual Fermion creation and annihilation operators).

General Treatment of ZFS

The zero-field splitting (ZFS) is typically the leading term in the Spin-Hamiltonian (SH) for transition metal complexes with a total ground state spin S>1/2 (for reviews and references see chapter *Publications Related to ORCA*). Its net effect is to introduce a splitting of the $2S+1$ M_S levels (which are exactly degenerate at the level of the Born-Oppenheimer Hamiltonian), even in the absence of an external magnetic field. Thus, an analysis and interpretation of the ZFS is imperative if the information content of the various physical methods that are sensitive to ZFS effects.

In 2007, we have developed a procedure that makes the ZFS calculation compatible with the language of analytic derivatives.[613] Perhaps the most transparent route is to start from the exact solutions of the Born-Oppenheimer Hamiltonian. To this end, we look at the second derivative of the ground state energy ($E = \left\langle 0^{SS} \middle| \hat{H} \middle| 0^{SS} \right\rangle$) with respect to a spin-dependent one-electron operator of the general form:

$$\hat{h}^{K;(m)} = x_K^{(m)} \sum_{pq} h_{pq}^K \hat{S}_{pq}^{(m)} \tag{7.364}$$

Where h_{pq}^K is the matrix of the K'th component of the spatial part of the operator (assumed to be imaginary Hermitian as is the case for the spatial components of the SOC operator) and $\hat{S}_{pq}^{(m)}$ is the second quantized form of the spin vector operator ($m = 0, \pm 1$). The quantity $x_K^{(m)}$ is a formal perturbation parameter. Using the exact eigenfunctions of the BO operator, the first derivative is:

$$\left.\frac{\partial E}{\partial x_K^{(m)}}\right|_{x_K^{(m)}=0} = \sum_{pq} h_{pq}^K P_{pq}^{(m)} \tag{7.365}$$

With the components of the spin density:

$$P_{pq}^{(m)} = \left\langle 0^{SS} | \hat{S}_{pq}^{(m)} | 0^{SS} \right\rangle \tag{7.366}$$

The second derivative with respect to a second component for $m' = -m$ is:

$$\left.\frac{\partial^2 E}{\partial x_K^{(m)} \partial x_L^{(-m)}}\right|_{x_K^{(m)}=x_L^{(-m)}=0} = \sum_{pq} h_{pq}^K \frac{\partial P_{pq}^{(m)}}{x_L^{(-m)}} \tag{7.367}$$

The derivative of the spin density may be written as:

$$\frac{\partial P_{pq}^{(m)}}{x_L^{(-m)}} = \left\langle 0_L^{SS(-m)} | \hat{S}_{pq}^{(m)} | 0^{SS} \right\rangle + \left\langle 0^{SS} | \hat{S}_{pq}^{(m)} | 0_L^{SS(-m)} \right\rangle \tag{7.368}$$

Expanding the perturbed wavefunction in terms of the unperturbed states gives to first order:

$$\left| 0_L^{SS(-m)} \right\rangle = -\sum n \neq 0 \Delta_n^{-1} |n\rangle \left\langle n \middle| \hat{h}^{L;(-m)} \middle| 0^{SS} \right\rangle \tag{7.369}$$

Where $|n\rangle$ is any of the $\left| b^{S'M'} \right\rangle$. Thus, one gets:

$$\frac{\partial^2 E}{\partial x_K^{(m)} \partial x_L^{(-m)}} = \sum_{pq} h_{pq}^K \frac{\partial P_{pq}^{(m)}}{x_L^{(-m)}} \tag{7.370}$$

$$= -\sum_{n \neq 0} \Delta_n^{-1} \left[\left\langle 0^{SS} | \hat{h}^{L;(-m)} | n \right\rangle \left\langle n | \hat{h}^{K;(m)} | 0^{SS} \right\rangle + \left\langle 0^{SS} | \hat{h}^{K;(m)} | n \right\rangle \left\langle n | \hat{h}^{L;(-m)} | 0^{SS} \right\rangle \right] \tag{7.371}$$

The equality holds for exact states. For approximate electronic structure treatments, the analytic derivative approach is more attractive since an infinite sum over states can never be performed in practice and the calculation of analytic derivative is computationally less demanding than the calculation of excited many electron states.

Using eq. (7.370), the components of the SOC-contribution to the **D**-tensor are reformulated as

$$D_{KL}^{\mathrm{SOC}-(0)} = \frac{1}{2S^2} \sum_{pq} h_{pq}^{K;\mathrm{SOC}} \frac{\partial P_{pq}^{(0)}}{\partial x_L^{(0)}} \tag{7.372}$$

$$D_{KL}^{\mathrm{SOC}-(-1)} = \frac{1}{S\left(2S-1\right)} \sum_{pq} h_{pq}^{K;\mathrm{SOC}} \frac{\partial P_{pq}^{(+1)}}{\partial x_L^{(-1)}} \tag{7.373}$$

$$D_{KL}^{\mathrm{SOC}-(+1)} = \frac{1}{\left(S+1\right)\left(2S+1\right)} \sum_{pq} h_{pq}^{K;\mathrm{SOC}} \frac{\partial P_{pq}^{(-1)}}{\partial x_L^{(+1)}} \tag{7.374}$$

These are general equations that can be applied together with any non-relativistic or scalar relativistic electronic structure method that can be cast in second quantized form. Below, the formalism is applied to the case of a self-consistent field (HF, DFT) reference state.

For DFT or HF ground states, the equations are further developed as follows:

The SCF energy is:

$$E_{\mathrm{SCF}} = V_{\mathrm{NN}} + \left\langle \mathbf{P}\mathbf{h}^{+} \right\rangle + \frac{1}{2} \int\int \frac{\rho\left(\mathbf{r}_1\right)\rho\left(\mathbf{r}_2\right)}{\left|\mathbf{r}_1 - \mathbf{r}_2\right|} d\mathbf{r}_1 d\mathbf{r}_2 - \frac{1}{2} a_{\mathrm{X}} \sum_{\mu\nu\kappa\tau\sigma} P_{\mu\kappa}^{\sigma} P_{\nu\tau}^{\sigma} \left(\mu\nu|\kappa\tau\right) + c_{\mathrm{DF}} E_{\mathrm{XC}}\left[\rho_\alpha, \rho_\beta\right] \tag{7.375}$$

Here V_{NN} is the nuclear repulsion energy and $h_{\mu\nu}$ is a matrix element of the one-electron operator which contains the kinetic energy and electron-nuclear attraction terms ($\langle \mathbf{ab} \rangle$ denotes the trace of the matrix product $\mathbf{ab}$). As usual, the molecular spin-orbitals ψ_p^σ are expanded in atom centered basis functions ($\sigma = \alpha, \beta$):

$$\psi_p^\sigma\left(\mathbf{r}\right) = \sum_\mu c_{\mu p}^\sigma \phi_\mu\left(\mathbf{r}\right) \tag{7.376}$$

with MO coefficients $c_{\mu p}^\sigma$. The two-electron integrals are defined as:

$$\left(\mu\nu|\kappa\tau\right) = \int\int \phi_\mu\left(\mathbf{r}_1\right)\phi_\nu\left(\mathbf{r}_1\right) r_{12}^{-1} \phi_\kappa\left(\mathbf{r}_2\right)\phi_\tau\left(\mathbf{r}_2\right) d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.377}$$

The mixing parameter a_{X} controls the fraction of Hartree-Fock exchange and is of a semi-empirical nature. $E_{\mathrm{XC}}\left[\rho_\alpha, \rho_\beta\right]$ represent the exchange-correlation energy. The parameter c_{DF} is an overall scaling factor that allows one to proceed from Hartree-Fock theory ($a_{\mathrm{X}} = 1, c_{\mathrm{DF}} = 0$) to pure DFT ($a_{\mathrm{X}} = 0, c_{\mathrm{DF}} = 1$) to hybrid DFT ($0 < a_{\mathrm{X}} < 1, c_{\mathrm{DF}} = 1$). The orbitals satisfy the spin-unrestricted SCF equations:

$$F_{\mu\nu}^\sigma = h_{\mu\nu} + \sum_{\kappa\tau} P_{\kappa\tau}\left(\mu\nu|\kappa\tau\right) - a_{\mathrm{X}} P_{\kappa\tau}^\sigma\left(\mu\kappa|\nu\tau\right) + c_{\mathrm{DF}}\left(\mu|V_{XC}^\alpha|\nu\right) \tag{7.378}$$

With $V_{XC}^\sigma = \frac{\delta E_{XC}}{\delta \rho_\sigma(\mathbf{r})}$ and $P_{\mu\nu} = P_{\mu\nu}^\alpha + P_{\mu\nu}^\beta$ being the total electron density. For the SOC perturbation it is customary to regard the basis set as perturbation independent. In a spin-unrestricted treatment, the first derivative is:

$$\frac{\partial E_{\mathrm{SCF}}}{\partial x_K^{(m)}} = \sum_{i_\alpha} \left(i_\alpha|h^K s_m|i_\alpha\right) + \sum_{i_\beta} \left(i_\beta|h^K s_m|i_\beta\right) = 0 \tag{7.379}$$

For the second derivative, the perturbed orbitals are required. However, in the presence of a spin-dependent perturbation they can no longer be taken as pure spin-up or spin-down orbitals. With respect to the L'th component of the perturbation for spin-component m, the orbitals are expanded as:

$$\psi_i^{\alpha;(m)L}\left(\mathbf{r}\right) = \sum_{a_\alpha} U_{a_\alpha i_\alpha}^{(m);L} \psi_a^\alpha\left(\mathbf{r}\right) + \sum_{a_\beta} U_{a_\beta i_\alpha}^{(m);L} \psi_a^\beta\left(\mathbf{r}\right) \tag{7.380}$$

$$\psi_i^{\beta;(m)L}\left(\mathbf{r}\right) = \sum_{a_\alpha} U_{a_\alpha i_\beta}^{(m);L} \psi_a^\alpha\left(\mathbf{r}\right) + \sum_{a_\beta} U_{a_\beta i_\beta}^{(m);L} \psi_a^\beta\left(\mathbf{r}\right) \tag{7.381}$$

Since the matrix elements of the spin-vector operator components are purely real and the spatial part of the SOC operator has purely imaginary matrix elements, it follows that the first order coefficients are purely imaginary. The

second derivative of the total SCF energy becomes:

$$\begin{aligned}\frac{\partial^2 E_{\text{SCF}}}{\partial x_K^{(m)} \partial x_L^{(-m)}} = &\sum_{i_\alpha a_\alpha} \left\{ U^{(-m);L*}_{a_\alpha i_\alpha} \left(a_\alpha | h^K s_m | i_\alpha\right) + U^{(-m);L}_{a_\alpha i_\alpha} \left(i_\alpha | h^K s_m | a_\alpha\right) \right\} \\ &+ \sum_{i_\alpha a_\beta} \left\{ U^{(-m);L*}_{a_\beta i_\alpha} \left(a_\beta | h^K s_m | i_\alpha\right) + U^{(-m);L}_{a_\beta i_\alpha} \left(i_\alpha | h^K s_m | a_\beta\right) \right\} \\ &+ \sum_{i_\beta a_\alpha} \left\{ U^{(-m);L*}_{a_\alpha i_\beta} \left(a_\alpha | h^K s_m | i_\beta\right) + U^{(-m);L}_{a_\alpha i_\beta} \left(i_\beta | h^K s_m | a_\alpha\right) \right\} \\ &+ \sum_{i_\beta a_\beta} \left\{ U^{(-m);L*}_{a_\beta i_\beta} \left(a_\beta | h^K s_m | i_\beta\right) + U^{(-m);L}_{a_\beta i_\beta} \left(i_\beta | h^K s_m | a_\beta\right) \right\}\end{aligned} \tag{7.382}$$

Examination of the three cases $m = 0, \pm 1$ leads to the following equations for the **D**-tensor components:

$$D^{(0)}_{KL} = -\frac{1}{4S^2} \sum_{\mu\nu} \frac{\partial P^{(0)}_{\mu\nu}}{\partial x^{(0)}_L} \left(\mu | h^{K;\text{SOC}} | \nu\right) \tag{7.383}$$

$$D^{(+1)}_{KL} = \frac{1}{2\left(S+1\right)\left(2S+1\right)} \sum_{\mu\nu} \left(\mu | h^{K;\text{SOC}} | \nu\right) \frac{\partial P^{(-1)}_{\mu\nu}}{\partial x^{(+1)}_L} \tag{7.384}$$

$$D^{(-1)}_{KL} = \frac{1}{2S\left(2S-1\right)} \sum_{\mu\nu} \left(\mu | h^{K;\text{SOC}} | \nu\right) \frac{\partial P^{(+1)}_{\mu\nu}}{\partial x^{(-1)}_L} \tag{7.385}$$

Where a special form of the perturbed densities has been chosen. They are given in the atomic orbital basis as:

$$\frac{\partial P^{(0)}_{\mu\nu}}{\partial x^{(0)}_L} = \sum_{i_\alpha a_\alpha} U^{(0);L}_{a_\alpha i_\alpha} c^\alpha_{\mu i} c^\alpha_{\nu a} + \sum_{i_\beta a_\beta} U^{(0);L}_{a_\beta i_\beta} c^\beta_{\mu i} c^\beta_{\nu a} \tag{7.386}$$

$$\frac{\partial P^{(+1)}_{\mu\nu}}{\partial x^{(-1)}_L} = \sum_{i_\alpha a_\beta} U^{(-1);L}_{a_\beta i_\alpha} c^\alpha_{\mu i} c^\beta_{\nu a} - \sum_{i_\beta a_\alpha} U^{(-1);L}_{a_\alpha i_\beta} c^\alpha_{\mu a} c^\beta_{\nu i} \tag{7.387}$$

$$\frac{\partial P^{(-1)}_{\mu\nu}}{\partial x^{(+1)}_L} = -\sum_{i_\alpha a_\beta} U^{(+1);L}_{a_\beta i_\alpha} c^\beta_{\mu a} c^\alpha_{\nu i} + \sum_{i_\beta a_\alpha} U^{(+1);L}_{a_\alpha i_\beta} c^\beta_{\mu i} c^\alpha_{\nu a} \tag{7.388}$$

The special form of the coupled perturbed equations are implemented in ORCArun as follows: The perturbation is of the general form $h^K \hat{s}_m$. The equations (7.383)-(7.388) and (7.389)-(7.394) below have been written in such a way that the spin integration has been performed but that the spin-dependent factors have been dropped from the right-hand sides and included in the prefactors of eqs. (7.383)-(7.385). The explicit forms of the linear equations to be solved are:

$m = 0$:

$$\left(\varepsilon^{(0)}_{a_\alpha} - \varepsilon^{(0)}_{i_\alpha}\right) U^{K(0)}_{a_\alpha i_\alpha} + a_\text{X} \sum_{j_\alpha b_\alpha} U^{K(m)}_{b_\alpha j_\alpha} \left\{ \left(b_\alpha i_\alpha | a_\alpha j_\alpha\right) - \left(j_\alpha i_\alpha | a_\alpha b_\alpha\right) \right\} = -\left(a_\alpha | h^K | i_\alpha\right) \tag{7.389}$$

$$\left(\varepsilon^{(0)}_{a_\beta} - \varepsilon^{(0)}_{i_\beta}\right) U^{K(0)}_{a_\beta i_\beta} + a_\text{X} \sum_{j_\beta b_\beta} U^{K(m)}_{b_\beta j_\beta} \left\{ \left(b_\beta i_\beta | a_\beta j_\beta\right) - \left(j_\beta i_\beta | a_\beta b_\beta\right) \right\} = -\left(a_\beta | h^K | i_\beta\right) \tag{7.390}$$

$m = +1$:

$$\left(\varepsilon^{(0)}_{a_\alpha} - \varepsilon^{(0)}_{i_\beta}\right) U^{K(+1)}_{a_\alpha i_\beta} + a_\text{X} \sum_{j_\alpha b_\alpha} U^{K(+1)}_{b_\beta j_\alpha} \left(b_\beta i_\beta | a_\alpha j_\alpha\right) - a_\text{X} \sum_{b_\alpha j_\beta} U^{K(+1)}_{b_\beta j_\alpha} \left(j_\beta i_\beta | a_\alpha b_\alpha\right) = -\left(a_\alpha | h^K | i_\beta\right) \tag{7.391}$$

$$\left(\varepsilon^{(0)}_{a_\beta} - \varepsilon^{(0)}_{i_\alpha}\right) U^{K(+1)}_{a_\beta i_\alpha} + a_\text{X} \sum_{j_\beta b_\alpha} U^{K(+1)}_{b_\alpha j_\beta} \left(b_\alpha i_\alpha | a_b j_\beta\right) - a_\text{X} \sum_{b_\beta j_\alpha} U^{K(+1)}_{b_\beta j_\alpha} \left(j_\alpha i_\alpha | a_\beta b_\beta\right) = 0 \tag{7.392}$$

$m = -1$:

$$\left(\varepsilon^{(0)}_{a_\beta} - \varepsilon^{(0)}_{i_\alpha}\right) U^{K(-1)}_{a_\beta i_\alpha} + a_\text{X} \sum_{j_\beta b_\alpha} U^{K(-1)}_{b_\alpha j_\beta} \left(b_\alpha i_\alpha | a_b j_\beta\right) - a_\text{X} \sum_{b_\beta j_\alpha} U^{K(-1)}_{b_\beta j_\alpha} \left(j_\alpha i_\alpha | a_\beta b_\beta\right) = -\left(a_\beta | h^K | i_\alpha\right) \tag{7.393}$$

$$\left(\varepsilon_{a_\alpha}^{(0)} - \varepsilon_{i_\beta}^{(0)}\right) U_{a_\alpha i_\beta}^{K(-1)} + a_X \sum_{j_\alpha b_\alpha} U_{b_\beta j_\alpha}^{K(-1)} \left(b_\beta i_\beta | a_\alpha j_\alpha\right) - a_X \sum_{b_\alpha j_\beta} U_{b_\beta j_\alpha}^{K(-1)} \left(j_\beta i_\beta | a_\alpha b_\alpha\right) = 0 \tag{7.394}$$

Note that these coupled-perturbed (CP) equations contain no contribution from the Coulomb potential or any other local potential such as the exchange-correlation potential in DFT. Hence, in the absence of HF exchange, the equations are trivially solved:

$$U_{a_\alpha i_\alpha}^{K(0)} = -\frac{\left(a_\alpha | h^K | i_\alpha\right)}{\varepsilon_{a_\alpha}^{(0)} - \varepsilon_{i_\alpha}^{(0)}} \tag{7.395}$$

$$U_{a_\beta i_\beta}^{K(0)} = -\frac{\left(a_\beta | h^K | i_\beta\right)}{\varepsilon_{a_\beta}^{(0)} - \varepsilon_{i_\beta}^{(0)}} \tag{7.396}$$

$$U_{a_\alpha i_\beta}^{K(+1)} = -\frac{\left(a_\alpha | h^K | i_\beta\right)}{\varepsilon_{a_\alpha}^{(0)} - \varepsilon_{i_\beta}^{(0)}} \tag{7.397}$$

$$U_{a_\beta i_\alpha}^{K(+1)} = 0 \tag{7.398}$$

$$U_{a_\beta i_\alpha}^{K(-1)} = -\frac{\left(a_\beta | h^K | i_\alpha\right)}{\varepsilon_{a_\beta}^{(0)} - \varepsilon_{i_\alpha}^{(0)}} \tag{7.399}$$

$$U_{a_\alpha i_\beta}^{K(-1)} = 0 \tag{7.400}$$

It is interesting that the "reverse spin flip coefficients" $U_{a_\beta i_\alpha}^{K(+1)}$ and $U_{a_\alpha i_\beta}^{K(-1)}$ are only nonzero in the presence of HF exchange. In a perturbation expansion of the CP equations they arise at second order ($V^2/\Delta\varepsilon^2$) while the other coefficients are of first order ($V/\Delta\varepsilon$; V represents the matrix elements of the perturbation). Hence, these contributions are of the order of α^4 and one could conceive dropping them from the treatment in order to stay consistently at the level of α^2. These terms were nevertheless kept in the present treatment.

Equations (7.389)-(7.394) are referred to as CP-SOC (coupled-perturbed spin-orbit coupling) equations. They can be solved by standard techniques and represent the desired analogue of the CP-SCF magnetic response equations solved for the determination of the g-tensor and discussed in detail earlier [663]. It is readily confirmed that in the absence of HF exchange, eqs. (7.395)-(7.400) inserted into eqs. (7.383)-(7.388) lead back to a modified Pederson-Khanna type treatment of the SOC contributions to the **D**-tensor [661]. In the framework of the formalism developed above, the Pederson-Khanna formula can be re-written in the form:

$$\begin{aligned} D_{KL}^{(\mathrm{SOC;PK})} = &\frac{1}{4S^2} \sum_{i_\beta, a_\beta} \left(\psi_i^\beta \left| h^{K;\mathrm{SOC}} \right| \psi_a^\beta\right) U_{a_\beta i_\beta}^{L;(0)} + \frac{1}{4S^2} \sum_{i_\alpha, a_\alpha} \left(\psi_i^\alpha \left| h^{K;\mathrm{SOC}} \right| \psi_a^\alpha\right) U_{a_\alpha i_\alpha}^{L;(0)} \\ &- \frac{1}{4S^2} \sum_{i_\alpha, a_\beta} \left(\psi_i^\alpha \left| h^{K;\mathrm{SOC}} \right| \psi_a^\beta\right) U_{a_\beta i_\alpha}^{L;(-1)} - \frac{1}{4S^2} \sum_{i_\beta, a_\alpha} \left(\psi_i^\alpha \left| h^{K;\mathrm{SOC}} \right| \psi_a^\alpha\right) U_{a_\alpha i_\beta}^{L;(+1)} \end{aligned} \tag{7.401}$$

This equation was derived from second-order non-self-consistent perturbation theory without recourse to spin-coupling. For the special case of no Hartree-Fock exchange, the main difference to the treatment presented here is that the correct prefactors from eqs. (7.372)-(7.374) occur in front of the spin-flip contributions rather than $\pm$ $1/(4S^2)$ in eq. (7.401). In the presence of HF exchange it is suggested that the consistent generalization of the PK method are eqs. (7.383)-(7.385) with the $\pm$ $1/(4S^2)$ prefactors and this way the method has been implemented as an option into the ORCA program.

For completeness, the evaluation of the spin-spin term in the SCF case proceeds conveniently through:

$$D_{KL}^{(\mathrm{SS})} = \frac{g_e^2}{4} \frac{\alpha^2}{S\left(2S-1\right)} \sum_{\mu\nu} \sum_{\kappa\tau} \left\{P_{\mu\nu}^{\alpha-\beta} P_{\kappa\tau}^{\alpha-\beta} - P_{\mu\kappa}^{\alpha-\beta} P_{\nu\tau}^{\alpha-\beta}\right\} \left\langle \mu\nu \left| r_{12}^{-5} \left\{3 r_{12,K} r_{12,L} - \delta_{KL} r_{12}^2\right\} \right| \kappa\tau \right\rangle \tag{7.402}$$

as derived by McWeeny and Mizuno and discussed in some detail by Sinnecker and Neese.[799] In this reference it was found that DFT methods tend to overestimate the spin-spin contribution if the calculations are based on a spin-unrestricted SCF treatment. A much better correlation with experiment was found for open-shell spin restricted calculations. The origin of this effect proved to be difficult to understand but it was shown in ref [354] that in the case of small spin-contamination, the results of ROKS calculations and of those that are obtained on the basis of the spin-unrestricted natural orbital (UNO) determinant are virtually indistinguishable. It is therefore optionally possible in the ORCA program to calculate the spin-spin term on the basis of the UNO determinant.

Spin-rotation constants

Spin-rotation constant calculations are implemented using perturbation-dependent atomic orbitals following [290]. As given in eq. 34 of that reference, the spin-rotation tensor of nucleus K, $\mathbf{M}_K$, is related to the nuclear shielding tensor computed with GIAOs, $\boldsymbol{\sigma}_K^{\mathrm{GIAO}}$, and the diamagnetic part of the shielding tensor with the gauge origin set at that nucleus, $\boldsymbol{\sigma}_K^{\mathrm{dia}}(\mathbf{R}_K)$:

$$\mathbf{M}_K = 2\gamma_K \left(\boldsymbol{\sigma}_K^{\mathrm{GIAO}} - \boldsymbol{\sigma}_K^{\mathrm{dia}}(\mathbf{R}_K)\right) \mathbf{I}^{-1} - \mathbf{M}_K^{\mathrm{nuc}}$$

where $\mathbf{M}_K^{\mathrm{nuc}}$ is the nuclear component (given in eq. 14 of the reference), $\mathbf{I}$ is the inertia tensor, and γ_K is the nuclear magnetogyric ratio. Accordingly, upon requesting spin-rotation constants, ORCA automatically computes the NMR shieldings with GIAOs as well. The following input shows an example calculation of $\mathbf{M}(^{17}\mathrm{O})$ in $\mathrm{H_2}^{12}\mathrm{C}^{17}\mathrm{O}$:

```
! HF pcSseg-1 Mass2016 Bohrs
*xyz 0 1
     O        -0.00000000    -0.00000000     1.13863731 M=16.999131
     C        -0.00000000    -0.00000000    -1.14131773 M=12.0
     H        -0.00000000     1.76770755    -2.24076285
     H         0.00000000    -1.76770755    -2.24076285
*
%eprnmr
  Nuclei = all O {srot, ist=17}
end
```

Note

- The magnetogyric ratio used can be changed either by choosing the correct isotope via `ist`, or by providing the nuclear g-factor directly via `ssgn`.
- The masses used to compute the inertia tensor are independent of the chosen isotopes! The example above requests atomic masses of the most abundant isotope (via the `Mass2016` keyword) and explicitly specifies those of ^{12}C (which is the default) and ^{17}O.

Cartesian Index Conventions for EPR and NMR Tensors

The NMR shielding tensor $\boldsymbol{\sigma}$ and the EPR g and A tensors are in general nonsymmetric matrices. It is therefore important to know the conventions used with regard to their cartesian indices. These conventions stipulate the order of the vector–matrix–vector multiplications in the respective spin Hamiltonians. Unless stated otherwise, ORCA adopts the following conventions:

For the NMR shielding tensor the nuclear Zeeman Hamiltonian assumes the form:

$$H_I = -g_N \beta_N \mathbf{B}(\mathbf{1} - \boldsymbol{\sigma})\mathbf{I},$$

where $\mathbf{B}$ is the applied magnetic field vector.

For the EPR g and A tensors the EPR spin Hamiltonian assumes the form:

$$H_S = \beta_e \mathbf{B}\mathbf{g}\mathbf{S} + \mathbf{S}\mathbf{A}\mathbf{I}. \tag{7.403}$$

MP2 level magnetic properties

Presently, hyperfine couplings (excluding the A_{orb} term), g-tensors, and shielding tensors without GIAOs can be calculated for both canonical and RI-MP2 and double-hybrid DFT without the frozen core approximation. The A_{orb} term of the hyperfine couplings is available only for RI-MP2 and double-hybrid DFT with and without frozen core approximation. In case the RIJCOSX approximation is used, the keywords `Z_GridX`, `Z_GridX_RHS`, `KCOpt`, `KC_GridX` and `KC_IntAccX` are relevant – see sections *RIJCOSX-RI-MP2 Gradients* and *MP2 and RI-MP2 Second Derivatives*. NMR shielding and g-tensor calculations with GIAOs are available for RI-MP2 and double-hybrid DFT with or without a frozen core. The implementation is described in detail in refs [828, 849] and the available options are shown in section *RI-MP2 and Double-Hybrid DFT Response Properties*. Note that for double-hybrid DFT the correct properties are printed after the density heading containing "Method : MP2" and "Level : Relaxed density". DLPNO-MP2 (and double-hybrid) shielding tensors are also available - see section *Local MP2 Response Properties*.

Nucleus-independent chemical shielding

Aromaticity is a fundamental concept in chemistry and much attention has been payed to its analysis in quantum chemistry. One possibility to gain insight into aromaticity is to sample the ring current effect in NMR close to the π-system. As this is not done by inspecting the chemical shielding of any of the atoms, this quantity is called nucleus independent chemical shielding (NICS). Usually, a "dummy" atom is placed in the center of the ring and/or at some distance away from it. In ORCA, one needs to use a ghost atom, e.g. "`H:`" to ensure that the program generates DFT or COSX grid points in the region on interest. For technical reasons, this atom must also have at least one basis function, which can be set with `NewGTO`. An s-function with a sufficiently large exponent will not overlap with any other basis function in the molecule and will thus have no effect on the results, but only satisfy the technical requirement (note that the extra grid points may change some of the calculation results by increasing the accuracy of the numerical integration). If RI is used, a dummy fitting function must also be added to the AuxJ, AuxJK, and/or AuxC basis set. A typical input for benzene looks like this:

```
! TightSCF NMR PBE def2-TZVP RI def2/J
* gzmt 0 1
   H:                              NewGTO S 1 1 1e6 1 end NewAuxJGTO S 1 1 2e6 1 end
   H:  1   1                       NewGTO S 1 1 1e6 1 end NewAuxJGTO S 1 1 2e6 1 end
   C   1   1.39    2   90
   C   1   1.39    2   90  3   60
   C   1   1.39    2   90  4   60
   C   1   1.39    2   90  5   60
   C   1   1.39    2   90  3  -60
   C   1   1.39    2   90  6   60
   H   3   1.09    4   120 1   180
   H   4   1.09    3   120 1   180
   H   5   1.09    4   120 1   180
   H   6   1.09    5   120 1   180
   H   7   1.09    3   120 1   180
   H   8   1.09    7   120 1   180
*
```

Shielding tensor orbital decomposition

It is possible to decompose the NMR shielding tensor into orbital- or orbital pair contributions. One such option is the Natural Chemical Shielding analysis (see Section *Natural Chemical Shielding Analysis (NCS)*), while another is presented here. The shielding tensor for nucleus A can be exactly decomposed as follows:

$$\boldsymbol{\sigma}^{A} = \sum_{p} \left(\boldsymbol{\sigma}_{p}^{A,\text{para}} + \boldsymbol{\sigma}_{p}^{A,\text{dia}} \right)$$

$$\boldsymbol{\sigma}_{i}^{A,\text{para/dia}} = \boldsymbol{\sigma}_{ii}^{A,\text{para/dia}} + \frac{1}{2} \sum_{j \neq i} \left(\boldsymbol{\sigma}_{ij}^{A,\text{para/dia}} + \boldsymbol{\sigma}_{ji}^{A,\text{para/dia}} \right) + \sum_{a} \left(\boldsymbol{\sigma}_{ia}^{A,\text{para/dia}} + \boldsymbol{\sigma}_{ai}^{A,\text{para/dia}} \right)$$

$$\boldsymbol{\sigma}_{a}^{A,\text{para/dia}} = \boldsymbol{\sigma}_{aa}^{A,\text{para/dia}} + \frac{1}{2} \sum_{b \neq a} \left(\boldsymbol{\sigma}_{ab}^{A,\text{para/dia}} + \boldsymbol{\sigma}_{ba}^{A,\text{para/dia}} \right)$$

$$\boldsymbol{\sigma}_{pq}^{A,\text{para}} = D_{pq}^{\mathbf{B}} h_{pq}^{\mathbf{M}_A}$$

$$\boldsymbol{\sigma}_{pq}^{A,\text{dia}} = D_{pq} h_{pq}^{\mathbf{B},\mathbf{M}_A}$$

$$D_{pq} = \sum_{\mu\nu\kappa\lambda} c_{\mu p} S_{\mu\nu} D_{\nu\kappa} S_{\kappa\lambda} c_{\lambda q}$$

$$D_{pq}^{\mathbf{B}} = \sum_{\mu\nu\kappa\lambda} c_{\mu p} S_{\mu\nu} D_{\nu\kappa}^{\mathbf{B}} S_{\kappa\lambda} c_{\lambda q}$$

$$h_{pq}^{\mathbf{M}_A} = \sum_{\mu\nu} c_{\mu p} h_{\mu\nu}^{\mathbf{M}_A} c_{\nu q}$$

$$h_{pq}^{\mathbf{B},\mathbf{M}_A} = \sum_{\mu\nu} c_{\mu p} h_{\mu\nu}^{\mathbf{B},\mathbf{M}_A} c_{\nu q}$$

Note that for SCF methods (HF or DFT), $\boldsymbol{\sigma}_{a}^{A,\text{para/dia}} = 0$. To request the analysis, a valid GBW file must be given with the keyword `LocOrbGBW`, which can contain any orthonormal MOs in the same basis, e.g. canonical or localized, although the decomposition above assumes that the Brillouin condition is fulfilled and may be misleading if performed over NBOs for example. Orbital contributions over 0.01 ppm are printed – this is currently not user-configurable. The separate orbital pair contributions can also be printed using `Printlevel=3`. The following example input calculates HF and RI-MP2 shieldings for formaldehyde and decomposes them over Pipek–Mezey localized orbitals (note that the virtual orbitals are likely not well localized - `AHFB` would be better suited there). The second sub-calculation just prints the LMOs for convenient visualization with Avogadro.

```
! RI-MP2 NMR TightSCF RIJK pcSseg-1 cc-pwCVDZ/C def2/JK
%base "H2CO_NMR_PM"
%loc
  LocMet PM    # Pipek-Mezey localization
  Occ    true  # localize both occupied
  Virt   true  # and virtual orbitals (better use AHFB for these!)
  T_core -1e5  # including the core
end
%eprnmr
  PrintLevel 3 # print orbital pair contributions (lots of output!)
  LocOrbGBW "H2CO_NMR_PM.loc" # use these orbitals
end
!xyzfile       # store the coordinates
*gzmt 0 1
  O
  C 1 1.2078
  H 2 1.1161 1 121.74
  H 2 1.1161 1 121.74 3 180
*
# read the localized orbitals and print them in the output
# for easy visualization with Avogadro
$new_job
! HF pcSseg-1 NoRI MORead NoIter PrintBasis PrintMOs
%moinp "H2CO_NMR_PM.loc"
*xyzfile 0 1 H2CO_NMR_PM.xyz
```

Treatment of Tau in Meta-GGA Functionals

For GIAO-based calculations with meta-GGAs, different options are available for the kinetic energy density τ. The current-independent τ_0 is not gauge-invariant. Ignoring the terms, which produce the gauge-dependence, leads to an ad-hoc gauge-invariant treatment (this was the default up to ORCA 5). A gauge-invariant definition τ_{MS}, containing an explicit dependence on the magnetic field, was proposed by Maximoff and Scuseria.[566] However, this ansatz produces unphysical paramagnetic contributions to the shielding tensor.[757] The last option, introduced by Dobson,[216] is gauge-invariant but requires the solution of the CP-SCF equations, even for pure density functionals. For a discussion and comparison of these alternatives see refs [79] (in the context of TDDFT) and [713, 757] (in the context of NMR shielding). Note that the calculated shieldings can differ substantially between the different approaches! Some other electronic structure programs use the MS ansatz by default, so be careful when comparing results between different codes. In ORCA the treatment of τ in GIAO-based calculations is chosed as follows

```
%eprnmr
  Tau 0      # gauge-variant
      GI     # ad-hoc gauge-invariant
      MS     # field-dependent, gauge-invariant version of Maximoff and Scuseria
      Dobson # (default) current density-dependent, gauge-invariant version of Dobson
  end
```

7.51.4 Paramagnetic NMR shielding tensors

For systems with spin $S > 0$, the nuclear shielding contains a contribution which arises from the paramagnetism of the unpaired electrons.[1] This contribution is temperature-dependent and is called the "paramagnetic shielding" ($\boldsymbol{\sigma}^{\mathrm{p}}$). It adds to the temperature-indendent contribution to the shielding, also called the "orbital" contribution:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathrm{orb}} + \boldsymbol{\sigma}^{\mathrm{p}}.$$

ORCA currently supports the calculation of $\boldsymbol{\sigma}^{\mathrm{p}}$ for systems whose paramagnetism can be described by the effective EPR spin Hamiltonian

$$H_S = \mathbf{SDS} + \beta_e \mathbf{BgS} + \mathbf{SAI}. \tag{7.404}$$

The theoretical background can be found in Refs. [809, 862]. We reproduce here the main equations.

For a spin state described by Eq. (7.404), the paramagnetic shielding tensor is given by

$$\boldsymbol{\sigma}^{\mathrm{p}} = -\frac{\beta_e S(S+1)}{g_N \beta_N 3kT} \mathbf{gZA}, \tag{7.405}$$

where $\mathbf{Z}$ is a dimensionless 3×3 matrix which is determined by the ZFS and the temperature, as follows: Diagonalization of the ZFS Hamiltonian $\mathbf{SDS}$ yields energy levels E_λ and eigenstates $|S\lambda a\rangle$, where a labels degenerate states if E_λ is degenerate. Then $\mathbf{Z}$ is defined as ($i, j = x, y, z$)

$$Z_{ij} = \frac{3}{S(S+1)} \frac{1}{Q_0} \sum_\lambda e^{-E_\lambda/kT} \left[\sum_{a,a'} \langle S\lambda a|S_i|S\lambda a'\rangle \langle S\lambda a'|S_j|S\lambda a\rangle \right.$$
$$\left. + 2kT \sum_{\lambda' \neq \lambda} \sum_{a,a'} \frac{\langle S\lambda a|S_i|S\lambda' a'\rangle \langle S\lambda' a'|S_j|S\lambda a\rangle}{E_{\lambda'} - E_\lambda} \right],$$

where $Q_0 = \sum_{\lambda',a} e^{-E_\lambda/kT}$ denotes the partition function. An important property of the $\mathbf{Z}$ matrix as defined above is that it goes to the identity matrix as $\mathbf{D}/kT$ goes to zero.

The orbital part of the shielding, $\boldsymbol{\sigma}^{\mathrm{orb}}$, is calculated in the same manner as for closed-shell molecules. It is available in ORCA for the unrestricted HF and DFT methods and for MP2 (see Section *MP2 level magnetic properties* for more information on the latter).

The `orca_pnmr` tool uses Eq. *(7.407)* to calculate $\boldsymbol{\sigma}^{\mathrm{p}}$. Usage of `orca_pnmr` is described in Section *orca_pnmr*.

[1] For a comprehensive review on paramagnetic NMR, see e.g. [657].

7.51.5 Calculating properties from existing densities

Occasionally, one may calculate a density matrix using an expensive correlated method such as CCSD and realize afterwards that a certain property such as the quadrupole moment or a hyperfine coupling constant (HFCC) is also required. Rather than start the whole calculation from scratch, one may wish to use the existing density matrix to calculate the properties. For this purpose, we have experimentally introduced a "properties only" calculation mode, whereby the MOs are read from an existing *BaseName.gbw* file and the densities are read from an existing *BaseName.densities* file and only the property calculations are performed. Note however, that this presents many possibilities for error, so **only use it as a last resort and be very careful with the results**!

Take, for example, this CCSD calculation:

```
! ccsd def2-svp
%base "BO-CCSD"
%mdci density unrelaxed end
*xyz 0 2
B 0 0 0
O 0 0 1.2049
*
```

This produces the files *BO-CCSD.gbw* and *BO-CCSD.densities*. To obtain the CCSD quadrupole moment and HFCCs without repeating the whole calculation, we can copy these two files into a new directory (highly recommended!) and start a second job with the `!PropertiesOnly` keyword. Note that the basename of the second job must be identical!

```
! ccsd def2-svp
%base "BO-CCSD"
%mdci density unrelaxed end
*xyz 0 2
B 0 0 0
O 0 0 1.2049
*
# Everything above must be the same as in the first job!

# Request the property calculations
! PropertiesOnly
%elprop
  quadrupole true
end
%eprnmr
  Nuclei = all B { Aiso, Adip }
  Nuclei = all O { Aiso, Adip }
end
```

7.51.6 Local Energy Decomposition

The DLPNO-CCSD(T) method provides very accurate relative energies and allows to successfully predict many chemical phenomena. In order to facilitate the interpretation of coupled cluster results, we have developed the Local Energy Decomposition (LED) analysis [33, 105, 767], which permit to divide the total DLPNO-CCSD(T) energy (including the reference energy) into physically meaningful contributions. A practical guide to the LED scheme is reported in Section *Local Energy Decomposition*. Examples of applications of this scheme can be found in Ref. [84, 106, 294, 536, 537, 903]

As a word of caution we emphasize that only the total energy is an observable and its decomposition is, to some extent, arbitrary. Nevertheless, the LED analysis appears to be physically well grounded and logical to us, it is straightforward to apply and comes typically at a negligible computational cost compared to DLPNO-CCSD(T) calculations. Starting from ORCA 4.1, the LED scheme is available for both closed shell and open shell calculations. The code has also been made parallel and more efficient.

The LED scheme makes no assumption about the strength of the intermolecular interaction and hence it remains valid and consistent over the entire potential energy surface. Alternative schemes, such as the popular symmetry

adapted perturbation theory, are perturbative in nature and hence are best applied to weakly interacting systems.

The idea of the LED analysis is rather simple. In local correlation methods occupied orbitals are localized and can be readily assigned to pre-defined fragments in the molecule. The same can be done for the correlation energy in terms of pair correlation energies that refer to pairs of occupied orbitals. In this way, both the correlation and the reference energy can be decomposed into intra- and interfragment contributions. The fragmentation is user defined. An arbitrary number of fragments can be defined. In the case that more than 2 fragments are defined, the interfragment interaction is printed for each fragment pair.

A very important feature of the LED scheme is the possibility to distinguish between dispersive and non-dispersive part of the DLPNO-CCSD(T) correlation energy. In brief, we exploit the fact that each CCSD pair correlation energy contribution can be expressed as a sum of double excitations from pairs of occupied orbitals into the virtual space. As in the DLPNO-CCSD(T) scheme the virtual space is spanned by Pair Natural Orbitals(PNOs) that are essentially local, the entire correlation energy can be decomposed into double excitations types, depending on where occupied and virtual orbitals are localized. For each pair of fragments, the sum of all excitations corresponding to the interaction of instantaneous local dipoles located on different fragments defines the so called "London dispersion" attraction between the two fragments in the LED framework.

For a system of two fragments, one can use as a reference point the geometrically and electronically relaxed fragments that constitute the interacting super-molecule. Relative to this reference state, the binding energy between the fragments can be written as:

$$\begin{aligned}\Delta E =&\Delta E_{geo-prep} \\ &+ \Delta E^{ref.}_{el-prep} + E^{ref.}_{elstat} + E^{ref.}_{exch} \\ &+ \Delta E^{C-CCSD}_{non-dispersion} + E^{C-CCSD}_{dispersion} \\ &+ \Delta E^{C-(T)}_{int}\end{aligned} \tag{7.406}$$

where $\Delta E_{geo-prep}$ is the energy needed to distort the fragments from their equilibrium configuration to the interacting geometry (also called "strain" in other energy decomposition schemes). The $\Delta E^{ref.}_{el-prep}$ term represents the electronic preparation energy and describes how much energy is necessary to bring the fragments into the electronic structure that is optimal for interaction. $E^{ref.}_{exch}$ is the inter-fragment exchange interaction (it always gives a binding contribution in our formalism) and $E^{ref.}_{elstat}$ is the electrostatic energy interaction between the distorted electronic clouds of the fragments. The sum of these terms gives the Hatree-Fock energy in the closed shell case and the energy of the QRO determinant in the open shell case. Finally, the correlation energy is decomposed into dispersive $E^{C-CCSD}_{dispersion}$ and non-dispersive $\Delta E^{C-CCSD}_{non-dispersion}$ contributions plus a triples correction term to the interaction energy $\Delta E^{C-(T)}_{int}$.

The $E^{C-CCSD}_{dispersion}$ term contains the London dispersion contribution from the strong pairs described above plus the interfragment component of the weak pairs, which is essentially dispersive in nature. The $\Delta E^{C-CCSD}_{non-dispersion}$ correlation term serves to correct the contributions to the binding energy approximately included in the reference energy, e.g it counteracts the overpolarization typical of the HF method. It contains the so called charge transfer excitations from the strong pairs $E^{CT(X\rightarrow Y)}_{C-SP} + E^{CT(X\leftarrow Y)}_{C-SP}$, which represent the sum of all double excitation contributions that do not conserve the charge within each fragment. Moreover, the non-dispersive term also includes the electronic preparation from strong ($\Delta E^{C-SP}_{el-prep}$) and weak ($\Delta E^{C-WP}_{el-prep}$) pairs. Finally, $\Delta E^{C-(T)}_{int}$ represents the triples correction contribution to the interaction energy between the fragments. In the LED scheme, this term can be further decomposed into intra- and interfragment components. This can be useful, for example, to estimate the London dispersion contribution from the triples correction term, as suggested in ref.[536].

7.52 Natural Bond Orbital (NBO) Analysis

A popular and useful method for population analysis is the natural bond orbital analysis due to Weinhold and co-workers. It is implemented in the **NBO** program which is distributed in older versions via the CCL list and in newer versions via the University of Wisconsin/Madison. Information about the NBO program can be found at http://www.chem.wisc.edu/~nbo7. In order to use it together with ORCA you need a version of the stand-alone executable. Starting with version 3.1.x ORCA can only be used with NBO6 or NBO7. To specify the NBO executable the environment variable NBOEXE=/full/name/of/nbo7-executable has to be set. As the NBO part of the interface is not independent of the integer data-type width (i4 or i8), the NBO executable which will be used together with ORCA has to be compiled using i4!

ORCA features two methods to interface with the **NBO** program: `! NBO` keyword and the `%nbo`-block. The following example illustrates the use for formaldehyde:

```
#
# Test the interface to the NBO program
#
! RHF SVP NBO

* xyz 0 1
  C     0.000000    0.000000    0.000000
  O     1.200000    0.000000    0.000000
  H    -0.550000    0.952628    0.000000
  H    -0.550000   -0.952628   -0.000000
*
```

This produces the following output:

```
Now starting NBO....

 ********************************** NBO 7.0 ***********************************
            N A T U R A L   A T O M I C   O R B I T A L   A N D
          N A T U R A L   B O N D   O R B I T A L   A N A L Y S I S
 ************************ development version (D000000) ************************
  (c) Copyright 1996-2018 Board of Regents of the University of Wisconsin System
      on behalf of the Theoretical Chemistry Institute.  All rights reserved.

          Cite this program [NBO 7.0.0 (15-Nov-2018)] as:

          NBO 7.0.  E. D. Glendening, J. K. Badenhoop, A. E. Reed,
          J. E. Carpenter, J. A. Bohmann, C. M. Morales, P. Karafiloglou,
          C. R. Landis, and F. Weinhold, Theoretical Chemistry Institute,
          University of Wisconsin, Madison, WI (2018)

       /NPA    / : Natural Population Analysis
       /NBO    / : Natural Bond Orbital Analysis
       /AONBO  / : Checkpoint the AO to NBO transformation
       /ARCHIVE/ : Write the archive file to lfn47

 Job title:  ORCA Job: NBO_1


 NATURAL POPULATIONS:  Natural atomic orbital occupancies

  NAO Atom No lang   Type(AO)    Occupancy      Energy
 ----------------------------------------------------------
    1    C  1  s      Cor( 1s)     1.99997     -11.34329
    2    C  1  s      Val( 2s)     1.01533      -0.17540
    3    C  1  s      Ryd( 3s)     0.00701       0.61376
    4    C  1  px     Val( 2p)     0.81697       0.08822
    5    C  1  px     Ryd( 3p)     0.01268       0.63900
```

(continues on next page)

(continued from previous page)

```
  6    C  1  py     Val( 2p)     1.09795      -0.01243
  7    C  1  py     Ryd( 3p)     0.00055       0.80803
  8    C  1  pz     Val( 2p)     0.66003      -0.03464
  9    C  1  pz     Ryd( 3p)     0.00283       0.62824
 10    C  1  dxy    Ryd( 3d)     0.00576       2.75039
 11    C  1  dxz    Ryd( 3d)     0.00375       2.25746
 12    C  1  dyz    Ryd( 3d)     0.00000       2.08566
 13    C  1  dx2y2  Ryd( 3d)     0.00337       2.74845
 14    C  1  dz2    Ryd( 3d)     0.00114       2.40647

 15    O  2  s      Cor( 1s)     1.99998     -20.56485
 16    O  2  s      Val( 2s)     1.70725      -0.92198
 17    O  2  s      Ryd( 3s)     0.00171       1.55322
 18    O  2  px     Val( 2p)     1.62177      -0.42255
 19    O  2  px     Ryd( 3p)     0.00079       1.29654
 20    O  2  py     Val( 2p)     1.91529      -0.46844
 21    O  2  py     Ryd( 3p)     0.00383       1.41052
 22    O  2  pz     Val( 2p)     1.32984      -0.28626
 23    O  2  pz     Ryd( 3p)     0.00011       1.30080
 24    O  2  dxy    Ryd( 3d)     0.00213       3.26414
 25    O  2  dxz    Ryd( 3d)     0.00340       3.20490
 26    O  2  dyz    Ryd( 3d)     0.00000       2.98918
 27    O  2  dx2y2  Ryd( 3d)     0.00406       3.55008
 28    O  2  dz2    Ryd( 3d)     0.00119       3.17511

 29    H  3  s      Val( 1s)     0.88576       0.07107
 30    H  3  s      Ryd( 2s)     0.00298       0.41181
 31    H  3  px     Ryd( 2p)     0.00030       2.18260
 32    H  3  py     Ryd( 2p)     0.00159       2.49146
 33    H  3  pz     Ryd( 2p)     0.00002       1.85643

 34    H  4  s      Val( 1s)     0.88576       0.07107
 35    H  4  s      Ryd( 2s)     0.00298       0.41181
 36    H  4  px     Ryd( 2p)     0.00030       2.18260
 37    H  4  py     Ryd( 2p)     0.00159       2.49146
 38    H  4  pz     Ryd( 2p)     0.00002       1.85643


Summary of Natural Population Analysis:

                                     Natural Population
           Natural    -----------------------------------------------
 Atom No    Charge        Core      Valence    Rydberg      Total
------------------------------------------------------------------------
   C  1    0.37265      1.99997     3.59028    0.03709     5.62735
   O  2   -0.59134      1.99998     6.57415    0.01720     8.59134
   H  3    0.10934      0.00000     0.88576    0.00489     0.89066
   H  4    0.10934      0.00000     0.88576    0.00489     0.89066
========================================================================
* Total * -0.00000      3.99995    11.93596    0.06408    16.00000

                                Natural Population
--------------------------------------------------------------
  Core                      3.99995 ( 99.9988% of    4)
  Valence                  11.93596 ( 99.4664% of   12)
  Natural Minimal Basis    15.93592 ( 99.5995% of   16)
  Natural Rydberg Basis     0.06408 (  0.4005% of   16)
--------------------------------------------------------------

   Atom No         Natural Electron Configuration
--------------------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
     C  1      [core]2s( 1.02)2p( 2.57)3s( 0.01)3p( 0.02)3d( 0.01)
     O  2      [core]2s( 1.71)2p( 4.87)3d( 0.01)
     H  3            1s( 0.89)
     H  4            1s( 0.89)


NATURAL BOND ORBITAL ANALYSIS:

                         Occupancies       Lewis Structure    Low   High
        Max    Occ     -------------------  -----------------   occ   occ
 Cycle  Ctr   Thresh    Lewis   non-Lewis     CR  BD  nC  LP    (L)   (NL)
=============================================================================
   1     2     1.90    15.89671   0.10329      2   4   0   2     0      0
-----------------------------------------------------------------------------

Structure accepted: No low occupancy Lewis orbitals

-------------------------------------------------------
  Core                     3.99995 ( 99.999% of   4)
  Valence Lewis           11.89676 ( 99.140% of  12)
 ==================      ============================
  Total Lewis             15.89671 ( 99.354% of  16)
 -------------------------------------------------------
  Valence non-Lewis        0.07835 (  0.490% of  16)
  Rydberg non-Lewis        0.02493 (  0.156% of  16)
 ==================      ============================
  Total non-Lewis          0.10329 (  0.646% of  16)
-------------------------------------------------------


    (Occupancy)   Bond orbital / Coefficients / Hybrids
------------------ Lewis ------------------------------------------------------
  1. (1.99997) CR ( 1) C  1            s(100.00%)
                                        1.0000  0.0000  0.0000  0.0000  0.0000
                                        0.0000  0.0000  0.0000  0.0000  0.0000
                                        0.0000  0.0000  0.0000  0.0000
  2. (1.99998) CR ( 1) O  2            s(100.00%)
                                        1.0000  0.0000  0.0000  0.0000  0.0000
                                        0.0000  0.0000  0.0000  0.0000  0.0000
                                        0.0000  0.0000  0.0000  0.0000
  3. (1.98853) LP ( 1) O  2            s( 56.22%)p 0.78( 43.73%)d 0.00(  0.05%)
                                        0.0000  0.7496 -0.0170  0.6612  0.0069
                                        0.0000  0.0000  0.0000  0.0000  0.0000
                                        0.0000  0.0000 -0.0201  0.0100
  4. (1.91757) LP ( 2) O  2            s(  0.00%)p 1.00( 99.89%)d 0.00(  0.11%)
                                        0.0000  0.0000  0.0000 -0.0000 -0.0000
                                        0.9994 -0.0098  0.0000 -0.0000 -0.0330
                                       -0.0000  0.0000  0.0000 -0.0000
  5. (1.99996) BD ( 1) C  1- O  2
             ( 33.33%)   0.5773* C  1 s(  0.00%)p 1.00( 99.44%)d 0.01(  0.56%)
                                        0.0000 -0.0000 -0.0000  0.0000 -0.0000
                                       -0.0000  0.0000  0.9951 -0.0652 -0.0000
                                        0.0750  0.0000 -0.0000  0.0000
             ( 66.67%)   0.8165* O  2 s(  0.00%)p 1.00( 99.75%)d 0.00(  0.25%)
                                        0.0000 -0.0000  0.0000  0.0000 -0.0000
                                       -0.0000  0.0000  0.9987 -0.0090  0.0000
                                       -0.0505  0.0000  0.0000 -0.0000
  6. (1.99975) BD ( 2) C  1- O  2
             ( 32.59%)   0.5709* C  1 s( 32.18%)p 2.09( 67.35%)d 0.01(  0.46%)
                                        0.0000  0.5628  0.0714  0.8149  0.0973
                                       -0.0000 -0.0000 -0.0000 -0.0000 -0.0000
```

(continues on next page)

(continued from previous page)

```
                                     -0.0000  0.0000  0.0618 -0.0286
             ( 67.41%)   0.8211* O  2 s( 43.84%)p 1.27( 55.85%)d 0.01(  0.31%)
                                      0.0000  0.6615  0.0284 -0.7470 -0.0215
                                     -0.0000  0.0000 -0.0000 -0.0000  0.0000
                                      0.0000  0.0000  0.0490 -0.0270
  7. (1.99548) BD ( 1) C  1- H  3
             ( 56.63%)   0.7526* C  1 s( 33.98%)p 1.94( 65.86%)d 0.00(  0.16%)
                                      0.0000  0.5826 -0.0192 -0.3995 -0.0029
                                      0.7063 -0.0087  0.0000 -0.0000 -0.0318
                                     -0.0000  0.0000 -0.0180 -0.0153
             ( 43.37%)   0.6585* H  3 s( 99.79%)p 0.00(  0.21%)
                                      0.9989 -0.0095  0.0184 -0.0416  0.0000
  8. (1.99548) BD ( 1) C  1- H  4
             ( 56.63%)   0.7526* C  1 s( 33.98%)p 1.94( 65.86%)d 0.00(  0.16%)
                                      0.0000  0.5826 -0.0192 -0.3995 -0.0029
                                     -0.7063  0.0087 -0.0000  0.0000  0.0318
                                      0.0000  0.0000 -0.0180 -0.0153
             ( 43.37%)   0.6585* H  4 s( 99.79%)p 0.00(  0.21%)
                                      0.9989 -0.0095  0.0184  0.0416 -0.0000
---------------- non-Lewis ----------------------------------------------------
  9. (0.00000) BD*( 1) C  1- O  2
             ( 66.67%)   0.8165* C  1 s(  0.00%)p 1.00( 99.44%)d 0.01(  0.56%)
             ( 33.33%)  -0.5773* O  2 s(  0.00%)p 1.00( 99.75%)d 0.00(  0.25%)
 10. (0.00000) BD*( 2) C  1- O  2
             ( 67.41%)   0.8211* C  1 s( 32.18%)p 2.09( 67.35%)d 0.01(  0.46%)
             ( 32.59%)  -0.5709* O  2 s( 43.84%)p 1.27( 55.85%)d 0.01(  0.31%)
 11. (0.03918) BD*( 1) C  1- H  3
             ( 43.37%)   0.6585* C  1 s( 33.98%)p 1.94( 65.86%)d 0.00(  0.16%)
                                      0.0000 -0.5826  0.0192  0.3995  0.0029
                                     -0.7063  0.0087 -0.0000  0.0000  0.0318
                                      0.0000  0.0000  0.0180  0.0153
             ( 56.63%)  -0.7526* H  3 s( 99.79%)p 0.00(  0.21%)
                                     -0.9989  0.0095 -0.0184  0.0416 -0.0000
 12. (0.03918) BD*( 1) C  1- H  4
             ( 43.37%)   0.6585* C  1 s( 33.98%)p 1.94( 65.86%)d 0.00(  0.16%)
                                      0.0000 -0.5826  0.0192  0.3995  0.0029
                                      0.7063 -0.0087  0.0000 -0.0000 -0.0318
                                     -0.0000  0.0000  0.0180  0.0153
             ( 56.63%)  -0.7526* H  4 s( 99.79%)p 0.00(  0.21%)
                                     -0.9989  0.0095 -0.0184 -0.0416  0.0000
 13. (0.00969) RY ( 1) C  1           s( 29.83%)p 2.30( 68.57%)d 0.05(  1.60%)
                                      0.0000 -0.0565  0.5432 -0.1169  0.8198
                                      0.0000 -0.0000  0.0000 -0.0000 -0.0000
                                      0.0000  0.0000  0.1087 -0.0648
 14. (0.00517) RY ( 2) C  1           s(  0.00%)p 1.00(  9.56%)d 9.46( 90.44%)
                                      0.0000  0.0000 -0.0000  0.0000 -0.0000
                                      0.0465  0.3057  0.0000 -0.0000  0.9510
                                      0.0000  0.0000 -0.0000 -0.0000
 15. (0.00001) RY ( 3) C  1           s( 20.02%)p 0.82( 16.47%)d 3.17( 63.52%)
 16. (0.00000) RY ( 4) C  1           s(  0.00%)p 1.00( 90.64%)d 0.10(  9.36%)
 17. (0.00000) RY ( 5) C  1           s(  0.00%)p 1.00(100.00%)d 0.00(  0.00%)
 18. (0.00000) RY ( 6) C  1           s( 42.72%)p 0.35( 15.02%)d 0.99( 42.26%)
 19. (0.00000) RY ( 7) C  1           s(  0.00%)p 1.00(  0.56%)d99.99( 99.44%)
 20. (0.00000) RY ( 8) C  1           s(  0.00%)p 0.00(  0.00%)d 1.00(100.00%)
 21. (0.00000) RY ( 9) C  1           s(  7.29%)p 0.09(  0.66%)d12.63( 92.05%)
 22. (0.00368) RY ( 1) O  2           s(  0.00%)p 1.00( 98.96%)d 0.01(  1.04%)
                                      0.0000 -0.0000 -0.0000  0.0000  0.0000
                                      0.0064  0.9948 -0.0000  0.0000 -0.1018
                                      0.0000  0.0000 -0.0000  0.0000
 23. (0.00014) RY ( 2) O  2           s( 35.10%)p 1.44( 50.60%)d 0.41( 14.30%)
                                      0.0000 -0.0178  0.5922  0.0556 -0.7091
```

(continues on next page)

(continued from previous page)

```
                                     0.0000 -0.0000 -0.0000  0.0000 -0.0000
                                     0.0000  0.0000  0.3336 -0.1780
 24. (0.00000) RY ( 3) O  2           s( 56.05%)p 0.25( 13.79%)d 0.54( 30.17%)
 25. (0.00000) RY ( 4) O  2           s(  0.00%)p 1.00(100.00%)d 0.00(  0.00%)
 26. (0.00000) RY ( 5) O  2           s(  0.00%)p 1.00(  1.14%)d86.35( 98.86%)
 27. (0.00000) RY ( 6) O  2           s(  0.00%)p 1.00(  0.25%)d99.99( 99.75%)
 28. (0.00000) RY ( 7) O  2           s(  0.00%)p 0.00(  0.00%)d 1.00(100.00%)
 29. (0.00000) RY ( 8) O  2           s(  6.72%)p 5.27( 35.42%)d 8.61( 57.85%)
 30. (0.00000) RY ( 9) O  2           s(  2.07%)p 0.30(  0.61%)d47.00( 97.32%)
 31. (0.00308) RY ( 1) H  3           s( 99.42%)p 0.01(  0.58%)
                                      0.0096  0.9970 -0.0710 -0.0281  0.0000
 32. (0.00002) RY ( 2) H  3           s(  0.22%)p99.99( 99.78%)
 33. (0.00002) RY ( 3) H  3           s(  0.00%)p 1.00(100.00%)
 34. (0.00001) RY ( 4) H  3           s(  0.57%)p99.99( 99.43%)
 35. (0.00308) RY ( 1) H  4           s( 99.42%)p 0.01(  0.58%)
                                      0.0096  0.9970 -0.0710  0.0281  0.0000
 36. (0.00002) RY ( 2) H  4           s(  0.22%)p99.99( 99.78%)
 37. (0.00002) RY ( 3) H  4           s(  0.00%)p 1.00(100.00%)
 38. (0.00001) RY ( 4) H  4           s(  0.57%)p99.99( 99.43%)


NHO DIRECTIONALITY AND BOND BENDING (deviation from line of nuclear centers at
                                  the position of maximum hybrid amplitude)

        [Thresholds for printing:  angular deviation  >  1.0 degree]
                                   p- or d-character  > 25.0%
                                   orbital occupancy  >  0.10e

                      Line of Centers        Hybrid 1             Hybrid 2
                      ---------------  -------------------  -------------------
          NBO           Theta   Phi    Theta   Phi    Dev   Theta   Phi    Dev
================================================================================
  3. LP ( 1) O  2          --     --     90.0    0.0   --      --     --    --
  4. LP ( 2) O  2          --     --     90.0   90.7   --      --     --    --
  5. BD ( 1) C  1- O  2   90.0    0.0     3.0    0.0  87.0   178.7  180.0  88.7


SECOND ORDER PERTURBATION THEORY ANALYSIS OF FOCK MATRIX IN NBO BASIS

    Threshold for printing:   0.50 kcal/mol
                                                          E(2) E(NL)-E(L) F(L,NL)
     Donor (L) NBO              Acceptor (NL) NBO      kcal/mol   a.u.    a.u.
================================================================================

within unit  1
   3. LP ( 1) O  2           13. RY ( 1) C  1           8.58    1.40   0.098
   4. LP ( 2) O  2           11. BD*( 1) C  1- H  3     26.11    1.16   0.156
   4. LP ( 2) O  2           12. BD*( 1) C  1- H  4     26.11    1.16   0.156
   4. LP ( 2) O  2           14. RY ( 2) C  1           5.72    3.06   0.118
   4. LP ( 2) O  2           26. RY ( 5) O  2           0.73    3.75   0.047
   7. BD ( 1) C  1- H  3     12. BD*( 1) C  1- H  4     0.74    1.42   0.029
   7. BD ( 1) C  1- H  3     22. RY ( 1) O  2           2.25    2.12   0.062
   8. BD ( 1) C  1- H  4     11. BD*( 1) C  1- H  3     0.74    1.42   0.029
   8. BD ( 1) C  1- H  4     22. RY ( 1) O  2           2.25    2.12   0.062


NATURAL BOND ORBITALS (Summary):

                                                      Principal Delocalizations
          NBO                 Occupancy    Energy   (geminal,vicinal,remote)
================================================================================
```

(continues on next page)

(continued from previous page)

```
 Molecular unit  1  (CH2O)
 ------ Lewis -------------------------------------
    1. CR ( 1) C  1             1.99997   -11.34329
    2. CR ( 1) O  2             1.99998   -20.56485
    3. LP ( 1) O  2             1.98853    -0.81352  13(v)
    4. LP ( 2) O  2             1.91757    -0.46975  11(v),12(v),14(v),26(g)
    5. BD ( 1) C  1- O  2       1.99996    -0.53505
    6. BD ( 2) C  1- O  2       1.99975    -1.23345
    7. BD ( 1) C  1- H  3       1.99548    -0.72703  22(v),12(g)
    8. BD ( 1) C  1- H  4       1.99548    -0.72703  22(v),11(g)
 ------ non-Lewis -----------------------------------
    9. BD*( 1) C  1- O  2       0.00000     0.20704
   10. BD*( 2) C  1- O  2       0.00000     0.95146
   11. BD*( 1) C  1- H  3       0.03918     0.69317
   12. BD*( 1) C  1- H  4       0.03918     0.69317
   13. RY ( 1) C  1             0.00969     0.58169
   14. RY ( 2) C  1             0.00517     2.58837
   15. RY ( 3) C  1             0.00001     1.75652
   16. RY ( 4) C  1             0.00000     0.96046
   17. RY ( 5) C  1             0.00000     0.64510
   18. RY ( 6) C  1             0.00000     1.49223
   19. RY ( 7) C  1             0.00000     2.24615
   20. RY ( 8) C  1             0.00000     2.08566
   21. RY ( 9) C  1             0.00000     2.49414
   22. RY ( 1) O  2             0.00368     1.39676
   23. RY ( 2) O  2             0.00014     1.56107
   24. RY ( 3) O  2             0.00000     2.18160
   25. RY ( 4) O  2             0.00000     1.30222
   26. RY ( 5) O  2             0.00000     3.27920
   27. RY ( 6) O  2             0.00000     3.20505
   28. RY ( 7) O  2             0.00000     2.98918
   29. RY ( 8) O  2             0.00000     2.69359
   30. RY ( 9) O  2             0.00000     3.12898
   31. RY ( 1) H  3             0.00308     0.41874
   32. RY ( 2) H  3             0.00002     2.57996
   33. RY ( 3) H  3             0.00002     1.85643
   34. RY ( 4) H  3             0.00001     2.06898
   35. RY ( 1) H  4             0.00308     0.41874
   36. RY ( 2) H  4             0.00002     2.57996
   37. RY ( 3) H  4             0.00002     1.85643
   38. RY ( 4) H  4             0.00001     2.06898
          -------------------------------
                 Total Lewis   15.89671  ( 99.3545%)
           Valence non-Lewis    0.07835  (  0.4897%)
           Rydberg non-Lewis    0.02493  (  0.1558%)
          -------------------------------
               Total unit  1   16.00000  (100.0000%)
              Charge unit  1    0.00000

 $CHOOSE
   LONE 2 2 END
   BOND D 1 2 S 1 3 S 1 4 END
 $END

 NBO analysis completed in 0.05 CPU seconds (0 wall seconds)
 Maximum scratch memory used by NBO was 297106 words (2.27 MB)
Stopping NBO...Storing NBOs: NBO_1.nbo

                               *** returned from  NBO  program ***
```

Thus, in this example the NBO analysis of formaldehyde shows that a single Lewis structure is dominant with single bonds between C and H, a double bond between C and O and two lone pairs at the oxygen – just as ordinary

chemical arguments would imply. In addition, the program produces the four corresponding valence antibonds. The remaining components of the basis set span the "Rydberg" space and lead to semilocalized, orthogonal orbitals that are assigned to single atoms (Note the nomenclature: BD = bond, BD* = antibond, LP = lone pair, CR = core orbital, RY= Rydberg orbital). The NPA analysis shows a patially negative oxygen and partially positive carbon and hydrogen atoms.

Additionally, the NBO orbitals are stored in the ORCA .gbw file format as `jobname.nbo`. This file can be used for further analysis and usage with ORCA e.g. for plotting orbitals via `orca_plot`.

The **NBO** program has many additional features and analysis tools. The features that are implemented in ORCA can be controlled via the `%nbo`-block

```
%nbo
NBOKEYLIST    = "$NBO     ... $END"
DELKEYLIST    = "$DEL     ... $END"
COREKEYLIST   = "$CORE    ... $END"
NRTSTRKEYLIST = "$NRTSTR  ... $END"
NPEPAKEYLIST  = "$NPEPA   ... $END"
end
```

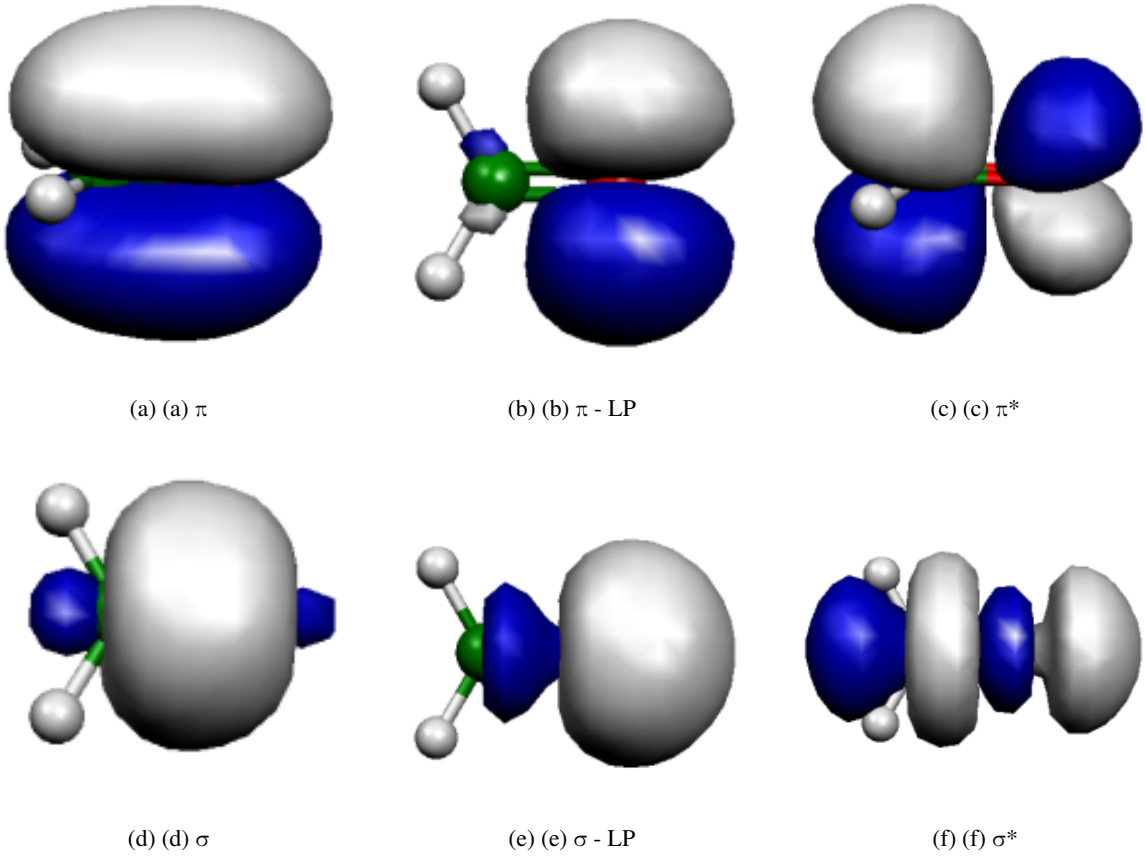
The syntax of the respective keylists is given by the NBO6.x/NBO7.x manual.

Specifying the single `! NBO` keyword corresponds to the `%nbo`-block

```
%nbo
NBOKEYLIST    = "$NBO NBO NPA AONBO=C ARCHIVE $END"
end
```

The full set of features beyond those which can be give via the `%nbo` block can be accessed using the file `FILE.47`, which is generated by the NBO program. This is an ascii file that can be edited with a text editor. Add or remove keywords in the corresponding blocks as needed and call the gennbo program like

```
gennbo < FILE.47 >jobname.nboout
```



(a) (a) π (b) (b) π - LP (c) (c) π*

(d) (d) σ (e) (e) σ - LP (f) (f) σ*

Fig. 7.63: Six NBOs of the H_2CO molecule. Shown are the occupied bonding π and σ orbitals (left) for C and O, the two oxygen lone pairs (middle) and the two π^* and σ^* antibonding orbitals (right).

The `FILE.47` file looks like:

```
$GENNBO  NATOMS=4  NBAS=38  UPPER  BODM  FORMAT  $END
 $NBO  $END
 $COORD
 ORCA Job: check
     6    6       0.000000       0.000000       0.000000
     8    8       2.267671       0.000000       0.000000
     1    1      -1.039349       1.800206       0.000000
     1    1      -1.039349      -1.800206       0.000000
 $END
 $BASIS
```

If you have no need for this (rather large) file, then you have to delete it manually!

7.52.1 NBO Deletions

An advanced feature, which has been implemented via the ORCA-NBO interface, is the possibility of using deletions.

```
! RHF 3-21G BOHRS TightSCF

%nbo
nbokeylist="$nbo nbo npa aonbo=c archive $end"
delkeylist="$del lewis delete 1 element 3 11 $end"
end
```

(continues on next page)

(continued from previous page)

```
*xyz 0 1
 C              1.4089705283        0.0210567401        0.0000000000
 N             -1.3645072652       -0.1355759321        0.0000000000
 H              1.9849776453        1.9986808971        0.0000000000
 H              2.1492280974       -0.9096841007        1.6818209547
 H              2.1492280974       -0.9096841007       -1.6818209547
 H             -2.0504340036        0.7268536543       -1.5583845544
 H             -2.0504340036        0.7268536543        1.5583845544
*
```

The DELKEYLIST provides NBO with the task to perform certain deletions of orbitals/interactions. Per deletion ORCA calculates a new Fock matrix on basis of an NBO density corresponding to the deletions:

```
Stopping NBO...Starting NBO again for $del instructions...

 LEWIS: Delete all non-Lewis NBOs
 Deletion of the following orbitals from the NBO Fock matrix:
   10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28

 Orbital occupancies:

      Orbital                          No deletions  This deletion    Change
 --------------------------------------------------------------------------------
    1. CR ( 1) C  1                      1.99978        2.00000       0.00022
    2. CR ( 1) N  2                      1.99983        2.00000       0.00017
    3. LP ( 1) N  2                      1.97796        2.00000       0.02204
    4. BD ( 1) C  1- N  2                1.99846        2.00000       0.00154
    5. BD ( 1) C  1- H  3                1.99858        2.00000       0.00142
    6. BD ( 1) C  1- H  4                1.99406        2.00000       0.00594
    7. BD ( 1) C  1- H  5                1.99406        2.00000       0.00594
    8. BD ( 1) N  2- H  6                1.99440        2.00000       0.00560
    9. BD ( 1) N  2- H  7                1.99440        2.00000       0.00560
   10. BD*( 1) C  1- N  2                0.00009        0.00000      -0.00009
   11. BD*( 1) C  1- H  3                0.01567        0.00000      -0.01567
   12. BD*( 1) C  1- H  4                0.00763        0.00000      -0.00763
   13. BD*( 1) C  1- H  5                0.00763        0.00000      -0.00763
   14. BD*( 1) N  2- H  6                0.00424        0.00000      -0.00424
   15. BD*( 1) N  2- H  7                0.00424        0.00000      -0.00424
   16. RY ( 1) C  1                      0.00094        0.00000      -0.00094
   17. RY ( 2) C  1                      0.00034        0.00000      -0.00034
   18. RY ( 3) C  1                      0.00020        0.00000      -0.00020
   19. RY ( 4) C  1                      0.00001        0.00000      -0.00001
   20. RY ( 1) N  2                      0.00114        0.00000      -0.00114
   21. RY ( 2) N  2                      0.00044        0.00000      -0.00044
   22. RY ( 3) N  2                      0.00034        0.00000      -0.00034
   23. RY ( 4) N  2                      0.00001        0.00000      -0.00001
   24. RY ( 1) H  3                      0.00163        0.00000      -0.00163
   25. RY ( 1) H  4                      0.00079        0.00000      -0.00079
   26. RY ( 1) H  5                      0.00079        0.00000      -0.00079
   27. RY ( 1) H  6                      0.00117        0.00000      -0.00117
   28. RY ( 1) H  7                      0.00117        0.00000      -0.00117

 NEXT STEP:  Perform one SCF cycle to evaluate the energy of the new density
             matrix constructed from the deleted NBO Fock matrix.

 --------------------------------------------------------------------------------
Copying NBO density...
Calculating new Fock-Matrix...
Calculating Fock-Matrix...done!
```

(continues on next page)

(continued from previous page)

```
New NBO energy via Fock-Matrix: -94.629937


Starting NBO again for $del/return energy instructions...

 ---------------------------------------------------------------------------
   Energy of deletion :        -94.629936711
     Total SCF energy :        -94.679444929
                         --------------------
        Energy change :          0.049508 a.u.,          31.067 kcal/mol
 ---------------------------------------------------------------------------
```

Multiple deletions can also be specified, as can be seen for this example. The output then also contains the additional energy values:

```
Starting NBO again for $del instructions...

 Deletion of the following NBO Fock matrix elements:
     3, 11;

 Orbital occupancies:

       Orbital                          No deletions  This deletion    Change
 ---------------------------------------------------------------------------
    1. CR ( 1) C  1                       1.99978        1.99978      -0.00000
    2. CR ( 1) N  2                       1.99983        1.99983      -0.00000
    3. LP ( 1) N  2                       1.97796        1.99348       0.01552
    4. BD ( 1) C  1- N  2                 1.99846        1.99860       0.00015
    5. BD ( 1) C  1- H  3                 1.99858        1.99845      -0.00014
    6. BD ( 1) C  1- H  4                 1.99406        1.99404      -0.00002
    7. BD ( 1) C  1- H  5                 1.99406        1.99404      -0.00002
    8. BD ( 1) N  2- H  6                 1.99440        1.99450       0.00011
    9. BD ( 1) N  2- H  7                 1.99440        1.99450       0.00011
   10. BD*( 1) C  1- N  2                 0.00009        0.00008      -0.00000
   11. BD*( 1) C  1- H  3                 0.01567        0.00042      -0.01525
   12. BD*( 1) C  1- H  4                 0.00763        0.00780       0.00017
   13. BD*( 1) C  1- H  5                 0.00763        0.00780       0.00017
   14. BD*( 1) N  2- H  6                 0.00424        0.00424       0.00000
   15. BD*( 1) N  2- H  7                 0.00424        0.00424       0.00000
   16. RY ( 1) C  1                       0.00094        0.00063      -0.00031
   17. RY ( 2) C  1                       0.00034        0.00034       0.00000
   18. RY ( 3) C  1                       0.00020        0.00032       0.00012
   19. RY ( 4) C  1                       0.00001        0.00002       0.00001
   20. RY ( 1) N  2                       0.00114        0.00115       0.00001
   21. RY ( 2) N  2                       0.00044        0.00044       0.00000
   22. RY ( 3) N  2                       0.00034        0.00034       0.00000
   23. RY ( 4) N  2                       0.00001        0.00001       0.00000
   24. RY ( 1) H  3                       0.00163        0.00092      -0.00072
   25. RY ( 1) H  4                       0.00079        0.00083       0.00005
   26. RY ( 1) H  5                       0.00079        0.00083       0.00005
   27. RY ( 1) H  6                       0.00117        0.00118       0.00000
   28. RY ( 1) H  7                       0.00117        0.00118       0.00000

 NEXT STEP:  Perform one SCF cycle to evaluate the energy of the new density
             matrix constructed from the deleted NBO Fock matrix.

 ---------------------------------------------------------------------------
Copying NBO density...
Calculating new Fock-Matrix...
Calculating Fock-Matrix...done!
```

(continues on next page)

(continued from previous page)

```
New NBO energy via Fock-Matrix: -94.668383


Starting NBO again for $del/return energy instructions...

 ---------------------------------------------------------------------------
   Energy of deletion :        -94.668383268
     Total SCF energy :        -94.679444929
                         --------------------
        Energy change :          0.011062 a.u.,           6.941 kcal/mol
 ---------------------------------------------------------------------------
```

NOTE: Deletions are only implemented for SCF methods!

7.52.2 NBO for Post-HF Densities

NBO analysis can be performed on all methods producing a density. In some methods the density generation has to be specified explictly, e. g. for MP2 calculations this would be:

```
! MP2 3-21G TightSCF BOHRS NBO

%MP2
density relaxed
end

*xyz 0 1
 C             1.4089705283        0.0210567401        0.0000000000
 N            -1.3645072652       -0.1355759321        0.0000000000
 H             1.9849776453        1.9986808971        0.0000000000
 H             2.1492280974       -0.9096841007        1.6818209547
 H             2.1492280974       -0.9096841007       -1.6818209547
 H            -2.0504340036        0.7268536543       -1.5583845544
 H            -2.0504340036        0.7268536543        1.5583845544
*
```

The output will contain both the NBO analysis of the SCF density as well as of the MP2 relaxed density. An NBO analysis of a density generated by the MDCI module can be specified as follows:

```
! CISD 3-21G TightSCF BOHRS NBO

%mdci
    density linearized
end

*xyz 0 1
 C        1.4089705283        0.0210567401        0.0000000000
 N       -1.3645072652       -0.1355759321        0.0000000000
 H        1.9849776453        1.9986808971        0.0000000000
 H        2.1492280974       -0.9096841007        1.6818209547
 H        2.1492280974       -0.9096841007       -1.6818209547
 H       -2.0504340036        0.7268536543       -1.5583845544
 H       -2.0504340036        0.7268536543        1.5583845544
*
```

Again, the output will contain both the NBO analysis of the SCF density as well es of the CISD linearized density.

7.52.3 Natural Chemical Shielding Analysis (NCS)

For closed-shell calculations of NMR chemical shielding at the SCF level (see sections *NMR Chemical Shifts* and *EPR and NMR properties*), the NCS analysis can be requested by adding `NCS` to the `NBOKEYLIST`. The `NCS` keyword accepts the arguments `U`, `I`, `CSA`, `XYZ`, and `MO` to analyze the "unperturbed", "induced", anisotropic, Cartesian, and canonical MO contributions to the shielding tensors, respectively, as well as a decimal number for the printing threshold (in ppm). For more information, consult the NBO manual and the original publication.[113]

```
! PBE def2-TZVP NMR
%nbo
  NBOKeyList = "$NBO NCS=0.01,I,U,XYZ $END"
end
* xyz 0 1
  H     0.00     0.00     0.00
  C     1.06     0.00     0.00
  N     2.23     0.00     0.00
*
```

```
Summary of isotropic NMR chemical shielding
 Total Lewis (L) and non-Lewis (NL) contributions: (ppm)

        NBO           H  1    C  2    N  3
 ---------------    ------- ------- -------
   1.  C 2(cr)   L    -0.18  200.26    0.18
               NL    -0.02    0.02    0.00
   2.  N 3(cr)   L    -0.03   -0.12  235.18
               NL     0.00    0.01    0.02
   3.  N 3(lp)   L     1.02  -33.00 -151.92
               NL    -1.04    1.81   12.18
   4.  H 1- C 2  L    25.75  -49.28  -20.30
               NL    -1.24    6.10    2.26
   5.  C 2- N 3  L     2.29   15.40   13.66
               NL     0.02    0.00   -0.00
   6.  C 2- N 3  L     2.29   15.40   13.66
               NL     0.02    0.00   -0.00
   7.  C 2- N 3  L     0.46  -77.94 -151.00
               NL     0.05   -4.41    0.95
 ---------------    ------- ------- -------
          Lewis       31.59   70.70  -60.53
      non-Lewis       -2.21    3.53   15.42
 ---------------    ------- ------- -------
          Total       29.38   74.23  -45.11
```

7.53 Population Analyses and Control of Output

At present ORCA knows three different ways of analyzing the computed SCF wavefunction that will be described below. All of these methods can produce a tremendous amount of output. However, this output can be precisely controlled by the user to his or her individual needs.

In general there is one compound key called `PrintLevel` which is there to choose reasonable amounts of output. All that `PrintLevel` does is to set certain flags in the array `Print` which holds the details about what to print and what not.

7.53.1 Controlling Output

The array Print allows the control of output. The general way of assigning elements of Print is:

```
%output
  PrintLevel     Normal
  Print[ Flag ]  0  # turn print off
                 1  # turn print on
                 n  # some flags are more sophisticated
end
```

The compound key PrintLevel can be used to select certain default settings for the print array. Specifying Print after PrintLevel can be used to modify these defaults.

```
%output
  PrintLevel     Nothing
                 Mini
                 Small
                 Normal
                 Maxi
                 Large
                 Huge
                 Debug
end
```

Print has presently the following elements that can be user controlled:

Flag	Action
P_InputFile	Echo the input file
P_Cartesian	Print the cartesian coordinates
P_Internal	Print the internal coordinates
P_Basis	$= 1$: Print the basis set information
	$= 2$: Also print the primitives in input format
P_OneElec	Print of the one electron matrix
P_Overlap	Print the overlap matrix
P_KinEn	Print the kinetic energy matrix
P_S12	Print the $S^{-1/2}$ matrix
P_GuessOrb	Print the initial guess orbitals
P_OrbEn	Print Orbital Energies
P_MOs	Print the MO coefficients on convergence
P_Density	Print the converged electron density
P_SpinDensity	Print the converged spin density
P_EHTDetails	Print initial guess extended Hückel details
P_SCFInfo	Print the SCF input flags
P_SCFMemInfo	Print the estimated SCF memory requirements
P_SCFIterInfo	$= 1$: print short iteration information
	$= 2$: print longer iteration information
	$= 3$: in a direct SCF also print integral progress
P_Fockian	Print Fockian matrix
P_DIISMat	Print DIIS matrix
P_DIISError	Print DIIS error
P_Iter_P	Print Density
P_Iter_C	Print MO coefficients
P_Iter_F	Print Fock matrix
P_Mayer	Print Mayer population analysis. Default = on.
P_NatPop	Print Natural population analysis. Default = off.
P_Hirshfeld	Print Hirshfeld population analysis. Default = off.
P_MBIS	Print MBIS population analysis. Default = off.

continues on next page

Table 7.31 – continued from previous page

Flag	Action
P_Mulliken	Print Mulliken population analysis. Default = on
P_AtCharges_M	Print Mulliken atomic charges
P_OrbCharges_M	Print Mulliken orbital charges
P_FragCharges_M	Print Mulliken fragment charges
P_FragBondOrder_M	Print Mulliken fragment bond orders
P_BondOrder_M	Print Mulliken bond orders
P_ReducedOrbPop_M	Print Mulliken reduced orb. charges
P_FragPopMO_M	Print Mulliken fragment population for each MO
P_FragOvlMO_M	Print Mulliken overlap populations per fragment pair
P_AtPopMO_M	Print Mulliken atomic charges in each MO
P_OrbPopMO_M	Print Mulliken orbital population for each MO
P_ReducedOrbPopMO_M	Print Mulliken reduced orbital population for each MO
P_Loewdin	Print Loewdin population analysis. Default = on.
P_AtCharges_L	Print Loewdin atomic charges
P_OrbCharges_L	Print Loewdin orbital charges
P_FragCharges_L	Print Loewdin fragment charges
P_FragBondOrder_L	Print Loewdin fragment bond orders
P_BondOrder_L	Print Loewdin bond orders
P_ReducedOrbPop_L	Print Loewdin reduced orb. charges
P_FragPopMO_L	Print Loewdin fragment population for each MO
P_FragOvlMO_L	Print Loewdin overlap populations per fragment pair
P_AtPopMO_L	Print Loewdin atomic charges in each MO
P_OrbPopMO_L	Print Loewdin orbital population for each MO
P_ReducedOrbPopMO_L	Print Loewdin reduced orbital population for each MO
P_NPA	Natural population analysis
P_NBO	Natural bond orbital analysis
P_Fragments	Print fragment information
P_GUESSPOP	Print initial guess populations
P_UNO_FragPopMO_M	Print Mulliken fragment population per UNO
P_UNO_OrbPopMO_M	Print Mulliken orbital pop. per UNO
P_UNO_AtPopMO_M	Print Mulliken atomic charges per UNO
P_UNO_ReducedOrbPopMO_M	Print Mulliken reduced orbital pop. per UNO
P_UNO_FragPopMO_L	Print Loewdin fragment population per UNO
P_UNO_OrbPopMO_L	Print Loewdin orbital pop. per UNO
P_UNO_AtPopMO_L	Print Loewdin atomic charges per UNO
P_UNO_ReducedOrbPopMO_L	Print Loewdin reduced orbital pop. per UNO
P_UNO_OccNum	Print occupation numbers per UNO
P_AtomExpVal	Print atomic expectation values
P_AtomBasis	Print atomic basis
P_AtomDensFit	Print electron density fit
P_Symmetry	Symmetry basic information
P_Sym_Salc	Symmetry process printing
P_SCFSTABANA	Information on progress, convergence, and results of the SCF stability analysis
P_DFTD	Print info on Grimme's dispersion correction
	print mini = 0
	print small = 1
	print normal = 1
	print maxi = 2
	print huge = 2
P_DFTD_GRAD	Print gradient info on Grimme's dispersion correction
	print mini = 0
	print small = 0
	print normal = 0
	print maxi = 1
	print huge = 2

continues on next page

Table 7.31 – continued from previous page

Flag	Action
P_G1EL2EL	Print one- and two-electron contributions to the g-tensor

The various choices for PrintLevel have the following defaults:

PrintLevel	Print settings	
Mini	P_OrbEn	= 1
	P_Cartesian	= 1
	P_InputFile	= 1
	P_SCFIterInfo	= 1
Small	all the previous plus	
	P_SCFInfo	= 1
	P_Mayer	= 1
	P_MULLIKEN	= 1
	P_AtCharges_M	= 1
	P_ReducedOrbPop_M	= 1
	P_Loewdin	= 1
	P_AtCharges_L	= 1
	P_ReducedOrbPop_L	= 1
	P_Fragments	= 1
	P_FragCharges_M	= 1
	P_FragBondOrder_M	= 1
	P_FragCharges_L	= 1
	P_FragBondOrder_L	= 1
Normal	all the previous plus	
	P_Internal	= 1
	P_BondOrder_L	= 1
	P_BondOrder_M	= 1
	P_FragPopMO_L	= 1
	P_ReducedOrbPopMO_L	= 1
	P_SCFIterInfo	= 2
Maxi	all the previous plus	
	P_GuessOrb	= 1
	P_MOs	= 1
	P_Density	= 1
	P_SpinDensity	= 1
	P_Basis	= 1
	P_FragOVLMO_M	= 1
	P_OrbPopMO_M	= 1
	P_OrbCharges_M	= 1
Huge	All the previous plus	
	P_OneElec	= 1
	P_Overlap	= 1
	P_S12	= 1
	P_AtPopMO_M	= 1
	P_OrbPopMO_M	= 1
	P_AtPopMO_L	= 1
	P_EHTDetails	= 1
Debug	print everything	

7.53.2 Mulliken Population Analysis

The Mulliken population analysis [598] is, despite all its known considerable weaknesses, the standard in most quantum chemical programs. It partitions the total density using the assignment of basis functions to given atoms in the molecules and the basis function overlap. If the total charge density is written as $\rho(\vec{r})$ and the total number of electrons is N we have:

$$\int \rho(\vec{r})\, d\vec{r} = N \tag{7.407}$$

and from the density matrix $\mathbf{P}$ and the basis functions ϕ:

$$\rho(\vec{r}) = \sum_{\mu\nu} P_{\mu\nu}\phi_\mu(\vec{r})\,\phi_\nu(\vec{r}) \tag{7.408}$$

therefore:

$$\int \rho(\vec{r})\, d\vec{r} = \sum_{\mu\nu} P_{\mu\nu} \underbrace{\int \phi_\mu(\vec{r})\,\phi_\nu(\vec{r})\, d\vec{r}}_{S_{\mu\nu}} \tag{7.409}$$

$$= \sum_{\mu\nu} P_{\mu\nu} S_{\mu\nu} \tag{7.410}$$

After assigning each basis function to a given center this can be rewritten:

$$= \sum_A \sum_B \sum_\mu {}^A \sum_\nu {}^B P^{AB}_{\mu\nu} S^{AB}_{\mu\nu} \tag{7.411}$$

$$= \sum_A \sum_\mu {}^A \sum_\nu {}^A P^{AA}_{\mu\nu} S^{AA}_{\mu\nu} + 2\sum_A \sum_{B<A} \sum_\mu {}^A \sum_\nu {}^B P^{AB}_{\mu\nu} S^{AB}_{\mu\nu} \tag{7.412}$$

Mulliken proposed to divide the second term equally between each pair of atoms involved and define the number of electrons on center A, N_A, as:

$$N_A = \sum_\mu {}^A \sum_\nu {}^A P^{AA}_{\mu\nu} S^{AA}_{\mu\nu} + \sum_{B\neq A} \sum_\mu {}^A \sum_\nu {}^B P^{AB}_{\mu\nu} S^{AB}_{\mu\nu} \tag{7.413}$$

such that $\sum_A N_A = N$. The charge of an atom in the molecule is then:

$$Q_A = Z_A - N_A \tag{7.414}$$

where Z_A is the core charge of atom A. The cross terms between pairs of basis functions centered on different atoms is the overlap charge and is used in ORCA to define the Mulliken bond order:

$$B_{AB} = 2\sum_\mu {}^A \sum_\nu {}^B P^{AB}_{\mu\nu} S^{AB}_{\mu\nu} \tag{7.415}$$

The Mulliken population analysis is turned on by using:

```
%output
  Print[ P_Mulliken ] 1  # default = on
end
```

A number of additional options can be specified to control the details of the Mulliken population analysis. By default the Mulliken population analysis is turned on.

```
%output
  Print[ P_AtCharges_M  ] 1   # Print atomic charges
  Print[ P_OrbCharges_M ] 1   # Print orbital charges
  Print[ P_FragCharges_M] 1   # Print fragment charges
  Print[ P_BondOrder_M  ] 1   # Print bond orders
```

(continues on next page)

(continued from previous page)

```
  Print[ P_FragBondOrder_M ] 1# Print fragment b.o.
  Print[ P_ReducedOrbPop_M ] 1# Print reduced orb. Charges
  Print[ P_AtPopMO_M    ] 1   # Print atomic charges in
                              # each MO
  Print[ P_OrbPopMO_M   ] 1   # Print orbital populaiton
                              # for each MO
  Print[ P_ReducedOrbPopMO_M] 1 # Print reduced orbital
                                # pop for each MO
  Print[ P_FragPopMO_M  ] 1     # Print the fragment
                                # population for for each MO
end
```

These options allow to get very detailed information about the computed wavefunctions and is much more convenient than to look at the MOs directly. A "reduced orbital population" is a population per angular momentum type. For example the sum of populations of each p_z orbital at a given atom is the reduced orbital population of the p_z function.

Note that for finite temperature HF or KS-DFT calculations (SmearTemp > 0 K, fractional occupation numbers or FOD analysis, see *Finite Temperature HF/KS-DFT*), only the Mulliken reduced orbital charges based on ρ^{FOD} will be printed. They can be used to get a first impression about the localization of hot electrons in the molecule without generating the corresponding FOD plot (see *FOD plots*). The following example shows the corresponding printout for the first carbon atom of *p*-benzyne based on a FOD analysis with default settings (see *Fractional Occupation Number Weighted Electron Density (FOD)*).

```
----------------------------------------
FOD BASED MULLIKEN REDUCED ORBITAL CHARGES
----------------------------------------
  0 C s       :     0.006371  s :     0.006371
      pz      :     0.016375  p :     0.030785
      px      :     0.009893
      py      :     0.004516
      dz2     :     0.004248  d :     0.010308
      dxz     :     0.000254
      dyz     :     0.004855
      dx2y2   :     0.000860
      dxy     :     0.000091
      f0      :     0.000006  f :     0.000378
      f+1     :     0.000014
      f-1     :     0.000309
      f+2     :     0.000002
      f-2     :     0.000006
      f+3     :     0.000010
      f-3     :     0.000032
```

If other population analysis printouts are wanted the user is referred to the Löwdin analysis (*Löwdin Population Analysis*) which is turned on by default using the total SCF density of the calculation, also in the case of finite electronic temperature.

7.53.3 Löwdin Population Analysis

The Löwdin analysis [839] is somewhat more straightforward than the Mulliken analysis. In the Löwdin method one changes to a basis where all overlap integrals vanish. This is accomplished via Löwdins symmetric orthogonalization matrix $\mathbf{S}^{-1/2}$. Using this transformation matrix the new basis functions are multicentered but are in a least square sense as similar as possible to the original, strictly localized, atomic basis functions. The similarity of the transformed functions and original functions is explored in the population analysis. The density matrix transforms as:

$$\mathbf{P}^L = \mathbf{S}^{1/2}\mathbf{P}\mathbf{S}^{1/2} \tag{7.416}$$

Then the atomic populations are:

$$N_A = \sum_{\mu}{}^{A} P^{L}_{\mu\mu} \tag{7.417}$$

The bond order is defined from the Wiberg index [888] that was first used in the context of semiempirical methods (that are formulated in the Löwdin basis right from the start):

$$B_{AB} = \sum_{\mu}{}^{A} \sum_{\nu}{}^{B} \left(P^{L}_{\mu\nu}\right)^2 \tag{7.418}$$

The output for the Löwdin population analysis (that I personally prefer over the Mulliken analysis) is closely similar. By default the Löwdin population analysis is turned on and provides some more detail than the Mulliken analysis.

```
%output
  Print[ P_Loewdin ] 1  # default = on
end
```

The flags to regulate the details are almost identical:

```
%output
  Print[ P_AtCharges_L  ] 1     # Print atomic charges
  Print[ P_OrbCharges_L ] 1     # Print orbital charges
  Print[ P_FragCharges_L] 1     # Print fragment charges
  Print[ P_BondOrder_L  ] 1     # Print bond orders
  Print[ P_FragBondOrder_L ] 1  # Print fragment b.o.
  Print[ P_ReducedOrbPop_L ] 1  # Print reduced orb. Charges
  Print[ P_AtPopMO_L    ] 1     # Print atomic charges in
                                # each MO
  Print[ P_OrbPopMO_L   ] 1     # Print orbital population
                                # for each MO
  Print[ P_ReducedOrbPopMO_L] 1 # Print reduced orbital
                                # pop for each MO
  Print[ P_FragPopMO_L  ] 1     # Print the fragment
                                # population for each MO
end
```

In addition one can set, in the method block, the threshold for the printing of the bond order.

```
%method
  LOEWDIN_BONDORDERTHRESH 0.05
end
```

7.53.4 Mayer Population Analysis

Mayers bonding analysis [567, 568, 569, 570] is another creative attempt to define chemically useful indices. The Mayer atomic charge is identical to the Mulliken charge. The Mayer bond order is defined as:

$$B_{AB} = \sum_{\mu}{}^{A} \sum_{\nu}{}^{B} \left(\mathbf{PS}\right)_{\mu\nu} \left(\mathbf{PS}\right)_{\nu\mu} + \left(\mathbf{RS}\right)_{\mu\nu} \left(\mathbf{RS}\right)_{\nu\mu} \tag{7.419}$$

Here $\mathbf{P}$ is the total electron density matrix and $\mathbf{R}$ is the spin-density matrix. These Mayer bond orders are very useful. Mayer's total valence for atom A is defined as:

$$V_A = 2N_A - \sum_{\mu}{}^{A} \sum_{\nu}{}^{A} \left(\mathbf{PS}\right)_{\mu\nu} \left(\mathbf{PS}\right)_{\nu\mu} \tag{7.420}$$

In normal bonding situations and with normal basis sets V_A should be reasonably close to the valence of atom A in a chemical sense (i.e. close to four for a carbon atom). The bonded valence is given by:

$$X_A = V_A - \sum_{B \neq A} B_{AB} \tag{7.421}$$

and finally the free valence (a measure of the ability to form further bonds) is given by:

$$F_A = V_A - X_A \tag{7.422}$$

The Mayer population analysis is turned on by:

```
%output
  Print[ P_Mayer ] 1  # default = on
end
```

The output is rather simple and short and can not be further controlled. By default the Mayer population analysis is turned on. In addition one can set, in the method block, the threshold for the printing of the bond order.

```
%method
  MAYER_BONDORDERTHRESH 0.1
end
```

7.53.5 Natural Population Analysis

A popular and useful method for population analysis is the natural population analysis due to Weinhold and co-workers. It is implemented in the NBO interface.

7.53.6 Local Spin Analysis

It is common practice in various areas of chemistry to think about the interaction of open-shell systems in terms of local spin states. For example, in dimeric or oligomeric transition metal clusters, the 'exchange coupling' between open shell ions that exist locally in high-spin states is a much studied phenomenon. Diradicals would be typical systems in organic chemistry that show this phenomenon. In quantum mechanics, however, the total spin is not a local property, but instead a property of the system as a whole. The total spin squared, S^2, and its projection onto the z-axis, S_z, commute with the non-relativistic Hamiltonian and hence, the eigenfunctions of the non-relativistic Hamiltonian can be classified according to good quantum numbers S and M according to:

$$\mathbf{S}^2 \left|\Psi^{SM}\right\rangle = S(S+1) \left|\Psi^{SM}\right\rangle$$

$$S_z \left|\Psi^{SM}\right\rangle = M \left|\Psi^{SM}\right\rangle$$

where $\left|\Psi^{SM}\right\rangle$ is an exact eigenfunction of the non-relativistic Hamiltonian or an approximation to it that conserves the total spin as a good quantum number. The total spin itself is given by the sum over the individual electron spins as:

$$\mathbf{S} = \sum_i \mathbf{s}(i)$$

And hence,

$$\mathbf{S}^2 = \sum\nolimits_{i,j} \mathbf{s}(i)\mathbf{s}(j)$$

is a two-electron property of the system. It is obviously not trivial to relate the chemically very meaningful concept of local spin to a rigorous quantum mechanical treatment. While there are various proposals of how to deal with this problem, we follow here a proposal of Clark and Davidson (Clark, A.E.; Davidson, E.R., J. Chem. Phys. 2001, 115, 7382-7392). The following equations are implemented in the SCF and CASSCF modules of Orca.

Clark and Davidson define fragment projection operators with the property:

$$P_A P_B = \delta_{AB} P_A$$

and:

$$\sum_A P_A = 1$$

Then using this identity:

$$\mathbf{S} = \sum_i \sum_A \mathbf{s}(i) P_{A(i)}$$

$$\mathbf{S} = \sum_A \sum_i \mathbf{s}(i) P_A(i)$$

$$= \sum_A \mathbf{S}_A(i)$$

they show that the local spin operators obey the standard relations for spin operators:

$$\mathbf{S}_A = \mathbf{S}_A^\dagger$$

$$\mathbf{S}_A \times \mathbf{S}_A = i\hbar \mathbf{S}_A$$

Hence

$$\mathbf{S}^2 = \sum_A \sum_B \mathbf{S}_A \mathbf{S}_B$$

But then importantly:

$$\mathbf{S}_A \mathbf{S}_B = \sum_i \sum_j \mathbf{s}(i)\mathbf{s}(j) P_A(i) P_B(j)$$

$$= \tfrac{3}{4}\delta_{AB} \sum_A P_A(i) + \sum_i \sum_{j>i} \mathbf{s}(i)\mathbf{s}(j)\{P_A(i)P_B(j) + P_A(j)P_B(i)\}$$

With the first- and second-order density matrix:

$$\gamma(\mathbf{x}, \mathbf{x}') = N \int \Psi(\mathbf{x}, \mathbf{x}_2, ..., \mathbf{x}_N)\Psi^*(\mathbf{x}', \mathbf{x}_2, ..., \mathbf{x}_N) d\mathbf{x}_2 ... d\mathbf{x}_N$$

$$\Gamma(\mathbf{x}_1, \mathbf{x'}_1; \mathbf{x}_2, \mathbf{x'}_2) = \binom{N}{2} \int \Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N)\Psi^*(\mathbf{x'}_1, \mathbf{x'}_2, ..., \mathbf{x}_N) d\mathbf{x}_3 ... d\mathbf{x}_N$$

(with $\binom{N}{2} = \frac{1}{2}N(N-1)$). Then:

$$\langle \mathbf{S}_A \mathbf{S}_B \rangle = \tfrac{3}{4}\delta_{AB} tr(\gamma P_A) + 2tr(P_A(1)P_B(2)\mathbf{s}(1)\mathbf{s}(2)\Gamma(1,1;2,2))$$

In terms of the number of electrons on site 'A'and the expectation value of S_z^A

$$\left\langle S_z^A \right\rangle = \tfrac{1}{2} tr(\gamma^{\alpha-\beta} P_A)$$

$$\left\langle N^A \right\rangle = tr(\gamma^{\alpha+\beta} P_A)$$

in terms of molecular orbitals:

$$\left\langle S_z^A \right\rangle = \tfrac{1}{2} \sum_{p,q} \gamma_{pq}^{\alpha-\beta} \langle p | P_A | q \rangle$$

$$\left\langle N^A \right\rangle = \sum_{p,q} \gamma_{pq}^{\alpha+\beta} \langle p | P_A | q \rangle$$

McWeeny and Kutzelnigg (McWeeny, R.; Kutzelnigg, W. Int. J. Quant. Chem. 1968, 11, 187-203) show that for the expectation value of s(1)s(2), the relevant irreducible part of the two-body density can be expressed in terms of the spinless density matrix of second order:

$$R_0^{(0)}(1,1';2,2') = -\tfrac{1}{3}\Gamma(1,1';2,2') - \tfrac{2}{3}\Gamma(2,1';1,2')$$

$$= -\tfrac{1}{3} \sum_{pqrs} \Gamma_{rs}^{pq} p(1) q(1') r(2) s(2') + 2\Gamma_{rs}^{pq} p(2) q(1') r(1) s(2')$$

$$= -\tfrac{1}{3}\sum_{pqrs}(\Gamma^{pq}_{rs} + 2\Gamma^{rq}_{ps})p(1)q(1')r(2)s(2')$$

with a normalization factor of $\frac{3}{4}$ after spin integration. Hence using this:

$$\langle \mathbf{S}_A \mathbf{S}_B \rangle = \tfrac{3}{4}\delta_{AB} tr(\gamma P_A) + \tfrac{6}{4} tr(P_A(1)P_B(2)R_0^{(0)}(1,1;2,2))$$

And then performing the integral:

$$\langle \mathbf{S}_A \mathbf{S}_B \rangle = \tfrac{3}{4}\delta_{AB} tr(\gamma P_A) - \underbrace{\tfrac{6}{4}\tfrac{1}{3}}_{\frac{1}{2}}\sum_{pqrs}(\Gamma^{pq}_{rs} + 2\Gamma^{rq}_{ps})P^A_{pq}P^B_{rs}$$

This is the final and perhaps most compact equation. The projection operator can be defined in very many different ways. The easiest is to Löwdin orthogonalize the basis set:

$$\left|\mu^A_L\right\rangle = \sum_{\nu^A}\left|\nu^A\right\rangle S^{-1/2}_{\mu\nu}$$

where L denotes the Löwdin basis. This means that molecular orbitals are expressed in the orthogonal basis as:

$$\mathbf{c}_L = \mathbf{S}^{+1/2}\mathbf{c}$$

and the density as:

$$\mathbf{P}_L = \mathbf{S}^{+1/2}\mathbf{P}\mathbf{S}^{+1/2}$$

The fragment projector is defined as:

$$P_A = \sum_{\mu_L \in A}|\mu_L\rangle\langle\mu_L|$$

Clark and Davidson suggest a slightly more elaborate projector in which first, the intra-fragment overlap is eliminated. This happens with a matrix U that for two fragments takes form:

$$\mathbf{U} = \begin{pmatrix} \mathbf{S}_A^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_B^{-1/2} \end{pmatrix}$$

where is the block of basis functions belonging to fragment A. Likewise:

$$\mathbf{U}^{-1} = \begin{pmatrix} \mathbf{S}_A^{+1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_B^{+1/2} \end{pmatrix}$$

Then the 'pre-overlap'is:

$$\bar{\mathbf{S}} = \mathbf{U}^\dagger\mathbf{S}\mathbf{U}$$

This contains the unit matrix in the intra-fragment blocks and non-zero elements elsewhere. This overlap matrix is the finally orthogonalized to obtain the globally orthogonal Löwdin basis. We finally transform the MO coefficients by the following transformation:

$$\mathbf{c}_L = \mathbf{S}^{+1/2}\mathbf{U}^{-1}\mathbf{c}$$

For the projectors, operating with the two MOs i and j gives:

$$\langle i|P_A|j\rangle = \sum_{\mu_L\in A}\sum_{\kappa^B_L\tau^C_L}\left\langle\kappa^B_L|\mu^A_L\right\rangle\left\langle\mu^A_L|\tau^C_L\right\rangle c^L_{\kappa i}c^L_{\tau j}$$

$$= \sum_{\mu_L\in A}\sum_{\kappa^B_L\tau^C_L}\delta_{AB}\delta_{AC}\delta_{\kappa\mu}\delta_{\tau\mu}c^L_{\kappa i}c^L_{\tau j}$$

$$= \sum_{\mu_L \in A} c^L_{\mu i} c^L_{\mu j}$$

Herrmann et al. (Herrmann, C.; Reiher, M.; Hess, B.A. J. Chem. Phys. 2005, 122, 34102) give the correct expression of the expectation values for a single spin-unrestricted determinant

$$\begin{aligned}\langle \mathbf{S}_A \mathbf{S}_B \rangle = & \tfrac{3}{4}\delta_{AB}\left\{\sum_i P^A_{ii} + \sum_{\bar{i}} P^A_{\bar{i}\bar{i}}\right\} \\ & + \tfrac{1}{4}\left\{\sum_{ij} P^A_{ii}P^B_{jj} + \sum_{\bar{i}\bar{j}} P^A_{\bar{i}\bar{i}}P^B_{\bar{j}\bar{j}} - \sum_{ij} P^A_{ij}P^B_{ij} - \sum_{\bar{i}\bar{j}} P^A_{\bar{i}\bar{j}}P^B_{\bar{i}\bar{j}} - \sum_{\bar{i}j} P^A_{\bar{i}\bar{i}}P^B_{jj} - \sum_{i\bar{j}} P^A_{ii}P^B_{\bar{j}\bar{j}}\right\} \\ & - \sum_{i\bar{j}} P^A_{i\bar{j}}P^B_{i\bar{j}}\end{aligned}$$

Which is used in the Orca implementation.

The use of the Local spin-implementation is very easy. All that is required is to divide the molecule into fragments. The rest happens automatically. For example, let us consider two nitrogen atoms at the dissociation limit. While the total spin state is S=0, the tow nitrogen atoms local exist in high-spin states (S=3/2). Consider the following test job:

```
! HF def2-SVP UHF TightSCF PModel

%scf brokensym 3,3 end

* xyz 0 1
N(1) 0 0 0
N(2) 0 0 1094
*
```

and the output:

```
-------------------
LOCAL SPIN ANALYSIS (Loewdin* projector)
-------------------


(1) A.E. Clark; E.R. Davison J. Chem. Phys. (2001), 115(16), pp 7382-7392
(2)  C. Herrmann, M. Reiher, B.A. Hess J. Chem. Phys. (2005) 122, art 034102-1

Number of fragments       = 2
Number of basis functions = 28
Number of atoms           = 2

   ... Fragment AO indices were mapped
   ... intra-fragment orthogonalization completed
   ... Global Loewdin orthogonalizer constructed
   ... Loewdin orthogonalized occupied orbitals constructed


 <SA*SB>        1            2
-----------------------------------
   1    :     3.7568
   2    :    -2.2500       3.7568


               <SzA>      Seff(A)
          -------------------------
   1    :     1.5000      1.5017
   2    :    -1.5000      1.5017
```

thus perfectly corresponding to the expectations. The same can be done at the CASSCF level:

```
! HF def2-SVP UHF TightSCF PModel

%casscf nel 6 norb 6 nroots 1 end

* xyz 0 1
N(1) 0 0 0
N(2) 0 0 1094
*
```

With the result:

```
<SA*SB>        1            2
---------------------------------
   1    :     3.7500
   2    :    -3.7500       3.7500


               <SzA>*      Seff(A)
         --------------------------
   1    :       n.a.      1.5000
   2    :       n.a.      1.5000

* = for a singlet state all <SzA> values are zero by definition
```

Thus, cleanly confirming the expectations.

As a less trivial example, consider a typical Fe(III) antiferromatically coupled transition metal dimer. An appropriate input may be:

```
! pbe def2-sv(p) tightscf kdiis soscf pmodel

%scf
  brokensym 5,5
end


* xyz -2 1
Fe(1)        -1.93818         0.53739        -0.00010
Fe(2)         1.06735         0.47031         0.00029
S(3)         -0.38935         2.59862        -0.00983
S(3)         -0.48170        -1.59050         0.01091
S(1)          2.68798         0.43924         1.99710
S(1)          2.68692         0.42704        -1.99712
S(2)         -3.55594         0.56407        -1.99889
S(2)         -3.55850         0.58107         1.99646
H(1)          3.91984         0.39462         1.47608
H(1)          3.91940         0.39536        -1.47662
H(2)         -4.78410         0.69179        -1.48280
H(2)         -4.78991         0.49249         1.47983
*
```

Where one of the bridging sulfurs was assigned to each site respectively.

```
<SA*SB>        1            2
---------------------------------
   1    :     7.7009
   2    :    -5.3721       7.7012


               <SzA>      Seff(A)
         --------------------------
   1    :     1.7579      2.3197
```

(continues on next page)

(continued from previous page)

```
   2    :    -1.7579      2.3198
```

Nice shows the expected results with the local site spins being close to their ideal value 2.5 which would hold for a high-spin Fe(III) ion.

7.53.7 UNO Orbital Printing

The analysis of UNO's can be controlled similarly. The flags together with their default values are shown below:

```
%output
  Print[ P_UNO_OccNum     ] = 1; # Occupation numbers
  Print[ P_UNO_AtPopMO_M  ] = 0; # Mulliken atom pop.
                                 # per UNO
  Print[ P_UNO_OrbPopMO_M] = 0;  # Mulliken orbital pop.
                                 # per UNO
  Print[ P_UNO_ReducedOrbPopMO_M] = 0;
                                 # Mulliken reduced orbital
                                 # pop. per UNO
  Print[ P_UNO_AtPopMO_L  ] = 0; # Loewdin atom pop.
                                 # per UNO
  Print[ P_UNO_OrbPopMO_L] = 0;  # Loewdin orbital pop.
                                 # per UNO
  Print[ P_UNO_ReducedOrbPopMO_L] = 0;
                                 # Loewdin reduced orbital
                                 # pop. per UNO
end
```

7.53.8 Hirshfeld Charges

The partitioning method by Hirshfeld is one of the most used approaches in the so-called atoms in molecules (AIM) methods.[394] In this case, the AIM density of atom A, $\rho_A(\vec{r})$ is written as:

$$\rho_A(\vec{r}) = \rho(\vec{r}) w_A(\vec{r}) \tag{7.423}$$

Here, $\rho(\vec{r})$ is the total charge density at position $\vec{r}$, and $w_A(\vec{r})$ a weighting function, that within the Hirshfeld method is equal to:

$$w_A(\vec{r}) = \frac{\rho_A^0(\vec{r})}{\rho^0(\vec{r})} \tag{7.424}$$

where $\rho_A^0(\vec{r})$ is the pro-atomic density of atom A and $\rho^0(\vec{r}) = \sum_A \rho_A^0(\vec{r})$ the pro-molecular density. The ratio in eq. (7.423) is known as *stockholder*. From eqs. (7.423) and (7.424) one can calculate the Hirshfeld charges as:

$$Q_A^{\text{Hirsh.}} = Z_A - \int \rho_A(\vec{r}) d\vec{r} \tag{7.425}$$

In ORCA, the pro-atomic density within the Hirshfeld method is calculated via density fitting with a set of Gaussian s-functions per element.

The calculation of the Hirshfeld charges in ORCA is requested by writing

```
! Hirshfeld
```

in the ORCA input file, or alternatively via the `%output` block:

```
%output
  Print[ P_Hirshfeld ] 1  # default = off
end
```

For instance, if we request the Hirshfeld charges for a water molecule:

```
!HF cc-pvdz tightscf Hirshfeld

%maxcore 4000

* xyz 0 1
  O    0.00000006589375       0.00157184228646       0.00000000004493
  H    0.77316868532439      -0.58666889665624      -0.00000000000005
  H   -0.77316876182122      -0.58666895650640      -0.00000000000005
*
```

ORCA prints the following information at the end of the output file:

```
------------------
HIRSHFELD ANALYSIS
------------------

Total integrated alpha density =      4.999998580
Total integrated beta density  =      4.999998580

  ATOM     CHARGE      SPIN
   0 O   -0.333756    0.000000
   1 H    0.166879    0.000000
   2 H    0.166879    0.000000

  TOTAL   0.000003    0.000000
```

7.53.9 MBIS Charges

The Minimal Basis Iterative Stockholder (MBIS) method is a variant of the Hirshfeld method.[869] The idea behind this approach is that the pro-atomic density $\rho_A^0(\vec{r})$ is expanded in a minimal set of atom-centered s-type Slater functions $\rho_{Ai}^0(\vec{r})$:

$$\rho_A^0(\vec{r}) = \sum_{i=1}^{m_A} \rho_{Ai}^0(\vec{r}) \tag{7.426}$$

with $\rho_{Ai}^0(\vec{r})$ equal to:

$$\rho_{Ai}^0(\vec{r}) = \frac{N_{Ai}}{\sigma_{Ai}^3 8\pi} \exp\left(-\frac{\left|\vec{r} - \vec{R_A}\right|}{\sigma_{Ai}}\right) \tag{7.427}$$

Here, m_A is the number of shells of atom A. The populations N_{Ai}, and the widths σ_{Ai} can be written as:

$$N_{Ai} = \int \rho(\vec{r}) \frac{\rho_{Ai}^0(\vec{r})}{\rho^0(\vec{r})} d\vec{r} \tag{7.428}$$

$$\sigma_{Ai} = \frac{1}{3N_{Ai}} \int \rho(\vec{r}) \frac{\rho_{Ai}^0(\vec{r})}{\rho^0(\vec{r})} \left|\vec{r} - \vec{R_A}\right| d\vec{r} \tag{7.429}$$

In order to compute the AIM densities $\rho_A(\vec{r})$, the MBIS method uses an iterative algorithm where: (1) an initial guess is generated for the set of N_{Ai} and σ_{Ai} and the pro-atomic densities are calculated through eqs. (7.426) and (7.427), (2) the new set of N_{Ai} and σ_{Ai} are obtained via eqs. (7.428) and (7.429), (3) if convergence is reached for $\rho_A(\vec{r})$, the iterative process stops, otherwise we go back to (1) but now one uses the last estimates for N_{Ai} and σ_{Ai}.

Once, the MBIS iterative process stops, the MBIS charges are calculated as:

$$Q_A^{\text{MBIS.}} = Z_A - \int \rho_A(\vec{r}) d\vec{r} \tag{7.430}$$

The calculation of the MBIS charges in ORCA is requested by writing

```
! MBIS
```

in the ORCA input file, or alternatively via the %output block:

```
%output
  Print[ P_MBIS ] 1  # default = off
end
```

If we request the MBIS charges for a HF calculation at the cc-pvdz level of a chloroform molecule:

```
!HF cc-pvdz tightscf MBIS

* xyz 0 1
  C   -0.00000997794639     -0.00091664148112      0.45499807439812
  H    0.00000069467312      0.00031189002174      1.53703126401237
  Cl   0.00003188789531      1.69433732001280     -0.08420513240263
  Cl   1.46635420502892     -0.84684178730039     -0.08421103795485
  Cl  -1.46637680965097     -0.84689178125304     -0.08420916805301
*
```

ORCA prints the following information at the end of the output file:

```
------------------
MBIS ANALYSIS
------------------

Convergence threshold (charges)      ...        1.000e-06
Number of iterations                 ...        46

Total integrated alpha density       ...     29.000001385
Total integrated beta density        ...     29.000001385

  ATOM     CHARGE    POPULATION     SPIN
   0 C    0.208633    5.791367    0.000000
   1 H    0.169417    0.830583    0.000000
   2 Cl  -0.126877   17.126877    0.000000
   3 Cl  -0.125586   17.125586    0.000000
   4 Cl  -0.125590   17.125590    0.000000

  TOTAL  -0.000003   58.000003    0.000000

 MBIS VALENCE-SHELL DATA:
  ATOM   POPULATION   WIDTH(A.U.)
   0 C    4.122213    0.508675
   1 H    0.830583    0.358785
   2 Cl   8.532439    0.524031
   3 Cl   8.531380    0.523959
   4 Cl   8.531381    0.523959
```

The second block corresponds to the valence Slater function, which is caracterized by its population $N_{A,v}$ and width $\sigma_{A,v}$.

The convergence threshold for the MBIS charges is set to 10^{-6}. However, it can be changed via the tag MBIS_CHARGETHRESH in the %method block:

```
%method
  MBIS_CHARGETHRESH 0.0001
end
```

ORCA can also print the following MBIS-related quantities: (1) atomic dipole moments, (2) atomic quadrupole moments, (3) atomic octupole moments, and (4) third radial moment of the MBIS density $\left(\langle r^3\rangle_A = \int \left|\vec{r} - \vec{R_A}\right|^3 \rho_A(\vec{r}) d\vec{r}\right)$. The printing of these properties is controlled by the tag MBIS_LARGEPRINT,

to be specified in the %method block:

```
%method
  MBIS_LARGEPRINT TRUE  # default = FALSE
end
```

If this option is activated, an extra iteration is performed after reaching the convergence threshold for the charges.

The origin for the calculation of the atomic dipole, quadrupole and octupole moments is the center of each atom (default). However, the user can also define a global origin (independent of the atom) through the tag MBIS_ORIGIN_MULT in the %method block:

```
%method
  MBIS_ORIGIN_MULT  CenterOfCoords     # origin of coordinate system (0,0,0)
                    CenterOfMass       # center of mass
                    CenterOfNucCharge  # center of nuclear charge
                    CenterXYZ          # arbitrary position, set coordinates with MBIS_ORIMULT_
↪XYZ
                    CenterOfEachAtom   # center of each atom (default)

  MBIS_ORIMULT_XYZ  x,y,z # set the coordinates, otherwise 0,0,0 (unit: Angstrom)
end
```

7.54 Orbital and Density Plots

There are two types of graphics output possible in ORCA - two dimensional contour plots and three dimensional surface plots. The quantities that can be plotted are the atomic orbitals, molecular orbitals, natural orbitals, the total electron density or the total spin density. The graphics is controlled through the block %plots.

7.54.1 Contour Plots

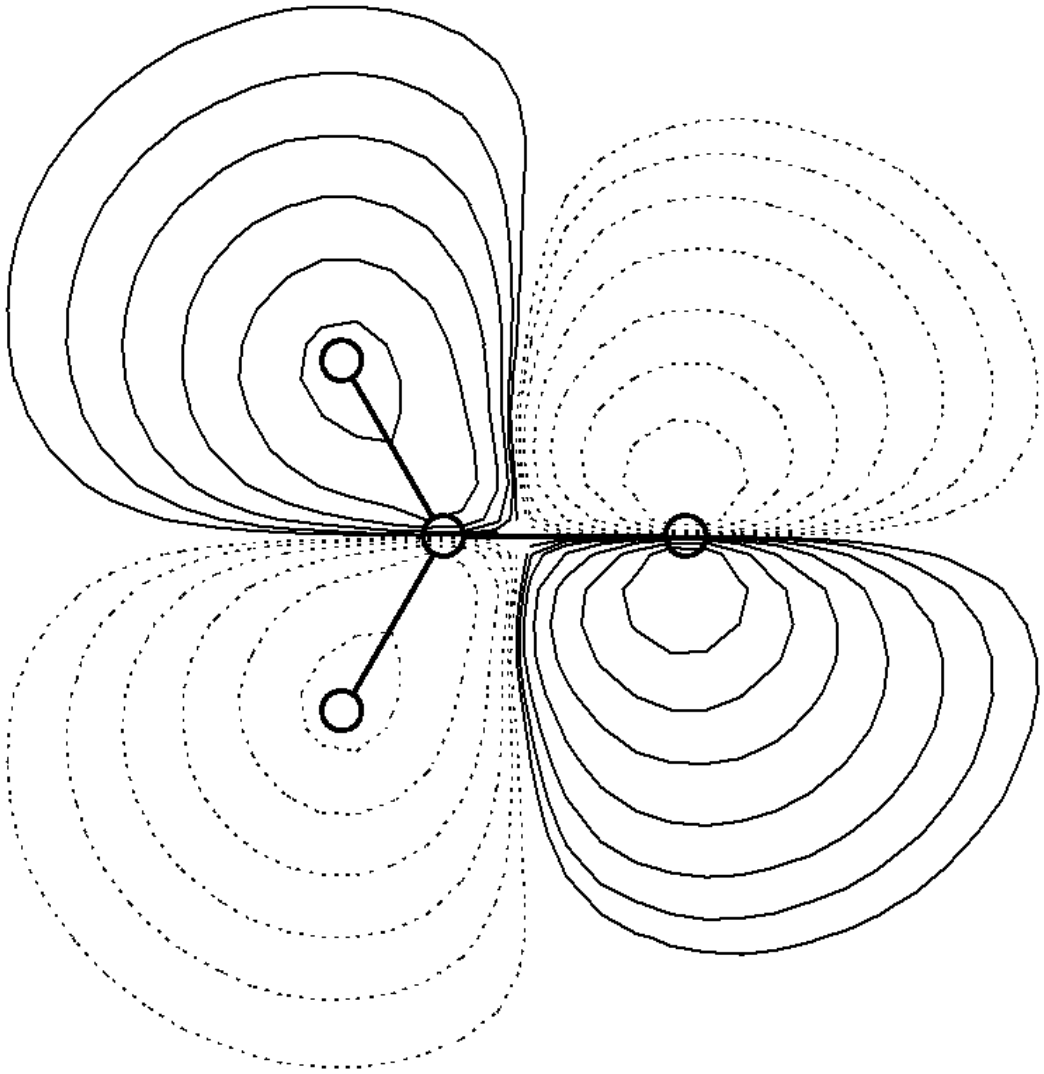
The contour plots are controlled via the following variables

```
%plots
  #*** the vectors defining the cut plane
  v1 0, 0, 0  # pointer to the origin
  v2 1, 0, 0  # first direction
  v3 0, 1, 0  # second direction
  #*** alternative to defining vectors. Use atom coordinates
  at1 0 # first atom defining v1
  at2 2 # second atom defining v2
  at3 4 # third atom defining v3
  #*** resolution of the contour
  dim1  45   # resolution in v2-direction
  dim2  45   # resolution in v3-direction
  #*** minimum and maximum values along v2 and v3
  min1  -7.0 # min value along v2 in bohr
  max1   7.0 # min value along v2 in bohr
  min2  -7.0 # min value along v3 in bohr
  max2   7.0 # max value along v3 in bohr
  #***
  UseCol   true   # Use color in the plot (blue=positive,
                  # red=negative)
  Skeleton true   # Draw Skeleton of the molecule of those
                  # atoms that are in or close to the cut
                  # plane
  Atoms    true   # Draw the atoms that are in the plane as
                  # circles
  NCont    200    # Number of contour levels.
```

(continues on next page)

(continued from previous page)

```
  ICont    0      # Draw NCont equally space contours
           1      # Start with 1/NCont and the double the
                  # value for each additional contour
  #*** the format of the output file
  Format   Origin   # straight ascii files
           HPGL     # plotter language files
  #*** the quantities to plot
  MO("MyOrbital-15xy.plt",15,0);  # orbital to plot
  v3= 0, 0, 1                     # change cut plane
  MO("MyOrbital-16xz.plt",16,0);  # orbital to plot
  ElDens("MyElDens.plt");         # Electron density
  SpinDens("MySpinDens");         # Spin density
end
```



The input was:
`v1 = 0, 0, 0; v2 = 1, 0, 0; v3 = 0, 1, 0; min1= -8; max1= 8; min2= -8; max2= 8; dim1= 50; dim2=50; Format = HPGL; NCont = 200; Icont = 1; Skeleton= true; Atoms = true; MO("Test-DFT-H2CO+-MO7xy.plt",7,1);`

NOTE:

- The command `MO("MyOrbital-15xy.plt",15,0);` is to be interpreted as follows: `MO` means that a MO is to be plotted. "`MyOrbital-15xy.plt`" is the file to be created. `15` is the number of the MO to be drawn (remember: counting starts at orbital 0!) and `0` is the operator the orbital belongs to. For a RHF (or RKS) calculation there is only one operator which has number 0. For a UHF (or UKS) calculation there are two operators - the spin-up orbitals belong to operator 0 and the spin-down orbitals belong to operator 1. For ROHF calculations there may be many operators but at the end all orbitals will be collected in one set of vectors. Thus the operator is always =0 in ROHF.
- The `ELDENS` (plot of the total electron density) and `SPINDENS` (plot of the total spin density) commands work analogous to the `MO` with the obvious difference that there is no MO or operator to be defined.
- Analogous to `ELDENS` and `SPINDENS`, post-HF densities can be selected using the keyword extended by the respective method. `ELDENSMDCI / SPINDENSMDCI` will plot the MDCI density, of course only if is available. `ELDENSMP2RE` and `SPINDENSMP2RE` will work with the MP2 relaxed density, while `ELDENSMP2UR` and `SPINDENSMP2UR` will yield the MP2 unrelaxed density. The OO-RI-MP2 densities can be requested by `ELDENSOO` or `SPINDENSOO`. Similarly, AutoCI relaxed densities can be plotted by using the `ELDENSAUTOCIRE` and `SPINDENSAUTOCIRE` keywords, and the unrelaxed densities by using `ELDENSAUTOCIUR` and `SPINDENSAUTOCIUR`.
- The `UNO` option plots natural orbitals of the UHF wavefunction (if they are available). No operator can be given for this command because there is only one set of UHF-NOs. Similarly, using `UCO` option can be used to plot the UHF corresponding orbitals.
- If the program cannot find the plot module ("Bad command or filename") try to use `ProgPlot="orca_plot.exe"` in the `%method` block or point to the explicit path.
- The defining vectors `v2` and `v3` are required to be orthonormal. The program will use a Schmidt orthonormalization of `v3` with respect to `v2` to ensure orthonormality. If you do not like this make sure that the input vectors are already orthogonal.
- `at1`, `at2` and `at3` can be used instead of `v1`, `v2` and `v3`. In this case say `v1` is taken as the coordinates of atom `at1`. Mixed definitions where say `v2` is explicitly given and say `v3` is defined through `at3` are possible. A value of -1 for `at1`, `at2` and `at3` signals that `at1`, `at2` and `at3` are not to be used. This type of definition may sometimes be more convenient.
- Variables can be assigned several times. The "actual" value a variable has is stored together with the command to generate a plot (`MO`, `ELDENS` or `SPINDENS`). Thus after each plot command the format or orientation of the plot can be changed for the next one.
- The `Origin` format produces a straightforward ASCII file with x, y and z values that can be read into your favorite contour plot program or you could write a small program that reads such files and converts them to whatever format is more appropriate for you.
- I usually use Word for Windows to open the `HPGL` files which appears to work fine. Double clicking on the graphics will allow modification of linewidth etc. For some reason that is not clear to me some graphics programs do not like the HPGL code that is produced by ORCA. If you are an HPGL expert and you have a suggestion - let me know.

7.54.2 Surface Plots

General Points

Surface plots can, for example, be created through an interface to Leif Laaksonen's ***gOpenMol*** program. This program can be obtained free of charge over the internet. It runs on a wide variety of platforms, is easy to use, produces high quality graphics and is easy to interface[1] - thank you Leif for making this program available!

The relevant `[PLOTS]` section looks like this:

[1] There were some reports of problems with the program on Windows platforms. Apparently it is better to choose the display settings as "true color 32 bit" rather than "high 16 bit". Thanks to Thomas Brunold!

```
%output
  XYZFile  true
end

%plots
  dim1  45   # resolution in x-direction
  dim2  45   # resolution in y-direction
  dim3  45   # resolution in z-direction
  min1  -7.0 # x-min value in bohr
  max1   7.0 # x-min value in bohr
  min2  -7.0 # y-min value in bohr
  max2   7.0 # y-max value in bohr
  min3  -7.0 # z-min value in bohr
  max3   7.0 # z-max value in bohr
  Format  gOpenMol_bin        # binary *.plt file
          gOpenMol_ascii      # ascii *.plt file
          Gaussian_Cube       # Gaussian-cube format
                              # (an ASCII file)
  MO("MyOrbital-15.plt",15,0);  # orbital to plot
  MO("MyOrbital-16.plt",16,0);  # orbital to plot
  UNO("MyUNO-48.plt",48);       # UHF-NO to plot
  ElDens("MyElDens.plt");       # Electron density
  SpinDens("MySpinDens.plt");   # Spin density
end
```

NOTE:

- it is admittedly inconvenient to manually input the dimension of the cube that is used for plotting. If you do nothing such that `min1 = max1 = min2 = max2 = min3 = max3=0` then the program will try to be smart and figure out a good cube size by itself. It will look at the minimum and maximum values of the coordinates and then add 7 bohrs to each dimension in the hope to properly catch all wavefunction tails.

Sometimes you will want to produce orbital plots after you looked at the output file and decided which orbitals you are interested in. In this case you can also run the `orca_plot` program in a crude interactive form by invoking it as:

```
orca_plot  MyGBWFile.gbw  -i
```

This will provide you with a subset of the capabilities of this program but may already be enough to produce the plots you want to look at. Note that for the name of the GBW-file you may as well input files that result from natural orbitals (normally `*.uno`), corresponding orbitals (normally `*.uco`) or localized orbitals (normally `*.loc`). Once in the interactive program, by entering '1' for 'Enter type of plot,' you will access a list of available plot capabilities relevant to your current calculation file (MyGBWFile.gbw):

```
-----------------------------------------------------------------------
Plot-Type is presently: 1
-----------------------------------------------------------------------
Searching for Ground State Electron or Spin Densities:              ...
-----------------------------------------------------------------------
     1 -   molecular orbitals
     2 -   (scf) electron density                          ......  (scfp                ␣
↪  )  => AVAILABLE
     3 -   (scf) spin density                              ......  (scfr                ␣
↪  )  => AVAILABLE
     4 -   natural orbitals
     5 -   corresponding orbitals
     6 -   atomic orbitals
     7 -   mdci electron density                          ......  (mdcip               ␣
↪  )  - NOT AVAILABLE
     8 -   mdci spin density                              ......  (mdcir               ␣
↪  )  - NOT AVAILABLE
     9 -   OO-RI-MP2 density                             ......  (pmp2re              ␣
```

(continues on next page)

(continued from previous page)

```
↪  )  - NOT AVAILABLE
    10 -   OO-RI-MP2 spin density                          ......  (pmp2ur              ␣
↪  )  - NOT AVAILABLE
    11 -   MP2 relaxed density                             ......  (pmp2re              ␣
↪  )  - NOT AVAILABLE
    12 -   MP2 unrelaxed density                           ......  (pmp2ur              ␣
↪  )  - NOT AVAILABLE
    13 -   MP2 relaxed spin density                        ......  (rmp2re              ␣
↪  )  - NOT AVAILABLE
    14 -   MP2 unrelaxed spin density                      ......  (rmp2ur              ␣
↪  )  - NOT AVAILABLE
    15 -   LED dispersion interaction density             ......  (ded21               ␣
↪  )  - NOT AVAILABLE
    16 -   Atom pair density
    17 -   Shielding Tensors
    18 -   Polarisability Tensor
    19 -   AutoCI relaxed density                          ......  (autocipre           ␣
↪  )  - NOT AVAILABLE
    20 -   AutoCI unrelaxed density                        ......  (autocipur           ␣
↪  )  - NOT AVAILABLE
    21 -   AutoCI relaxed spin density                     ......  (autocirre           ␣
↪  )  - NOT AVAILABLE
    22 -   AutoCI unrelaxed spin density                   ......  (autocirur           ␣
↪  )  - NOT AVAILABLE
------------------------------------------------------------------------
Searching for State or Transition State AO Electron Densities:      ...
------------------------------------------------------------------------
    23 -   CIS unrelaxed transition AO density             ......  (Tdens-CIS           ␣
↪  )  - NOT AVAILABLE
    24 -   ROCIS unrelaxed transition AO density          ......  (Tdens-ROCIS         ␣
↪  )  - NOT AVAILABLE
    25 -   CAS unrelaxed transition AO density            ......  (Tdens-CAS           ␣
↪  )  - NOT AVAILABLE
    26 -   ICE unrelaxed transition AO density            ......  (Tdens-ICE           ␣
↪  )  - NOT AVAILABLE
    27 -   MRCI unrelaxed transition AO density           ......  (Tdens-MRCI          ␣
↪  )  - NOT AVAILABLE
    28 -   LFT unrelaxed transition AO density            ......  (Tdens-LFT           ␣
↪  )  - NOT AVAILABLE
------------------------------------------------------------------------
Searching for State or Transition State MO Electron Densities:      ...
------------------------------------------------------------------------
    29 -   CIS unrelaxed transition MO density             ......  (Tdens-CISMO         ␣
↪  )  - NOT AVAILABLE
    30 -   ROCIS unrelaxed transition MO density          ......  (Tdens-ROCISMO       ␣
↪  )  - NOT AVAILABLE
    31 -   CAS unrelaxed transition MO density            ......  (Tdens-CASMO         ␣
↪  )  - NOT AVAILABLE
    32 -   ICE unrelaxed transition MO density            ......  (Tdens-ICEMO         ␣
↪  )  - NOT AVAILABLE
    33 -   MRCI unrelaxed transition MO density           ......  (Tdens-MRCIMO        ␣
↪  )  - NOT AVAILABLE
    34 -   LFT unrelaxed transition MO density            ......  (Tdens-LFTMO         ␣
↪  )  - NOT AVAILABLE
------------------------------------------------------------------------
Searching for State or Transition State QDPT AO Electron Densities: ...
------------------------------------------------------------------------
    35 -   CAS QDPT unrelaxed transition AO density        ......  (Tdens-CASQDSOC      ␣
↪  )  - NOT AVAILABLE
    36 -   DCDCAS QDPT unrelaxed transition AO density     ......  (Tdens-CASDCDQDSOC   ␣
↪  )  - NOT AVAILABLE
```

(continues on next page)

(continued from previous page)

```
   37 -   CAS CUSTOM E QDPT unrelaxed transition AO density   ......  (Tdens-CASCUSTOMEQDSOC
↪  )  - NOT AVAILABLE
   38 -   NEVPT2 QDPT unrelaxed transition AO density        ......  (Tdens-CASPTQDSOC
↪  )  - NOT AVAILABLE
   39 -   QDNEVPT2 QDPT unrelaxed transition AO density      ......  (Tdens-CASQDPTQDSOC
↪  )  - NOT AVAILABLE
   40 -   MRCI QDPT unrelaxed transition AO density          ......  (Tdens-MRCIQDSOC
↪  )  - NOT AVAILABLE
   41 -   ROCIS QDPT unrelaxed transition AO density         ......  (Tdens-ROCISQDSOC
↪  )  - NOT AVAILABLE
   42 -   LFT QDPT unrelaxed transition AO density           ......  (Tdens-LFTQDSOC
↪  )  - NOT AVAILABLE
```

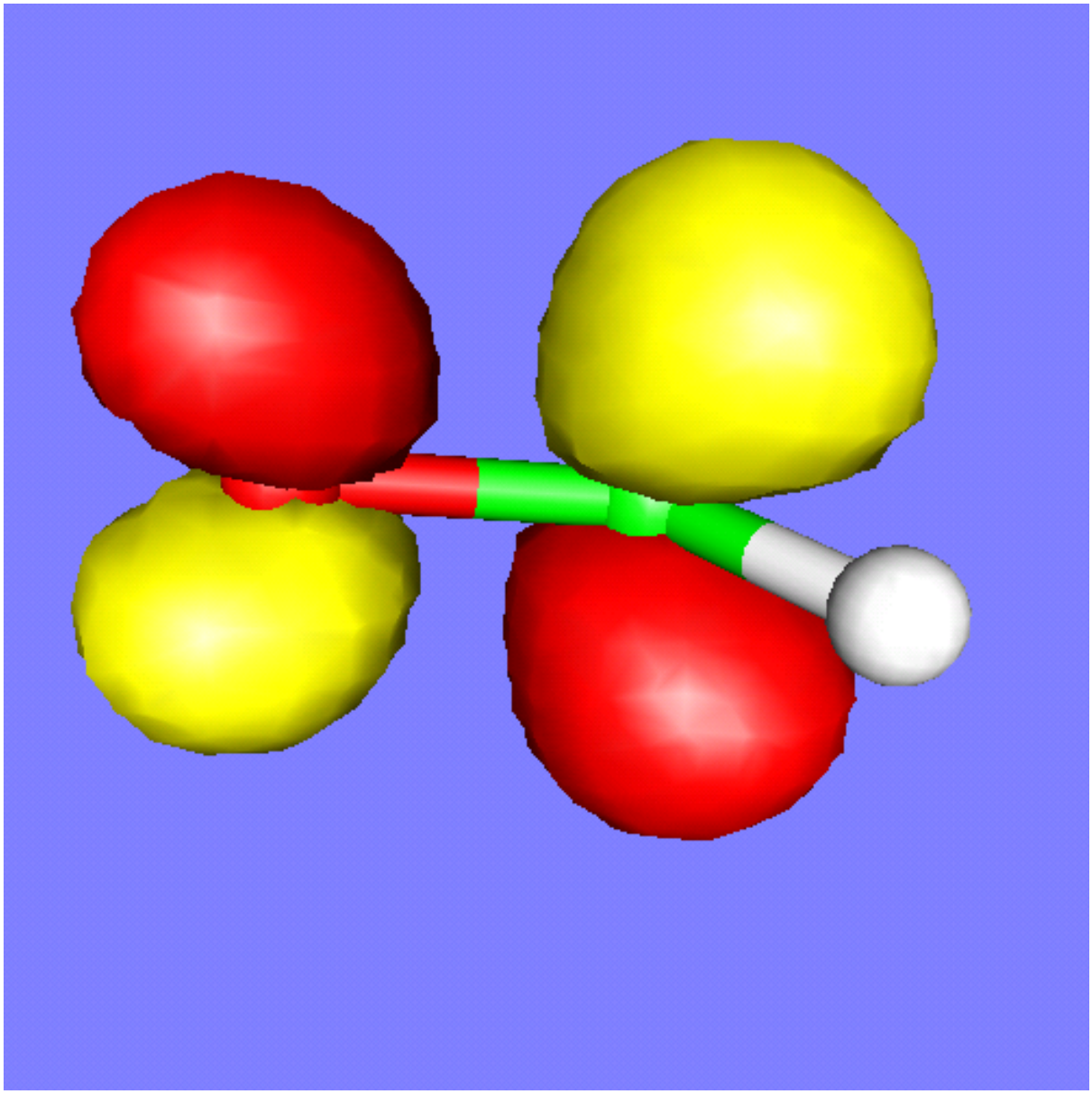


Fig. 7.64: The π^* orbital of H_2CO as calculated by the RI-BP/VDZP method.

FOD plots

The fractional occupation number weighted electron density (ρ^{FOD}, see *Fractional Occupation Numbers*) can be plotted in 3D for a pre-defined contour surface value which, after extensive testing, was set to the default value of $\sigma = 0.005$ e/Bohr^3. In order to allow comparison of various systems this value should be kept fix (in critical cases, one may also check the FOD plot with a a smaller value of $\sigma = 0.002$ e/Bohr^3 for comparison). The FOD is strictly positive in all space and resembles orbital densities (e.g., π-shape in large polyenes) or the total charge density for an ideal 'metal' with complete orbital degeneracy in trivial cases. FOD plots represent a cost-effective and robust way to identify the 'hot' (strongly correlated) electrons in a molecule and to choose appropriate approximate QC methods for a subsequent computational study of the systems in question. Based on our experience, the following rules of thumb can be derived:

- no significant ρ^{FOD}: use (double)-hybrid functionals or (DLPNO-)CCSD(T) (single-reference electronic structure)
- significant but rather localized ρ^{FOD}: use semi-local GGA functionals (or hybrid functional with low Fock-exchange, avoid HF or MP2; slight multi-reference character)
- significant *and* delocalized ρ^{FOD}: use multi-reference methods (or finite temperature DFT; strong multi-reference character)

Basically, ρ^{FOD} can be plotted analogously to an electron density calculated with ORCA using `Basename.scfp_fod` instead of `Basename.scfp`. The required `Basename.scfp_fod` is stored in the `Basename.densities` container. To print all available densities use the `9 - List all available densities` in `orca_plot`:

```
---------------------
List of density names
---------------------

Index:                                                    Name of Density
------------------------------------------------------------------------
    0:                                                      orca.scfp_fod   <--- required for␣
↪FOD plot
    1:                                                          orca.scfp
    2:                                                          orca.scfr
```

Note that producing `*.cube` files with `orca_plot` (see *orca_plot*) may take a considerable amount of time for larger molecules, particularly if high quality plots for publication purposes (i.e., 120x120x120 resolution) are wanted. An example FOD plot (singlet ground sate of *p*-benzyne, see *Fractional Occupation Numbers* for the corresponding ORCA input) is shown in Fig. 7.65. It has been produced with the ***UCSF CHIMERA*** program (this program can be obtained free of charge over the internet: https://www.cgl.ucsf.edu/chimera/) using the `*.cube` file generated with `orca_plot`:

```
orca_plot  pbenzyne.gbw  -i

user input:
1 (type of plot)
2 (electron density)
n (default name: no)
pbenzyne.scfp_fod (name of the FOD file)
4 (number of grid intervals)
120 (NGrid)
5 (output file format)
7 (cube)
10 (generate plot)
11 (exit)
```

It is also possible to generate `*.cube` files from ρ^{FOD} (analogously to electron density plots) with other programs that can read ORCA`BaseName.gbw` and electron density files by simply using the `Basename.scfp_fod` file instead of the `Basename.scfp` file.

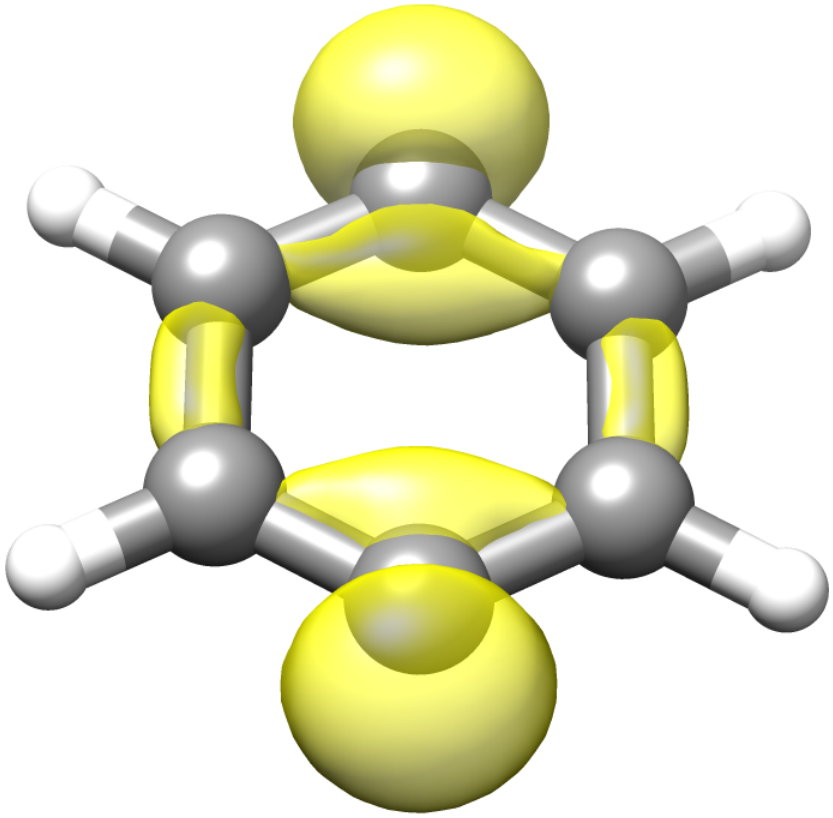


Fig. 7.65: FOD plot at $\sigma = 0.005$ e/Bohr^3 (TPSS/def2-TZVP (T = 5000 K) level) for the ^1A_g ground state of p-benzyne (FOD depicted in yellow).

The significant and rather delocalized FOD for p-benzyne (^1A_g) indicates that multi-reference methods would be needed for reliable computational study of this molecule (category c)). More examples of FOD plots generated with the same setup and programs can be found in the original publication and corresponding supplementary information.[327]

Interface to gOpenMol

Here is a short summary of how to produce these plots with gOpenMol:

- First of all the molecular geometry must be save by choosing `XYZFile=true` in the `[OUTPUT]` block. This will produce a straightforward ascii file containing the molecular geometry (or simply `! XYZFile`).
- After having produced the plot files start ***gOpenMol*** and choose ***File-Import-Coords*** . In the dialog choose the *XYZ* format and select the file. Then press ***apply*** and ***dismiss*** . The molecule should now be displayed in the graphics window.
- You can change the appearance by choosing ***View-Atom type*** .
- The color of the background can be changed with ***Colour-Background*** .
- After having done all this choose ***Plot-Contour*** and select the ***Browse*** button to select the appropriate file. Then press ***Import File*** to read it in. NOTE: you can only directly read files that were produced in `gOpenMol_bin` format. If you have chosen `gOpenMol_ascii` you must first use the ***gOpenMol*** file conversion utility under ***Run-Pltfile (conversion)*** to produce the binary plt file.
- After having read the plt file choose the appropriate isocontour value (one for the positive and one for the negative lobes of an orbital) and select suitable colors via ***Colour(n)*** to the right of the isocontour value. The ***Details*** button allows you to choose between solid and mesh representation and other things.

- Once the plot looks the way you like, use ***File-Hardcopy*** to produce a publication quality postscript or bitmap picture that can be imported into any word processing or graphics software.

Interface to Molekel

The Molekel program (http://ugovaretto.github.io/molekel/) is another beautiful and easy-to-use graphics tool that is recommended in combination with ORCA. You may even find it a little easier to use than gOpenMol but this may be a matter of personal taste. In order to produce plots with Molekel follow the following procedure:

- Produce Gaussian-Cube files (and optionally also an XYZ file) with ORCA as described above.
- Start Molekel and use the right mouse button to obtain the **Load** menu.
- Choose the format **xyz** to load your coordinates
- Use the right mouse button again to select the **Surface** menu
- Choose the format "Gaussian Cube" and click **Load Surface**
- Click on **Both Signs** if you visualize an orbital or spin density
- Select a suitable contour value in the **Cutoff** box.
- Click on **Create Surface**. That's it!
- In the **Color** menu (also available via the right mouse button) you can adjust the colors and in the main menu the display options for your molecule. Default settings are in a startup file that you can modify to suit your taste. More details are in the Molekel manual – check it out; it can do many other useful things for you too!

7.55 Utility Programs

7.55.1 orca_mapspc

This utility program is used to turn calculated spectra into a format that can be plotted with standard graphics programs. The usage is simple (for output examples see for example sections *Semiempirical Methods*, *IR Spectra*, *Raman Spectra* and *Resonant Inelastic Scattering Spectroscopy*):

```
orca_mapspc  file spectrum options

  file    = name of an ORCA output file
            name of an ORCA Hessian file (for IR and Raman)

  spectrum= abs    - Absorption spectra
            cd     - CD spectra
            ir     - IR spectra
            raman  - Raman spectra

  options  -x0value: Start of the x-axis for the plot
           -x1value: End of the x-axis for the plot
           -wvalue : Full-width at half-maximum height in
                     cm**-1 for each transition
           -nvalue : Number of points to be used
```

The exact abilities of `orca_mapspc` can be seen by simply executing the command in a terminal

```
orca_mapspc
```

Then one gets:

```
-------------------------------------------------------------------------------
usage: orca_mapspc Output-file { ABS, ABSV, ABSQ, ABSOI, CD, IR, RAMAN, NRVS, VDOS,
                       MCD, SOCABS, SOCABSQ,SOCABSOI, XES, XESV, XESQ, XESOI,
                       XAS, XASV, XASQ, XASOI, XESSOC, XESSOCQ, XESSOCOI,
                       XASSOC, RIXS, RIXSSOC} -options
-------------------------------------------------------------------------------
----General options ----
-o  output file
-cm use cm**-1 (default)
-eV use eV (default cm**-1)
-g  use Gaussian lineshape default
-l  use Lorentz  lineshape (only for Absorption and Emission like spectra)
-v  use Voigt lineshape    (only for Absorption and Emission like spectra)
-x0 initial point of spectrum
-x1 final   point of spectrum
-w  line width for Gaussian/Lorenzian linewidth
-q  line width for the Gaussian part of Voigt linewidth
-kw coeffitient for the line width calculated as kw*sqrt(energy)
-n  number of points
----The following additional options are for RIXS and RIXSSOC calculations----
-x2 initial point of the spectrum along y axis
-x3 final   point of the spectrum along yaxis
-g  line width for Gaussian/Lorenzian linewidth along y axis
-m  number of points for the emission spectrum
-eaxis plot option for the emission axis: (1) for Energy transfer
                                         (2) for emission spectrum
-uex  number of user defined cuts at constant Excitation Energy axis
-udw  number of user defined cuts at constant Emission/ Energy Transfer axis
-dx   number for shifting the spectra along the Excitation /Emission Energy axis
-kg coeffitient for the line width calculated as kg*sqrt(energy)
----Using external files----
paras.inp: a list of energy ranges with desired broadening parameters
for x axis: E_start E_stop Width
for y axis: 0 0 0 E_start E_stop Width
for xy axis: E_start1 E_stop1 Width1 E_start2 E_stop2 Width2
udex.inp: list of energies for taking cuts
          at constant Excitation Energy axis (RIXS/RIXSOC)
udem.inp: list of energies for taking cuts
          at constant Emission/ Energy Transfer axis (RIXS/RIXSOC)
gfsp.inp: list of ground-final state pairs to generate
          individual state pair RIXS planes
          and respective analysis planes (ROCIS RIXS/RIXSOC)
-------------------------------------------------------------------------------
```

NOTE:

- The input to this program can either be a normal output file from an ORCA calculation or a ORCA .hess file if IR or Raman spectra are desired
- Unless it is specified otherwise the default lineshape is always assumed to be a Gaussian
- There will be two output files:
  - `Input-file.spc.dat` (spc=abs-like or cd or ir or raman): This file contains the data to be plotted
  - `Input-file.spc.stk`: This file contains the individual transitions (wavenumber and intensity)
- The absorption plot has five columns: The first is the wavenumber in reciprocal centimeters, the second the total intensity and the third to fifth are the individual polarizations (i.e. assuming that the electric vector of the incoming beam is parallel to either the input x-, or y- or z-axis respectively). The last three columns are useful for interpreting polarized single crystal spectra.
- Generation of multiple spectra. When more than one spectra of the same kind are available the program will try to plot them. For example in the case of a CASSCF calculation with the NEVPT2 flag on, there will be two Absorption spectra (CASSCF and NEVPT2) that can be ploted

For example:

```
orca_mapspc My-CASSCF/NEVPT2-Output.out SOCABS -x07000 -x18000 -eV -n10000 -w2.0 -l
```

```
Mode is SOCABS
Entering SOC-ABS reading
Using eV units
Using Lorentzian shape
Multiple SOCABS (2) spectra detected ...
---------------------------------
Plotting SOCABS Spectrum 0
---------------------------------
Cannot read the paras.inp file ...
taking the line width parameter from the command line
Number of peaks            ...   4455
Start energy [eV]          ...  7000.00
Stop energy [eV]           ...  8000.00
Peak FWHM [eV]             ...     2.00
Number of points           ...   10000
---------------------------------
Plotting SOCABS Spectrum 1
---------------------------------
Cannot read the paras.inp file ...
taking the line width parameter from the command line
Number of peaks            ...   4455
Start energy [eV]          ...  7000.00
Stop energy [eV]           ...  8000.00
Peak FWHM [eV]             ...     2.00
Number of points           ...   10000
```

This will generate two kind of spectra one for the CASSCF and one for the NEVPT2 calculation

```
CASSCF:
       My-CASSCF/NEVPT2-Output.out.0.socabs.dat
       My-CASSCF/NEVPT2-Output.out.0.socabs.stk
NEVPT2:
       My-CASSCF/NEVPT2-Output.out.1.socabs.dat
       My-CASSCF/NEVPT2-Output.out.1.socabs.stk
```

Other Absorption or CD spectra can also be generated in the same way.

7.55.2 orca_chelpg

This program calculates CHELPG atomic charges according to Breneman and Wiberg[124]. The atomic charges are fitted to reproduce the electrostatic potential on a regular grid around the molecule, while constraining the sum of all atomic charges to the molecule's total charge. An additional constraint can be added, so the CHELPG charges also reproduce the total dipole moment of the molecule.

The program works with default values in the following way:

```
orca_chelpg  MyJob.gbw
```

The program uses three adjustable parameters, which can also be set in a separate chelpg input block

```
%chelpg
  GRID 0.3       # Spacing of the regular grid in Angstroems
  RMAX 2.8       # Maximum distance of all atoms to any gridpoint in Angstroems
  VDWRADII COSMO # VDW Radii of Atoms, COSMO default
           BW    # Breneman, Wiberg radii
  DIPOLE FALSE   # If true, then the charges also reproduce the total dipole moment
end
```

In this case, ORCA automatically calculates the CHELPG charges at the end of the calculation. Automatic calculation of CHELPG charges using the default values can also be achieved by specifying

```
! CHELPG
```

in the simple input section. By default the program uses the COSMO VDW radii for the exclusion of gridpoints near the nuclei, as these are defined for all atoms. The BW radii are similar, but only defined for very few atom types.

The charges may exhibit some dependence on the molecule's orientation in space, or some artificial variations in symmetric molecules. These effects can be minimized by increasing the CHELPG grid size, either by setting the GRID parameter in the CHELPG block, or in the one-liner via

```
! CHELPG(LARGE)
```

If one wants that the calculated CHELPG charges reproduce the total dipole moment of the molecule, as well as the electrostatic potential, then the following tag has to be added to the %chelpg block:

```
%chelpg
  DIPOLE TRUE  # The default is set to FALSE
end
```

In particular, the constraint affects the x,y,z components of the total dipole moment, so they reproduce the exact 3 components of the total dipole moment calculated via one-electron integrals.

7.55.3 orca_pltvib

This program is used in conjunction with gOpenMol (or xmol) to produce animations or plots of vibrational modes following a frequency run. The usage is again simple and described in section *Animation of Vibrational Modes* together with a short description of how to produce these plots in gOpenMol.

The program produces 20 frames of animation, where first and last frame correspond to the TS, all others calculated as $sin(2\pi frame/20 - 1) * displacement$.

7.55.4 orca_vib

This is a small "standalone" program to perform vibrational analysis. The idea is that the user has some control over things like the atomic masses that enter the prediction of vibrational frequencies but are independent of the electronic structure calculation as such.

The program takes a "`.hess`" file as input and produces essentially the same output as follows the frequency calculation. The point is that the "`.hess`" is a user-editable textfile that can be manually changed to achieve isotope shift predictions and the like. The usage together with an example is described in section *Isotope Shifts*. If you pipe the output from the screen into a textfile you should also be able to use `orca_mapspc` to plot the modified IR, Raman and NRVS spectra.

7.55.5 orca_loc

Localization is a widely used technique nowadays. By defining different functionals, various localization methods are established. The most favorable localization methods are developed by Foster-Boys and Pipek-Mezey. In ORCA there are four different localization methods available, the Pipek-Mezey method (PM), the Foster-Boys method (FB), the intrinsic atomic orbitals (IAO) based PM method and the IAO based FB method.

For Foster-Boys localization there are three different algorithms: first there is the conventional algorithm (FB). Second, there is an alternative algorithm (NEWBOYS), which is faster and could be used, for example, to localize the virtual MOs of a large system.

The third Foster-Boys algorithm is based on an augmented Hessian procedure (AHFB). It is particularly suited to obtain very tightly converged orbitals if an appropriate tolerance is requested (useful for local correlation).

Furthermore, it systematically converges towards a local minimum, rather than a different type of stationary point. The method proceeds in three stages. An initial set of localized orbitals is obtained through the NEWBOYS method. This is followed by an augmented Hessian maximization (rational function optimization) either with direct or with Davidson diagonalization, depending on the number of orbitals. Efficiency is therefore achieved for small and large systems alike. If the optimization fails to proceed but the augmented Hessian has got the correct eigenvalue structure, a Newton-Raphson maximization is triggered as the third stage. Currently, the only user-adjustable parameter of the AHFB method is the tolerance Tol. Convergence is signalled when the eigenvalue structure is correct, and the largest element of the orbital gradient, $4\langle i|\mathbf{r}|j\rangle\left(\langle j|\mathbf{r}|j\rangle - \langle i|\mathbf{r}|i\rangle\right)$, is below Tol. This is different from the other localization methods, which take the difference in the localization sum between two successive iterations as the convergence criterion.

The intrinsic atomic orbitals and intrinsic bond orbitals **(IAOIBO)** localization method is developed by Gerald Knizia, see Ref. [449]. In IAOIBO method, the occupied MOs are projected to a minimal basis set to get the IAOs, firstly. In ORCA different from original IAOIBO method, the converged SCF MO of atoms are used instead of Huzinaga MINI or STO-3G. However, the IAO charges computed by our method are quite similar to original IAO. Then, Pipek-Mezey functional is employed to localize these IAOs to IBOs. Finally,IBOs will be backtransformed to their original basis set. The IAO partial charges of canonical MOs for each atom is also printed out before the IAOIBO localization. But make sure you have included all occupied MOs in the IAOIBO localization. Otherwise, the IAO charges are meaningless. We further improved the original IAOIBO method by using the FB functional instead of PM functional. The computational time of the new method named **IAOBOYS** should be faster than the standard FB method for large systems. However, the IAO based method can only be used for the localization of occupied MOs.

There are two ways to do the MO localization in ORCA . The simpler way is to request the localization at the end of any ORCA calculation input file. Details are set in the %loc block.

```
%loc
LocMet      PM      # Localization method e.g. PIPEK-MEZEY
            FB      # FOSTER-BOYS
            IAOIBO  # IAOIBO
            IAOBOYS # IAOBOYS
            NEWBOYS # FOSTER-BOYS
            AHFB    # Augmented Hessian Foster-Boys
Tol         1e-6    # absolute convergence tolerance for the localization sum
                    # default value is 1e-6
                    # In the case of AHFB, however, this is the gradient threshold!
Random      0       # Always take the same seed for start for localization
                    # (For testing/debug purpose,optional)
            1       # Take a random seed for start of localization (default)
PrintLevel  2       # Amount of printing
MaxIter     64      # Max number of iterations
T_Bond      0.85    # Thresh that classifies orbitals in bond-like at the printing
T_Strong    0.95    # Thresh that classifies orbitals into strongly-localized at
                    # the printing
OCC         true    # Localize the occupied space
T_CORE      -99.9   # The Energy window for the first OCC MO to be localized (in a.u.)
                    # Here, we localize all occupied MOs including core orbitals.
VIRT        true    # Localize the virtual space
end
```

The localized MOs are obtained iteratively. Convergence is achieved when the localization functional value is self-consistent (contraled by Tol). Setting the flags OCC/VIRT to true will request a localization of the subspace. If both flags are set, two consecutive localizations are performed. The localized orbitals are stored in the form of a standard GBW file named *.loc*. Keep in mind that the localization of the occupied orbitals might change the total energy depending on what type of calculation you want to perform thereafter. For RHF and UHF there shouldn't be any problems, but for CASSCF the keyword OCC is not sufficient. **CASSCF is not invariant to rotation of *all* the occupied orbitals.**

The other way to do the localization is calling the `orca_loc` program directly from shell, which is more general. The `orca_loc` program requires an input of its own. The input is a textfile containing the necessary parameters. If no input is specified, `orca_loc` returns a help-file with a description of the necessary input-parameters. You need to specify in/output gbw-files, along with orbital ranges and the localization method to be used. A source of

confusion is the operator line op (alpha = 0 or beta = 1). For RHF(ROHF) and CASSCF, this should be set to zero. The input file usually looks like,

```
Myjob.gbw     # input orbitals
Myjob.loc.gbw # output orbitals
10            # orbital window: first orbital to be localized e.g. first active
15            # orbital window: last orbital to be localized e.g. last active
0             # localization method:
              # 1=PIPEK-MEZEY,2=FOSTER-BOYS,3=IAO-IBO,4=IAO-BOYS,5=NEW-BOYS,6=AHFB
# The following parameters are optional
# However, if you want to change one of them, all preceding ones have to be set, too.
0             # operator: 0 for alpha, 1 for beta
128           # maximum number of iterations
1e-6          # convergence tolerance of the localization functional value
0.0           # relative convergence tolerance of the localization functional value
0.95          # printing thresh to call an orbital strongly localized
0.85          # printing thresh to call an orbital bond-like
2             # printlevel
1             # use Cholesky Decomposition (0=false, 1=true)
1             # randomize seed for localization (0=false, 1=true)
```

If the input file is called myloc.inp, running "orca_loc myloc.inp" will produce the Myjob.loc.gbw file containing the localized orbitals. Please make sure the Myjob.gbw is in the same directory as myloc.inp.

7.55.6 orca_blockf

This utility program allows the canonicalization of orbitals (.gbw file) for arbitrary subspaces. With canonicalization we refer to the block diagonalization of the Fock matrix. Note that the necessary Fock matrix must be generated and be available on disk prior calling `orca_blockf`. The program is described in section *Local Zero-Field Splitting*, where the Local ZFS decomposition is discussed.

7.55.7 orca_plot

Orca_plot is a utility program that can be used to generate 2D and 3D graphics of various types of orbitals and densities during an ORCA run. It can in principle called in two ways:

1) Within a calculation input via the %block section. This is described in section *Orbital and Density Plots*. In this way it can be used to create graphics (2D) or (3D) data for visualization.

2) It is also possible to run this program interactively. The input parameters are:

```
gbwfile     # name of gbw-file
-i          # interactive use of orca_plot
-m 256      # max. memory in MB (if needed)
```

You will then get a simple, self-explaining menu that will allow you to generate a variety of files (such as `.plt` and `.cube`) directly from the `.gbw` files without restarting or running a new job. If needed, the -m-option allows to control the memory usage of your plotting job.

The listed utilities are printed by writing in the terminal

```
orca_plot my.gbw -i
```

```
1 - Enter type of plot
2 - Enter no of orbital to plot
3 - Enter operator of orbital (0=alpha,1=beta)
4 - Enter number of grid intervals
5 - Select output file format
6 - Plot CIS/TD-DFT difference densities
7 - Plot CIS/TD-DFT transition densities
```

(continues on next page)

(continued from previous page)

```
8 - Set AO(=1) vs MO(=0) to plot
9 - List all available densities
10 - Perform Density Algebraic Operations

11 - Generate the plot
12 - exit this program
```

Perform Orbital Plots

Let's assume the pyridine molecule in the following input:

```
!def2-SVP

*xyz 0 1
C        0.690940233      0.417992301     -1.170801378
C        0.690940233      1.616339301     -0.458357378
C        0.690940233      1.560238301      0.936438622
N        0.690940233      0.417992301      1.635468622
C        0.690940233     -0.724253699      0.936438622
C        0.690940233     -0.780354699     -0.458357378
H        0.690940233      0.417992301     -2.257043378
H        0.690940233      2.574997301     -0.967574378
H        0.690940233      2.478336301      1.521602622
H        0.690940233     -1.642351699      1.521602622
H        0.690940233     -1.739012699     -0.967574378
*
```

We can select to plot the HOMO from the list of Occupied Orbitals

```
----------------
ORBITAL ENERGIES
----------------

 NO   OCC          E(Eh)            E(eV)
  0   2.0000     -15.563726      -423.5105
...
Occupied Orbitals Manifold
...
 20   2.0000      -0.349834        -9.5195
...
Unoccupied Orbitals Manifold
...
 21   0.0000       0.111722         3.0401
...
```

For this we modify options 2, 3, 4 and 8 as:

```
Enter a number: 2
Enter MO: 20

Enter a number: 3
Enter OP: 0

Enter a number: 4
Enter NGRID: 80

Enter a number: 5
File-Format is presently: 7
(7 -  3D   Gaussian cube)
```

(continues on next page)

(continued from previous page)

```
Enter a number: 8
Enter 0(MO) or 1(AO): 0
```

And we generate the plot as:

```
11 - Generate the plot
Enter a number: 11 =>

PlotType       ... MO-PLOT
MO/Operator    ... 20 0
Output file    ... pyridine_scf.mo20a.cube
Format         ... Grid3d/Cube
Resolution     ... 80 80 80

Calling PlotGrid3d with ATOM-A,B=0,0
Entering PlotGrid3d with Plottype =1
                 *** PLOTTING FINISHED ***
Output file: pyridine_scf.mo20a.cube
```

We can now use any visualization software e.g. chimera to plot the generated HOMO 20 orbital:(Figure: Fig. 7.66)

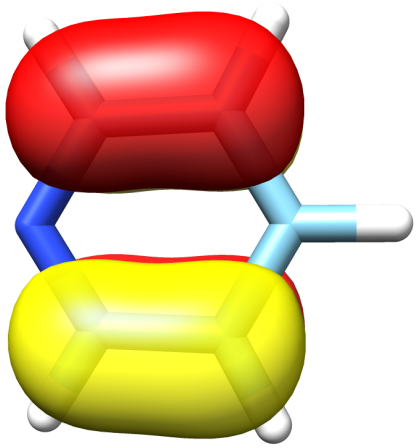


Fig. 7.66: Pyridine HOMO

List of Density Plots

If a density instead of an orbital plot is required options 1 and 9 can be used to list the available densities.

For example option 1 in the above example will provide the computed available densities.

```
Enter a number: 1
------------------------------------------------------------------------
Reading Over   2 Saved Densities                                    ...
------------------------------------------------------------------------
------------------------------------------------------------------------
Plot-Type is presently: 1
------------------------------------------------------------------------
Searching for Ground State Electron or Spin Densities:              ...
------------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
     1 -   molecular orbitals
     2 -   (scf) electron density                          ......  (scfp                ␣
↪  )  => AVAILABLE
     3 -   (scf) spin density                              ......  (scfr                ␣
↪  )  - NOT AVAILABLE
     4 -   natural orbitals
     5 -   corresponding orbitals
     6 -   atomic orbitals
     7 -   mdci electron density                          ......  (mdcip               ␣
↪  )  - NOT AVAILABLE
     8 -   mdci spin density                              ......  (mdcir               ␣
↪  )  - NOT AVAILABLE
     9 -   OO-RI-MP2 density                             ......  (pmp2re              ␣
↪  )  - NOT AVAILABLE
    10 -   OO-RI-MP2 spin density                        ......  (pmp2ur              ␣
↪  )  - NOT AVAILABLE
    11 -   MP2 relaxed density                           ......  (pmp2re              ␣
↪  )  - NOT AVAILABLE
    12 -   MP2 unrelaxed density                         ......  (pmp2ur              ␣
↪  )  - NOT AVAILABLE
    13 -   MP2 relaxed spin density                      ......  (rmp2re              ␣
↪  )  - NOT AVAILABLE
    14 -   MP2 unrelaxed spin density                    ......  (rmp2ur              ␣
↪  )  - NOT AVAILABLE
    15 -   LED dispersion interaction density           ......  (ded21               ␣
↪  )  - NOT AVAILABLE
    16 -   Atom pair density
    17 -   Shielding Tensors
    18 -   Polarisability Tensor
    19 -   AutoCI relaxed density                        ......  (autocipre           ␣
↪  )  - NOT AVAILABLE
    20 -   AutoCI unrelaxed density                      ......  (autocipur           ␣
↪  )  - NOT AVAILABLE
    21 -   AutoCI relaxed spin density                   ......  (autocirre           ␣
↪  )  - NOT AVAILABLE
    22 -   AutoCI unrelaxed spin density                 ......  (autocirur           ␣
↪  )  - NOT AVAILABLE
```

In this case the (scf) electron density is available and can be chosen to be visualized in a similar process as described above for the orbitals.

```
Enter Type: 2
The default name of the density would be: pyridine_scf.scfp
Is this the one you want (y/n)?
```

While Option 9 will give us the name of this density:

```
---------------------
List of density names
---------------------

Index:                                                Name of Density
-------------------------------------------------------------------------
    0:                                              pyridine_scf.scfp
```

Following the above steps we can visualize the SCF electron density of pyridine

```
PlotType       ... DENSITY-PLOT
ElDens File    ... pyridine_scf.scfp
Output file    ... MyElDens
Format         ... Grid3d/Cube
```

(continues on next page)

(continued from previous page)

```
Resolution     ... 80 80 80

Calling PlotGrid3d with ATOM-A,B=0,0
Entering PlotGrid3d with Plottype =2
                 *** PLOTTING FINISHED ***
 Output file: pyridine_scf.eldens.cube
```

(Figure: Fig. 7.67)

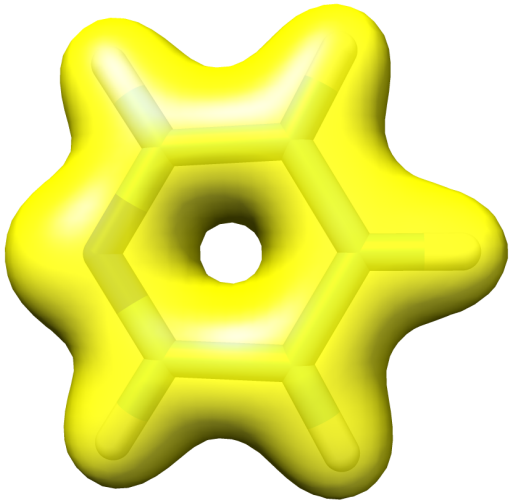


Fig. 7.67: Pyridine SCF Electron Density

Perform Algebraic Operations

Starting from ORCA 6 one can perform simple Algebraic Operations with the computed densities. The presently available operations are listed in menu option `10`:

```
Enter a number: 10
---------------------------------------------------------------------
Available Algebraic Operations:
---------------------------------------------------------------------
       1 - Pair Densities Addition
       2 - Pair Densities Subtraction
       3 - Pair Densities Multiplication
       4 - Pair Densities Division
       5 - Density Normalization
       6 - Make Natural Transition Orbitals
       7 - Make Natural Difference Orbitals
       8 - Leave Section - Return to Main Menu
```

The important step to these processes is to make sure that the desired State or Transition State Densities are available in the Densities file. To achieve this, one should request the desired densities to be stored in the disk after the calculation is executed by the commands `!KeepDens` or `!KeepTransDensity`. Please NOTE that storage of several hundreds or thousands of these densities need to be done with care from the user's perspective as they might occupy several hundreds of GB disk space!

Let us in addition on top of the SCF calculation of pyridine perform a canonical CCSD calculation with the following input:

```
!CCSD def2-SVP

*xyz 0 1
C        0.690940233      0.417992301     -1.170801378
C        0.690940233      1.616339301     -0.458357378
C        0.690940233      1.560238301      0.936438622
N        0.690940233      0.417992301      1.635468622
C        0.690940233     -0.724253699      0.936438622
C        0.690940233     -0.780354699     -0.458357378
H        0.690940233      0.417992301     -2.257043378
H        0.690940233      2.574997301     -0.967574378
H        0.690940233      2.478336301      1.521602622
H        0.690940233     -1.642351699      1.521602622
H        0.690940233     -1.739012699     -0.967574378
*
```

Option 1 in the `orca_plot` menu will now provide us with both the scf and mdci electron densities.

```
-----------------------------------------------------------------------
Plot-Type is presently: 1
-----------------------------------------------------------------------
Searching for Ground State Electron or Spin Densities:              ...
-----------------------------------------------------------------------
     1 -   molecular orbitals
     2 -   (scf) electron density                              ......  (scfp                ␣
↪  )  => AVAILABLE
     3 -   (scf) spin density                                  ......  (scfr                ␣
↪  )  - NOT AVAILABLE
     4 -   natural orbitals
     5 -   corresponding orbitals
     6 -   atomic orbitals
     7 -   mdci electron density                               ......  (mdcip               ␣
↪  )  => AVAILABLE
```

Hence we can proceed and in option 2 of the `Available Algebraic Operations` menu take their difference to produce the respective CCSD - HF electron correlation electron density as:

```
---------------------
List of density names
---------------------

Index:                                                  Name of Density
-----------------------------------------------------------------------
    0:                                              pyridine_ccsd.scfp
    1:                                             pyridine_ccsd.P0.tmp
    2:                                              pyridine_ccsd.mdcip


-----------------------------------------------------------------------
Performing Algebraic Operations Over Densities: => SUBTRACTION
-----------------------------------------------------------------------
Number of Densities to be Processed from the List =>   2:
-----------------------------------------------------------------------
Enter FileName for Density[  0]: pyridine_ccsd.scfp
Provide a Scale Factor for Density[  0] (Default => 1.00): 1
Enter FileName for Density[  1]: pyridine_ccsd.mdcip
Provide a Scale Factor for Density[  1] (Default => 1.00): 1
-----------------------------------------------------------------------
INTERPRETTING EQUATION:
-----------------------------------------------------------------------
1/sqrt(N) * [(1.00) * {pyridine_ccsd.scfp} - (1.00) * {pyridine_ccsd.mdcip}]
-----------------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
PlotType       ... DENSITY-PLOT
ElDens File    ... pyridine_ccsd.scfp_minus_pyridine_ccsd.mdcip
Output file    ... pyridine_ccsd.scfp_minus_pyridine_ccsd.mdcip.cube
Format         ... Grid3d/Cube
Resolution     ... 80 80 80
```

We can directly visualize the produced `pyridine_ccsd.scfp_minus_pyridine_ccsd.mdcip.cube` (Figure: Fig. 7.68)

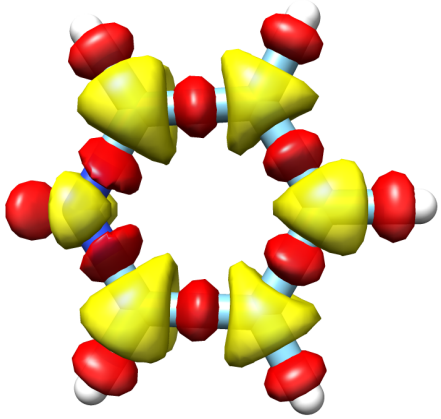


Fig. 7.68: Pyridine CCSD-HF Electron Density

NOTE: The generated density is stored in the Density Container so that it can be further processed or properly stored for future use

```
---------------------
List of density names
---------------------

Index:                                                  Name of Density
-------------------------------------------------------------------------
    0:                                                pyridine_ccsd.scfp
    1:                                               pyridine_ccsd.P0.tmp
    2:                                                pyridine_ccsd.mdcip
    3:                          pyridine_ccsd.scfp_minus_pyridine_ccsd.mdcip
```

Another useful utility option of `orca_plot` is that is able to generate Natural Transition Orbitals (NTOs) of Natural Difference Orbitals (NDOs) from any theory level available State or Transition Density. Let us show case an example. We now perform a SA-CASSCF (7,8) calculation on the pyridine molecule according to the input where we make sure that all relevant densities are kept on disk after the calculation by adding the `KeepTransDensity` in the keyword list.

```
!def2-SVP KeepTransDensity

%casscf
 nel 8
 norb 7
```

(continues on next page)

(continued from previous page)

```
 mult 1
 nroots 10
end


*xyz 0 1
C        0.690940233      0.417992301     -1.170801378
C        0.690940233      1.616339301     -0.458357378
C        0.690940233      1.560238301      0.936438622
N        0.690940233      0.417992301      1.635468622
C        0.690940233     -0.724253699      0.936438622
C        0.690940233     -0.780354699     -0.458357378
H        0.690940233      0.417992301     -2.257043378
H        0.690940233      2.574997301     -0.967574378
H        0.690940233      2.478336301      1.521602622
H        0.690940233     -1.642351699      1.521602622
H        0.690940233     -1.739012699     -0.967574378
*
```

Indeed after the calculation we now have besides the CASSCF electron densities all the relevant CASSCF State and Transition densities become now available.

```
-------------------------------------------------------------------------
Plot-Type is presently: 1
-------------------------------------------------------------------------
Searching for Ground State Electron or Spin Densities:              ...
-------------------------------------------------------------------------
     1 -   molecular orbitals
     2 -   (scf) electron density                              ......  (scfp                ␣
↪  )  => AVAILABLE
     3 -   (scf) spin density                                  ......  (scfr                ␣
↪  )  - NOT AVAILABLE
     ...
-------------------------------------------------------------------------
Searching for State or Transition State AO Electron Densities:      ...
-------------------------------------------------------------------------
    23 -   CIS unrelaxed transition AO density                 ......  (Tdens-CIS           ␣
↪  )  - NOT AVAILABLE
    24 -   ROCIS unrelaxed transition AO density               ......  (Tdens-ROCIS         ␣
↪  )  - NOT AVAILABLE
    25 -   CAS unrelaxed transition AO density                 ......  (Tdens-CAS           ␣
↪  )  => AVAILABLE
    ...
-------------------------------------------------------------------------
Searching for State or Transition State MO Electron Densities:      ...
-------------------------------------------------------------------------
    29 -   CIS unrelaxed transition MO density                 ......  (Tdens-CISMO         ␣
↪  )  - NOT AVAILABLE
    30 -   ROCIS unrelaxed transition MO density               ......  (Tdens-ROCISMO       ␣
↪  )  - NOT AVAILABLE
    31 -   CAS unrelaxed transition MO density                 ......  (Tdens-CASMO         ␣
↪  )  => AVAILABLE
   ...
```

We can now set ourselves to generate the NTOs and NDOs dominating the computed State 1

```
ROOT   1:  E=    -246.3609192216 Eh  5.176 eV  41749.8 cm**-1
      0.82391 [   346]: 2212100
      0.06447 [   295]: 2112200
```

Hence by using options `6 =>NTOs` or `7 =>NDOs` of the `Available Algebraic Operations` menu, and processing the respective:

1) AO Transition densities D_{01} for NTOs

```
-------------------------------------------------------------------------
Performing Algebraic Operations Over Densities: => MAKE_NTOS
-------------------------------------------------------------------------
Enter FileName for Density[  0]: Tdens-CAS-0-0-0-1
Provide a Scale Factor for Density[  0] (Default => 1.00): 1
-------------------------------------------------------------------------
NATURAL TRANSITION ORBITALS GENERATION:
-------------------------------------------------------------------------
Warning: The one-electron matrix doesn't exist - is recalculated (SHARK)
Calculating the overlap matrix            ... done!

------------------------------------------------
NATURAL TRANSITION ORBITALS FOR STATE     1 1A
------------------------------------------------

STATE    1 1A  :  E=   0.190226 au      5.176 eV    41749.8 cm**-1

Threshold for printing occupation numbers 1.0000e-04

     0  : n=  1.28519446
     1  : n=  0.03997952
     2  : n=  0.01505089
     3  : n=  0.00336119

=> Natural Transition Orbitals (donor   ) were saved in Tdens-CAS-0-0-0-1.1-1A_nto-donor.gbw
=> Natural Transition Orbitals (acceptor) were saved in Tdens-CAS-0-0-0-1.1-1A_nto-acceptor.gbw
-------------------------------------------------------------------------
Provide a Number of NTOs to plot (Default => 1): 1
-------------------------------------------------------------------------
-------------------------------------------------------------------------
Reading Donor NTO-file   : Tdens-CAS-0-0-0-1.1-1A_nto-donor.gbw
-------------------------------------------------------------------------
Generating cube file for Donor NTO[0]:=> Tdens-CAS-0-0-0-1.1-1A_nto-donor.gbw.0.cube
-------------------------------------------------------------------------
Reading Acceptor NTO-file   : Tdens-CAS-0-0-0-1.1-1A_nto-acceptor.gbw
-------------------------------------------------------------------------
Generating cube file for Acceptor NTO[0]:=> Tdens-CAS-0-0-0-1.1-1A_nto-acceptor.gbw.0.cube
Current-settings:

PlotType       ... MO-PLOT
MO/Operator    ... 0 0
Output file    ... Tdens-CAS-0-0-0-1.1-1A_nto-acceptor.gbw.0.cube
Format         ... Grid3d/Cube
Resolution     ... 80 80 80
```

2) AO State densities $D_{00} - D_{11}$ for NDOs

```
-------------------------------------------------------------------------
Performing Algebraic Operations Over Densities: => MAKE_NDOS
-------------------------------------------------------------------------
Enter FileName for Density[  0]: Tdens-CAS-0-0-0-0
Provide a Scale Factor for Density[  0] (Default => 1.00): 1
Enter FileName for Density[  1]: Tdens-CAS-0-0-1-1
Provide a Scale Factor for Density[  1] (Default => 1.00): 1
-------------------------------------------------------------------------
NATURAL DIFFERENCE ORBITALS GENERATION:
-------------------------------------------------------------------------
Warning: The one-electron matrix doesn't exist - is recalculated (SHARK)
Calculating the overlap matrix            ... done!

------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
NATURAL DIFFERENCE ORBITALS FOR STATE     1 1A
-----------------------------------------------

STATE    1 1A  :  E=   0.190226 au      5.176 eV    41749.8 cm**-1

Threshold for printing occupation numbers 1.0000e-04

     0  : n=  0.49881235
     1  : n=  0.08492368
     2  : n=  0.06708141
     3  : n=  0.01254920
     4  : n=  0.00667407
     5  : n=  0.00006162

=> Natural Difference Orbitals (donor   ) were saved in Tdens-CAS-0-0-0-0-Tdens-CAS-0-0-1-1.1-
↪1A_ndo-donor.gbw
=> Natural Difference Orbitals (acceptor) were saved in Tdens-CAS-0-0-0-0-Tdens-CAS-0-0-1-1.1-
↪1A_ndo-acceptor.gbw
------------------------------------------------------------------------
Provide a Number of NDOs to plot (Default => 1): 1
------------------------------------------------------------------------
------------------------------------------------------------------------
Reading Donor NDO-file   : Tdens-CAS-0-0-0-0-Tdens-CAS-0-0-1-1.1-1A_ndo-donor.gbw
------------------------------------------------------------------------
Generating cube file for Donor NDO[0]:=> Tdens-CAS-0-0-0-0-Tdens-CAS-0-0-1-1.1-1A_ndo-donor.
↪gbw.0.cube
------------------------------------------------------------------------
Reading Acceptor NDO-file   : Tdens-CAS-0-0-0-0-Tdens-CAS-0-0-1-1.1-1A_ndo-acceptor.gbw
------------------------------------------------------------------------
Generating cube file for Acceptor NDO[0]:=> Tdens-CAS-0-0-0-0-Tdens-CAS-0-0-1-1.1-1A_ndo-
↪acceptor.gbw.0.cube
Current-settings:

PlotType       ... MO-PLOT
MO/Operator    ... 0 0
Output file    ... Tdens-CAS-0-0-0-0-Tdens-CAS-0-0-1-1.1-1A_ndo-acceptor.gbw.0.cube
Format         ... Grid3d/Cube
Resolution     ... 80 80 80
```

It is possible to produce the corresponding Donor and Acceptor NTOs and NDOs orbital pairs which can be readily visualized in:

(Figure: Fig. 7.69)

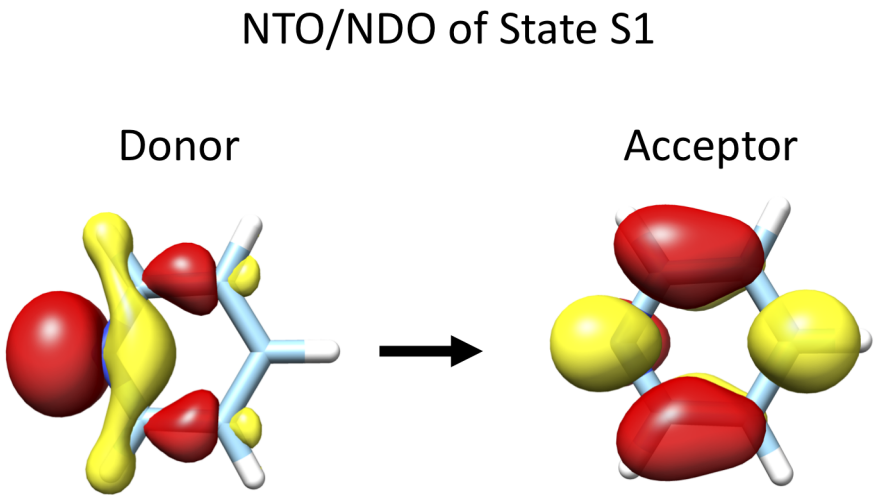


Fig. 7.69: Pyridine SA-CASSCF(8,7) NTO/NDO donor/acceptor orbital pairs for State 1

NOTE: It is beneficial and more correct to always process the AO basis Densities. In particular it is not possible nor is correct to process reduced MO basis Densities.

Additional TIPS regarding Density Plots

1) As an alternative to menu option 9 one can check, which densities are available, directly in the calculation directory by reading out the list of all densities contained in the densities-file:

```
orca_plot mymol.densities
```

2) When `orca_plot` is used for very expensive plots, it even can be called in parallel mode:

```
mpirun -np 4 /orca_path/orca_plot_mpi MyGBWFile.gbw -i MyGBWFile
```

3) It is possible to use `orca_plot` to create difference densities between the ground and excited states from CIS or TD-DFT calculations directly from the .cis calculation information. This is implemented as an extra interactive menu point that is (hopefully) self-explanatory. Starting from ORCA 6 one can use the `Algebraic Operations` utility menu to carry out these operations directly from the generated State and Transition Densities.

7.55.8 orca_2mkl: Old Molekel as well as Molden inputs

This little utility program can be used to convert gbw files into mkl files which are of ASCII format. This is useful since molekel can read these files and use them for plotting and the like. The contents of the mkl file is roughly the same as the gbw file (except for the internal flags of ORCA) but this is an ASCII file which can also be read for example by your own programs. It would therefore be a good point for developing an interface. It is likely that this functionality will be further expanded in the future.

```
orca_2mkl BaseName
      (will produce BaseName.mkl from BaseName.gbw)
orca_2mkl BaseName -molden
      (writes a file in molden format)
orca_2mkl BaseName -mkl
      (writes a file in MKL format)
```

We have recently also added the capability to convert any gbw type file into MKL or Molden format. Thus, you can use this device to vizualize QRO or UNO or UCO orbitals or any type of natural orbitals:

```
orca_2mkl gbw_type_file.extension mkl_file.extension -mkl -anyorbs
# or
orca_2mkl gbw_type_file.extension molden_file.extension -molden -anyorbs
```

You also have the opportunity to run `orca_2mkl` backwards in order to produce gbw type files. You can use this device in order to import orbitals from other sources into ORCA. This is not a frequently used option and it has limited capabilities. Hence, it is documented here only in a cursory way in order for you to be able to experiment. Note that the CASSCF tutorial, that supplements the manual, shows how to edit the molecular orbitals using `orca_2mkl`.

```
orca_2mkl BaseName -gbw
      (will produce BaseName.gbw out of BaseName.mkl)
```

7.55.9 orca_2aim

This utility program reads a `.gbw` file and creates a `.wfn` and `.wfx` file that can be used for topological analysis of the electron density by other programs. This works for open-shell and closed-shell wave functions. The usage is very simple – just type `AIM` in the simple input line of your input file, or use

```
orca_2aim BaseName
      (will produce BaseName.wfn and BaseName.wfx from BaseName.gbw)
```

7.55.10 orca_vpot

This program calculates the electrostatic potential at a given set of user defined points. It can be used with either an input file or in interactive mode. It needs four arguments:

```
orca_vpot  MyJob.gbw  MyJob.scfp  MyJob.vpot.xyz  MyJob.vpot.out
```

First: The `gbw` file containing the correct geometry and basis set

Second: The desired density matrix in this basis (perhaps use the `KeepDens` keyword)

Third: an ASCII file with the target positions in AU, e.g.

```
6                 (number_of_points)
  5.0 0.0  0.0    (XYZ coordinates)
 -5.0 0.0  0.0
  0.0 5.0  0.0
  0.0-5.0  0.0
  0.0 0.0  5.0
  0.0 0.0 -5.0
```

Fourth: The target file which will then contain the electrostatic potential, e.g.

```
6                 (number of points)
 VX1  VY1  VZ1    (potential for first point)
 VX2  VY2  VZ2    (potential for second point)
  etc.
```

It should be straightforward for you to read this file and use the potential for whatever purpose.

There are some special ways to call orca_vpot: A call for use in parallel mode

```
mpirun -np 4 /full_path/orca_vpot_mpi MyJob.gbw MyJob.scfp MyJob.pot.xyz MyJob.pot.out
```

A call to check, which densities are available in MyJob.densities,

```
orca_vpot  MyJob.densities
```

will give you a listing of all densities contained in the file.

In the case that basename of gbw- and densities-file do not match, you have to pass the densities' name as fifth argument to orca_vpot.

```
orca_vpot  MyJob.gbw  OtherJob.scfp  MyJob.pot.xyz  MyJob.pot.out  OtherJob
```

A call to orca_vpot without any arguments will display a help message.

7.55.11 orca_euler

This utility program is used to calculate the relative orientation between calculated hyperfine coupling (HFC)/nuclear quadrupole coupling (NQC) tensors and a reference tensor (the calculated molecular g-/D-tensor). The `orca_euler` program is run by default in an ORCA job after the calculation of HFCs or NQCs, if g- or D-tensor are also calculated in the same job. The utility program can also be run as a stand-alone program. In this case the `.property.txt` file of a previous NQC/HFC- and D- or g-tensor calculation must be available.

The orientation between the tensors is calculated in terms of a 3x3 rotational matrix R. This is parametrized by the three so-called Euler angles α, β and γ. These angles define the relative orientation between two tensors A and B by three successively applied rotations around different axes in order to align A with B. In the commonly used z-y-z convention these three rotations are:

- **Rotate $\mathbf{A}_{xyz}$ counterclockwise around its z axis by α to give $\mathbf{A}_{x'y'z'}$.**
- **Rotate $\mathbf{A}_{x'y'z'}$ counterclockwise around its y' axis by β to give $\mathbf{A}_{x''y''z''}$.**
- **Rotate $\mathbf{A}_{x''y''z''}$ counterclockwise around its z'' axis by γ to align with B.**

```
orca_euler prop-file options

  file    = basename of an ORCA .property.txt file

  options
   -refg/-refD:         Reference tensor (g-tensor or D-tensor, default is -refg)
   -conv zyz/-conv zxz: Euler rotation convention (default is zyz)
   -order:              Ordering of the reference tensor (x, y, z) with respect to
                        ORCA output (min, mid, max)
   -plotA:              plot the HFC-tensors
   -plotQ:              plot the NQC-tensors
   -detail:             print detailed information
```

NOTE:

- By default the D-tensor is used as reference tensor only if $S > \frac{1}{2}$ and if D>0.3 cm^{-1}; in all other cases the g-tensor is used as reference tensor. The user can manually select the reference tensor – if the information is available in the prop-file – by using –`refg` or -`refD`.
- By default the Euler rotation in the z-y-z convention is used. The z-x-z convention can be selected manually by using the option –`conv zxz`.
- By default the axes of the g- or D-tensor are assigned depending on their magnitude. $g_{min} \rightarrow g_x$, $g_{mid} \rightarrow g_y$, $g_{max} \rightarrow g_z$ (similarly for D). This ordering can be modified manually when running the standalone program as shown in the following examples:

-order 3 2 1:	min → z
	mid → y
	max → x
-order 1 -2 3:	min → x
	mid → y (flipped in the orientation)
	max → z

- The nuclear hyperfine and quadrupole coupling tensors can be plotted (in the xyz-file format) by the `orca_euler` program using –`plotA` or –`plotQ`. The HFC tensor for atom 3 (counting starts at zero) is e.g. stored in the file `prop-file.3.A.xyz`, the respective NQC tensor is stored in `prop-file.3.Q.xyz`. In these xyz files the position of four atoms (He, Ne, Ar, Kr) is given. The x-, y- and z-direction of the tensor are in the direction of the vectors between He-Ne, He-Ar and He-Kr.
- The actual definition of the used rotation matrix and more information on the relative orientation can be printed by using the option –`detail`.

7.55.12 orca_exportbasis

A small utility program to print out the basis sets used by ORCA. Its usage requires at least the name of the basis set, as specified in the simple input line of ORCA. Additional parameters like an ECP basis set, a list of specific atoms or the name of an ouput file are accepted. The output is stored in ASCII format, it can be inspected and modified. The user can choose to print the basis sets in either `ORCA` format, which then can be copied into the input file, or in `GAMESS-US` format, which can be read via the %basis block as externally specified basis. **NOTE:** Basis set names containing special characters may need a pair of enclosing “ or ‘ to be recognized.

```
USAGE: orca_exportbasis keywords options

       -b, --basis  : name of basis set
                      def2-svp
                      'def2-tzvp(-f)' - string to be passed with literals

EXAMPLE: orca_exportbasis -b svp
```

```
Additional Options:

       -e, --ecp    : ecp basis
                      sdd
                      default        - ECP-part of basis (if present)

       -f, --format : output format
                      ORCA           - to be read via %basis NewGTO
                      GAMESS-US      - to be read as %basis GTOName 'mybasis.bas'
                      default        - ORCA
       -a, --atoms  : list of elements
                      Cu             - single element
                      Ga Ge As Se    - list of elements separated by blanks
                      default        - whole periodic table is printed

       -o, --outfile: name of outputfile
                      mybasis.bas
                      default        - derived name
```

(continues on next page)

(continued from previous page)

```
EXAMPLE: orca_exportbasis -b svp -e sdd -a Ag -f GAMESS-US -o mybasis.bas
```

The output stored in GAMESS-US format can be used in the %basis block of the next ORCA calculation.

```
%basis
GTOName      "mybasis.bas"
GTOAuxJName  "myauxjbasis.bas"
GTOAuxJKName "myauxjkbasis.bas"
GTOAuxCName  "myauxjcbasis.bas"
end
```

7.55.13 orca_eca

This utility program makes use of the calculated exchange coupling constants to compute relative energies of all possible spin states through diagonalization of spin Hamiltonian. The absolute and relative energies of the spin states are printed in the `*.en` and `*.en0` files respectively. The on-site spin expectation values are also printed in a `*.sp` file.The following example calculates the spin ladder for a system with exchange coupling constant of -152.48 cm**-1 between Mn(III) and Mn(IV).

```
%sim
ms_bs 0.5 # Arbitrary spin state
end

# specification of spin centers
$spins 2
1 2.0  # Spin on first manganese
2 1.5  # Spin on second manganese

# Exchange coupling constant (H = -2J S1 S2)
$ecc 1
1 -152.48

$aiso_bs 2 # A false segment just to print the *.sp file
1 0.00
2 0.00
```

7.55.14 orca_pnmr

`orca_pnmr` calculates the paramagnetic contribution to the NMR shielding tensor from EPR g, A, and D tensors (see Section *Paramagnetic NMR shielding tensors* for theoretical background). It is a standalone program which you can invoke on the command line after the main ORCA calculation has finished. Alternatively, it can read user-provided $g/A/D$ tensors from an input file (option `-i`). Note that `orca_pnmr` expects g and A tensors that conform to the convention described in Section *Cartesian Index Conventions for EPR and NMR Tensors*.

```
USAGE:    orca_pnmr BaseName [-i] [-v]

OPTIONS:  -i  : read from the input file "BaseName.pnmr.inp"
          -v  : print more output (Z matrices)
```

When called without options, `orca_pnmr` will attempt to extract g, D, and A tensors from the property file `BaseName.property.txt` and use these to calculate the paramagnetic shieldings at 298 K. Note that this functionality is not sophisticated: it only recognizes EPR tensors calculated by the EPRNMR module. A more flexible way to use the capabilities of `orca_pnmr` is to manually edit the textfile `BaseName.pnmr.inp` and then run `orca_pnmr` with the `-i` option. The file has the following format:

```
3    # Spin multiplicity (2S+1)
298  # Temperature range minimum [K]
300  # Temperature range maximum [K]
1    # Temperature step [K]
1    # Have g-tensor? 0 or 1
  2.004689   0.000000  -0.000000 # Cartesian g-tensor
  0.000000   2.004689   0.000000 # (if available)
  0.000000   0.000000   2.002123
1    # Have D-tensor? 0 or 1
 -0.514341  -0.000000  -0.000000 # Cartesian D-tensor [cm-1]
 -0.000000  -0.514341  -0.000000 # (if available)
 -0.000000  -0.000000   1.028682
2    # Number of A-tensors
00   # First nucleus (index,element)
-72.358765                      # Prefactor [MHz]
-78.989514   0.000000  -0.000000 # Cartesian A-tensor [MHz]
  0.000000 -78.989514  -0.000000
 -0.000000  -0.000000  53.478581
10   # Next nucleus (index,element)
-72.358765                      # Prefactor [MHz]
-78.989516   0.000000  -0.000000 # Cartesian A-tensor [MHz]
  0.000000 -78.989516  -0.000000
 -0.000000  -0.000000  53.478580
# ... Further nuclei
```

- D and g tensors are optional: if they are not supplied `orca_pnmr` will assume the isotropic free-electron value for g, and D to be zero.
- A tensors, however, are not optional; without an A tensor for a given nucleus, the pNMR shielding cannot be calculated for that nucleus.

7.55.15 orca_lft

Starting from ORCA 5.0, ORCA features a standalone multiplet program called orca_lft.

- Orca_lft is dedicated to experimental spectroscopists.
- It is able to run an arbitrary number of spectra simulations with emphasis on X-ray spectroscopies.

In this section we briefly review the main functionalities of orca_lft. For a more detail description and examples discussion please refer to the orca_lft tutorial.

1. The goal is to be able to compute various spectroscopic properties of a given LFT center (ion) if one can manually pass the information of 1 and 2 electron integrals in the form of e.g the diagonal elements of the LFT matrix (LFT orbital energies) and the Slater-Condon parameters of a given LFT problem.
2. This will allow the experimental spectroscopist to perform a massive amount of spectra simulations during the actual running experiments

Any LFT problem can be parametrized in terms 1-electron H_{LF} matrix elements and the Slater-Condon 2-electron integrals F0, F2, F4, (or the Racah parameters A, B, C, of the d-shell). (Figure: Fig. 7.70)

$$\hat{H}_{LF} = \overbrace{\sum_i \hat{v}_{LF}(i)}^{1-electron\,integrals} + \overbrace{\sum_{i<j} \hat{G}_{LF}(i,j)}^{2-electron\,integrals}$$

Fig. 7.70: Definition of an LFT problem in terms of 1-electron energies and Slater Condon parameters (SCPs)

In practice we need to know:

1. the Slater Condon parameters of a given LFT problem
2. the H_{LF} matrix elements or the relation of them (ligand field splitting, 10Dq, AOM model)

The design workflow of orca_lft is the following:

- Solve the General CI problem on a User-specified LFT problem: Type of Ion, Number of Electrons, Involved Shells, Involved Multiplicities
- Compute All possible Non Relativistic States/Multiplicity
- Compute the Transition Densities on the CSFs basis
- Compute the Needed transition Moments in the given LFT basis (i.e. 2p3d)
- Compute various properties with emphasis to X-ray spectroscopy (ABS, XAS, XES, RIXS) at the Non Relativistic Limit
- Compute the respective Relativistically corrected States on the Quasi Degenerate Perturbation Theory basis
- Provide access to various relativistically corrected properies (ABS, XAS, XES, RIXS, MCD, XMCD, GTensors, Dtensors, Hyperfine Couplings, Electric Field Gradients)

orca_lft requires its own input. By simply executing it from the terminal

```
orca_lft
```

one gets printings of the Usage:

```
****************************************************************************************************
Simulate or Fit Spectra
****************************************************************************************************

================================================================================
Usage: orca_lft BaseName.lft.inp [options]
================================================================================

--------------------------------------------------------------------------------
↪-------------
[Options]:
--------------------------------------------------------------------------------
↪-------------
-sim                          Simulate Spectra
-fit                          Fit Spectra (This is not yet availiable!)
--------------------------------------------------------------------------------
↪-------------

****************************************************************************************************
Generate Initial Input
****************************************************************************************************

================================================================================
Usage: orca_lft BaseName [options]
================================================================================
```

various Run Options:

```
--------------------------------------------------------------------------------
↪-------------
[Options]:
--------------------------------------------------------------------------------
↪-------------
-p_case                       Requests  p-shell case
-d_case                       Requests  d-shell case
-f_case                       Requests  f-shell case
-sp_case                      Requests sp-shell case
```

(continues on next page)

(continued from previous page)

```
-ps_case                      Requests sp-shell case
-sd_case                      Requests sd-shell case
-ds_case                      Requests ds-shell case
-sf_case                      Requests sf-shell case
-fs_case                      Requests fs-shell case
-pd_case                      Requests pd-shell case
-dp_case                      Requests dp-shell case
-pf_case                      Requests pf-shell case
-fp_case                      Requests fp-shell case
-df_case                      Requests pf-shell case
-fd_case                      Requests fp-shell case
-spd_case                     Requests spd-shell case
-spf_case                     Requests spf-shell case
-sdf_case                     Requests sdf-shell case
-pdf_case                     Requests pdf-shell case
-soc                          Requests Input with SOC Constants
-atnoN                        Sets Atomic Number N
-------------------------------------------------------------------------------------------
-Special Cases for known Elements. LFT Parameters are filled from an internal AILFT NEVPT2␣
↪database

-------------------------------------Supported Oxidation States-----------------------------
Presently Default Oxidation States are supported except for Fe:
Main Elements                 -> 0
TM Elements (Fe(26))          -> I,II,III
TM Elements (All other)       -> II
Lanthanide/Actinide Elements  -> III
-----------------------------------------Valence Cases--------------------------------------
-atnoN  -2s2p_case            Requests 2s2p-shell case for element with atomic number N (4-9)
-atnoN  -3s3p_case            Requests 3s3p-shell case for element with atomic number N (12-18)
-atnoN  -4s4p_case            Requests 4s4p-shell case for element with atomic number N (20,31-
↪36)
-atnoN  -4s4p_case            Requests 5s5p-shell case for element with atomic number N (38,49-
↪54)
-atnoN  -3d4s_case            Requests 3d4s-shell case for element with atomic number N (22-29)
-atnoN  -4d5s_case            Requests 4d5s-shell case for element with atomic number N (40-47)
-atnoN  -5d6s_case            Requests 4d6s-shell case for element with atomic number N (72-79)
-atnoN  -4f5d_case            Requests 4d5d-shell case for element with atomic number N (59-70)
-atnoN  -5f6d_case            Requests 5d6d-shell case for element with atomic number N (91-
↪102)
---------------------------------------Core-Valence Cases-----------------------------------
-atnoN  -1s3d_case            Requests 1s3d-shell case for element with atomic number N (22-29)
-atnoN  -2p3d_case            Requests 2p3d-shell case for element with atomic number N (22-29)
-atnoN  -3p3d_case            Requests 3p3d-shell case for element with atomic number N (22-29)
-------------------------------------Core-Valence XES Cases-------------------------------
-atnoN  -1s2p3d_case          Requests 1s2p3d-shell case for element with atomic number N (22-
↪29)
-atnoN  -1s3p3d_case          Requests 1s3p3d-shell case for element with atomic number N (22-
↪29)
-------------------------------------------------------------------------------------------
```

and various Spectra Simulation Options:

```
*******************************************************************************************************
Spectra Simulation Options for the BaseName.lft.inp:
```

(continues on next page)

(continued from previous page)

```
*****************************************************************************************************************

-----------------------------------------------------------------------------------------------
↪-------------
TIP: Switch ON a Property calculation as DoProperty true:
-----------------------------------------------------------------------------------------------
↪-------------

-----------------------------------------------------------------------------------------------
↪-------------
General Input Parameters:
-----------------------------------------------------------------------------------------------
↪-------------
NEl                           Sets the number of electrons
Shell_PQN                     Sets the principle quantum number per type of shells (s,p,d,f)
LFTCase                       !!!ALTERNATIVE TO Shell_PQN!!! Sets the given LFT problem (2p3d,␣
↪1s3p3d, ...)
-----------------------------------------------------------------------------------------------
↪------------
LFTCase WILL replace Shell_PQN
-----------------------------------------------------------------------------------------------
↪------------
(e.g. Shell_PQN = 0,2,3,0 for a 2p3d calculation)
Mult                          Sets the Multiplicity/Multiplicities
NRoots                        Sets the number of Roots/Multiplicity
TMultiplets (0.01)            Threshold for the Multiplets grouping in eV
DoeV                          All values in eV. This is default. If set false the cm-1 unit is␣
↪used throughout
DoRAS                         Requests a RASCI calculation
RAS(nel: m1 h/ m2 / m3 p)     Computes the X-Ray Emission Spectra

RAS-reference with nel electrons
m1= number orbitals in RAS-1
h = max. number of holes in RAS-1
m2= number of orbitals in RAS-2 (any number of electrons or holes)
m3= number of orbitals in RAS-3
p = max. number of particles in RAS-3

DoElastic                     Computes in addition the Elastic Scattering terms in RIXS/
↪RIXSSOC calculations
-----------------------------------------------------------------------------------------------
↪-------------
Non Relativistic Spectroscopic Properties:
-----------------------------------------------------------------------------------------------
↪-------------
DoABS/DoXAS                   Computes the Absorption like Spectra
DoCD                          Computes the CD Spectra
DoXES                         Computes the X-Ray Emission Spectra
DoRIXS                        Computes the RIXS Spectra
DoQuadrupole                  Computes the ABS,XAS,RIXS Spectra beyond dipole approximation
-----------------------------------------------------------------------------------------------
↪-------------
Relativistically Corrected Spectroscopic Properties:
-----------------------------------------------------------------------------------------------
↪-------------
DoSOC                         Requests the Spin Orbit Coupling Calculations
Note that this turned on automatically if zeta SOC
constant are provided
-----------------------------------------------------------------------------------------------
↪-------------
DoABS/DoXAS                   Computes the SOC Corrected Absorption like Spectra
```

(continues on next page)

(continued from previous page)

```
DoCD                          Computes the SOC Corrected CD Spectra
DoMCD/DoXMCD                  Computes the SOC Corrected MCD/XMCD Spectra
DoXESSOC:                     Computes the SOC Corrected XES Spectra
DoRIXSSOC:                    Computes the SOC Corrected RIXS Spectra
DoQuadrupole                  Computes the SOC Corrected ABS,XAS, XES RIXS Spectra beyond␣
↪dipole approximation
-------------------------------------------------------------------------------------------
↪-------------
Magnetic Properties:
-------------------------------------------------------------------------------------------
↪-------------
DoMagnetization               Computes the Magnetization
DoSusceptibility              Computes the Susceptibilities
DoGTensor                     Computes the g-Tensors/Matrices
DoDTensor                     Computes the Zero-Field Splittings
DoATensor                     Computes the Hyperfine Tensors
DoEFGTensor                   Computes the Electric Field Gradient Tensors
(also Moesbauer Parameters in the presence of Fe centers)
-------------------------------------------------------------------------------------------
↪-------------
Variable Parameters (if they are not given, default values are used)
-------------------------------------------------------------------------------------------
↪-------------
Temperature(300)              Temperature to be used in the SOC calculations
MagneticField(0)              Magnetic Field (in Gauss)
NPointsPsi(10)                Solid Angle Integration points Psi for MCD and XMCD
NPointsPhi(10)                Solid Angle Integration points Phi for MCD and XMCD
NPointsTheta(10)              Solid Angle Integration points Theta for MCD and XMCD
-------------------------------------------------------------------------------------------
↪-------------
Variable Parameters needed for Magnetization and Susceptibility calculations
-------------------------------------------------------------------------------------------
↪-------------
LebedevIntegrationPoints(26)  Number for Integration points for Lebedev)
LebedevPrec(5)                Precision of the grid for different field directions
(meaningful values range from 1 (smallest) to 10 (largest)
nPointsFStep (5)              Number of steps for numerical differentiation
(def: 5, meaningful values are 3, 5 7 and 9)
MAGTemperatureMIN(4.0)        Minimum Temperature (K) for Magnetization
MAGTemperatureMAX(4.0)        Maximum Temperature (K) for Magnetization
MAGTemperatureNPoints(1)      Number of Temperature points for Magnetization
MAGFieldStep(100.0)           Size of Field step for numerical differentiation (def: 100 Gauss)
MAGFieldMin(0.0)              Minimum Field (Gauss) for Magnetization
MAGFieldMax(70000.0)          Minimum Field (Gauss) for Magnetization
MAGNPoints(15)                Number of Field points for Magnetization
SUSTempMin(1)                 Minimum Temperature (K) for Susceptibility
SUSTempMxn(300.0)             Maximum Temperature (K) for Susceptibility
SUSNPoints(300)               Number of Temperature points for Susceptibility
SUSStatFieldMIN(0.0)          Minimum Static Field (Gauss) for Susceptibility
SUSStatFieldMAX(1)            Maximum Static field (Gauss) for Susceptibility
SUSStatFieldNPoints(1)        Number of Static Fields for Susceptibility
-------------------------------------------------------------------------------------------
↪-------------
```

Orca_lft can run standalone by processing an input file (Basename.lft.inp) with orca_lft

```
orca_lft BaseName.lft.inp -sim
```

Alternatively, one can call the main orca program like

```
orca BaseName.lft.inp
```

The different benefits of the two runs are provided in the orca_lft tutorial

Orca_lft can also be used to automatically generate initial proper input files. For example, by running

```
orca_lft BaseName -pd_case
```

will generate an initial basename.lft_pd.inp input.

```
-------------------------------------------------------------------
Initial Input: BaseName.lft_pd.inp for Orca_LFT has been generated:
-------------------------------------------------------------------
```

This will generate a bare input. One has to fill in the required LFT parameters and the desired spectroscopic properies and start the simulations like in every common multiplet program.

```
%lft


#-----Parameters------
NEl= 0
Shell_PQN= 0, 2, 3, 0
Mult= 0
NRoots= -1
#--------------------


#---Slater-Condon Parameters---
#---All Values in eV---
PARAMETERS
F0pp =  0.00
F2pp =  0.00
F0dd =  0.00
F2dd =  0.00
F4dd =  0.00
F0pd =  0.00
F2pd =  0.00
G1pd =  0.00
G3pd =  0.00
end
#--------------------


#---Diagonal LFT-Matrix Elelemnts---
#---All Values in eV---
FUNCTIONS
0   0 " 0.00"
1   1 " 0.00"
2   2 " 0.00"
3   3 " 0.00"
4   4 " 0.00"
5   5 " 0.00"
6   6 " 0.00"
7   7 " 0.00"
end
#--------------------


#---SPECTRA/PROPERTIES---
DoABS true
#---------
end
```

(continues on next page)

(continued from previous page)

```
*xyz 0 0
Atom 0.00 0.00 0.00
*
```

Special initial inputs based on an internal NEVPT2 data base can also be generated. For the 2p3d LFT case of NiII this will look like this

```
orca_lft BaseName -atno28 -2p3d_case
```

```
-----------------------------------------------------------------------------
Creating input for Atom Ni(II) ...
-----------------------------------------------------------------------------


---------------------------------------------------------------
Initial Input: BaseName.lft_pd.inp for Orca_LFT has been generated:
---------------------------------------------------------------
```

This will generate the following imput where the LFT parameters are filled in from an internal NEVPT2 database from precomputed CASCI/NEVPT2 AILFT LFT parameters

In the case of Ni^{II} this will look like this. This is basically a ready to run input.

```
%lft


#-----Parameters------
NEl= 14
LFTCase 2p3d
Mult= 3, 1
NRoots= 25, 30
#--------------------


#---Slater-Condon Parameters---
#---All Values in eV---
PARAMETERS
F0pp =  85.88
F2pp =  54.77
F0dd =  23.31
F2dd =  13.89
F4dd =  9.14
F0pd =  33.03
F2pd =  7.76
G1pd =  6.42
G3pd =  2.11
end
#--------------------


#---Diagonal LFT-Matrix Elelemnts---
#---All Values in eV---
FUNCTIONS
0   0 " 0.00"
1   1 " 0.00"
2   2 " 0.00"
3   3 "1138.35"
4   4 "1138.35"
5   5 "1138.35"
6   6 "1138.35"
7   7 "1138.35"
end
```

(continues on next page)

(continued from previous page)

```
#--------------------


#---SPECTRA/PROPERTIES---
DoABS true
#---------
end


*xyz 2 3
Ni 0.00 0.00 0.00
*
```

Alternatively as discussed in the Abinitio Ligand Field Theory section (*1- and 2-shell Abinitio Ligand Field Theory*) one may actually run a 2-shell AILFT calculation and produce the respective *nevpt2.lft.inp file

```
!NoIter NEVPT2 def2-SVP def2-SVP/C


%method
frozencore fc_none
end

#--------------------
#Rotate Orbitals
#--------------------
%scf
rotate
{2,6,90}
{3,7,90}
{4,8,90}
end
end

#--------------------
#General Options
#--------------------
%casscf
nel 14
norb 8
mult 3,1
nroots 100,100

LFTCase 2p3d
rel
dosoc true
end
end


*xyz 2 3
Ni      0.0000000000      0.0000000000      0.0000000000
*
```

The structure of an orca_lft input is the following:

It contains:

- The General Parameters Block where the LFT problem is defined

  ——Parameters—— NEl= 14 LFTCase 2p3d #Shells_PQN 0,2,3,0, Alternative definition using s,p,d,f main quantum numbers Mult= 3, 1 NRoots= 25, 30 ————————

- The PARAMETERS Block where the SCPs and SOC constant parameters are defined

  —Slater-Condon Parameters— —All Values in eV— PARAMETERS F0pp = 85.88 F2pp = 54.77 F0dd = 23.31 F2dd = 13.89 F4dd = 9.14 F0pd = 33.03 F2pd = 7.76 G1pd = 6.42 G3pd = 2.11 end ————————

  ...

  —SOC-CONSTANTS— —All Values in eV— PARAMETERS ZETA_P = 10.68 ZETA_D = 0.08 end ————————

- The FUNCTIONS Block where the LFT matrix is defined
- The Properties Block where the desire simulation properties are specified

  —SPECTRA/PROPERTIES— DoABS true ———

- The xyz Block where the ion, charge, multiplicity and coordinates (0. 0. 0.) are defined

  *xyz 2 3 Ni 0.0000000000 0.0000000000 0.0000000000 *

It should be emphasized the orca_lft via the FUNCTIONS and PARAMETERS blocks provides

- an arbitary paramterization of the 1-electron LFT matrix and the Slater Condon Parameters
- a powerfull parameters scanability

which helps to performs any kind of simulation without any symmetry restrictions. Details regarding the use of the FUNCTIONS and PARAMETERS blocks with examples are provided in the orca_lft tutorial

Let us perform the Ni^{2+} L-edge XAS spectrum simulation using the following input

```
%lft


  #-----Parameters------
  NEl= 14
  LFTCase 2p3d
  Mult= 3, 1
  NRoots= 25, 30
  #--------------------


  #---Slater-Condon Parameters---
  #---All Values in eV---
  PARAMETERS
    F0pp =  85.88
    F2pp =  54.77
    F0dd =  23.31
    F2dd =  13.89
    F4dd =  9.14
    F0pd =  33.03
    F2pd =  7.76
    G1pd =  6.42
    G3pd =  2.11
  end
  #--------------------


  #---Diagonal LFT-Matrix Elelemnts---
  #---All Values in eV---
  FUNCTIONS
    0   0 " 0.00"
    1   1 " 0.00"
    2   2 " 0.00"
    3   3 "1138.35"
    4   4 "1138.35"
    5   5 "1138.35"
```

(continues on next page)

(continued from previous page)

```
    6   6 "1138.35"
    7   7 "1138.35"
  end
  #--------------------

  #---SOC-CONSTANTS---
  #---All Values in eV---
  PARAMETERS
    ZETA_P = 11.341
    ZETA_D = 0.085
  end
  #--------------------

  #---SPECTRA/PROPERTIES---
  DoABS true
  DoSOC true
  #---------
end


*xyz 2 3
Ni 0.00 0.00 0.00
*
```

In a first step the definition of the LFT problem is performed:

```
------------------------------------------------------------------------------
L I G A N D    F I E L D    T H E O R Y
------------------------------------------------------------------------------


Number of electrons        = 14
Multiplicities             = 3 1
Roots                      = 25 30
Shells included            = 0 2 3 0


Definition of the ligand field basis set:
0  = pz
1  = px
2  = py
3  = dz2
4  = dxz
5  = dyz
6  = dx2y2
7  = dxy


Definition of the static ligand field by the user:
There are 9 ligand field parameters

Nr.               Name       Initial Value
-----------------------------------------
1                F0PP          85.880000
2                F2PP          54.770000
3                F0DD          23.310000
4                F2DD          13.890000
5                F4DD           9.140000
6                F0PD          33.030000
7                F2PD           7.760000
8                G1PD           6.420000
```

(continues on next page)

(continued from previous page)

```
9               G3PD           2.110000


Definition of the ligand field functions by the user:
There are 8 ligand field functions

Nr.      H-element   value   function
----------------------------------------
1   H(0,0)      0.000000000   0.00
2   H(1,1)      0.000000000   0.00
3   H(2,2)      0.000000000   0.00
4   H(3,3)   1138.350000000  1138.35
5   H(4,4)   1138.350000000  1138.35
6   H(5,5)   1138.350000000  1138.35
7   H(6,6)   1138.350000000  1138.35
8   H(7,7)   1138.350000000  1138.35


Defining the one-electron LFT matrix                    ... done
Defining the two-electron LFT matrix                    ... done
```

In following the CI problem is defined:

```
Defining the CI spaces and setting up the CI             ...
Making Checks...
Multiplicty =   3, #(configurations) =     28 #(CSF's) =     28 #(Roots) =     25
NRoots<NCSFs Adjussting  ==> (CSF's) =     25
Setting up CI...
Multiplicty =   3, #(configurations) =     28 #(CSF's) =     25 #(Roots) =     25
Making Checks...
Multiplicty =   1, #(configurations) =     36 #(CSF's) =     36 #(Roots) =     30
NRoots<NCSFs Adjussting  ==> (CSF's) =     30
Setting up CI...
Multiplicty =   1, #(configurations) =     36 #(CSF's) =     30 #(Roots) =     30
CI setup done
```

And the CI problem is solved:

```
LOWEST ROOT (ROOT 0, MULT 3) =    460.113216547 Eh 12520.317 eV

STATE ROOT MULT  DE/a.u.     DE/eV    DE/cm**-1
1:   1    3   0.000000     0.000      0.0
2:   2    3   0.000000     0.000      0.0
3:   3    3   0.000000     0.000      0.0
4:   4    3   0.000000     0.000      0.0
5:   5    3   0.000000     0.000      0.0
6:   6    3   0.000000     0.000      0.0
7:   0    1   0.086266     2.347  18933.2
8:   1    1   0.086266     2.347  18933.2
9:   2    1   0.086266     2.347  18933.2
10:   3    1   0.086266     2.347  18933.2
11:   4    1   0.086266     2.347  18933.2
12:   7    3   0.099001     2.694  21728.2
13:   8    3   0.099001     2.694  21728.2
14:   9    3   0.099001     2.694  21728.2
15:   5    1   0.132624     3.609  29107.7
16:   6    1   0.132624     3.609  29107.7
17:   7    1   0.132624     3.609  29107.7
18:   8    1   0.132624     3.609  29107.7
19:   9    1   0.132624     3.609  29107.7
20:  10    1   0.132624     3.609  29107.7
```

(continues on next page)

(continued from previous page)

```
21:  11    1   0.132624     3.609  29107.7
22:  12    1   0.132624     3.609  29107.7
23:  13    1   0.132624     3.609  29107.7
24:  14    1   0.331652     9.025  72789.3
25:  15    1  29.897078   813.541 6561650.2
26:  16    1  29.897078   813.541 6561650.2
27:  17    1  29.897078   813.541 6561650.2
28:  18    1  29.897078   813.541 6561650.2
29:  19    1  29.897078   813.541 6561650.2
30:  10    3  29.915627   814.046 6565721.2
31:  11    3  29.915627   814.046 6565721.2
32:  12    3  29.915627   814.046 6565721.2
33:  13    3  29.915627   814.046 6565721.2
34:  14    3  29.915627   814.046 6565721.2
35:  15    3  29.915627   814.046 6565721.2
36:  16    3  29.915627   814.046 6565721.2
37:  17    3  29.978158   815.747 6579445.1
38:  18    3  29.978158   815.747 6579445.1
39:  19    3  29.978158   815.747 6579445.1
40:  20    3  29.978158   815.747 6579445.1
41:  21    3  29.978158   815.747 6579445.1
42:  22    3  30.016020   816.777 6587754.9
43:  23    3  30.016020   816.777 6587754.9
44:  24    3  30.016020   816.777 6587754.9
45:  20    1  30.087356   818.719 6603411.3
46:  21    1  30.087356   818.719 6603411.3
47:  22    1  30.087356   818.719 6603411.3
48:  23    1  30.106270   819.233 6607562.6
49:  24    1  30.106270   819.233 6607562.6
50:  25    1  30.106270   819.233 6607562.6
51:  26    1  30.106270   819.233 6607562.6
52:  27    1  30.106270   819.233 6607562.6
53:  28    1  30.106270   819.233 6607562.6
54:  29    1  30.106270   819.233 6607562.6
```

accompanied by a multiplet analysis:

```
-------------------------------------------------
Atomic calculation : Multiplet analysis (LFT)
-------------------------------------------------
11 multiplets found (Threshold = 0.01 eV)
0 3F | E =     0.000 eV | 2p(6)3d(8)
1 3P | E =     2.694 eV | 2p(6)3d(8)
2 3F | E =   814.046 eV | 2p(5)3d(9)
3 3D | E =   815.747 eV | 2p(5)3d(9)
4 3P | E =   816.777 eV | 2p(5)3d(9)
5 1D | E =     2.347 eV | 2p(6)3d(8)
6 1G | E =     3.609 eV | 2p(6)3d(8)
7 1S | E =     9.025 eV | 2p(6)3d(8)
8 1D | E =   813.541 eV | 2p(5)3d(9)
9 1P | E =   818.719 eV | 2p(5)3d(9)
10 1F | E =   819.233 eV | 2p(5)3d(9)
-------------------------------------------------
```

In following SOC is computed on the QDPT framework

```
************************************
Doing QDPT with ONLY SOC!
************************************


----------------------------------
```

(continues on next page)

(continued from previous page)

```
NONZERO SOC MATRIX ELEMENTS (cm**-1)
------------------------------------

Bra                       Ket
<Block Root  S    Ms  | HSOC |  Block Root  S    Ms>    =  Real-part     Imaginary part
-------------------------------------------------------------------------------------
0    1  1.0  1.0              0     0  1.0  1.0            0.000           -685.571
0    3  1.0  1.0              0     2  1.0  1.0            0.000            342.781
0    4  1.0  1.0              0     2  1.0  1.0            0.000              0.134
0    4  1.0  1.0              0     3  1.0  1.0            0.000             -2.946
0    5  1.0  1.0              0     2  1.0  1.0            0.000              0.987
0    5  1.0  1.0              0     3  1.0  1.0            0.000              7.920
0    5  1.0  1.0              0     4  1.0  1.0            0.000          -1022.995
0    6  1.0  1.0              0     2  1.0  1.0            0.000             -3.379

...

  1   29  0.0  0.0              0    17  1.0 -1.0        -2850.628           3039.103
1   29  0.0  0.0              0    18  1.0 -1.0        -16745.172           2885.219
1   29  0.0  0.0              0    19  1.0 -1.0        -1533.311           4672.283
1   29  0.0  0.0              0    20  1.0 -1.0          105.943            -43.140
1   29  0.0  0.0              0    21  1.0 -1.0           14.076             39.149

Note: In the following the full <I|HBO+SOC|J> are printed in the CI Basis.
I,J are compound indices for |Block/Mult, Ms, Root>, where the states
are ordered first by MultBlock, then Ms and finally Root.

...
```

The corrected SOC states are then printed

```
The threshold for printing is 0.0010
Eigenvectors:
Weight      Real          Image    : Block Root    Spin   Ms
STATE   0:       0.0000
0.194318     0.035977    -0.439345 :     0    0    1    1
0.193821     0.438951     0.033797 :     0    1    1    1
0.001621    -0.008938    -0.039257 :     0    4    1    0
0.248122     0.000664    -0.498118 :     0    5    1    0
0.180615     0.034848     0.423557 :     0    0    1   -1
0.180625     0.423745    -0.032633 :     0    1    1   -1

STATE   1:       0.0000
0.180266     0.018153     0.424189 :     0    0    1    1
0.180623    -0.424680     0.016424 :     0    1    1    1
0.245822     0.492098    -0.060506 :     0    4    1    0
0.001648    -0.040383    -0.004177 :     0    5    1    0
0.002114     0.045969    -0.001038 :     0    6    1    0
0.195326     0.087330     0.433242 :     0    0    1   -1
0.192550     0.430084    -0.087049 :     0    1    1   -1

...

STATE 104: 6683223.2237
0.031604     0.157946     0.081591 :     0   17    1    1
0.054142    -0.232659    -0.003404 :     0   18    1    1
0.064958    -0.076463     0.243129 :     0   19    1    1
0.040339     0.047206     0.195219 :     0   24    1    1
0.127180    -0.000013     0.356624 :     0   20    1    0
0.001970    -0.000039    -0.044381 :     0   21    1    0
0.080521    -0.000008     0.283763 :     0   23    1    0
```

(continues on next page)

(continued from previous page)

```
0.031582     0.157887    -0.081573 :     0   17    1   -1
0.054144    -0.232664     0.003379 :     0   18    1   -1
0.064929    -0.076436    -0.243078 :     0   19    1   -1
0.040357     0.047226    -0.195260 :     0   24    1   -1
0.031231    -0.176723     0.000065 :     1   21    0    0
0.376805     0.613845     0.000034 :     1   22    0    0
```

Followed by the SOC corrected Absorption (Here XAS) spectrum

```
-------------------------------------------------------------------------------------------
SOC CORRECTED ABSORPTION SPECTRUM
-------------------------------------------------------------------------------------------
States   Energy   Wavelength   fosc          T2        |TX|       |TY|       |TZ|
(cm-1)     (nm)                   (D**2)      (D)        (D)        (D)
-------------------------------------------------------------------------------------------
0  1       0.0      0.0   0.000000000   0.00006   0.00748   0.00248   0.00014
0  2       0.0      0.0   0.000000000   0.08561   0.15114   0.25053   0.00001
0  3       0.0      0.0   0.000000000   0.08072   0.24632   0.14159   0.00001
0  4       0.0   5379590440159.6   0.000000000   0.00010   0.00622   0.00793   0.00050
0  5       0.0   5308337586647.6   0.000000000   0.00021   0.01379   0.00399   0.00008
0  6       0.0   3910043783335.5   0.000000000   0.02400   0.04644   0.14780   0.00056
0  7       0.0   3891062988270.8   0.000000000   0.03327   0.05783   0.17298   0.00064
0  8       0.0   3626963690424.4   0.000000000   0.04657   0.21455   0.02332   0.00029
0  9    1386.7   7211.2   0.000000723   0.00998   0.01613   0.09858   0.00007
0 10    1386.7   7211.2   0.000000741   0.01022   0.09943   0.01828   0.00001
0 11    1386.7   7211.2   0.000000001   0.00002   0.00200   0.00345   0.00006
0 12    1386.7   7211.2   0.000000018   0.00024   0.00512   0.01469   0.00011
0 13    1386.7   7211.2   0.000000006   0.00009   0.00899   0.00263   0.00002
0 14    1386.7   7211.2   0.000000000   0.00000   0.00044   0.00179   0.00001

...

2 101 6677152.3      1.5   0.000693465   0.00199   0.02052   0.03956   0.00095
2 102 6683223.2      1.5   0.000000000   0.00000   0.00000   0.00000   0.00000
2 103 6683223.2      1.5   0.000000000   0.00000   0.00000   0.00000   0.00000
2 104 6683223.2      1.5   0.000000000   0.00000   0.00000   0.00000   0.00000
```

By processing the *.out file as usual with orca_mapspc *orca_mapspc* the *.dat and *.stk files are generated resulting in Fig. 7.71.

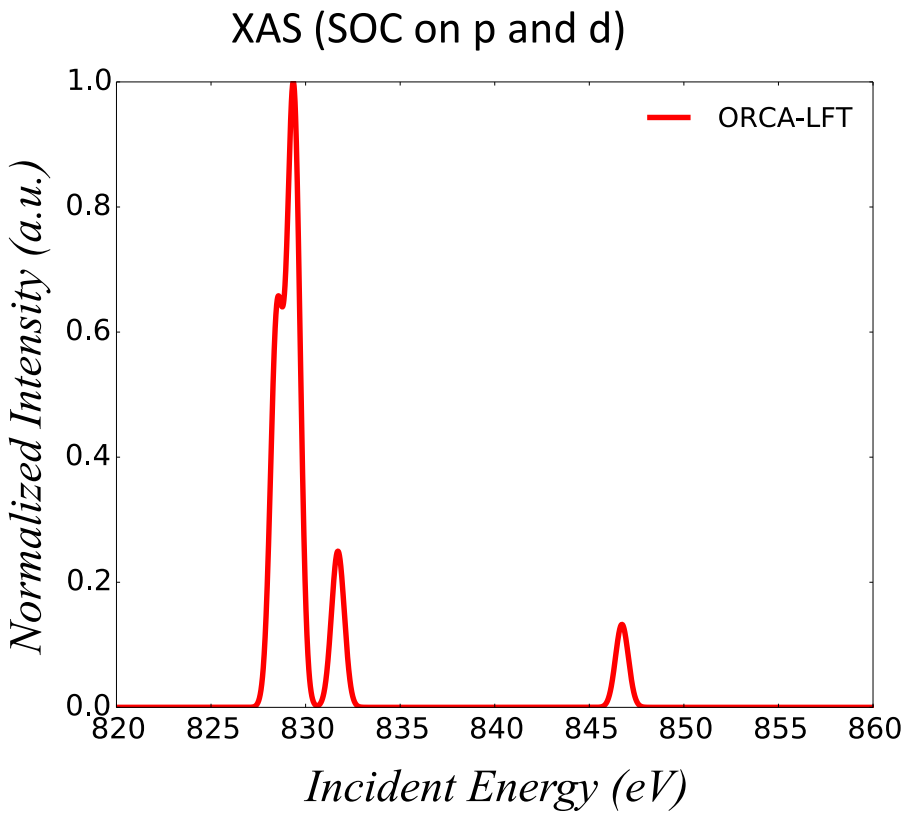


Fig. 7.71: orca_lft simulated Ni^{2+} L-edge XAS spectrum

7.55.16 orca_crystalprep

ORCA 5.0 fetures a utility program that can process crystallographic files (.cif) or .xyz supercell files and produce proper inputs for the embedded cluster calculations It is named orca_crystalprep tool.

Performing an embedded cluster calculation conventionally or within the Ionic-Crystal-QMMM one needs to define basically 3 regions

1. The quantum cluster QC that will be treated quantum mechanically
2. The point charges region PC that represents the solids environment
3. A boundary region BR or ECP that is located between the QC and PC with the main role to prevent charge communication between the QC and PC regions.

This implies that in a first step one needs to generate a SuperCell (.xyz) structure and separate the different regions according to the calculation design. In a second step one needs to charge balance the system. All this is then need to be combined into a proper calculation input.

This is clearly a multistep and many times multiplatform process that is

1. Complicated
2. Time consuming
3. Not user friendly

The orca_crystalprep utility is designed to automatically generate proper inputs for ORCA embedded cluster calculations with the aim to allow to a wide range of experienaced and not experienced users the ability to setup an embedded cluster calculation with a minimal effort.

orca_crystalprep requires its own input. By simply executing it from the terminal

```
orca_crystalprep
```

one gets printings of the Usage:

```
****************************************************************************************************************************
Generate initial ORCA CrystalPrep Input
****************************************************************************************************************************


============================================================================================================================
Usage: orca_crystalprep [Basename Input] [options]
============================================================================================================================


----------------------------------------------------------------------------------------------------
↪-------------
[Options]:
----------------------------------------------------------------------------------------------------
↪-------------
-geninput                      Generate Initial Input
----------------------------------------------------------------------------------------------------
↪-------------

****************************************************************************************************************************
Generate ORCA Embedding Cluster Inputs using the CrystalPrep Utility
****************************************************************************************************************************


============================================================================================================================
Usage: orca_crystalprep [CrystalPrep Input]
============================================================================================================================
```

and the different Options:

```
----------------------------------------------------------------------------------------------------
↪-------------
[CrystalPrep Input Options]:
----------------------------------------------------------------------------------------------------
↪-------------
General Definitions
----------------------------------------------------------------------------------------------------
↪-------------
DoCIF                          true                    This will process a .cif file
DoXYZ                          true                    This will process a .xyz file
InputCIF                       "CIFFileName"           Set the name of the CIFFileName
InputXYZ                       "XYZFileName"           Set the name of the XYZFileName
----------------------------------------------------------------------------------------------------
↪-------------
SuperCell Construction Definitions
----------------------------------------------------------------------------------------------------
↪-------------
DoSuperCell                    true                    Flag to generate a SuperCell
SCDimension                    "axbxc"                 The Dimension of the SuperCell (e.g.
↪"1x1x1")
InputCIF                       "CIFFileName"           Set the name of the CIFFileName
InputXYZ                       "XYZFileName"           Set the name of the XYZFileName
----------------------------------------------------------------------------------------------------
↪-------------
Embedding Cluster Definitions
----------------------------------------------------------------------------------------------------
↪-------------
DoEmbedding                     true                   Flag to generate the files for the␣
↪embedding approach
UseVolumeCriterion              true                   Volume Criterion to generate layers
UseDistanceCriterion            true                   Distance Criterion to generate layers
```

(continues on next page)

(continued from previous page)

```
CellVolumeFraction          value                   Cell(UniCell/SuperCell) fraction (default␣
↪1.0)
DoSimpleInput               true                    Flag to generate a conventional Embedding␣
↪Cluster input
DoICQMMMInput               true                    Flag to generate a Ionic-Crystal-QMMM␣
↪input
WritePDB                    true                    Flag to run a Ionic-Crystal-QMMM input␣
↪from a PDF file
QCCharge                    Charge Number           Specify the total QC Charge
QCMult                      Multiplicity Number     Specify the total Multiplicity
-------------------------------------------------------------------------------------------
↪--------------
Special Tasks on Embedding Cluster Construction/Definition
-------------------------------------------------------------------------------------------
↪--------------
DoLayers                    true                    Request Layers Definition
1) Layers Definition. There are 2 Options:
a) The Differnt regions are build in layers as multipoles of the UnitCell
b) The Differnt regions are build in layers around a predefined QC cluster via a QCAtom List
QCLayers                    QC Layers Number        Specify the number of the QC Layers
ECPLayers                   ECP Layers Number       Specify the number of the ECP Layers
PCLayers                    PC Layers Number        Specify the number of the PC Layers
HFLayers                    HF Layers Number        Specify the number of the HF Layers
Example Input
QCLayers 1
2) Atoms Definition. This is alternative to Layers Definition (e.g. DoLayers false)
NQCAtoms                    QC Atoms  Number        Specify the number of the QC Atoms
NHFAtoms                    HF Atoms  Number        Specify the number of the HF Atoms
NECPAtoms                   ECP Atoms  Number       Specify the number of the ECP Atoms
NPCAtoms                    PC Atoms  Number        Specify the number of the PC Atoms
QCAtoms                     QC Atoms  List          Specify the List   of the QC Atoms (e.g.␣
↪0,1,4,10...
HFAtoms                     HF Atoms  List          Specify the List   of the HF Atoms (e.g.␣
↪0,1,4,10...
ECPAtoms                    ECP Atoms  List         Specify the List   of the ECP Atoms (e.g.␣
↪0,1,4,10...
PCAtoms                     PC Atoms  List          Specify the List   of the PC Atoms (e.g.␣
↪0,1,4,10...
Example Input
NQAtoms= 4
QCAtoms 0,1,4,10
-------------------------------------------------------------------------------------------
↪--------------
Request an Explicit Atom Definion in Ionic-Crystal-QMMM
HINT: This is automatically Set to true if an Atom List (QC Atoms, ECP Atoms, ...) is provided␣
↪by the user
-------------------------------------------------------------------------------------------
↪--------------
SetQCAtoms                  true                    Set explicitely the QC Atoms
SetHFAtoms                  true                    Set explicitely the HF Atoms
SetECPAtoms                 true                    Set explicitely the ECP Atoms
SetPCAtoms                  true                    Set explicitely the PC Atoms
SetPC2Atoms                 true                    Set explicitely the PC2 Atoms
-------------------------------------------------------------------------------------------
↪--------------
Redefine SuperCell Origin. This is for shifting the center origin to a desired atom during the␣
↪SC Construction
This Helps to automatically construct desired Cluster structures using the Layers Definition
-------------------------------------------------------------------------------------------
↪--------------
ShiftOrigin                 true                    Request Origin Shift  to a particular Atom
```

(continues on next page)

(continued from previous page)

```
ChosenAtom                    an Atom Number          The Number of the Chosen atom (e.g. 2)
-------------------------------------------------------------------------------------------
Neutralize the Embedded Cluster in the Simple Input
Note that in the Ionic-Crustal QMMM Input this step is taken care on demand during the QM/MM␣
↪run
-------------------------------------------------------------------------------------------
ChargePoints                  the charge points       Specify the number of the charge points␣
↪(default 1000)
ChargeStep                    the charge points       Define the step during the iterations␣
↪(default 0.01)
ChargeThres                   the threshold           Define the neutralization threshold␣
↪(default 0.01)
Neutralization Schemes:
NeutralizingScheme1           true                ->  Neutralization is based on All Charges
NeutralizingScheme2           true                ->  Neutralization is based on different␣
↪charges
Equiping the ECP and PC regions
NeutralizingScheme3           true                ->  Neutralization is based on PC2 Region
-------------------------------------------------------------------------------------------
Special Definitions for the .metainfo File in the Ionic-Crustal QMMM
This helps Setting Charge/Spin in a specific atom type
It can also aid the neutralization step of the Embedded Cluster in the Simple Input
-------------------------------------------------------------------------------------------
NAtomTypes                    Number of Atom Types    Specify The Atom types that will be␣
↪defined
Example Input
#-------------------------
#Atom Type Charge Spin
#-------------------------
AtomTypes 2
Co    1    2.0   1.5
Co    2    3.0   0
end
This specifies the local Spin and Charge of a Td HS Co2+ center (Type 1) and a OH Co3+ center␣
↪(Type 2)
During the Embedding cluster construction
-------------------------------------------------------------------------------------------
```

In a first step the orca_crystalprep tool can be used to generate its own input. So by running:

```
orca_crystalprep crystalprep.inp -geninput
```

it will generate the following initial input:

```
-----------------------------------------------------------------------------
Initial Input: crystalprep.inp for Orca_CrystalPrep has been generated. All Done!
-----------------------------------------------------------------------------
```

which looks like the following:

```
%crystalprep


#************************
#Read CIF/XYZ
#************************
```

(continues on next page)

(continued from previous page)

```
DoCIF true

#-----------------------
#INPUT CIF/XYZ
#-----------------------
#InputCIF "CIFName.cif"
#-----------------------
#Set Special Tasks
#-----------------------
#SpaceGroupNumber 200

#************************
#Generate SuperCell
#************************
DoSuperCell true
SCDimension "1x1x1"

#************************
#Setup Embedding Approach
#************************
DoEmbedding true
DoLayers true

#-----------------------
#Atom Type Charge Spin
#-----------------------
#NAtomTypes 2
#Co 1 2.0  1.5
#Co 2 3.0  0.0

#************************
#Generate Inputs
#************************
#DoSimpleInput true
#DoICQMMMInput true

#Neutralize true
#QCCharge 0
#QCMult 1
#-----------------------
end
```

By providing names for the *.cif or *.xyz files that are desired to be processed and different options it is possible to generate a ready to run embedding cluster input as is shown in Fig. 7.72.

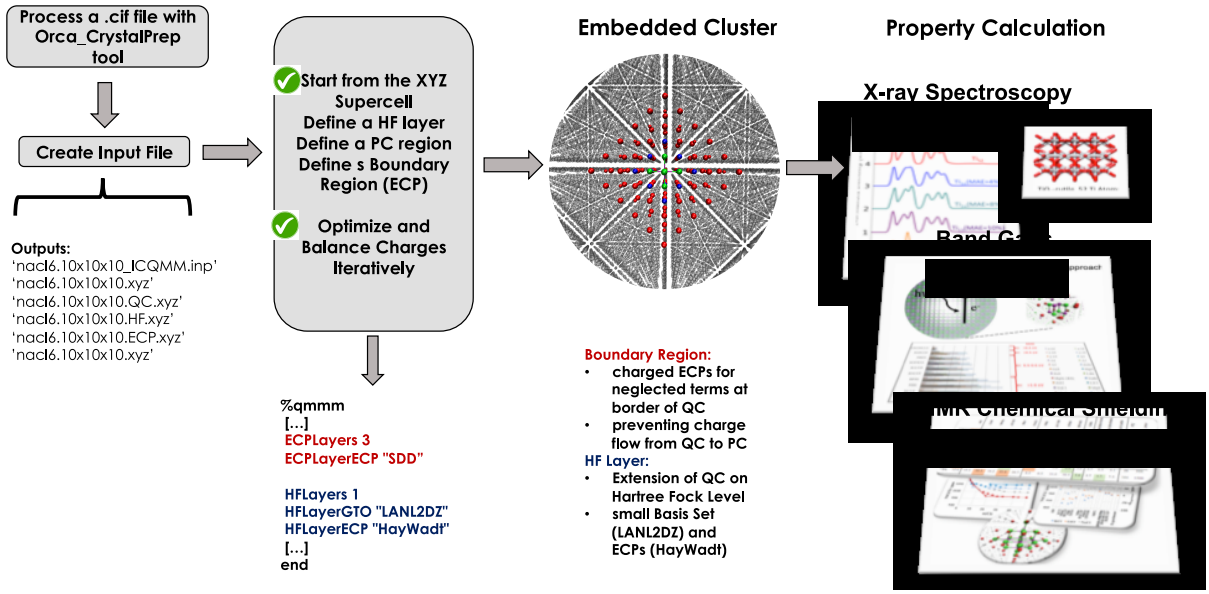


Fig. 7.72: Embedded cluster IC-QM/MM Input generation

For the construction of the embedded cluster structure by default a layers approach is performed in which the different structural layers are constructed as multiplets of the unitcell or a fraction of the unitcell

The unitcell fractions in terms of volume units are specified by the following keyword

```
CellVolumeFraction N #by default N=1
```

As an example let us discuss in detail the case of NaCl.

Let us assume that we want to generate an embedding cluster input

1. with a 20x20x20 supercell starting from the nacl.cif file
2. create an embedding cluster with 1 QC and 1ECPs layers
3. create an IC-QM/MM embedding cluster input

for this purpose the following input is used

```
%crystalprep


#************************
#Read CIF/XYZ
#************************
DoCIF true

#-----------------------
#INPUT CIF/XYZ
#-----------------------
InputCIF "nacl.cif"

#************************
#Generate SuperCell
#************************
DoSuperCell true
SCDimension "20x20x20"

#************************
#Setup Embedding Approach
#************************
```

(continues on next page)

(continued from previous page)

```
DoEmbedding true
DoLayers true

#------------------------
#Atom Type Charge Spin
#------------------------
NAtomTypes 2
Na 0  1.0  0.0
Cl 1 -1.0  0.0

#************************
#Generate Inputs
#************************
DoICQMMMInput true
QCCharge 0
QCMult 1
#------------------------
end
```

In a first step the orca_crystalprep will process the nacl.cif file and will create the unitcell and the requested 20x20x20 supercell

```
------------------------------------------------------------
Reading Information from the provided CIF file: nacl.cif
------------------------------------------------------------

------------------------Unit Cell Parameters-----------------
Hermann-Mauguin Space Group:     P1
Space Group ID:                 1
Unit Cell Symmetry:
Unit Cell Volume:              46.09
Unit Cell alpha angle:         60.00
Unit Cell beta angle:          60.00
Unit Cell gamma angle:         60.00
Unit Cell alpha length:        4.024
Unit Cell beta length:         4.024
Unit Cell gamma length:        4.024
Atom   Type   AO    x     y     z    occ
Na     0     11 0.000 0.000 0.000  1.00
Cl     1     17 0.500 0.500 0.500  1.00
Done
------------------------------------------------------------

----------Making a SuperCell with 20x20x20 dimensions-----------
Unit Cell:
0          1          2
0       4.024000   0.000000   0.000000
1       0.000000   4.024000   0.000000
2       0.000000   0.000000   4.024000

Transformation Matrix:
0          1          2
0      80.480000   0.000000   0.000000
1       0.000000  80.480000   0.000000
2       0.000000   0.000000  80.480000


--------------------------------------------------------------------
Saving xyz file: nacl4.cif_20x20x20.xyz                   ...Done
--------------------------------------------------------------------
```

In a following step the costruction of the embedding cluster structure will be initiated

```
------------Making a Embedding Cluster Input ------------
------------Using the Layers Approach with: ------------
QC_Layers: 1
ECP_Layers: 1
PC_Layers: 1
--------------
Preparing Inputs ...
```

In a next step the center of the .xyz supercell will be assigned to the closest atom:

```
-------------------------------
The Center of XYZ Coordinates
-------------------------------
41.246, 41.246, 41.246
-------------------------------
Closest Atom to Center
-------------------------------
Na(4630) 40.240, 40.240, 40.240
-------------------------------
Shifting Origin to closest atom
-------------------------------
40.240, 40.240, 40.240
```

In following an automatic layers generation is performed

```
-------------------------------------
Saving Generated Layers XYZ Files ...
-------------------------------------
Saving Layer 0 XYZ File: nacl.cif_20x20x20.xyz_0.xyz
Saving Layer 1 XYZ File: nacl.cif_20x20x20.xyz_1.xyz
Saving Layer 2 XYZ File: nacl.cif_20x20x20.xyz_2.xyz
Saving Layer 3 XYZ File: nacl.cif_20x20x20.xyz_3.xyz
Saving Layer 4 XYZ File: nacl.cif_20x20x20.xyz_4.xyz
Saving Layer 5 XYZ File: nacl.cif_20x20x20.xyz_5.xyz
Saving Layer 6 XYZ File: nacl.cif_20x20x20.xyz_6.xyz
Saving Layer 7 XYZ File: nacl.cif_20x20x20.xyz_7.xyz
Saving Layer 8 XYZ File: nacl.cif_20x20x20.xyz_8.xyz
Saving Layer 9 XYZ File: nacl.cif_20x20x20.xyz_9.xyz

-------------------------------------
Saving Generated Layers PDB File ...
-------------------------------------
```

The embedding cluster is then constructed:

```
-------------------------------------
Saving Generated Cluster XYZ Files ...
-------------------------------------
Saving QC XYZ File: nacl.cif_20x20x20.xyz_QC.xyz
Saving ECP Region XYZ File: nacl.cif_20x20x20.xyz_ECP.xyz
Saving PC Region XYZ File: nacl.cif_20x20x20.xyz_PC.xyz
```

Finally the IC-QM/MM embedding cluster will be generated

```
-------------------------------------
Saving Embedding Cluster Inputs ...
-------------------------------------
Saving ICQMMM Input: nacl.cif_20x20x20.xyz.ICQMMM.inp
Done
----------------------------------------------------------------
```

Ionic-Crystal QM/MM requires the generation of a simple force field *.prms For details see section: *ORCA Multiscale Implementation* The needed information including charge and spin is taken from a *metainfo file. The

orca_crystalprep provides the possibility to externally set chatge and spin in the *.metainfo file

This is acheived by the following process:

At first processing of the .cif files assignes atom types to all the detected atoms in the asymetric unit

```
Atom   Type   AO    x     y     z    occ
Na     0      11 0.000 0.000 0.000  1.00
Cl     1      17 0.500 0.500 0.500  1.00
```

In the orca_crystalprep input it is possible to assign specific initial charge and spin to atoms of a paricular atom type in the following way:

```
NAtomTypes 2
Na 0  1.0  0.0
Cl 1 -1.0  0.0
```

This information is then passed in the *.metainfo file

```
18522
# atom nr. - element - atom type - formal charge - formal spin - molecule nr
0  Na   0  1  0           1
1  Na   0  1  0           1
2  Na   0  1  0           1
...

9261  Cl   1 -1  0        1
9262  Cl   1 -1  0        1
9263  Cl   1 -1  0        1
...
```

The constructed embedded cluster is shown in Fig. 7.73.

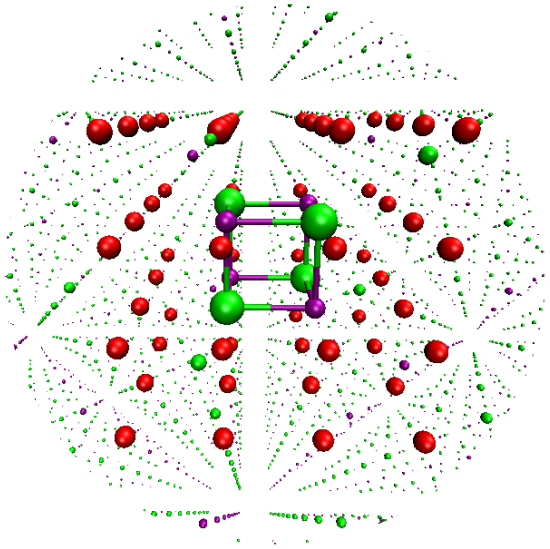


Fig. 7.73: Generated Embedded cluster. QC: Na_4Cl_4, ECP region red dots, PC region small green and purple dots

while the generated IC-QM/MM embeding cluster input is provided below

```
!Ionic-Crystal-QMMM
#Include Method
```

```
%qmmm
#--------Define the Cluster---------
ORCAFFFilename= ""
Use_QM_InfoFromPDB true
Use_QM3_InfoFromPDB true
ECPLayerECP= "SDD"
#--------Charge Convergence---------
CONV_Charges false
ENFORCETOTALCHARGE true
CHARGE_TOTAL    0
PrintLevel 4
end
#---------------------------------

*pdbfile 0  1 nacl.cif_20x20x20.pdb
```

Note: that the information regarding the QC, ECP and PC regions is read from the generated *.pdb file

- Further information regarding the IC-QM/MM and the QM/MM module in general is provided in section *ORCA Multiscale Implementation*
- Further information and examples regarding the orca_crystalprep tool and the embedding approach is provided in the Treating Solids with the Embedding Cluster approach tutorial.

7.56 Compound Methods

7.56.1 Commands

Below is a list of all available commands available in *Compound*

———————————— Dataset Related ————————————

- Dataset(*Dataset*)
- MakeReferenceFromDir(*D.MakeReferenceFromDir*)
- Print(*D.Print*)

———————————— File Handling Related ————————————

- CloseFile (*CloseFile*)
- OpenFile (*OpenFile*)

————————————- *For* Loop Related ————————————-

- Break (*Break*)
- Continue (*Continue*)
- EndFor (*EndFor*)
- For (*For*)

———————————— Geometry Related ————————————-

- Geometry(*Geometry*)
- BohrToAngs(*G.BohrToAngs*)
- CreateBSSE(*G.CreateBSSE*)
- FollowNormalMode(*G.FollowNormalMode*)
- GetAtomicNumbers(*G.GetAtomicNumbers*)
- GetBondDistance(*G.GetBondDistance*)

- GetCartesians(*G.GetCartesians*)
- GetGhostAtoms(*G.GetGhostAtoms*)
- GetNumOfAtoms(*G.GetNumOfAtoms*)
- MoveAtomToCenter(*G.MoveAtomToCenter*)
- Read(*G.Read*)
- RemoveAtoms(*G.RemoveAtoms*)
- RemoveElements(*G.RemoveElements*)
- WriteXYZFile(*G.WriteXYZFile*)

———————— *If* block Related ————————

- If (*If*)

————————- Linear Algebra Related ————————-

- Diagonalize(*Diagonalize*)
- InvertMatrix(*InvertMatrix*)
- Mat_p_Mat(*Mat_p_Mat*)
- Mat_x_Mat(*Mat_x_Mat*)
- Mat_x_Scal(*Mat_x_Scal*)

———————— ORCA calculation Related ————————

- ReadMOs(*ReadMOs*)

———————— Program flow Related ————————

- Abort (*Abort*)
- End (*End*)
- EndRun (*EndRun*)
- GoTo (*GoTo*)

———————— Property File Related ————————-

- GetNumOfInstances(*GetNumOfInstances*)
- Read(*Read*)
- ReadProperty(*ReadProperty*)

———————— String Handling Related ————————

- GetBasename (*S.GetBasename*)
- GetChar (*S.GetChar*)
- GetSuffix (*S.GetSuffix*)
- Print (*Print*)
- Write2File (*Write2File*)
- Write2String (*Write2String*)

———————— Step Related ————————

- & (*&*)
- Alias (*Alias*)

———————— Timer Related ————————

- Timer (*Timer*)

- Last (*T.Last*)
- Reset (*T.Reset*)
- Start (*T.Start*)
- Stop (*T.Stop*)
- Total (*T.Total*)

———————— Variables Related ————————-

- Variables (*Variables - General*)
- Variables - Assignment (*Variables - Assignment*)
- Variables - Declaration (*Variables - Declaration*)
- Variables - Functions (*Variables - Functions*)
- Variable - With (*With*)
- GetBool() *V.GetBool()*
- GetDim1() *V.GetDim1()*
- GetDim2() *V.GetDim2()*
- GetDouble() *V.GetDouble()*
- GetInteger() *V.GetInteger()*
- GetSize() *V.GetSize()*
- GetString() *V.GetString()*
- PrintMatrix() *V.PrintMatrix()*
- Write2File (*Write2File*)
- Write2String (*Write2String*)

&

The *&* symbol has a special meaning in the compound block. Using this symbol inside the *New_Step - Step_End* block the user can use variables that are defined outside the block. Both string and numerical variables are allowed.

Syntax:

&{variable}

Example

```
# ------------------------------------------
# This script checks the options for the
#             '&' symbol
# ------------------------------------------
%Compound

  Variable method   = "BP86";              #string variable'
  Variable basis    = "def2-TZVP def2/J";
  Variable name     = "base";
  Variable number   = 0;                   #integer variable
  Variable distance = 0.8;                 #double variable
  New_step
    ! &{method} &{basis} TightSCF
    %base "&{name}_&{number}"              #combination of variables
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 &{distance}
```

(continues on next page)

(continued from previous page)

```
    *
  Step_End
  # ----------------------------------------
  # Add some printing
  # ----------------------------------------
  print("SUMMARY OF VARIABLES\n");
  print("Method:   %s\n",    method);
  print("Basis:    %s\n",    basis);
  print("Name:     %s\n",    name);
  print("Number:   %d\n",    number);
  print("Distance: %.2lf\n", distance);
End
```

Abort

Abort is used when the user wants to exit the program instantly. **Syntax:**
Abort;
or alternatively:
Abort

Example:

```
%Compound
for i from 0 to 4 do
  print("i: %d\n", i);
  if (i=2) then
    abort;
  endif
endfor
End
```

Alias

Alias is used to replace an integer number with a more representative string. It is useful when one performs more than one calculations and the step numbers become too complicated to evaluate. In this case using Alias_Step after the Step_End command will connect the preceeding calculation step number with the provided name.

Syntax:
Alias $name$;
or alternatively:
Alias $name$
Example:

```
# ---------------------------------
# This script checks 'alias' keyword
# ---------------------------------
Variable numOfSteps = 20;
Variable Range      = 4.0;
Variable distStart  = 0.4;
Variable Step       = Range/numOfSteps;
Variable distance;
Variable Energies[numOfSteps];
Variable simpleInput  = "BP86 def2-SVP def2/J";

For index from 0 to numOfSteps-1 Do
  Distance =  distStart+ index*Step;
  New_Step
    !&{simpleInput}
```

(continues on next page)

(continued from previous page)

```
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 &{Distance}
    *
  Step_End
  Alias currStep;
  print("Current step: %d\n", currStep);
  Read Energies[index] = JOB_INFO_TOTAL_EN[currStep];
EndFor

print("---------------------------------------\n");
print("       Compound Printing               \n");
print("%s  %12s  %16s \n","Step",  "Distance", "Energy");
print("---------------------------------------\n");
For index from 0 to numOfSteps - 1 Do
  Distance = distStart + index*Step;
  print("%4d  %12.4lf   %16.8lf \n", index, Distance, Energies[index]);
EndFor

End
```

NOTE for the **Alias** command the final ';' is optional.

Break

Break can be used inside a *For* loop (see (see *For*)) when one needs to break the loop under certain conditions. The syntax is the following:

Syntax:
For variable From Start value To End value Do
commands
break ;
commands
EndFor;

NOTE both versions *break* and *break;* are legal.

What break actually does is to set the running index of the loop to the last allowed value and then jump to the *EndFor* (see *EndFor*)).

Example:

```
#
# This a script to check Compound 'break' command
#
%Compound
print(" Test for 'break'\n");
print(" It should print 0, 1 and 2\n");
for i from 0 to 6 Do
  if (2*i > 4) then
    break
  endIF
  print("index: %d\n", i);
EndFor

print("Continued outside the 'for' loop\n");

End
```

CloseFile

When a file is opened in *Compound* using the *openFile* command (see *OpenFile*), then it must be closed using the *closeFile* command.

Syntax:
closeFile(file);

file is the file pointer created from the *openFile* command. For an example see paragraph *OpenFile*.

Continue

Continue can be used inside a *For* loop (see (see *For*)) when one needs to skip the current step of the loop and proceed to the next one. The syntax is the following:

Syntax:
For variable From Start value To End value Do
commands
continue ;
commands
EndFor;

NOTE both versions *continue* and *continue;* are legal.

What continue actually does is to jump to the *EndFor* (see *EndFor*)).

Example:

```
# ------------------------------------------------------
# This is a script to check Compound 'continue' command
# ------------------------------------------------------
%Compound
print(" Test for 'continue'\n");
print(" It should print 0, 1, 2 and 4\n");
for i from 0 to 4 Do
  if ( i=3) then
    continue;
  endIf
  print("index: %d\n", i);
EndFor
End
```

Dataset

In *Compound* we have *Dataset* objects. These objects can be treated like normal variables of type '*compDataset*'. An important difference between normal variables and *Dataset* variables is the declaration. Instead of the normal:

```
Variable x;
```

we excplicitly have to declare that this is a dataset. So the syntax for a dataset declaration is:

Syntax:

```
Dataset mySet;
```

NOTE in the case of datasets we do not allow multiple dataset declarations per line.

Below is a list of functions that work on *Dataset*.

- MakeReferenceFromDir(*D.MakeReferenceFromDir*)
- Print(*D.Print*)

Example:

```
# ----------------------------------------------------
# This is an example script for dataset definition
# ----------------------------------------------------
%Compound
  dataset mySet;
  mySet.Print();
End
```

D.MakeReferenceFromDir

MakeReferenceFromDir command acts on a dataset object (see *Dataset*). It creates a json reference file based on the *xyz files of the current folder. **NOTE** By default all charges and multplicities will be set to 0 and 1 respectively.

Syntax:
mySet.MakeReferenceFromDir(dirName);

Where:

- *mySet* is a dataset object that is already declared
- *dirName* The name of a directory that should contain some xyz files.

Example:

```
%Compound
  # ------------------------------------
  #   This is a compound script that will
  #   check dataset.MakeReferenceFromDir
  #   function
  #   NOTE: The script assumes that some
  #         xyz files rest in the current
  #         directory
  # ------------------------------------

  # First some definitions
  Variable name = "mySet";
  Variable numOfMolecules = 0;
  dataset mySet;
  Variable myDir="./";
  mySet.MakeReferenceFromDir(myDir);
  mySet.ReadReferenceFile();
  mySet.Print();
End
```

D.Print

Print command acts on a dataset object (see *Dataset*). It prints all details of the specific dataset object.

Syntax:
mySet.Print();

Where:

- *mySet* is a dataset object that is already declared

Example:

```
# ----------------------------------------------------
# This is an example script for dataset definition
# ----------------------------------------------------
%Compound
  dataset mySet;
```

(continues on next page)

(continued from previous page)

```
  mySet.Print();
End
```

Diagonalize

Compound can peform matrix algebraic operations, one of the available algebraic operation is matrix diagonalization. Be carefull that the matrix, that is to be diagonalized, **must be** a square symmetric matrix. It is also important to remember that only the upper triangle part of the matrix will be used for the diagonalization. If everything proceeds smoothly then the function will return the eigenvectors and eigenvalues of the matrix.

Syntax:
A.Diagonalize(eigenValues, eigenVectors);

Where:

- *A:* The matrix to be diagonalized.
- *eigenValues:* The vector with the eigenvalues
- *eigenVectors:* The square matrix with the eigenvectros of the initial matrix.

Example:

```
# ------------------------------------------------------
# This is an example script for diagonalization
# ------------------------------------------------------
%Compound
  Variable Dim=3;
  Variable A[Dim][Dim];
  Variable eigenVal;
  Variable eigenVec;
  for i from 0 to Dim-1 Do
    for j from 0 to Dim-1 Do
      if (i<=j) then
        A[i][j] = i+j+1;
      else
        A[i][j] =  A[j][i];
      EndIf
    EndFor
  EndFor
  A.Diagonalize(eigenVal, eigenVec);
  A.PrintMatrix();
  eigenVal.PrintMatrix();
  eigenVec.PrintMatrix();
End
```

End

End is the final command each compound script must have (unless there is an *EndRun* command (see *EndRun*)). After *Compound* executes what is written in the script then it passes control again to normal ORCA input reading. ORCA will continue analyze the input that rests after the *Compound* part but it will not run any calculation.

EndRun

#EndRun* is an alternative to the *End* command (see *EndRun*) for ending the execution of a *Compound* script. The difference between *end* and *EndRun* is that *EndRun* ignores everything after the *Compound* block. This makes it even easier to use *Compound* as a full workflow run.

EndFor

All *For* loops (see *For*) must finish with *EndFor*. The syntax and an example is shown in the *For* section (see *For*).

NOTE For *EndFor* both *EndFor* and *EndFor;* are possible.

For

For loops are used to perform repetitive tasks. The syntax is the following:

Syntax:

For *variable* **From** *Start value* **To** *End value* **Do**

commands

EndFor or **EndFor;**

Variable should be a variable name not previously defined. *Start value* and *End value* should be integers defining the start and end value of the *variable*. *Start value* and *End value* can be numbers, predefined variables or functions of previously defined variables. The only requirement is that they should be **integers**. Keep in mind that the loop will be performed from the first value to the *End value*, including the *End value*.

Example:

```
# ---------------------------------------
# This is a script to check 'for' loops
# ---------------------------------------

# ---------------------------------
# Some necessary initial definitions
# ---------------------------------
Variable x = {0.0, 1.0, 2.0, 3.0, 4.0};
Variable f;
Variable loopStart;
Variable upLimit;

# ---------------------------------
# Case 1.
# Constant Start / Constant End
# ---------------------------------
print(" --------------  Case 1 --------------\n");
print("     Constant Start / Constant End    \n");
print("          f = index*x[index]          \n");
print( "      for index from 0 to 4 Do     \n");
print(" ------------------------------------\n");
for index from 0 to 4 Do
  f = index*x[index];
  print("Index: %3d    x[index]:  %.2lf   f:   %.2lf\n", index, x[index], f);
EndFor

loopStart = 0;
upLimit   = 4;
print(" --------------  Case 2 --------------\n");
print("     Variable Start / Variable End    \n");
print("          f = index*x[index]          \n");
```

(continues on next page)

(continued from previous page)

```
print( "   for index from loopStart to upLimit Do\n");
print(" ------------------------------------\n");
for index from loopStart to upLimit Do
  f = index*x[index];
  print("Index: %3d    x[index]:  %.2lf   f:   %.2lf\n", index, x[index], f);
EndFor


loopStart  = 1;
upLimit    = 3;
print(" --------------  Case 3 --------------    \n");
print("     function Start / function End        \n");
print("          f = index*x[index]              \n");
print( "   for index from start-1 to upLimit+1 Do\n");
print(" ------------------------------------    \n");
for index from loopStart-1 to upLimit+1 Do
  f = index*x[index];
  print("Index: %3d    x[index]:  %.2lf   f:   %.2lf\n", index, x[index], f);
EndFor

End
```

Geometry

In *Compound* we have *Geometry* objects. These objects can be treated like normal variables of type '*compGeometry*'. An important difference between normal variables and *Geometry* variables is the declaration. Instead of the normal:

```
Variable myGeom;
```

we excplicitly have to declare that this is a geometry. So the syntax for a geometry declaration is:

Syntax:

```
Geometry myGeom;

Geometry myGeom1, myGeom2;
```

Using the second definition one can define two geometry objects in the same line.

Below is a list of functions that work on *Geometry* objects.

- BohrToAngs(*G.BohrToAngs*)
- CreateBSSE (*G.CreateBSSE*)
- FollowNormalMode (*G.FollowNormalMode*)
- GetAtomicNumbers(*G.GetAtomicNumbers*)
- GetBondDistance(*G.GetBondDistance*)
- GetCartesians(*G.GetCartesians*)
- GetGhostAtoms(*G.GetGhostAtoms*)
- GetNumOfAtoms(*G.GetNumOfAtoms*)
- MoveAtomToCenter(*G.MoveAtomToCenter*)
- Read(*G.Read*)
- RemoveAtoms(*G.RemoveAtoms*)
- RemoveElements(*G.RemoveElements*)

- WriteXYZFile(*G.WriteXYZFile*)

G.BohrToAngs

BohrToAngs command acts on a geometry object (see see *Geometry*). It will transform the geometry of the loaded geometry object from Bohr to Angstroms. Practically it will just multiply the coordinates with the factor *0.529177249*.

Syntax:
myGeom.BohrToAngs();

Where:

- *myGeom* is a geometry object that already contains a geometry

Example:

```
# ------------------------------------------------------
# This is a script to check the BohrToAngs function
# ------------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable CC;
  New_Step
    !BP86
  Step_End
  myGeom.Read();
  myGeom.BohrToAngs();
  CC = myGeom.GetCartesians();
  CC.PrintMatrix();
End
```

G.CreateBSSE

CreateBSSE command acts on a geometry object (see see *Geometry*). In the case that the geometry object contains *ghost* atoms then *CreateBSSE* will create five new files:

- myFilename_FragmentA.xyz
- myFilename_MonomerA.xyz
- myFilename_FragmentB.xyz
- myFilename_MonomerB.xyz
- myFilename_Total.xyz

Syntax:
myGeom.CreateBSSE(filename=myFilename);

Where:

- *myGeom* is a geometry object that already contains a geometry
- *filename* is a base filename for the created files.

Example:

```
# ------------------------------------------------------
# This is a scrtipt to check the geom.CreateBSSE command
# ------------------------------------------------------

*xyz 0 1
o:      -1.69296787   -0.05579265    0.00556629
h:      -2.01296504    0.84704339   -0.01586469
h:      -0.73325076    0.04238910    0.00084302
o        1.23009925    0.02698440   -0.00375550
h        1.60672086   -0.41139567    0.76236888
h        1.60236356   -0.44922858   -0.74915800
*

%Compound
  Geometry monomerA;
  variable myFilename     = "BSSE";
  Variable method         = "BP86";

  # --------------------------------------
  # Calculation for Fragment A
  # --------------------------------------
  New_Step
    !&{method}
  Step_End

  # --------------------------------------
  # Read the geometry of Fragment A
  # --------------------------------------
  monomerA.Read();

  # --------------------------------------
  # Create the missing xyz files
  # --------------------------------------
  monomerA.CreateBSSE(filename=myFilename);

End
```

NOTE The files will contain XYZ geometries in **BOHRS**.

G.FollowNormalMode

FollowNormalMode command acts on a geometry object (see see *Geometry*). It will displace the loaded geometry following a chosen normal mode of vibtation

Syntax:
myGeom.FollowNormalMode(vibrationSN=myVibration, [ScalingFactor=myScalingFactor]);

Where:

- *myGeom* is a geometry object that already contains a geometry
- *vibrationSN* is the serial number of the vibration. **NOTE** Please remember that counting starts with 1.
- *scalingFactor* is the scaling of the normal mode of vibration. This argument is optional.

Example:

```
# ------------------------------------------------------
# This is a script to check the followNormalMode function
# ------------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
```

(continues on next page)

(continued from previous page)

```
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable CC, normalModes;
  Variable res = -1;
  New_Step
    !BP86 Freq
  Step_End
  myGeom.Read();
  myGeom.FollowNormalMode(vibrationSN=7, scalingFactor=0.8);
  CC = myGeom.GetCartesians();
  CC.PrintMatrix();
End
```

G.GetAtomicNumbers

Function *GetAtomicNumbers* acts on geometry objects and returns and array with the atomic numbers of the elements in the working geometry.

Syntax: *atomNumbers = geom.GetAtomicNumbers()*

atomNumbers A variable that will be filled with the values of the atomic numbers

geom A geometry object that should already be loaded.

Example:

```
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable atomicNumbers;

  New_Step
    !BP86
  Step_End
  myGeom.Read();
  atomicNumbers = myGeom.GetAtomicNumbers();
  print("\nCompound \n");
  for i from 0 to atomicNumbers.GetSize()-1 Do
    print("Atom '%d': atomic number: %d\n", i, atomicNumbers[i]);
  EndFor
End
```

G.GetBondDistance

Function *GetBondDistance* acts on geometry objects and returns the distance between two atoms **in Bohrs**.

Syntax:

res = geom.GetBondDistance(atomA, atomB)

Where:

res The distance between atoms atomA and atomB.

geom A geometry object previously loaded.

atomA The index of atomA in the geometry.

atomB The index of atomB in the geometry.

NOTE indices start counting from 0

Example:

```
# ------------------------------------
# This is to test Geometry function
#    GetBondDistance
# ------------------------------------
%Compound
  Variable dist;
  Geometry myGeom;
  New_Step
    !BP86
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
  myGeom.Read();  #Reads teh geometry of the previous step
  print( " ------------------------------------------------------\n");
  print( "  Compound Geometry functions test (GetBondDistance) \n");
  print( "  It should print 1.5118\n");
  print( " ------------------------------------------------------\n");
  print( " The distance between atom %d and atom %d  is: %.4lf Bohr\n",
           0, 1, myGeom.GetBondDistance(0,1));

End
```

G.GetCartesians

Function *GetCartesians* acts on geometry objects (see *Geometry*) and returns the distance xyz cartesian coordinates. Please remember that it alsways returns the cooridnates in **BOHRS**.

Syntax:

coords = geom.GetCartesians()

Where:

coords: A (*nAtoms*,3) array with the cartesian coordinates in **BOHRS**.

geom: A geometry object previously loaded.

Example:

```
# ------------------------------------------------
# This is a script to check the GetCartesians function
# NOTE: It always return it in Bohrs!
```

(continues on next page)

(continued from previous page)

```
# ------------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable CC;
  New_Step
    !BP86
  Step_End
  myGeom.Read();
  CC = myGeom.GetCartesians();
  for i from 0 to CC.GetDim1()-1 Do
    print("%12.9lf  %12.9lf  %12.9lf\n",
    CC[i][0], CC[i][1], CC[i][2]);
  endFor
End
```

G.GetGhostAtoms

Function *GetGhostAtoms* acts on geometry objects (see *Geometry*). It returns a vector of size nAtoms where for each atom the value will be -1 if it is a ghost atom, otherwise the atomic number of the element

Syntax:

ghostAtoms = geom.GetGhostAtoms()

Where:

ghostAtoms: A (*nAtoms*,1) integer vector with values -1 or the atomic number of the atom, in case it is not a ghost atom.

geom: A geometry object previously loaded.

Example:

```
# ------------------------------------------------------
# This is a script to check the getGhostAtoms function
# ------------------------------------------------------
*xyz 0 1
o:      -1.69296787   -0.05579265    0.00556629
h:      -2.01296504    0.84704339   -0.01586469
h:      -0.73325076    0.04238910    0.00084302
o        1.23009925    0.02698440   -0.00375550
h        1.60672086   -0.41139567    0.76236888
h        1.60236356   -0.44922858   -0.74915800
*

%Compound
  Geometry myGeom;
  Variable ghostAtoms;
  New_Step
    !BP86
  Step_End
  myGeom.Read();
  ghostAtoms = myGeom.GetGhostAtoms();
  ghostAtoms.PrintMatrix();
End
```

G.GetNumOfAtoms

GetNumOfAtoms returns an integer with the number of atoms of the working geometry.

Syntax:

res = geom.GetNumOfAtoms();

Where:

res is the resulting number of atoms

geom is the name of a geometry variable (see *Geometry*) we are using.

Example:

```
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable numOfAtoms = 0;

  New_Step
    !BP86
  Step_End
  Alias currStep;
  myGeom.Read(currStep);
  numOfAtoms = myGeom.GetNumOfAtoms();
  print("\nCompound \n");
  print("Number of atoms: %d (it should print 3)\n", numOfAtoms);

End
```

G.MoveAtomToCenter

Function *MoveAtomToCenter* acts on geometry objects (see *Geometry*). It will adjust the cartesian coordinates so that the chosen atom will rest at (0.0 0.0 0.0).

Syntax:

geom.MoveAtomToCenter(atom serial nubmer);

Where:

geom: A geometry object previously loaded.

atom serial number: The serial number of the atom in the geometry.

Example:

```
# ------------------------------------------------------
# This is a script to check the moveAtomToCenter function
# ------------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
```

(continues on next page)

(continued from previous page)

```
  Variable CC;
  New_Step
    !BP86
  Step_End
  myGeom.Read();
  myGeom.MoveAtomToCenter(0);
  myGeom.BohrToAngs();
  CC = myGeom.GetCartesians();
  CC.PrintMatrix();
End
```

NOTE Please remember that counting **starts with 0**, meaning that the first atom is 0 and not 1!

G.Read

Function *Read* acts on geometry objects (see *Geometry*) and reads a geometry from a property file related to a previous step.

Syntax:

geom.Read([stepID=myStepID], propertySN=myPropertySN)

Where:

geom: A geometry object that will be updated.

stepID: The step from which we are going to read the geometry. If not given the previous step will be used.

propertySN: The serial number the geometry in the property file. If not given the last available geometry will be used.

Example:

```
# ------------------------------------------------------
# This is a script to check the Read function for
#   geometries
# ------------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable CC;
  New_Step
    !BP86 opt
  Step_End
  myGeom.Read(propertySN=2);
End
```

G.RemoveAtoms

Function *RemoveAtoms* acts on geometry objects (see *Geometry*). It accepts a list of atoms and removes them from the loaded geometry. In the end the geometry object will be updated.

Syntax:

geom.RemoveAtoms(atom1, atom2, . . .);

Where:

geom: A geometry object previously loaded.

atom1, atom2, . . .: The serial number of the atoms in the geometry.

Example:

```
# ------------------------------------------------
# This is a script to check the RemoveAtoms function
# ------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable numOfAtoms;
  New_Step
    !BP86
  Step_End
  myGeom.Read();
  numOfAtoms = myGeom.GetNumOfAtoms();
  print("Number of atoms before: %d (It should print 3)\n", numOfAtoms);
  myGeom.RemoveAtoms(0);  #Remove the first atom
  numOfAtoms = myGeom.GetNumOfAtoms();
  print("Number of atoms after : %d (It should print 2)\n", numOfAtoms);
End
```

NOTE Please remember that counting **starts with 0**, meaning that the first atom is 0 and not 1!

G.RemoveElements

Function *RemoveElements* acts on geometry objects (see *Geometry*). It will remove from the loaded geometry all atoms with an atomic number given in the list.

Syntax:

geom.RemoveElements(atomNumber1, atomicNumber2, . . .);

Where:

geom: A geometry object previously loaded.

atomicNumber1, atomicNumber2, . . .: The atomic number of elements to be removed from the current geometry.

Example:

```
# ------------------------------------------------
# This is a script to check the RemoveElements function
# ------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
```

(continues on next page)

(continued from previous page)

```
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable numOfAtoms;
  Variable CC;
  New_Step
    !BP86
  Step_End
  myGeom.Read();
  numOfAtoms = myGeom.GetNumOfAtoms();
  CC = myGeom.GetCartesians();
  CC.PrintMatrix();
  print("Number of atoms before: %d (It should print 3)\n", numOfAtoms);
  myGeom.RemoveElements(8);  #Remove the oxygen
  numOfAtoms = myGeom.GetNumOfAtoms();
  print("Number of atoms after : %d (It should print 2)\n", numOfAtoms);
  CC = myGeom.GetCartesians();
  CC.PrintMatrix();
End
```

G.WriteXYZFile

Function *WriteXYZFile* acts on geometry objects (see *Geometry*) and writes on disc an xyz file with the coordinates of the current goemetry object. Please remember that it alsways writes the coordinates in **BOHRS**.

Syntax:

res = geom.WriteXYZFile(filename=myFilename)

Where:

res: An integer that returns '0' if everything worked smoothly.

myFilename The name of the file that will contain the coordinates.

geom: A geometry object previously loaded.

Example:

```
# ------------------------------------------------------
# This is a script to check the WriteXYZ function
# NOTE: It always write the coordinates in Bohrs!
# ------------------------------------------------------
*xyz 0 1
  O      -1.69296787   -0.05579265    0.00556629
  H      -2.01296504    0.84704339   -0.01586469
  H      -0.73325076    0.04238910    0.00084302
*

%Compound
  Geometry myGeom;
  Variable res=-1;
  New_Step
    !BP86
  Step_End
  myGeom.Read();
  res = myGeom.WriteXYZFile(filename="myGeom.xyz");
End
```

GetNumOfInstances

The *GetNumOfInstances* returns the number of instances of a specific object in a propertyfile.

Syntax:

[res=] GetNumOfInstances(propertyName=myName, [step=myStep], [filename=myFilename], [baseProperty=true/false])

Where:

res: An integer that returns the number of instances of the required property in the property file.

propertyName: A string alias that defines the variable the user wants to read.

step: The step from which we want to read the property. If not given the property file from the last step will be read.

filename: A filename of a property file. If a filename and at the same time a step are provided the program will ignore the step and try to read the property file with the given filename.

NOTE please note that in the end of filename the extension *.property.txt* will be added.

baseProperty: A true/false boolean. The default value is set to false. If the value is set to true then a generic property of the type asked will be read. This means if dipole moment is asked, it will return the last dipole moment, irrelevant if and MP2 or SCF one wad defined.

Example

```
# ------------------------------------------------------
# This is an example script for readNumOfInstances
# ------------------------------------------------------
%Compound
  Variable res = 0;
  Variable myProperty="MP2_DIPOLE_TOTAL";
  Variable myBaseProperty="DIPOLE_MOMENT_TOTAL";
  New_Step
    !MP2
    #%mp2
    #  density relaxed
    #end
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
  # First read the MP2 dipole moment
  res = GetNumOfInstances(propertyName=myProperty);
  print("Num of MP2 dipole moments   : %d\n", res);
  res = GetNumOfInstances(propertyName=myBaseProperty, Property_Base=true);
  print("Num of total dipole moments : %d\n", res);
End
```

GoTo

The *GoTo* command allows the 'jump' inside the normal flow of a *compound* script. The syntax of the command can be best presented through an example.

Example:

```
# ------------------------------------------------------
# This is an example script for GoTo
# (It should print only 0,1,2,3)
# ------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
%Compound
  Variable TCut=3;
  Variable Done;
  for i from 0 to 6 Do
    print("Index: %d\n", i);
    if (i >= TCut) then
      GoTo Done;
    EndIf
  EndFor
  Done:
    print("Done\n");
End
```

Please note that the variable we use as a label for the *GoTo* command should be previously defined like a normal variable.

If

The *if* block allows the user to make decisions. The syntax in *Compound* is the following:
Syntax:

If *(expression)* **Then**

actions

Else if *(expression)* **Then**

actions

Else

actions

Endif

Below is an example of the usage of *if block* in *compound*.

```
# ------------------------------------------------------------
# This is to check all available ways of 'if blocks'
# ------------------------------------------------------------
Variable x1 = 10.0;
Variable y1 = 20.0;
Variable b1 = False;
Variable b2 = True;
Variable s1 = "alpha";
Variable s2 = "beta";
Variable s3 = "alpha";
print( " ------------------------------------------------------------ \n");
print( " ----------      SUMMARY OF IF CASES        -------------- \n");
print( " ------------------------------------------------------------ \n");

print(" x1: %.1lf\n", x1);
print(" y1: %.1lf\n", y1);
print(" b1: %s\n", b1.GetString());
print(" b2: %s\n", b2.GetString());
print(" s1: %s\n", s1);
print(" s2: %s\n", s2);
print(" s3: %s\n", s3);
# ****************************************************************
#                            DOUBLES
# ****************************************************************
print(" -------------------     Doubles     ---------------------- \n");
print("      Variable/constant / One operator  / if (x1>5)         \n");
```

(continues on next page)

(continued from previous page)

```
print("                    No else if/No else                    \n");
if (x1>5) then
  print(" %.2lf > 5 \n", x1);
endif
# ------------------------------------------------------------------
print("      function / Variable / One operator  / if (3*x1>y1)    \n");
print("                   no else if /   else                      \n");
if (3*x1>y1) then
  print(" 3*%.1lf > %.1lf\n", x1, y1);
else
  print(" 3*%.1lf < %.1lf\n", x1, y1);
endif
# ------------------------------------------------------------------
print("     function / function / One operator  / if (x1-y1>-10.0) \n");
print("                    else if/else                       \n");
if (x1-y1>-10.0) then
  print(" %.2lf - %.2lf > -10.0\n", x1, y1);
else if (x1-y1 < -10.0) then
  print(" %.2lf - %.2lf < -10.0\n", x1, y1);
else
  print(" %.2lf - %.2lf = -10.0\n", x1, y1);
endif
# ******************************************************************
#                           BOOLEANS
# ******************************************************************
print(" ---------------      Booleans      ---------------------------\n");
print("       Variable  /  No operator  /  if (b1)                   \n");
if (b1) then
  print("b1 is True\n");
else
  print("b1 is False\n");
endIf
# ------------------------------------------------------------------------
# ------------------------------------------------------------------------
print("       Constant  / No operator  / if (true)                   \n");
if (True) then
  print( "True\n");
else
  print( "False");
endIf
# ------------------------------------------------------------------------
# ------------------------------------------------------------------------
print("       Variable/Variable / AND operator / if (b1 and b2)                 \n");
if (b1 and b2) then
  print("(%s and %s) is true\n", b1.GetString(), b2.GetString());
else
  print("(%s and %s) is not true\n", b1.GetString(), b2.GetString());
endIf
# ------------------------------------------------------------------------
# ------------------------------------------------------------------------
print("       Variable/Variable / OR operator / if (b1 or b2)                \n");
if (b1 OR b2) then
  print("(%s or %s) is true\n", b1.GetString(), b2.GetString());
else
  print("(%s or %s) is not true\n", b1.GetString(), b2.GetString());
endIf
# ------------------------------------------------------------------------
# ------------------------------------------------------------------------
print("       Bool /Doubles Function / AND operator / if (b1 and x1>y1 )\n");
if (b1 and y1>x1) then
  print("(%s and %.1lf>%.1lf) is True\n", b1.GetString(), x1, y1);
```

(continues on next page)

(continued from previous page)

```
else
  print("(%s and %.1lf>%.1lf) is False\n", b1.GetString(), x1, y1);
endIf
# ------------------------------------------------------------------------
# ------------------------------------------------------------------------
print("       Nested if / if (b2) then if (y1>x1)\n");
if (b2) then
  if (y1 > x1) then
    print ( "(%s is True) and (%.1lf>%.1lf)\n", b2.GetString(), y1, x1);
  else
    print ( "(%s is True) and (%.1lf<%.1lf)\n", b2.GetString(), y1, x1);
  endIf
else
    print ( "%s is False\n", b2.GetString());
endIf
# ------------------------------------------------------------------------
# ------------------------------------------------------------------------
print(" ----------------      Strings     --------------------------\n");
print(" ------------------------------------------------------------\n");
print(" ------------------------------------------------------------\n");
print("       Variable/Variable / if s1=s2\n");
if (s1=s2) then
  print("%s is same as %s \n", s1, s2);
else
  print("%s is not same as %s \n", s1, s2);
EndIf

# ------------------------------------------------------------------------
# ------------------------------------------------------------------------
print(" ----------------      Strings     --------------------------\n");
print(" ------------------------------------------------------------\n");
print(" ------------------------------------------------------------\n");
print("       Variable/constant / if s1=\"alpha\"\n");
if (s1="alpha") then
  print("%s is same as %s \n", s1, "alpha");
else
  print("%s is not same as %s \n", s1, "alpha");
EndIf
End
```

Some comments about the syntax:

The *Else if* or *Else* blocks are not obligatory.

The numerical operators that can be used are: *'>', '<', '>=', '<=', '='*.

The available logical operators are: *'and' and 'or'*.

Unfortunately in the current version multi-parentheses are not allowed.

There is now the possibility to compare strings.

InvertMatrix

Compound can peform matrix algebraic operations, one of the available algebraic operation is the inversion of a matrix. Be carefull that the matrix, whose the invert we are looking for, **must be** a real, square matrix.

Syntax:
AInvert = A.InvertMatrix();

Where:

- *A:* The matrix to be inverted.
- *AInvert:* The invert of A. It can be A itself and then A will just be updated.

Example:

```
# ------------------------------------------------------
# This is an example script for matrix inversion
# ------------------------------------------------------
%Compound
  Variable Dim=3;
  Variable A[Dim][Dim];
  Variable invertA, C;
  Variable res=-1;
  for i from 0 to Dim-1 Do
    for j from 0 to Dim-1 Do
      if (i=j) then
        A[i][j] = i+1;
      else
        A[i][j] =  0.0;
      EndIf
    EndFor
  EndFor
  invertA = A.invertMatrix();
  A.PrintMatrix();
  invertA.PrintMatrix();
  C = Mat_x_Mat(A,invertA,false, false, 1.0, 1.0);
  C.PrintMatrix();
End
```

Mat_p_Mat

Compound can peform matrix algebraic operations, one of the available algebraic operation is matrix addittion. In order to add two matrices they **must** have the same dimensions.

Syntax:
C=Mat_p_Mat(alpha, A, beta, B);

Where:

- *C:* The resulting matrix.
- *alpha:* The coefficient for matrix A.
- *A:* The left matrix of the addition.
- *beta:* The coefficient for matrix B.
- *B:* The right matrix of the addition.

Example:

```
# ------------------------------------------------------
# This is an example script for matrix addition
# ------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
%Compound
  Variable Dim=3;
  Variable A[Dim][Dim];
  Variable B[Dim][Dim];
  Variable C;
  Variable res=-1;
  for i from 0 to Dim-1 Do
    for j from 0 to Dim-1 Do
      A[i][j] = 1.0;
      B[i][j] = 2.0;
    EndFor
  EndFor
  A.PrintMatrix();
  B.PrintMatrix();
  C = Mat_p_Mat(2.0, A, 3.0, B);
  C.PrintMatrix();
End
```

Mat_x_Mat

Compound can peform matrix algebraic operations, one of the available algebraic operation is matrix multiplication. In general we can multiply each matrix with constants *alpha* and *beta* so that the general multiplication is:

*C=(alpha*A)*(beta*B)

In addition each of matrices A and B are allowed to be transposed.

Syntax:
C=Mat_x_Mat(A, B, [transposeA], [transposeB], [alpha], [beta]);

Where:

- *C:* The resulting matrix.
- *A:* The left matrix of the multiplication.
- *B:* The right matrix of the multiplication.
- *transposeA:* A boolean to state if matrix A should be transposed before the mutliplication (default: False).
- *transposeB:* A boolean to state if matrix B should be transposed before the multiplication (default: False).
- *alpha:* A scalar to multiply matrix A before the mutliplication (default 1.0).
- *beta:* A scalar to mutliply matrix B before the multiplication (default 1.0).

Example:

```
# ------------------------------------------------------
# This is an example script for matrix multiplication
# ------------------------------------------------------
%Compound
  Variable Dim=3;
  Variable A[Dim][Dim];
  Variable invertA;
  Variable C;
  Variable D;
  Variable res=-1;
  for i from 0 to Dim-1 Do
    for j from 0 to Dim-1 Do
      if (i=j) then
        A[i][j] = i+1;
      else
```

(continues on next page)

(continued from previous page)

```
        A[i][j] =  0.0;
      EndIf
    EndFor
  EndFor
  A.invertMatrix(invertA);
  A.PrintMatrix();
  invertA.PrintMatrix();
  C = Mat_x_Mat(A,invertA,false, false, 1.0, 1.0);
  C.PrintMatrix();
  C = Mat_x_Mat(A, invertA, true, true, 1.0, 1.0);
  C.PrintMatrix();
  C = Mat_x_Mat(A, invertA, false, false, 2.0);
  C.PrintMatrix();
  C = Mat_x_Mat(A, invertA, false, false, 2.0, 3.0);
  C.PrintMatrix();
End
```

Mat_x_Scal

Compound can peform matrix algebraic operations, one of the available algebraic operation is multiplication of the elements of a matrix with a scalar. The function returns the multiplied matrix that can be the one that we use as an argument in the parenthesis, meaning it is updated, or a different one.

Syntax:
C=Mat_x_Scal(alpha, A);

Where:

- *C:* The resulting matrix.
- *alpha:* A scalar to multiply the elements of matrix A.
- *A:* The matrix to be mutliplied.

Example:

```
# ------------------------------------------------------
# This is an example script for matrix times scalar
# ------------------------------------------------------
%Compound
  Variable Dim=3;
  Variable A[Dim][Dim];
  Variable alpha=2.0;
  Variable C;
  for i from 0 to Dim-1 Do
    for j from 0 to Dim-1 Do
        A[i][j] = i+j;
    EndFor
  EndFor
  A.PrintMatrix();
  C = Mat_x_Scal(alpha,A);
  A = Mat_x_Scal(alpha,A);
  A.PrintMatrix();
  C.PrintMatrix();
End
```

New_Geom

New_Geom is a platform for geometry manipulation. The basic idea is to have functions that can read a geometry and then produce one or more new geometries with some characteristics that we need. For the moment under the umbrella of *New_Geom* fall 4 different functions, and these are: Displace, Remove_Atom, Remove_Element.

Read

The *Read* command reads a property from the property file.

NOTE This is the old syntax to read the property file and it **will be deprecated** in the next version of ORCA. For the new syntax please use the **ReadProperty** command (see *ReadProperty*)

Syntax

Read myVar = propertyName[stepID]

Where:

myVar: The variable that will be updated

propertyName: The alias for the property we need to read

stepID: The step to which we refer.

Example

```
# ------------------------------------------------------
# This is an example script for readProperty
# ------------------------------------------------------
%Compound
  Variable enDirect=0.0;
  New_Step
    !BP86
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
  alias currStep;
  # First read the energy directly
  Read enDirect = DFT_Total_en[currStep];
  print("DFT Energy : %.12lf\n", enDirect);
End
```

ReadProperty

One of the fundamental features of *Compound* is the ability to easily read ORCA calculated values from the property file.

Syntax:

[res=] readProperty(propertyName=myName, [step=myStep], [filename=myFilename], [baseProperty=true/false])

Where: res: An integer that returns the index of the found property if the property was found in the property file, -1 if the property does not exist. This is not obligatory.

propertyName: A string alias that defines the variable the user wants to read.

step: The step from which we want to read the property. If not given the property file from the last step will be read.

filename: A filename of a property file. If a filename and at the same time a step are provided the program will ignore the step and try to read the property file with the given filename.

NOTE please note that in the end of filename the extension *.property.txt* will be added.

baseProperty: A true/false boolean. The default value is set to false. If the value is set to true then a generic property of the type asked will be read. This means if dipole moment is asked, it will return the last dipole moment, irrelevant if and MP2 or SCF one wad defined.

Example

```
# ------------------------------------------------------
# This is an example script for readProperty
# ------------------------------------------------------
%Compound
  Variable enDirect=0.0;
  Variable enFilename=0.0;
  Variable myProperty="DFT_Total_en";
  Variable basename="compound_example_properertFile_readProperty";
  Variable newBasename ="newFilename";
  Variable res = -1;
  New_Step
    !BP86
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
  # First read the energy directly (returning res)
  res = enDirect.ReadProperty(propertyName=myProperty);
  print("res : %d\n", res);
  # Now read the same energy through filename (not returning res)
  sys_cmd("cp %s_Compound_1.property.txt %s.property.txt", basename, newBasename);
  enFilename.ReadProperty(propertyName=myProperty, filename=newBasename);
  print("Difference between 2 energies : %.12lf\n", enDirect-enFilename);
End
```

Displace

The idea behind "*Displace*" is to have a structure, perform an analytical frequncy calculation on it (currently we do not store numerical frequencies in the property file) and then read the Hessian from this calculation to adjust the geometry based on a normal mode of vibration that we choose. The syntax of this command is:

syntax: New_Geom = (Displace, step, hessian, frequency, scaling)

where:
step is the step from which we choose the original geometry
hessian is the hessian read from a property file
frequency defines which normal mode we will use and
scaling is a factor of how severe we want the displacement to be.

We should note that *Displace* is the only command of the *new_geom* family of commands that will not store a geometry on disk but only internally pass the new geometry to the next calculation. An example of the usage of this command can be found in the script 'iterativeOptimization' that is given with ORCA. The relevant part is as follows: **example:**

```
# Define variables
                Variable MaxNTries    =  25;
                Variable CutOff       = -50;
                Variable displacement = 0.6;
                Variable NNegative =   0;
                Variable freqs[];
                Variable modes[];
                Variable NFreq;
                Variable limit;
```

(continues on next page)

(continued from previous page)

```
                Variable done;
                Variable FinalEnergy;

                # ============================================================
                # Start a for loop over number of tries
                # ============================================================
                For itry From 1 To maxNTries Do

                # ---------------------------------
                # Run a geometry optimization
                # ---------------------------------
                New_Step
                ! tightopt freq verytightscf nopop def2-TZVP xyzfile
                Step_End
                Read freqs  = THERMO_FREQS[itry];
                Read modes  = HESSIAN_MODES[itry];
                Read NFreq  = THERMO_NUM_OF_FREQS[itry];
                limit = NFreq - 1;
                # ---------------------------------
                #  check for sufficeintly negative
                #  frequencies
                # ---------------------------------
                NNegative = 0;
                For ifreq From 0 to limit Do
                if ( freqs[ifreq] < CutOff )  then
                New_Geom = (Displace, itry,  modes,  ifreq, displacement);
```

OpenFile

Compound can write text files on disk. In order to write to a file a filepointer must be created. For this in *Compound* exists the command *OpenFile*.

Syntax:
filePtr = *OpenFile*(*Filename*, "*open mode*");

filePtr is a variable previously declared.

Filename can be a string or a variable of string type and represents the name of the file on disk.

There are two available *opening modes*:

- 'w'. In this mode a new file will be created and the user can write on it. If an old file with the same name exists, it's contents will be deleted.
- 'a'. In this mode if a file already exists, the user will append to what already exists in the file.

Example:

```
%Compound

# ------------------------------------------------------------
#      This is to check all available open and close
#               file options
# ------------------------------------------------------------
Variable myFilename    = "myFile.txt";
Variable file;

# ---------------------------
# First open for writing
# ---------------------------
file = openFile(myFilename, "w");
write2File(file, "This is the first time we write.\n");
```

(continues on next page)

(continued from previous page)

```
closeFile(file);

# --------------------------
# Re-open to append
# --------------------------
file = openFile(myFilename, "a");
write2File(file, "This is the second time we write.\n");
closeFile(file);

End
```

Remove_Atom

Remove_atom removes an atom from a geometry given its index. We should point out that counting of atoms in ORCA starts with 0. After this command is executed it will store on a disk a new geomety in a xyz format where only the atom with the given index will be missing.

Syntax:

New_Geom = (Remove_atom, atomIndex, "filename", stepIndex, [geometry Index]);

where:
atom Index is the number of the atom we want to remove. It can be an integer number or a variable.
filename is the name of the file that we want to use for the new xyz file. It can be a string in quotation marks or a variable already defined before. In the name the xyz extension will be automatically appended. *step Index* the number of the step from which we will get the initial geometry. It has to be an integer number. *geometry Index* In case there are more than one geometries in the corresponding property file we can choose one. If no number is given but default the program will use the last one.

example:

if we use the normal ORCA input file:

```
*xyz 0 1
O         2.220871067         0.026716792         0.000620476
H         2.597492682        -0.411663274         0.766744858
H         2.593135384        -0.449496183        -0.744782026
*

%Compound "removeAtom.cmp"
```

together with the compound file "removeAtom.cmp" :

```
Variable filename  = "newGeom";
Variable atomIndex = 0;

New_Step
  !BP86
Step_End

New_Geom = ( Remove_atom, atomIndex, filename, 1);

end
```

then the xyz file 'newGeom.xyz' will be created that should look like:

```
2

   H         2.5974927       -0.4116633        0.7667449
   H         2.5931354       -0.4494962       -0.7447820
```

where the atom with *atomIndex = 0* meaning the first atom, meaning the oxygen is removed.

Remove_Element

Remove_element is similar to the Remove_atom but instead of using the index of the atom we use its atomic number. Thus the syntax is:

Syntax: New_Geom = (Remove_Element, atomic number, "filename", stepIndex, [geometry Index]);

where:
atomic number is the atomic number of the atom we want to remove. It can be an integer number or a variable. *filename* is the name of the file that we want to use for the new xyz file. It can be a string in quotation marks or a variable already defined before. In the name the xyz extension will be automatically appended. *step Index* the number of the step from which we will get the initial geometry. It has to be an integer number. *geometry Index* In case there are more than one geometries in the corresponding property file we can choose one. If no number is given but default the program will use the last one.

example:

if we use again the input from paragraph *1.1.7.2* but instead of asking the compound file "removeAtom.cmp" we ask for the compound file "removeElement.cmp" that looks like:

```
Variable filename  = "newGeom";

New_Step
  !BP86
Step_End

New_Geom = ( Remove_element, 8, filename, 1);

end
```

then we will get again the same xyz file that was crated in paragraph *Remove_Atom* since the atom with atomic number 8 (meaning the Oxygen) will be removed from the original geometry.

New_Step

New_Step signals the beginning of a new ORCA input.

Syntax:

New_Step
...Normal ORCA input commands
Step_End

There is no restriction in the input of ORCA, except of course that it should not include another Compound block. It is important to remember that a *New_Step* command should always end with a *Step_End* command. Below we show a simple example.

Example:

```
New_Step
! BP86 def2-SVP
Step_End
```

There is only a basic fundamental difference with a normal ORCA input. Inside the *New_Step* block it is not necessary to include a geometry. ORCA will automatically try to read the geometry from the previous calculation. Of course a geometry can be given and then ORCA will use it.

Print

Printing in the ORCA outpug can be customized using the *print* command. The syntax of the *print* command closely follows the corresponding printf command from C/C++. So the usage of the *print* command is:

Syntax:
print(format string, [variables]);

For each variable there can be specifiers and flags for the specifiers. Currently *print* command supports three datatypes namely integers, doubles and strings.

A format specifier follows this prototype: *%[flags][width][.precision]specifier*

where details for the specifiers and flags can be found in table Table 7.33

Table 7.33: compound print Specifiers

Specifier	
s	strings
d	integers
lf	doubles
Flags	
number	width
.number	number of decimal digits
-	left alignement (by default is right)

Example:

```
# ------------------------------------------------------------
# This is to check all available print defintions
# ------------------------------------------------------------
Variable x1    = 2.0;
Variable x2    = {10.0, 20.0, 30.0, 40.0};
Variable x3    = {10,   20,   30,   40};
Variable x4    = {"ten", "twenty", "thirty", "fourty"};
Variable x5 = "test";
Variable index = 2;

print( " ------------------------------------------- \n");
print( " ------ SUMMARY OF PRINT DEFINITIONS --------- \n");
print( " ------------------------------------------- \n");
# ---------- No variables ------------
print( " No variables: \n" );
# ------------- Doubles --------------
print( " -------- Doubles --------------\n");
print( " constant double ( no format)             : %lf\n", 3.5);
```

(continues on next page)

(continued from previous page)

```
print( " constant double (defined width)          : %16lf\n",3.5);
print( " constant double (defined width/accuracy) : %16.8lf\n", 3.5);
print( " variable double (x1)                     : %lf\n", x1);
print( " function double 2*x1*x1                  : %lf\n", 2*x1*x1);
print( " array element double                     : %lf\n", x2[2]);
print( " array element double with var index      : %lf\n", x2[index]);
#print( " array element double with function index : %lf\n", x2[index + 1];
# ------------ Integers --------------
print( " -------- Integers --------------\n");
print( " constant integer ( no format)            : %d\n", 3);
print( " constant integeer (defined width)        : %8d\n",3);
print( " variable integer (index)                 : %d\n", index);
print( " function integer 2*index*index           : %d\n", 2*index*index);
print( " array element                            : %d\n", x3[2]);
print( " array element intteger with var index    : %lf\n",x3[index]);
# ------------  Strings --------------
print( " -------- Strings --------------\n");
print( " constant string ( no format)             : %s\n", "test");
print( " constant string (defined width)          : %8s\n","test");
print( " variable string                          : %s\n", x5);
print( " array element                            : %s\n", x4[2]);
print( " array element intteger with var index    : %s\n",x4[index]);
print(" ------------------------------------------------\n");
print(" ------------------------------------------------\n");
End
```

Read

There are three ways to assign a value to a variable. The first one is through the *"Read"* directive. Read works only for a set of predefined variables that the program stores in the property file, during each ORCA calculation, and can then retrieve from there. A list with the available known variables is given in Table *Variables, known to the compound block, with short explanation*.

Syntax:
Read VariableName = KnownVariable[StepIndex] ;
Example:

```
variable Scale, ZPE, ZPEScaled;

#------------------------------------------------------------
# (Calculation 1)
# the ZPE correction from HF
New_Step
HF 6-31G(d) VeryTightSCF TightOpt Freq
STEP_END
read ZPE   = THERMO_ZPE[1];
```

The *VariableName* should be the name of a variable already defined. The *KnownVariable* should be one of the variables defined in Table *Variables, known to the compound block, with short explanation*. The StepIndex defines for which calculation we should work. It can be either an integer or, if we have already use an *Step_Alias* before, the string of the alias. It should be noted that number counting starts from 1, meaning that the first ORCA calculation, defined through *New_Step* corresponds to number 1.

NOTE: Please do not forget the final *;*.

Read_Geom

Read_Geom will read the geometry from a previous step.

Syntax:
Read_Geom $number$

Here number is the number of the job that we want to read the geometry from. The directive should be positioned before a *New_Step - Step_End* block.

Example:

```
#Compound Job 1
New_Step
!BP86 def2-SVP
Step_End

#Compound Job 2
New_Step
!BP86 def2-SVP opt
Step_End

#Compound Job 3
Read_Geom 1
New_Step
!CCSD  def2-SVP
Step_End

End  #Final End
```

In this case the third calculation, through the *Read_MGeom 1* command, will read the geometry from the first calculation.

ReadMOs

ReadMOs reads the molecular orbitals from a previous step.

Syntax:
ReadMOs(stepNumber);

Where:

- *stepNumber:* is the number of the step from which we want to read the orbitals.

Example:

```
# ------------------------------------------------
# This is an example script for ReadMOs
# ------------------------------------------------
%Compound
  Variable step = 1;
  New_Step
    !BP86
    *xyzfile 0 1 h2o.xyz
  Step_End

  ReadMOs(step);
  New_Step
    !BP86
  Step_End
End
```

S.GetBasename

In *Compound* strings have all the functionality of a normal variable. In addition they have some additional functions that act only on strings. On of these functions is the function *GetBasename*. This function searches the string and if it contains a dot it will return the part of the string before the dot.

Syntax:

result = source.GetBasename();

Where:

result is the returned string.

source is the original string.

NOTE If the original string contains no dot then the result string will be a copy of the source one.

Example:

```
# ------------------------------------------------------
# This is an example script for string related functions:
#    - GetBasename
#    - GetSuffix
#    - GetChar
# ------------------------------------------------------
%Compound
  Variable original = "lala.xyz";
  Variable basename, suffix;
  Variable constructed = "";
  basename = original.GetBasename();
  suffix   = original.GetSuffix();
  for i from 0 to original.stringlength()-1 Do
    write2String(constructed,"%s%s", constructed, original.GetChar(i));
  endfor
  print("Original    : %s\n", original);
  print("Basename    : %s\n", basename);
  print("Sufix       : %s\n", suffix);
  print("Constructed : %s\n", constructed);
End
```

S.GetChar

In *Compound* strings have all the functionality of a normal variable. In addition they have some additional functions that act only on strings. On of these functions is the function *GetChar*. This function searches the string and if it contains a dot it will return the part of the string after the dot.

Syntax:

result = source.GetChar(index);

Where:

result is the returned string.

index is the index of the curracter in the string. Keep in mind that counting starts with **0** and not 1.

source is the original string.

NOTE If the index is larger than the size of the string or negative then the program will exit.

Example:

```
# ------------------------------------------------------
# This is an example script for string related functions:
#    - GetBasename
```

(continues on next page)

(continued from previous page)

```
#    - GetSuffix
#    - GetChar
# ------------------------------------------------------
%Compound
  Variable original = "lala.xyz";
  Variable basename, suffix;
  Variable constructed = "";
  basename = original.GetBasename();
  suffix   = original.GetSuffix();
  for i from 0 to original.stringlength()-1 Do
    write2String(constructed,"%s%s", constructed, original.GetChar(i));
  endfor
  print("Original    : %s\n", original);
  print("Basename    : %s\n", basename);
  print("Sufix       : %s\n", suffix);
  print("Constructed : %s\n", constructed);
End
```

S.GetSuffix

In *Compound* strings have all the functionality of a normal variable. In addition they have some additional functions that act only on strings. On of these functions is the function *GetSuffix*. This function searches the string and if it contains a dot it will return the part of the string after the dot.

Syntax:

result = source.GetSuffix();

Where:

result is the returned string.

source is the original string.

NOTE If the original string contains no dot then the result string will be an empty string.

Example:

```
# ------------------------------------------------------
# This is an example script for string related functions:
#    - GetBasename
#    - GetSuffix
#    - GetChar
# ------------------------------------------------------
%Compound
  Variable original = "lala.xyz";
  Variable basename, suffix;
  Variable constructed = "";
  basename = original.GetBasename();
  suffix   = original.GetSuffix();
  for i from 0 to original.stringlength()-1 Do
    write2String(constructed,"%s%s", constructed, original.GetChar(i));
  endfor
  print("Original    : %s\n", original);
  print("Basename    : %s\n", basename);
  print("Sufix       : %s\n", suffix);
  print("Constructed : %s\n", constructed);
End
```

Step_End

Step_End signals the end of an ORCA Input. It should always be the last directive of an ORCA input inside the compound block that starts with *New_Step* (see paragraph *New_Step*)

Sys_cmd

Sys_cmd will read a system command and execute it.

Syntax:
Sys_cmd *command*

Example:

```
SYS_CMD "orca_mapspc test.out SOCABS -x0700 -x1900 -w0.5 -eV -n10000 "
```

Timer

A *timer* is an object that can keep time for tasks in compound. Before a timer object is used it has to be declared. The declaration of a *timer* is slightly different that the rest of variables, because it has to explicitly declare it's type.

Syntax

timer myTimer;

where

timer is used instead of the normal *variable* command to explicitly set the variable type to *compTimer*.

myTimer is a normal instance of the object.

Example:

```
# ------------------------------------------------
# This is to test timer functions.
# --------------------------------
timer tm;
Variable x = 0.0;
tm.start();
for index from 0 to 100000 Do
  x = x + 0.1;
EndFor
tm.stop();
x = tm.Total();
print( "-----------------------------\n");
print( "   Compound - Timer Results   \n");
print( "-----------------------------\n");
print( " First total time:   %.2lf\n", x);

x = tm.total();
tm.Reset();
tm.Start();
for index from 0 to 200000 Do
  x = x + 0.1;
EndFor
tm.Stop();
x = tm.total();

print( " Second total time:   %.2lf\n", x);
End
```

Below is a lit of functions that work exclusively on *Timer* objects.

- Last (*T.Last*)
- Reset (*T.Reset*)
- Start (*T.Start*)
- Stop (*T.Stop*)
- Total (*T.Total*)

T.Last

Last is a function that works on a *timer* object. It returns, as a real number, the last value of the timer.

Syntax

myTimer.Last();

Where:

myTimer: is the *timer* object initialized before.

Example:

```
# ------------------------------------------------
# This is to test timer functions.
# --------------------------------
timer tm;
Variable x = 0.0;
tm.start();
for index from 0 to 100000 Do
  x = x + 0.1;
EndFor
tm.stop();
x = tm.Total();
print( "------------------------------\n");
print( "   Compound - Timer Results   \n");
print( "------------------------------\n");
print( " First total time:   %.2lf\n", x);

x = tm.total();
tm.Reset();
tm.Start();
for index from 0 to 200000 Do
  x = x + 0.1;
EndFor
tm.Stop();
x = tm.total();

print( " Second total time:   %.2lf\n", x);
End
```

NOTE Before using last the *timer* object must, beside defined, be also initializee, using *Start* (see *T.Start*)

T.Reset

Reset is a function that works on a *timer* object. It resets the *timer* object to its initial state.

Syntax

myTimer.Reset();

Where:

myTimer: is the *timer* object initialized before.

Example:

```
# ------------------------------------------------
# This is to test timer functions.
# --------------------------------
timer tm;
Variable x = 0.0;
tm.start();
for index from 0 to 100000 Do
  x = x + 0.1;
EndFor
tm.stop();
x = tm.Total();
print( "------------------------------\n");
print( "   Compound - Timer Results   \n");
print( "------------------------------\n");
print( " First total time:   %.2lf\n", x);

x = tm.total();
tm.Reset();
tm.Start();
for index from 0 to 200000 Do
  x = x + 0.1;
EndFor
tm.Stop();
x = tm.total();

print( " Second total time:   %.2lf\n", x);
End
```

T.Start

Start is a function that works on a *timer* object. It retunrs the *timer* object to its initial state.

Syntax

myTimer.Start();

Where:

myTimer: is the *timer* object initialized before.

Example:

```
# ------------------------------------------------
# This is to test timer functions.
# --------------------------------
timer tm;
Variable x = 0.0;
tm.start();
for index from 0 to 100000 Do
  x = x + 0.1;
```

(continues on next page)

(continued from previous page)

```
EndFor
tm.stop();
x = tm.Total();
print( "------------------------------\n");
print( "   Compound - Timer Results   \n");
print( "------------------------------\n");
print( " First total time:   %.2lf\n", x);

x = tm.total();
tm.Reset();
tm.Start();
for index from 0 to 200000 Do
  x = x + 0.1;
EndFor
tm.Stop();
x = tm.total();

print( " Second total time:   %.2lf\n", x);
End
```

T.Stop

Stop is a function that works on a *timer* object. It stops the timer from counting.

Syntax

myTimer.Stop();

Where:

myTimer: is the *timer* object initialized before.

Example:

```
# ------------------------------------------------
# This is to test timer functions.
# --------------------------------
timer tm;
Variable x = 0.0;
tm.start();
for index from 0 to 100000 Do
  x = x + 0.1;
EndFor
tm.stop();
x = tm.Total();
print( "------------------------------\n");
print( "   Compound - Timer Results   \n");
print( "------------------------------\n");
print( " First total time:   %.2lf\n", x);

x = tm.total();
tm.Reset();
tm.Start();
for index from 0 to 200000 Do
  x = x + 0.1;
EndFor
tm.Stop();
x = tm.total();

print( " Second total time:   %.2lf\n", x);
End
```

T.Total

Total is a function that works on a *timer* object. It returns a real number with the total time.

Syntax

myTimer.Total();

Where:

myTimer: is the *timer* object initialized before.

Example:

```
# ------------------------------------------------
# This is to test timer functions.
# --------------------------------
timer tm;
Variable x = 0.0;
tm.start();
for index from 0 to 100000 Do
  x = x + 0.1;
EndFor
tm.stop();
x = tm.Total();
print( "------------------------------\n");
print( "   Compound - Timer Results   \n");
print( "------------------------------\n");
print( " First total time:   %.2lf\n", x);

x = tm.total();
tm.Reset();
tm.Start();
for index from 0 to 200000 Do
  x = x + 0.1;
EndFor
tm.Stop();
x = tm.total();

print( " Second total time:   %.2lf\n", x);
End
```

Variables - General

Everything in the *Compound* language is based on variables. Their meaning and usage are similar to those in any programming language: you need to declare a variable and then assign a value to it. Notably, in *Compound*, a variable must be declared before it is assigned a value, following the syntax rules of languages like C. This differs from languages like Python, where you can assign a value to a variable without prior declaration. The only exception to this rule in Compound is the index in a for loop, which does not require prior declaration.

In *Compound* we support the following data types for variables:

- Integer
- Double
- String
- Boolean
- File pointer

In addition to these data types *Compound* supports also variables of type *Geometry* and *Timer* but these are treated separately (see *Geometry* and *Timer*).

For each variable in *Compound* there are 3 major categories of usage:

- The declaration (see *Variables - Declaration*)
- The assignement (see *Variables - Assignment*) and
- Variable functions (see *Variables - Functions*)

Variables - Declaration

There are currently 6 different ways to declare a variable in *Compound*. Their syntax is the following:

Syntax:

A. *Variable name;*

B. *Variable name1, name2;*

C. *Variable name=value;*

D. *Variable name[n];*

E. *Variable name[$n1$][$n2$];*

F. *Variable name={value1, value2, ...}*;

NOTE In previous versions of *Compound* for variables that were matrices but the size was not known one had to declare the variable using the following syntax:

Variable name = [];

This is now changed and the empty brackets are no longer needed, so that this variable can be defined like a normal variable as in case *A*.

Example

```
# ------------------------------------------------------------
# This is to check all available ways of
#   Variable declaration in Compound
# ------------------------------------------------------------
Variable size = 5;
Variable x1;                            #caseA
Variable x2,x3;                         #caseB
Variable x4 = 1.0;                      #caseC
Variable x5 = 0;                        #caseC
Variable x6 = "Test";                   #caseC
Variable x7 = True;                     #caseC
Variable x8 = 2*x4;                     #caseC
Variable x9 = x6;                       #caseC
Variable x10 = x7;                      #caseC
Variable x11;                           #caseA
Variable x12[3];                        #caseD
Variable x12b[size-2];                  #caseD
Variable x12c[size][size-1];            #caseE
x12b[1] = 4.0;
x12c[2][2] = 7.0;
Variable x13[3][3];                     #caseE
Variable x14 = {2.0, 4*x4, 2, "lala"};  #caseF
Variable x15 = {0, 1, 2, 3, 4};
Variable x15b = {0, 1, 2, 3, 4};
Variable x15c[x15[x15b[2]-1]+2];
Variable x16, x17=x10, x18;
Variable x19=2.0, x20, x21[2], x23[size][size];
print( " ------------------------------------------- \n");
print( " ---------- SUMMARY OF DEFINITIONS ---------- \n");
print( " ------------------------------------------- \n");
print( " x4 (1.0)          :  %.2lf\n", x4);
print( " x5 (0)            :  %.2d\n",  x5);
```

(continues on next page)

(continued from previous page)

```
print( " x6 (\"Test\")       :  %s\n",    x6);
if (x7) then
  print(" x7 (True)         :  TRUE\n");
else
  print(" x7 (True)         :  FALSE\n");
endIf
print(" x8 (2*x4)         :  %.2lf\n", x8);
print(" x9 (x6)           :  %s\n", x9);
print(" x12b[1] (4.0)     :  %lf\n", x12b[1]);
print(" x12c[2][2] (7.0)  :  %lf\n", x12c[2][2]);
print(" Variable x14 = {2.0, 4*x4, 2, \"lala\"};\n");
print(" x14[0]            :  %lf\n", x14[0]);
print(" x14[1]            :  %lf\n", x14[1]);
print(" x14[2]            :  %d\n",  x14[2]);
print(" x14[3]            :  %s\n",  x14[3]);
print(" Variable x15  = {0, 1, 2, 3, 4};\n");
print(" Variable x15b = {0, 1, 2, 3, 4};\n");
print(" Variable x15c[x15[x15b[2]-1]+2];\n");
print(" x15c.GetSize()    :  %d\n",  x15c.GetSize());
print(" ------------------------------------------------\n");
print(" ------------------------------------------------\n");
End
```

Some comments for the different cases of variable declaration.
Case A is the simplest one were we just declare the name of a variable.

Case B is similar to *Case A* but here more than one variables are declared simultaneously.

In **Case C** we combine the variable declaration with the assignment of a value to the variable. It worth noting that in this case *Compound* automatically deducts the type of the variable based on the given value.

Case D declares a 1-Dimensional array of a defined size.

Case E declares a 2-Dimensional array of defined size. For this case, and also accordingly for *Case E*, one can use previously defined integer variables instead of numbers.

Case F defines an array based on a list of given values. The array will automatically define it's size based on the size of the list. The values in the list do not have to be all of the same type.

NOTE In the past for *Case F* empty brackets were needed after the name of the variable. This is no longer necessary.

NOTE It is important not to forget the final **;** symbol in the end of each declaration because the result of omitting it is undefined.

Variables - Assignment

Assigning a value to a variable has a rather straightforward syntax.

Syntax:
VariableName = CustomFunction;

Where:

VariableName is a variable already declared.

CustomFunction a mathematical expression.

Example:

```
# ---------------------------------------------------------------
# This is to check all available ways of variable assignement
#   (It does not take care of 'with' we will have a separate
#         file for this)
# ---------------------------------------------------------------
```

(continues on next page)

(continued from previous page)

```
# ----------------------------------
# Some necessary initial declarations
# ----------------------------------
Variable x1, x2, x3, x4;
Variable y1, y2, y3, y4;
Variable x5[4];
Variable y5[4];
Variable x6[3][3];
Variable y6[3][3];

# ----------------------------------
# Now the assignements
# ----------------------------------
#Scalars doubles
x1 = 1.0;
y1 = 2*x1;
#Scalars integers
x2 = 1;
y2 = 2*x2;
#Scalars strings
x3 = "test";
y3 = x3;
#Scalars bools
x4 = True;
y4 = x4;
#1D Arrays
x5[0] = 2.0;
y5[0] = x5[0];
x5[1] = 1;
y5[1] = 2*x5[1];
x5[2] = "test";
y5[2] = x5[2];
x5[3] = True;
y5[3] = x5[3];
#2D Arrays
x6[0][0] = 1.0;
y6[0][0] = 2*x6[0][0];
print( " ------------------------------------------- \n");
print( " ---------- SUMMARY OF ASSIGNMENTS ---------- \n");
print( " ------------------------------------------- \n");
print( " ---------------- Scalars -------------------- \n");
print( " x1         :  %.2lf\n", x1);
print( " y1         :  %.2lf\n", y1);
print( " x2         :  %d\n", x2);
print( " y2         :  %d\n", y2);
print( " x3         :  %s\n", x3);
print( " y3         :  %s\n", y3);
print( " ---------------- 1D - Arrays----------------- \n");
print( " x5[0]      :  %.2lf\n", x5[0]);
print( " y5[0]      :  %.2lf\n", y5[0]);
print( " x5[1]      :  %d\n", x5[1]);
print( " y5[1]      :  %d\n", y5[1]);
print( " x5[2]      :  %s\n", x5[2]);
print( " y5[2]      :  %s\n", y5[2]);
print( " ---------------- 2D - Arrays----------------- \n");
print( " x6[0][0]   :  %lf\n", x6[0][0]);
print( " y6[0][0]   :  %lf\n", y6[0][0]);
print(" ---------------------------------------------\n");
print(" ---------------------------------------------\n");
End
```

NOTE It is important to remember to finish the variable assignment using the **';'** symbol.

Variables - Functions

Variables in *Compound* have a small number of functions that can help exctract information about them. In the current version of *Compound* these functions are the following:

Syntax:
VariableName.Function();

where *VariableName* is a variable that is already declared. Then the function will return a value that depends on the *Function* that we used.

Currenlty *Compound* supports the following functions:

- GetBool() *V.GetBool()*
- GetDim1() *V.GetDim1()*
- GetDim2() *V.GetDim2()*
- GetDouble() *V.GetDouble()*
- GetInteger() *V.GetInteger()*
- GetSize() *V.GetSize()*
- GetString() *V.GetString()*
- PrintMatrix() *V.PrintMatrix()*

Example:

```
# ------------------------------------------------------------
# This is to check all available ways of variable assignement
#   (It does not take care of 'with' we will have a separate
#         file for this)
# ------------------------------------------------------------

# ----------------------------------
# Some necessary initial declarations
# ----------------------------------
Variable x1, x2, x3, x4;
Variable y1, y2, y3, y4;
Variable x5[4];
Variable y5[4];
Variable x6[3][3];
Variable y6[3][3];

# ----------------------------------
# Now the assignements
# ----------------------------------
#Scalars doubles
x1 = 1.0;
y1 = 2*x1;
#Scalars integers
x2 = 1;
y2 = 2*x2;
#Scalars strings
x3 = "test";
y3 = x3;
#Scalars bools
x4 = True;
y4 = x4;
#1D Arrays
x5[0] = 2.0;
```

(continues on next page)

(continued from previous page)

```
y5[0] = x5[0];
x5[1] = 1;
y5[1] = 2*x5[1];
x5[2] = "test";
y5[2] = x5[2];
x5[3] = True;
y5[3] = x5[3];
#2D Arrays
x6[0][0] = 1.0;
y6[0][0] = 2*x6[0][0];
print( " ---------------------------------------------- \n");
print( " ---------- SUMMARY OF ASSIGNMENTS ---------- \n");
print( " ---------------------------------------------- \n");
print( " ---------------- Scalars -------------------- \n");
print( " x1         :  %.2lf\n", x1);
print( " y1         :  %.2lf\n", y1);
print( " x2         :  %d\n", x2);
print( " y2         :  %d\n", y2);
print( " x3         :  %s\n", x3);
print( " y3         :  %s\n", y3);
print( " ---------------- 1D - Arrays----------------- \n");
print( " x5[0]      :  %.2lf\n", x5[0]);
print( " y5[0]      :  %.2lf\n", y5[0]);
print( " x5[1]      :  %d\n", x5[1]);
print( " y5[1]      :  %d\n", y5[1]);
print( " x5[2]      :  %s\n", x5[2]);
print( " y5[2]      :  %s\n", y5[2]);
print( " ---------------- 2D - Arrays----------------- \n");
print( " x6[0][0]   :  %lf\n", x6[0][0]);
print( " y6[0][0]   :  %lf\n", y6[0][0]);
print(" ----------------------------------------------\n");
print(" ----------------------------------------------\n");
End
```

V.GetBool()

This function will return a boolean value in case the variable is boolean or integer. For integers it will return *false* for 0 and *true* for all other integer values. In all other cases the program will crash providing a relevant message.

Syntax:

myVar.GetBool();

where:

myVar is an already initialized variable.

Example

```
# ------------------------------------------------------
# This is an example script for
#  Variable functions
# ------------------------------------------------------
%Compound
  Variable double=1.0;
  Variable integer=2;
  Variable iToBool = integer.GetBool();
  Variable boolean=false;

  print("----------------------------------------\n");
  print("     Results for translation functions \n");
```

(continues on next page)

(continued from previous page)

```
  print("Double  to integer : %d    (it should print 1)\n", double.GetInteger());
  print("Integer to double  : %.2lf (it should print 2.00)\n", integer.GetDouble());
  print("Boolean to string  : %s    (it should print FALSE)\n", boolean.GetString());
  print("Integer to boolean : %s    (it should print TRUE)\n", iToBool.GetString());
  print("Double  to string  : %s    (it should print 1.0000000000000000000e+00)\n", double.
↪GetString());
  print("Integer to string  : %s    (it should print 2)\n", integer.GetString());
End
```

V.GetDim1()

This function works on a variable. It will return the size of the first dimension of an array. If the variable is a scalar it will return 1.

Syntax:

myVar.GetDim1();

where:

myVar is an already initialized variable.

Example

```
# ------------------------------------------------------
# This is an example script for
#  Variable functions
# ------------------------------------------------------
%Compound
  Variable dim1, dim2, size;
  Variable A;
  Variable B[3];
  Variable C[3][2];

  print("----------------------------------------\n");
  print("          Results for scalar  \n");
  print("Dim1 : %d (it should print 1)\n", A.GetDim1());
  print("Dim2 : %d (it should print 1)\n", A.GetDim2());
  print("Size : %d (it should print 1)\n", A.GetSize());
  print("----------------------------------------\n");
  print("          Results for 1D-Array  \n");
  print("Dim1 : %d (it should print 3)\n", B.GetDim1());
  print("Dim2 : %d (it should print 1)\n", B.GetDim2());
  print("Size : %d (it should print 3)\n", B.GetSize());
  print("----------------------------------------\n");
  print("          Results for 2D-Array  \n");
  print("Dim1 : %d (it should print 3)\n", C.GetDim1());
  print("Dim2 : %d (it should print 2)\n", C.GetDim2());
  print("Size : %d (it should print 6)\n", C.GetSize());
End
```

V.GetDim2()

This function works on a variable. It will return the size of the second dimension of an array. If the variable is a scalar or a 1-Dimensional array it will return 1.

Syntax:

myVar.GetDim2();

where:

myVar is an already initialized variable.

Example

```
# ------------------------------------------------------
# This is an example script for
#  Variable functions
# ------------------------------------------------------
%Compound
  Variable dim1, dim2, size;
  Variable A;
  Variable B[3];
  Variable C[3][2];

  print("----------------------------------------\n");
  print("          Results for scalar  \n");
  print("Dim1 : %d (it should print 1)\n", A.GetDim1());
  print("Dim2 : %d (it should print 1)\n", A.GetDim2());
  print("Size : %d (it should print 1)\n", A.GetSize());
  print("----------------------------------------\n");
  print("          Results for 1D-Array  \n");
  print("Dim1 : %d (it should print 3)\n", B.GetDim1());
  print("Dim2 : %d (it should print 1)\n", B.GetDim2());
  print("Size : %d (it should print 3)\n", B.GetSize());
  print("----------------------------------------\n");
  print("          Results for 2D-Array  \n");
  print("Dim1 : %d (it should print 3)\n", C.GetDim1());
  print("Dim2 : %d (it should print 2)\n", C.GetDim2());
  print("Size : %d (it should print 6)\n", C.GetSize());
End
```

V.GetDouble()

This function works on a variable. It will return a double value in case the variable is integer or double. In all other cases the program will crash providing a relevant message.

Syntax:

myVar.GetDouble();

where:

myVar is an already initialized variable.

Example

```
# ------------------------------------------------------
# This is an example script for
#  Variable functions
# ------------------------------------------------------
%Compound
  Variable double=1.0;
  Variable integer=2;
```

(continues on next page)

(continued from previous page)

```
  Variable iToBool = integer.GetBool();
  Variable boolean=false;

  print("----------------------------------------\n");
  print("     Results for translation functions \n");
  print("Double  to integer : %d    (it should print 1)\n", double.GetInteger());
  print("Integer to double  : %.2lf (it should print 2.00)\n", integer.GetDouble());
  print("Boolean to string  : %s    (it should print FALSE)\n", boolean.GetString());
  print("Integer to boolean : %s    (it should print TRUE)\n", iToBool.GetString());
  print("Double  to string  : %s    (it should print 1.0000000000000000000e+00)\n", double.
↪GetString());
  print("Integer to string  : %s    (it should print 2)\n", integer.GetString());
End
```

V.GetInteger()

This function works on a variable. It will return an integer value in case the variable is integer or double. In all other cases the program will crash providing a relevant message.

Syntax:

myVar.GetInteger();

where:

myVar is an already initialized variable.

Example

```
# ------------------------------------------------------
# This is an example script for
#  Variable functions
# ------------------------------------------------------
%Compound
  Variable double=1.0;
  Variable integer=2;
  Variable iToBool = integer.GetBool();
  Variable boolean=false;

  print("----------------------------------------\n");
  print("     Results for translation functions \n");
  print("Double  to integer : %d    (it should print 1)\n", double.GetInteger());
  print("Integer to double  : %.2lf (it should print 2.00)\n", integer.GetDouble());
  print("Boolean to string  : %s    (it should print FALSE)\n", boolean.GetString());
  print("Integer to boolean : %s    (it should print TRUE)\n", iToBool.GetString());
  print("Double  to string  : %s    (it should print 1.0000000000000000000e+00)\n", double.
↪GetString());
  print("Integer to string  : %s    (it should print 2)\n", integer.GetString());
End
```

V.GetSize()

This function works on a variable. If the variable is a scalar it will return 1. If the variable is a 1-Dimensional array it will return the size of the array which is the same with the *GetDim1()* . If the variable is a 2-Dimensional array it will return the results Dim1*Dim2.

Syntax:

myVar.GetSize();

where:

myVar is an already initialized variable.

Example

```
# ------------------------------------------------------
# This is an example script for
#  Variable functions
# ------------------------------------------------------
%Compound
  Variable dim1, dim2, size;
  Variable A;
  Variable B[3];
  Variable C[3][2];

  print("----------------------------------------\n");
  print("          Results for scalar  \n");
  print("Dim1 : %d (it should print 1)\n", A.GetDim1());
  print("Dim2 : %d (it should print 1)\n", A.GetDim2());
  print("Size : %d (it should print 1)\n", A.GetSize());
  print("----------------------------------------\n");
  print("          Results for 1D-Array  \n");
  print("Dim1 : %d (it should print 3)\n", B.GetDim1());
  print("Dim2 : %d (it should print 1)\n", B.GetDim2());
  print("Size : %d (it should print 3)\n", B.GetSize());
  print("----------------------------------------\n");
  print("          Results for 2D-Array  \n");
  print("Dim1 : %d (it should print 3)\n", C.GetDim1());
  print("Dim2 : %d (it should print 2)\n", C.GetDim2());
  print("Size : %d (it should print 6)\n", C.GetSize());
End
```

V.GetString()

This function works on a variable. It will return a string of the value of the variable. It works for doubles, integers and booleans.

Syntax:

myVar.GetString();

where:

myVar is an already initialized variable.

Example

```
# ------------------------------------------------------
# This is an example script for
#  Variable functions
# ------------------------------------------------------
%Compound
  Variable double=1.0;
```

(continues on next page)

(continued from previous page)

```
  Variable integer=2;
  Variable iToBool = integer.GetBool();
  Variable boolean=false;

  print("----------------------------------------\n");
  print("     Results for translation functions \n");
  print("Double  to integer : %d    (it should print 1)\n", double.GetInteger());
  print("Integer to double  : %.2lf (it should print 2.00)\n", integer.GetDouble());
  print("Boolean to string  : %s    (it should print FALSE)\n", boolean.GetString());
  print("Integer to boolean : %s    (it should print TRUE)\n", iToBool.GetString());
  print("Double  to string  : %s    (it should print 1.0000000000000000000e+00)\n", double.
↪GetString());
  print("Integer to string  : %s    (it should print 2)\n", integer.GetString());
End
```

V.PrintMatrix()

This function works on variables. It will print print an array on a format with 8 columns.

Syntax: *myVar.PrintMatrix([NCols=numOfColumns]);*

where:

myVar is an already initialized variable.

numOfColumns is the desired number of columns for the printing. This is not obligatory and if not used then by default ORCA will print using **4** columns.

Example

Example:

```
# ------------------------------------------------
#  A script to check PrintMatrix
# ------------------------------------------------
%Compound
  Variable Dim1 = 5;
  Variable Dim2 = 16;
  Variable x[Dim1][Dim2];
  for i from 0 to Dim1-1 Do
    for j from 0 to Dim2-1 Do
      x[i][j] = i+j;
    EndFor;
  EndFor;

  x.PrintMatrix();        # This should print with 4 columns
  x.PrintMatrix(NCols=8); # This should print with 8 columns
EndRun
```

NOTE In case of scalars it will only print the header without any values.

NOTE It only works for arrays of type *'double'* or type *'integer'*. With all variables of other types the program will exit providing an error message.

With

The purpose of the "with" command is to add the ability to call compound while adjusting some of the variables that are already defined in the compound file. This means that if there is a variable defined in the compound file and a value is assigned to it, we can during the call change the assigned value of this variable.

One can pass numbers, string or boolean variables.

It should be noted that it is not possible to call array variables this way. Beside this restriction, the syntax of the variable assignment in the case of with is the same with the variable assignment in a normal *Compound* script.

An important note here is that in case we use the *With* command the *%Compound* block should end with an *'End'* even if we call a *Compound* script file.
Syntax:
% compound *"filename"*
With
var1 = *val1*;
var2 = *val2*;
End
Example:

```
# ---------------------------------------------------------------
# This is to check all available ways of variable assignement
#   in combination with the 'with' calls.
# ---------------------------------------------------------------

# ---------------------------------
# Some necessary initial definitions
# ---------------------------------
Variable x1, x2, x3, x4;

# -----------------------------------
# Now the assignments
# -----------------------------------
#Scalars doubles
x1 = 1.0;
#Scalars integers
x2 = 1;
#Scalars strings
x3 = "test";
#Scalars bools
x4 = True;
print( " ---------------------------------------------- \n");
print( " ------- SUMMARY OF WITH ASSIGNMENTS --------- \n");
print( " ---------------------------------------------- \n");
print( " The calling input:\n");
print("%Compound \"0975.cmp\"\n");
print("  with\n");
print("    x1       = 3.0;\n");
print("    x2       = 2;\n");
print("    x3       = \"with\";\n");
print("    x4       = False;\n");
print("end\n");
print( " ---------------- Scalars -------------------- \n");
print( " x1 (1.0)      :  %.2lf\n", x1);
print( " x2 (1)        :  %d\n", x2);
print( " x3 (\"test\")   :  %s\n", x3);
print( " x4 (True)     :  %s\n", x4.GetString());
#print( " x6      :  %s\n",    x6);
#if (x4) then
#  print(" x4      : TRUE\n");
#else
```

(continues on next page)

(continued from previous page)

```
#  print(" x4   : FALSE\n");
#endIfo
End
```

Write2File

With the *Print* command (see *Print*) one can write in the ORCA output. Nevertheless it might be that one would prefer to write to a different file. In *Compound* one can achieve this using the *write2File* command. The syntax follows closely the syntax of 'fprintf' command of the programming language C. The arguments definition and the syntax is identical with the syntax of the *Compound 'Print'* command with the addition that one should define a file object to send the printing.

Syntax:
Write2File(file variable, format string, [variables]);

Where:

file variable: is a predefined variable corresponding to an already open, through the *OpenFile* command, file.

format string and *variables* follow exactly the syntax of the *Print* command, so for more details please refere to section *Print*.

NOTE Please remember once everything is writen to the file to close the file, using the *CloseFile* command (see *CloseFile*).

Example:

```
%Compound
  # ------------------------------------------------------------
  #      This is to check all available write2String and
  #                 write2File options
  # ------------------------------------------------------------
  Variable xS    = "test_";
  Variable xI    = 1;
  Variable final;
  Variable fp;
  Variable myFilename = "0955.txt";


  #Create also a file object
  fp = OpenFile(myFilename, "w");
  write2String(final, "   ----- Test ----- \n");
  write2File(fp, "%s", final);
  CloseFile(fp);

  print( " ---------------------------------------------- \n");
  print( " ------ SUMMARY OF WRITE2STRING AND  --------- \n");
  print( " ------          WRITE2FILE          --------- \n");
  print( " ---------------------------------------------- \n");
  write2String(final, "%s", "constant" );
  print( " Final      : %s\n", final);
  write2String(final,"%s", "constant" );      #No space before the quotation marks
  print( " Final      : %s\n", final);
  write2String(final,  "%s", "constant" );    #More than one spaces before
  print( " Final      : %s\n", final);
  write2String(final,"    %s", "constant" );  #No spaces before but  more afterwards
  print( " Final      : %s\n", final);
  write2String(final,   "  %s", "constant" ); #More spaces before and more afterwards
  print( " Final      : %s\n", final);
  write2String(final, "%s", xS);
  print( " Final      : %s\n", final);
```

(continues on next page)

(continued from previous page)

```
  write2String(final, "%s_%d", xS, xI);
  print( " Final      : %s_%d\n", final, xI);
  write2String(final, "%s_%d", xS,2*xI+1);
  print( " Final      : %s\n", final);
End
```

Write2String

In case one needs to construct a string using some variables, *Compound* provides the *Write2String* command. The syntax of the command is identical with the *Write2File* (see *Write2File*) command with the only exception that instead of a file we should provide the name of a variable that is already declared in the file. The syntax of the format and the variables used is identical with the *Print* command (please refer to *Print*.)

Syntax:

Write2String(variable, format string, [variables]);

where:

variable: is the name of a variable that should already be declared.

format string and *variables* follow exactly the syntax of the *Print* command, so for more details please refer to section *Print*.

Example:

```
%Compound
  # ------------------------------------------------------------
  #      This is to check all available write2String and
  #                write2File options
  # ------------------------------------------------------------
  Variable xS    = "test_";
  Variable xI    = 1;
  Variable final;
  Variable fp;
  Variable myFilename = "0955.txt";


  #Create also a file object
  fp = OpenFile(myFilename, "w");
  write2String(final, "   ----- Test ----- \n");
  write2File(fp, "%s", final);
  CloseFile(fp);

  print( " ------------------------------------------- \n");
  print( " ------ SUMMARY OF WRITE2STRING AND  --------- \n");
  print( " ------          WRITE2FILE          --------- \n");
  print( " ------------------------------------------- \n");
  write2String(final, "%s", "constant" );
  print( " Final      : %s\n", final);
  write2String(final,"%s", "constant" );      #No space before the quotation marks
  print( " Final      : %s\n", final);
  write2String(final,  "%s", "constant" );    #More than one spaces before
  print( " Final      : %s\n", final);
  write2String(final,"    %s", "constant" );  #No spaces before but  more afterwards
  print( " Final      : %s\n", final);
  write2String(final,   "  %s", "constant" ); #More spaces before and more afterwards
  print( " Final      : %s\n", final);
  write2String(final, "%s", xS);
  print( " Final      : %s\n", final);
  write2String(final, "%s_%d", xS, xI);
```

(continues on next page)

(continued from previous page)

```
  print( " Final      : %s_%d\n", final, xI);
  write2String(final, "%s_%d", xS,2*xI+1);
  print( " Final      : %s\n", final);
End
```

7.56.2 List of known Properties

The name and a sort explanation of all the known variables that can be automatically recovered, from the property file, are given in the next table

Table 7.34: Variables, known to the compound block, with short e

==	==========
==	==========AUTOC
==	==========
++	++++++++++Energies++++
AUTOCI_REF_ENERGY	AutoCI Reference Energy
AUTOCI_CORR_ENERGY	AutoCI Correlatioin Energy
AUTOCI_TOTAL_ENERGY	AutoCI Total Energy
++	++++++++++ ENERGY Gradie
AUTOCI_NUCLEAR_GRADIENT	AutoCI Energy nuclear gradient
AUTOCI_NUCLEAR_GRADIENT_NORM	AutoCI Norm of the nuclear gradient
AUTOCI_NUCLEAR_GRADIENT_ATOM_NUMBERS	AutoCI The atomic numbers of the atoms in the gradient
++	++++++++++ Electric Properties (Dipole mon
AUTOCI_DIPOLE_MAGNITUDE	AutoCI The value of the dipole moment
AUTOCI_DIPOLE_ELEC_CONTRIB	AutoCI The electronic contribution to the dipole moment
AUTOCI_DIPOLE_NUC_CONTRIB	AutoCI The nuclear contribution to the dipole moment
AUTOCI_DIPOLE_TOTAL	AutoCI The total dipole moment
SCF_ENERGY	SCF Energy
++	++++++++++ Electric Properties (Polarizability
AUTOCI_POLAR_ISOTROPIC	AutoCI The polarizability isotropic value
AUTOCI_POLAR_RAW	AutoCI The raw polarizability tensor
AUTOCI_POLAR_DIAG_TENSOR	AutoCI The polarizability diagonalized tensor
AUTOCI_POLAR_ORIENTATION	AutoCI The polarizability orientation (eigenvectors)
++	++++++++++ Electric Properties (Quadrupole mom
AUTOCI_QUADRUPOLE_MOMENT_ISOTROPIC	AutoCI The quadrupole moment isotropic value
AUTOCI_QUADRUPOLE_MOMENT_DIAG_TENSOR	AutoCI The quadrupole moment diagonalized tensor
AUTOCI_QUADRUPOLE_MOMENT_ELEC_CONTRIB	AutoCI The elctronic contribution to the quadrupole moment tensor
AUTOCI_QUADRUPOLE_MOMENT_NUC_CONTRIB	AutoCI The nuclear contribution to the quadrupole moment tensor
AUTOCI_QUADRUPOLE_MOMENT_TOTAL	AutoCI The total quadrupole moment
++	++++++++++ Magnetic Properties (D Tenso
AUTOCI_D_TENSOR_EIGENVALUES	AutoCI The D Tensor eigenvalues
AUTOCI_D_TENSOR_EIGENVECTORS	AutoCI The D Tensor eigenvectors
AUTOCI_D_TENSOR_RAW	AutoCI The Raw D Tensor
AUTOCI_D_TENSOR_D	AutoCI The final D value for the D Tensor
AUTOCI_D_TENSOR_E	AutoCI The final E value for the D Tensor
AUTOCI_D_TENSOR_MULTIPLICITY	AutoCI The spin-multiplicity used for the D Tensor calculation
++	++++++++++ Magnetic Properties (G Tenso
AUTOCI_G_TENSOR_RAW	AutoCI The Raw G Tensor
AUTOCI_G_TENSOR_ELEC	AutoCI The Electronic part of the G Tensor
AUTOCI_G_TENSOR_TOT	AutoCI The Total G Tensor
AUTOCI_G_TENSOR_ISO	AutoCI The isotropic g value
AUTOCI_G_TENSOR_ORIENTATION	AutoCI The G Tensor orientation (eigenvectors)
SCF_ENERGY	SCF Energy

Table 7.34 – continued from previous page

VDW_CORRECTION	van der Waals correction
SCF Electric properties	
SCF_DIPOLE_MAGNITUDE_DEBYE	SCF dipole moment (debye)
SCF_DIPOLE_ELEC_CONTRIB	SCF Electronic contribution to dipole moment
SCF_DIPOLE_NUC_CONTRIB	SCF Nuclear contribution to dipole moment
SCF_DIPOLE_TOTAL	SCF Total dipole moment
SCF_QUADRUPOLE_ISOTROPIC	SCF isotropic quadrupole moment
SCF_QUADRUPOLE_DIAG_TENSOR	SCF quadrupole moment diagonalised tensor
SCF_QUADRUPOLE_ELEC_CONTRIB	SCF electronic contribution to the quadrupole moment
SCF_QUADRUPOLE_NUC_CONTRIB	SCF nuclear contribution to the quadrupole moment
SCF_QUADRUPOLE_TOTAL	SCF total quadrupole moment
SCF_POLAR_ISOTROPIC	SCF isotropic polarizability
SCF_POLAR_RAW	SCF polarizability raw tensor
SCF_POLAR_DIAG_TENSOR	SCF diagonaised polarizability tensor
DFT	
DFT_NUM_OF_ALPHA_EL	Number of alpha electrons
DFT_NUM_OF_BETA_EL	Number of beta electrons
DFT_NUM_OF_TOTAL_EL	Total number of electrons
DFT_TOTAL_EN	DFT Total energy
DFT_EXCHANGE_EN	DFT Exchange energy
DFT_CORR_EN	DFT Correlation Energy
DFT_XC_EN	DFT Exchange-Correlation Energy
DFT_NON_LOC_EN	DFT Non-Local correlation
DFT_EMBED_CORR	DFT Embedding correction
MP2	
MP2_REF_ENERGY	Reference SCF Energy
MP2_CORR_ENERGY	MP2 Correlation energy
MP2_TOTAL_ENERGY	Total Energy (SCF + MP2)
MP2 Electric properties	
MP2_DIPOLE_MAGNITUDE_DEBYE	MP2 dipole moment (debye)
MP2_DIPOLE_ELEC_CONTRIB	MP2 Electronic contribution to dipole moment
MP2_DIPOLE_NUC_CONTRIB	MP2 Nuclear contribution to dipole moment
MP2_DIPOLE_TOTAL	MP2 Total dipole moment
MP2_QUADRUPOLE_ISOTROPIC	MP2 isotropic quadrupole moment
MP2_QUADRUPOLE_DIAG_TENSOR	MP2 quadrupole moment diagonalised tensor
MP2_QUADRUPOLE_ELEC_CONTRIB	MP2 electronic contribution to the quadrupole moment
MP2_QUADRUPOLE_NUC_CONTRIB	MP2 nuclear contribution to the quadrupole moment
MP2_QUADRUPOLE_TOTAL	MP2 total quadrupole moment
MP2_POLAR_ISOTROPIC	MP2 isotropic polarizability
MP2_POLAR_RAW	MP2 polarizability raw tensor
MP2_POLAR_DIAG_TENSOR	MP2 diagonaised polarizability tensor
MDCI	
MDCI_REF_ENERGY	Reference SCF Energy
MDCI_CORR_ENERGY	Total Correlation Energy
MDCI_TOTAL_ENERGY	Total Energy (SCF + Correlation)

Table 7.34 – continued from previous page

MDCI_ALPHA_ALPHA_CORR_ENERGY	Correlation energy from $\alpha\alpha$ electron pairs
MDCI_BETA_BETA_CORR_ENERGY	Correlation energy from $\beta\beta$ electron pairs
MDCI_ALPHA_BETA_CORR_ENERGY	Correlation energy from $\alpha\beta$ electron pairs
MDCI_DSINGLET_CORR_ENERGY	Correlation energy from singlet electron pairs (only for closed-shell
MDCI_DTRIPLET_CORR_ENERGY	Correlation energy from triplet electron pairs (only for closed-shell)
MDCI_SSINGLET_CORR_ENERGY	Correlation energy from singlet electron pairs (only for closed-shell
MDCI_STRIPLET_CORR_ENERGY	Correlation energy from triplet electron pairs (only for closed-shell)
MDCI_TRIPLES_ENERGY	Perturbative triples correlation energy
MDCI_ALL_ELECTRONS	Total number of electrons
MDCI_CORR_ELECTRONS	Number of correlated electrons
MDCI_CORR_ALPHA_ELECTRONS	Number of correlated α electrons
MDCI_CORR_BETA_ELECTRONS	Number of correlated β electrons
MDCI Electric properties	
MDCI_DIPOLE_MAGNITUDE_DEBYE	MDCI dipole moment (debye)
MDCI_DIPOLE_ELEC_CONTRIB	MDCI Electronic contribution to dipole moment
MDCI_DIPOLE_NUC_CONTRIB	MDCI Nuclear contribution to dipole moment
MDCI_DIPOLE_TOTAL	MDCI Total dipole moment
MDCI_QUADRUPOLE_ISOTROPIC	MDCI isotropic quadrupole moment
MDCI_QUADRUPOLE_DIAG_TENSOR	MDCI quadrupole moment diagonalised tensor
MDCI_QUADRUPOLE_ELEC_CONTRIB	MDCI electronic contribution to the quadrupole moment
MDCI_QUADRUPOLE_NUC_CONTRIB	MDCI nuclear contribution to the quadrupole moment
MDCI_QUADRUPOLE_TOTAL	MDCI total quadrupole moment
MDCI_POLAR_ISOTROPIC	MDCI isotropic polarizability
MDCI_POLAR_RAW	MDCI polarizability raw tensor
MDCI_POLAR_DIAG_TENSOR	MDCI diagonaised polarizability tensor
CASSCF	
CASSCF_NUM_OF_MULTS	The number of CASSCF spin multiplicities
CASSCF_NUM_OF_IRREPS	The number of CASSCF irreps
CASSCF_FINAL_ENERGY	The CASSCF final energy
PT2_NUM_OF_MULTS	The CASPT2 spin multiplicities
PT2_NUM_OF_IRREPS	The number of CASPT2 irreps
PT2_FINAL_ENERGY	The CASPT2 Energy
DCDCAS_NUM_OF_MULTS	The number of DCDCAS spin multiplicities
DCDCAS_NUM_OF_IRREPS	The number of DCDCAS irreps
DCDCAS_FINAL_ENERGY	The DCDCAS Energy
CASSCF_ABS_SPECTRUM	The CASSCF Absorption spectrum
CASSCF_ABS_SPECTRUM_INFO	Information about the excitations of the CASSCF spectrum
CASSCF_ABS_SPECTRUM_NROOTS	The number of Roots
CASSCF_CD_SPECTRUM	The CASSCF CD spectrum
CASSCF_CD_SPECTRUM_INFO	Information about the excitations of the CASSCF CD spectrum
CASSCF_CD_SPECTRUM_NROOTS	The number or roots
CASPT2_ABS_SPECTRUM	The CASPT2 Absorption spectrum
CASPT2_ABS_SPECTRUM_INFO	Information about the excitations of the CASPT2 spectrum
CASPT2_ABS_SPECTRUM_NROOTS	The number of roots
CASPT2_CD_SPECTRUM	The CASPT2 CD spectrum
CASPT2_CD_SPECTRUM_INFO	Information about the excitations of the CASPT2 CD spectrum
CASPT2_CD_SPECTRUM_NROOTS	The number of roots
CAS_CUSTOM_ABS_SPECTRUM	The Custom CASSCF Absorption spectrum
CAS_CUSTOM_ABS_SPECTRUM_INFO	Information about the excitations of the custom CASSCF absorption
CAS_CUSTOM_ABS_SPECTRUM_NROOTS	The number of roots
CAS_CUSTOM_CD_SPECTRUM	The Custom CASSCF CD spectrum

Table 7.34 – continued from previous page

CAS_CUSTOM_CD_SPECTRUM_INFO	Information about the excitations of the custom CASSCF CD spectr
CAS_CUSTOM_CD_SPECTRUM_NROOTS	The number of roots
DCDCAS_ABS_SPECTRUM	The DCDCAS Absorption spectrum
DCDCAS_ABS_SPECTRUM_INFO	Information about the excitations of the DCDCAS absorption spectr
DCDCAS_ABS_SPECTRUM_NROOTS	The number of roots
CASSCF_DTENSOR_EIGENVALUES	CASSCF D Tensor eigenvalues
CASSCF_DTENSOR_RAW_EIGENVECTORS	CASSCF D Tensor Raw eigenvectors
CASSCF_DTENSOR_D	D value of CASSCF ZFS
CASSCF_DTENSOR_E	E value of CASSCF ZFS
CASSCF_DTENSOR_MULTIPLICITY	Spin multiplicity
CASPT2_DTENSOR_EIGENVALUES	CASPT2 D Tensor eigenvalues
CASPT2_DTENSOR_RAW_EIGENVECTORS	CASPT2 D Tensor raw eigenvectors
CASPT2_DTENSOR_D	D value of CASPT2 ZFS
CASPT2_DTENSOR_E	E value of CASPT2 ZFS
CASPT2_DTENSOR_MULTIPLICITY	Spin multiplicity
CAS_CUSTOM_DTENSOR_EIGENVALUES	custom CASSCF D Tensor eigenvalues
CAS_CUSTOM_DTENSOR_RAW_EIGENVECTORS	custom CASSCF D Tensor Raw eigenvectors
CAS_CUSTOM_DTENSOR_D	D value of custom CASSCF ZFS
CAS_CUSTOM_DTENSOR_E	E value of custom CASSCF ZFS
CAS_CUSTOM_DTENSOR_MULTIPLICITY	Spin multiplicity
CIPSI	
CIPSI_SPIN_MULTIPLICITY	The CIPSI spin multiplicity
CIPSI_NUM_OF_ROOTS	The CIPSI number of roots
CIPSI_FINAL_ENERGY	The CIPSI Final energy
CIPSI_ENERGIES	The CIPSI Energies
CIS	
CIS_FINAL_ENERGY	The final total energy
CIS_ESCF	The SCF Energy
CIS_E0	The Energy of the ground state
CIS_ENERGIES	The singlet energies
CIS_ENERGIESP1	The triplet energies
CIS_MODE	One of the CIS modes
CIS_NUM_OF_ROOTS	The number of roots
CIS_ROOT	State to be optimized
CIS_ABS_SPECTRUM_NROOTS	The number of roots
CIS_ABS_SPECTRUM	The CIS absorption spectrum
CIS_ABS_SPECTRUM_VELOCITY	The CIS absorptioin spectrum in velocity representation
CIS_ABS_SOC_SPECTRUM_NROOTS	The number or roots
CIS_ABS_SOC_SPECTRUM	The CIS absorption spectrum including SOC
CIS_CD_SPECTRUM_NROOTS	The number of roots
CIS_CD_SPECTRUM	The CIS CD spectrum
CIS_CD_SOC_SPECTRUM_NROOTS	The number of roots
CIS_CD_SOC_SPECTRUM	The CIS CD spectrum including SOC
ROCIS	
ROCIS_STATE	ROCIS State
ROCIS_REF_ENERGY	ROCIS Reference energy
ROCIS_CORR_ENERGY	ROCIS correlation energy
ROCIS_TOTAL_ENERGY	ROCIS total energy
ROCIS_ABS_SPECTRUM_NROOTS	Number of roots

Table 7.34 – continued from previous page

ROCIS_ABS_SPECTRUM	ROCIS Absorption spectrum
ROCIS_ABS_SOC_SPECTRUM_NROOTS	Number of roots
ROCIS_ABS_SOC_SPECTRUM	ROCIS absorption spectrum including SOC
ROCIS_CD_SPECTRUM_NROOTS	Number of roots
ROCIS_CD_SPECTRUM	ROCIS CD spectrum
ROCIS_CD_SOC_SPECTRUM_NROOTS	Number of roots
ROCIS_CD_SOC_SPECTRUM	ROCIS CD spectrum including SOC
MRCI	
MRCI_ABS_SPECTRUM	The MRCI absorption spectrum
MRCI_ABS_SPECTRUM_INFO	Information about the absorption spectrum
MRCI_ABS_SPECTRUM_NROOTS	The number of roots
MRCI_CD_SPECTRUM	The MRCI CD spectrum
MRCI_CD_SPECTRUM_INFO	Information about the MRCI CD spectrum
MRCI_CD_SPECTRUM_NROOTS	The number of roots
MRCI_DIPOLE_MOMENTS	The MRCI dipole moments
MRCI_DIPOLE_MOMENTS_INFO	Information about the MRCI dipole moments
MRCI_DTENSOR_EIGENVECTORS	The eigenvectors of the MRCI D tensor
MRCI_DTENSOR_EIGENVALUES	The eigenvalues of the MRCI D tensor
MRCI_DTENSOR_RAW_EIGENVECTORS	The raw eigenvectors of the MRCI D tensor
MRCI_DTENSOR_D	The MRCI D value for the ZFS
MRCI_DTENSOR_E	The MRCI E value for the ZFS
MRCI_DTENSOR_MULTIPLICITY	The MRCI spin multiplicity
EXTRAPOLATION	
EXTRAP_SCF_ENERGIES	The SCF energies with the different basis sets
EXTRAP_CBS_SCF	The extrapolated SCF energy
EXTRAP_CORR_ENERGIES	The correlation energies with the different basis sets
EXTRAP_CBS_CORR	The extrapolated correlatioin energy
EXTRAP_CBS_TOTAL	The extrapolated total energy
EXTRAP_CCSDT_X	The (T) contribution to the energy
EXTRAP_NUM_OF_ENERGIES	The number of energies (basis sets) used for the extrapolation
THERMOCHEMISTRY	
THERMO_TEMPERATURE	Temperature (^oK)
THERMO_PRESSURE	Pressure (Atm)
THERMO_TOTAL_MASS	Total Mass of the molecule (AMU)
THERMO_SPIN_DEGENERACY	Electronic degeneracy
THERMO_ELEC_ENERGY	Electronic energy (Eh)
THERMO_TRANS_ENERGY	Translational energy (Eh)
THERMO_ROT_ENERGY	Rotational energy (Eh)
THERMO_VIB_ENERGY	Vibrational energy (Eh)
THERMO_NUM_OF_FREQS	The number of vibrational frequencies
THERMO_FREQS	Frequencies
THERMO_ZPE	Zero point energy (Eh)
THERMO_INNER_ENERGY_U	Inner Energy (Eh)
THERMO_ENTHALPY_H	Enthalpy (Eh)
THERMO_ELEC_ENTROPY	(Electronic Entropy)*T (Eh)
THERMO_ROR_ENTROPY	(Rotational Entropy)*T (Eh)
THERMO_VIB_ENTROPY	(Vibrational Entropy)*T (Eh)
THERMO_TRANS_ENTROPY	(Translational Entropy)*T (Eh)
THERMO_ENTROPY_S	(Total Entropy)*T (Eh)

Table 7.34 – continued from previous page

THERMO_FREE_ENERGY_G	Free Energy (Eh)
EPR-NPR Spin-Spin coupling	
EPRNMR_SSC_NUM_OF_NUC_PAIRS	Number of nuclear pairs to calculate something
EPRNMR_SSC_NUM_OF_NUC_PAIRS_DSO	Number of nuclear pairs to calculate DSO terms
EPRNMR_SSC_NUM_OF_NUC_PAIRS_PSO	Number of nuclear pairs to calculate PSO terms
EPRNMR_SSC_NUM_OF_NUC_PAIRS_FC	Number of nuclear pairs to calculate FC terms
EPRNMR_SSC_NUM_OF_NUC_PAIRS_SD	Number of nuclear pairs to calculate SD terms
EPRNMR_SSC_NUM_OF_NUC_PAIRS_SD_FC	Number of nuclear pairs to calculate SD/FC terms
EPRNMR_SSC_NUM_OF_NUCLEI_PSO	Number of nuclei to calculate PSO perturbations
EPRNMR_SSC_NUM_OF_NUCLEI_FC	Number of nuclei to calculate SD/FC perturbations
Solvation	
SOLVATION_EPSILON	Dielectric constant
SOLVATION_REFRAC	Refractive index
SOLVATION_RSOLV	Solvent probe radius
SOLVATION_SURFACE_TYPE	Cavity surface
SOLVATION_CPCM_DIEL_ENERGY	Total energy including the CPCM dielectric correction
SOLVATION_NPOINTS	Number of points for the Gaussian surface
SOLVATION_SURFACE_AREA	Surface area
General Job Information	
JOB_INFO_MULT	Job Multiplicity
JOB_INFO_CHARGE	Job Charge
JOB_INFO_NUM_OF_ATOMS	Total number of atoms
JOB_INFO_NUM_OF_EL	Total number of electrons
JOB_INFO_NUM_OF_FC_EL	Number of frozen core electrons
JOB_INFO_NUM_OF_CORR_ELC	Number of correlated electrons
JOB_INFO_NUM_OF_BASIS_FUNCS	Number of basis functions
JOB_INFO_NUM_OF_AUXC_BASIS_FUNCS	Number of auxilliary C basis functions
JOB_INFO_NUM_OF_AUXJK_BASIS_FUNCS	Number of auxilliary J basis functions
JOB_INFO_NUM_OF_AUX_CABS_BASIS_FUNCS	Number of auxilliary JK basis functions
JOB_INFO_NUM_OF_AUX_CABS_BASIS_FUNCS	Number of auxilliary CABS basis functions
JOB_INFO_TOTAL_EN	Final energy
HESSIAN	
HESSIAN_MODES	The hessian
Math Functions	
ABS	Absolute value
COS	Cosine
SIN	Sine
TAN	Tangent
ACOS	Inverse cosine
ASIN	Inverse sine
ATAN	Inverse tangent
COSH	Hyperbolic cosine
SINH	Hyperbolic sine
TANH	Hyperbolic tangent
EXP	Exponential

Table 7.34 – continued from previous page

LOG	Common logarithm
LN	Natural logarithm
SQRT	Square root
ROUND	Round down to nearest integer

7.56.3 List of known Simple input commands

The name and a sort explanation of all compound protocols that are know through the simple input line, are given in the next table. The syntax is always:

Syntax:
! compound[*protocol name*]

Table 7.35: Protocols, known to the simple input line, with short expla

General Printing	
PRINT-ALLMETHODS	Print all available know protocols
PRINT-ALLDESCRIPTIONS	Print only the description of all protocols without the actual
Extrapolation Schemes	
EXTRAPOLATE-EP1-MDCI	Direct two-point extrapolation scheme with MDCI energies
COMPOUND[EXTRAPOLATE-EP1-MP2	SCF isotropic quadrupole moment
EXTRAPOLATE-EP2-DLPNO	EP2 type extrapolation with DLPNO-CCSD(T) as secondar
EXTRAPOLATE-EP2-MP2	EP2 type extrapolation with MP2 as secondary method
EXTRAPOLATE-EP3-DLPNO	EP3 type extrapolation with DLPNO-CCSD(T) as secondar
EXTRAPOLATE-EP3-MP2	EP3 type extrapolation with MP2 as secondary method
EXTRAPOLATE-PETERSON	Extrapolation based on the scheme of Peterson
EXTRAPOLATE-XANTHEAS-FELLER	Extrapolation based on the scheme of Xantheas and Feller
Wn Protocols	
W2-2	The W2-2 version of the Wn protocols
Gn Protocols	
G2-MP2	The G2-MP2 protocol
G2-MP2-ATOM	The G2-MP2 protocol for atoms
G2-MP2-SV	The G2-MP2-SV protocol
G2-MP2-SV-ATOM	The G2-MP2-SV protocol for atoms
G2-MP2-SVP	The G2-MP2-SVP protocol
G2-MP2-SVP-ATOM	The G2-MP2-SVP protocol for atoms
ccCA Protocols	
CCCA-DZ-QCISD-T	The CCCA-DZ-QCISD-T protocol
CCCA-DZ-QCISD-T-ATOM	The CCCA-DZ-QCISD-T protocol for atoms
CCCA-TZ-QCISD-T	The CCCA-TZ-QCISD-T protocol
CCCA-TZ-QCISD-T-ATOM	The CCCA-TZ-QCISD-T protocol for atoms
CCCA-ATZ-QCISD-T	The CCCA-ATZ-QCISD-T protocol
CCCA-ATZ-QCISD-T-ATOM	The CCCA-ATZ-QCISD-T protocol for atoms

Table 7.35 – continued from previous page

CCCA-CBS-1	The CCCA-CBS-1 protocol
CCCA-CBS-1-ATOM	The CCCA-CBS-1 protocol for atoms
CCCA-CBS-2	The CCCA-CBS-2 protocol
CCCA-CBS-2-ATOM	The CCCA-CBS-2 protocol for atoms
Accurate Energies	
EXTRAPOLATE-PNO	A custom PNO extrapolation scheme to reach the complete
DLPNO-CC-ENERGY	A protocol for accurate energies
Ab Initio ligand field	
AILFT_1SHELL	Ab-initio ligand field theory for 1 shell
AILFT_2SHELL	Ab-initio ligand field theory for 2 shells
X-Ray spectroscopy	
MRCI-XAS	X-Ray absorption spectroscopy with MRCI
CASCI-NEVPT2-XAS-XMCD	X-Ray and XMCD spectroscopies with CASCI-NEVPT2
MREOM-XAS	X-Ray absorption spectroscopy with MREOM..
Solvation	
DFT-SOLVATION-ENERGY	Calculation of solvation energy
COMPOUND[DFT-DIPOLE-MOMENT-SOLVENT-INDUCTION	Calculate the effect of solvent in the dipole moment
Geometry optimizations	
ITERATIVE_OPTIMIZATION	Iterative Optimization protocol to find structure with no neg

7.57 Compound Examples

7.57.1 Introduction

A library of compound scripts exist in page https://github.com/ORCAQuantumChemistry/CompoundScripts .

7.57.2 Hello World

Introduction

This is the simplest script that nevertheless points to an important feature of *Compound*. That is the fact that *Compound* does not have to run an actual 'normal' ORCA calculation but it can also be used as a driver for various tasks, in this case to just print a message.

Filename

helloWorld.inp

SCRIPT

```
%Compound
  print("Hellow World!\n");
EndRun
```

7.57.3 New Job

Introduction

One of the features of ORCA that will be deprecated in the future and should not be used any more is the *'New_Job'* feature. The current script is a simple example how *Compound* can be used to just run a series of calculations.

Filename

replaceNewJob.inp

SCRIPT

```
# This is a small script thas shows how
# 'Compound' can replace the previous
# ORCA '$New_Job' feature
%Compound
  # ------------------------------------
  # First job
  # ------------------------------------
  New_Step
    !BP86
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
  # ------------------------------------
  # Second job with same goemetry
  # but different functional
  # ------------------------------------
  New_Step
    !B3LYP
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
EndRun
```

Comments

From the *Compound* point of view the syntax in this script is not the most efficient one. It can be rewritten in more compact, cleaner, general way. Neverteless this is meant only as an exmample of how *Compound* can replace older ORCA calculations that used the, to be deprecated, *'New_Job'* feature.

7.57.4 High Accuracy

Introduction

This is a script that utilizes the scheme by N. J. DeYonker, T. R. Cundari, and A. K. Wilson published on: J. Chem. Phys. 124, 114104 (2006). The script calculates accurate total energies of molecules.

Filename

ccCA_CBS_2.cmp

SCRIPT

```
# This is a small script thas shows how
# 'Compound' can replace the previous
# ORCA '$New_Job' feature
%Compound
  # ------------------------------------
  # First job
  # ------------------------------------
  New_Step
    !BP86
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
  # ------------------------------------
  # Second job with same goemetry
  # but different functional
  # ------------------------------------
  New_Step
    !B3LYP
    *xyz 0 1
      H 0.0 0.0 0.0
      H 0.0 0.0 0.8
    *
  Step_End
EndRun
```

Comments

It is interesting that in this scheme the total energy is treated and there is not separation in extrapolation between HF energy and correlation energy.

7.57.5 Scan

Introduction

This is an example script for a 1-Dimensional geometry scan. It is set up for the Ne-Ne bond distance but can be modified to suit the user's specific needs.

Filename

scan_1D_1M_1P.cmp

SCRIPT

```
# Author : Dimitrios G. Liakos
# Date   : May of 2024
#
# This is a script that will calculate and potentially
#   plot ONE property(1P) along a scan in ONE dimesion (1D)
#   using only ONE method (1M)
#
#   It is part of a series of scripts for different
#     combinations of scans for dimensions, methods,
#     and properties
#
# Here as an example we use for:
#    - dimension: the Ne-Ne bond (dist)
#    - method   : "HF" (method)
```

(continues on next page)

(continued from previous page)

```
#    - property : the SCF energy (propName)
#
# The script creates a csv file with the absolute energies
#   and an additional one with the potential energies  in
#   kcal/mol. Both will be saved on disk.
#
# If 'DoPython' is set to true it will also create a python
#   script that plots the generated values and then run
#   it. The python script will be saved on disk and thus one
#   can afterwards manipulate it.
#
# NOTE The boolean option plotPotential will choose between
#      plotting absolute values or potential.
#
# NOTE The boolean obtion doKcal if set to true multiplies
#      the potential values with the HartreeToKcal factor.
#
# NOTE In case the doPython is set to true the script expects
#   that python3 is avaiable and also the following libraries:
#   - pandas
#   - seaborn
#   - matplotlib.pyplot
#
# ------------------------------------------------------------
#
# ----------------    Variables to change (e.g. through 'with')   --------------------------
Variable method       = "HF";             # The methods of the calculation
Variable basis        = "cc-pVDZ";        # The basis set of the calculation
Variable restOfInput  = "TightSCF";       # Maybe something common for the simple input
Variable charge       = 0;                # Charge
Variable mult         = 1;                # Spin multiplicity
Variable myPropName   = "SCF_Energy";     # The properties we want to read
#
Variable lowerLimit   = 2.5;              # Lower limit value
Variable UpperLimit   = 5.0;              # Upper limit value
Variable NSteps       = 13;               # Number of steps for the grid
Variable baseFilename = "myPotential";    # The basename for the created files
Variable plotPotential= true;             # Plot the potential instead of absolute values
Variable DoKcal       = true;             # Multiply the potential values with the␣
↪HartreeToKcal factor
Variable removeFiles  = true;             # Remove *_Compound_*, *bas* files
# ------------------  python plot relevant variables ---------------------------------------
Variable DoPython     = true;             # if we want python or not
Variable lw           = 4;                # The line width in case we plot with python
Variable marker       = "o";              # The type of markers
Variable markerSize   = 10;               # The size of the markers in case we plot
Variable fontSize     = 18;
#
# ------------------------      Rest of the variables       ---------------------------------
#
Variable HartreeToKcal = 627.5096080305927;                  # Hartree to kcal/mol conversion␣
↪factor
Variable stepSize      = (UpperLimit-LowerLimit)/(NSteps-1);  # The stepsize of the grid
Variable calcValues[NSteps];                                 # An array to store the␣
↪calculated values
Variable  res, dist, calcValue;
Variable myFilename, csvFilename;
Variable fPtr;                                               # A file to write

# ------------------------------------------------
#  Open and Write file header for the absolute values
```

(continues on next page)

(continued from previous page)

```
# ------------------------------------------------
write2String(csvFilename, "%s_absValues.csv", baseFilename);
fPtr = OpenFile(csvFilename, "w");
write2File(fptr, "distance,method,property,calcValue\n");

# ------------------------------------------------
#  Perform the calculations and update the file
# ------------------------------------------------
for iStep from 0 to NSteps-1 Do
  dist = lowerLimit + (iStep)*stepSize;
  New_Step
    !&{method} &{basis} &{restOfInput}
    *xyz &{charge} &{mult}
      Ne 0.0 0.0 0.0
      Ne 0.0 0.0 &{dist}
    *
    Step_end
    res = calcValue.readProperty(propertyName=myPropName);
    write2File(fPtr, "%.4lf,%20s,%20s,%20.10lf\n", dist, method,myPropName, calcValue);
    calcValues[iStep]=calcValue;
EndFor
CloseFile(fPtr);   # Close the file

# ------------------------------------------------
# Evaluate and write the relative values
# ------------------------------------------------
write2String(csvFilename, "%s_relValues.csv", baseFilename);
fPtr = OpenFile(csvFilename, "w");
write2File(fPtr, "distance,method,property,calcValue\n");
for iStep from 0 to NSteps-1 Do
  dist = lowerLimit + (iStep)*stepSize;
  if (DoKcal) then
    calcValue = (calcValues[iStep]-calcValues[NSteps-1])*HartreeToKcal;
  else
    calcValue = calcValues[iStep]-calcValues[NSteps-1];
  EndIf
  write2File(fPtr, "%.4lf,%20s,%20s,%20.10lf\n", dist, method,myPropName, calcValue);
EndFor
CloseFile(fPtr);   # Close the file

if (removeFiles) then
  sys_cmd("rm *_Compound_* *.bas*");
EndIf

# ------------------------------------------------
#  Create a python file and run it
# ------------------------------------------------
if (DoPython) then
  if (plotPotential) then
    write2String(csvFilename, "%s_relValues.csv", baseFilename);
  else
    write2String(csvFilename, "%s_absValues.csv", baseFilename);
  endIf

  write2String(myFilename, "%s.py", baseFilename);
  fPtr = openFile(myFilename, "w");
  # Import necessary libraries
  write2File(fPtr, "import pandas as pd\n");
  write2File(fPtr, "import seaborn as sns\n");
  write2File(fPtr, "import matplotlib.pyplot as plt\n");
  # Read the csv file
```

(continues on next page)

(continued from previous page)

```
  write2File(fPtr, "df = pd.read_csv('%s')\n", csvFilename);
  #Make a lineplot
  write2File(fPtr, "sns.lineplot(data=df, x=\"distance\", y=\"calcValue\", hue=\"property\", \n
                       lw=%d, markers=True, marker='%s', markersize=%d, dashes=False)\n", lw,␣
↪marker, markersize);
  write2File(fPtr, "plt.axhline(y=0, color='black', linestyle='-', linewidth=1)\n");
  write2File(fPtr, "plt.title(\"Energy Potential\", fontsize=%d)\n", fontsize+4);
  write2File(fPtr, "plt.xlabel(\"Ne-Ne Distance\", fontsize=%d)\n", fontsize);
  write2File(fPtr, "plt.ylabel(\"Energy (kcal/mol)\", fontsize=%d)\n", fontsize);
  write2File(fPtr, "plt.xticks(fontsize=%d)\n", fontSize);
  write2File(fPtr, "plt.yticks(fontsize=%d)\n", fontSize);
  write2File(fPtr, "plt.show()\n");
  closeFile(fPtr);
  sys_cmd("python3 %s", myFilename);
EndIf

End
```

Comments

This script has some interesing features. It contains two variables *removeFiles* and *DoPython*. If the first of them is set to *true* then the script will use a system command to remove files that are not needed anymore after the end of the calculation. The latter, *DoPython*, if set to *true* will read the *.csv* file that is created and write a *python* file to make a plot of the results. Then it will run the python script to actually make the plot.

7.57.6 Numerical polarizabilities

Introduction

This script calculates numerically the polarizability of the molecule using single point calculations with an electric field.

Filename

numericalPolarizability.cmp

SCRIPT

```
# Authors: Dimitrios G. Liakos / Frank Neese / Zikuan Wang
# Date   : May of 2024
#
# This is a compound script that calculates the
#  dipole-dipole polarizability tensor numerically
#  using the double derivative of energy.
#
# The idea is the following:
#
# 1 Perform a field free calculation
#
# 2 Loop over directions I=X,Y,Z
#
# 3 Loop over directions J=X,Y,Z
#
#    - put a small Q-field in directions I and J
#    - Solve equations to get the energy for each combination
#    - Polarizability alpha(I,J) =- ( E(+I,+J) - E(+I,-J)-E(-I,+J)+ E(-i,-j)/(4*Field^2)
# 4 Print polarisability
#
# ------------------------------------------------------------------------
# ----------------------    Variables    ---------------------------------
# --- Variables to be adjusted (e.g. using 'with' ----------------------
```

(continues on next page)

(continued from previous page)

```
Variable molecule    = "h2o.xyz";
Variable charge      = 0;
Variable mult        = 1;
Variable method      = "HF";
Variable basis       = " ";
Variable restOfInput = "VeryTightSCF";
Variable blocksInput = " ";
Variable E_Field     = 0.0001;
Variable enPropName  = "JOB_Info_Total_En";
Variable removeFiles = true;
# -------------- Rest of the variables --------------------------------
Variable FField[3];
Variable EFree, EPlusPlus, EPlusMinus, EMinusPlus, EMinusMinus, a[3][3];
Variable FFieldStringPlusPlus, FFieldStringPlusMinus;
Variable FFieldStringMinusPlus, FFieldStringMinusMinus;
Variable aEigenValues, aEigenVectors;

# ----------------------------------------
# Calculation without field
# ----------------------------------------
New_Step
  !&{method} &{basis} &{restOfInput}
  &{blocksInput}
  *xyzfile &{charge} &{mult} &{molecule}
Step_End
EFree.ReadProperty(propertyName=enPropName);

# ------------------------------------------------------------
# Loop over the x, y, z directions
# ------------------------------------------------------------
for i from 0 to 2 Do
  for j from 0 to 2 Do
    # ------------------------------------------------------------
    # Create the appropriate direction oriented field string
    # ------------------------------------------------------------
    # ---------------------- (++) -----------------------------
    for k from 0 to 2 Do
      FField[k] = 0.0;
    EndFor
    FField[i] = FField[i] + E_Field;
    FField[j] = FField[j] + E_Field;
    write2String(FFieldStringPlusPlus,   " %lf,  %lf,  %lf",
    FField[0], FField[1], FField[2]);
    #
    # ---------------------- (+-) -----------------------------
    for k from 0 to 2 Do
      FField[k] = 0.0;
    EndFor
    FField[i] = FField[i] + E_Field;
    FField[j] = FField[j] - E_Field;
    write2String(FFieldStringPlusMinus,   " %lf,  %lf,  %lf",
    FField[0], FField[1], FField[2]);
    #
    # ---------------------- (-+) -----------------------------
    for k from 0 to 2 Do
      FField[k] = 0.0;
    EndFor
    FField[i] = FField[i] - E_Field;
    FField[j] = FField[j] + E_Field;
    write2String(FFieldStringMinusPlus,   " %lf,  %lf,  %lf",
    FField[0], FField[1], FField[2]);
```

(continues on next page)

(continued from previous page)

```
#
# ---------------------- (--) -----------------------------
for k from 0 to 2 Do
  FField[k] = 0.0;
EndFor
FField[i] = FField[i] - E_Field;
FField[j] = FField[j] - E_Field;
write2String(FFieldStringMinusMinus,   " %lf,  %lf,  %lf",
FField[0], FField[1], FField[2]);

# ----------------------------------------
# Perform the calculations.
# The plus_plus (++) one
# ----------------------------------------
ReadMOs(1);
New_Step
  !&{method} &{basis} &{restOfInput}
  %SCF
    EField = &{FFieldStringPlusPlus}
  End
  &{blocksInput}
Step_End
EPlusPlus.readProperty(propertyName=enPropName);
# ----------------------------------------
# The plus_minus (+-) one
# ----------------------------------------
ReadMOs(1);
New_Step
  !&{method} &{basis} &{restOfInput}
   %SCF
    EField = &{FFieldStringPlusMinus}
  End
  &{blocksInput}
Step_End
EPlusMinus.readProperty(propertyName=enPropName);
# ----------------------------------------
# The minus_plus (-+) one
# ----------------------------------------
ReadMOs(1);
New_Step
  !&{method} &{basis} &{restOfInput}
  %SCF
    EField = &{FFieldStringMinusPlus}
  End
  &{blocksInput}
Step_End
EMinusPlus.readProperty(propertyName=enPropName);
# ----------------------------------------
# And the minus_minus (--) one
# ----------------------------------------
ReadMOs(1);
New_Step
  !&{method} &{basis} &{restOfInput}
   %SCF
    EField = &{FFieldStringMinusMinus}
  End
  &{blocksInput}
Step_End
EMinusMinus.readProperty(propertyName=enPropName);

a[i][j] = -(EPlusPlus-EPlusMinus-EMinusPlus+EMinusMinus)/(4*E_Field*E_Field);
```

(continues on next page)

(continued from previous page)

```
  EndFor
EndFor

# ----------------------------------------
# Diagonalize
# ----------------------------------------
a.Diagonalize(aEigenValues, aEigenVectors);

# ----------------------------------------
# Do some printing
# ----------------------------------------
print( "\n\n");
print( " ------------------------------------------------------\n");
print( "                   COMPOUND                           \n");
print( " Numerical calculation of dipole polarizability\n");
print( " ------------------------------------------------------\n");
print( " Molecule    : %s\n", molecule);
print( " charge      : %d\n", charge);
print( " Mult        : %d\n", mult);
print( " Method      : %s\n", method);
print( " Basis       : %s\n", basis);
print( " RestOfInput : %s\n", restOfInput);
print( " BlocksInput : %s\n", blocksInput);
print( " The electric field perturbation used was:    %.5lf a.u.\n", E_Field);
print( " \n\n");

print( " ------------------------------------------------------\n");
print( " Raw electric dipole  polarizability tensor is:\n");
print( " ------------------------------------------------------\n");
For i from 0 to 2 Do
  print("%13.8lf  %13.8lf  %13.8lf\n", a[i][0], a[i][1], a[i][2]);
EndFor
print( " ------------------------------------------------------\n");
print("\n");

print( " ------------------------------------------------------\n");
print( " Raw electric dipole polarizability Eigenvalues\n");
print( " ------------------------------------------------------\n");
print("%13.8lf  %13.8lf  %13.8lf\n", aEigenValues[0], aEigenValues[1], aEigenValues[2]);
print( " ------------------------------------------------------\n");
print("\n");

print( " ------------------------------------------------------\n");
print( " Raw electric dipole polarizability Eigenvectors\n");
print( " ------------------------------------------------------\n");
For i from 0 to 2 Do
  print("%13.8lf  %13.8lf  %13.8lf\n", aEigenVectors[i][0], aEigenVectors[i][1],␣
↪aEigenVectors[i][2]);
EndFor

print( "\n a isotropic value : %.5lf\n", (aEigenValues[0]+aEigenValues[1]+aEigenValues[2])/3.
↪0);
print( " ------------------------------------------------------\n");
print("\n\n");
#
#
# ------------------------------------------------
#  Maybe remove unneccesary files
# ------------------------------------------------
if (removeFiles) then
  sys_cmd("rm *_Compound_* *.bas* ");
```

(continues on next page)

(continued from previous page)

```
EndIf
#
End
```

Comments

In this script we also use the linear algebra *diagonalize* function that is available in *Compound*.

7.57.7 Iterative optimization

Introduction

This is a script that will perform a geometry optimization, then run a frequency calculation and in case there are negative frequencies it will adjust the geometry, based on the Hessian, and optimize again.

Filename

iterativeOptimization.cmp

SCRIPT

```
# Author: Dimitrios G. Liakos and Franke Neese
# Date  : May/June of 2024
#
# ************************************* DESCRIPTION␣
↪*********************************************
# iterative Optimization protocol to find structure with no negative
# frequencies (e.g. real minima)
#
# Step 1. Run a single point calculation (we need it for the first property file)
#
# Step 1. Loop and perform calculations with (optimization and frequencies)
#
# Step 2. Check the frequencies. If there are negative ones use the hessian
#         of the appropriate normal mode to adjust the geometry
#
# ------ Variables to adjust (e.g. using 'with') -------------------
Variable method          = "HF"; #"HF-3c";
Variable MaxNTries       =  25;   # Number of maximum tries
Variable CutOff          = -10.0; # CutOff for a negative frequency
Variable scaling         = 0.6;   # Scaling factor for normal mode
Variable NNegativeTarget = 0;     # Number of negative frequencies we allow
Variable myFilename      = "xyzInput.xyz";
Variable charge          = 0;
Variable multiplicity    = 2;
# -----------------------------------------------------------------
# ------            Rest of variables            -------------------
Geometry myGeom;
Variable freqs, modes;
Variable res = -1;
Variable NNegative =   0;
Variable OptDone;

# ---------------------------------------------------------
# Perform a single point calculation. We need it for
# the initial geometry from the property file
# ---------------------------------------------------------
New_Step
  !&{method}
Step_End
myGeom.Read();
myGeom.WriteXYZFile(filename=myFilename);
```

(continues on next page)

(continued from previous page)

```
# ------------------------------------------------------------
# Start a for loop over number of tries
# ------------------------------------------------------------
For itry From 1 To maxNTries Do
  # -----------------------------------------
  # Perform a geometry optimization/Frequency calculation
  # -----------------------------------------
  New_Step
    ! &{method} freq Opt
    *xyzfile &{charge} &{multiplicity} &{myFilename}
  Step_End
  res = freqs.readProperty(propertyName = "THERMO_FREQS");
  res = modes.readProperty(propertyName = "HESSIAN_MODES");
  myGeom.Read();

  # -----------------------------------------
  #  check for sufficiently negative frequencies
  # -----------------------------------------
  NNegative = 0;
  For ifreq From 0 to freqs.GetSize()-1 Do
    if ( freqs[ifreq] < CutOff )  then
       myGeom.FollowNormalMode(vibrationSN=ifreq, scalingFactor=scaling);
       NNegative = NNegative + 1;
    endif
  endfor
  myGeom.WriteXYZFile(filename=myFilename);
  If ( NNegative <= NNegativeTarget ) then
    goto OptDone;
  endif
endfor

# ------------------------------------------------------------------
# Either found correct geometry or reached maximum number of tries.
# ------------------------------------------------------------------
OptDone :
if (NNegative > NNegativeTarget) then
  print("ERROR The program did not find a structure with the desired\n number of imaginary␣
↪frequencies.\n There are %9.3lf negative frequencies after %3d steps", NNegative,itry);
else
  print("\nSUCCESS optimized structure with (%d) negative\n frequencies found after %3d steps",
↪ NNegative, itry);
endif

End
```

7.57.8 Gradient extrapolation

Introduction

This script extrapolates the gradient of a molecule. It uses a two point extrapolation where the Hartree-Fock and correlation parts of the gradient are extrapolated separately. This opens the way for geometry optimizations with extrapolated gradients.

Filename

gradientExtrapolation.cmp

SCRIPT

```
#   Author: Dimitrios G. Liakos and Frank Neese
#   Date  : May of 2024
#
#   This is a compound file that extrapolates the
#   energy gradients to Complete Basis Set Limit (CBS).
#
#   STEPS:
#   Step1 : Run HF calculation with small basis set
#           Read scfGradX and scfEnX
#   Step2 : Run Correlation calculation with small basis set
#           Read totalGradX and totalEnX
#   Step3 : Calculate the gradient differene to get
#           the corrGradX (only the correlation part)
#   Step4 : Run HF calculation with big basis set
#           Read scfGradY and scfEnY
#   Step5 : Run correlation calculation with big basis set
#           Read totalGradY and totalEnY
#   Step6 : Calculate the gradient difference with the
#           big basis set to get corrGradY
#   Step7 : Evaluate scfGradCBS and scfEnCBS
#           using scfGradX and scfGradY
#   Step8 : Evaluate corrGradCBS using
#           corrGradX and corrGradY
#   Step9 : Add scfGradCBS and corrGradCBS to get
#           totalGradCBS
#   Step10: If needed, create an ORCA engrad file
#
#
#   NOTE: It works with an xyz file the name of which we should provide.
#         using the variable initialXYZFilename.
#
#   We extrapolate the SCF part using the scheme
#     proposed in: J. Phys. Chem. 129, 184116, 2008
#       E_SCF(X) = E_SCF(CBS)+Aexp(-a*SQRT(X))
#
#   We extrapolate the correlation part using the schem
#     proposed in: J. Chem. Phys. 1997, 106, 9639
#   E_CBS(CORR)=(X^b*E_X(CORR)-Y^b*E_Y(CORR))/(X^b-Y^b)

#   We use alpha and beta exponents proposed in:
#     J. Chem. Theory Comput., 7, 33-43 (2011)
# ----------------------    Variables    -------------------------------
# --- Variables to be adjusted (e.g. using 'with' ----------------------
Variable Molecule           = "initial.xyz";   # xyz file of the initial structure
Variable charge             = 0;               # Charge
Variable multiplicity       = 1;               # Spin multiplicity
Variable method             = "MP2";           # The method we use for the calculation
Variable LowerBasis         = "cc-pVDZ";       # Small basis set
Variable UpperBasis         = "cc-pVTZ";       # Big basis set
Variable restOfInput        = "EnGrad ";       # The rest of the simple input
Variable addCorrelation     = true;            # If we have a correlation part
Variable scfEnPropName      = "MP2_Ref_Energy";  # The name of the property for the SCFenergy
Variable corrEnPropName     = "MP2_Corr_Energy"; # The name of the property for the␣
↪correlation energy
Variable LowerCardinal      = 2;               # Cardinal number of small basis set
Variable UpperCardinal      = 3;               # Cardinal number of big basis set
Variable alpha              = 4.420;           # Exponent for SCF extrapolation
Variable beta               = 2.460;           # Exponent for corrleation extrapolation
Variable enGradFilename     = "result.engrad"; # Filename of the ORCA engrad file
Variable produceEnGradFile  = true;            # Produce an ORCA engrad file
# ----------------------------------------------------------------------
# -------------- Rest of the variables -------------------------------
```

(continues on next page)

(continued from previous page)

```
Geometry myGeom;
Variable scfGradX, scfGradY;                    # SCF Gradients
Variable scfEnX, scfEnY, scfEnCBS;              # SCF energies
Variable corrEnX, corrEnY, corrEnCBS;           # Correlation enegies
Variable totalGradX, totalGradY;                # Total Gradients
Variable eX = 0.0;
Variable eY = 0.0;
Variable res = -1;

Variable denominator = 0.0;
Variable gradX = 0.0, gradY = 0.0, gradCBS=0.0;
Variable nAtoms = 0;
Variable EnGradFile;
Variable Cartesians, AtomicNumbers;

# ------------------------------------------------------------------
# Step 1. SCF Calculation with small basis set (X)
# ------------------------------------------------------------------
New_Step
  ! HF &{LowerBasis} &{restOfInput}
  *xyzfile &{charge} &{multiplicity} &{Molecule}
Step_end
res = scfEnX.readProperty(propertyName="SCF_Energy");
res = scfGradX.readProperty(propertyName="Nuclear_Gradient", Property_Base=true);
myGeom.Read();
nAtoms = myGeom.GetNumOfAtoms();

# ------------------------------------------------------------------
# Step 2. Initialize rest of the variables
# ------------------------------------------------------------------
Variable corrGradX[3*nAtoms];   # Correlation part of gradient with basis X
Variable corrGradY[3*nAtoms];   # Correlation part of gradient with basis Y
Variable corrGradCBS[3*nAtoms]; # CBS estimation of correlation part of the gradient
Variable scfGradCBS[3*nAtoms];  # CBS estimation of SCF part of the gradient
Variable totalGradCBS[3*nAtoms];# CBS estimation of total gradient

# ------------------------------------------------
# Step3. Correlation Calculation with small basis set (X)
# ------------------------------------------------
if (addCorrelation) then
  New_Step
    ! &{method} &{LowerBasis} &{restOfInput}
  Step_end
  res = scfEnX.readProperty(propertyName=scfEnPropName);
  res = corrEnX.readProperty(propertyName=corrEnPropName);
  res = totalGradX.readProperty(propertyName="Nuclear_Gradient", Property_Base=true);

  # ------------------------------------------------
  # Evaluate correlation gradient with small basis set (X)
  # ------------------------------------------------
  corrGradX =mat_p_mat(1, totalGradX, -1, scfGradX);
EndIf

# ------------------------------------------------
# Step4. SCF Calculation with large basis set (Y)
# ------------------------------------------------
New_Step
  !HF &{UpperBasis} &{restOfInput}
Step_End
res = scfEnY.readProperty(propertyName="SCF_Energy");
res = scfGradY.readProperty(propertyName="Nuclear_Gradient", Property_Base=true);
```

(continues on next page)

(continued from previous page)

```
# ------------------------------------------------
# Step5. Correlation calculation with large basis set (Y)
# ------------------------------------------------
if (addCorrelation) then
  New_Step
    ! &{method} &{UpperBasis} &{restOfInput}
  Step_end
  res = scfEnY.readProperty(propertyName=scfEnPropName);
  res = corrEnY.readProperty(propertyName=corrEnPropName);
  res = totalGradY.readProperty(propertyName="Nuclear_Gradient", Property_Base=true);

  # ------------------------------------------------
  # Evaluate correlation gradient with big basis set Y
  # ------------------------------------------------
  corrGradY = mat_p_mat(1, totalGradY, -1, scfGradY);
EndIf

# ------------------------------------------------
# Step6. Extrapolate the SCF part of the gradient and energy
# ------------------------------------------------
eX               = exp(-alpha * sqrt(LowerCardinal));
eY               = exp(-alpha * sqrt(UpperCardinal));
denominator      = eY-eX;

scfEnCBS    = (scfEnX*eY - scfEnY*eX)/(eY-eX);
for i from 0 to scfGradX.GetSize()-1 Do
  gradX = scfGradX[i];
  gradY = scfGradY[i];

  scfGradCBS[i]  = (gradX * eY - gradY * eX)/denominator;
endFor

if (addCorrelation) then
  # ------------------------------------------------
  # Step7. Extrapolate the correlation part of the gradient and energy
  # ------------------------------------------------
  denominator = LowerCardinal^(beta)-(UpperCardinal)^(beta);

  corrEnCBS =  (LowerCardinal^(beta)*corrEnX-(UpperCardinal)^(beta)*corrEnY)/denominator;
  for i from 0 to scfGradX.GetSize()-1 Do
    gradX = corrGradX[i];
    gradY = corrGradY[i];

    corrGradCBS[i] = (LowerCardinal^(beta)*gradX-(UpperCardinal)^(beta)*gradY)/denominator;
  endFor

  # ------------------------------------------------
  # Add SCF and correlation part to get total CBS extrapolated values
  # ------------------------------------------------
  totalGradCBS = mat_p_mat( 1, scfGradCBS, 1, corrGradCBS);
EndIf


# ------------------------------------------------
# Step8. Present the results
# ------------------------------------------------
print( "\n\n\n");
print( "------------------------------------------------------------\n");
print( "             Compound Extrapolation of Gradient         \n");
print( "------------------------------------------------------------\n");
```

(continues on next page)

(continued from previous page)

```
print( "Number of atoms       : %d\n", nAtoms);
print( "Lower basis set       : %s\n", LowerBasis);
print( "Upper basis set       : %s\n", UpperBasis);
print( "Alpha                 : %.2lf\n", alpha);
print( "Beta                  : %.2lf\n", beta);
print( "Lower Cardinal number : %d\n", LowerCardinal);
print( "Upper Cardinal number : %d\n", UpperCardinal);
print( "Method                : %s\n", method);
print( "AddCorrelation        : %s\n", AddCorrelation.GetString());
print( "Produce EnGrad File   : %s\n", produceEnGradFile.GetString());
print( "\n\n");
print( "SCF Energy with small basis set         : %.12e\n", scfEnX);
print( "SCF Energy with big basis set           : %.12e\n", scfEnY);
print( "Extrapolated SCF energy                 : %.12e\n", scfEnCBS);
print("\n\n");
if (addCorrelation) then
  print( "Correlation Energy with small basis set : %.12e\n", corrEnX);
  print( "Correlation Energy with big basis set   : %.12e\n", corrEnY);
  print( "Extrapolated correlation  energy        : %.12e\n", corrEnCBS);
  print("\n\n");
  print( "Total Energy with small basis set : %.12e\n", scfEnX   + corrEnX);
  print( "Total Energy with big basis set   : %.12e\n", scfEnY   + corrEnY);
  print( "Extrapolated Total energy         : %.12e\n", scfEnCBS + corrEnCBS);
  print("\n\n");
else
  print( "Total Energy with small basis set : %.12e\n", scfEnX);
  print( "Total Energy with big basis set   : %.12e\n", scfEnY);
  print( "Extrapolated Total energy         : %.12e\n", scfEnCBS);
  print("\n\n");
EndIf

print( "-----------------------------------------------------------------\n");
print( "SCF Gradient with basis set: %s\n", LowerBasis );
print( "-----------------------------------------------------------------\n");
print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
for i from 0 to nAtoms-1 Do
  print("%4d     %20lf     %20lf     %20lf\n", i, scfGradX[3*i], scfGradX[3*i+1],␣
↪scfGradX[3*i+2]);
EndFor
if (addCorrelation) then
  print( "-----------------------------------------------------------------\n");
  print( "Correlation Gradient with basis set: %s\n", LowerBasis );
  print( "-----------------------------------------------------------------\n");
  print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
  for i from 0 to nAtoms-1 Do
    print("%4d     %20lf     %20lf     %20lf\n", i, corrGradX[3*i], corrGradX[3*i+1],␣
↪corrGradX[3*i+2]);
  EndFor

  print( "-----------------------------------------------------------------\n");
  print( "Total Gradient with basis set: %s\n", LowerBasis );
  print( "-----------------------------------------------------------------\n");
  print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
  for i from 0 to nAtoms-1 Do
    print("%4d     %20lf     %20lf     %20lf\n", i, totalGradX[3*i], totalGradX[3*i+1],␣
↪totalGradX[3*i+2]);
  EndFor
EndIf

print( "-----------------------------------------------------------------\n");
print( "SCF Gradient with basis set: %s\n", UpperBasis );
```

(continues on next page)

(continued from previous page)

```
print( "-----------------------------------------------------------------\n");
print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
for i from 0 to nAtoms-1 Do
  print("%4d     %20lf     %20lf     %20lf\n", i, scfGradY[3*i], scfGradY[3*i+1],␣
↪scfGradY[3*i+2]);
EndFor

if (addCorrelation) then
  print( "-----------------------------------------------------------------\n");
  print( "Correlation gradient with basis set: %s\n", UpperBasis );
  print( "-----------------------------------------------------------------\n");
  print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
  for i from 0 to nAtoms-1 Do
    print("%4d     %20lf     %20lf     %20lf\n", i, corrGradY[3*i], corrGradY[3*i+1],␣
↪corrGradY[3*i+2]);
  EndFor
  print( "-----------------------------------------------------------------\n");
  print( "Total Gradient with basis set: %s\n", UpperBasis );
  print( "-----------------------------------------------------------------\n");
  print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
  for i from 0 to nAtoms-1 Do
    print("%4d     %20lf     %20lf     %20lf\n", i, totalGradY[3*i], totalGradY[3*i+1],␣
↪totalGradY[3*i+2]);
  EndFor
EndIf

print( "-----------------------------------------------------------------\n");
print( "Extrapolated SCF part of the Gradient:\n" );
print( "-----------------------------------------------------------------\n");
print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
for i from 0 to nAtoms-1 Do
  print("%4d     %20lf     %20lf     %20lf\n", i, scfGradCBS[3*i], scfGradCBS[3*i+1],␣
↪scfGradCBS[3*i+2]);
EndFor

if (addCorrelation) then
  print( "-----------------------------------------------------------------\n");
  print( "Correlation Gradient with basis set:\n" );
  print( "-----------------------------------------------------------------\n");
  print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
  for i from 0 to nAtoms-1 Do
    print("%4d     %20lf     %20lf     %20lf\n", i, corrGradCBS[3*i], corrGradCBS[3*i+1],␣
↪corrGradCBS[3*i+2]);
  EndFor
  print( "-----------------------------------------------------------------\n");
  print( "Total Gradient with basis set:\n" );
  print( "-----------------------------------------------------------------\n");
  print( "Atom     %20s     %20s     %20s\n", "X", "Y", "Z");
  for i from 0 to nAtoms-1 Do
    print("%4d     %20lf     %20lf     %20lf\n", i, totalGradCBS[3*i], totalGradCBS[3*i+1],␣
↪totalGradCBS[3*i+2]);
  EndFor
EndIf
print( "-----------------------------------------------------------------\n");


if (produceEnGradFile) then
  # ----------------------------------------
  # Read the geometry of the last calculation
  # ----------------------------------------
  myGeom.Read();
```

(continues on next page)

(continued from previous page)

```
  Cartesians = myGeom.GetCartesians();
  atomicNumbers = myGeom.GetAtomicNumbers();
  EnGradFile = openFile(enGradFilename, "w");
  Write2File(EnGradFile, "\n\n\n");
  Write2File(EnGradFile, " %d\n", nAtoms);
  Write2File(EnGradFile, "\n\n\n");
  if (addCorrelation) then
    Write2File(EnGradFile, " %.12lf\n", scfEnCBS + corrEnCBS);
  else
    Write2File(EnGradFile, " %.12lf\n", scfEnCBS);
  EndIf
  Write2File(EnGradFile, "\n\n\n");
  for i from 0 to 3*nAtoms-1 Do
    if (addCorrelation) then
      Write2File(EnGradFile, "     %20.12lf\n", totalGradCBS[i]);
    else
      Write2File(EnGradFile, "     %20.12lf\n", scfGradCBS[i]);
    EndIf
  EndFor
  Write2File(EnGradFile, "\n\n\n");
  for i from 0 to nAtoms-1 Do
    Write2File(EnGradFile, "%5d  %12.8lf  %12.8lf   %12.8lf\n", atomicNumbers[i],␣
↪cartesians[i][0], cartesians[i][1], cartesians[i][2]);
  EndFor
  closeFile(EnGradFile);

EndIf

End
```

Comments

7.57.9 BSSE Optimization

Introduction

This script optimizes the geometry of a molecule using gradients corrected for Basis Set Superposition Error (BSSE) correction. The basic step is the usage of a second script that calculates BSSE corrected gradients.

Filename

BSSEOptimization.cmp

SCRIPT

```
# Author:  Frank Neese and Dimitrios G. Liakos
# Date  :  May of 2024
# ---------------------------------------------------
#
# This is a script that will use a compound script to
#   calculate BSSE corrected gradients and use them
#   in combination with ORCA External Optimizer to
#   perform a geometry optimization.
#
# We perform the following steps.
# 1. Choose a compound script that calculates the BSSE
#    corrected gradient.
#    We achieve this with the compoundFilename
#
# 2. Create a script to run an ORCA calculalation with
#    the external optimizer and the BSSE cprrected
```

(continues on next page)

(continued from previous page)

```
#    gradient. We do that by running a script that runs
#    an ORCA calculation that calculates the gradient
#    and then copy this gradient file back to the expected
#    name
#
# 3. Make a normal ORCA New_Step that calls the external
#    optimizer
#
# NOTE: Depending on the chosen method the property names of
#       myPropName has to be adjusted. For the gradient we do
#       not have this problem because we read the last
#       available in the corresponding property file.
#
# NOTE: Variable baseFilename should have the name of the calling
#       orca input file!
#
# ----------------------    Variables    ------------------------------
# --- Variables to be adjusted (e.g. using 'with' ----------------------
Variable molecule           = "01.xyz";         # xyz file of the initial structure
Variable method             = "BP86";            # The method we use for the calculation
Variable basis              = " ";               # The basis set
Variable restOfInput        = "";                # The rest of the simple input
Variable charge             = 0;                 # Charge
Variable mult               = 1;                 # Spin multiplicity
Variable myPropName         = "SCF_Energy";      # The name of the property for the energy
variable myFilename         = "compoundBSSE";    # Name for the created xyz files
Variable baseFilename       = "run";
Variable gradCreateFile     = "BSSEGradient.cmp";# The compound script that extrapolates the␣
↪gradient
Variable DoOptimization     = false;             # Optimize the monomers or not
Variable produceEnGradFile  = true;              # Produce an ORCA engrad file
Variable enGradFilename     = "result.engrad";   # Filename of the ORCA engrad file
# ---------------------------------------------------------
#
#           Variables for the driver script
Variable createDriverScript = true;            # The shell script driver
Variable driverScript;                         # A script to create the extrapolated energy␣
↪gradient
Variable driverScriptName   = "runningScript";
Variable submitCommand      = "orca";
# ---------------------------------------------------------

# ---------------------------------------------------------
#
#           Variables for the ORCA input
Variable createORCAInput = true;
Variable orcaInput;                           # The ORCA input for the gradient extrapolation
Variable orcaInputName  = "runGradient.inp";
# ---------------------------------------------------------

# -------------------------------------------------
# 1. Maybe Create the necessary driver script
#    for the external optimizer and make it executable
#    NOTE: This will depenend on the operating system
# -------------------------------------------------
if (createDriverScript) then
  driverScript = openFile(driverScriptName, "w");
  write2File(driverScript, "source ~/.bashrc\n");
  write2File(driverScript, "%s %s\n", submitCommand, orcaInputName );
  write2File(driverScript, "cp %s  %s_Compound_1_EXT.engrad\n", engradFilename, baseFilename);
  closeFile(driverScript);
```

(continues on next page)

(continued from previous page)

```
  sys_cmd("chmod +x %s",driverScriptName);
EndIf

# ------------------------------------------------
# 2. Maybe Create the ORCA input that will run the
#    compound script for the gradient extrapolation
# ------------------------------------------------
if (createORCAInput) then
  orcaInput = openFile(orcaInputName, "w");
  Write2File(orcaInput, "%%Compound \"%s\"\n", gradCreateFile);
  Write2File(orcaInput, "  with\n");
  Write2File(orcaInput, "    molecule          =\"%s_Compound_1_EXT.xyz\"\;\n", baseFilename);
  Write2File(orcaInput, "    charge            = %d\;\n",     charge);
  Write2File(orcaInput, "    mutliplicity      = %d\;\n",     mult);
  Write2File(orcaInput, "    method            = \"%s\"\;\n", method);
  Write2File(orcaInput, "    basis             = \"%s\"\;\n", basis);
  Write2File(orcaInput, "    restOfInput       = \"%s\"\;\n", restOfInput);
  Write2File(orcaInput, "    myPropName        = \"%s\"\;\n", myPropName);
  Write2File(orcaInput, "    myFilename        = \"%s\"\;\n", myFilename);
  Write2File(orcaInput, "    removeFiles       = false\;\n");
  Write2File(orcaInput, "    DoOptimization    = %s\;\n",     DoOptimization.GetString());
  Write2File(orcaInput, "    produceEnGradFile = %s\;\n",     produceEnGradFile.GetString());
  Write2File(orcaInput, "    enGradFilename    = \"%s\"\;\n", enGradFilename);
  Write2File(orcaInput, "End\n");
  closeFile(orcaInput);
EndIf

# ------------------------------------------------
# 3. Copy the initial XYZ file to the one needed
#    for the external optimizer
# ------------------------------------------------
sys_cmd("cp %s %s_Compound_1_EXT.xyz", molecule, baseFilename);


# ------------------------------------------------
# 1. Run the driver ORCA input file that calls the
#    External optimizer
# ------------------------------------------------
New_Step
  !ExtOpt Opt
  *xyzfile &{charge} &{mult} &{baseFilename}_Compound_1_EXT.xyz
  %method
    ProgExt "./&{driverScriptName}"
  End
Step_End

End
```

Comments

The initial structure should contain some ghost atoms.

7.57.10 Umbrella script

Introduction

This script calculates the potential for the "umbrella effect" in NH_3. In addition it locates the minima and maxima in the potential surface.

Filename

Umbrella.cmp

SCRIPT

```
# ----------------------------------------------
# Umbrella coordinate mapping for NH3
# Author: Frank Neese
# ----------------------------------------------
variable JobName = "NH3-umbrella";
variable amin    = 50.0;
variable amax    = 130.0;
variable nsteps  = 21;
Variable energies[21];

Variable angle;
Variable JobStep;
Variable JobStep_m;
variable step;

Variable method = "BP86";
Variable basis  = "def2-SVP def2/J";

step  = 1.0*(amax-amin)/(nsteps-1);

# Loop over the number of steps
# -----------------------------
for iang from 0 to nsteps-1 do
  angle    = amin + iang*step;
  JobStep  = iang+1;
  JobStep_m= JobStep-1;
  if (iang>0) then
    Read_Geom(JobStep_m);
    New_step
      ! &{method} &{basis} TightSCF Opt
      %base "&{JobName}.step&{JobStep}"
      %geom constraints
        {A 1 0 2 &{angle} C}
        {A 1 0 3 &{angle} C}
        {A 1 0 4 &{angle} C}
        end
      end

    Step_End
  else
    New_step
      ! &{method} &{basis} TightSCF Opt
      %base "&{JobName}.step&{JobStep}"
      %geom constraints
        {A 1 0 2 &{angle} C}
        {A 1 0 3 &{angle} C}
        {A 1 0 4 &{angle} C}
        end
      end

      * int 0 1
```

(continues on next page)

(continued from previous page)

```
      N 0 0 0 0.0 0.0 0.0
      DA 1 0 0 2.0 0.0 0.0
      H 1 2 0 1.06 &{angle} 0.0
      H 1 2 3 1.06 &{angle} 120.0
      H 1 2 3 1.06 &{angle} 240.0
      *
    Step_End
  endif
   energies[iang].readProperty(propertyName="SCF_ENERGY");
   print(" index: %3d Angle %6.2lf Energy: %16.12lf Eh\n", iang, angle, energies[iang]);
EndFor

# Print a summary at the end of the calculation
# ---------------------------------------------
print("////////////////////////////////////////////////////////\n");
print("// POTENTIAL ENERGY RESULT\n");
print("////////////////////////////////////////////////////////\n");
variable minimum,maximum;
variable Em,E0,Ep;
variable i0,im,ip;
for iang from 0 to nsteps-1 do
  angle   = amin + 1.0*iang*step;
  JobStep = iang+1;
  minimum = 0;
  maximum = 0;
  i0      = iang;
  im      = iang-1;
  ip      = iang+1;
  E0      = energies[i0];
  Em      = E0;
  Ep      = E0;
  if (iang>0 and iang<nsteps-1) then
    Em = energies[im];
    Ep = energies[ip];
  endif
  if (E0<Em and E0<Ep) then minimum=1; endif
  if (E0>Em and E0>Ep) then maximum=1; endif
  if (minimum = 1 ) then
    print(" %3d  %6.2lf %16.12lf (-)\n",JobStep,angle, E0 );
  endif
  if (maximum = 1 ) then
    print(" %3d  %6.2lf %16.12lf (+)\n",JobStep,angle, E0 );
  endif
  if (minimum=0 and maximum=0) then
    print(" %3d  %6.2lf %16.12lf    \n",JobStep,angle, E0 );
  endif
endfor
print("////////////////////////////////////////////////////////\n");

End # end of compound block
```

7.57.11 Multi reference

Introduction

This is a script that calculates the atomic electron denstities in free atoms and makes a library of them.

Filename

atomDensities.inp

SCRIPT

```
# FN 07/2024
#
# A compound script to run calculation on free atoms
# in order to generate a library of electron densities
#
%compound
   Variable Element = {" ",
                        "H",                                    "He",
                        "Li","Be","B" ,"C" ,"N" ,"O" ,"F" ,"Ne"
                        };
   Variable Nact    = {" ",
                        "1" ,                                   "0",
                        "1" ,"0" ,"3" ,"4" ,"5" ,"6" ,"7" ,"0"
                        };

   Variable Norb    = {" ",
                        "1" ,                                   "0",
                        "1" ,"0" ,"4" ,"4" ,"4" ,"4" ,"4" ,"0"
                        };

   Variable Nroots  = {" ",
                        "1" ,                                   "0",
                        "1" ,"0" ,"3" ,"3" ,"1" ,"3" ,"3" ,"0"
                        };

   Variable Charge  = {" ",
                       "0" ,                                "0",
                       "0" ,"0" ,"0" ,"0" ,"0" ,"0" ,"0" ,"0"
                       };

   Variable Mult    =  {" ",
                         "2" ,                                  "1",
                         "2" ,"1" ,"2" ,"3" ,"4" ,"3" ,"2" ,"1"
                         };

   Variable HFTyp   = {" ",
                         "UHF",                                                                  "RHF
↪",
                         "CASSCF","RHF","CASSCF" ,"CASSCF" ,"CASSCF" ,"CASSCF" ,"CASSCF" ,"RHF"
                         };

   Variable el;
   for el from 1 to Element.GetSize()-1 do
      if (HFTyp[el]="CASSCF") then
        New_Step
          ! cc-pVDZ VeryTightSCF Conv
          %base "atom_&{Element[el]}_&{Charge[el]}_&{Mult[el]}"
          %casscf  nel    = &{Nact[el]};
                   norb   = &{Norb[el]};
                   nroots = &{Nroots[el]};
                   mult   = &{Mult[el]};
                   end
```

(continues on next page)

(continued from previous page)

```
         * xyz &{Charge[el]} &{Mult[el]}
          &{Element[el]} 0.0  0.0   0.0
         *
        Step_end
      else
        New_Step
         ! &{HFTyp[el]} cc-pVDZ VeryTightSCF Conv
          %base "atom_&{Element[el]}_&{Charge[el]}_&{Mult[el]}"
         * xyz &{Charge[el]} &{Mult[el]}
          &{Element[el]} 0.0  0.0   0.0
         *
        Step_end
      endif
   endfor

endrun;
```

Comments

Here, it's interesting to note that depending on the selected atom, the script either performs a CASSCF calculation, which provides details such as the number of electrons and number of roots, among other parameters, or it carries out a simple Hartree-Fock calculation.

7.57.12 GoTo

Introduction

This is a brief example demonstrating how the *GoTo* command can be used in *Compound*.

Filename

goTo_Example.inp

SCRIPT

```
# Compound Example on GoTo usage
# Efficient ON/OFF switch

%Compound
  Variable switch="OFF";
  Variable turnOff, turnOn, loopEnd;
  Variable maxIter = 10;
  for i from 0 to maxIter do
    if (switch="ON") then
      GoTo turnOff;
    else
      GoTo turnOn;
    endIf
    turnOff:
      print("Switch: %s\n", switch);
      switch="OFF";
      GoTo loopEnd;
    turnON:
      print("Switch: %s\n", switch);
      switch="ON";
      GoTo loopEnd;
    loopEnd:
  EndFor
End
```

7.58 orca_2json

This utility program supports the exchange of external ORCA data like geometry, orbitals and basisets but also of internal ORCA data like 1-electron and 2-electron integals with other programs.

7.58.1 Export ORCA data

The program reads information like geometries, basis sets, MOs etc. stored in the `.gbw` file or other equivalent ones as `.uno`, `.mp2nat`, `.qro` etc. and calculates integrals to export them in JSON standard output formats. For density information the `.densities` file must also be available. The program is called as a standalone via command line.

Syntax:

```
orca_2json BaseName.gbw -options
or
orca_2json BaseName.mp2nat -options
or
orca_2json BaseName.uno -options
```

The following ASCII and binary JSON-formats are available as command line options. It is possible to specify more than one format option.

```
  -json             Write ASCII JSON file (default)
  -bson             Write binary JSON file
  -ubjson           Write universal binary JSON specification file
  -msgpack          Write MessagePack file
```

In addition two more options are availble. The first of them is used to translate a basename.property.txt file to a corresponding one in JSON format (see *Property File*).

```
   -property        Translate a *.txt property file to a *.json one
```

Finally, orca_2json has the ability to create a .gbw file from a json file. For this one needs to use the '-gbw' option (see *Import JSON data into ORCA*).

```
   -gbw        Create a GBW file from a json one
```

7.58.2 Configuration file

The data stored in the json file can be configured more individually. Some information like atom information, geometry and charge are always written in the outputfile (see *Basic Information*). Other data can be requested via keywords in a JSON-formatted configuration file that either exists for every ORCA output file or for all files in a directory. Without these configuration files **ALL** available data are stored except for the densities and the integrals. Because of the huge amount of data these are only available when explicitly requested in the configuration file.

You can specify a basename-dependend configuration file

```
BaseName.json.conf
```

or a global file used for all requests in a directory.

```
orca.json.conf
```

Most keywords in the configuration file can be activated or deactivated with true or false but some keywords like densities or output formats have more options and require a list of values. If an option is not specifically selected it is omitted.

Structure of the configuration file:

```
{
        "keyword": true/false,
        ...
        "keyword": true/false,
        "keyword": ["option",...."option"]
        ...
}
```

Example

```
{
  "MOCoefficients":false,
  "BasisSet":false
}
```

Using the above configuration file in the working directory, *orca_2json* will not export the molecular coefficients and the basis set information.

Example

Here is an example configuration file with most available keywords where everything is disabled except for the basis set information and the specified Integrals, all densities stored in the density file are requested and the output format should be ascii json and binary json.

```
{
        "MOCoefficients": false,
        "Basisset": true,
        "1elPropertyIntegrals": ["dipole","quadrupole","velosity", "printLinMom", "angular_
↪momentum", "higherMoment"],
        "1elPropertyRelIntegrals": ["dipole","quadrupole"],
        "1elIntegrals": ["H","S", "T", "V", "HMO"],
        "1elIntegralsRel": ["H","S","T", "V"],
        "Vaux": false,
        "AuxBasisType": "AuxC"
        "FullTrafo": false,
        "OrbWin": [0,0,0,0,0,0,0,0].
        "2elIntegrals": [ "MO_IJKL", "RI_IAV", "RI_IJKL"],
        "2elNonRedIntegrals": false,
        "2elNonRedRIIntegrals": false,
        "MullikenCharge": false,
        "LoewdinCharge": false,
        "Densities": ["all"],   <--- here you specify the names like "scfp" or "scfr" or all
        "JSONFormats": ["json", "bson"]
}
```

7.58.3 Available information

Property File

Beside all information that we will see orca_2json can create, it can also translate the property file of ORCA (basename.property.txt) to a JSON file. This option gives access to all properties stored in the property file. (For more information on property file see *Property File*).

The **syntax** is:

```
orca_2json basename -property
```

Where

basename is the name of the property file **without** the extension *property.txt*.

Example

If we use the following ORCA input (with the name test.inp):

```
!HF
*xyz 0 1
  H 0.0 0.0 0.0
  H 0.0 0.0 0.8
*
```

ORCA will create and store on disk, a file named *"test.property.txt"*. The start of the file will look like this:

```
************************************************
****************** ORCA 6.0  *******************
************************************************
$Calculation_Status
   &GeometryIndex 1
   &ListStatus       OUT
   &VERSION [&Type "String"] "6.0"
   &PROGNAME [&Type "String"] "LeanSCF"
   &STATUS [&Type "String"] "NORMAL TERMINATION"
$End
$Geometry
   &GeometryIndex 1
   &ListStatus       OUT
   &NATOMS [&Type "Integer"] 2
   &NCORELESSECP [&Type "Integer"] 0
   &NGHOSTATOMS [&Type "Integer"] 0
   &CartesianCoordinates [&Type "Coordinates", &Dim(2,4), &Units "Bohr"]
              H      0.000000000000    0.000000000000    0.000000000000
              H      0.000000000000    0.000000000000    1.511780907137
$End
$SCF_Energy
   &GeometryIndex 1
   &ListStatus       OUT
   &SCF_ENERGY [&Type "Double"]      -1.1271129220233238e+00
$End
```

Then running *orca_2json* in the following way:

```
orca_2json test -property
```

ORCA will create a new file on disk, named *"test.property.json"*. The start of this file will look like this:

```
{
"Calculation_Status" : {
  "PropertyName" : "Calculation_Status",
  "GeometryIndex" : 1,
  "VERSION" : "6.0" ,
  "PROGNAME" : "LeanSCF" ,
  "STATUS" : "NORMAL TERMINATION"
},
"Geometry_1" : {
"Geometry" : {
  "PropertyName" : "Geometry",
  "GeometryIndex" : 1,
  "NATOMS" : 2 ,
  "NCORELESSECP" : 0 ,
  "NGHOSTATOMS" : 0 ,
  "Coordinates" : {
    "Type":  "Cartesians",
    "Units": "Bohr",
    "Cartesians": [
      ["H ",    0.000000000000,    0.000000000000,    0.000000000000],
      ["H ",    0.000000000000,    0.000000000000,    1.511780907137]
```

(continues on next page)

(continued from previous page)

```
    ]
  }
},
"SCF_Energy" : {
  "PropertyName" : "SCF_Energy",
  "GeometryIndex" : 1,
  "SCF_ENERGY" :      -1.1271129220233238e+00
},
```

Basic Information

Some basic information will always be written into the JSON-file as

per Atom	-	Coords
	-	ElementLabel
	-	ElementNumber
	-	Idx
	-	NuclearCharge
per molecule	-	BaseName
	-	Charge
	-	CoordinateUnits
	-	HFTyp
	-	Multiplicity
	-	PointGroup

For example using the following configuration file where we set everything to false:

```
{
  "MOCoefficients": false,
  "Basisset": false,
  "MullikenCharge": false,
  "LoewdinCharge": false,
  "JSONFormats": ["json"]
}
```

will still produce a json file which for the case of a H2 molecule should look like:

```
{
  "Molecule": {
    "Atoms": [
      {
        "Coords": [
          0.0,
          0.0,
          0.0
```

(continues on next page)

(continued from previous page)


```
        ],
        "ElementLabel": "H",
        "ElementNumber": 1,
        "Idx": 0,
        "NuclearCharge": 1.0
      },
      {
        "Coords": [
          0.0,
          0.0,
          0.8
        ],
        "ElementLabel": "H",
        "ElementNumber": 1,
        "Idx": 1,
        "NuclearCharge": 1.0
      }
    ],
    "BaseName": "test",
    "Charge": 0,
    "CoordinateUnits": "Angs",
    "HFTyp": "RHF",
    "Multiplicity": 1,
    "PointGroup": "C1"
  },
  "ORCA Header": {
    "Date": "2024-06-03 00:06:37 +0200",
    "Git": "548015a5a0",
    "Version": "                          Program Version 6.0  -  CURRENT   -\n"
  }
}
```


Densities

orca_2json can also export calculated densities in json format.

Densities	-	densities as available in the .densities file

Syntax *"Densities" : [list of densities]*

Where *list of densities* should be a list of strings with the expected densities.

NOTE By default densities, due to their potential size, are not exported to a json file.

NOTE An empty bracket syntax (*"Densities" : []*) will cause the program to crash.

NOTE There is the string "All" available where the program will export all available densities.

Electron Integrals

The list of available electron integrals is shown in the next table.

1elIntegrals	-	1-electron integrals
1elPropertyIntegrals	-	1-electron property integrals
1elIntegralsRel	-	relativistic 1-electron integrals
1elPropertyRelIntegrals	-	relativistic 1-electron property integrals
2elIntegrals	-	two-electron integrals
2elNonRedIntegrals	-	non-redundant two-electron integrals
2elNonRedRIIntegrals	-	non-redundant two-electron RI integrals

1-electron integrals

For 1-electron integrals we use the following notation:

H	-	one electron matrix
HMO	-	one electron matrix in MO basis
S	-	overlap matrix
T	-	kinetic energy matrix
V	-	nuclear attraction matrix

Example

```
{
  "1elIntegrals": ["H", "S"],
  "JSONFormats": ["json"]
}
```

will produce a json file were only the H-Matrix and the Overlap matrix are printed (beside the basic information). Please note that for the one electron relativistic integrals there is a separate variable (see *1-electron relativistic integrals*)

1-electron property integrals

Also available are 1-electron property integrals.

1elPropertyIntegrals	-	1-electron property integrals

Currently the following options are valid:

angular_momentum	-	Angular momentum integrals
dipole	-	Dipole moment integrals
higherMoment	-	Octupole moment integrals
quadrupole	-	Quadrupole moment integrals
velocity	-	Velocity integrals

Example

```
{
  "1elPropertyIntegrals": ["dipole", "quadrupole"],
  "JSONFormats": ["json"]
}
```

1-electron relativistic integrals

1-electron relativistic integrals follow the same notation with the corresponding non-relativistic ones (see *1-electron integrals*).

Example

```
{
  "1elIntegralsRel":["H", "S"],
  "JSONFormats": ["json"]
}
```

This example will produce and store in the corresponding json file the relativistic H-Matrix and S-Matrix.

1-electron relativistic property integrals

Also available are 1-electron relativistic property integrals similar to the non-relativistic ones but with reduced options.

1elPropertyRelIntegrals	-	relativistic 1-electron property integrals

Currently the following options are valid:

dipole	-	Dipole moment integrals
quadrupole	-	Quadrupole moment integrals

Example

```
{
    "1elPropertyIntegrals": ["dipole","quadrupole"]
}
```

Origin setting

The origin of the electric property is per default the Cartesian origin but also the center of mass and the center of nuclear charges can be selected. Additionally an arbitrary position can be given as x,y,z coordinates when ori_el = 3 is chosen. Currently the following options are valid:

ori_el	-	0 - Cartesian origin
	-	1 - center of mass
	-	2 - center of nuclear charge
	-	3 - arbitrary position
ori_el_xyz	-	position of the origin(x,y,z)

Example

```
{
    "1elPropertyIntegrals": ["dipole","quadrupole"],
    "ori_el": 3,
    "ori_el_xyz": [0.0, 1.0, 1.0]
}
```

2-electron integrals

ORCA_2json can produce and write on disk three main categories of 2-electron integrals.

1. Two-electron integrals in atomic basis (*2-electron integrals in AO basis*)
2. Two-electron integrals in molecular basis (*2-electron integrals in MO basis*)
3. Two-electron integrals using the resolution of identity approximation (RI). (*RI 2-electron integrals*)

2-electron integrals in AO basis

In atomic basis the two-electron integrals can be saved in Coulomb order or in Exchange order. The keywords for the two options are shown in the next table.

AO_PQRS	-	AO basis integrals in Coulomb order
AO_PRQS	-	AO basis integrals in Exchange order

2-electron integrals in MO basis

In molecular basis, ORCA follows the accepted notation where by I,J,K and L we specify "internal" orbitals, meaning occupied in the reference wavefunction while by A,B,C and D we specify "external" orbitals, meaning orbitals that are empty in the reference wavefunction. With P,Q,R and S we specify all possible orbitals, meaning both "internal" and "external". The available keywords in *"orca_2json"* for the two-electron integrals in the molecular basis are the ones shown in the table below:

MO_IJKL	-	Coulomb 0-external
MO_IJKA	-	Coulomb 1-external
MO_IJAB	-	Coulomb 2-external
MO_IABC	-	Coulomb 3-external
MO_ABCD	-	Coulomb 4-external
MO_PQRS	-	Coulomb ALL integrals
MO_IKJL	-	Exchange 0-external
MO_IKJA	-	Exchange 1-external
MO_IAJB	-	Exchange 2-external
MO_IBAC	-	Exchange 3-external
MO_ACBD	-	Exchange 4-external
MO_PRQS	-	Exchange ALL integrals

Example

```
{
  "2elIntegrals":["MO_IJKL", "MO_ABCD"],
  "Thresh": 1e-8
}
```

RI 2-electron integrals

Using the Resolution of Identity (RI) one can create the integrals in a more efficient way. There are two main categories of RI integrals: the 3-index integrals, where only only half of the transformation has taken place, and the 4-index integrlas where the integrals are totally transformed in the molecular basis. The notation of the integrals follows the one we just described for the two-electron integrals in the molecular basis.

RI_IJV	-	RI 3-index 0-external
RI_IAV	-	RI 3-index 1-external
RI_ABV	-	RI 3-index 2-external
RI_IJKL	-	RI 4-index Coulomb 0-external
RI_IJKA	-	RI 4-index Coulomb 1-external
RI_IJAB	-	RI 4-index Coulomb 2-external
RI_IABC	-	RI 4-index Coulomb 3-external
RI_ABCD	-	RI 4-index Coulomb 4-external
RI_IKJL	-	RI 4-index Exchange 0-external
RI_IKJA	-	RI 4-index Exchange 1-external
RI_IAJB	-	RI 4-index Exchange 2-external
RI_IBAC	-	RI 4-index Exchange 3-external
RI_ACBD	-	RI 4-index Exchange 4-external

In addition to the integrals in case of RI integrals also the used RI Metric is available through the option Vaux:

Vaux	-	true/false

Example

```
{
  "2elIntegrals":["RI_IJKL", "RI_IJV"],
  "Vaux":true
}
```

Full Integral Transformation

The full transformation integrals can be selected via the FullTrafo keyword.

FullTrafo	-	true/false

More 2-electron integrals

Also the non-redundant 2-electron integrals are available for the RI and the nonRI case. Therefore the options 2elNonRedIntegrals or 2elNonRedRIIntegrals must be specified.

2elNonRedIntegrals	-	true/false
2elNonRedRIIntegrals	-	true/false

Orbital Windows, AuxBasisType and Threshold

The orbital window can either be selected automatically by the transformation routine or given by the user via the OrbWin keyword. The internal and external space (i0,i1,a0,a1) is defined via an integer list.

for example:

OrbWin	-	[0,8,9,15]
	-	[0,12,13,30,0,12,13,30]

The default AUX basis type is AuxC but can be changed with the keyword AuxBasisType. Please keep in mind that only those basis types used during the ORCA run can be selected.

AuxBasisType	-	AuxJ
	-	AuxJK
	-	AuxC

To reduce the number of the integrals the keyword Thresh can be used to decrease the selected integrals to save disk space. This effects ONLY the printing and not the accuracy of the generated integrals. The default print threshold is 1.e-15.

Thresh	-	printout threshold (default 1.e-15)

Example

```
{
        "OrbWin": [0,7,8,85,0,0,0,0],
        "Thresh": 1e-8,
        "AuxBasisType": "AuxC",
        "Vaux": true,
        "2elIntegrals": ["RI_IJKL"]
}
```

TDDFT amplitude data (CIS/RPA)

The TDDFT amplitudes and root informations can be requested (no triplet information yet). The available options are:

CIS	-	TDDFT amplitudes
CISNRoots	-	informations of all roots
CISROOT	-	List of roots

Example

```
{
        "CIS": true,
        "CISNRoots": false,
        "CISRoot": [1,4,7,12]
}
```

Example

```
{
        "CIS": true,
        "CISNRoots": true
}
```

JSON Format

The JSON Format that is created can also be defined in the configuration file through the *JSONFormats* variable.

JSONFormats	-	JSON output format

The available JSON formats are:

json	-	ASCII JSON format
bson	-	binary JSON format
ubjson	-	universal binary json format
msgpck	-	MessagePack format

Example

```
{
  "JSONFormats":["json","bson","ubjson","msgpck"]
}
```

MO Coefficients

MOCoefficients	-	molecule orbital information (true\false)

Example

```
{
  "MOCoefficients": true
}
```

7.58.4 Import JSON data into ORCA

Some information like geometry, basis set and Molecular orbitals stored in the json format written by orca_2json can be used to create a new gbw-file to be specified as an orbital file in the orca input. The program is called as a standalone via command line.

```
orca_2json BaseName.json -gbw
```

Basic Information

In order to create a functional gbw-file the following information must be provided:

per Atom	-	Basis
	-	Coords
	-	ElementNumber
	-	NuclearCharge
per molecule	-	Charge
	-	CoordinateUnits
	-	HFTyp
	-	Multiplicity
	-	MolecularOrbitals

NOTE

Please keep in mind that *MolecularOrbitals* is a composite of 2 different components, namely *"EnergyUnit"* and *"MOs"*. Then *"MOs"* contains *"MOCoefficients"*, *"Occupancy"*, *"OrbitalEnergy"*, "OrbitalSymLabel"* and "OrbitalSymmetry"*.

Example

The following file (let's call it *filename.json*) is a json file for H_2 molecule with STO-3G basis set.

```
{
  "Molecule": {
    "Atoms": [
      {"Basis": [
          {"Coefficients": [ 0.1543289707029839,
              0.5353281424384732,0.44463454202535485],
            "Exponents": [3.42525091,0.62391373,0.1688554],
            "Shell": "s"}],
        "Coords": [0.0,0.0,0.0],
        "ElementNumber": 1,
        "NuclearCharge": 1.0},
      {"Basis": [
          {"Coefficients": [0.1543289707029839,
              0.5353281424384732,0.44463454202535485],
            "Exponents": [3.42525091,0.62391373,0.1688554],
            "Shell": "s"}],
        "Coords": [0.0,0.0,0.8],
        "ElementNumber": 1,
        "NuclearCharge": 1.0}],
    "Charge": 0,
    "CoordinateUnits": "Angs",
    "HFTyp": "RHF",
    "MolecularOrbitals": {
      "EnergyUnit": "Eh",
      "MOs": [{"MOCoefficients": [
            -0.5554171364661243,-0.5554171364661241],
          "Occupancy": 2.0,
```

(continues on next page)

(continued from previous page)

```
          "OrbitalEnergy": -0.5544958795933514,
          "OrbitalSymLabel": "A",
          "OrbitalSymmetry": 0},
        {"MOCoefficients": [
            -1.1482994800696493,1.1482994800696493],
          "Occupancy": 0.0,
          "OrbitalEnergy": 0.6126180830925017,
          "OrbitalSymLabel": "A",
          "OrbitalSymmetry": 0}]},
    "Multiplicity": 1}
}
```

running then the command

```
orca_2json filename.json -gbw
```

should create a gbw file that ORCA can read.

Definition of the real solid harmonic Gaussian orbitals

When integrals over real solid harmonic Gaussian orbitals are issued into a JSON file, the precise definition of these orbitals becomes important. ORCA uses its own peculiar conventions for the arrangement and the phases of the individual components of the orbital shells.

The definitions of the angular parts of all shell components up to angular momentum k ($\ell = 8$) are documented below. These correspond to the real solid harmonics $S_{\ell,m}(x,y,z)$ that are normalized the following way:

$$\int_0^\pi \sin\vartheta \, d\vartheta \int_{-\pi}^{\pi} d\varphi \, r^{-2\ell} S^2_{\ell,m}(x,y,z) = 1$$

$R_\ell(r)$ is the common radial part of a shell with angular momentum ℓ that consists of a specific basis set dependent linear combination of Gaussian primitives. It is normalized independently:

$$\int_0^\infty r^{2\ell+2} R^2_\ell(r) \, dr = 1$$

The factors r^2 and $\sin\vartheta$ in these integrals arise from the volume element in spherical polar coordinates:

$$\begin{aligned} x &= r\sin\vartheta\cos\varphi \\ y &= r\sin\vartheta\sin\varphi \\ z &= r\cos\vartheta \\ dx\,dy\,dz &= r^2\sin\vartheta\,dr\,d\vartheta\,d\varphi \end{aligned}$$

Angular momentum s ($\ell = 0$)

$$s = \frac{1}{2\sqrt{\pi}} R_s(r)$$

Angular momentum p ($\ell = 1$)

$$
\begin{aligned}
N_p &= \frac{1}{2}\sqrt{\frac{3}{\pi}} \\
p^{(0)} = p_0 \quad &= N_p \, z \, R_p(r) \\
p^{(1)} = p_{+1} &= N_p \, x \, R_p(r) \\
p^{(2)} = p_{-1} &= N_p \, y \, R_p(r)
\end{aligned}
$$

Angular momentum d ($\ell = 2$)

$$
\begin{aligned}
N_d &= \frac{1}{2}\sqrt{\frac{15}{\pi}} \\
d^{(0)} = d_0 \quad &= \frac{\sqrt{3}}{6} \, N_d \left(3z^2 - r^2\right) R_d(r) \\
d^{(1)} = d_{+1} &= N_d \, xz \, R_d(r) \\
d^{(2)} = d_{-1} &= N_d \, yz \, R_d(r) \\
d^{(3)} = d_{+2} &= N_d \, \frac{x^2 - y^2}{2} \, R_d(r) \\
d^{(4)} = d_{-2} &= N_d \, xy \, R_d(r)
\end{aligned}
$$

Angular momentum f ($\ell = 3$)

$$
\begin{aligned}
N_f &= \frac{1}{2}\sqrt{\frac{105}{\pi}} \\
f^{(0)} = f_0 \quad &= \frac{\sqrt{15}}{30} \, N_f \, z \left(5z^2 - 3r^2\right) R_f(r) \\
f^{(1)} = f_{+1} &= \frac{\sqrt{10}}{20} \, N_f \, x \left(5z^2 - r^2\right) R_f(r) \\
f^{(2)} = f_{-1} &= \frac{\sqrt{10}}{20} \, N_f \, y \left(5z^2 - r^2\right) R_f(r) \\
f^{(3)} = f_{+2} &= \frac{1}{2} \, N_f \left(x^2 - y^2\right) z \, R_f(r) \\
f^{(4)} = f_{-2} &= N_f \, xyz \, R_f(r) \\
f^{(5)} = f_{+3} &= -\frac{\sqrt{6}}{12} \, N_f \, x \left(x^2 - 3y^2\right) R_f(r) \\
f^{(6)} = f_{-3} &= -\frac{\sqrt{6}}{12} \, N_f \, y \left(3x^2 - y^2\right) R_f(r)
\end{aligned}
$$

Angular momentum g ($\ell = 4$)

$$
\begin{aligned}
N_g &= \frac{3}{2}\sqrt{\frac{35}{\pi}} \\
g^{(0)} = g_0 \quad &= \frac{\sqrt{35}}{280}\, N_g \left(35z^4 - 30z^2r^2 + 3r^4\right) R_g(r) \\
g^{(1)} = g_{+1} &= \frac{\sqrt{14}}{28}\, N_g\, xz \left(7z^2 - 3r^2\right) R_g(r) \\
g^{(2)} = g_{-1} &= \frac{\sqrt{14}}{28}\, N_g\, yz \left(7z^2 - 3r^2\right) R_g(r) \\
g^{(3)} = g_{+2} &= \frac{\sqrt{7}}{28}\, N_g \left(x^2 - y^2\right) \left(7z^2 - r^2\right) R_g(r) \\
g^{(4)} = g_{-2} &= \frac{\sqrt{7}}{14}\, N_g\, xy \left(7z^2 - r^2\right) R_g(r) \\
g^{(5)} = g_{+3} &= -\frac{\sqrt{2}}{4}\, N_g\, x \left(x^2 - 3y^2\right) z\, R_g(r) \\
g^{(6)} = g_{-3} &= -\frac{\sqrt{2}}{4}\, N_g\, y \left(3x^2 - y^2\right) z\, R_g(r) \\
g^{(7)} = g_{+4} &= -\frac{1}{8}\, N_g \left(x^4 - 6x^2y^2 + y^4\right) R_g(r) \\
g^{(8)} = g_{-4} &= -\frac{1}{2}\, N_g\, xy \left(x^2 - y^2\right) R_g(r)
\end{aligned}
$$

Angular momentum h ($\ell = 5$)

$$
\begin{aligned}
N_h &= \frac{1}{2}\sqrt{\frac{11}{\pi}} \\
h^{(0)} = h_0 \quad &= \frac{1}{8}\, N_h\, z \left(63z^4 - 70z^2r^2 + 15r^4\right) R_h(r) \\
h^{(1)} = h_{+1} &= \frac{\sqrt{15}}{8}\, N_h\, x \left(21z^4 - 14z^2r^2 + r^4\right) R_h(r) \\
h^{(2)} = h_{-1} &= \frac{\sqrt{15}}{8}\, N_h\, y \left(21z^4 - 14z^2r^2 + r^4\right) R_h(r) \\
h^{(3)} = h_{+2} &= \frac{\sqrt{105}}{4}\, N_h \left(x^2 - y^2\right) z \left(3z^2 - r^2\right) R_h(r) \\
h^{(4)} = h_{-2} &= \frac{\sqrt{105}}{2}\, N_h\, xyz \left(3z^2 - r^2\right) R_h(r) \\
h^{(5)} = h_{+3} &= -\frac{\sqrt{70}}{16}\, N_h\, x \left(x^2 - 3y^2\right) \left(9z^2 - r^2\right) R_h(r) \\
h^{(6)} = h_{-3} &= -\frac{\sqrt{70}}{16}\, N_h\, y \left(3x^2 - y^2\right) \left(9z^2 - r^2\right) R_h(r) \\
h^{(7)} = h_{+4} &= -\frac{3\sqrt{35}}{8}\, N_h \left(x^4 - 6x^2y^2 + y^4\right) z\, R_h(r) \\
h^{(8)} = h_{-4} &= -\frac{3\sqrt{35}}{2}\, N_h\, xy \left(x^2 - y^2\right) z\, R_h(r) \\
h^{(9)} = h_{+5} &= \frac{3\sqrt{14}}{16}\, N_h\, x \left(x^4 - 10x^2y^2 + 5y^4\right) R_h(r) \\
h^{(10)} = h_{-5} &= \frac{3\sqrt{14}}{16}\, N_h\, y \left(5x^4 - 10x^2y^2 + y^4\right) R_h(r)
\end{aligned}
$$

Angular momentum i ($\ell = 6$)

$$
\begin{aligned}
N_i &= \frac{1}{2}\sqrt{\frac{13}{\pi}} \\
i^{(0)} = i_0 \quad &= \frac{1}{16} N_i \left(231z^6 - 315z^4r^2 + 105z^2r^4 - 5r^6\right) R_i(r) \\
i^{(1)} = i_{+1} &= \frac{\sqrt{21}}{8} N_i\, xz \left(33z^4 - 30z^2r^2 + 5r^4\right) R_i(r) \\
i^{(2)} = i_{-1} &= \frac{\sqrt{21}}{8} N_i\, yz \left(33z^4 - 30z^2r^2 + 5r^4\right) R_i(r) \\
i^{(3)} = i_{+2} &= \frac{\sqrt{210}}{32} N_i \left(x^2 - y^2\right) \left(33z^4 - 18z^2r^2 + r^4\right) R_i(r) \\
i^{(4)} = i_{-2} &= \frac{\sqrt{210}}{16} N_i\, xy \left(33z^4 - 18z^2r^2 + r^4\right) R_i(r) \\
i^{(5)} = i_{+3} &= -\frac{\sqrt{210}}{16} N_i\, x \left(x^2 - 3y^2\right) z \left(11z^2 - 3r^2\right) R_i(r) \\
i^{(6)} = i_{-3} &= -\frac{\sqrt{210}}{16} N_i\, y \left(3x^2 - y^2\right) z \left(11z^2 - 3r^2\right) R_i(r) \\
i^{(7)} = i_{+4} &= -\frac{3\sqrt{7}}{16} N_i \left(x^4 - 6x^2y^2 + y^4\right) \left(11z^2 - r^2\right) R_i(r) \\
i^{(8)} = i_{-4} &= -\frac{3\sqrt{7}}{4} N_i\, xy \left(x^2 - y^2\right) \left(11z^2 - r^2\right) R_i(r) \\
i^{(9)} = i_{+5} &= \frac{3\sqrt{154}}{16} N_i\, x \left(x^4 - 10x^2y^2 + 5y^4\right) z\, R_i(r) \\
i^{(10)} = i_{-5} &= \frac{3\sqrt{154}}{16} N_i\, y \left(5x^4 - 10x^2y^2 + y^4\right) z\, R_i(r) \\
i^{(11)} = i_{+6} &= \frac{\sqrt{462}}{32} N_i \left(x^2 - y^2\right) \left(x^4 - 14x^2y^2 + y^4\right) R_i(r) \\
i^{(12)} = i_{-6} &= \frac{\sqrt{462}}{16} N_i\, xy \left(3x^4 - 10x^2y^2 + 3y^4\right) R_i(r)
\end{aligned}
$$

Angular momentum j ($\ell = 7$)

$$
\begin{aligned}
N_j &= \frac{1}{2}\sqrt{\frac{15}{\pi}} \\
j^{(0)} = j_0 \quad &= \frac{1}{16}\, N_j\, z \left(429z^6 - 693z^4r^2 + 315z^2r^4 - 35r^6\right) R_j(r) \\
j^{(1)} = j_{+1} &= \frac{\sqrt{7}}{32}\, N_j\, x \left(429z^6 - 495z^4r^2 + 135z^2r^4 - 5r^6\right) R_j(r) \\
j^{(2)} = j_{-1} &= \frac{\sqrt{7}}{32}\, N_j\, y \left(429z^6 - 495z^4r^2 + 135z^2r^4 - 5r^6\right) R_j(r) \\
j^{(3)} = j_{+2} &= \frac{\sqrt{42}}{32}\, N_j \left(x^2 - y^2\right) z \left(143z^4 - 110z^2r^2 + 15r^4\right) R_j(r) \\
j^{(4)} = j_{-2} &= \frac{\sqrt{42}}{16}\, N_j\, xyz \left(143z^4 - 110z^2r^2 + 15r^4\right) R_j(r) \\
j^{(5)} = j_{+3} &= -\frac{\sqrt{21}}{32}\, N_j\, x \left(x^2 - 3y^2\right) \left(143z^4 - 66z^2r^2 + 3r^4\right) R_j(r) \\
j^{(6)} = j_{-3} &= -\frac{\sqrt{21}}{32}\, N_j\, y \left(3x^2 - y^2\right) \left(143z^4 - 66z^2r^2 + 3r^4\right) R_j(r) \\
j^{(7)} = j_{+4} &= -\frac{\sqrt{231}}{16}\, N_j \left(x^4 - 6x^2y^2 + y^4\right) z \left(13z^2 - 3r^2\right) R_j(r) \\
j^{(8)} = j_{-4} &= -\frac{\sqrt{231}}{4}\, N_j\, xy \left(x^2 - y^2\right) z \left(13z^2 - 3r^2\right) R_j(r) \\
j^{(9)} = j_{+5} &= \frac{\sqrt{231}}{32}\, N_j\, x \left(x^4 - 10x^2y^2 + 5y^4\right) \left(13z^2 - r^2\right) R_j(r) \\
j^{(10)} = j_{-5} &= \frac{\sqrt{231}}{32}\, N_j\, y \left(5x^4 - 10x^2y^2 + y^4\right) \left(13z^2 - r^2\right) R_j(r) \\
j^{(11)} = j_{+6} &= \frac{\sqrt{6006}}{32}\, N_j \left(x^2 - y^2\right) \left(x^4 - 14x^2y^2 + y^4\right) z\, R_j(r) \\
j^{(12)} = j_{-6} &= \frac{\sqrt{6006}}{16}\, N_j\, xy \left(3x^2 - y^2\right) \left(x^2 - 3y^2\right) z\, R_j(r) \\
j^{(13)} = j_{+7} &= -\frac{\sqrt{429}}{32}\, N_j\, x \left(x^6 - 21x^4y^2 + 35x^2y^4 - 7y^6\right) R_j(r) \\
j^{(14)} = j_{-7} &= -\frac{\sqrt{429}}{32}\, N_j\, y \left(7x^6 - 35x^4y^2 + 21x^2y^4 - y^6\right) R_j(r)
\end{aligned}
$$

Angular momentum k ($\ell = 8$)

$$
\begin{aligned}
N_k &= \frac{1}{2}\sqrt{\frac{17}{\pi}} \\
k^{(0)} = k_0 \quad &= \frac{1}{128} N_k \left(6435z^8 - 12012z^6r^2 + 6930z^4r^4 - 1260z^2r^6 + 35r^8\right) R_k(r) \\
k^{(1)} = k_{+1} &= \frac{3}{32} N_k\, xz \left(715z^6 - 1001z^4r^2 + 385z^2r^4 - 35r^6\right) R_k(r) \\
k^{(2)} = k_{-1} &= \frac{3}{32} N_k\, yz \left(715z^6 - 1001z^4r^2 + 385z^2r^4 - 35r^6\right) R_k(r) \\
k^{(3)} = k_{+2} &= \frac{3\sqrt{70}}{64} N_k \left(x^2 - y^2\right) \left(143z^6 - 143z^4r^2 + 33z^2r^4 - r^6\right) R_k(r) \\
k^{(4)} = k_{-2} &= \frac{3\sqrt{70}}{32} N_k\, xy \left(143z^6 - 143z^4r^2 + 33z^2r^4 - r^6\right) R_k(r) \\
k^{(5)} = k_{+3} &= -\frac{\sqrt{1155}}{32} N_k\, x \left(x^2 - 3y^2\right) z \left(39z^4 - 26z^2r^2 + 3r^4\right) R_k(r) \\
k^{(6)} = k_{-3} &= -\frac{\sqrt{1155}}{32} N_k\, y \left(3x^2 - y^2\right) z \left(39z^4 - 26z^2r^2 + 3r^4\right) R_k(r) \\
k^{(7)} = k_{+4} &= -\frac{3\sqrt{77}}{64} N_k \left(x^4 - 6x^2y^2 + y^4\right) \left(65z^4 - 26z^2r^2 + r^4\right) R_k(r) \\
k^{(8)} = k_{-4} &= -\frac{3\sqrt{77}}{16} N_k\, xy \left(x^2 - y^2\right) \left(65z^4 - 26z^2r^2 + r^4\right) R_k(r) \\
k^{(9)} = k_{+5} &= \frac{3\sqrt{1001}}{32} N_k\, x \left(x^4 - 10x^2y^2 + 5y^4\right) z \left(5z^2 - r^2\right) R_k(r) \\
k^{(10)} = k_{-5} &= \frac{3\sqrt{1001}}{32} N_k\, y \left(5x^4 - 10x^2y^2 + y^4\right) z \left(5z^2 - r^2\right) R_k(r) \\
k^{(11)} = k_{+6} &= \frac{\sqrt{858}}{64} N_k \left(x^2 - y^2\right) \left(x^4 - 14x^2y^2 + y^4\right) \left(15z^2 - r^2\right) R_k(r) \\
k^{(12)} = k_{-6} &= \frac{\sqrt{858}}{32} N_k\, xy \left(x^2 - 3y^2\right) \left(3x^2 - y^2\right) \left(15z^2 - r^2\right) R_k(r) \\
k^{(13)} = k_{+7} &= -\frac{3\sqrt{715}}{32} N_k\, x \left(x^6 - 21x^4y^2 + 35x^2y^4 - 7y^6\right) z\, R_k(r) \\
k^{(14)} = k_{-7} &= -\frac{3\sqrt{715}}{32} N_k\, y \left(7x^6 - 35x^4y^2 + 21x^2y^4 - y^6\right) z\, R_k(r) \\
k^{(15)} = k_{+8} &= -\frac{3\sqrt{715}}{128} N_k \left(x^8 - 28x^6y^2 + 70x^4y^4 - 28x^2y^6 + y^8\right) R_k(r) \\
k^{(16)} = k_{-8} &= -\frac{3\sqrt{715}}{16} N_k\, xy \left(x^2 - y^2\right) \left(x^4 - 6x^2y^2 + y^4\right) R_k(r)
\end{aligned}
$$

7.59 Property File

One of the files that ORCA produces, during a calculation, is the *property file*. The name of the file is *basename.property.txt*, where basename is the basename of the input file. As we will see later ORCA can also produce the property file with the extension *.json*.

Property file has mainly two usages in ORCA. The first usage is to work as basis for the *Compound* scripting language. *Compound* reads all its information concerning properties through the property file and not through parsing of the ORCA output. The second usage of *property file* is to make it easier for other programs, or potential GUIs, to create interfaces with ORCA.

The advantage of *Property file* compared to the normal ORCA output is that it's syntax is well defined and so it is easier to parse it.

7.59.1 txt format

After any ORCA calculation a property file is created with the extension *.property.txt*. The file is a text file and one can read it and edit it with any available text editor. Below we will analyze the syntax of the file.

The file always starts with the following three lines:

```
*****************************************************
****************** ORCA 6.0.x ******************
*****************************************************
```

where, obviously the version of ORCA changes.

Then, the file consists of a list of properties. Each property starts with the symbol "*$*" followed by the name of the property and ends with the symbol "*$*" followed by "*End*".

For example:

```
$SCF_Energy
   &GeometryIndex 1
   &ListStatus       OUT
   &SCF_ENERGY [&Type "Double"]      -1.1271129230772137e+00
$End
```

Each property consists of components. Each component starts with the symbol "*&*" and has not ending symbol.

For example:

```
   &SCF_ENERGY [&Type "Double"]      -1.1271129230772137e+00
```

Before proceeding to details about the property specific components, there are two components that exist in every property and always in the same order.

The first one is the *GeometryIndex* component. The syntax for this is quite simple, the normal "*&*" component start symbol, followed by "*GeometryIndex*" and then an integer.

For example:

```
&GeometryIndex 1
```

This is an easy way to know the geometry that the current property belongs to.

Then for all properties follows the *ListStatus* component. The syntax for this component is again the "*&*" component start symbol, followed by "*ListStatus*" follow by one the following 5 options:

1. "*IN*" when the property is inside a list properties
2. "*OUT*" when the property is not inside a list of properties
3. "*FIRST*" when the property is the first in a list of properties
4. "*LAST*" when the property is the last in a list of properties
5. "*UNIQUE*" when the property is the only one in a list of properties

For example:

```
 &ListStatus       OUT
```

Then for each property follows a series of *components*. Each component has the following syntax:

First the start of component symbol "*&*".

Then follows the name of the component.

Then a bracket opens with various bracket information about the component. For details on the syntax of the bracket information please check *Bracket information*.

After the bracket there are different options.

If the type is a “Double” or an “Integer” then a number of the appropriate type is expected.

For example:

```
&NATOMS [&Type "Integer"] 2
```

If the type is a “String” then a string is expected starting with quotation marks.

For example:

```
&PROGNAME [&Type "String"] "ProgMDCI"
```

Finally if the type is a kind or Array then, unless there is a comment, that is enclosed inside quotation marks, an array is written starting from the next line. We should note here that after the column header there is an empty line. In addition there is always a first column with an integer giving just the row of the array.

For example:

```
   &DIPOLETOTAL [&Type "ArrayOfDoubles", &Dim (3,1)] "Total"
                                                          0

0                                      0.0000000000000000e+00
1                                      0.0000000000000000e+00
2                                     -5.1833121128553384e-12
```

Bracket information

Bracket information is a list of information separated with ‘,’. The first and most important bracket component is the “*Type*”. Type can be one of the following:

1. “Double”
2. “Integer
3. “Boolean”
4. “String”
5. “ArrayOfDoubles”
6. “ArrayOfIntegers”
7. “ArrayOfBooleans”
8. “MixedMatrix”
9. “Coordinates”

For example:

```
&NATOMS [&Type "Integer"] 2
```

Then in case the “*Type*” is a kind of array the bracket must contain the dimension of the array, using the “*Dim*” component.

For example:

```
&ATNO [&Type "ArrayOfIntegers", &Dim (2,1)]
```

7.59.2 JSON format

The property file can be also produced in a JSON format. Internally this happens through transformation of the *txt format* to JSON format. There are two ways to create a JSON property file.

The first way is through the normal ORCA input using the *WriteJSONPropertyFile* command.

```
%Method
  WriteJSONPropertyfile True
End
```

this will create a *basename.property.txt* and in addition a *basename.property.json* file.

The second way is through the ORCA_2JSON command. For this one first has to run a normal ORCA input, that will create a *basename.property.txt* file, and then use the command:

```
orca_2json basename property
```

where *basename* is the name of the ORCA input.

Property File in JSON format is readable by any JSON library.

CHAPTER

EIGHT

SOME TIPS AND TRICKS

8.1 Input

For calculations on open-shell systems we recommend to use the keywords `!UNO !UCO` in the input line. This will generate quasi-restricted molecular orbitals `QRO`, unrestricted natural spin-orbitals `UNSO`, unrestricted natural orbitals `UNO` and unrestricted corresponding orbitals `UCO`. Moreover, it will print the UCO overlaps in the output, which can provide very clear information about the spin-coupling in the system. Below an example of the input and section of the output is provided.

```
!B3LYP def2-SVP UNO UCO TightSCF
```

The UCO overlap section in the output will look like:

```
***UHF Corresponding Orbitals were saved in MyJob.uco***

----------------------
Orbital    Overlap(*)
----------------------
.
.
.
96:        0.99968
97:        0.99955
98:        0.99947
99:        0.99910
100:       0.99873
101:       0.99563
102:       0.74329
103:       0.00000
```

The overlap corresponds to a value usually less than 0.85 denotes a spin-coupled pair. Whereas, values close to 1.00 and 0.00 refers to the doubly occupied and singly occupied orbitals respectively.

8.2 Cost versus Accuracy

A difficult but important subject in electronic structure theory is to balance the price/accuracy ratio of the calculations. This ratio is governed by: (a) the method used, (b) the basis set used and (c) the cutoffs and tolerances used. There are certainly differing opinions among scientists and I merely quote a few general, subjective points:

- Calculations with minimal basis sets are always unreliable and are only good for explorations. This is also true for small split-valence basis sets like 3-21G, 3-21GSP and perhaps also 4-22GSP. These basis sets are significantly more reliable than their minimal basis counterparts but are not capable of delivering quantitatively reliable results. They may, however, be the only choice if very large molecules are targeted.
- In our own research we almost exclusively use the basis sets of the Karlsruhe group for non-relativistic calculations. They have been updated to the "def2" set that is more consistent than the older basis sets.

- Def2-SV(P) is the smallest and computationally efficient split-valence basis set and is largely identical to the old SV(P), except for the transition metals which have more consistent polarization sets.
- Def2-TZVP is different from the old TZVP. It has been realized that if one invests into an accurate triple-zeta description of the valence region it makes limited sense to only employ a single polarization function. The accuracy is then limited by the polarization set and is not much better than what one gets from SV(P). Hence, def2-TZVP contains a single p-set for hydrogens but is otherwise very similar to the old TZVPP basis set, e.g. it contains 2d1f polarization for main group elements and much more extensive polarization sets for transition metals. The highest polarization function (f for main group) does add substantially to the computational effort. Hence, we often use def2-TZVP without the f polarization function. In order to do that one can use the keyword def2-TZVP(-f). Together with RI or RIJCOSX this is still computationally economic enough for most studies.
- Def2-TZVPP is a fully consistent triple-zeta basis set that provides excellent accuracy for SCF calculations (HF and DFT) and is still pretty good for correlated calculations. It is a good basis set to provide final single point energies.
- Def2-QZVPP is a high accuracy basis set for all kinds of calculations. It provides SCF energies near the basis set limit and correlation energies that are also excellent. It is computationally expensive but with RI and RIJCOSX in conjunction with parallelization it can often still be applied for final single-point energy calculations. In conjunction with such large basis sets one should also increase the accuracy of the integration grids in DFT and RIJCOSX — it would be a shame to limit the accuracy of otherwise very accurate calculations by numerical noise due to the grid.
- Correlation consistent basis sets provide good correlation energies but poor to very poor SCF energies. For the same size, the ano-pVDZ basis sets are much more accurate but are also computationally more expensive. Except for systematic basis set extrapolation we see little reason to use the cc bases.
- Pople basis sets are somewhat old fashioned and also much less consistent across the periodic table than the basis from the Karlsruhe group. Hence, we generally prefer the latter.
- For scalar relativistic calculations (ZORA and DKH) we strongly recommend to use the SARC bases in conjunction with the ZORA or DKH recontractions of the Karlsruhe bases. They are also flexible enough in the core region for general purpose and spectroscopic applications.
- Effective core potentials lead to some savings (but not necessarily spectacular ones) compared to all-electron relativistic calculations. For accurate results, small core ECPs should be used. They are generally available for the def2 Karlsruhe type basis sets for elements past krypton. In general we prefer Stuttgart–Dresden ECPs over LANL ones. For the first transition row, the choices are more meager. Here Karlsruhe basis sets do not exist in conjunction with ECPs and you are bound to either SDD or LANL of which we recommend the former. Geometries and energies are usually good from ECPs, but for property calculations we strongly recommend to switch to all electron scalar relativistic calculations using ZORA (magnetic properties) or DKH (electric properties).
- You can take advantage of a built-in basis set (printed using `!PrintBasis` or `orca_exportbasis`) and then modify it by uncontracting primitives, adding steeper functions etc. (fully uncontracted bases are generated via `uncontract` in `%basis`) Alternatively, some basis sets exist that are of at least double-zeta quality in the core region including the DZP and Dunning basis sets. For higher accuracy you may want to consider the `aug-` series of basis sets. See section *Choice of Basis Set* for more about basis set input.
- Likewise, if you are doing calculations on anions in the gas phase it is advisable to include diffuse functions in the basis set. Having these diffuse functions, however, makes things much more difficult as the locality of the basis set is significantly reduced. If these functions are included it is advisable to choose a small value for `Thresh` (10^{-12} or lower). This is automatically done if the smallest eigenvalue of the overlap matrix is below DiffSThresh (which is 1e-6 by default). Also, diffuse functions tend to introduce basis set linear dependency issues, which can be solved by setting `Sthresh` to a larger value than the default 10^{-7} (see Section *Linear Dependence*). Any value of `Sthresh` beyond 1e-6 has to be used carefully, specially if one is running geometry optimizations, were different basis might be cut off during different geometry steps, or when comparing different conformers since there could be some discontinuity on the final basis set.
- The integration grids used in DFT should be viewed together with the basis set. If large basis set calculations are converged to high accuracy it is advisable to also use large DFT integration grids (like `! DEFGRID3`). For "unlimited" accuracy (i.e. benchmark calculations) it is probably best to use product grids (`Grid=0`)

with a large value for IntAcc (perhaps around 6.0). The default grids have been chosen such that they provide adequate accuracy at the lowest possible computational cost, but for all-electron calculations on heavy elements in conjunction with scalar relativistic Hamiltonians you should examine the grid dependency very carefully and adjust these parameters accordingly to minimize errors. You should be aware that for large molecules the exchange-correlation integration is usually *not* the dominating factor (not even in combination with RI-J).

- Similarly important is the value of Thresh that will largely determine the tunaround time for direct SCF calculations. It may be possible to go to values of 10^{-6}–10^{-8} which will result in large speed-ups. However, the error in the final energy may then be 3 orders of magnitude larger than the cutoff or, sometimes, your calculation will fail to converge, due to the limited integral accuracy. In general it will not be possible to converge a direct SCF calculation to better than Thresh (the program will also not allow this). For higher accuracy values of maybe 10^{-10}–10^{-12} may be used with larger molecules requiring smaller cutoffs. In cases where the SCF is almost converged but then fails to finally converge (which is very annoying) decreasing Thresh and switch to TRAH SCF is recommended. In general, TCut should be around 0.01 $\times$Thresh in order to be on the safe side.
- DFT calculations have many good features and in many cases they produce reliable results. In particular if you study organic molecules it is nevertheless a good idea to check on your DFT results using MP2. MP2 in the form of RI-MP2 is usually affordable and produces reliable results (in particular for weaker interactions where DFT is less accurate). In case of a large mismatch between the MP2 and DFT results the alarm rings — in many such cases MP2 is the better choice, but in others (e.g. for redox processes or transition metal systems) it is not. Remember that SCS-MP2 (RI-SCS-MP2) and double hybrid functionals will usually produce more accurate results than MP2 itself.
- Coupled-cluster calculations become more and more feasible and should be used whenever possible. The LPNO-CCSD, DLPNO-CCSD and DLPNO-CCSD(T) calculations are available for single-point calculations and provide accurate results. However, a coupled-cluster study does require careful study of basis set effects because convergence to the basis set limit is very slow. The established basis set extrapolation schemes may be very helpful here. For open-shell molecules and in particular for transition metals one cannot be careful enough with the reference. You have to carefully check that the Hartree-Fock calculation converged to the desired state in order to get coupled-cluster results that are meaningful. Orbital optimized MP2, CASSCF or DFT orbitals may help but we have often encountered convergence difficulties in the coupled-cluster equations with such choices.
- Generally speaking, CEPA is often better than CCSD and approaches the quality of CCSD(T). It is, however, also a little less robust than CC methods because of the less rigorous treatment of the single excitations in relation to electronic relaxation.
- Don't forget: "Computers don't solve problems – people do". Not denying the importance and desire to obtain accurate numbers: don't forget that in the end it is the molecule and its chemistry or spectroscopy that we want to learn something about. The fact that you may be able to compute one or the other number a little more accurate doesn't mean that this helps us understanding the physics and chemistry of our target system any better. The danger of getting locked into technicalities and miss the desired insight is real!

8.3 Converging SCF Calculations

Despite all efforts you may still find molecules where SCF convergence is poor. These are almost invariably related to open-shell situations and the answer is almost always to provide "better" starting orbitals. Here is my standard strategy to deal with this (assuming a DFT calculation):

- Perform a small basis set (SV) calculation in using the LSD or BP functional and RI approximation with a cheap auxiliary basis set. Set Convergence=Loose and MaxIter=200 or so. The key point is to use a large damping factor and damp until the DIIS comes into a domain of convergence. This is accomplished by SlowConv or even VerySlowConv. If you have an even more pathological case you may need to set DampFac larger and DampErr smaller than chosen by these defaults. This calculation is quite crude and may take many cycles to converge. It will however be rather quick in terms of wall clock time. If the DIIS gets stuck at some error 0.001 or so the SOSCF (or even better TRAH) could be put in operation from this point on.

- Use the orbitals of this calculation and `GuessMode=CMatrix` to start a calculation with the target basis set. In DFT we normally use a pure GGA functional (e.g. BP86). This calculation normally converges relatively smoothly.
- Use the target functional, grid etc. to get the final calculation converged. In many cases this should converge fairly well now.

Here are a few other things that can be tried:

- Try to start from the orbitals of a related closed-shell species. In general closed-shell MO calculations tend to converge better. You then hope to reach the convergence radius of another converger for the open-shell case.
- Try to start from the orbitals of a more positive cation. Cation calculations tend to converge better.
- Try to start from a calculation with a smaller basis set. Smaller basis sets converge better. Then you have the choice of `GuessMode=CMatrix` or `GuessMode=FMatrix` which will affect the convergence behavior.
- Use large level shifts. This increases the number of iterations but stabilizes the converger. (`shift shift 0.5 erroff 0 end`)
- If you are doing DFT calculations try to start from a Hartree-Fock solution for your molecule. HF calculations tend to converge somewhat better because they have a larger HOMO-LUMO gap (there are of course exceptions).
- Carefully look at the starting orbitals (`Print[P_GuessOrb]=1`) and see if they make sense for your molecule. Perhaps you have to reorder them (using `Rotate`) to obtain smooth convergence.
- Most of the time the convergence problems come from "unreasonable" structures. Did you make sure that your coordinates are in the correct units (Angström or Bohrs?) and have been correctly recognized as such by the program?
- If you have trouble with UHF calculations try ROHF (especially SAHF or CAHF) first and then go to the UHF calculation.
- Fool around with `Guess=Hueckel`, `PAtom` or even `HCore`.
- It may sometimes be better to converge to an undesired state and then take the orbitals of this state, reorder them (using `Rotate`) and try to converge to the desired state.
- Similarly, bad orbitals may be manipulated using the SCF stability analysis (section *SCF Stability Analysis*) to provide a new guess.
- Try to start the calculation with a large damping factor (`DampFac=0.90`; or even larger) and specify a relatively small DIIS error to turn damping off (say `DampErr=0.02`;). This will increase the number of cycles but may guide you into a regime were the calculation actually converges.
- The advices above mostly apply to Hartree-Fock and DFT. For CASSCF, the available options and how they can aid to overcome convergence problems are described in the CASSCF manual section. In many cases modifying the initial guess or adding a level shift will help. Do not hesitate to use large level-shifts (e.g 2.0 or even 3.0). The manual is accompanied by CASSCF tutorial that goes through many details of the process including practical advices on convergence. The choice of initial guess is crucial. Some guesses work better for organic molecules while others excel for transition-metal complexes. The tutorial therefore discusses various initial guess options available in ORCA.
- If nothing else helps, stop, grab a friend and go to the next pub (you can also send me an unfriendly e-mail but this will likely not make your calculation converge any quicker; ◡̈).

8.4 Choice of Theoretical Method

The array of available functionals makes it perhaps difficult to decide which one should be used. While this is a matter of ongoing research and, in the end, can only be answered by experimentation and comparison to experimental results or high-level *ab initio* calculations, I may attempt to give some guidelines.

The simplest density functionals (and in general the least accurate) are the local functionals (`Functional=LSD`). Although several variants of the local DFT exist in ORCA there is little to choose among them — they give more or less the same result.

The gradient corrected functionals are (very slightly) more expensive because the gradient of the electron density at each point in space must be computed, but they are also significantly more accurate for structures and energetics of molecules. The various gradient corrected functionals (GGA functionals) are generally similar in their behavior. The BP functional is probably the most widely used in transition metal chemistry. The BLYP, PBE or PW91 functionals may also be considered. PWP has been shown to be rather good for hyperfine coupling predictions of light nuclei in radicals. In addition, since no Hartree-Fock exchange is used you have the ability to speed up the calculation by a factor of 4–40 if the RI approximation is employed. This approximation is really advisable for the LSD and GGA functionals since it leads to very little or no loss in accuracy while giving large speedups. It is, in fact, automatically chosen to be operative when you use pure functionals.

In addition, meta-GGAs (TPSS) are available in ORCA and may provide superior results for certain properties compared to standard GGAs. They are somewhat but not much more expensive to evaluate than standard GGAs.

For many properties (but not necessarily for geometries), more accurate results are usually given by the hybrid density functionals that incorporate part of the HF exchange. The computational effort for these is higher than for the GGA functionals because the HF exchange needs to be computed exactly. Very large speedups result if this is done via the RIJCOSX approximation. Nevertheless for energetics, properties and for predictions of charge and spin densities the hybrids appear to be the best choice. The prototype functional of this kind is B3LYP, which has been very widely used throughout chemistry and is successful for a wide range of molecular properties. Other hybrids have been less well tested but maybe a good choice in specific situations, for example the PBE0 functional has been advertised for NMR chemical shift predictions and other properties. From my personal experience I can also recommend PBE0 and PWP1 as two functionals that give good predictions for EPR g-values and hyperfine couplings. The TPSSh meta-GGA hybrid is also very succesful in this area.[1]

Together with DFT, it is often observed that the atom-pairwise dispersion correction of Stefan Grimme (DFT-D3, and especially the newer DFT-D4) substantially improves the results at no extra cost.

Don't forget that in present days the MP2 method becomes affordable for molecules of significant size and there are quite a number of instances where MP2 (and particularly SCS-MP2) will do significantly better than DFT even if it takes a little longer (the RI approximation is also highly recommended here). The perturbatively corrected functionals (B2PLYP) may also be a very good choice for many problems (at comparable cost to MP2; note that even for large molecules with more than 1000 basis functions the MP2 correction only takes about 10-20% of the time required for the preceding SCF calculation if the RI approximation is invoked. For even larger molecules one has the option of speeding up the MP2 part even further by the DLPNO approximation).

Beyond DFT and (SCS-)MP2 there are coupled-cluster methods and their implementation in ORCA is efficient. With the local pair natural orbital methods you can even study molecules of substantial size and with appealing turnaround times.

When to go to multireference methods is a more complicated question. Typically, this will be the case if multiplets are desired, pure spin functions for systems with several unpaired electrons, in bond breaking situations or for certain classes of excited states (loosely speaking: whenever there are weakly interacting electrons in the system). However, whenever you decide to do so, please be aware that this require substantial insight into the physics and chemistry of the problem at hand. An uneducated use of CASSCF or MRCI/MRPT method likely yields numbers that are nonsensical and that at tremendous computational cost. Here, there is no substitute for experience (and patience ⌣̈).

[1] Some researchers like to adjust the amount of Hartree-Fock exchange according to their needs or what they think is "better" than the standard. This increases the semiempirical character of the calculations and may represent fixes that only work for a given class of compounds and/or properties while worsening the results for others. With this caveat in mind it is one of the things that you are free to try if you like it. However, we do not recommend it since it will deteriorate the comparability of your results with those of other workers the vast majority of which use standard functionals. An alternative to changing the amount of HF exchange could be to simply construct a linear regression for a number of known cases and then use the linear regression.

APPENDIX

A

PUBLICATIONS RELATED TO ORCA

The generic references for ORCA are:

- Frank Neese. The orca program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2(1):73–78, 2012. doi:http://doi.wiley.com/10.1002/wcms.81.
- Frank Neese. Software update: the orca program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 8(1):e1327, 2018. doi:http://doi.wiley.com/10.1002/wcms.1327.
- Frank Neese, Frank Wennmohs, Ute Becker, and Christoph Riplinger. The orca quantum chemistry program package. *J. Chem. Phys.*, 152(22):224108, 2020. doi:https://aip.scitation.org/doi/10.1063/5.0004608.

Please do not only cite the above generic reference, but also cite in addition the original papers that report the development and implementation of the methods you have used in your studies! The following publications describe functionality implemented in . We would highly appreciate if you cite them when you use the program.

A.1 Method development

A.1.1 2017

1. Achintya Kumar Dutta, Frank Neese, and Róbert Izsák. A simple scheme for calculating approximate transition moments within the equation of motion expectation value formalism. *J. Chem. Phys.*, 146(21):214111, 2017.
2. Achintya Kumar Dutta, Marcel Nooijen, Frank Neese, and Róbert Izsák. Automatic active space selection for the similarity transformed equations of motion coupled cluster method. *J. Chem. Phys.*, 146(7):074103, 2017.
3. Stefan Grimme, Christoph Bannwarth, Sebastian Dohm, Andreas Hansen, Jana Pisarek, Philipp Pracht, Jakob Seibert, and Frank Neese. Fully automated quantum-chemistry-based computation of Spin–Spin-Coupled nuclear magnetic resonance spectra. *Angew. Chem. Int. Ed.*, 56(46):14763–14769, 2017. doi:10.1002/anie.201708266.
4. Yang Guo, Kantharuban Sivalingam, Edward F. Valeev, and Frank Neese. Explicitly correlated N-electron valence state perturbation theory (NEVPT2-F12). *J. Chem. Phys.*, 147(6):064110, 08 2017. doi:10.1063/1.4996560.
5. L .M. J. Huntington, M. Krupička, F. Neese, and R. Izsák. Similarity transformed equation of motion coupled-cluster theory based on an unrestricted hartree-fock reference for applications to high-spin open-shell systems. *J. Chem. Phys.*, 147:174104, 2017.
6. Jaroslaw Kalinowski, Frank Wennmohs, and Frank Neese. Arbitrary angular momentum electron repulsion integrals with graphical processing units: application to the resolution of identity hartree–fock method. *J. Chem. Theory Comput.*, 13(7):3160–3170, 2017.
7. M. Krupička, K. Sivalingam, L. Huntington, A. A. Auer, and F. Neese. A toolchain for the automatic generation of computer codes for correlated wavefunction calculations. *J. Comput. Chem.*, 38:1853, 2017.

8. Dimitrios Maganas, Serena DeBeer, and Frank Neese. A restricted open configuration interaction with singles method to calculate valence-to-core resonant x-ray emission spectra: a case study. *Inorg. Chem.*, 56(19):11819–11836, 2017.

9. Sebastian Mai, Felix Plasser, Mathias Pabst, Frank Neese, Andreas Köhn, and Leticia González. Surface hopping dynamics including intersystem crossing using the algebraic diagrammatic construction method. *J. Chem. Phys.*, 147(18):184109, 2017.

10. Shubhrodeep Pathak, Lucas Lang, and Frank Neese. A dynamic correlation dressed complete active space method: Theory, implementation, and preliminary applications. *J. Chem. Phys.*, 147:234109, 2017.

11. Fabijan Pavošević, Chong Peng, Peter Pinski, Christoph Riplinger, Frank Neese, and Edward F Valeev. Sparsemaps—a systematic infrastructure for reduced scaling electronic structure methods. v. linear scaling explicitly correlated coupled-cluster method with pair natural orbitals. *J. Chem. Phys.*, 146(17):174108, 2017.

12. M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, and F. Neese. *J. Chem. Phys.*, 146:164105, 2017.

13. Manuel Sparta, Marius Retegan, Peter Pinski, Christoph Riplinger, Ute Becker, and Frank Neese. Multilevel approaches within the local pair natural orbital framework. *J. Chem. Theory Comput.*, 13(7):3198–3207, 07 2017. URL: https://pubs.acs.org/doi/10.1021/acs.jctc.7b00260, arXiv:28590754, doi:10.1021/acs.jctc.7b00260.

14. Georgi L. Stoychev, Alexander A. Auer, and Frank Neese. Automatic generation of auxiliary basis sets. *J. Chem. Theory Comput.*, 13(2):554, 2017. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b01041, doi:https://doi.org/10.1021/acs.jctc.6b01041.

15. Libor Veis, Andrej Antalík, Jiri Brabec, Frank Neese, Ors Legeza, and Jiri Pittner. Coupled cluster method with single and double excitations tailored by matrix product state wave functions. *J. Phys. Chem. Lett.*, 7(20):4072–4078, 2016.

A.1.2 2018

1. G. Bistoni, I. Polyak, M. Sparta, W. Thiel, and F. Neese. Toward accurate QM/MM reaction barriers with large QM regions using domain based pair natural orbital coupled cluster theory. *J. Chem. Theory Comput.*, 14(7):3524–3531, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000438654500015, doi:https://doi.org/10.1021/acs.jctc.8b00348.

2. J. Brabec, J. Lang, M. Saitow, J. Pittner, F. Neese, and O. Demel. Domain-based local pair natural orbital version of mukherjee's state-specific coupled cluster method. *J. Chem. Theory Comput.*, 14(3):1370–1382, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000427661400021, doi:10.1021/acs.jctc.7b01184.

3. Bernardo de Souza, Frank Neese, and Robert Izsak. On the theoretical prediction of fluorescence rates from first principles using the path integral approach. *J. Chem. Phys.*, 148(3):034104, 2018. URL: http://aip.scitation.org/doi/abs/10.1063/1.5010895 (visited on 2018-01-31), doi:10.1063/1.5010895.

4. A. K. Dutta, F. Neese, and R. Izsak. Accelerating the coupled-cluster singles and doubles method using the chain-of-sphere approximation. *Mol. Phys.*, 116(11):1428–1434, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000431017000002, doi:10.1080/00268976.2017.1416201.

5. A. K. Dutta, M. Nooijen, F. Neese, and R. Izsak. Exploring the accuracy of a low scaling similarity transformed equation of motion method for vertical excitation energies. *J. Chem. Theory Comput.*, 14(1):72–91, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000419998300008, doi:10.1021/acs.jctc.7b00802.

6. A. K. Dutta, M. Saitow, C. Riplinger, F. Neese, and R. Izsak. A nearlinear scaling equation of motion coupled cluster method for ionized states. *J. Chem. Phys.*, 148(24):13, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000437190300049, doi:10.1063/1.5029470.

7. Yang Guo, Ute Becker, and Frank Neese. Comparison and combination of "direct" and fragment based local correlation methods: Cluster in molecules and domain based local pair natural orbital perturbation and coupled cluster theories. *J. Chem. Phys.*, 148(12):124117, 2018. doi:10.1063/1.5021898.

8. Yang Guo, Christoph Riplinger, Ute Becker, Dimitrios G. Liakos, Yury Minenkov, Luigi Cavallo, and Frank Neese. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.*, 148(1):011101, 2018. doi:10.1063/1.5011798.

9. D. Maganas, S. DeBeer, and F. Neese. Pair natural orbital restricted open-shell configuration interaction (PNO-ROCIS) approach for calculating x-ray absorption spectra of large chemical systems. *J. Phys. Chem. A*, 122(5):1215–1227, 2018. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jpca.7b10880, doi:https://doi.org/10.1021/acs.jpca.7b10880.

10. M. C. R. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. C. Maia, G. B. Rocha, J. V. Ribeiro, J. E. Stone, F. Neese, K. Schulten, and Z. Luthey-Schulten. NAMD goes quantum: an integrative suite for hybrid simulations. *Nat. Methods*, 15(5):351–+, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000431372700019, doi:10.1038/nmeth.4638.

11. Frank Neese. Software update: the orca program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 8(1):e1327, 2018. doi:http://doi.wiley.com/10.1002/wcms.1327.

12. Peter Pinski and Frank Neese. Communication: Exact analytical derivatives for the domain-based local pair natural orbital MP2 method (DLPNO-MP2). *J. Chem. Phys.*, 148:031101, 2018. doi:10.1063/1.5011204.

13. Masaaki Saitow and Frank Neese. Accurate spin-densities based on the domain-based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.*, 149:034104, 2018. doi:10.1063/1.5027114.

14. Avijit Sen, Bernardo de Souza, Lee M. J. Huntington, Martin Krupička, Frank Neese, and Róbert Izsák. An efficient pair natural orbital based configuration interaction scheme for the calculation of open-shell ionization potentials. *J. Chem. Phys.*, 149(11):114108, 2018.

15. Georgi L Stoychev, Alexander A Auer, Róbert Izsák, and Frank Neese. Self-consistent field calculation of nuclear magnetic resonance chemical shielding constants using gauge-including atomic orbitals and approximate two-electron integrals. *J. Chem. Theory Comput.*, 14(2):619–637, 2018. URL: http://pubs.acs.org/doi/10.1021/acs.jctc.7b01006, doi:10.1021/acs.jctc.7b01006.

16. Georgi L. Stoychev, Alexander A. Auer, and Frank Neese. Efficient and accurate prediction of nuclear magnetic resonance shielding tensors with double-hybrid density functional theory. *J. Chem. Theory Comput.*, 14(9):4756–4771, 09 2018. URL: http://pubs.acs.org/doi/10.1021/acs.jctc.8b00624, doi:10.1021/acs.jctc.8b00624.

A.1.3 2019

1. A. Altun, F. Neese, and G. Bistoni. Open-shell variant of the london dispersion-corrected hartree-fock method hfld for the quantification and analysis of noncovalent interaction energies. *J. Chem. Theory Comput.*, 18(4):2292–2307, 2022. doi:10.1021/acs.jctc.1c01295.

2. Ahmet Altun, Frank Neese, and Giovanni Bistoni. HFLD: A nonempirical london dispersion-corrected Hartree–Fock method for the quantification and analysis of noncovalent interaction energies of large molecular systems. *J. Chem. Theory Comput.*, 15(11):5894–5907, 2019. URL: https://doi.org/10.1021/acs.jctc.9b00425, arXiv:https://doi.org/10.1021/acs.jctc.9b00425, doi:10.1021/acs.jctc.9b00425.

3. Bernardo de Souza, Giliandro Farias, Frank Neese, and Robert Izsak. Efficient simulation of overtones and combination bands in Resonant Raman spectra. *J. Chem. Phys.*, 150(21):accepted – still waiting for publication, 2019. URL: https://aip.scitation.org/doi/10.1063/1.5099247, doi:10.1063/1.5099247.

4. A. K. Dutta, M. Saitow, B. Demoulin, F. Neese, and R. Izsak. A domain-based local pair natural orbital implementation of the equation of motion coupled cluster method for electron attached states. *J. Chem. Phys.*, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000466698700028, doi:10.1063/1.5089637.

5. M. Garcia-Ratés and F. Neese. Efficient implementation of the analytical second derivatives of hartree-fock and hybrid dft energies within the framework of the conductor-like polarizable continuum model. *J. Comput. Chem.*, 40(20):1816–1828, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000470923100002, doi:10.1002/jcc.25833.

6. S. Haldar, C. Riplinger, B. Demoulin, F. Neese, R. Izsak, and A. K. Dutta. Multilayer approach to the ip-eom-dlpno-ccsd method: theory, implementation, and application. *J. Chem. Theory Comput.*, 15(4):2265–2277, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000464475500015, doi:10.1021/acs.jctc.8b01263.

7. Benjamin Helmich-Paris. CASSCF linear response calculations for large open-shell molecules. *J. Chem. Phys.*, 150(17):174121, 2019. URL: https://doi.org/10.1063/1.5092613, doi:10.1063/1.5092613.

8. Christian Kollmar, Kantharuban Sivalingam, Benjamin Helmich-Paris, Celestino Angeli, and Frank Neese. A perturbation-based super-CI approach for the orbital optimization of a CASSCF wave function. *J. Comput. Chem.*, 40:1463–1470, 2019. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.25801, doi:10.1002/jcc.25801.

9. A. Kumar, F. Neese, and E. Valeev. Near-linear scaling explicitly correlated coupled cluster singles and doubles method based on an open-shell domain-based local pair natural orbitals. *Abstr. Pap. Am. Chem. Soc.*, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000525061504143.

10. J. Lang, J. Brabec, M. Saitow, J. Pittner, F. Neese, and O. Demel. Perturbative triples correction to domain-based local pair natural orbital variants of mukherjee's state specific coupled cluster method. *Phys. Chem. Chem. Phys.*, 21(9):5022–5038, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000461722700029, doi:10.1039/c8cp03577f.

11. Lucas Lang and Frank Neese. Spin-dependent properties in the framework of the dynamic correlation dressed complete active space method. *J. Chem. Phys.*, 150:104104, 2019.

12. Dimitrios Maganas, Joanna K. Kowalska, Marcel Nooijen, Serena DeBeer, and Frank Neese. Comparison of multireference ab initio wavefunction methodologies for X-ray absorption edges: A case study on [Fe(II/III)Cl4]2–/1– molecules. *J. Chem. Phys.*, 150(10):104106, 2019. URL: https://pubs.aip.org/aip/jcp/article-abstract/150/10/104106/1058894/, doi:https://doi.org/10.1063/1.5051613.

13. Peter Pinski and Frank Neese. Analytical gradient for the domain-based local pair natural orbital second order Møller-Plesset perturbation theory method (DLPNO-MP2). *J. Chem. Phys.*, 150:164102, 2019.

14. M. Saitow, A. K. Dutta, and F. Neese. Accurate ionization potentials, electron affinities and electronegativities of single-walled carbon nanotubes by state-of-the-art local coupled-cluster theory. *Bull. Chem. Soc. Jpn.*, 92(1):170–174, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000455409900003, doi:10.1246/bcsj.20180254.

A.1.4 2020

1. A. Altun, F. Neese, and G. Bistoni. Extrapolation to the limit of a complete pair natural orbital space in local coupled-cluster calculations. *J. Chem. Theory Comput.*, 16(10):6142–6149, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000580954000015, doi:https://doi.org/10.1021/acs.jctc.0c00344.

2. A. A. Auer, V. A. Tran, B. Sharma, G. L. Stoychev, D. Marx, and F. Neese. A case study of density functional theory and domain-based local pair natural orbital coupled cluster for vibrational effects on epr hyperfine coupling constants: vibrational perturbation theory versus ab-initio molecular dynamics. *Mol. Phys.*, 118:16, 2020. URL: https://www.tandfonline.com/doi/full/10.1080/00268976.2020.1797916, doi:https://doi.org/10.1080/00268976.2020.1797916.

3. D. Datta, M. Saitow, B. Sandhofer, and F. Neese. Fe-57 mossbauer parameters from domain based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.*, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000596027300001, doi:10.1063/5.0022215.

4. A. Dittmer, G. L. Stoychev, D. Maganas, A. A. Auer, and F. Neese. Computation of nmr shielding constants for solids using an embedded cluster approach with dft, double-hybrid dft, and mp2. *J. Chem. Theory Comput.*, 16(11):6950–6967, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000592392800018, doi:10.1021/acs.jctc.0c00067.

5. M. Garcia-Ratés and F. Neese. Effect of the solute cavity on the solvation energy and its derivatives within the framework of the gaussian charge scheme. *J. Comput. Chem.*, 41(9):922–939, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000504862400001, doi:10.1002/jcc.26139.

6. Y. Guo, C. Riplinger, D. G. Liakos, U. Becker, M. Saitow, and F. Neese. Linear scaling perturbative triples correction approximations for open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory dlpno-ccsd(t-0/t). *J. Chem. Phys.*, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000519813300005, doi:10.1063/1.5127550.

7. Christian Kollmar, Kantharuban Sivalingam, and Frank Neese. An alternative choice of the zeroth-order Hamiltonian in CASPT2 theory. *J. Chem. Phys.*, 152(21):214110, 06 2020. doi:10.1063/5.0010019.

8. A. Kumar, F. Neese, and E. F. Valeev. Explicitly correlated coupled cluster method for accurate treatment of open-shell molecules with hundreds of atoms. *J. Chem. Phys.*, 153(9):17, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000570176400001, doi:10.1063/5.0012753.

9. Lucas Lang, Mihail Atanasov, and Frank Neese. Improvement of Ab Initio Ligand Field Theory by Means of Multistate Perturbation Theory. *J. Phys. Chem. A*, 124(5):1025–1037, 02 2020. doi:10.1021/acs.jpca.9b11227.

10. Lucas Lang, Enrico Ravera, Giacomo Parigi, Claudio Luchinat, and Frank Neese. Solution of a puzzle: High-level quantum-chemical treatment of pseudocontact chemical shifts confirms classic semiempirical theory. *J. Phys. Chem. Lett.*, 11(20):8735–8744, 2020.

11. Lucas Lang, Kantharuban Sivalingam, and Frank Neese. The combination of multipartitioning of the Hamiltonian with canonical Van Vleck perturbation theory leads to a Hermitian variant of quasidegenerate N-electron valence perturbation theory. *J. Chem. Phys.*, 152(1):014109, 01 2020. doi:10.1063/1.5133746.

12. D. G. Liakos, Y. Guo, and F. Neese. Comprehensive benchmark results for the domain based local pair natural orbital coupled cluster method (dlpno-ccsd(t)) for closed- and open-shell systems. *J. Phys. Chem. A*, 124(1):90–100, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000507151000012, doi:10.1021/acs.jpca.9b05734.

13. Frank Neese, Frank Wennmohs, Ute Becker, and Christoph Riplinger. The orca quantum chemistry program package. *J. Chem. Phys.*, 152(22):224108, 2020. doi:https://aip.scitation.org/doi/10.1063/5.0004608.

14. Julian D. Rolfes, Frank Neese, and Dimitrios A. Pantazis. All-electron scalar relativistic basis sets for the elements Rb–Xe. *J. Comput. Chem.*, 41:1842–1849, 2020. doi:10.1002/jcc.26355.

15. Van Anh Tran and Frank Neese. Double-hybrid density functional theory for g-tensor calculations using gauge including atomic orbitals. *J. Chem. Phys.*, 153(5):054105, 08 2020. URL: https://doi.org/10.1063/5.0013799, doi:10.1063/5.0013799.

A.1.5 2021

1. Vijay Gopal Chilkuri and Frank Neese. Comparison of many-particle representations for selected-CI I: A tree based approach. *J. Comput. Chem.*, 42(14):982–1005, 2021. doi:10.1002/jcc.26518.

2. Vijay Gopal Chilkuri and Frank Neese. Comparison of Many-Particle Representations for Selected Configuration Interaction: II. Numerical Benchmark Calculations. *J. Chem. Theory Comput.*, 17(5):2868–2885, 05 2021. doi:10.1021/acs.jctc.1c00081.

3. M. Garcia-Ratés, U. Becker, and F. Neese. Implicit solvation in domain based pair natural orbital coupled cluster (dlpno-ccsd) theory. *J. Comput. Chem.*, 42(27):1959–1973, 2021. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.26726, doi:10.1002/jcc.26726.

4. Y. Guo, W. Li, and S. Li. Improved cluster-in-molecule local correlation approach for electron correlation calculation of large systems. *J. Phys. Chem. A*, 118(39):8996–9004, 2014. doi:https://doi.org/10.1021/jp501976x.

5. Benjamin Helmich-Paris. A trust-region augmented Hessian implementation for restricted and unrestricted Hartree–Fock and Kohn–Sham methods. *J. Chem. Phys.*, 154(16):164104, 2021. URL: https://doi.org/10.1063/5.0040798, doi:10.1063/5.0040798.

6. Benjamin Helmich-Paris, Bernardo de Souza, Frank Neese, and Róbert Izsák. An improved chain of spheres for exchange algorithm. *J. Chem. Phys.*, 155(10):104109, 2021. doi:10.1063/5.0058766.

7. Georgi L. Stoychev, Alexander A. Auer, Jürgen Gauss, and Frank Neese. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *J. Chem. Phys.*, 154(16):164110, 2021. URL: https://aip.scitation.org/doi/10.1063/5.0047125, doi:10.1063/5.0047125.

A.2 Relevant applications, benchmarks and reviews

A.2.1 2017

1. Giovanni Bistoni, Alexander A. Auer, and Frank Neese. Understanding the role of dispersion in frustrated lewis pairs and classical lewis adducts: A domain-based local pair natural orbital coupled cluster study. *Chem. Eur. J.*, 23(4):865–873, 2017.

2. Giovanni Bistoni, Christoph Riplinger, Yury Minenkov, Luigi Cavallo, Alexander A Auer, and Frank Neese. Treating subvalence correlation effects in domain based pair natural orbital coupled cluster calculations: an out-of-the-box approach. *J. Chem. Theory Comput.*, 2017. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.7b00352, doi:https://doi.org/10.1021/acs.jctc.7b00352.

3. Octav Caldararu, Martin A Olsson, Christoph Riplinger, Frank Neese, and Ulf Ryde. Binding free energies in the sampl5 octa-acid host–guest challenge calculated with dft-d3 and ccsd (t). *J. Comput.-Aided Mol. Des.*, 31(1):87–106, 2017. URL: https://link.springer.com/article/10.1007/s10822-016-9957-5, doi:https://doi.org/10.1007/s10822-016-9957-5.

4. Uttam Chakraborty, Serhiy Demeshko, Franc Meyer, Christophe Rebreyend, Bas de Bruin, Mihail Atanasov, Frank Neese, Bernd Mühldorf, and Robert Wolf. Electronic Structure and Magnetic Anisotropy of an Unsaturated Cyclopentadienyl Iron(I) Complex with 15 Valence Electrons. *Angew. Chem. Int. Ed.*, 56(27):7995–7999, 06 2017. doi:10.1002/anie.201702454.

5. Vijay Gopal Chilkuri, Serena DeBeer, and Frank Neese. Revisiting the Electronic Structure of FeS Monomers Using ab Initio Ligand Field Theory and the Angular Overlap Model. *Inorg. Chem.*, 56(17):10418–10436, 09 2017. doi:10.1021/acs.inorgchem.7b01371.

6. Julie Jung, Mihail Atanasov, and Frank Neese. Ab Initio Ligand-Field Theory Analysis and Covalency Trends in Actinide and Lanthanide Free Ions and Octahedral Complexes. *Inorg. Chem.*, 56(15):8802–8816, 08 2017. doi:10.1021/acs.inorgchem.7b00642.

7. Adam Kubas, Johannes Noak, Annette Trunschke, Robert Schlögl, Frank Neese, and Dimitrios Maganas. A combined experimental and theoretical spectroscopic protocol for determination of the structure of heterogeneous catalysts: developing the information content of the resonance raman spectra of m1 movo x. *Chem. Sci.*, 8(9):6338–6353, 2017.

8. Yury Minenkov, Giovanni Bistoni, Christoph Riplinger, Alexander A Auer, Frank Neese, and Luigi Cavallo. Pair natural orbital and canonical coupled cluster reaction enthalpies involving light to heavy alkali and alkaline earth metals: the importance of sub-valence correlation. *Phys. Chem. Chem. Phys.*, 19(14):9374–9391, 2017.

9. Frank Neese. High-level spectroscopy, quantum chemistry, and catalysis: not just a passing fad. *Angew. Chem. Int. Ed.*, 56(37):11003–11010, 2017.

10. Frank Neese. *Quantum Chemistry and EPR Parameters*, pages 1–22. American Cancer Society, 2017. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470034590.emrstm1505, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470034590.emrstm1505, doi:https://doi.org/10.1002/9780470034590.emrstm1505.

11. Kasper S Pedersen, Daniel N Woodruff, Saurabh Kumar Singh, Alain Tressaud, Etienne Durand, Mihail Atanasov, Panagiota Perlepe, Katharina Ollefs, Fabrice Wilhelm, Corine Mathonière, and others. [osf6] x-: molecular models for spin-orbit entangled phenomena. *Chem. Eur. J.*, 23(47):11244–11248, 2017.

12. Saurabh Kumar Singh, Julien Eng, Mihail Atanasov, and Frank Neese. Covalency and chemical bonding in transition metal complexes: An ab initio based ligand field perspective. *Coordin. Chem. Rev.*, 344:2–25, 08 2017. doi:10.1016/j.ccr.2017.03.018.

13. Elizaveta A Suturina, Joscha Nehrkorn, Joseph M Zadrozny, Junjie Liu, Mihail Atanasov, Thomas Weyhermüller, Dimitrios Maganas, Stephen Hill, Alexander Schnegg, Eckhard Bill, and others. Magneto-structural correlations in pseudotetrahedral forms of the [co (sph) 4] 2–complex probed by magnetometry, mcd spectroscopy, advanced epr techniques, and ab initio electronic structure calculations. *Inorg. Chem.*, 56(5):3102–3118, 2017.

A.2.2 2018

1. A. Altun, F. Neese, and G. Bistoni. Local energy decomposition analysis of hydrogen-bonded dimers within a domain-based pair natural orbital coupled cluster study. *Beilstein J. Org. Chem.*, 14:919–929, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000430878300002, doi:10.3762/bjoc.14.79.

2. G. Bistoni, I. Polyak, M. Sparta, W. Thiel, and F. Neese. Toward accurate QM/MM reaction barriers with large QM regions using domain based pair natural orbital coupled cluster theory. *J. Chem. Theory Comput.*, 14(7):3524–3531, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000438654500015, doi:https://doi.org/10.1021/acs.jctc.8b00348.

3. P. C. Bunting, M. Atanasov, E. Damgaard-Moller, M. Perfetti, I. Crassee, M. Orlita, J. Overgaard, J. van Slageren, F. Neese, and J. R. Long. A linear cobalt(II) complex with maximal orbital angular momentum from a non-Aufbau ground state. *Science*, 362(6421):1378–+, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000453845000052, doi:10.1126/science.aat7319.

4. A. Chantzis, J. K. Kowalska, D. Maganas, S. DeBeer, and F. Neese. Ab initio wave function-based determination of element specific shifts for the efficient calculation of x-ray absorption spectra of main group elements and first row transition metals. *J. Chem. Theory Comput.*, 14(7):3686–3702, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000438654500028, doi:10.1021/acs.jctc.8b00249.

5. L. R. Collins, M. van Gastel, F. Neese, and A. Furstner. Enhanced electrophilicity of heterobimetallic bi-rh paddlewheel carbene complexes: A combined experimental, spectroscopic, and computational study. *J. Am. Chem. Soc.*, 140(40):13042–13055, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000447354800056, doi:10.1021/jacs.8b08384.

6. G. David, F. Wennmohs, F. Neese, and N. Ferre. Chemical tuning of magnetic exchange couplings using broken-symmetry density functional theory. *Inorg. Chem.*, 57(20):12769–12776, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000447680400039, doi:10.1021/acs.inorgchem.8b01970.

7. T. Gatzenmeier, M. Turberg, D. Yepes, Y. W. Xie, F. Neese, G. Bistoni, and B. List. Scalable and highly diastereo- and enantioselective catalytic diels-alder reaction of alpha,beta-Unsaturated methyl esters. *J. Am. Chem. Soc.*, 140(40):12671–12676, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000447354800004, doi:10.1021/jacs.8b07092.

8. H. C. Gottschalk, A. Poblotzki, M. A. Suhm, M. M. Al-Mogren, J. Antony, A. A. Auer, L. Baptista, D. M. Benoit, G. Bistoni, F. Bohle, R. Dahmani, D. Firaha, S. Grimme, A. Hansen, M. E. Harding, M. Hochlaf, C. Holzer, G. Jansen, W. Klopper, W. A. Kopp, L. C. Kroger, K. Leonhard, H. Mouhib, F. Neese, M. N. Pereira, I. S. Ulusoy, A. Wuttke, and R. A. Mata. The furan microsolvation blind challenge for quantum chemical methods: First steps. *J. Chem. Phys.*, 148(1):13, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000419394500013, doi:10.1063/1.5009011.

9. Adam Kubas, Max Verkamp, Josh Vura-Weis, Frank Neese, and Dimitrios Maganas. Restricted open-shell configuration interaction singles study on m-and l-edge x-ray absorption spectroscopy of solid chemical systems. *J. Chem. Theory Comput.*, 14(8):4320–4334, 2018.

10. Q. Lu, F. Neese, and G. Bistoni. Formation of agostic structures driven by london dispersion. *Angew. Chem. Int. Ed.*, 57(17):4760–4764, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000430165700058, doi:10.1002/anie.201801531.

11. B. Mondal, F. Neese, E. Bill, and S. F. Ye. Electronic structure contributions of non-herne oxo-iron(v) complexes to the reactivity. *J. Am. Chem. Soc.*, 140(30):9531–9544, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000440877000033, doi:10.1021/jacs.8b04275.

12. D. H. Moseley, S. E. Stavretis, K. Thirunavukkuarasu, M. Ozerov, Y. Q. Cheng, L. L. Daemen, J. Ludwig, Z. G. Lu, D. Smirnov, C. M. Brown, A. Pandey, A. J. Ramirez-Cuesta, A. C. Lamb, M. Atanasov, E. Bill, F. Neese, and Z. L. Xue. Spin-phonon couplings in transition metal complexes with slow

magnetic relaxation. *Nature Comm.*, 9:11, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000437101700002, doi:10.1038/s41467-018-04896-0.

13. C. Romelt, S. F. Ye, E. Bill, T. Weyhermuller, M. van Gastel, and F. Neese. Electronic structure and spin multiplicity of iron tetraphenylporphyrins in their reduced states as determined by a combination of resonance raman spectroscopy and quantum chemistry. *Inorg. Chem.*, 57(4):2141–2148, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000426014800048, doi:10.1021/acs.inorgchem.7b03018.

14. B. Scheibe, C. Pietzonka, O. Mustonen, M. Karppinen, A. J. Karttunen, M. Atanasov, F. Neese, M. Conrad, and F. Kraus. The U2F12 (2-) anion of sr U2F12. *Angew. Chem. Int. Ed.*, 57(11):2914–2918, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000426490700027, doi:10.1002/anie.201800743.

15. R. G. Shirazi, F. Neese, and D. A. Pantazis. Accurate spin-state energetics for aryl carbenes. *J. Chem. Theory Comput.*, 14(9):4733–4746, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000444792700020, doi:10.1021/acs.jctc.8b00587.

16. Saurabh Kumar Singh, Mihail Atanasov, and Frank Neese. Challenges in multireference perturbation theory for the calculations of the g-tensor of first-row transition-metal complexes. *J. Chem. Theory Comput.*, 14(9):4662–4677, 2018.

17. C. Van Stappen, D. Maganas, S. DeBeer, E. Bill, and F. Neese. Investigations of the magnetic and spectroscopic properties of V(III) and V(IV) complexes. *Inorg. Chem.*, 57(11):6421–6438, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000434491700027, doi:10.1021/acs.inorgchem.8b00486.

18. K. Yamamoto, J. K. Li, J. A. O. Garber, J. D. Rolfes, G. B. Boursalian, J. C. Borghs, C. Genicot, J. Jacq, M. van Gastel, F. Neese, and T. Ritter. Palladium-catalysed electrophilic aromatic C-H fluorination. *Nature*, 554(7693):511–514, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000425597400043, doi:10.1038/nature25749.

A.2.3 2019

1. Ahmet Altun, Frank Neese, and Giovanni Bistoni. Effect of electron correlation on intermolecular interactions: A pair natural orbitals coupled cluster based local energy decomposition study. *J. Chem. Theory Comput.*, 15(1):215–228, 2019. URL: https://doi.org/10.1021/acs.jctc.8b00915, doi:https://doi.org/10.1021/acs.jctc.8b00915.

2. R. Berraud-Pache, F. Neese, G. Bistoni, and R. Izsak. Computational design of near-infrared fluorescent organic dyes using an accurate new wave function approach. *J. Phys. Chem. Lett.*, 10(17):4822–4828, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000484884300010, doi:https://doi.org/10.1021/acs.jpclett.9b02240.

3. H. C. Chang, Y. H. Lin, C. Werle, F. Neese, W. Z. Lee, E. Bill, and S. F. Ye. Conversion of a fleeting open-shell iron nitride into an iron nitrosyl. *Angew. Chem. Int. Ed.*, 58(49):17589–17593, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000491791300001, doi:10.1002/anie.201908689.

4. H. C. Chang, B. Mondal, H. Y. Fang, F. Neese, E. Bill, and S. F. Ye. Electron paramagnetic resonance signature of tetragonal low spin Iron(V)-Nitrido and -Oxo complexes derived from the electronic structure analysis of heme and non-heme archetypes. *J. Am. Chem. Soc.*, 141(6):2421–2434, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459222100035, doi:10.1021/jacs.8b11429.

5. Anneke Dittmer, Róbert Izsák, Frank Neese, and Dimitrios Maganas. Accurate band gap predictions of semiconductors in the framework of the similarity transformed equation of motion coupled cluster theory. *Inorg. Chem.*, 58(14):9303–9315, 2019. URL: https://doi.org/10.1021/acs.inorgchem.9b00994, doi:10.1021/acs.inorgchem.9b00994.

6. Benjamin Helmich-Paris. Benchmarks for electronically excited states with CASSCF methods. *J. Chem. Theory Comput.*, 15(7):4170–4179, 2019. URL: https://doi.org/10.1021/acs.jctc.9b00325, doi:10.1021/acs.jctc.9b00325.

7. M. Keilwerth, J. Hohenberger, F. W. Heinemann, J. Sutter, A. Scheurer, H. Y. Fang, E. Bill, F. Neese, S. F. Ye, and K. Meyer. A series of iron nitrosyl complexes Fe-NO(6-9) and a fleeting Fe-NO(10) intermediate en route to a metalacyclic iron nitrosoalkane. *J. Am. Chem. Soc.*, 141(43):17217–17235, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000493866300028, doi:10.1021/jacs.9b08053.

8. V. Krewald, F. Neese, and D. A. Pantazis. Implications of structural heterogeneity for the electronic structure of the final oxygen-evolving intermediate in photosystem II. *J. Inorg. Biochem.*, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000488146900034, doi:10.1016/j.jinorgbio.2019.110797.

9. Q. Lu, F. Neese, and G. Bistoni. London dispersion effects in the coordination and activation of alkanes in sigma-complexes: a local energy decomposition study. *Phys. Chem. Chem. Phys.*, 21(22):11569–11577, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000472218500051, doi:10.1039/c9cp01309a.

10. Dimitrios Maganas, Joanna K. Kowalska, Marcel Nooijen, Serena DeBeer, and Frank Neese. Comparison of multireference ab initio wavefunction methodologies for X-ray absorption edges: A case study on [Fe(II/III)Cl4]2–/1– molecules. *J. Chem. Phys.*, 150(10):104106, 2019. URL: https://pubs.aip.org/aip/jcp/article-abstract/150/10/104106/1058894/, doi:https://doi.org/10.1063/1.5051613.

11. F. Neese, M. Atanasov, G. Bistoni, D. Maganas, and S. F. Ye. Chemistry and quantum mechanics in 2019: Give us insight and numbers. *J. Am. Chem. Soc.*, 141(7):2814–2824, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459642000006, doi:10.1021/jacs.8b13313.

12. M. Saitow, A. K. Dutta, and F. Neese. Accurate ionization potentials, electron affinities and electronegativities of single-walled carbon nanotubes by state-of-the-art local coupled-cluster theory. *Bull. Chem. Soc. Jpn.*, 92(1):170–174, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000455409900003, doi:10.1246/bcsj.20180254.

13. C. A. M. Salla, J. T. dos Santos, G. Farias, A. J. Bortoluzi, S. F. Curcio, T. Cazati, R. Izsak, F. Neese, B. de Souza, and I. H. Bechtold. New Boron(III) blue emitters for all-solution processed OLEDs: Molecular design assisted by theoretical modeling. *Eur. J. Inorg. Chem.*, pages 2247–2257, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000471302400001, doi:10.1002/ejic.201900265.

14. R. G. Shirazi, F. Neese, D. A. Pantazis, and G. Bistoni. Physical nature of differential spin-state stabilization of carbenes by hydrogen and halogen bonding: A domain-based pair natural orbital coupled cluster study. *J. Phys. Chem. A*, 123(24):5081–5090, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000472800600009, doi:10.1021/acs.jpca.9b01051.

15. A. Sirohiwal, F. Neese, and D. A. Pantazis. Microsolvation of the redox-active tyrosine-d in photosystem II: Correlation of energetics with EPR spectroscopy and oxidation-induced proton transfer. *J. Am. Chem. Soc.*, 141(7):3217–3231, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459642000056, doi:10.1021/jacs.8b13123.

16. S. E. Stavretis, Y. Q. Cheng, L. L. Daemen, C. M. Brown, D. H. Moseley, E. Bill, M. Atanasov, A. J. Ramirez-Cuesta, F. Neese, and Z. L. Xue. Probing magnetic excitations in co-ii single-molecule magnets by inelastic neutron scattering. *Eur. J. Inorg. Chem.*, pages 1119–1127, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459919200008, doi:10.1002/ejic.201801088.

17. M. K. Thomsen, A. Nyvang, J. P. S. Walsh, P. C. Bunting, J. R. Long, F. Neese, M. Atanasov, A. Genoni, and J. Oyergaard. Insights into single-molecule-magnet behavior from the experimental electron density of linear two-coordinate iron complexes. *Inorg. Chem.*, 58(5):3211–3218, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000460600300035, doi:10.1021/acs.inorgchem.8b03301.

A.2.4 2020

1. K. Chakarawet, M. Atanasov, J. Marbey, P. C. Bunting, F. Neese, S. Hill, and J. R. Long. Strong electronic and magnetic coupling in m-4 (m = ni, cu) clusters via direct orbital interactions between low-coordinate metal centers. *J. Am. Chem. Soc.*, 142(45):19161–19169, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000588273900023, doi:10.1021/jacs.0c08460.

2. Vijay Gopal Chilkuri, Serena DeBeer, and Frank Neese. Ligand Field Theory and Angular Overlap Model Based Analysis of the Electronic Structure of Homovalent Iron–Sulfur Dimers. *Inorg. Chem.*, 59(2):984–995, 01 2020. doi:10.1021/acs.inorgchem.9b00974.

3. A. Dittmer, G. L. Stoychev, D. Maganas, A. A. Auer, and F. Neese. Computation of nmr shielding constants for solids using an embedded cluster approach with dft, double-hybrid dft, and mp2. *J. Chem. Theory Comput.*, 16(11):6950–6967, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000592392800018, doi:10.1021/acs.jctc.0c00067.

4. B. M. Floser, Y. Guo, C. Riplinger, F. Tuczek, and F. Neese. Detailed pair natural orbital-based coupled cluster studies of spin crossover energetics. *J. Chem. Theory Comput.*, 16(4):2224–2235, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000526313000020, doi:10.1021/acs.jctc.9b01109.

5. H. C. Gottschalk, A. Poblotzki, M. Fatima, D. A. Obenchain, C. Perez, J. Antony, A. A. Auer, L. Baptista, D. M. Benoit, G. Bistoni, F. Bohle, R. Dahmani, D. Firaha, S. Grimme, A. Hansen, M. E. Harding, M. Hochlaf, C. Holzer, G. Jansen, W. Klopper, W. A. Kopp, L. C. Kroger, K. Leonhard, M. M. Al-Mogren, H. Mouhib, F. Neese, X. N. Pereira, M. Prakash, I. S. Ulusoy, R. A. Mata, M. A. Suhm, and M. Schnell. The first microsolvation step for furans: New experiments and benchmarking strategies. *J. Chem. Phys.*, 152(16):17, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000531832400001, doi:10.1063/5.0004465.

6. J. Hillenbrand, M. van Gastel, E. Bill, F. Neese, and A. Furstner. Isolation of a homoleptic non-oxo Mo(V) alkoxide complex: Synthesis, structure, and electronic properties of penta-tert-butoxymolybdenum. *J. Am. Chem. Soc.*, 142(38):16392–16402, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000575684100032, doi:10.1021/jacs.0c07073.

7. J. Jung, S. T. Loffler, J. Langmann, F. W. Heinemann, E. Bill, G. Bistoni, W. Scherer, M. Atanasov, K. Meyer, and F. Neese. Dispersion forces drive the formation of uranium-alkane adducts. *J. Am. Chem. Soc.*, 142(4):1864–1870, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000510531900033, doi:10.1021/jacs.9b10620.

8. D. Maganas, J. K. Kowalska, C. Van Stappen, S. DeBeer, and F. Neese. Mechanism of L-2,L-3-edge x-ray magnetic circular dichroism intensity from quantum chemical calculations and experiment-A case study on V-(IV)/V-(III) complexes. *J. Chem. Phys.*, 152(11):15, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000521227700001, doi:10.1063/1.5129029.

9. F. Meyer and F. Neese. Impact of modern spectroscopy in inorganic chemistry. *Inorg. Chem.*, 59(19):13805–13806, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000580381700001, doi:10.1021/acs.inorgchem.0c02755.

10. Julian D. Rolfes, Frank Neese, and Dimitrios A. Pantazis. All-electron scalar relativistic basis sets for the elements Rb–Xe. *J. Comput. Chem.*, 41:1842–1849, 2020. doi:10.1002/jcc.26355.

11. R. G. Shirazi, D. A. Pantazis, and F. Neese. Performance of density functional theory and orbital-optimised second-order perturbation theory methods for geometries and singlet-triplet state splittings of aryl-carbenes. *Mol. Phys.*, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000535113400001, doi:10.1080/00268976.2020.1764644.

12. A. Sirohiwal, F. Neese, and D. A. Pantazis. Protein matrix control of reaction center excitation in photosystem II. *J. Am. Chem. Soc.*, 142(42):18174–18190, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000580559000041, doi:10.1021/jacs.0c08526.

13. Abhishek Sirohiwal, Romain Berraud-Pache, Frank Neese, Róbert Izsák, and Dimitrios A. Pantazis. Accurate computation of the absorption spectrum of chlorophyll a with pair natural orbital coupled cluster methods. *J. Phys. Chem. B*, 124(40):8761–8771, 2020. URL: https://doi.org/10.1021/acs.jpcb.0c05761, doi:10.1021/acs.jpcb.0c05761.

14. N. Spiller, V. G. Chilkuri, S. DeBeer, and F. Neese. Sulfur vs. Selenium as bridging ligand in di-iron complexes: A theoretical analysis. *Eur. J. Inorg. Chem.*, 2020(15-16):1525–1538, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000520715700001, doi:10.1002/ejic.202000033.

15. D. Yepes, F. Neese, B. List, and G. Bistoni. Unveiling the delicate balance of steric and dispersion interactions in organocatalysis using high-level computational methods. *J. Am. Chem. Soc.*, 142(7):3613–3625, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000515214000044, doi:10.1021/jacs.9b13725.

A.2.5 2021

1. M. E. Beck, C. Riplinger, F. Neese, and G. Bistoni. Unraveling individual host-guest interactions in molecular recognition from first principles quantum mechanics: Insights into the nature of nicotinic acetylcholine receptor agonist binding. *J. Comput. Chem.*, 42(5):293–302, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000591776900001, doi:https://doi.org/10.1002/jcc.26454.

2. R. Berraud-Pache, E. Santamaria-Aranda, B. de Souza, G. Bistoni, F. Neese, D. Sampedro, and R. Izsak. Redesigning donor-acceptor Stenhouse adduct photoswitches through a joint experimental and computational study. *Chem. Sci.*, 12(8):2916–2924, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000625958900018, doi:https://doi.org/10.1039/d0sc06575g.

3. C. Daniel, L. Gonzalez, and F. Neese. Quantum theory: The challenge of transition metal complexes. *Phys. Chem. Chem. Phys.*, 23(4):2533–2534, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000614634000001, doi:10.1039/d0cp90278k.

4. C. E. Schulz, R. G. Castillo, D. A. Pantazis, S. DeBeer, and F. Neese. Structure-spectroscopy correlations for intermediate q of soluble methane monooxygenase: Insights from QM/MM calculations. *J. Am. Chem. Soc.*, 143(17):6560–6577, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000648704100032, doi:10.1021/jacs.1c01180.

5. C. E. Schulz, M. van Gastel, D. A. Pantazis, and F. Neese. Converged structural and spectroscopic properties for refined qm/mm models of azurin. *Inorg. Chem.*, 60(10):7399–7412, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000653539100063, doi:10.1021/acs.inorgchem.1c00640.

6. A. Sirohiwal, F. Neese, and D. A. Pantazis. Chlorophyll excitation energies and structural stability of the cp47 antenna of photosystem ii: a case study in the first-principles simulation of light-harvesting complexes. *Chem. Sci.*, 12(12):4463–4476, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000635768300025, doi:10.1039/d0sc06616h.

7. A. Sirohiwal, F. Neese, and D. A. Pantazis. How can we predict accurate electrochromic shifts for biochromophores? a case study on the photosynthetic reaction center. *J. Chem. Theory Comput.*, 17(3):1858–1873, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000629135700044, doi:10.1021/acs.jctc.0c01152.

8. M. Tarrago, C. Romelt, J. Nehrkorn, A. Schnegg, F. Neese, E. Bill, and S. F. Ye. Experimental and theoretical evidence for an unusual almost triply degenerate electronic ground state of ferrous tetraphenylporphyrin. *Inorg. Chem.*, 60(7):4966–4985, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000637850300081, doi:10.1021/acs.inorgchem.1c00031.

A.3 Classification by topic

A.3.1 *Ab initio* Ligand Field Analyis

1. Atanasov, M.; Ganyushin, D.; Sivalingam, K.; Neese, F.)in Molecular Electronic Structures of Transition Metal Complexes II (eds. Mingos, D. M. P., Day, P. Dahl, J. P.) 149–220 (Springer Berlin Heidelberg, **2011**).

A.3.2 Absorption, Resonance Raman and Fluorescence Spectra

1. Sirohiwal, A.; Berraud-Pache, R.; Neese, F.; Izsak, R.; Pantazis, D. A. (**2020**) Accurate Computation of the Absorption Spectrum of Chlorophyll alpha with Pair Natural Orbital Coupled Cluster Methods, *J. Phys. Chem. B*, 124, 8761-8771.

2. Petrenko, T.; Neese F. (**2012**) Efficient and automatic calculation of optical band shapes and resonance Raman spectra for larger molecules within the independent mode displaced harmonic oscillator model, *J. Chem. Phys.*, 137, 234107.

3. Petrenko, T.; Neese, F. (**2007**) A general efficient quantum chemical method for predicting absorption bandshapes, resonance Raman spectra and excitation profiles for larger molecules. *J. Chem. Phys.*, 127, 164319.

4. Petrenko, T.; Krylova, O.; Neese, F. Sokolowski, M. (**2009**) Optical Absorption and Emission Properties of Rubrene: Insight by a Combined Experimental and Theoretical Study. *New J. Phys.*, 11, 015001.

A.3.3 Analytic Hessian Implementation

1. Bykov, D.; Petrenko, T.; Izsák, R.; Kossmann, S.; Becker, U.; Valeev, E.; Neese, F. (**2015**) Efficient implementation of the analytic second derivatives of Hartree-Fock and hybrid DFT energies: a detailed analysis of different approximations, *Mol. Phys.*, 113, 1961.
2. Garcia-Ratés, M.; Neese, F. (**2019**) Efficient implementation of the analytical second derivatives of hartree-fock and hybrid DFT energies within the framework of the conductor-like polarizable continuum model, *J. Comp. Chem.*, 40, 1816-1828.

A.3.4 ANO basis sets

1. Neese, F.; Valeev, E. F. (**2011**) Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated ab initio Methods?, *J. Chem. Theory Comput.*, 7, 33–43.

A.3.5 Applications

1. Mondal, B.; Neese, F.; Ye, S. F. (**2016**) Toward Rational Design of 3d Transition Metal Catalysts for CO2 Hydrogenation Based on Insights into Hydricity-Controlled Rate-Determining Steps, *Inorg. Chem.*, 55, 5438-5444.
2. Mondal, B.; Neese, F.; Ye, S. F. (**2015**) Control in the Rate-Determining Step Provides a Promising Strategy To Develop New Catalysts for CO2 Hydrogenation: A Local Pair Natural Orbital Coupled Cluster Theory Study, *Inorg. Chem.*, 54, 7192-7198.
3. Krewald, V.; Retegan, M.; Cox, N.; Messinger, J.; Lubitz, W.; DeBeer, S.; Neese, F.; Pantazis, D. A. (**2015**) Metal oxidation states in biological water splitting, *Chem. Sci.*, 6, 1676-1695.
4. Kochem, A.; Weyhermuller, T.; Neese, F.; van Gastel, M. (**2015**) EPR and Quantum Chemical Investigation of a Bioinspired Hydrogenase Model with a Redox-Active Ligand in the First Coordination Sphere, *Organometallics*, 34, 995-1000.
5. Kochem, A.; Bill, E.; Neese, F.; van Gastel, M. (**2015**) Mossbauer and computational investigation of a functional NiFe hydrogenase model complex, *Chem. Comm.*, 51, 2099-2102.

A.3.6 Approximate FCI

1. Chilkuri, V. G.; Neese, F. (**2021**) Comparison of many-particle representations for selected-CI I: A tree based approach, *J.Comp. Chem.*, **2021**, 42, 982-1005.
2. Chilkuri, V. G.; Neese, F. (**2021**) Comparison of Many-Particle Representations for Selected Configuration Interaction: II. Numerical Benchmark Calculations, *J. Chem. Theo. Comp.*, 17, 2868-2885.

A.3.7 CASSCF/NEVPT2/MRCI & Magnetism

1. Chilkuri, V.G., DeBeer, S., and Neese, F. (**2017**). Revisiting the Electronic Structure of FeS Monomers Using ab Initio Ligand Field Theory and the Angular Overlap Model., *Inorg. Chem.*, 56, 10418.
2. Singh, S.K., Eng, J., Atanasov, M., and Neese, F. (**2017**). Covalency and chemical bonding in transition metal complexes: An ab initio based ligand field perspective., *Coordination Chemistry Reviews*, 344, 225.
3. Jung, J., Atanasov, M., and Neese, F. (**2017**). Ab Initio Ligand-Field Theory Analysis and Covalency Trends in Actinide and Lanthanide Free Ions and Octahedral Complexes., *Inorg. Chem.*, 56, 8802.
4. Werncke, C. G.; Suturina, E.; Bunting, P. C.; Vendier, L.; Long, J. R.; Atanasov, M.; Neese, F.; Sabo-Etienne, S.; Bontemps, S. (**2016**) Homoleptic Two-Coordinate Silylamido Complexes of Chromium(I), Manganese(I), and Cobalt(I), *Chem. Eur. J.*, 22, 1668-1674.
5. Rechkemmer, Y.; Breitgoff, F. D.; van der Meer, M.; Atanasov, M.; Hakl, M.; Orlita, M.; Neugebauer, P.; Neese, F.; Sarkar, B.; van Slageren, J. (**2016**) A four-coordinate cobalt(II) single-ion magnet with coercivity and a very high energy barrier, *Nat. Commun.*, 7, 10467.
6. Aravena, D.; Atanasov, M.; Neese, F. (**2016**) Periodic Trends in Lanthanide Compounds through the Eyes of Multireference ab Initio Theory, *Inorg. Chem.*, 55, 4457-4469.
7. Suturina, E. A.; Maganas, D.; Bill, E.; Atanasov, M.; Neese, F. (**2015**) Magneto-Structural Correlations in a Series of Pseudotetrahedral Co-II(XR)(4) (2-) Single Molecule Magnets: An ab Initio Ligand Field Study, *Inorg. Chem.*, 54, 9948-9961.
8. Schweinfurth, D.; Sommer, M. G.; Atanasov, M.; Demeshko, S.; Hohloch, S.; Meyer, F.; Neese, F.; Sarkar, B. (**2015**) The Ligand Field of the Azido Ligand: Insights into Bonding Parameters and Magnetic Anisotropy in a Co(II)-Azido Complex, *J. Am. Chem. Soc.*, 137, 1993-2005.

A.3.8 Corresponding Orbital Transformation

1. Neese, F. (**2004**) Definition of Corresponding Orbitals and the Diradical Character in Broken Symmetry DFT Calculations on Spin Coupled Systems. *J. Phys. Chem. Solids*, 65, 781–785.

A.3.9 COSMO Implementation

1. Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F. (**2006**) Calculation of Solvent Shifts on Electronic G-Tensors with the Conductor-Like Screening Model (COSMO) and its Self-Consistent Generalization to Real Solvents (COSMO-RS), *J. Phys. Chem. A*, 110, 2235–2245.

A.3.10 COSX

1. Dutta, A. K.; Neese, F.; Izsak, R. (**2016**) Speeding up equation of motion coupled cluster theory with the chain of spheres approximation, *J. Chem. Phys.*, 144, 034102.
2. Christian, G. J.; Neese, F.; Ye, S. F. (**2016**) Unravelling the Molecular Origin of the Regiospecificity in Extradiol Catechol Dioxygenases, *Inorg. Chem.*, 55, 3853-3864.

A.3.11 Coupled-Cluster and Coupled Pair Implementation (MDCI module)

1. Pavošević, F.; Neese, F.; Valeev, E. F. (**2014**) Geminal-spanning orbitals make explicitly correlated reduced-scaling coupled-cluster methods robust, yet simple, *J. Chem. Phys.*, 141, 054106
2. Kollmar, C.; Neese, F. (**2010**) The coupled electron pair approximation: variational formulation and spin adaptation, *Mol. Phys.*, 108, 2449–2458.
3. Neese, F.; Wennmohs, F.; Hansen, A.; Grimme, S. (**2009**) Accurate Theoretical Chemistry with Coupled Electron Pair Models *Acc. Chem. Res.*, 42(5), 641–648.
4. Wennmohs, F.; Neese, F. (**2008**) A Comparative Study of Single Reference Correlation Methods of the Coupled-Pair Type, *Chem. Phys.* (70^{th} birthday issue for Prof. Peyerimhoff), 343, 217–230.

A.3.12 EPR/NMR

1. Schulz, C. E.; van Gastel, M.; Pantazis, D. A.; Neese, F. (**2021**) Converged Structural and Spectroscopic Properties for Refined QM/MM Models of Azurin, *Inorg. Chem.*, 60, 7399-7412.
2.
3. Stoychev, G. L., Auer, A. A., Gauss, J. and Neese, F. (**2021**) DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings, *J. Chem. Phys.*, 154, 164110.
4.
5. Tran, V. A.; Neese, F. (**2020**) Double-hybrid density functional theory for g-tensor calculations using gauge including atomic orbitals, *J. Chem. Phys.*, 153, 13.
6. Dittmer, A.; Stoychev, G. L.; Maganas, D.; Auer, A. A.; Neese, F. (**2020**) Computation of NMR Shielding Constants for Solids Using an Embedded Cluster Approach with DFT, Double-Hybrid DFT, and MP2, *J. Chem. Theo. Comp.*, 16, 6950-6967.
7. Auer, A. A.; Tran, V. A.; Sharma, B.; Stoychev, G. L.; Marx, D.; Neese, F. (**2020**) A case study of density functional theory and domain-based local pair natural orbital coupled cluster for vibrational effects on EPR hyperfine coupling constants: vibrational perturbation theory versusab initiomolecular dynamics, *Mol. Phys.*, 16.
8. Lang, L.; Ravera, E.; Parigi, G.; Luchinat, C.; Neese, F. (**2020**) Solution of a Puzzle: High-Level Quantum-Chemical Treatment of Pseudocontact Chemical Shifts Confirms Classic Semiempirical Theory, *J. Phys. Chem. Lett.*, 11, 8735-8744.
9. Sirohiwal, A.; Neese, F.; Pantazis, D. A. (**2019**) Microsolvation of the Redox-Active Tyrosine-D in Photosystem II: Correlation of Energetics with EPR Spectroscopy and Oxidation-Induced Proton Transfer, *J. Am. Chem. Soc.*, 141, 3217-3231.

A.3.13 ESD Module

Fluorescence

1. de Souza, B.; Neese F.; Izsak (**2018**) On the theoretical prediction of fluorescence rates from first principles using the path integral approach, *J. Chem. Phys.*, 148, 034104.

Phosphorescence

1. de Souza, B.; Farias, G.; Neese F.; Izsak (**2019**) Predicting Phosphorescence Rates of Light Organic Molecules Using Time-Dependent Density Functional Theory and the Path Integral Approach to Dynamics, *J. Chem. Theo. Comp.*, 15, 1896.

Resonance Raman

1. De Souza, B.; Farias, G.; Neese, F.; Izsak, R. (**2019**) Efficient simulation of overtones and combination bands in Resonant Raman spectra, *J. Chem. Phys.*, accepted - waiting for publication.

A.3.14 DFT/Hartree–Fock Theory of EPR Parameters

1. Sandhoefer, B.; Neese, F. (**2012**) One-electron contributions to the g-tensor for second-order Douglas–Kroll–Hess theory, *J. Chem. Phys.*, 137, 094102.
2. Ganyushin, D.; Gilka, N.; Taylor, P. R.; Marian, C. M.; Neese, F. (**2010**) The resolution of the identity approximation for calculations of spin-spin contribution to zero-field splitting parameters, *J. Chem. Phys.*, 132, 144111.
3. Duboc, C.; Ganyushin, D.; Sivalingam, K.; Collomb, M. N.; Neese, F. (**2010**) Systematic Theoretical Study of the Zero-Field Splitting in Coordination Complexes of Mn(III). Density Functional Theory versus Multireference Wave Function Approaches, *J. Phys. Chem. A*, 114, 10750–10758.
4. Neese, F. (**2009**) First principles approach to Spin-Hamiltonian Parameters, invited chapter in Misra, S.K. Multifrequency EPR: Theory and Applications, Wiley-VCH, pp. 297–326.
5. Pantazis, D. A.; Orio, M.; Petrenko, T.; Messinger, J.; Lubitz, W.; Neese, F. (**2009**) A new quantum chemical approach to the magnetic properties of oligonuclear transition metal clusters: Application to a model for the tetranuclear manganese cluster of Photosystem II *Chem. Eur. J.*, 15(20), 5108–5123.
6. Riplinger, C.; Kao, J.P.Y.; Rosen, G.M.; Kathirvelu, V.; Eaton, G.R.; Eaton, S.S.; Kutateladze, A.; Neese F. (**2009**) Interaction of Radical Pairs Through-Bond and Through-Space: Scope and Limitations of the Point-Dipole Approximation in Electron Paramagnetic Resonance Spectroscopy, *J. Am. Chem. Soc.*, 131, 10092–10106.
7. Cirera, J.; Ruiz, E.; Alvarez, S.; Neese, F.; Kortus, J. (**2009**) How to Build Molecules with Large Magnetic Anisotropy. *Chem. Eur. J.*, 15(16), 4078–4087.
8. Neese, F. (**2008**) Spin Hamiltonian Parameters from First Principle Calculations: Theory and Application. In Hanson, G.; Berliner, L. (Eds), *Biological Magnetic Resonance*, pp. 175–232.
9. Koßmann, S.; Kirchner, B.; Neese, F. (**2007**) Performance of modern density functional theory for the prediction of hyperfine structure: meta-GGA and double hybrid functionals, *Molec. Phys.* (Arthur Schweiger memorial issue), 105, 2049–2071.
10. Neese, F. (**2007**) Calculation of the Zero-Field Splitting Tensor Using Hybrid Density Functional and Hartree-Fock Theory. *J. Chem. Phys.*, 127, 164112.
11. Neese, F. (**2006**) Importance of Direct Spin-Spin Coupling and Spin-Flip Excitations for the Zero-Field Splittings of Transition Metal Complexes: A Case Study, *J. Am. Chem. Soc.*, 128, 10213–10222.
12. Sinnecker, S.; Neese, F. (**2006**) Spin-Spin Contributions to the Zero-Field Splitting Tensor in Organic Triplets, Carbenes and Biradicals – A Density Functional and *ab initio* Study. *J. Phys. Chem. A*, 110, 12267–12275.
13. Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F. (**2006**) Calculation of Solvent Shifts on Electronic G-Tensors with the Conductor-Like Screening Model (COSMO) and its Self-Consistent Generalization to Real Solvents (COSMO-RS), *J. Phys. Chem. A*, 110, 2235–2245.
14. Neese, F.; Wolf, A.; Reiher, M.; Fleig, T.; Hess, B.A. (**2005**) Higher Order Douglas-Kroll Calculation of Electric Field Gradients, *J. Chem. Phys.*, 122, 204107.

15. Neese, F. (**2005**) Efficient and Accurate Approximations to the Molecular Spin-Orbit Coupling Operator and their use in Molecular g-Tensor Calculations, *J. Chem. Phys.*, 122, 034107.

16. Ray, K.; Begum, A; Weyhermüller, T.; Piligkos, S.; van Slageren, J.; Neese, F.; Wieghardt, K. (**2005**) The Electronic Structure of the Isoelectronic, Square Planar Complexes $[Fe^{II}(L)_2]^{2-}$ and $[Co^{III}(L^{Bu})_2]^{-}$ (L^{2-} and $(L^{Bu})^{2-}$ = benzene-1,2-dithiolates): an Experimental and Density Functional Theoretical Study, *J. Am. Chem. Soc.*, 127, 4403–4415.

17. Neese, F. (**2003**) Metal and Ligand Hyperfine Couplings in Transition Metal Complexes. The Effect of Spin-Orbit Coupling as Studied by Coupled Perturbed Kohn-Sham Theory and Hybrid Density Functionals, *J. Chem. Phys.*, 117, 3939–3948.

18. Neese, F. (**2001**) Prediction of Electron Paramagnetic Resonance g-values by Coupled Perturbed Hartree-Fock and Kohn-Sham Theory. *J. Chem. Phys.*, 115, 11080–11096.

19. Neese, F. (**2001**) Theoretical Study of Ligand Superhyperfine Structure. Application to Cu(II) Complexes. *J. Phys. Chem. A*, 105, 4290–4299.

A.3.15 Dispersion Corrections to DFT

(not originally implemented in but the implementation is based on the code described in these papers)

1. Grimme, S. (**2004**) *J. Comput. Chem.*, 25, 1463–1476.

2. Grimme, S. (**2006**) *J. Comput. Chem.*, 27, 1787–1799.

3. Grimme, S.; Antony, J. Ehrlich, S. Krieg, H. (**2010**) *J. Chem. Phys.*, 132, 154104.

4. Grimme, S.; Ehrlich, S.; Goerigk, L. (**2011**) *J. Comput. Chem.*, 32, 1456-1465.

A.3.16 DLPNO

1. Stoychev, G. L., Auer, A. A., Gauss, J. and Neese, F. (**2021**) DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings, *J. Chem. Phys.*, 154, 164110.

2. Ni, Z. G.; Guo, Y.; Neese, F.; Li, W.; Li, S. H. (**2021**) Cluster-in-Molecule Local Correlation Method with an Accurate Distant Pair Correction for Large Systems, *J. Chem. Theo. Comp.*, 17, 756-766.

3. Liakos, D. G.; Guo, Y.; Neese, F. (**2020**) Comprehensive Benchmark Results for the Domain Based Local Pair Natural Orbital Coupled Cluster Method (DLPNO-CCSD(T)) for Closed- and Open-Shell Systems, *J. Phys. Chem. A*, 124, 90-100.

4. Guo, Y.; Riplinger, C.; Liakos, D. G.; Becker, U.; Saitow, M.; Neese, F. (**2020**) Linear scaling perturbative triples correction approximations for open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory DLPNO-CCSD(T-0/T), *J. Chem. Phys.*, 152.

5. Altun, A.; Neese, F.; Bistoni, G. (**2020**) Extrapolation to the Limit of a Complete Pair Natural Orbital Space in Local Coupled-Cluster Calculations, *J. Chem. Theo. Comp.*, 16, 6142-6149.

6. Pinski, P.; Neese, F. (**2019**) Analytical gradient for the domain-based local pair natural orbital second order Møller-Plesset perturbation theory method (DLPNO-MP2), *J. Chem. Phys.*, 150, 164102.

7. Pinski, P.; Neese, F. (**2018**) Communication: Exact analytical derivatives for the domain-based local pair natural orbital MP2 method (DLPNO-MP2), *J. Chem. Phys.*, 148, 031101.

8. Sparta, M.; Retegan, M.; Pinski, P.; Riplinger, C.; Becker, U.; Neese, F. (**2017**) Multilevel Approaches within the Local Pair Natural Orbital Framework, *J. Chem. Theory Comput.*, 13, 3198-3207.

9. Pavosevic, F; Peng, C.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E.F. (**2017**) SparseMaps-A systematic infrastructure for reduced scaling electronic structure methods. V. Linear scaling explicitly correlated coupled-cluster method with pair natural orbitals, *J. Chem. Phys.*, 146, 174108.

10. Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. (**2016**) Decomposition of Intermolecular Interaction Energies within the Local Pair Natural Orbital Coupled Cluster Framework, *J. Chem. Theory Comput.*, 12, 4778-4792.

11. Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. (**2016**) Sparse maps-A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory, *J. Chem. Phys.*, 144, 024109.

12. Pavosevic, F.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. (**2016**) SparseMaps-A systematic infrastructure for reduced-scaling electronic structure methods. IV. Linear-scaling second-order explicitly correlated energy with pair natural orbitals, *J. Chem. Phys.*, 144, 144109.

13. Kubas, A.; Berger, D.; Oberhofer, H.; Maganas, D.; Reuter, K.; Neese, F. (**2016**) Surface Adsorption Energetics Studied with "Gold Standard" Wave Function-Based Ab Initio Methods: Small-Molecule Binding to TiO2(110), *J. Phys. Chem. Lett.*, 7, 4207-4212.

14. Isegawa, M.; Neese, F.; Pantazis, D. A. (**2016**) Ionization Energies and Aqueous Redox Potentials of Organic Molecules: Comparison of DFT, Correlated ab Initio Theory and Pair Natural Orbital Approaches, *J. Chem. Theory Comput.*, 12, 2272-2284.

15. Guo, Y.; Sivalingam, K.; Valeev, E. F.; Neese, F. (**2016**) SparseMaps-A systematic infrastructure for reduced-scaling electronic structure methods. III. Linear-scaling multireference domain-based pair natural orbital N-electron valence perturbation theory, *J. Chem. Phys.*, 144, 094111.

16. Dutta, A. K.; Neese, F.; Izsak, R. (**2016**) Towards a pair natural orbital coupled cluster method for excited states, *J. Chem. Phys.*, 145, 034102.

17. Datta, D.; Kossmann, S.; Neese, F. (**2016**) Analytic energy derivatives for the calculation of the first-order molecular properties using the domain-based local pair-natural orbital coupled-cluster theory, *J. Chem. Phys.*, 145, 114101.

18. Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. (**2016**) Sparse maps-A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals, *J. Chem. Phys.*, 143, 034108.

19. Mondal, B.; Neese, F.; Ye, S. F. (**2015**) Control in the Rate-Determining Step Provides a Promising Strategy To Develop New Catalysts for CO2 Hydrogenation: A Local Pair Natural Orbital Coupled Cluster Theory Study, *Inorg. Chem.*, 54, 7192-7198.

20. Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. (**2015**) Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory, *J. Chem. Theory Comput.*, 11, 1525-1539.

21. Liakos, D. G.; Neese, F. (**2015**) Domain Based Pair Natural Orbital Coupled Cluster Studies on Linear and Folded Alkane Chains, *J. Chem. Theory Comput.*, 11, 2137-2143.

22. Liakos, D. G.; Neese, F. (**2015**) Is It Possible To Obtain Coupled Cluster Quality Energies at near Density Functional Theory Cost? Domain-Based Local Pair Natural Orbital Coupled Cluster vs Modern Density Functional Theory, *J. Chem. Theory Comput.*, 11, 4054-4063.

23. Demel, O.; Pittner, J.; Neese, F. (**2015**) A Local Pair Natural Orbital-Based Multireference Mukherjee's Coupled Cluster Method, *J. Chem. Theory Comput.*, 11, 3104-3114.

A.3.17 Double hybrid density functionals

1. Grimme, S.; Neese, F. (**2007**) Double Hybrid Density Functional Theory for Excited States of Molecules, *J. Chem. Phys.*, 127, 154116.

2. Neese, F.; Schwabe, T.; Grimme, S. (**2007**) Analytic Derivatives for Perturbatively Corrected 'Double Hybrid' Density Functionals, *J. Chem. Phys.*, 126, 124115.

3. Koßmann, S.; Kirchner, B.; Neese, F. (**2007**) Performance of modern density functional theory for the prediction of hyperfine structure: meta-GGA and double hybrid functionals, *Mol. Phys.* (Arthur Schweiger memorial issue), 105, 2049–2071.

A.3.18 Excited States Methods and Resonance Raman Spectra

1. Schapiro, I.; Sivalingam K.; Neese, F. (**2015**) Assessment of n-Electron Valence State Perturbation Theory for Vertical Excitation Energies, *J. Chem. Theory Comput.*, 9, 3567.
2. Roemelt, M.; Neese, F. (**2013**) Excited States of Large Open-Shell Molecules: An Efficient, General, and Spin-Adapted Approach Based on a Restricted Open-Shell Ground State Wave function, *J. Phys. Chem. A*, 117, 3069.
3. Petrenko, T.; Kossmann, S.; Neese, F. (**2011**) Efficient time-dependent density functional theory approximations for hybrid density functionals: Analytical gradients and parallelization, *J. Chem. Phys.*, 134, 054116.
4. Petrenko, T.; Krylova, O.; Neese, F.; Sokolowski, M. (**2009**) Optical Absorption and Emission Properties of Rubrene: Insight by a Combined Experimental and Theoretical Study. *New J. Phys.*, 11, 015001.
5. Grimme, S.; Neese, F. (**2007**) Double Hybrid Density Functional Theory for Excited States of Molecules, *J. Chem. Phys.*, 127, 154116.
6. Petrenko, T.; Ray, K.; Wieghardt, K.; Neese, F. (**2006**) Vibrational Markers for the Open-Shell Character of Metal bis-Dithiolenes: An Infrared, resonance Raman and Quantum Chemical Study. *J. Am. Chem. Soc.*, 128, 4422–4436.
7. Neese, F.; Olbrich, G. (**2002**) Efficient use of the Resolution of the Identity Approximation in Time-Dependent Density Functional Calculations with Hybrid Functionals, *Chem. Phys. Lett.*, 362, 170–178.

A.3.19 F12

1. Pavosevic, F.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. (**2016**) SparseMaps-A systematic infrastructure for reduced-scaling electronic structure methods. IV. Linear-scaling second-order explicitly correlated energy with pair natural orbitals, *J. Chem. Phys.*, 144, 144109.
2. Guo, Y.; Sivalingam, K.; Valeev, E.F.and Neese, F. (**2017**), Explicitly correlated N-electron valence state perturbation theory (NEVPT2-F12), *J. Chem. Phys.*, 147, 064110.
3. Liakos D. G.; Izsák R.; F.; Valeev, E. F.; Neese, F. (**2013**) What is the most efficient way to reach the canonical MP2 basis set limit?, *Mol. Phys.*, 111, 2653

A.3.20 FOD analysis and FOD plots

1. Grimme, S.; Hansen, A. (**2015**) A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects, *Angew. Chem. Int. Ed.*, 54, 12308.

A.3.21 Gaussian Charge Scheme (C-PCM) Implementation

1. Garcia-Ratés, M.; Neese, F. (**2020**) Effect of the Solute Cavity on the Solvation Energy and its Derivatives within the Framework of the Gaussian Charge Scheme, *J. Comput. Chem.*, 41, 922-939.

A.3.22 gCP correction to HF and DFT

(ORCA implementation is based upon the code used in this paper)

1. Kruse, H; Grimme, S. (**2012**) *J. Chem. Phys.*, 126, 154101.

A.3.23 HF-3c method

1. Sure, R; Grimme, S. (**2013**) *J. Comput. Chem.*, 34, 1672-1685.

A.3.24 Internally Contracted Multireference CI

1. Sivalingam, K.; Krupicka, M.; Auer, A. A.; Neese, F. (**2016**) Comparison of fully internally and strongly contracted multireference configuration interaction procedures. *J. Chem. Phys.*, 145, 54104.

A.3.25 Magnetic Circular Dichroism Spectra

1. Westphal, A.; Broda, H.; Kurz, P.; Neese, F.; Tuczek, F. (**2012**) Magnetic Circular Dichroism Spectrum of the Molybdenum(V) Complex (Mo(O)Cl_3dppe): C-Term Signs and Intensities for Multideterminant Excited Doublet States, *Inorg. Chem.*, 51, 5748–5763.
2. van Slageren, J.; Piligkos, S.; Neese, F. (**2010**) Magnetic circular dichroism spectroscopy on the Cr(8) antiferromagnetic ring, *Dalton Trans.*, 39, 4999–5004.
3. Sundararajan, M.; Ganyushin, D.; Ye, S.; Neese, F. (**2009**) Multireference *ab initio* studies of Zero-Field Splitting and Magnetic Circular Dichroism Spectra of Tetrahedral Co(II) Complexes, *Dalton Trans*., 30, 6021–6036.
4. Piligkos, S.; Slep, L.; Weyhermüller, T.; Chaudhuri, P.; Bill, E.; Neese, F. (**2009**) Magnetic Circular Dichroism Spectroscopy of weakly exchange coupled dimers. A model study. *Coord. Chem. Rev.,* 253, 2352–2362.
5. Ganyushin, D.; Neese, F. (**2008**) First principles calculation of magnetic circular dichroism spectra, *J. Chem. Phys.*, 128, 114117.
6. Neese, F.; Solomon, E.I. (**1999**) MCD C-term Signs, Saturation Behavior and Determination of Band Polarizations in Randomly Oriented Systems with Spin $S \geqslant 1/2$. Applications to $S = 1/2$ and $S = 5/2$. *Inorg. Chem.*, 38, 1847–1865.

A.3.26 MCD

1. Ye, S. F.; Kupper, C.; Meyer, S.; Andris, E.; Navratil, R.; Krahe, O.; Mondal, B.; Atanasov, M.; Bill, E.; Roithova, J.; Meyer, F.; Neese, F. (**2016**) Magnetic Circular Dichroism Evidence for an Unusual Electronic Structure of a Tetracarbene-Oxoiron(IV) Complex, *J. Am. Chem. Soc.*, 138, 14312-14325.
2. Ye, S. F.; Xue, G. Q.; Krivokapic, I.; Petrenko, T.; Bill, E.; Que, L.; Neese, F. (**2015**) Magnetic circular dichroism and computational study of mononuclear and dinuclear iron(IV) complexes, *Chem. Sci.*, 6, 2909-2921.

A.3.27 MDCI

1. Veis, L.; Antalik, A.; Brabec, J.; Neese, F.; Legeza, O.; Pittner, J. (**2016**) Coupled Cluster Method with Single and Double Excitations Tailored by Matrix Product State Wave Functions, *J. Phys. Chem. Lett.*, 7, 4072-4078.
2. Dutta, A. K.; Neese, F.; Izsak, R. (**2016**) Speeding up equation of motion coupled cluster theory with the chain of spheres approximation, *J. Chem. Phys.*, 144, 034102.

A.3.28 Mössbauer Spectroscopy

1. Datta, D.; Saitow, M.; Sandhofer, B.; Neese, F. (**2020**) Fe-57 Mossbauer parameters from domain based local pair-natural orbital coupled-cluster theory, *J. Chem. Phys.*, 153.
2. Römelt, M.; Ye, S.; Neese, F. (**2009**) Calibration of Mössbauer Isomer Shift Calculations for Modern Density Functional Theory: meta-GGA and Double Hybrid Functionals *Inorg. Chem.*, 48, 784–785.
3. Sinnecker, S.; Slep, L.; Bill, E.; Neese, F. (**2005**) Performance of Nonrelativistic and Quasirelativistic Hybrid DFT for the Prediction of Electric and Magnetic Hyperfine Parameters in ^{57}Fe Mössbauer Spectra, *Inorg. Chem.*, 44, 2245–2254.
4. Neese, F. (**2002**) Prediction and Interpretation of Isomer Shifts in ^{57}Fe Mössbauer Spectra by Density Functional Theory. *Inorg. Chim. Acta* (special Karl Wieghardt honorary issue), 337C, 181–192.

A.3.29 Multireference CI Module and its application to EPR properties and optical spectra

1. Lang, L.; Sivalingam, K.; Neese, F. (**2020**) The combination of multipartitioning of the Hamiltonian with canonical Van Vleck perturbation theory leads to a Hermitian variant of quasidegenerate N-electron valence perturbation theory, *J. Chem. Phys.*, 152.
2. Demel, O.; Pittner, J.; Neese, F. (**2015**) A Local Pair Natural Orbital-Based Multireference Mukherjee's Coupled Cluster Method, *J. Chem. Theory Comput.*, 11, 3104.
3. Nooijen, M.; Demel, O.; Datta, D.; Kong, L.; Shamasundar K. R.; Lotrich, V.; Huntington, L. M.; Neese, F. (**2014**), *J. Chem. Phys.*, 140, 081102.
4. Ganyushin, D.; Neese, F. (**2013**) A fully variational spin-orbit coupled complete active space self-consistent field approach: Application to electron paramagnetic resonance g-tensors, *J. Chem. Phys.*, 138, 104113.
5. Atanasov, M.; Zadrozny, J. M.; Long, J. R.; Neese, F. (**2013**) A theoretical analysis of chemical bonding, vibronic coupling, and magnetic anisotropy in linear iron(II) complexes with single-molecule magnet behavior, *Chem. Sci.*, 4, 139–156.
6. Atanasov, M.; Ganyushin, D.; Pantazis, D. A.; Sivalingam, K.; Neese, F. (**2011**) Detailed Ab Initio First-Principles Study of the Magnetic Anisotropy in a Family of Trigonal Pyramidal Iron(II) Pyrrolide Complexes, *Inorg. Chem,* 50, 7460–7477.
7. Neese, F.; Pantazis, D. A. (**2011**) What is not required to make a single molecule magnet, *Faraday Discussions*, 148, 229–238.
8. Duboc, C.; Ganyushin, D.; Sivalingam, K.; Collomb, M. N.; Neese, F. (**2010**) Systematic Theoretical Study of the Zero-Field Splitting in Coordination Complexes of Mn(III). Density Functional Theory versus Multireference Wave Function Approaches, *J. Phys. Chem. A*, 114, 10750–10758.
9. Sundararajan, M.; Ganyushin, D.; Ye, S.; Neese, F. (**2009**) Multireference *ab initio* studies of Zero-Field Splitting and Magnetic Circular Dichroism Spectra of Tetrahedral Co(II) Complexes, *Dalton Trans*., 30, 6021–6036.
10. Ganyushin, D.; Neese, F. (**2008**) First principles calculation of magnetic circular dichroism spectra, *J. Chem. Phys.*, 128, 114117.
11. Petrenko, T.; Neese, F. (**2007**) A general efficient quantum chemical method for predicting absorption bandshapes, resonance Raman spectra and excitation profiles for larger molecules. *J. Chem. Phys.*, 127, 164319.
12. Neese, F. (**2007**) Analytic Derivative Calculation of Electronic g-Tensors based on Multireference Configuration Interaction Wavefunctions. *Mol. Phys.* (honorary issue for Prof. Peter Pulay), 105, 2507–2514.
13. Neese, F.; Petrenko, T.; Ganyushin, D.; Olbrich, G. (**2007**) Advanced Aspects of *ab initio* Theoretical Spectroscopy of Open-Shell Transition Metal Ions. *Coord. Chem. Rev.*, 205, 288–327.
14. Neese, F. (**2006**) Importance of Direct Spin-Spin Coupling and Spin-Flip Excitations for the Zero-Field Splittings of Transition Metal Complexes: A Case Study, *J. Am. Chem. Soc.*, 128, 10213–10222.

15. Chalupský, J.; Neese, F.; Solomon, E.I.; Ryde, U.; Rulíšek, L. (**2006**) Identification of intermediates in the reaction cycle of multicopper oxidases by quantum chemical calculations of spectroscopic parameters, *Inorg. Chem.*, 45, 11051–11059.

16. Neese, F. (**2006**) Theoretical spectroscopy of model-nonheme $[Fe^{IV}O(NH_3)_5]^{2+}$ complexes with triplet and quintet ground states using multireference *ab initio* and density functional theory methods. *J. Inorg. Biochem.* (special issue on high-valent Fe(IV)), 716–726.

17. Ganyushin, D.; Neese, F. (**2006**) First Principle Calculation of Zero-Field Splittings, *J. Chem. Phys.*, 125, 024103.

18. Ray, K.; Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2005**) Electronic Structure of Square-Planar Bis(benzene-1,2-dithiolate)metal Complexes $[M(L)_2]^z$ ($z = 2-, 1-, 0$; M = Ni, Pd, Pt, Cu, Au): An experimental, Density Functional and Correlated *ab initio* Study. *Inorg. Chem.*, 44, 5345–5360.

19. Schöneboom, J.; Neese, F.; Thiel, W. (**2005**) Towards Identification of the Compound I Reactive Intermediate in Cytochrome P450 Chemistry: A QM/MM Study of its EPR and Mössbauer Parameters, *J. Am. Chem. Soc.*, 127, 5840–5853.

20. Wanko, M.; Hoffmann, M.; Strodel, P.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. (**2005**) Calculating Absorption Shifts for Retinal Proteins: Computational Challenges *J. Phys. Chem. B*, 109, 3606–3615.

21. Neese, F. (**2004**) Sum Over States Based Multireference *ab initio* Calculation of EPR Spin Hamiltonian Parameters for Transition Metal Complexes. A Case Study *Mag. Res. Chem.*, 42, S187–S198.

22. Neese, F. (**2003**) Correlated *ab Initio* Calculation of Electronic g-Tensors Using a Sum Over States Formulation. *Chem. Phys. Lett.*, 380/5–6, 721–728.

23. Neese, F. (**2003**) A Spectroscopy Oriented Configuration Interaction Procedure, *J. Chem. Phys.*, 119, 9428–9443.

24. Neese, F. (**2001**) Configuration Interaction Calculation of Electronic g-Tensors in Transition Metal Complexes, *Int. J. Quant. Chem.*, 83, 104–114.

25. Neese, F.; Solomon, E.I. (**1998**) Calculation of Zero-Field Splittings, g-values and the Relativistic Nephelauxetic Effect in Transition Metal Complexes. Application to High Spin Ferric Complexes. *Inorg. Chem.*, 37, 6568–6582.

A.3.30 NRVS

1. Ogata, H.; Kramer, T.; Wang, H. X.; Schilter, D.; Pelmenschikov, V.; van Gastel, M.; Neese, F.; Rauchfuss, T. B.; Gee, L. B.; Scott, A. D.; Yoda, Y.; Tanaka, Y.; Lubitz, W.; Cramer, S. P. (**2015**) Hydride bridge in NiFe -hydrogenase observed by nuclear resonance vibrational spectroscopy, *Nat. Commun.*, 6, 7890.

A.3.31 Nuclear Resonance Vibrational Spectra

1. Petrenko T,; Sturhahn, W.; Neese, F. (**2008**) First principles calculation of Nuclear Resonance Vibrational Spectra, *Hyperfine Interactions*, 175, 165–174.

2. DeBeer-George, S.; Petrenko, T.; Aliaga-Alcade, N.; Bill, E.; Mienert, B.; Sturhan, W.; Ming, Y.; Wieghardt, K.; Neese, F. (**2007**) Characterization of a Genuine Iron(V)Nitrido Species by Nuclear Resonant Vibrational Spectroscopy Coupled to Density Functional Calculations, *J. Am. Chem. Soc.*, 129, 11053–11060.

A.3.32 Orbital Optimized MP2

1. Shirazi, R. G.; Pantazis, D. A.; Neese, F. (**2020**) Performance of density functional theory and orbital-optimised second-order perturbation theory methods for geometries and singlet-triplet state splittings of aryl-carbenes, *Mol. Phys.*, 118.
2. Sandhoefer, B.; Kossmann, S.; Neese, F. (**2013**) Derivation and assessment of relativistic hyperfine-coupling tensors on the basis of orbital-optimized second-order Moller-Plesset perturbation theory and the second-order Douglas–Kroll–Hess transformation, *J. Chem. Phys.*, 138, 104102.
3. Kossmann, S.; Neese, F. (**2010**) Correlated ab Initio Spin Densities for Larger Molecules: Orbital-Optimized Spin-Component-Scaled MP2 Method, *J. Phys. Chem. A*, 114, 11768-11781.
4. Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. (**2009**) Assessment of Orbital Optimized, Spin-Component Scaled Second Order Many Body Perturbation Theory for Thermochemistry and Kinetics. *J. Chem. Theory Comput.*, 5, 3060-3073.

A.3.33 Pair Natural Orbital Local Correlation Methods

1. Riplinger, C.; Sandhoefer B.; Hansen, A.; Neese, F. (**2013**) Natural triple excitations in local coupled cluster calculations with pair natural orbitals, *J. Chem. Phys.*, 139, 134101.
2. Riplinger, C.; Neese, F. (**2013**) An efficient and near linear scaling pair natural orbital based local coupled cluster method, *J. Chem. Phys.*, 138, 034106.
3. Liakos, D. G.; Neese, F. (**2012**) Improved Correlation Energy Extrapolation Schemes Based on Local Pair Natural Orbital Methods, *J. Phys. Chem. A*, 116, 4801–4816.
4. Huntington, L. M. J.; Hansen, A.; Neese, F.; Nooijen, M. (**2012**) Accurate thermochemistry from a parameterized coupled-cluster singles and doubles model and a local pair natural orbital based implementation for applications to larger systems, *J. Chem. Phys.*, 136, 064101.
5. Izsák, R.; Hansen, A.; Neese, F. (**2012**) The resolution of identity and chain of spheres approximations for the LPNO-CCSD singles Fock term, *Mol. Phys.*, 110, 2413–2417.
6. Liakos, D. G.; Hansen, A.; Neese, F. (**2011**) Weak Molecular Interactions Studied with Parallel Implementations of the Local Pair Natural Orbital Coupled Pair and Coupled Cluster Methods, *J. Chem. Theory Comput.*, 7, 76–87.
7. Hansen, A.; Liakos, D. G.; Neese, F. (**2011**) Efficient and accurate local single reference correlation methods for high-spin open-shell molecules using pair natural orbitals, *J. Chem. Phys.*, 135, 214102.
8. Kollmar, C.; Neese, F. (**2011**) An orbital-invariant and strictly size extensive post-Hartree-Fock correlation functional, *J. Chem. Phys.*, 135, 084102.
9. Kollmar, C.; Neese, F. (**2011**) The relationship between double excitation amplitudes and Z vector components in some post-Hartree-Fock correlation methods, *J. Chem. Phys.*, 135, 064103.
10. Neese, F.; Liakos, D.; Hansen, A. (**2009**) Efficient and accurate local approximations to the coupled cluster singles and doubles method using a truncated pair natural orbital basis *J. Chem. Phys.*, 131, 064103.
11. Neese, F.; Wennmohs, F.; Hansen, A. (**2009**) Efficient and accurate local approximations to coupled electron pair approaches. An attempt to revive the pair-natural orbital method *J. Chem. Phys.*, 130, 114108.

A.3.34 PBEh-3c method

1. Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. (**2015**) Consistent structures and interactions by density functional theory with small atomic orbital basis sets, *J. Chem. Phys.*, 143, 054107 .

A.3.35 QM/MM

1. Rokhsana, D.; Large, T. A. G.; Dienst, M. C.; Retegan, M.; Neese, F. (**2016**) A realistic in silico model for structure/function studies of molybdenum-copper CO dehydrogenase Journal of Biological, *Inorg. Chem.*, 21, 491-499.
2. Retegan, M.; Krewald, V.; Mamedov, F.; Neese, F.; Lubitz, W.; Cox, N.; Pantazis, D. A. (**2016**) A five-coordinate Mn(IV) intermediate in biological water oxidation: spectroscopic signature and a pivot mechanism for water binding, *Chem. Sci.*, 7, 72-84.
3. Sundararajan, M.; Neese, F. Distal (**2015**) Histidine Modulates the Unusual O-Binding of Nitrite to Myoglobin: Evidence from the Quantum Chemical Analysis of EPR Parameters, *Inorg. Chem.*, 54, 7209-7217.

A.3.36 QM/MM calculations with ORCA

1. Schulz, C. E.; van Gastel, M.; Pantazis, D. A.; Neese, F. (**2021**) Converged Structural and Spectroscopic Properties for Refined QM/MM Models of Azurin, *Inorg. Chem.*, 60, 7399-7412.
2.
3. Schulz, C. E.; Castillo, R. G.; Pantazis, D. A.; DeBeer, S.; Neese, F. (**2021**) Structure-Spectroscopy Correlations for Intermediate Q of Soluble Methane Monooxygenase: Insights from QM/MM Calculations, *J. Am. Chem. Soc.*, 143, 6560-6577.
4. Sundararajan, M.; Neese, F. (**2012**) Detailed QM/MM study of the Electron Paramagnetic Resonance Parameters of Nitrosyl Myoglobin, *J. Chem. Theory Comput.*, 8, 563–574.
5. Riplinger, C.; Neese, F. (**2011**) The reaction mechanism of Cytochrome P450 NO Reductase: A Detailed Quantum Mechanics/Molecular Mechanics Study, *ChemPhysChem*, 12, 3192–3203.
6. Radoul, M.; Sundararajan, M.; Potapov, A.; Riplinger, C.; Neese, F.; Goldfarb, D. (**2010**) Revisiting the nitrosyl complex of myoglobin by high-field pulse EPR spectroscopy and quantum mechanical calculations, *Phys. Chem. Chem. Phys.*, 12, 7276–7289.
7. Sundararajan, M.; Riplinger, C.; Orio, M.; Wennmohs, F.; Neese, F. (**2009**) Spectroscopic Properties of Protein-Bound Cofactors: Calculation by Combined Quantum Mechanical/Molecular Mechanical (QM/MM) Approaches, *Encyc. Inorg. Chem.*, DOI: 10.1002/9781119951438.eibc0371.
8. Altun, A.; Kumar, D.; Neese, F.; Thiel, W. (**2008**) Multi-reference Ab Initio QM/MM Study on Intermediates in the Catalytic Cycle of Cytochrome P450cam, *J. Phys. Chem.*, 112, 12904–12910.
9. Chalupský, J.; Neese, F.; Solomon, E.I.; Ryde, U.; Rulíšek, L. (**2006**) Identification of intermediates in the reaction cycle of multicopper oxidases by quantum chemical calculations of spectroscopic parameters, *Inorg. Chem.*, 45, 11051–11059.
10. Sinnecker, S.; Neese, F. (**2006**) QM/MM Calculations with DFT for Taking into Account Protein Effects on the EPR and Optical Spectra of Metalloproteins. Plastocyanin as a Case Study. *J. Comp. Chem. (Special issue on Theoretical Bioinorganic Chemistry)*, 27, 1463–1475.
11. Wanko, M.; Hoffmann, M.; Strodel, P.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. (**2005**) Calculating Absorption Shifts for Retinal Proteins: Computational Challenges *J. Phys. Chem. B*, 109, 3606–3615.
12. Schöneboom, J.; Neese, F.; Thiel, W. (**2005**) Towards Identification of the Compound I Reactive Intermediate in Cytochrome P450 Chemistry: A QM/MM Study of its EPR and Mössbauer Parameters, *J. Am. Chem. Soc.*, 127, 5840–5853.

A.3.37 Relativity and SARC Basis Sets

1. Rolfes, J. D.; Neese, F.; Pantazis, D. A. (**2020**) All-electron scalar relativistic basis sets for the elements Rb–Xe, *J. Comput. Chem.*, 41, 1842-1849.
2. Aravena, D.; Neese, F.; Pantazis, D. A. (**2016**) Improved Segmented All-Electron Relativistically Contracted Basis Sets for the Lanthanides, *J. Chem. Theory Comput.*, 12, 1148-1156.
3. Pantazis, D. A.; Neese, F. (**2012**) All-Electron Scalar Relativistic Basis Sets for the 6p Elements, *Theor. Chem. Acc.*, 131, 1292.
4. Pantazis, D. A.; Neese, F. (**2011**) All-Electron Scalar Relativistic Basis Sets for the Actinides, *J. Chem. Theory Comput.*, 7, 677–684.
5. Pantazis, D. A.; Neese, F. (**2009**) All-Electron Scalar Relativistic Basis Sets for the Lanthanides, *J. Chem. Theory Comput.*, 5, 2229–2238.
6. Bühl, M.; Reimann, C.; Pantazis, D. A.; Bredow, T.; Neese, F. (**2008**) Geometries of Third-row Transition-Metal Complexes from Density-Functional Theory *J. Chem. Theory Comput.*, 4, 1449–1459.
7. Pantazis, D. A.; Chen, X.-Y.; Landis, C.R.; Neese, F. (**2008**) All Electron Scalar Relativistic Basis Sets for Third Row Transition Metal Atoms. *J. Chem. Theory Comput.*, 4, 908–919.

A.3.38 Resonance Raman

1. Maganas, D.; Trunschke, A.; Schlogl, R.; Neese, F. (**2016**) A unified view on heterogeneous and homogeneous catalysts through a combination of spectroscopy and quantum chemistry, *Farad. Discuss.*, 188, 181-197.
2. De Souza, B.; Farias, G.; Neese, F.; Izsak, R. (**2019**) Efficient simulation of overtones and combination bands in Resonant Raman spectra, *J. Chem. Phys.*, accepted - waiting for publication.

A.3.39 SOC on TD-DFT

1. de Souza, B.; Farias, G.; Neese F.; Izsak (**2019**) Predicting Phosphorescence Rates of Light Organic Molecules Using Time-Dependent Density Functional Theory and the Path Integral Approach to Dynamics, *J. Chem. Theo Comp.*, 15, 1896.

A.3.40 SOSCF Method

1. Neese, F. (**2000**) Approximate Second Order Convergence for Spin Unrestricted Wavefunctions. *Chem. Phys. Lett.*, 325, 93–98.

A.3.41 The Split-J, Split-RI-J, RIJCOSX and RI-JK methods

1. Izsák, R.; Neese, F. (**2013**) Speeding up spin-component-scaled third-order pertubation theory with the chain of spheres approximation: the COSX-SCS-MP3 method, *Mol. Phys.*, 111, 1190.
2. Izsák, R.; Neese, F. (**2011**) An overlap fitted chain of spheres exchange method, *J. Chem. Phys.*, 135, 144105.
3. Kossmann, S.; Neese, F. (**2010**) Efficient Structure Optimization with Second-Order Many-Body Perturbation Theory: The RIJCOSX-MP2 Method, *J. Chem. Theory Comput.*, 6, 2325-2338.
4. Kossmann, S.; Neese, F. (**2009**) Comparison of Two Efficient Approximate Hartree–Fock Approaches. *Chem. Phys. Lett.*, 481, 240-243.
5. Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. (**2009**) Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange, *Chem. Phys.*, 356, 98–109.

6. Neese, F. (**2003**) An Improvement of the Resolution of the Identity Approximation for the Calculation of the Coulomb Matrix, *J. Comp. Chem.,* 24, 1740–1747.

A.3.42 sTDA and sTD-DFT approaches for electronic spectra

1. Grimme, S. (**2013**) *J. Chem. Phys.*, 138, 244104 .
2. Risthaus, T.; Hansen, A.; Grimme, S. (**2014**) *Phys. Chem. Chem. Phys.*, 16, 14408 .
3. Bannwarth, C.; Grimme, S. (**2014**) *Comput. Theor. Chem.*, 1040-1041, 45-53 .

A.3.43 XAS/XES

1. Van Kuiken, B. E.; Hahn, A. W.; Maganas, D.; DeBeer, S. (**2016**) Measuring Spin-Allowed and Spin-Forbidden d-d Excitations in Vanadium Complexes with 2p3d Resonant Inelastic X-ray Scattering, *Inorg. Chem.*, 55, 11497-11501.
2. Rees, J. A.; Wandzilak, A.; Maganas, D.; Wurster, N. I. C.; Hugenbruch, S.; Kowalska, J. K.; Pollock, C. J.; Lima, F. A.; Finkelstein, K. D.; DeBeer, S. (**2016**) Experimental and theoretical correlations between vanadium K-edge X-ray absorption and K emission spectra, *J. Biol. Inorg. Chem.*, 21, 793-805.
3. Martin-Diaconescu, V.; Chacon, K. N.; Delgado-Jaime, M. U.; Sokaras, D.; Weng, T. C.; DeBeer, S.; Blackburn, N. J. (**2016**) K beta Valence to Core X-ray Emission Studies of Cu(I) Binding Proteins with Mixed Methionine - Histidine Coordination. Relevance to the Reactivity of the M- and H-sites of Peptidylglycine Monooxygenase, *Inorg. Chem.*, 55, 3431-3439.
4. Maganas, D.; Trunschke, A.; Schlogl, R.; Neese, F. (**2016**) A unified view on heterogeneous and homogeneous catalysts through a combination of spectroscopy and quantum chemistry, *Farad. Discuss.*, 188, 181-197.
5. Kowalska, J. K.; Hahn, A. W.; Albers, A.; Schiewer, C. E.; Bjornsson, R.; Lima, F. A.; Meyer, F.; DeBeer, S. (**2016**) X-ray Absorption and Emission Spectroscopic Studies of L2Fe2S2 (n) Model Complexes: Implications for the Experimental Evaluation of Redox States in Iron-Sulfur Clusters, *Inorg. Chem.*, 55, 4485-4497.
6. Rees, J. A.; Martin-Diaconescu, V.; Kovacs, J. A.; DeBeer, S. (**2015**) X-ray Absorption and Emission Study of Dioxygen Activation by a Small-Molecule Manganese Complex, *Inorg. Chem.*, 54, 6410-6422.
7. Rees, J. A.; Bjornsson, R.; Schlesier, J.; Sippel, D.; Einsle, O.; DeBeer, S. (**2015**) The Fe-V Cofactor of Vanadium Nitrogenase Contains an Interstitial Carbon Atom, *Angew. Chem. Int. Ed.*, 54, 13249-13252.

A.3.44 X-Ray Absorption and X-Ray Emission Spectra

1. Maganas, D.; Kowalska, J. K.; Nooijen, M.; DeBeer, S.; Neese, F. (**2019**) Comparison of multireference ab initio wavefunction methodologies for X- ray absorption edges: A case study on Fe(II/III)Cl-4 (2-/1-) molecules, *J. Chem. Phys.*, 150.
2. Roemelt, M.; Beckwith, M. A.; Duboc, C.; Collomb, M.-N.; Neese, F.; DeBeer, S. (**2012**) Manganese K-Edge X-Ray Absorption Spectroscopy as a Probe of the Metal-Ligand Interactions in Coordination Compounds, *Inorg. Chem.*, 51, 680–687.
3. Chandrasekaran, P.; Stieber, S. C. E.; Collins, T. J.; Que, L.; Neese, F.; DeBeer, S. (**2011**) Prediction of high-valent iron K-edge absorption spectra by time-dependent Density Functional Theory, *Dalton Trans.*, 40, 11070–11079.
4. Beckwith, M. A.; Roemelt, M.; Collomb, M. N.; Duboc, C.; Weng, T. C.; Bergmann, U.; Glatzel, P.; Neese, F.; DeBeer, S. (**2011**) Manganese K beta X-ray Emission Spectroscopy As a Probe of Metal-Ligand Interactions, *Inorg. Chem,* 50, 8397–8409.
5. Lee, N.; Petrenko, T.; Bergmann, U.; Neese, F.; DeBeer, S. (**2010**) Probing Valence Orbital Composition with Iron K beta X-ray Emission Spectroscopy, *J. Am. Chem. Soc.*, 132, 9715–9727.

6. DeBeer-George, S.; Neese, F. (**2010**) Calibration of Scalar Relativistic Density Functional Theory for the Calculation of Sulfur K-Edge X-ray Absorption Spectra, *Inorg. Chem,* 49, 1849–1853.

7. DeBeer-George, S.; Petrenko, T.; Neese, F. (**2008**) Prediction of Iron- K-edge Absorption Spectra using Time-Dependent Density Functional Theory, *J. Phys. Chem. A.*, 112, 12936–12943.

8. DeBeer-George, S.; Petrenko, T.; Neese, F. (**2008**) Time-dependent density functional calculations of ligand K-edge X-ray absorption spectra, *Inorg. Chim. Acta* (60^{th} birthday issue of Prof. E.I. Solomon), 361, 965–972.

A.4 Applications that make use of include the following

1. Kruse, H.; Mladek, A.; Gkionis, K.; Hansen, A.; Grimme, S.; Sponer J. (**2015**) Quantum Chemical Benchmark Study on 46 RNA Backbone Families Using a Dinucleotide Unit, *J. Chem. Theory Comput.*, 11, 4972.

2. Qu, Z.-W.; Hansen, A.; Grimme, S. (**2015**) Co-C Bond Dissociation Energies in Cobalamin Derivatives and Dispersion Effects: Anomaly or Just Challenging?, *J. Chem. Theory Comput.*, 11, 1037.

3. A. Hansen, A.; Bannwarth, C.; Grimme, S.; Petrović, P.; Werlé, C.; Djukic, J.-P. (**2014**) The Thermochemistry of London Dispersion-Driven Transition Metal Reactions: Getting the 'Right Answer for the Right Reason', *ChemistryOpen*, 3, 177.

4. Krewald, V.; Neese, F.; Pantazis, D. A. (**2013**) On the magnetic and spectroscopic properties of high-valent Mn_3CaO_4 cubanes as structural units of natural and artificial water oxidizing catalysts, *J. Am. Chem. Soc.*, 135, 5726–5739.

5. Kampa, M.; Pandelia, M.-E.; Lubitz, W.; van Gastel, M.; Neese, F. (**2013**) A Metal-Metal Bond in the Light-Induced State of [NiFe] Hydrogenases with Relevance to Hydrogen Evolution, *J. Am. Chem. Soc.*, 135, 3915–3925.

6. Pandelia, M.-E.; Bykov, D.; Izsák, R.; Infossi, P.; Giudici-Orticoni, M.-T.; Bill, E.; Neese, F.; Lubitz, W. (**2013**) Electronic structure of the unique [4Fe-3S] cluster in O_2-tolerant hydrogenases characterized by Fe-57 Mossbauer and EPR spectroscopy, *Proc. Natl. Acad. Sci. USA*, 110, 483–488.

7. Atanasov, M.; Surawatanawong, P.; Wieghardt, K.; Neese, F. (**2013**) A theoretical study of zero-field splitting in Fe(IV)S_6 (S = 1) and Fe(III)S_6 (S = 1/2) core complexes, $[Fe^{IV}(Et_2dtc)_{3-n}(mnt)_n]^{(n-1)-}$ and $[Fe^{III}(Et_2dtc)_{3-n}(mnt)_n]^{n-}$ (n = 0, 1, 2, 3): The origin of the magnetic anisotropy, *Coord. Chem. Rev.*, 257(1), 27–41.

8. Retegan, M.; Collomb, M.-N.; Neese, F.; Duboc, C. (**2013**) A combined high-field EPR and quantum chemical study on a weakly ferromagnetically coupled dinuclear Mn(III) complex. A complete analysis of the EPR spectrum beyond the strong coupling limit, *Phys. Chem. Chem. Phys.*, 15, 223–234.

9. Zadrozny, J. M.; Atanasov, M.; Bryan, A. M.; Lin, C. Y.; Rekken, B. D.; Power, P. P.; Neese, F.; Long, J. R. (**2013**) Slow magnetization dynamics in a series of two-coordinate iron(II) complexes, *Chem. Sci.*, 4, 125–138.

10. Weber, K.; Krämer, T.; Shafaat, H. S.; Weyhermüller, T.; Bill, E.; van Gastel, M.; Neese, F.; Lubitz, W. (**2012**) A Functional [NiFe]-Hydrogenase Model Compound That Undergoes Biologically Relevant Reversible Thiolate Protonation, *J. Am. Chem. Soc.*, 134, 20745–20755.

11. Kampa, M.; Lubitz, W.; van Gastel, M.; Neese, F.; (**2012**) Computational study of the electronic structure and magnetic properties of the Ni-C state in [NiFe] hydrogenases including the second coordination sphere, *J. Biol. Inorg. Chem.*, 17, 1269–1281.

12. Atanasov, M.; Comba, P.; Helmle, S.; Müller, D.; Neese, F. (**2012**) Zero-Field Splitting in a Series of Structurally Related Mononuclear Ni^{II}-Bispidine Complexes, *Inorg. Chem.*, 51, 12324–12335.

13. Shafaat, H. S.; Weber, K.; Petrenko, T.; Neese, F.; Lubitz, W. (**2012**) Key Hydride Vibrational Modes in [NiFe] Hydrogenase Model Compounds Studied by Resonance Raman Spectroscopy and Density Functional Calculations, *Inorg. Chem.*, 51, 11787–11797.

14. Argirevic, T.; Riplinger, C.; Stubbe, J.; Neese, F.; Bennati, M. (**2012**) ENDOR Spectroscopy and DFT Calculations: Evidence for the Hydrogen-Bond Network Within $\alpha 2$ in the PCET of *E. coli* Ribonucleotide Reductase, *J. Am. Chem. Soc.*, 134, 17661–17670.

15. Albrecht, C.; Shi, L. L.; Perez, J. M.; van Gastel, M.; Schwieger, S.; Neese, F.; Streubel, R. (**2012**) Deoxygenation of Coordinated Oxaphosphiranes: A New Route to P=C Double-Bond Systems, *Chem. Eur. J.*, 18, 9780–9783.

16. Maganas, D.; Krzystek, J.; Ferentinos, E.; Whyte, A. M.; Robertson, N.; Psycharis, V.; Terzis, A.; Neese, F.; Kyritsis, P. (**2012**) Investigating Magnetostructural Correlations in the Pseudooctahedral trans-[Ni^{II}{($OPPh_2$) ($EPPh_2$)N}$_2$(sol)$_2$] Complexes (E = S, Se; sol = DMF, THF) by Magnetometry, HFEPR, and ab Initio Quantum Chemistry, *Inorg. Chem.*, 51, 7218–7231.

17. Ye, S. F.; Neese, F. (**2012**) How Do Heavier Halide Ligands Affect the Signs and Magnitudes of the Zero-Field Splittings in Halogenonickel(II) Scorpionate Complexes? A Theoretical Investigation Coupled to Ligand-Field Analysis, *J. Chem. Theory Comput.*, 8, 2344–2351.

18. Nesterov, V.; Ozbolat-Schon, A.; Schnakenburg, G.; Shi, L. L.; Cangonul, A.; van Gastel, M.; Neese, F.; Streubel, R. (**2012**) An Unusal Case of Facile Non-Degenerate P-C Bond Making and Breaking, *Chem. Asian J.*, 7, 1708–1712.

19. Bykov, D.; Neese, F. (**2012**) Reductive activation of the heme iron-nitrosyl intermediate in the reaction mechanism of cytochrome c nitrite reductase: a theoretical study, *J. Biol. Inorg. Chem.*, 17, 741–760.

20. Lancaster, K. M.; Zaballa, M.E.; Sproules, S.; Sundararajan, M.; DeBeer, S.; Richards, J. H.; Vila, A. J.; Neese, F.; Gray, H. B. (**2012**) Outer-Sphere Contributions to the Electronic Structure of Type Zero Copper Proteins, *J. Am. Chem. Soc.*, 134, 8241–8253.

21. Benkhauser-Schunk, C.; Wezisla, B.; Urbahn, K.; Kiehne, U.; Daniels, J.; Schnakenburg, G.; Neese, F.; Lutzen, A. (**2012**) Synthesis, Chiral Resolution, and Absolute Configuration of Functionalized Troger's Base Derivatives: Part II, *ChemPlusChem*, 77, 396–403.

22. Ye, S. F.; Riplinger, C.; Hansen, A.; Krebs, C.; Bollinger, J. M.; Neese, F. (**2012**) Electronic Structure Analysis of the Oxygen-Activation Mechanism by Fe^{II}- and α-Ketoglutarate (αKG)-Dependent Dioxygenases, *Chem. Eur. J.*, 18, 6555–6567.

23. Desrochers, P. J.; Sutton, C. A.; Abrams, M. L.; Ye, S. F.; Neese, F.; Telser, J.; Ozarowski, A.; Krzystek, J. (**2012**) Electronic Structure of Nickel(II) and Zinc(II) Borohydrides from Spectroscopic Measurements and Computational Modeling, *Inorg. Chem.*, 51, 2793–2805.

24. Torres-Alacan, J.; Krahe, O.; Filippou, A. C.; Neese, F.; Schwarzer, D.; Vöhringer, P. (**2012**) The Photochemistry of [$Fe^{III}N_3$(cyclam-ac)]PF_6 at 266 nm, *Chem. Eur. J.*, 18, 3043–3055.

25. Maekawa, M.; Römelt, M.; Daniliuc, C. G.; Jones, P. G.; White, P. S.; Neese, F.; Walter, M. D. (**2012**) Reactivity studies on [Cp′MnX(thf)]$_2$: manganese amide and polyhydride synthesis *Chem. Sci.*, 3, 2972–2979.

26. Pantazis, D. A.; Ames, W.; Cox, N.; Lubitz, W.; Neese, F. (**2012**) Two interconvertible structures that explain the spectroscopic properties of the oxygen-evolving complex of photosystem II in the S_2 state, *Angew. Chem. Int. Ed.*, 51, 9935–9940. (selected as cover article and VIP paper)

27. Vennekate, H.; Schwarzer, D.; Torres-Alacan, J.; Krahe, O.; Filippou, A. C.; Neese, F.; Vöhringer, P. (**2012**) Ultrafast primary processes of an iron-(III) azido complex in solution induced with 266 nm light, *Phys. Chem. Chem. Phys.*, 14, 6165–6172.

28. Christian, G. J.; Ye, S.; Neese, F. (**2012**) Oxygen activation in extradiol catecholate dioxygenases – a density functional study, *Chem. Sci.*, 3, 1600–1611.

29. Cowley, R. E.; Christian, G. J.; Brennessel, W. W.; Neese, F.; Holland, P. L. (**2012**) A Reduced (beta-Diketiminato)iron Complex with End-On and Side-On Nitriles: Strong Backbonding or Ligand Non-Innocence? *Eur. J. Inorg. Chem.*, 479–483.

30. Thiessen, A.; Wettach, H.; Meerholz, K.; Neese, F.; Hoger, S.; Hertel, D. (**2012**) Control of electronic properties of triphenylene by substitution, *Organic Electronics*, 13, 71–83.

31. Lancaster, K. M.; Roemelt, M.; Ettenhuber, P.; Hu, Y. L.; Ribbe, M. W.; Neese, F.; Bergmann, U.; DeBeer, S. (**2011**) X-ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor, *Science*, 334, 974–977.

32. Ames, W.; Pantazis, D. A.; Krewald, V.; Cox, N.; Messinger, J.; Lubitz, W.; Neese, F. (**2011**) Theoretical Evaluation of Structural Models of the S_2 State in the Oxygen Evolving Complex of Photosystem II: Protonation States and Magnetic Interactions, *J. Am. Chem. Soc.*, 133, 19743–19757.

33. Antony, J.; Grimme, S.; Liakos, D. G.; Neese, F. (**2011**) Protein-Ligand Interaction Energies with Dispersion Corrected Density Functional Theory and High-Level Wave Function Based Methods, *J. Phys. Chem. A*, 115, 11210–11220.

34. Radoul, M.; Bykov, D.; Rinaldo, S.; Cutruzzola, F.; Neese, F.; Goldfarb, D. (**2011**) Dynamic Hydrogen-Bonding Network in the Distal Pocket of the Nitrosyl Complex of Pseudomonas aeruginosa cd(1) Nitrite Reductase, *J. Am. Chem. Soc.*, 133, 3043–3055.

35. Liakos, D. G.; Neese, F. (**2011**) Interplay of Correlation and Relativistic Effects in Correlated Calculations on Transition-Metal Complexes: The $Cu_2O_2^{2+}$ Core Revisited, *J. Chem. Theory Comput.*, 7, 1511–1523.

36. Riplinger, C.; Neese, F. (**2011**) The Reaction Mechanism of Cytochrome P450 NO Reductase: A Detailed Quantum Mechanics/Molecular Mechanics Study, *ChemPhysChem*, 12, 3192–3203.

37. Maganas, D.; Sottini, S.; Kyritsis, P.; Groenen, E. J. J.; Neese, F. (**2011**) Theoretical Analysis of the Spin Hamiltonian Parameters in $Co^{II}S_4$ Complexes, Using Density Functional Theory and Correlated ab initio Methods, *Inorg. Chem,* 50, 8741–8754.

38. Surawatanawong, P.; Sproules, S.; Neese, F.; Wieghardt, K. (**2011**) Electronic Structures and Spectroscopy of the Electron Transfer Series Fe(NO)L_2^z ($z = 1+, 0, 1-, 2-, 3-$; L = Dithiolene), *Inorg. Chem,* 50, 12064–12074.

39. Cox, N.; Ames, W.; Epel, B.; Kulik, L. V.; Rapatskiy, L.; Neese, F.; Messinger, J.; Wieghardt, K.; Lubitz, W. (**2011**) Electronic Structure of a Weakly Antiferromagnetically Coupled Mn(II)Mn(III) Model Relevant to Manganese Proteins: A Combined EPR, ^{55}Mn-ENDOR, and DFT Study, *Inorg. Chem,* 50, 8238–8251.

40. Rota, J. B.; Knecht, S.; Fleig, T.; Ganyushin, D.; Saue, T.; Neese, F.; Bolvin, H. (**2011**) Zero field splitting of the chalcogen diatomics using relativistic correlated wave-function methods, *J. Chem. Phys.*, 135, 114106.

41. Atanasov, M.; Ganyushin, D.; Pantazis, D. A.; Sivalingam, K.; Neese, F. (**2011**) Detailed Ab Initio First-Principles Study of the Magnetic Anisotropy in a Family of Trigonal Pyramidal Iron(II) Pyrrolide Complexes, *Inorg. Chem,* 50, 7460–7477.

42. Cox, N.; Rapatskiy, L.; Su, J. H.; Pantazis, D. A.; Sugiura, M.; Kulik, L.; Dorlet, P.; Rutherford, A. W.; Neese, F.; Boussac, A.; Lubitz, W.; Messinger, J. (**2011**) Effect of Ca^{2+}/Sr^{2+} Substitution on the Electronic Structure of the Oxygen-Evolving Complex of Photosystem II: A Combined Multifrequency EPR, ^{55}Mn-ENDOR, and DFT Study of the S_2 State, *J. Am. Chem. Soc.*, 133, 3635–3648.

43. Maurice, R.; Sivalingam, K.; Ganyushin, D.; Guihery, N.; de Graaf, C.; Neese, F. (**2011**) Theoretical Determination of the Zero-Field Splitting in Copper Acetate Monohydrate, *Inorg. Chem,* 50, 6229–6236.

44. Su, J. H.; Cox, N.; Ames, W.; Pantazis, D. A.; Rapatskiy, L.; Lohmiller, T.; Kulik, L. V.; Dorlet, P.; Rutherford, A. W.; Neese, F.; Boussac, A.; Lubitz, W.; Messinger, J. (**2011**) The electronic structures of the S_2 states of the oxygen-evolving complexes of photosystem II in plants and cyanobacteria in the presence and absence of methanol, *Biochim. Biophys. Acta-Bioenergetics*, 1807, 829–840.

45. Bykov, D.; Neese, F. (**2011**) Substrate binding and activation in the active site of cytochrome c nitrite reductase: a density functional study, *J. Biol. Inorg. Chem,* 16, 417–430.

46. Gennari, M.; Orio, M.; Pecaut, J.; Bothe, E.; Neese, F.; Collomb, M. N.; Duboc, C. (**2011**) Influence of Mixed Thiolate/Thioether versus Dithiolate Coordination on the Accessibility of the Uncommon +I and +III Oxidation States for the Nickel Ion: An Experimental and Computational Study, *Inorg. Chem,* 50, 3707–3716.

47. Ye, S. F.; Neese, F. (**2011**) Nonheme oxo-iron(IV) intermediates form an oxyl radical upon approaching the C-H bond activation transition state, *Proc. Natl. Acad. Sci. USA*, 108, 1228–1233.

48. Gennari, M.; Pecaut, J.; DeBeer, S.; Neese, F.; Collomb, M. N.; Duboc, C. (**2011**) A Fully Delocalized Mixed-Valence Bis-μ-(Thiolato) Dicopper Complex: A Structural and Functional Model of the Biological Cu(A) Center, *Angew. Chem., Int. Ed.*, 50, 5661–5665.

49. Gennari, M.; Retegan, M.; DeBeer, S.; Pecaut, J.; Neese, F.; Collomb, M. N.; Duboc, C. (**2011**) Experimental and Computational Investigation of Thiolate Alkylation in Ni(II) and Zn(II) Complexes: Role of the Metal on the Sulfur Nucleophilicity, *Inorg. Chem,* 50, 10047–10055.

50. Atanasov, M.; Delley, B.; Neese, F.; Tregenna-Piggott, P. L.; Sigrist, M. (**2011**) Theoretical Insights into the Magnetostructural Correlations in Mn(3)-Based Single-Molecule Magnets, *Inorg. Chem,* 50, 2112–2124.

51. Neese, F.; Pantazis, D. A. (**2011**) What is not required to make a single molecule magnet, *Faraday Discussions*, 148, 229–238.

52. Lassalle-Kaiser, B.; Hureau, C.; Pantazis, D. A.; Pushkar, Y.; Guillot, R.; Yachandra, V. K.; Yano, J.; Neese, F.; Anxolabéhère-Mallart, E. (**2010**) Activation of a water molecule using a mononuclear Mn complex: from Mn-aquo, to Mn-hydroxo, to Mn-oxyl via charge compensation, *Energy Environ. Sci.*, 3, 924–938.

53. Pantazis, D. A.; Krewald, V.; Orio, M.; Neese, F. (**2010**) Theoretical magnetochemistry of dinuclear manganese complexes: broken symmetry density functional theory investigation on the influence of bridging motifs on structure and magnetism, *Dalton Trans.*, 39, 4959–4967.

54. Woertink, J. S.; Tian, L.; Maiti, D.; Lucas, H. R.; Himes, R. A.; Karlin, K. D.; Neese, F.; Wurtele, C.; Holthausen, M. C.; Bill, E.; Sundermeyer, J.; Schindler, S.; Solomon, E. I. (**2010**) Spectroscopic and Computational Studies of an End-on Bound Superoxo-Cu(II) Complex: Geometric and Electronic Factors That Determine the Ground State, *Inorg. Chem,* 49, 9450–9459.

55. McNaughton, R. L.; Roemelt, M.; Chin, J. M.; Schrock, R. R.; Neese, F.; Hoffman, B. M. (**2010**) Experimental and Theoretical EPR Study of Jahn–Teller-Active HIPTN(3)N MoL Complexes (L = N_2, CO, NH_3), *J. Am. Chem. Soc.*, 132, 8645–8656.

56. Geng, C. Y.; Ye, S. F.; Neese, F. (**2010**) Analysis of Reaction Channels for Alkane Hydroxylation by Nonheme Iron(IV)-Oxo Complexes, *Angew. Chem., Int. Ed.*, 49, 5717–5720.

57. Orio, M.; Jarjayes, O.; Kanso, H.; Philouze, C.; Neese, F.; Thomas, F. (**2010**) X-Ray Structures of Copper(II) and Nickel(II) Radical Salen Complexes: The Preference of Galactose Oxidase for Copper(II), *Angew. Chem., Int. Ed.*, 49, 4989–4992.

58. Gennari, M.; Orio, M.; Pecaut, J.; Neese, F.; Collomb, M. N.; Duboc, C. (**2010**) Reversible Apical Coordination of Imidazole between the Ni(III) and Ni(II) Oxidation States of a Dithiolate Complex: A Process Related to the Ni Superoxide Dismutase, *Inorg. Chem,* 49, 6399–6401.

59. Ye, S. F.; Price, J. C.; Barr, E. W.; Green, M. T.; Bollinger, J. M.; Krebs, C.; Neese, F. (**2010**) Cryoreduction of the NO-Adduct of Taurine:alpha-Ketoglutarate Dioxygenase (TauD) Yields an Elusive FeNO Species, *J. Am. Chem. Soc.*, 132, 4739–4751.

60. Maganas, D.; Grigoropoulos, A.; Staniland, S. S.; Chatziefthimiou, S. D.; Harrison, A.; Robertson, N.; Kyritsis, P.; Neese, F. (**2010**) Tetrahedral and Square Planar Ni(SPR_2)$_2N_2$ complexes, R = Ph & iPr Revisited: Experimental and Theoretical Analysis of Interconversion Pathways, Structural Preferences, and Spin Delocalization, *Inorg. Chem,* 49, 5079–5093.

61. Anoop, A.; Thiel, W.; Neese, F. (**2010**) A Local Pair Natural Orbital Coupled Cluster Study of Rh Catalyzed Asymmetric Olefin Hydrogenation, *J. Chem. Theory Comput.*, 6, 3137–3144.

62. Duboc, C.; Collomb, M. N.; Neese, F. (**2010**) Understanding the Zero-Field Splitting of Mononuclear Manganese(II) Complexes from Combined EPR Spectroscopy and Quantum Chemistry, *Appl. Magn. Res.*, 37, 229–245.

63. Ye, S. F.; Neese, F. (**2010**) The Unusual Electronic Structure of Dinitrosyl Iron Complexes, *J. Am. Chem. Soc.*, 132, 3646–3647.

64. Kochem, A.; Orio, M.; Jarjayes, O.; Neese, F.; Thomas, F. (**2010**) Unsymmetrical one-electron oxidized Ni(II)-bis(salicylidene) complexes: a protonation-induced shift of the oxidation site, *Chem. Commun.*, 46, 6765–6767.

65. Ozbolat-Schon, A.; Bode, M.; Schnakenburg, G.; Anoop, A.; van Gastel, M.; Neese, F.; Streubel, R. (**2010**) Insights into the Chemistry of Transient P-Chlorophosphanyl Complexes, *Angew. Chem., Int. Ed.*, 49, 6894–6898.

66. Vancoillie, S.; Chalupsky, J.; Ryde, U.; Solomon, E. I.; Pierloot, K.; Neese, F.; Rulisek, L. (**2010**) Multireference Ab Initio Calculations of g tensors for Trinuclear Copper Clusters in Multicopper Oxidases, *J. Phys. Chem. B*, 114, 7692–7702.

67. Grote, D.; Finke, C.; Kossmann, S.; Neese, F.; Sander, W. (**2010**) 3,4,5,6-Tetrafluorophenylnitren-2-yl: A Ground-State Quartet Triradical, *Chem. Eur. J.*, 16, 4496–4506.

68. Ye, S. F.; Neese, F.; Ozarowski, A.; Smirnov, D.; Krzystek, J.; Telser, J.; Liao, J. H.; Hung, C. H.; Chu, W. C.; Tsai, Y. F.; Wang, R. C.; Chen, K. Y.; Hsu, H. F. (**2010**) Family of V(III)-Tristhiolato Complexes Relevant to Functional Models of Vanadium Nitrogenase: Synthesis and Electronic Structure Investigations by Means of High-Frequency and -Field Electron Paramagnetic Resonance Coupled to Quantum Chemical Computations, *Inorg. Chem,* 49, 977–988.

69. Hegele, P.; Santhamma, B.; Schnakenburg, G.; Frohlich, R.; Kataeva, O.; Nieger, M.; Kotsis, K.; Neese, F.; Dotz, K. H. (**2010**) Hydroquinoid Chromium Complexes Bearing an Acyclic Conjugated Bridge: Chromium-Templated Synthesis, Molecular Structure, and Haptotropic Metal Migration, *Organometallics*, 29, 6172–6185.

70. Ye, S. F.; Neese, F. (**2010**) Accurate Modeling of Spin-State Energetics in Spin-Crossover Systems with Modern Density Functional Theory, *Inorg. Chem,* 49, 772–774.

71. Orio, M.; Philouze, C.; Jarjayes, O.; Neese, F.; Thomas, F. (**2010**) Spin Interaction in Octahedral Zinc Complexes of Mono- and Diradical Schiff and Mannich Bases, *Inorg. Chem,* 49, 646–658.

72. Pantazis, D. A.; Orio, M.; Petrenko, T.; Zein, S.; Lubitz, W.; Messinger, J.; Neese, F. (**2009**) Structure of the Oxygen-Evolving Complex of Photosystem II: Information on the S_2 state through Quantum Chemical Calculation of its Magnetic Properties. *Phys. Chem. Chem. Phys.*, 11, 6788–6798.

73. Baffert, C.; Orio, M.; Pantazis, D. A.; Duboc, C.; Blackman, A.G.; Blondin, G.; Neese, F.; Deronzier,A.; Collomb, M-N. (**2009**) A trinuclear terpyridine frustrated spin system with a $Mn_3^{IV}O_4$ core: synthesis, physical characterization and quantum chemical modeling of its magnetic properties. *Inorg. Chem.*, 48, 10281–10288.

74. Liakos, D.; Neese, F. (**2009**) A multiconfigurational *ab initio* study of the zero-field splitting in the di- and trivalent hexaquo-chromium complexes. *Inorg. Chem.*, 48, 10572–10580.

75. Astashkin, A.V.; Klein, E.C.; Ganyushin, D.; Johnson.Winters, K.; Neese, F.; Kappler, U.; Enemark, J.H. (**2009**) Exchangeable oxygens in the vicinity of the molybdenum center of the high-pH form of sulfite oxidase and sulfite dehydrogenase. *Phys. Chem. Chem. Phys.*, 11, 6733–6742.

76. Orio, M.; Pantazis, D. A.; Petrenko, T.; Neese, F. (**2009**) Magnetic and spectroscopic properties of mixed valence manganese(III,IV) dimers: a systematic study using broken symmetry density functional theory, *Inorg. Chem.*, 48, 7251–7260.

77. Klein, E.L.; Astashkin, A.V.; Ganyushin, D.; Johnson-Winters, K.; Wilson, H.L.; Rajagopalan, K. V.; Neese, F.; Enemark, J.H. (**2009**) Direct Detection and Characterization of Chloride in the Active Site of the Low-pH Form of Sulfite Oxidase Using ESEEM Spectroscopy, Isotopic Labeling, and DFT Calculations, *Inorg. Chem.*, 48(11), 4743–4752.

78. Vancoillie, S.; Rulisek, L.; Neese, F.; Pierloot, K. (**2009**) Theoretical description of the structure and magnetic properties of nitroxide-Cu(II)-nitroxide spin triads, *J. Phys. Chem.*, 113, 6149–6157.

79. Cowley, R.E.; Bill, E.; Neese, F.; Brennessel,W.W.; Holland, P.L. (**2009**) Iron(II) Complexes With Redox-Active Tetrazene (RNNNNR) Ligands, *Inorg. Chem.*, 48, 4828–4836.

80. Gansäuer, A.; Fleckhaus, A.; Lafon, A.; Okkel, M.; Anakuthil, A.; Kotsis, K.; Neese, F. (**2009**) Catalysis via Homolytic Substitutions with C-O and Ti-O Bonds: Oxidative Additions and Reductive Eliminations in Single Electron Steps. *J. Am. Chem. Soc.*, 131, 16989–16999.

81. Ye, S.; Neese, F. (**2009**) Quantum Chemical Studies of C-H Activation Reactions by High-Valent Nonheme Iron Centers *Curr. Op. Chem. Biol.*, 13(1), 89–98.

82. Krahe, O.; Neese, F.; Streubel, R. (**2009**) The quest for ring-opening of oxaphosphirane complexes: a coupled cluster and density functional study of CH_3PO isomers and their $Cr(CO)_5$ complexes *Chem. Eur. J.*, 15, 2594–2601.

83. Romain, S.; Duboc, C.; Neese, F.; Riviere, E.; Hanton, L. R.; Blackman, A. G.; Philouze, C.; Lepretre, J. C.; Deronzier, A.; Collomb, M. N. (**2009**) An Unusual Stable Mononuclear Mn(III) Bis-terpyridine Complex Exhibiting Jahn-Teller Compression: Electrochemical Synthesis, Physical Characterisation and Theoretical Study, *Chem. Eur. J.*, 15, 980–988

84. Zein, S.; Neese, F. (**2008**) *Ab initio* and Coupled Perturbed DFT Calculation of Zero-Field Splittings in Mn(II) Transition Metal complexes. *J. Phys. Chem. A*, 112, 7976–7983.

85. Ye, S.; Tuttle, T.; Bill, E.; Gross, Z.; Thiel, W.; Neese, F. (**2008**) The Noninnocence of Iron Corroles: A combined Experimental and Quantum Chemical Study. *Chem. Eur. J.* (selected as very important paper), 34, 10839–10851.

86. Duboc, C.; Collomb, M.-N.; Pecaut, J.; Deronzier, A.; Neese, F. (**2008**) Definition of Magneto-Structural Correlations for the Mn(II) Ion. *Chem. Eur. J.*, 21, 6498–6509.

87. Berry, J.F.; DeBeer-George, S.; Neese, F. (**2008**) Electronic Structure and Spectroscopy of "Superoxidized" Iron Centers in Model Systems: Theoretical and Experimental Trends. *Phys. Chem. Chem. Phys.*, 10, 4361–4374.

88. Sander, W.; Grote, D.; Kossmann, S.; Neese, F. (**2008**) 2.3.5.6-Tetrafluorophenylnitren-4-yl: EPR Spectroscopic Characterization of a Quartet Ground State Nitreno Radical, *J. Am. Chem. Soc.*, 130, 4396–4403.

89. Scheifele, Q.; Riplinger, C.; Neese, F.; Weihe, H.; Barra, A.L.; Jurany, F.; Podlesnyak, A.; Tregenna-Piggot, P.W.L. (**2008**) Spectroscopic and Theoretical Study of a Mononuclear Mn(III) Bioinorganic Complex Exhibiting a Compressed Jahn-Teller Octahedron, *Inorg. Chem.*, 47, 439–447.

90. Zein, S.; Kulik, L.V.; Yano, J.; Kern, J.; Zouni, A.; Yachandra, V.K.; Lubitz, W.; Neese, F.; Messinger, J. (**2008**) Focussing the View on Nature's Water Splitting Catalyst *Phil. Trans. Roy. Soc. London B*, 363, 1167–1177.

91. Zein, S.; Duboc, C.; Lubitz, W.; Neese, F. (**2008**) Theoretical Characterization of zero-Field Splittings in Mn(II) Complexes. *Inorg. Chem.*, 47, 134–142.

92. Parker, D.J.; Hammond, D.; Davies, E.S.; Garner, C.D.; Benisvy, L.; McMaster, J.; Wilson, C.; Neese, F.; Bothe, E.; Bittl, R.; Teutloff, C. (**2007**) A stable H-bonded *ortho*-Thioether Phenoxyl-Radical: A Chemical and Spectroscopic Analogue of $^{\bullet}\mathrm{Tyr}_{272}$ in *apo*-Galactose Oxidase, *J. Biol. Inorg. Chem.* (Ed Stiefel memorial issue), 101, 1859–1864.

93. Chlopek, K.; Muresan, N.; Neese, F.; Wieghardt, K. (**2007**) Electronic Structures of Five-Coordinate Complexes of Iron Containing Zero, One, or Two π Radical Ligands: A Broken Symmetry Density Functional Theoretical Study, *Chem. Eur. J.*, 13, 8391–8403.

94. Muresan, N.; Chlopek, K.; Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2007**) Bis(α-diimine)nickel Complexes: Molecular and Electronic Structure of Three Members of the Electron-Transfer Series $[\mathrm{Ni(L)}_2]^z$ ($z = 0, 1+, 2+$) (L = 2-Phenyl-1,4-bis(isopropyl)-1,4-diazabutadiene). A Combined Experimental and Theoretical Study, *Inorg. Chem.*, 46, 4905–4916.

95. Ray, K.; Petrenko, T.; Wieghardt, K.; Neese, F. (**2007**) Joint Spectroscopic and Theoretical Investigations of Transition Metal Complexes Involving Non-Innocent Ligands. *Dalton Trans.*, 1552 (selected for cover picture).

96. Sinnecker, S.; Svensen, N.; Barr, E.W.; Ye, S.; Bollinger, J.M.; Neese, F.; Krebs, C. (**2007**) Spectroscopic and Theoretical Evaluation of the Structure of the High-Spin Fe(IV)-Oxo Intermediates in Taurine:α-Ketoglutarate Dioxygenase from *Escherichia coli* and its His99Ala Ligand Variant *J. Am. Chem. Soc.*, 129, 6168–6179.

97. Duboc, C.; Phoeung, T; Zein, S.; Pécaut, J.; Collomb, M.-N.; Neese. F. (**2007**) Origin of the zero field splitting in mononuclear dihalide Mn(II) complexes: an investigation by multifrequency high-field EPR and density functional theory (DFT), *Inorg. Chem.*, 46, 4905–4916.

98. DeBeer-George, S.; Petrenko, T.; Aliaga-Alcade, N.; Bill, E.; Mienert, B.; Sturhan, W.; Ming, Y.; Wieghardt, K.; Neese, F. (**2007**) Characterization of a Genuine Iron(V)Nitrido Species by Nuclear Resonant Vibrational Spectroscopy Coupled to Density Functional Calculations, *J. Am. Chem. Soc.*, 129, 11053–11060.

99. Lehnert, N.M; Cornelissen, U.; Neese, F.; Ono, T.; Noguchi, Y.; Okamoto, K.-I.; Fujisawa, K. (**2007**) Synthesis and Spectroscopic Characterization of Cu(II)-Nitrite Complexes with Hydrotris(pyrazolyl)borate and Related Ligands. *Inorg. Chem.*, 46, 3916–3933.

100. Carmieli, R.; Larsen, T.; Reed, G.H.; Zein, S.; Neese, F.; Goldfarb, D. (**2007**) The Catalytic Mn^{2+} Sites in the Enolase-Inhibitor Complex - Crystallography, Single Crystal EPR and DFT calculations. *J. Am. Chem. Soc.*, 129, 4240–4252.

101. Kokatam, S.; Ray, K.; Pap, J.; Bill, E.; Geiger, W.E.; LeSuer, R.J.; Rieger, P.H.; Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2007**) Molecular and Electronic Structure of Square Planar Gold Complexes Containing Two 1,2-di(4-*tert*-butylphenyl)ethylene-1,2-dithiolato Ligands: $[Au(L)_2]^{1+/0/1-/2-}$. A Combined Experimental and Computational Study, *Inorg. Chem.*, 46, 1100–1111.

102. Ray, K.; DeBeer-George, S.; Solomon, E.I.; Wieghardt, K.; Neese, F. (**2007**) Description of the Ground State Covalencies of the Bis(dithiolato)Transition Metal Complexes Using X-ray Absorption Spectral and Time-Dependent-Density-Functional Studies. *Chem. Eur. Journal*, 13(10), 2753 (selected for cover picture).

103. Chalupský, J.; Neese, F.; Solomon, E.I.; Ryde, U.; Rulíšek, L. (**2006**) Identification of intermediates in the reaction cycle of multicopper oxidases by quantum chemical calculations of spectroscopic parameters, *Inorg. Chem.*, 45, 11051–11059.

104. Bart, S.C.; Chłopek, K.; Bill, E., Bouwkamp, B.W.; Lobkovsky, E.; Neese, F.; Wieghardt, K.; Chirik, P.J. (**2006**) Electronic Structure of Bis(imino)pyridine Iron Dichloride, Monochloride and Neutral Ligand Complexes: A Combined Structural, Spectroscopic and Computational Study, *J. Am. Chem. Soc.*, 128, 13901–13912.

105. Patra, A.K.; Bill, E.; Bothe, E.; Chlopek, K.; Neese, F.; Weyhermüller, T.; Stobie, K.; Ward, M.D.; McCleverty, J.A.; Wieghardt, K. (**2006**) The Electronic Structure of Mononuclear Bis(1,2-diaryl-1,2-ethylenedithiolate)iron Complexes Containing a Fifth Cyanide or Phosphite Ligand: A Combined Experimental and Computational Study, *Inorg. Chem.*, 45, 7877–7890.

106. Berry, J.F.; Bill, E.; Bothe, E.; DeBeer-George, S.; Mienert, B.; Neese, F.; Wieghardt, K. (**2006**) An Octahedral Coordination Complex of Iron(VI) – One Step Ahead of Nature?, *Science*, 312, 1937–1941.

107. Petrenko, T.; Ray, K.; Wieghardt, K.; Neese, F. (**2006**) Vibrational Markers for the Open-Shell Character of Metal bis-Dithiolenes: An Infrared, resonance Raman and Quantum Chemical Study. *J. Am. Chem. Soc.*, 128, 4422–4436.

108. Chłopek, K.; Bothe, E.; Neese, F.; Weyhermüller, T.; Wieghardt, K. (**2006**) The Molecular and Electronic Structures of Tetrahedral Complexes of Nickel and Cobalt Containing N, N'-Disubstituted, Bulky *o*-Diiminobenzosemiquinonate(1-) π-Radical Ligands, *Inorg. Chem.*, 45, 6298–6307.

109. Kababya, S.; Nelson. J.; Calle, C.; Neese, F.; Goldfarb, D. (**2006**) The electronic structure of bi-nuclear mixed valent copper azacryptates derived from integrated advanced EPR and DFT calculations. *J. Am. Chem. Soc.*, 128, 2017–2029.

110. Berry, J.F.; Bill, E.; Neese, F.; Garcia-Serres, R.; Weyhermüller, T.; Wieghardt, K. (**2006**) Effect of N-Methylation of Macrocyclic Amine Ligands on the Spin State of Fe(III): A Tale of Two Fluoro Complexes. *Inorg. Chem.*, 45, 2027–2037.

111. Kapre, R.; Ray, K.; Sylvestre, I.; Weyhermüller, T.; DeBeer-George, S.; Neese, F.; Wieghardt, K. (**2006**) The Molecular and Electronic Structure of Oxo-bis(benzene-1,2-dithiolato)chromate(V) Monoanions. A Combined Experimental and Density Functional Study. *Inorg. Chem.*, 45, 3499–3509.

112. Zhu, W.; Marr, A.C.; Wang, Q.; Neese, F.; Spencer, J.E.; Blake, A.J.; Cooke, P.A.; Wilson, C.; Schröder, M. (**2005**) Modulation of the Electronic Structure and the Ni-Fe Distance in Heterobimetallic Models for the Active Site in [NiFe]Hydrogenase: Is there a Ni-Fe Bond? *Proc. Natl. Acad. Sci. (USA)*, 102, 18280–18285.

113. Astashkin, A.V.; Neese, F.; Raitsimaring, A.M.; Cooney, J.J.A.; Bultman, E.; Enemark, J.H. (**2005**) Pulsed EPR investigation of systems modelling molybdenum enzymes: hyperfine and quadrupole parameters of oxo-^{17}O in $[Mo^{17}O(SPh)_4]^-$, *J. Am. Chem. Soc.,* 127, 16713–16722.

114. Benisvy, L.; Bittl, R.; Bothe, E.; Garner, C.D.; McMaster, J.; Ross, S.; Teutloff, C.; Neese, F. (**2005**) Phenoxyl Radicals Hydrogen-Bonded to Imidazolium – Analogues of Tyrosyl $D^\bullet$ of Photosystem II: High-Field EPR and DFT Studies. *Angew. Chem. Int. Ed.*, 44, 5314–5317.

115. Praneeth, V.K.K.; Neese, F.; Lehnert, N. (**2005**) Spin Density Distribution in Five- and Six-Coordinate Iron(II)-Porphyrin NO Complexes Evidenced by Magnetic Circular Dichroism Spectroscopy. *Inorg. Chem.*, 44, 2570–2572.

116. Sinnecker, S.; Neese, F.; Lubitz, W. (**2005**) Dimanganese Catalase – Spectroscopic Parameters from Broken Symmetry Density Functional Theory of the Superoxidized Mn^{III}/Mn^{IV} state, *J. Biol. Inorg. Chem.*, 10, 231–238.

117. Blanchard, S.; Neese, F.; Bothe, E.; Bill, E.; Weyhermüller, T.; Wieghardt, K. (**2005**) Square Planar vs. Tetrahedral Coordination in Diamagnetic Complexes of Nickel(II) Containing Two Bidentate π Radical Monoanions, *Inorg. Chem.*, 44, 3636–3656.

118. Mader-Cosper, M.; Neese, F.; Astashkin, A.V.; Carducci, M.A.; Raitsimring, A.M.; Enemark, J.H. (**2005**) Determination of the Magnitude and Orientation of the g-Tensors for *cis,trans*-(L-N_2S_2)Mo^{V}OX (X=Cl, SCH_2Ph) by Single Crystal EPR and Molecular Orbital Calculations, *Inorg. Chem.*, 44, 1290–1301.

119. Fouqeau, A.; Casida, M.E.; Lawson, L.M.; Hauser, A.; Neese, F. (**2005**) Comparison of Density Functionals for Energy and Structural Differences Between the High-[$^5T_{2g}$: $(t_{2g}^4)(e_g^2)$] and Low-[$^1A_{1g}$: $(t_{2g}^6)(e_g^0)$] Spin States of Iron(II) Coordination Compounds: II. Comparison of Results for More than Ten Modern Functionals with Ligand Field Theory and *Ab Initio* Results for Hexaquoferrous Dication, $[Fe(H_2O)_6]^{2+}$ and Hexaminoferrous Dication $[Fe(NH_3)_6]^{2+}$, *J. Chem. Phys.*, 122, 044110.

120. Aliaga-Alcade, N.; DeBeer George, S.; Bill, E.; Wieghardt, K.; Neese, F. (**2005**) The Geometric and Electronic Structure of [(Cyclam-acetato)Fe(N)]+: a Genuine Iron(V) Species with Ground State Spin $S = 1/2$. *Angew. Chem. Int. Ed.*, 44, 2908–2912.

121. Bill, E.; Bothe, E.; Chaudhuri, P.; Chlopek, K.; Herebian, D.; Kokatam, S.; Ray, K. Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2004**) Molecular and Electronic Structure of Four- and Five-Coordinate Cobalt Complexes Containing Two *o*-Phenylendiamine- or Two *o*-Aminophenol-Type Ligands at Various Oxidation Levels: An Experimental, Density Functional and Correlated *ab initio* Study. *Chem. Eur. J.*, 11, 204–224.

122. Paine, T.; Bothe, W.; Bill, E.; Weyhermüller, T.; Slep, L.; Neese, F.; Chaudhuri, P. (**2004**) Nonoxo Vanadium(IV) and Vanadyl(V) Complexes with Mixed O,X,O-Donor Ligand (X = S, Se, P, PO), *Inorg. Chem.*, 43, 7324–7338.

123. Baute, D.; Arieli, D.; Zimmermann, H.; Neese, F.; Weckhuysen, B.; Goldfarb, D. (**2004**) The Structure of Copper Histidine Complexes in Solution and in Zeolite Y: A Combined X- and W-Band Pulsed EPR/ENDOR and DFT Study, *J. Am. Chem. Soc.*, 126, 11733–11745.

124. Garcia Serres R.; Grapperhaus, C.A.; Bothe, E.; Bill, E.; Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2004**) Structural, Spectroscopic and Computational Study of an Octahedral, Non-heme $FeNO^{6,7,8}$ Series: $[Fe(NO)(cyclam\text{-}ac)]^{2+/1+/0}$, *J. Am. Chem. Soc.*, 126, 5138–5153.

125. Sinnecker, S.; Noodleman, L.; Neese, F.; Lubitz, W. (**2004**) Calculation of the EPR Parameters of a Mixed Valence Mn(III)/Mn(IV) Model Complex with Broken Symmetry Density Functional Theory. *J. Am. Chem. Soc.*, 126, 2613–2622.

126. Sinnecker, S.; Neese, F.; Lubitz, W. (**2004**) Benzosemichinone Solvent Interactions. A Density Functional Study of Electric and Magnetic Properties for Probing Hydrogen Bond Strengths and Geometries. *J. Am. Chem. Soc.*, 126, 3280–3290.

127. van Gastel, M.; Fichtner, C.; Neese, F.; Lubitz, W. (**2005**) EPR Experiments to Elucidate the Structure of the Ready and Unready States of the [NiFe] Hydrogenase of *Desulfovibrio vulgaris* Miyazaki F. *Biochem. Soc. Trans.*, 33, 7–11.

128. van Gastel, M.; Lassman, G.; Lubitz, W.; Neese, F. (**2004**) The unusual EPR parameters of the cysteine radical: a DFT and correlated *ab initio* study *J. Am. Chem. Soc.,* 126, 2237–2246.

129. Fouqueau, A.; Mer, S.; Casida, M.E.; Daku, L.M.L.; Hauser, A.; Mieva, T.; Neese, F. (**2004**) Comparison of Density Functionals for Energy and Structural Differences between the High [$^5T_{2g}$: $t_{2g}^4e_g^2$] and Low [$^1A_{1g}$: t_{2g}^6] Spin States of the Hexaquo-Ferrous Ion, $[Fe(H_2O)_6]^{2+}$, *J. Chem. Phys.*, 120, 9473–9486.

130. Slep, L.D.; Mijovilovich, A.; Meyer-Klaucke, W.; Weyhermüller, T.; Bill, E.; Bothe, E.; Neese, F.; Wieghardt, K. (**2003**) The Mixed-valent $Fe^{IV}(\mu\text{-O})(\mu\text{-carboxylato})_2Fe^{III3+}$ Core. *J. Am. Chem. Soc.*, 125, 15554–15570.

131. Herebian, D.; Wieghardt, K.; Neese, F. (**2003**) Analysis and Interpretation of Metal-Radical Coupling in a Series of Square Planar Nickel Complexes. Correlated *Ab Initio* and Density Functional Investigation of [Ni(L^{ISQ})$_2$] (L^{ISQ} =3,5-di-tert-butyl-*o*diiminobenzosemquinone). *J. Am. Chem. Soc.*, 125, 10997–11005.

132. Herebian, D.; Bothe, E.; Neese, F.; Weyhermüller, T.; Wieghardt, K. (**2003**) The Molecular and Electronic Structures of Bis(*o*-diiminobenzosemiquinonato)metal(II) Complexes (Ni, Pd, Pt), their Monocations and Anions, and their Dimeric Dications Containing Weak Metal-Metal Bonds. *J. Am. Chem. Soc.*, 125, 9116–9128.

133. Ghosh, P.; Bill, E.; Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2003**) The non-Innocence of the Ligand Glyoxal-bis (2-mercaptoanil). The Electronic Structures of [Fe(gma)]$_2$, [Fe(gma)(py)]$^{\bullet}$py, [Fe(gma)(CN)]$^{1-/0}$, [Fe(gma)I], [Fe(gma)(PR_3)$_n$] ($n = 1, 2$). Experimental and Theoretical Evidence for 'Excited State' Coordination. *J. Am. Chem. Soc.*, 125, 1293–1308.

134. Einsle, O.; Messerschmidt, A.; Huber, R.; Kroneck, P.M.H.; Neese, F. (**2002**) Mechanism of the Six Electron Reduction of Nitrite to Ammonia by Cytochrome c Nitrite Reductase (CCNIR). *J. Am. Chem. Soc.*, 124, 11737–11745.

135. Sun, X.; Chun, H.; Hildenbrand, K.; Bothe, E.; Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2002**) *o*-Iminobenzosemiquinonato(1-) and *o*-Amidophenolato(2-) Complexes of Palladium(II) and Plantinum(II): A Combined Experimental and Density Functional Theoretical Study, *Inorg. Chem.*, 41, 4295–4303.

136. Li, M.; Bonnet, D.; Bill, E.; Neese, F.; Weyhermüller, T.; Blum, N.; Sellmann, D.; Wieghardt, K. (**2002**) Tuning the Electronic Structure of Octahedral Iron Complexes [FeL(X)] (L = 1-alkyl-4,7-bis(4-tert-butyl-2-mercaptobenzyl)-1,4,7-triazacyclo-nonane, X = Cl, CH_3O, CN, CO). The $S = 1/2 \Leftrightarrow S = 3/2$ Spin-Equilibrium of [FeL^{Pr}(NO)]. *Inorg. Chem.*, 41, 3444–3456.

137. Lehnert, N.; Neese, F.; Ho, R.Y.N.; Que Jr., L.; Solomon, E.I. (**2002**) Electronic Structure and Reactivity of Low-Spin Fe(III)-Hydroperoxo Complexes: Comparison to Activated Bleomycin. *J. Am. Chem. Soc.*, 124, 10810–10822.

138. Grapperhaus, C.A.; Bill, E.; Weyhermüller, T.; Neese, F.; Wieghardt, K. (**2001**) Electronic and Geometric Structure and Spectroscopy of a High Valent Manganese(V) Nitrido Complex. An Experimental and DFT Study. *Inorg. Chem.*, 41, 4191–4198.

139. Neese, F., Solomon, E.I. (**1998**) Detailed Spectroscopic and Theoretical Studies on [Fe(EDTA)(O_2)]$^{3-}$: the Electronic Structure of the Side-On Ferric Peroxide Bond and its Relevance to Reactivity. *J. Am. Chem. Soc.*, 120, 12829–12848.

A.5 Reviews of interest

1. Atanasov, M.; Aravena, D.; Suturina, E.; Bill, E.; Maganas, D.; Neese, F. (**2015**) First principles approach to the electronic structure, magnetic anisotropy and spin relaxation in mononuclear 3d-transition metal single molecule magnets, *Coord. Chem. Rev.*, 289, 177-214.

2. Neese, F.; Liakos, D. G.; Ye, S. F. (**2011**) Correlated Wavefunction Methods in Bioinorganic Chemistry, *J. Biol. Inorg. Chem,* 16, 821–829.

3. Neese, F.; Ames, W.; Christian, G.; Kampa, M.; Liakos, D. G.; Pantazis, D. A.; Roemelt, M.; Surawatanawong, P.; Ye, S. F. (**2010**) Dealing with Complexity in Open-Shell Transition Metal Chemistry from a Theoretical Perspective: Reaction Pathways, Bonding, Spectroscopy, and Magnetic Properties, *Adv. Inorg. Chem.*, 62, 301–349.

4. Orio, M.; Pantazis, D. A.; Neese, F. (**2009**) Density Functional Theory, *Photosynth. Res.*, 102, 443–453.

5. Neese, F. (**2009**), Density Functional Theory and EPR Spectroscopy: a guided tour. *EPR Newsletter*, 18(4), Pro & Contra section.

6. Neese, F. (**2009**) Prediction of Molecular Spectra and Molecular Properties with Density Functional Theory: from Fundamental Theory to Exchange Coupling. *Coord. Chem. Rev.*, 253, 526–563.

7. Neese, F. (**2009**) Spin Hamiltonian Parameters from First Principle Calculations: Theory and Application. In: Hanseon, G.; Berliner, L. (Eds.) *Biological Magnetic Resonance. Vol 28*, pp 175–232.

8. Ray, K.; Petrenko, T.; Wieghardt, K.; Neese, F. (**2007**) Joint Spectroscopic and Theoretical Investigations of Transition Metal Complexes Involving Non-Innocent Ligands. *Dalton Trans.*, 1552. (selected for cover picture)

9. Kirchner, B.; Wennmohs, F.; Ye, S.; Neese, F. (**2007**) Theoretical Bioinorganic Chemistry: Electronic Structure Makes a Difference, *Curr. Op. Chem. Biol.*, 11, 131–141.

10. Neese, F.; Petrenko, T.; Ganyushin, D.; Olbrich, G. (**2007**) Advanced Aspects of *ab initio* Theoretical Spectroscopy of Open-Shell Transition Metal Ions. *Coord. Chem. Rev.*, 205, 288–327.

11. Ye, S.; Neese, F. (**2006**) Combined Quantum Chemical and Spectroscopic Studies on Transition Metal Complexes with Coordinating Radicals. *Chemtracts* (Special Volume on Computational Inorganic Chemistry), 19, 77–86.

12. Sinnecker, S.; Neese, F. (**2006**) Theoretical Bioinorganic Spectroscopy, Invited Chapter in the Series *Current Topics in Chemistry*, Editor M. Reiher, Springer, Heidelberg.

13. Neese, F. (**2006**) Quantum Chemical Approaches to Spin-Hamiltonian Parameters. *Specialist Periodical Reports on EPR Spectroscopy Vol. 20*, (Ed. B. Gilbert) Royal Scoiety Press.

14. Neese, F. (**2006**) A Critical Evaluation of DFT, including Time-Dependent DFT, Applied to Bioinorganic Chemistry. *J. Biol. Inorg. Chem.*, (commentary on invitation), 11, 702–711.

15. Neese, F.; Munzarova, M.L. (**2004**) Historical Aspects of EPR Parameter Calculations. In: Kaupp, M.; Bühl, M.; Malkin, V. (Eds) *Calculation of NMR and EPR Parameters. Theory and Applications*. Wiley-VCH, pp 21–32.

16. Neese, F. (**2004**) Zero-Field Splitting. In: Kaupp, M.; Bühl, M.; Malkin, V. (Eds) *Calculation of NMR and EPR Parameters. Theory and Applications*. Wiley-VCH, pp 541–564.

17. Neese, F. (**2004**) Application of EPR Parameter Calculations in Bioinorganic Chemistry. In: Kaupp, M.; Bühl, M.; Malkin, V. (Eds) *Calculation of NMR and EPR Parameters. Theory and Applications*. Wiley-VCH, pp 581–591.

18. Neese, F. (**2003**) Quantum Chemical Calculations of Spectroscopic Properties of Metalloproteins and Model Compounds: EPR and Mössbauer Properties. *Curr. Op. Chem. Biol.*, 7, 125–135.

19. Neese, F.; Solomon, E.I. (**2003**) Calculation and Interpretation of Spin-Hamiltonian Parameters in Transition Metal Complexes. Invited review, (Wiley series: Magnetoscience – From Molecules to Materials edited by J.S. Miller and M. Drillon), Volume IV, p 345–466.

APPENDIX

B

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] National institute of standards and technology (NIST) atomic spectra database.

[2] https://sites.google.com/site/orcainputlibrary/.

[3] V. Ásgeirsson. *Development and Evaluation of Computational Methods for Studies of Chemical Reactions*. University of Iceland, 2021.

[4] V. Ásgeirsson, Birgirsson B.O., R. Bjornsson, U. Becker, C. Riplinger, F. Neese, and H. Jónssson. Nudged elastic band method for molecular reactions using energy-weighted springs combined with eigenvector following. *J. Chem. Theory Comput.*, 17:4929, 2021. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.1c00462, doi:https://doi.org/10.1021/acs.jctc.1c00462.

[5] Vilhjálmur Ásgeirsson, Andri Arnaldsson, and Hannes Jónsson. Efficient evaluation of atom tunneling combined with electronic structure calculations. *J. Chem. Phys.*, 148(10):102334, 2018. URL: https://doi.org/10.1063/1.5007180, doi:10.1063/1.5007180.

[6] J. Řezáč, K. E. Riley, and P. Hobza. *J. Chem. Theory Comput.*, 7:2427, 2011.

[7] C. Adamo and V. Barone. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: the mpw and mpw1pw models. *J. Chem. Phys.*, 108:664, 1998. URL: https://pubs.aip.org/aip/jcp/article-abstract/108/2/664/182704/, doi:https://doi.org/10.1063/1.475428.

[8] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: the pbe0 model. *J. Chem. Phys.*, 110:6158, 1999. URL: https://pubs.aip.org/aip/jcp/article-abstract/110/13/6158/476177/, doi:https://doi.org/10.1063/1.478522.

[9] C. Adamo, A. di Matteo, and V. Barone. From classical density functionals to adiabatic connection methods. the state of the art. *Adv. Quant. Chem.*, 36:45, 2000. URL: https://www.sciencedirect.com/science/article/abs/pii/S0065327608604785, doi:https://doi.org/10.1016/S0065-3276(08)60478-5.

[10] R Ahlrichs, H Lischka, V Staemmler, and W Kutzelnigg. Pno–ci (pair natural orbital configuration interaction) and cepa–pno (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. i. outline of the method for closed-shell states. *The Journal of Chemical Physics*, 62(4):1225–1234, 1975.

[11] R. Ahlrichs. Many body perturbation calculations and coupled electron pair models. *Comp. Phys. Comm.*, 17:31, 1979. URL: https://www.sciencedirect.com/science/article/abs/pii/0010465579900675, doi:https://doi.org/10.1016/0010-4655(79)90067-5.

[12] R. Ahlrichs. In P. v. R. Schleyer, editor, *Encyclopedia of Computational Chemistry*, pages 3123. John Wiley and Sons, 1998. URL: https://onlinelibrary.wiley.com/page/book/10.1002/0470845015/homepage/editorscontributors.html, doi:https://doi.org/10.1002/0470845015.

[13] R. Ahlrichs, M. Bär, H. P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häser, H. Horn, C. Huber, U. Huniar, M. Kattanek, C. Kölmel, M. Kollwitz, K. May, C. Ochsenfeld, H. Öhm, A. Schäfer, U. Schneider, O. Treutler, M. von Arnim, F. Weigend, P. Weis, and H. Weiss. *TurboMole - Program System for Ab Initio Electronic Structure Calculations, Version 5.2*. Universität Karlsruhe, Karlsruhe, Germany, 2000. URL: https://www.turbomole.org/.

[14] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel. Electronic structure calculations on workstation computers: the program system turbomole. *Chem. Phys. Lett.*, 162:165, 1989. URL: https://www.sciencedirect.com/science/article/abs/pii/0009261489851188, doi:https://doi.org/10.1016/0009-2614(89)85118-8.

[15] R. Ahlrichs and P. Scharf. In K. P. Lawley, editor, *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry, Part I*, Advances in Chemical Physics. Wiley, 1987.

[16] R. Ahlrichs, P. Scharf, and C. Ehrhardt. The coupled pair functional (cpf). a size consistent modification of the ci(sd) based on an energy functional. *J. Chem. Phys.*, 82:890, 1985. URL: https://pubs.aip.org/aip/jcp/article-abstract/82/2/890/153875, doi:https://doi.org/10.1063/1.448517.

[17] R. Ahlrichs and coworkers. Unpublished.

[18] D. R. Alcoba, L. Lain, A. Torre, and R. C. Bochicchio. Local spin: a treatment beyond single determinant wave functions. *Chem. Phys. Lett.*, 470:136, 2009. URL: https://www.sciencedirect.com/science/article/abs/pii/S000926140900044X, doi:https://doi.org/10.1016/j.cplett.2009.01.034.

[19] J. Almlöf. Direct Methods in Electronic Structure Theory. In D. R. Yarkony, editor, *Modern Electronic Structure Theory*, pages 110. World Scientific, 1995. URL: https://www.worldscientific.com/worldscibooks/10.1142/1957#t=aboutBook, doi:https://doi.org/10.1142/1957.

[20] J. Almlöf and P. R. Taylor. Computational Aspects of Direct SCF and MCSCF Methods. In C. E. Dykstra, editor, *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*, pages 107. Springer, 1984. URL: https://link.springer.com/content/pdf/10.1007/978-94-009-6451-8.pdf#page=115, doi:https://doi.org/10.1007/97S-94-009-6451-S.

[21] K. Almlöf, J. Faegri and K. Korsell. Principles for a direct scf approach to licao\textendashmoab -initio calculations. *J. Comput. Chem.*, 3:385, 1982. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540030314, doi:https://doi.org/10.1002/jcc.540030314.

[22] A. Altun, F. Neese, and G. Bistoni. Local energy decomposition analysis of hydrogen-bonded dimers within a domain-based pair natural orbital coupled cluster study. *Beilstein J. Org. Chem.*, 14:919–929, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000430878300002, doi:10.3762/bjoc.14.79.

[23] A. Altun, F. Neese, and G. Bistoni. Extrapolation to the limit of a complete pair natural orbital space in local coupled-cluster calculations. *J. Chem. Theory Comput.*, 16(10):6142–6149, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000580954000015, doi:https://doi.org/10.1021/acs.jctc.0c00344.

[24] A. Altun, F. Neese, and G. Bistoni. Open-shell variant of the london dispersion-corrected hartree-fock method hfld for the quantification and analysis of noncovalent interaction energies. *J. Chem. Theory Comput.*, 18(4):2292–2307, 2022. doi:10.1021/acs.jctc.1c01295.

[25] Ahmet Altun, Miquel Garcia-Ratés, Frank Neese, and Giovanni Bistoni. Unveiling the complex pattern of intermolecular interactions responsible for the stability of the dna duplex. *Chemical Science*, 12(38):12785–12793, 2021. URL: https://pubs.rsc.org/en/content/articlehtml/2021/sc/d1sc03868k, doi:https://doi.org/10.1039/D1SC03868K.

[26] Ahmet Altun, Soumen Ghosh, Christoph Riplinger, Frank Neese, and Giovanni Bistoni. Addressing the system-size dependence of the local approximation error in coupled-cluster calculations. *The Journal of Physical Chemistry A*, 125(45):9932–9939, 2021. URL: https://pubs.acs.org/doi/full/10.1021/acs.jpca.1c09106, doi:https://doi.org/10.1021/acs.jpca.1c09106.

[27] Ahmet Altun, Róbert Izsák, and Giovanni Bistoni. Local energy decomposition of coupled-cluster interaction energies: Interpretation, benchmarks, and comparison with symmetry-adapted perturbation theory. *Int. J. Quantum Chem.*, 121(3):e26339, 2021. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.26339, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.26339, doi:10.1002/qua.26339.

[28] Ahmet Altun, Frank Neese, and Giovanni Bistoni. Local energy decomposition analysis of hydrogen-bonded dimers within a domain-based pair natural orbital coupled cluster study. *Beilstein J. Org. Chem.*, 14:919, 2018. URL: https://pubs.acs.org/doi/full/10.1021/acs.jctc.8b01145, doi:https://doi.org/10.1021/acs.jctc.8b01145.

[29] Ahmet Altun, Frank Neese, and Giovanni Bistoni. Effect of electron correlation on intermolecular interactions: A pair natural orbitals coupled cluster based local energy decomposition study.

J. Chem. Theory Comput., 15(1):215–228, 2019. URL: https://doi.org/10.1021/acs.jctc.8b00915, doi:https://doi.org/10.1021/acs.jctc.8b00915.

[30] Ahmet Altun, Frank Neese, and Giovanni Bistoni. HFLD: A nonempirical london dispersion-corrected Hartree–Fock method for the quantification and analysis of noncovalent interaction energies of large molecular systems. *J. Chem. Theory Comput.*, 15(11):5894–5907, 2019. URL: https://doi.org/10.1021/acs.jctc.9b00425, arXiv:https://doi.org/10.1021/acs.jctc.9b00425, doi:10.1021/acs.jctc.9b00425.

[31] Ahmet Altun, Frank Neese, and Giovanni Bistoni. Extrapolation to the limit of a complete pair natural orbital space in local coupled-cluster calculations. *Journal of Chemical Theory and Computation*, 16(10):6142–6149, 2020. URL: https://pubs.acs.org/doi/full/10.1021/acs.jctc.0c00344, doi:https://doi.org/10.1021/acs.jctc.0c00344.

[32] Ahmet Altun, Christoph Riplinger, Frank Neese, and Giovanni Bistoni. Exploring the accuracy limits of pno-based local coupled-cluster calculations for transition-metal complexes. *Journal of Chemical Theory and Computation*, 19(7):2039–2047, 2023.

[33] Ahmet Altun, Masaaki Saitow, Frank Neese, and Giovanni Bistoni. Local energy decomposition of open-shell molecular systems in the domain-based local pair natural orbital coupled cluster framework. *J. Chem. Theory Comput.*, 15(3):1616–1632, 2019. URL: https://doi.org/10.1021/acs.jctc.8b01145, doi:https://doi.org/10.1021/acs.jctc.8b01145.

[34] A. T. Amos and G. G. Hall. Single determinant wave functions. *Proc. R. Soc. Ser. A.*, 263:483, 1961. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1961.0175, doi:https://doi.org/10.1098/rspa.1961.0175.

[35] Hans C. Andersen. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.*, 52(1):24–34, 1983. URL: https://www.sciencedirect.com/science/article/abs/pii/0021999183900141, doi:https://doi.org/10.1016/0021-9991(83)90014-1.

[36] W. P. Anderson, T. R. Cundari, R. S. Drago, and M. C. Zerner. Utility of the semiempirical indo/1 method for the calculation of the geometries of second-row transition-metal species. *Inorg. Chem.*, 29:3, 1990. URL: https://pubs.acs.org/doi/pdf/10.1021/ic00326a001, doi:https://doi.org/10.1021/ic00326a001.

[37] W. P. Anderson, T. R. Cundari, and M. C. Zerner. An intermediate neglect of differential overlap model for second-row transition metal species. *Int. J. Quant. Chem.*, 39:31, 1991. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.560390106, doi:https://doi.org/10.1002/qua.560390106.

[38] W. P. Anderson, W. D. Edwards, and M. C. Zerner. Calculated spectra of hydrated ions of the first transition-metal series. *Inorg. Chem.*, 25:2728, 1986. URL: https://pubs.acs.org/doi/pdf/10.1021/ic00236a015, doi:https://doi.org/10.1021/ic00236a015.

[39] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. Second-order perturbation theory with a casscf reference function. *J. Phys. Chem.*, 94:5483–5488, 1990. URL: https://pubs.acs.org/doi/pdf/10.1021/j100377a012, doi:https://doi.org/10.1021/j100377a012.

[40] Kerstin. Andersson, Per Aake. Malmqvist, Bjoern O. Roos, Andrzej J. Sadlej, and Krzysztof. Wolinski. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.*, 94:5483–5488, 07 1990. doi:https://doi.org/10.1021/j100377a012.

[41] D. Andrae, U. Häu\ssermann , M. Dolg, H. Stoll, and H. Preuss. Energy-adjusted ab\textendashinitio pseudopotentials for the second and third row transition elements. *Theor. Chim. Acta*, 77:123–141, 1990. URL: https://www.scirp.org/(S(lz5mqp453edsnp55rrgjct55))/reference/referencespapers.aspx?referenceid=2566997, doi:https://doi.org/10.1007/BF01114537.

[42] J. Andzelm, M. Klobukowski, E. Radzio-Andzelm, Y. Sakai, and H. Tatewaki. In S. Huzinaga, editor, *Gaussian Basis Sets for Molecular Calculations*. Elsevier, 1984. URL: https://books.google.com/books?hl=en&lr=&id=2lBK9AgvnmIC&oi=fnd&pg=PP1&dq=Andzelm,+J.+and+Klobukowski,+M.+and+Radzio-Andzelm,+E.+and+Sakai,+Y.+and+Tatewaki,+H.&ots=NBQHZHfC3A&sig=btp_M0PZvYGZJhCLv8T5htEZVEM.

[43] C. Angeli, S. Borini, M. Cestari, and R. Cimigraglia. A quasidegenerate formulation of the second order n-electron valence state perturbation theory approach. *J. Chem. Phys.*, 121:4043, 2004. URL: https://pubs.aip.org/aip/jcp/article-abstract/121/9/4043/186965/, doi:https://doi.org/10.1063/1.1778711.

[44] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.P. Malrieu. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.*, 114:10252–10264, 2001. URL: https://pubs.aip.org/aip/jcp/article-abstract/114/23/10252/183833/, doi:https://doi.org/10.1063/1.1361246.

[45] C. Angeli, R. Cimiraglia, and J.P. Malrieu. N-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chem. Phys. Lett.*, 350:297–305, 2001. URL: https://www.sciencedirect.com/science/article/abs/pii/S0009261401013033, doi:https://doi.org/10.1016/S0009-2614(01)01303-3.

[46] C. Angeli, R. Cimiraglia, and J.P. Malrieu. N-electron valence state perturbation theory: a spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.*, 117:9138–9153, 2002. URL: https://pubs.aip.org/aip/jcp/article-abstract/117/20/9138/464690/, doi:https://doi.org/10.1063/1.1515317.

[47] Celestino Angeli. On the nature of the π \textrightarrow π* ionic excited states: The V state of ethene as a prototype. *J. Comput. Chem.*, 30:1319–1333, 2009. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21155, doi:https://doi.org/10.1002/jcc.21155.

[48] Celestino Angeli, Stefano Borini, Mirko Cestari, and Renzo Cimiraglia. A quasidegenerate formulation of the second order n-electron valence state perturbation theory approach. *J. Chem. Phys.*, 121:4043–4049, 08 2004. URL: https://pubs.aip.org/aip/jcp/article-abstract/121/9/4043/186965/, doi:https://doi.org/10.1063/1.1778711.

[49] Victor Anisimov and James JP Stewart. *Introduction to the Fast Multipole Method: Topics in Computational Biophysics, Theory, and Implementation*. CRC Press, 2019.

[50] Francesco Aquilante, Per-Å ke Malmqvist, Thomas Bondo Pedersen, Abhik Ghosh, and Björn Olof Roos. Cholesky Decomposition-Based Multiconfiguration Second-Order Perturbation Theory (CD-CASPT2): Application to the Spin-State Energetics of CoIII(diiminato)(NPh). *J. Chem. Theory Comput.*, 4:694–702, 05 2008. URL: https://pubs.acs.org/doi/abs/10.1021/ct700263h, doi:https://doi.org/10.1021/ct700263h.

[51] Juan Arago, Enrique Orti, and Juan C. Sancho-Garcia. Nonlocal van der Waals Approach Merged with Double-Hybrid Density Functionals: Toward the Accurate Treatment of Noncovalent Interactions. *J. Chem. Theory Comput.*, 9(8):3437–3443, 08 2013. URL: https://doi.org/10.1021/ct4003527 (visited on 2020-09-15), doi:10.1021/ct4003527.

[52] D. Aravena, F. Neese, and Dimitrios A. Pantazis. Improved segmented all-electron relativistically contracted basis sets for the lanthanides. *J. Chem. Theory Comput.*, 12:1148–1156, 2016. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.5b01048, doi:https://doi.org/10.1021/acs.jctc.5b01048.

[53] Daniel Aravena, Mihail Atanasov, and Frank Neese. Periodic Trends in Lanthanide Compounds through the Eyes of Multireference ab Initio Theory. *Inorg. Chem.*, 55(9):4457–4469, 05 2016. URL: https://pubs.acs.org/doi/abs/10.1021/acs.inorgchem.6b00244, doi:10.1021/acs.inorgchem.6b00244.

[54] Behnam Assadollahzadeh and Peter Schwerdtfeger. A systematic search for minimum structures of small gold clusters Aun (n=2–20) and their electronic properties. *The Journal of Chemical Physics*, 131(6):064306, August 2009. Publisher: American Institute of Physics. URL: https://aip.scitation.org/doi/full/10.1063/1.3204488 (visited on 2022-04-22), doi:https://doi.org/10.1063/1.3204488.

[55] X. Assfeld and J.-L. Rivail. Quantum chemical computations on parts of large molecules: the ab initio local self consistent field method. *Chem. Phys. Lett.*, 263:100–106, 12 1996. doi:10.1016/S0009-2614(96)01165-7.

[56] Mihail Atanasov, Dmitry Ganyushin, Dimitrios A. Pantazis, Kantharuban Sivalingam, and Frank Neese. Detailed Ab Initio First-Principles Study of the Magnetic Anisotropy in a Family of Trigonal Pyramidal Iron(II) Pyrrolide Complexes. *Inorg. Chem.*, 50(16):7460–7477, 08 2011. URL: https://pubs.acs.org/doi/abs/10.1021/ic200196k, doi:10.1021/ic200196k.

[57] Mihail Atanasov, Dmitry Ganyushin, Kantharuban Sivalingam, and Frank Neese. A Modern First-Principles View on Ligand Field Theory Through the Eyes of Correlated Multireference Wavefunctions. In David Michael P. Mingos, Peter Day, and Jens Peder Dahl, editors, *Molecular Electronic Structures of Transition Metal Complexes II*, number 143 in Structure and Bonding, pages 149–220. Springer Berlin Heidelberg, 2011. URL: https://link.springer.com/chapter/10.1007/430_2011_57, doi:https://doi.org/10.1007/430_2011_57.

[58] J. E. Atkins, E. G. Boman, and B. Hendrickson. A spectral algorithm for seriation and the consecutive ones problem. *SIAM J. Computing.*, 28(1):297–310, 1998. URL: https://epubs.siam.org/doi/abs/10.1137/S0097539795285771, doi:https://doi.org/10.1137/S0097539795285771.

[59] A. A. Auer and Stanton J. F. Gauss, J. Quantitative prediction of gas-phase ^{13}c nuclear magnetic shielding constants. *J. Chem. Phys.*, 118:10407, 2003. URL: https://pubs.aip.org/aip/jcp/article-abstract/118/23/10407/844083/, doi:https://doi.org/10.1063/1.1574314.

[60] A. A. Auer, V. A. Tran, B. Sharma, G. L. Stoychev, D. Marx, and F. Neese. A case study of density functional theory and domain-based local pair natural orbital coupled cluster for vibrational effects on epr hyperfine coupling constants: vibrational perturbation theory versus ab-initio molecular dynamics. *Mol. Phys.*, 118:16, 2020. URL: https://www.tandfonline.com/doi/full/10.1080/00268976.2020.1797916, doi:https://doi.org/10.1080/00268976.2020.1797916.

[61] M. C. Böhm and R. Gleiter. A cndo/indo molecular orbital formalism for the elements h to br. theory. *Theor. Chim. Acta*, 59:127 & 153, 1981. URL: https://link.springer.com/article/10.1007/BF00552536, doi:https://doi.org/10.1007/BF00552536.

[62] M. Bühl, C. Reimann, D. A. Pantazis, T. Bredow, and F. Neese. Geometries of third-row transition-metal complexes from density-functional theory. *J. Chem. Theory Comput.*, 4:1449–1459, 2008. URL: https://pubs.acs.org/doi/abs/10.1021/ct800172j, doi:https://doi.org/10.1021/ct800172j.

[63] A. D. Bacon and M. C. Zerner. An intermediate neglect of differential overlap theory for transition metal complexes: fe, co and cu chlorides. *Theor. Chim. Acta*, 53:21, 1979. URL: https://link.springer.com/article/10.1007/BF00547605, doi:https://doi.org/10.1007/BF00547605.

[64] George B. Bacskay. A quadratically convergent Hartree–Fock (QC-SCF) method. Application to closed shell systems. *Chem. Phys.*, 61(3):385–404, 1981. URL: https://www.sciencedirect.com/science/article/abs/pii/0301010481851567, doi:https://doi.org/10.1016/0301-0104(81)85156-7.

[65] E. J. Baerends, D. E. Ellis, and P. Ros. Self-consistent molecular hartree—fock—slater calculations i. the computational procedure. *Chem. Phys.*, 2:41, 1973. URL: https://www.sciencedirect.com/science/article/abs/pii/030101047380059X, doi:https://doi.org/10.1016/0301-0104(73)80059-X.

[66] Alberto Baiardi, Julien Bloino, and Vincenzo Barone. A general time-dependent route to Resonance-Raman spectroscopy including Franck-Condon, Herzberg-Teller and Duschinsky effects. *J. Chem. Phys.*, 141(11):114108, 09 2014. URL: http://aip.scitation.org/doi/10.1063/1.4895534, doi:https://doi.org/10.1063/1.4895534.

[67] J. Baker. An algorithm for the location of transition states. *J. Comput. Chem.*, 7:385, 1986. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540070402, doi:https://doi.org/10.1002/jcc.540070402.

[68] Jon Baker. Constrained optimization in delocalized internal coordinates. *J. Comput. Chem.*, 18(8):1079–1095, 06 1997. URL: http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(19970618:8\T1\textless{}1079::AID-JCC12\T1\textgreater{}3.0.CO;2-8/abstract, doi:10.1002/(SICI)1096-987X(199706)18:8<1079::AID-JCC12>3.0.CO;2-8.

[69] C. Bannwarth and S. Grimme. A simplified time-dependent density functional theory approach for electronic ultraviolet and circular dichroism spectra of very large molecules. *Comp. Theor. Chem.*, 1040–1041:45–53, 2014. URL: https://www.sciencedirect.com/science/article/abs/pii/S2210271X14000942, doi:https://doi.org/10.1016/j.comptc.2014.02.023.

[70] Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.*, 15(3):1652–1671, 2019. doi:10.1021/acs.jctc.8b01176.

[71] Giuseppe M. J. Barca, Andrew T. B. Gilbert, and Peter M. W. Gill. Simple Models for Difficult Electronic Excitations. *Journal of Chemical Theory and Computation*, 14(3):1501–1509, March 2018. Publisher: American Chemical Society. doi:10.1021/acs.jctc.7b00994.

[72] G. Barcza, Ö. Legeza, K. H. Marti, and M. Reiher. Quantum-information analysis of electronic states of different molecular structures. *Phys. Rev. A*, 83:012508, 2011. URL: https://link.aps.org/doi/10.1103/PhysRevA.83.012508, doi:https://doi.org/10.1103/PhysRevA.83.012508.

[73] Alessandro Barducci, Giovanni Bussi, and Michele Parrinello. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.*, 100(2):020603, 2008. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.020603, doi:10.1103/PhysRevLett.100.020603.

[74] Loïc Barnes, Baptiste Schindler, Isabelle Compagnon, and Abdul-Rahman Allouche. Fast and accurate hybrid QM//MM approach for computing anharmonic corrections to vibrational frequencies. *J. Mol. Model.*, 22(11):285, 11 2016. URL: https://doi.org/10.1007/s00894-016-3135-5 (visited on 2020-04-14), doi:10.1007/s00894-016-3135-5.

[75] V. Barone. In D. P. Chong, editor, *Recent Advances in Density Functional Methods, Part I*. World Scientific, 1996. URL: https://books.google.com/books?hl=en&lr=&id=Uu7sCgAAQBAJ&oi=fnd&pg=PR5&dq=Recent+Advances+in+Density+Functional+Methods,+Part+%7B%7BI%7D%7D&ots=Lo2ioOfVLp&sig=XRnsD7SEqIHDUQYPrb7GH2jizE4.

[76] V. Barone and M. Cossi. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A*, 102:1995–2001, 1998. URL: https://pubs.acs.org/doi/abs/10.1021/jp9716997, doi:https://doi.org/10.1021/jp9716997.

[77] Vincenzo Barone, Malgorzata Biczysko, and Julien Bloino. Fully anharmonic IR and Raman spectra of medium-size molecular systems: accuracy and interpretation. *Phys. Chem. Chem. Phys.*, 16(5):1759–1787, 2014. URL: http://dx.doi.org/10.1039/C3CP53413H, doi:10.1039/C3CP53413H.

[78] Albert P Bartók and Jonathan R Yates. Regularized SCAN functional. *J. Chem. Phys.*, 150(16):161101, 2019. URL: https://aip.scitation.org/doi/abs/10.1063/1.5094646, doi:https://doi.org/10.1063/1.5094646.

[79] Jefferson E. Bates and Filipp Furche. Harnessing the meta-generalized gradient approximation for time-dependent density functional theory. *J. Chem. Phys.*, 137(16):164105, 10 2012. URL: http://aip.scitation.org/doi/10.1063/1.4759080, arXiv:23126693, doi:https:/doi.org/10.1063/1.4759080.

[80] R. Bauernschmitt and R. Ahlrichs. Stability analysis for solutions of the closed shell kohn\textendashsham equation. *J. Chem. Phys.*, 104:9047, 1996. URL: https://pubs.aip.org/aip/jcp/article-abstract/104/22/9047/180688/, doi:https://doi.org/10.1063/1.471637.

[81] C. W. Jr. Bauschlicher, S. R. Langhoff, and L. A. Barnes. Theoretical studies of the first and second –row transition –metal methyls and their positive ions. *J. Chem. Phys.*, 91:2399, 1989. URL: https://pubs.aip.org/aip/jcp/article-abstract/91/4/2399/223082/, doi:https://doi.org/10.1063/1.456998.

[82] Krzysztof B. Bec and Christian W. Huck. Breakthrough Potential in Near-Infrared Spectroscopy: Spectra Simulation. A Review of Recent Developments. *Front. Chem.*, 7:48, 02 2019. URL: https://www.frontiersin.org/article/10.3389/fchem.2019.00048/full (visited on 2020-04-14), doi:https://doi.org/10.3389/fchem.2019.00048.

[83] M. E. Beck, C. Riplinger, F. Neese, and G. Bistoni. Unraveling individual host-guest interactions in molecular recognition from first principles quantum mechanics: Insights into the nature of nicotinic acetylcholine receptor agonist binding. *J. Comput. Chem.*, 42(5):293–302, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000591776900001, doi:https://doi.org/10.1002/jcc.26454.

[84] Michael Edmund Beck, Christoph Riplinger, Frank Neese, and Giovanni Bistoni. Unraveling individual host–guest interactions in molecular recognition from first principles quantum mechanics: Insights into the nature of nicotinic acetylcholine receptor agonist binding. *J. Comput. Chem.*, 42(5):293–302, 2021. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.26454, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.26454, doi:10.1002/jcc.26454.

[85] A. D. Becke. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.*, 88:2547, 1988. URL: https://pubs.aip.org/aip/jcp/article-abstract/88/4/2547/91134/, doi:https://doi.org/10.1063/1.454033.

[86] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098, 1988. URL: https://journals.aps.org/pra/abstract/10.1103/PhysRevA.38.3098, doi:https://doi.org/10.1103/PhysRevA.38.3098.

[87] A. D. Becke. A new mixing of hartree-fock and local density-functional theories. *J. Chem. Phys.*, 98:1372, 1993. URL: https://pubs.aip.org/aip/jcp/article-abstract/98/2/1372/821432/, doi:https://doi.org/10.1063/1.464304.

[88] A. D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.*, 98:5648, 1993. URL: https://ui.adsabs.harvard.edu/abs/1993JChPh..98.5648B/abstract, doi:https://doi.org/10.1063/1.464913.

[89] A. D. Becke. Density-functional thermochemistry. v. systematic optimization of exchange-correlation functionals. *J. Chem. Phys.*, 107:8554, 1997. URL: https://pubs.aip.org/aip/jcp/article-abstract/107/20/8554/478232/, doi:https://doi.org/10.1063/1.475007.

[90] A. D. Becke and E. R. Johnson. A density-functional model of the dispersion interaction. *J. Chem. Phys.*, 123:154101, 2005. URL: https://pubs.aip.org/aip/jcp/article-abstract/123/15/154101/347819/, doi:https://doi.org/10.1063/1.2065267.

[91] A. Bencini and D. Gatteschi. X.alpha.-sw calculations of the electronic structure and magnetic properties of weakly coupled transition-metal clusters. the [cu2cl6]2- dimers. *J. Am. Chem. Soc.*, 108:5763, 1980. URL: https://pubs.acs.org/doi/pdf/10.1021/ja00279a017, doi:https://doi.org/10.1021/ja00279a017.

[92] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81(8):3684–3690, 1984. URL: https://pubs.aip.org/aip/jcp/article-abstract/81/8/3684/565473/, doi:https://doi.org/10.1063/1.448118.

[93] G. Berghold, J. Hutter, and M. Parinello. Grid-free dft implementation of local and gradient-corrected xc functionals. *Theor. Chem. Acc.*, 99:344, 1998. URL: https://link.springer.com/article/10.1007/s002140050344, doi:https://doi.org/10.1007/s002140050344.

[94] A. Bergner, M. Dolg, W. Küchle, H. Stoll, and H. Preuss. Ab initio energy-adjusted pseudopotentials for elements of groups 13–17. *Mol. Phys.*, 80:1431–1441, 1993. URL: https://www.tandfonline.com/doi/abs/10.1080/00268979300103121, doi:https://doi.org/10.1080/00268979300103121.

[95] S. Bernadotte, A. J. Atkins, and C. R. Jacob. *J. Chem. Phys.*, 137:204106, 2012.

[96] D. E. Bernholdt and R. J. Harrison. Large-scale correlated electronic structure calculations: the ri-mp2 method on parallel computers. *Chem. Phys. Lett.*, 250:477, 1996. URL: https://www.sciencedirect.com/science/article/abs/pii/0009261496000541, doi:https://doi.org/10.1016/0009-2614(96)00054-1.

[97] A. Berning, M. Schweizer, H.J. Werner, P. J. Knowles, and P. Palmieri. Spin-orbit matrix elements for internally contracted multireference configuration interaction wavefunctions. *Mol. Phys.*, 98:1823–1833, 2000. URL: https://www.tandfonline.com/doi/abs/10.1080/00268970009483386, doi:https://doi.org/10.1080/00268970009483386.

[98] R. Berraud-Pache, F. Neese, G. Bistoni, and R. Izsak. Computational design of near-infrared fluorescent organic dyes using an accurate new wave function approach. *J. Phys. Chem. Lett.*, 10(17):4822–4828, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000484884300010, doi:https://doi.org/10.1021/acs.jpclett.9b02240.

[99] R. Berraud-Pache, E. Santamaria-Aranda, B. de Souza, G. Bistoni, F. Neese, D. Sampedro, and R. Izsak. Redesigning donor-acceptor Stenhouse adduct photoswitches through a joint experimental and computational study. *Chem. Sci.*, 12(8):2916–2924, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000625958900018, doi:https://doi.org/10.1039/d0sc06575g.

[100] Romain Berraud-Pache, Frank Neese, Giovanni Bistoni, and Róbert Izsák. Unveiling the photophysical properties of boron-dipyrromethene dyes using a new accurate excited state coupled cluster method. *J. Chem. Theory Comput.*, 16(1):564–575, 2020. URL: https://doi.org/10.1021/acs.jctc.9b00559, doi:https://doi.org/10.1021/acs.jctc.9b00559.

[101] K. Bhaskaran-Nair, O. Demel, J. Šmydke, and J. Pittner. *J. Chem. Phys.*, 134:154106, 2011.

[102] J. S. Binkley, J. A. Pople, and P. A. Dobosh. The calculation of spin-restricted single-determinant wavefunctions. *Mol. Phys.*, 28:1423, 1974. URL: https://www.tandfonline.com/doi/abs/10.1080/00268977400102701, doi:https://doi.org/10.1080/00268977400102701.

[103] J. S. Binkley, J. A. Pople, and W. J. Hehre. Self-consistent molecular orbital methods. 21. small split-valence basis sets for first-row elements. *J. Am. Chem. Soc.*, 102:939, 1980. URL: https://pubs.acs.org/doi/pdf/10.1021/ja00523a008, doi:https://doi.org/10.1021/ja00523a008.

[104] G. Bistoni, I. Polyak, M. Sparta, W. Thiel, and F. Neese. Toward accurate QM/MM reaction barriers with large QM regions using domain based pair natural orbital coupled cluster theory. *J. Chem. Theory Com-*

put., 14(7):3524–3531, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000438654500015, doi:https://doi.org/10.1021/acs.jctc.8b00348.

[105] Giovanni Bistoni. Finding chemical concepts in the Hilbert space: Coupled cluster analyses of noncovalent interactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 10(3):e1442, 2020. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1442, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1442, doi:10.1002/wcms.1442.

[106] Giovanni Bistoni, Alexander A. Auer, and Frank Neese. Understanding the role of dispersion in frustrated lewis pairs and classical lewis adducts: A domain-based local pair natural orbital coupled cluster study. *Chem. Eur. J.*, 23(4):865–873, 2017.

[107] Giovanni Bistoni, Christoph Riplinger, Yury Minenkov, Luigi Cavallo, Alexander A Auer, and Frank Neese. Treating subvalence correlation effects in domain based pair natural orbital coupled cluster calculations: an out-of-the-box approach. *J. Chem. Theory Comput.*, 2017. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.7b00352, doi:https://doi.org/10.1021/acs.jctc.7b00352.

[108] Erik Bitzek, Pekka Koskinen, Franz Gähler, Michael Moseler, and Peter Gumbsch. Structural relaxation made simple. *Phys. Rev. Lett.*, 97(17):170201, 10 2006. URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.170201, doi:https://doi.org/10.1103/PhysRevLett.97.170201.

[109] Ragnar Bjornsson and Michael Bühl. Modeling molecular crystals by QM/MM: Self-consistent electrostatic embedding for geometry optimizations and molecular property calculations in the solid. *J. Chem. Theory Comput.*, 8(2):498–508, 2012. URL: https://doi.org/10.1021/ct200824r, doi:https://doi.org/10.1021/ct200824r.

[110] J.P. Blaudeau, M. P. McGrath, L. A. Curtiss, and L. Radom. Extension of gaussian-2 (g2) theory to molecules containing third-row atoms k and ca. *J. Chem. Phys.*, 107:5016, 1997. URL: https://pubs.aip.org/aip/jcp/article-abstract/107/13/5016/476733/, doi:https://doi.org/10.1063/1.474865.

[111] P. M. Boerrigter, G. Te Velde, and E. J. Baerends. Three-dimensional numerical integration for electronic structure calculations. *Int. J. Quant. Chem.*, 33:87, 1988. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.560330204, doi:https://doi.org/10.1002/qua.560330204.

[112] J. M. Bofill, H. Bono, and J. Rubio. Analysis of the convergence of the general coupling operator method for one-configuration-type wave functions. *J. Comput. Chem.*, 19:368, 1998. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1096-987X(199802)19:3%3C368::AID-JCC10%3E3.0.CO;2-E, doi:https://doi.org/10.1002/(SICI)1096-987X(199802)19:3%3C368::AID-JCC10%3E3.0.CO;2-E.

[113] Jonathan A. Bohmann, Frank Weinhold, and Thomas C. Farrar. Natural chemical shielding analysis of nuclear magnetic resonance shielding tensors from gauge-including atomic orbital calculations. *J. Chem. Phys.*, 107(4):1173–1184, 1997. URL: https://doi.org/10.1063/1.474464, doi:10.1063/1.474464.

[114] A. Bondi. *J. Phys. Chem.*, 68:441–451, 1964.

[115] G. E. P. Box and Mervin E. Muller. A note on the generation of random normal deviates. *Ann. Math. Statist.*, 29(2):610–611, 1958. doi:10.1214/aoms/1177706645.

[116] Éric Brémond and Carlo Adamo. Seeking for parameter-free double-hybrid functionals: the PBE0-DH model. *J. Chem. Phys.*, 135(2):024106, 2011.

[117] Éric Brémond, Ángel José Pérez-Jiménez, Juan Carlos Sancho-García, and Carlo Adamo. *J. Chem. Phys.*, 150:201102, 2019.

[118] Éric Brémond, Juan Carlos Sancho-García, Ángel José Pérez-Jiménez, and Carlo Adamo. *J. Chem. Phys.*, 141:031101, 2014.

[119] Éric Brémond, Marika Savarese, Ángel José Pérez-Jiménez, Juan Carlos Sancho-García, and Carlo Adamo. *J. Chem. Theory Comput.*, 14:4052–4062, 2018.

[120] J. Brabec, J. Lang, M. Saitow, J. Pittner, F. Neese, and O. Demel. Domain-based local pair natural orbital version of mukherjee's state-specific coupled cluster method. *J. Chem. Theory Comput.*, 14(3):1370–1382, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000427661400021, doi:10.1021/acs.jctc.7b01184.

[121] J. G. Brandenburg, C. Bannwarth, A. Hansen, and S. Grimme. B97-3c: A revised low-cost variant of the B97-D density functional method. *J. Chem. Phys.*, 148(6):064104, 2018.

[122] Jan Gerit Brandenburg, Christoph Bannwarth, Andreas Hansen, and Stefan Grimme. B97-3c: A revised low-cost variant of the B97-D density functional method. *J. Chem. Phys.*, 148(6):064104, 2018. doi:10.1063/1.5012601.

[123] B.H." Brandow. *Effective Interactions and Operators in Nuclei*. Volume 40. Springer, 1974.

[124] C. M. Breneman and K. B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *J. Comput. Chem.*, 11:361–373, 1990. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540110311, doi:https://doi.org/10.1002/jcc.540110311.

[125] F. W. Brobowicz and W. A. Goddard. In H. F. Schaefer III, editor, *Methods of Electronic Structure Theory*, pages 79. Plenum Press, 1977.

[126] P. Bunting, M. Atanasov, E. Damgaard-Moller, M. Perfetti, I. Crassee, M. Orlita, J. Oyergaard, J. van Slageren, F. Neese, and J. Long. Linear cobalt(II) dialkyl complex with a non-Aufbau ground state and very large magnetic anisotropy. *Abstr. Pap. Am. Chem. Soc.*, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000525061501422.

[127] P. C. Bunting, M. Atanasov, E. Damgaard-Moller, M. Perfetti, I. Crassee, M. Orlita, J. Overgaard, J. van Slageren, F. Neese, and J. R. Long. A linear cobalt(II) complex with maximal orbital angular momentum from a non-Aufbau ground state. *Science*, 362(6421):1378–+, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000453845000052, doi:10.1126/science.aat7319.

[128] Markus Bursch, Hagen Neugebauer, Sebastian Ehlert, and Stefan Grimme. Dispersion corrected r^2SCAN based global hybrid functionals: r^2SCANh, r^2SCAN0, and r^2SCAN50. *J. Chem. Phys.*, 156(13):134105, 2022. URL: https://pubs.aip.org/aip/jcp/article-abstract/156/13/134105/2840951/, doi:https://doi.org/10.1063/5.0086040.

[129] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126(1):014101, 2007. URL: https://doi.org/10.1063/1.2408420, doi:10.1063/1.2408420.

[130] Octav Caldararu, Martin A Olsson, Christoph Riplinger, Frank Neese, and Ulf Ryde. Binding free energies in the sampl5 octa-acid host–guest challenge calculated with dft-d3 and ccsd (t). *J. Comput.-Aided Mol. Des.*, 31(1):87–106, 2017. URL: https://link.springer.com/article/10.1007/s10822-016-9957-5, doi:https://doi.org/10.1007/s10822-016-9957-5.

[131] Eike Caldeweyher, Christoph Bannwarth, and Stefan Grimme. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.*, 147(3):034112, 2017. URL: https://pubs.aip.org/aip/jcp/article-abstract/147/3/034112/595235/, doi:https://doi.org/10.1063/1.4993215.

[132] Eike Caldeweyher, Sebastian Ehlert, Andreas Hansen, Hagen Neugebauer, Sebastian Spicher, Christoph Bannwarth, and Stefan Grimme. A generally applicable atomic-charge dependent London dispersion correction. *J. Chem. Phys.*, 150(15):154122, 2019. URL: https://doi.org/10.1063/1.5090222, arXiv:https://doi.org/10.1063/1.5090222, doi:https://doi.org/10.1063/1.5090222.

[133] R. Cammi. Quantum cluster theory for the polarizable continuum model. i. the ccsd level with analytical first and second derivatives. *J. Chem. Phys*, 131:164104, 2009. URL: https://pubs.aip.org/aip/jcp/article-abstract/131/16/164104/71420/, doi:https://doi.org/10.1063/1.3245400.

[134] Roberto Cammi, Benedetta Mennucci, and Jacopo Tomasi. Fast Evaluation of Geometries and Properties of Excited Molecules in Solution: A Tamm-Dancoff Model with Application to 4-Dimethylaminobenzonitrile. *J. of Phys. Chem. A*, 104(23):5631–5637, 06 2000. URL: https://doi.org/10.1021/jp000156l (visited on 2020-05-12), doi:10.1021/jp000156l.

[135] Marco Campetella and Juan Sanz García. Following the evolution of excited states along photochemical reaction pathways. *Journal of Computational Chemistry*, 41(12):1156–1164, 2020. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.26162, doi:10.1002/jcc.26162.

[136] E. Cancès, Y. Maday, and B. Stamm. Domain decomposition for implicit solvation models. *J. Chem. Phys*, 139:054111, 2013. doi:https://doi.org/10.1063/1.4816767.

[137] X. Cao and M. Dolg. Valence basis sets for relativistic energy-consistent small-core lanthanide pseudopotentials. *J. Chem. Phys.*, 115:7348, 2001. URL: https://pubs.aip.org/aip/jcp/article-abstract/115/16/7348/458396/, doi:https://doi.org/10.1063/1.1406535.

[138] X. Cao and M. Dolg. Segmented contraction scheme for small-core lanthanide pseudopotential basis sets. *J. Mol. Struct.: THEOCHEM*, 581:139, 2002. doi:https://doi.org/10.1016/S0166-1280(01)00751-5.

[139] X. Cao and M. Dolg. Segmented contraction scheme for small-core actinide pseudopotential basis sets. *J. Mol. Struct.: THEOCHEM*, 673:203, 2004. doi:https://doi.org/10.1016/j.theochem.2003.12.015.

[140] X. Cao, M. Dolg, and H. Stoll. Valence basis sets for relativistic energy-consistent small-core actinide pseudopotentials. *J. Chem. Phys.*, 118:487, 2003. doi:https://doi.org/10.1063/1.1521431.

[141] R. Carbo and J. M. Riera. *A General SCF Theory. Lecture Notes in Chemistry*. Springer Verlag, 1978.

[142] M. Caricato. Ccsd-pcm: improving upon the reference reaction field approximation at no cost. *J. Chem. Phys*, 135:074113, 2011. doi:https://doi.org/10.1063/1.3624373.

[143] Kevin Carter-Fenk and John M. Herbert. State-Targeted Energy Projection: A Simple and Robust Approach to Orbital Relaxation of Non-Aufbau Self-Consistent Field Solutions. *Journal of Chemical Theory and Computation*, 16(8):5067–5082, 2020. Publisher: American Chemical Society. doi:10.1021/acs.jctc.0c00502.

[144] David Casanova and Anna I. Krylov. Spin-flip methods in quantum chemistry. *Phys. Chem. Chem. Phys.*, 22(8):4326–4342, 02 2020. URL: https://pubs.rsc.org/en/content/articlelanding/2020/cp/c9cp06507e (visited on 2020-10-20), doi:10.1039/C9CP06507E.

[145] M. Casanova-Páez and L. Goerigk. Assessing the tamm–dancoff approximation, singlet–singlet, and singlet–triplet excitations with the latest long-range corrected double-hybrid density functionals. *J. Chem. Phys.*, 153:064106, 2020.

[146] Marcos Casanova-Páez, Michael B. Dardis, and Lars Goerigk. ωB2PLYP & ωB2GPPLYP: the first two double-hybrid density functionals with long-range correction optimized for excitation energies. *J. Chem. Theory Comput.*, 15:4735, 2019. URL: http://dx.doi.org/10.1021/acs.jctc.9b00013, doi:10.1021/acs.jctc.9b00013.

[147] Marcos Casanova-Páez and Lars Goerigk. Time-dependent long-range-corrected double-hybrid density functionals with spin-component and spin-opposite scaling: A comprehensive analysis of singlet-singlet and singlet-triplet excitation energies. *J. Chem. Theory Comput.*, 17(8):5165–5186, 2021. doi:10.1021/acs.jctc.1c00535.

[148] M. E. Casida. Time-dependent density functional response theory for molecules. In D. P. Chong, editor, *Recent Advances in Density Functional Methods*, volume 1, pages 155–192. World Scientific, 1995.

[149] Javier Cerezo, José Zuniga, Alberto Requena, Francisco J. Avila Ferrer, and Fabrizio Santoro. Harmonic Models in Cartesian and Internal Coordinates to Simulate the Absorption Spectra of Carotenoids at Finite Temperatures. *J. Chem. Theory Comput.*, 9(11):4947–4958, 11 2013. URL: http://dx.doi.org/10.1021/ct4005849, doi:10.1021/ct4005849.

[150] G. Chaban, M. W. Schmidt, and M. S. Gordon. Approximate second order method for orbital optimization of scf and mcscf wavefunctions. *Theor. Chem. Acc.*, 97:88–95, 1997. doi:https://doi.org/10.1007/s002140050241.

[151] J.-D. Chai and M. Head-Gordon. *J. Chem. Phys.*, 128:084106, 2008.

[152] Jeng-Da Chai and Martin Head-Gordon. Long-range corrected double-hybrid density functionals. *J. Chem. Phys.*, 131(17):174105, 2009. URL: https://doi.org/10.1063/1.3244209, doi:10.1063/1.3244209.

[153] K. Chakarawet, M. Atanasov, J. Marbey, P. C. Bunting, F. Neese, S. Hill, and J. R. Long. Strong electronic and magnetic coupling in m-4 (m = ni, cu) clusters via direct orbital interactions between low-coordinate metal centers. *J. Am. Chem. Soc.*, 142(45):19161–19169, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000588273900023, doi:10.1021/jacs.0c08460.

[154] Uttam Chakraborty, Serhiy Demeshko, Franc Meyer, Christophe Rebreyend, Bas de Bruin, Mihail Atanasov, Frank Neese, Bernd Mühldorf, and Robert Wolf. Electronic Structure and Magnetic Anisotropy of an Unsaturated Cyclopentadienyl Iron(I) Complex with 15 Valence Electrons. *Angew. Chem. Int. Ed.*, 56(27):7995–7999, 06 2017. doi:10.1002/anie.201702454.

[155] G. K.-L. Chan. DMRG homepage. URL: http://www.princeton.edu/chemistry/chan/software/dmrg/.

[156] G. K.-L. Chan. An algorithm for large scale density matrix renormalization group calculations. *J. Chem. Phys.*, 120:3172, 2004. doi:https://doi.org/10.1063/1.1638734.

[157] G. K.-L. Chan and M. Head-Gordon. Highly correlated calculations with a polynomial cost algorithm: a study of the density matrix renormalization group. *J. Chem. Phys.*, 116:4462–4476, 2002. doi:https://doi.org/10.1063/1.1449459.

[158] G. K.-L. Chan and S. Sharma. *Ann. Rev. Phys. Chem.*, 62:465, 2011.

[159] H. C. Chang, Y. H. Lin, C. Werle, F. Neese, W. Z. Lee, E. Bill, and S. F. Ye. Conversion of a fleeting open-shell iron nitride into an iron nitrosyl. *Angew. Chem. Int. Ed.*, 58(49):17589–17593, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000491791300001, doi:10.1002/anie.201908689.

[160] H. C. Chang, B. Mondal, H. Y. Fang, F. Neese, E. Bill, and S. F. Ye. Electron paramagnetic resonance signature of tetragonal low spin Iron(V)-Nitrido and -Oxo complexes derived from the electronic structure analysis of heme and non-heme archetypes. *J. Am. Chem. Soc.*, 141(6):2421–2434, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459222100035, doi:10.1021/jacs.8b11429.

[161] A. Chantzis, J. K. Kowalska, D. Maganas, S. DeBeer, and F. Neese. Ab initio wave function-based determination of element specific shifts for the efficient calculation of x-ray absorption spectra of main group elements and first row transition metals. *J. Chem. Theory Comput.*, 14(7):3686–3702, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000438654500028, doi:10.1021/acs.jctc.8b00249.

[162] Koushik Chatterjee and Alexander Yu. Sokolov. Extended Second-Order Multireference Algebraic Diagrammatic Construction Theory for Charged Excitations. *J. Chem. Theory Comput.*, 16(10):6343–6357, 10 2020. URL: https://doi.org/10.1021/acs.jctc.0c00778, doi:10.1021/acs.jctc.0c00778.

[163] A. Chatzis, J. K. Kowalska, D. Maganas, S. DeBeer, and F. Neese. Ab initio wave function-based determination of element specific shifts for the efficient calculation of x-ray absorption spectra of main group elements and first row transition metals. *J. Chem. Theory Comput.*, 14(7):3686–3702, 2018. URL: https://pubs.acs.org/doi/10.1021/jacs.8b13313, doi:https://dx.doi.org/10.1021/acs.jctc.8b00249.

[164] Houxian Chen, Menglin Liu, and Tianying Yan. Molecular multipoles and (hyper)polarizabilities from the buckingham expansion: revisited. *Communications in Theoretical Physics*, 72(7):075503, jun 2020. URL: https://dx.doi.org/10.1088/1572-9494/ab8a0d, doi:10.1088/1572-9494/ab8a0d.

[165] Lan Cheng and Jürgen Gauss. Analytic energy gradients for the spin-free exact two-component theory using an exact block diagonalization for the one-electron Dirac Hamiltonian. *J. Chem. Phys.*, 135(8):084114, aug 2011. URL: http://aip.scitation.org/doi/10.1063/1.3624397, doi:10.1063/1.3624397.

[166] Lan Cheng and Jürgen Gauss. Analytic second derivatives for the spin-free exact two-component theory. *J. Chem. Phys.*, 135(24):244104, dec 2011. URL: https://doi.org/10.1063/1.3667202http://aip.scitation.org/doi/10.1063/1.3667202, doi:10.1063/1.3667202.

[167] Lan Cheng, Jürgen Gauss, and John F. Stanton. Treatment of scalar-relativistic effects on nuclear magnetic shieldings using a spin-free exact-two-component approach. *J. Chem. Phys.*, 139(5):054105, aug 2013. URL: https://doi.org/10.1063/1.4816130http://aip.scitation.org/doi/10.1063/1.4816130, doi:10.1063/1.4816130.

[168] L. F. Chibotaru and L. Ungur. Ab initio calculation of anisotropic magnetic properties of complexes. I. Unique definition of pseudospin Hamiltonians and their derivation. *J. Chem. Phys.*, 137:064112, 2012.

[169] Vijay Gopal Chilkuri, Serena DeBeer, and Frank Neese. Revisiting the Electronic Structure of FeS Monomers Using ab Initio Ligand Field Theory and the Angular Overlap Model. *Inorg. Chem.*, 56(17):10418–10436, 09 2017. doi:10.1021/acs.inorgchem.7b01371.

[170] Vijay Gopal Chilkuri, Serena DeBeer, and Frank Neese. Ligand Field Theory and Angular Overlap Model Based Analysis of the Electronic Structure of Homovalent Iron–Sulfur Dimers. *Inorg. Chem.*, 59(2):984–995, 01 2020. doi:10.1021/acs.inorgchem.9b00974.

[171] Vijay Gopal Chilkuri and Frank Neese. Comparison of many-particle representations for selected-CI I: A tree based approach. *J. Comput. Chem.*, 42(14):982–1005, 2021. doi:10.1002/jcc.26518.

[172] Vijay Gopal Chilkuri and Frank Neese. Comparison of Many-Particle Representations for Selected Configuration Interaction: II. Numerical Benchmark Calculations. *J. Chem. Theory Comput.*, 17(5):2868–2885, 05 2021. doi:10.1021/acs.jctc.1c00081.

[173] D. P. Chong and S. R. Langhoff. A modified coupled pair functional approach. *J. Chem. Phys.*, 84:5606–5610, 1986. doi:https://doi.org/10.1063/1.449920.

[174] Ove Christiansen, Poul Jø rgensen, and Christof Hättig. Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quantum Chem.*, 68:1–52, 1998. doi:https://doi.org/10.1002/(SICI)1097-461X(1998)68:1%3C1::AID-QUA1%3E3.0.CO;2-Z.

[175] D. W. Clack. Indo mo calculations for first row transition metal complexes. *Mol. Phys.*, 27:1513–1519, 1974. doi:https://doi.org/10.1080/00268977400101281.

[176] D. W. Clack, N. S. Hush, and J. R. Yandle. All\textendashvalence electron cndo calculations on transition metal complexes. *J. Chem. Phys.*, 57:3503, 1972. URL: https://pubs.aip.org/aip/jcp/article-abstract/57/8/3503/778222/, doi:https://doi.org/10.1063/1.1678785.

[177] D. W. Clack and W. Smith. Clack, d.w., smith, w. molecular orbital calculations on transition metal complexes. *Theor. Chim. Acta*, 36:87–92, 1974. doi:https://doi.org/10.1007/BF00554339.

[178] Aurora E. Clark and Ernest R. Davidson. P-Benzyne Derivatives That Have Exceptionally Small Singlet-Triplet Gaps and Even a Triplet Ground State. *J. Org. Chem.*, 68(9):3387–3396, 05 2003. URL: https://doi.org/10.1021/jo026824b (visited on 2020-10-20), doi:10.1021/jo026824b.

[179] T. Clark, J. Chandrasekhar, Spltznagel W. G., and P. v. R. Schleyer. Efficient diffuse function-augmented basis sets for anion calculations. iii. the 3-21+g basis set for first-row elements, li\textendashf . *J. Comput. Chem.*, 4:294, 1983. doi:https://doi.org/10.1002/jcc.540040303.

[180] P. Claverie, J. P. Daudey, J. Langlet, B. Pullman, D. Plazzola, and M. J. Huron. Studies of solvent effects. 1. discrete, continuum, and discrete –continuum models and their comparison for some simple cases: ammonium(1+) ion, methanol, and substituted ammonium(1+) ion. *J. Phys. Chem.*, 82:405–418, 1978. doi:https://doi.org/10.1021/j100493a008.

[181] E. Clementi and D. Raimondi. *IBM Res. Note*, 1963.

[182] L. R. Collins, M. van Gastel, F. Neese, and A. Furstner. Enhanced electrophilicity of heterobimetallic birh paddlewheel carbene complexes: A combined experimental, spectroscopic, and computational study. *J. Am. Chem. Soc.*, 140(40):13042–13055, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000447354800056, doi:10.1021/jacs.8b08384.

[183] M. G. Cory and M. C. Zerner. Metal-ligand exchange coupling in transition-metal complexes. *Chem. Rev.*, 91:813, 1991. doi:https://doi.org/10.1021/cr00005a009.

[184] Hector H. Corzo, Ali Abou Taka, Aurora Pribram-Jones, and Hrant P. Hratchian. Using projection operators with maximum overlap methods to simplify challenging self-consistent field optimization. *Journal of Computational Chemistry*, 43(6):382–390, 2022. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.26797, doi:10.1002/jcc.26797.

[185] E. Coulaud, J.-P. Malrieu, N. Guihéry, and N. Ferré. Additive decomposition of the physical components of the magnetic coupling from broken symmetry density functional theory calculations. *J. Chem. Theory Comput.*, 9(8):3429–3436, 2013. doi:10.1021/ct400305h.

[186] Jeanne Crassous. *Circularly Polarized Luminescence in Helicene and Helicenoid Derivatives*, pages 53–97. Springer Singapore, Singapore, 2020. URL: https://doi.org/10.1007/978-981-15-2309-0_4, doi:10.1007/978-981-15-2309-0_4.

[187] D. Cremer. In P. v. R. Schleyer, editor, *Encyclopedia of Computational Chemistry*, pages 1706. John Wiley and Sons, 1998.

[188] L. A. Curtiss, M. P. McGrath, J.-P. Blandeau, N. E. Davis, R. C. Binning Jr., and L. Radom. Extension of gaussian –2 theory to molecules containing third –row atoms ga –kr. *J. Chem. Phys.*, 103:6104–6113, 1995. doi:https://doi.org/10.1063/1.470438.

[189] Larry A. Curtiss, Krishnan Raghavachari, and John A. Pople. Gaussian-2 theory using reduced møller-plesset orders. *J. Chem. Phys.*, 98(2):1293–1298, 1993. URL: https://ui.adsabs.harvard.edu/abs/1993JChPh..98.1293C/abstract, doi:10.1063/1.464297.

[190] C. Daniel, L. Gonzalez, and F. Neese. Quantum theory: The challenge of transition metal complexes. *Phys. Chem. Chem. Phys.*, 23(4):2533–2534, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000614634000001, doi:10.1039/d0cp90278k.

[191] D. Datta, S. Kossmann, and F. Neese. Analytic energy derivatives for the calculation of the first-order molecular properties using the domain-based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.*, 175:114101, 2016. doi:https://doi.org/10.1063/1.4962369.

[192] D. Datta, M. Saitow, B. Sandhofer, and F. Neese. Fe-57 mossbauer parameters from domain based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.*, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000596027300001, doi:10.1063/5.0022215.

[193] Dipayan Datta, Liguo Kong, and Marcel Nooijen. A state-specific partially internally contracted multireference coupled cluster approach. *J. Chem. Phys.*, 134(21):214116, 2011. doi:https://doi.org/10.1063/1.3592494.

[194] Dipayan Datta and Marcel Nooijen. Multireference equation-of-motion coupled cluster theory. *J. Chem. Phys.*, 137(20):204107, 2012. doi:https://doi.org/10.1063/1.4766361.

[195] G. David, F. Wennmohs, F. Neese, and N. Ferre. Chemical tuning of magnetic exchange couplings using broken-symmetry density functional theory. *Inorg. Chem.*, 57(20):12769–12776, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000447680400039, doi:10.1021/acs.inorgchem.8b01970.

[196] E. R. Davidson and A. E. Clark. Analysis of wave functions for open-shell molecules. *Phys. Chem. Chem. Phys.*, 9:1881–1894, 2007. doi:https://doi.org/10.1039/B616481C.

[197] Bernardo de Souza, Giliandro Farias, Frank Neese, and Robert Izsak. Efficient simulation of overtones and combination bands in Resonant Raman spectra. *J. Chem. Phys.*, 150(21):accepted – still waiting for publication, 2019. URL: https://aip.scitation.org/doi/10.1063/1.5099247, doi:10.1063/1.5099247.

[198] Bernardo de Souza, Giliandro Farias, Frank Neese, and Robert Izsak. Predicting phosphorescence rates of light organic molecules using time-dependent density functional theory and the path integral approach to dynamics. *J. Chem. Theory Comput.*, 15(3):1896–1904, 2019. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.8b00841, doi:10.1021/acs.jctc.8b00841.

[199] Bernardo de Souza, Frank Neese, and Robert Izsak. On the theoretical prediction of fluorescence rates from first principles using the path integral approach. *J. Chem. Phys.*, 148(3):034104, 2018. URL: http://aip.scitation.org/doi/abs/10.1063/1.5010895 (visited on 2018-01-31), doi:10.1063/1.5010895.

[200] S. DeBeer-George, T. Petrenko, and F. Neese. *J. Phys. Chem. A*, 112:12936, 2008.

[201] S. DeBeer-George, T. Petrenko, and F. Neese. *Inorg. Chim. Acta*, 361:965, 2008.

[202] O. Demel and J. Pittner. *J. Chem. Phys.*, 124:144112, 2006.

[203] Ondrej Demel, Dipayan Datta, and Marcel Nooijen. Additional global internal contraction in variations of multireference equation of motion coupled cluster theory. *J. Chem. Phys.*, 138(13):134108, 2013.

[204] J. des Cloizeaux. Extension d'une formule de lagrange \a des probl\emes de valeurs propres. *Nucl. Phys.*, 20:321–346, 1960. doi:https://doi.org/10.1016/0029-5582(60)90177-2.

[205] M. J. S. Dewar, J. A. Hashmall, and C. G. Venier. *J. Am. Chem. Soc.*, 90:1953, 1968.

[206] M. J. S. Dewar and W. Thiel. *Theor. Chim. Acta*, 46:89, 1977.

[207] M. J. S. Dewar and W. Thiel. *J. Am. Chem. Soc.*, 99:4899, 1977.

[208] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. P. Stewart. *J. Am. Chem. Soc.*, 107:3902, 1985.

[209] J. D. Dill and J. A. Pople. *J. Chem. Phys.*, 62:2921, 1975.

[210] P. A. M. Dirac. *Proc. Camb. Phil. Soc.*, 26:376, 1930.

[211] R. Ditchfield. *J. Chem. Phys.*, 56:5688, 1972.

[212] A. Dittmer, G. L. Stoychev, D. Maganas, A. A. Auer, and F. Neese. Computation of nmr shielding constants for solids using an embedded cluster approach with dft, double-hybrid dft, and mp2. *J. Chem. Theory Comput.*, 16(11):6950–6967, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000592392800018, doi:10.1021/acs.jctc.0c00067.

[213] Anneke Dittmer, Róbert Izsák, Frank Neese, and Dimitrios Maganas. Accurate band gap predictions of semiconductors in the framework of the similarity transformed equation of motion coupled cluster theory. *Inorg. Chem.*, 58(14):9303–9315, 2019. URL: https://doi.org/10.1021/acs.inorgchem.9b00994, doi:10.1021/acs.inorgchem.9b00994.

[214] K. D. Dobbs and W. J. Hehre. *J. Comput. Chem.*, 7:359, 1986.

[215] K. D. Dobbs and W. J. Hehre. *J. Comput. Chem.*, 8:861 & 880, 1987.

[216] John F Dobson. Alternative expressions for the Fermi hole curvature. *J. Chem. Phys.*, 98(11):8870–8872, 06 1993. URL: https://doi.org/10.1063/1.464444, doi:10.1063/1.464444.

[217] M. Dolg, P. Fulde, W. Küchle, C.-S. Neumann, and H. Stoll. *J. Chem. Phys.*, 94:3011–3017, 1991.

[218] M. Dolg, H. Stoll, H.-J. Flad, and H. Preuss. *J. Chem. Phys.*, 97:1162–1173, 1992.

[219] M. Dolg, H. Stoll, and H. Preuss. *J. Chem. Phys.*, 90:1730–1734, 1989.

[220] M. Dolg, H. Stoll, and H. Preuss. *Theor. Chim. Acta*, 85:441–450, 1993.

[221] M. Dolg, H. Stoll, H. Preuss, and R. M. Pitzer. *J. Phys. Chem.*, 97:5852–5859, 1993.

[222] M. Dolg, H. Stoll, A. Savin, and H. Preuss. *Theor. Chim. Acta*, 75:173–194, 1989.

[223] M. Dolg, U. Wedig, H. Stoll, and H. Preuss. *J. Chem. Phys.*, 86:866–872, 1987.

[224] A. Domingo, M.-A. Carvajal, C. de Graaf, K. Sivalingam, F. Neese, and C. Angeli. *Theor. Chem. Acc.*, 131(9):1264, 2012.

[225] B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin. *J. Chem. Phys.*, 71:3396, 1979.

[226] T. H. Dunning Jr. *J. Chem. Phys.*, 90:1007, 1989.

[227] M. Dupuis and H. F. King. Molecular symmetry and closed-shell scf calculations. i. *Int. J. Quantum Chem.*, XI(4):613–625, 1977. URL: https://doi.org/10.1002/qua.560110408.

[228] F. Duschisnky. *Acta Physicochim. URSS*, 7:551, 1937.

[229] A. K. Dutta, F. Neese, and R. Izsak. Accelerating the coupled-cluster singles and doubles method using the chain-of-sphere approximation. *Mol. Phys.*, 116(11):1428–1434, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000431017000002, doi:10.1080/00268976.2017.1416201.

[230] A. K. Dutta, M. Nooijen, F. Neese, and R. Izsak. Exploring the accuracy of a low scaling similarity transformed equation of motion method for vertical excitation energies. *J. Chem. Theory Comput.*, 14(1):72–91, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000419998300008, doi:10.1021/acs.jctc.7b00802.

[231] A. K. Dutta, M. Saitow, B. Demoulin, F. Neese, and R. Izsak. A domain-based local pair natural orbital implementation of the equation of motion coupled cluster method for electron attached states. *J. Chem. Phys.*, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000466698700028, doi:10.1063/1.5089637.

[232] A. K. Dutta, M. Saitow, C. Riplinger, F. Neese, and R. Izsak. A nearlinear scaling equation of motion coupled cluster method for ionized states. *J. Chem. Phys.*, 148(24):13, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000437190300049, doi:10.1063/1.5029470.

[233] Achintya Kumar Dutta, Frank Neese, and Róbert Izsák. Speeding up equation of motion coupled cluster theory with the chain of spheres approximation. *J. Chem. Phys.*, 144(3):034102, 2016.

[234] Achintya Kumar Dutta, Frank Neese, and Róbert Izsák. Towards a pair natural orbital coupled cluster method for excited states. *J. Chem. Phys.*, 145(3):034102, 2016.

[235] Achintya Kumar Dutta, Frank Neese, and Róbert Izsák. A simple scheme for calculating approximate transition moments within the equation of motion expectation value formalism. *J. Chem. Phys.*, 146(21):214111, 2017.

[236] Achintya Kumar Dutta, Marcel Nooijen, Frank Neese, and Róbert Izsák. Towards a pair natural orbital coupled cluster method for excited states. *J. Chem. Phys.*, 2017.

[237] Achintya Kumar Dutta, Marcel Nooijen, Frank Neese, and Róbert Izsák. Automatic active space selection for the similarity transformed equations of motion coupled cluster method. *J. Chem. Phys.*, 146(7):074103, 2017.

[238] K. G. Dyall. *J. Chem. Phys.*, 102:4909–4918, 1995.

[239] Anatoly Y. Dymarsky and Konstantin N. Kudin. Computation of the pseudorotation matrix to satisfy the Eckart axis conditions. *J. Chem. Phys.*, 122(12):124103, 03 2005. URL: http://scitation.aip.org/content/aip/journal/jcp/122/12/10.1063/1.1864872, doi:10.1063/1.1864872.

[240] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66(5):052301, 08 2002. URL: https://link.aps.org/doi/10.1103/PhysRevB.66.052301, arXiv:0205527 [cond-mat], doi:10.1103/PhysRevB.66.052301.

[241] F. Eckert, P. Pulay, and H. J. Werner. *J. Comput. Chem.*, 12:1473, 1997.

[242] M. Eden and M. H. Levitt. *J. Magn. Res.*, 132:220, 1998.

[243] W. D. Edwards and M. C. Zerner. *Theor. Chim. Acta*, 72:347, 1987.

[244] Sebastian Ehlert, Uwe Huniar, Jinliang Ning, James W. Furness, Jianwei Sun, Aaron D. Kaplan, John P. Perdew, and Jan Gerit Brandenburg. r2SCAN-D4: Dispersion corrected meta-generalized gradient approximation for general chemical applications. *J. Chem. Phys.*, 154(6):061101, 2021. URL: https://doi.org/10.1063/5.0041008, arXiv:https://doi.org/10.1063/5.0041008, doi:10.1063/5.0041008.

[245] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs. *Chem. Phys. Lett.*, 240:283, 1995.

[246] K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs. *Theor. Chem. Acc.*, 97:119, 1997.

[247] Dennis M. Elking, Lalith Perera, Robert Duke, Thomas Darden, and Lee G. Pedersen. A finite field method for calculating molecular polarizability tensors for arbitrary multipole rank. *Journal of Computational Chemistry*, 32(15):3283–3295, 2011. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21914, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.21914, doi:https://doi.org/10.1002/jcc.21914.

[248] M. Ernzerhof. In D. P. Joubert, editor, *Density Functionals: Theory and Applications*. Springer Verlag, 1998.

[249] Francesco A. Evangelista and Jürgen Gauss. An orbital-invariant internally contracted multireference coupled cluster approach. *J. Chem. Phys.*, 134(11):114102, 03 2011. doi:10.1063/1.3559149.

[250] W. Förner, J. Ladik, P. Otto, and J. Čížek. Coupled-cluster studies. II. The role of localization in correlation calculations on extended systems. *Chem. Phys.*, 97:251–262, 1985. doi:https://doi.org/10.1016/0301-0104(85)87035-X.

[251] Molecular Science Computing Facility. Extensible computational chemistry environment basis set database. 2000. URL: http://www.emsl.pnl.gov:2080/forms/basisform.html.

[252] Shervin Fatehi, Ethan Alguire, Yihan Shao, and Joseph E. Subotnik. Analytic derivative couplings between configuration-interaction-singles states with built-in electron-translation factors for translational invariance. *J. Chem. Phys.*, 135(23):234105, 12 2011. URL: https://aip.scitation.org/doi/10.1063/1.3665031 (visited on 2020-06-08), doi:10.1063/1.3665031.

[253] Ignacio Fdez. Galván, Morgane Vacher, Ali Alavi, Celestino Angeli, Francesco Aquilante, Jochen Autschbach, Jie J. Bao, Sergey I. Bokarev, Nikolay A. Bogdanov, Rebecca K. Carlson, Liviu F. Chibotaru, Joel Creutzberg, Nike Dattani, Mickaël G. Delcey, Sijia S. Dong, Andreas Dreuw, Leon Freitag, Luis Manuel Frutos, Laura Gagliardi, Frédéric Gendron, Angelo Giussani, Leticia González, Gilbert Grell, Meiyuan Guo, Chad E. Hoyer, Marcus Johansson, Sebastian Keller, Stefan Knecht, Goran Kovačević, Erik Källman, Giovanni Li Manni, Marcus Lundberg, Yingjin Ma, Sebastian Mai, João Pedro Malhado, Per Åke Malmqvist, Philipp Marquetand, Stefanie A. Mewes, Jesper Norell, Massimo Olivucci, Markus Oppel, Quan Manh Phung, Kristine Pierloot, Felix Plasser, Markus Reiher, Andrew M. Sand, Igor Schapiro, Prachi Sharma, Christopher J. Stein, Lasse Kragh Sørensen, Donald G. Truhlar, Mihkel Ugandi, Liviu Ungur, Alessio Valentini, Steven Vancoillie, Valera Veryazov, Oskar Weser, Tomasz A. Wesołowski, Per-Olof Widmark, Sebastian Wouters, Alexander Zech, J. Patrick Zobel, and Roland Lindh. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.*, 15(11):5925–5964, 2019.

[254] J. Fernandez-Rico, J. M. Garcia de la Vega, and J. I. Fernandez Alonso. *J. Comput. Chem.*, 1:33, 1983.

[255] J. Fernandez-Rico, R. Lopez, A. Aguado, I. Ema, and G. Ramirez. *J. Comput. Chem.*, 11:1284, 1998.

[256] J. Fernandez-Rico, M. Paniagua, J. I. Fernandez Alonso, and P. Fantucci. *J. Comput. Chem.*, 1:41, 1983.

[257] N. Ferre, N. Guihery, and J.-P. Malrieu. Spin decontamination of broken-symmetry density functional theory calculations: deeper insight and new formulations. *Phys. Chem. Chem. Phys.*, 17(22):14375–14382, 2015. doi:10.1039/C4CP05531D.

[258] M. Feyereisen, G. Fitzerald, and A. Komornicki. *Chem. Phys. Lett.*, 208:359, 1993.

[259] M. Fiedler. *Czech. Math. J.*, 23:298, 1973.

[260] M. Fiedler. *Czech. Math. J.*, 25:619, 1975.

[261] D. Figgen, K. A. Peterson, M. Dolg, and H. Stoll. *J. Chem. Phys.*, 130:164108, 2009.

[262] D. Figgen, G. Rauhut, M. Dolg, and H. Stoll. *Chem. Phys.*, 311:227, 2005.

[263] J. Finley, P.-A. Malmqvist, B. O. Roos, and L. Serrano-Andres. *Chem. Phys. Lett.*, 288:299, 1998.

[264] T. H. Fischer and J. Almlöf. *J. Phys. Chem.*, 96:9768, 1992.

[265] D. Flaig, M. Maurer, M. Hanni, K. Braunger, L. Kick, M. Thubauville, and C. Ochsenfeld. *J. Chem. Theory Comput.*, 10:572, 2014.

[266] R. Flores-Moreno, R. J. Alvares-Mendez, A. Vela, and A. M. Köster. *J. Comput. Chem.*, 27:1009, 2006.

[267] B. M. Floser, Y. Guo, C. Riplinger, F. Tuczek, and F. Neese. Detailed pair natural orbital-based coupled cluster studies of spin crossover energetics. *J. Chem. Theory Comput.*, 16(4):2224–2235, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000526313000020, doi:10.1021/acs.jctc.9b01109.

[268] Nicolás Foglia, Bernardo De Souza, Dimitrios Maganas, and Frank Neese. Including vibrational effects in magnetic circular dichroism spectrum calculations in the framework of excited state dynamics. *J. Chem. Phys.*, 158(15):154108, 04 2023. URL: https://doi.org/10.1063/5.0144845, doi:10.1063/5.0144845.

[269] Nicolás O. Foglia, Dimitrios Maganas, and Frank Neese. Going beyond the electric-dipole approximation in the calculation of absorption and (magnetic) circular dichroism spectra including scalar relativistic and spin–orbit coupling effects. *J. Chem. Phys.*, 157(8):084120, 2022. URL: https://doi.org/10.1063/5.0094709, arXiv:https://doi.org/10.1063/5.0094709, doi:10.1063/5.0094709.

[270] Niclas Forsberg and Per-Å ke Malmqvist. Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett.*, 274:196–204, 08 1997. doi:https://doi.org/10.1016/S0009-2614(97)00669-6.

[271] M. M. Francl, W. J. Petro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, and J. A. Pople. *J. Chem. Phys.*, 77:3654, 1982.

[272] Yannick J Franzke and Florian Weigend. NMR Shielding Tensors and Chemical Shifts in Scalar-Relativistic Local Exact Two-Component Theory. *J. Chem. Theory Comput.*, 15(2):1028–1043, feb 2019. URL: https://pubs.acs.org/sharingguidelineshttps://pubs.acs.org/doi/10.1021/acs.jctc.8b01084, doi:10.1021/acs.jctc.8b01084.

[273] Yannick J. Franzke, Fabian Mack, and Florian Weigend. NMR Indirect Spin–Spin Coupling Constants in a Modern Quasi-Relativistic Density Functional Framework. *J. Chem. Theory Comput.*, 17(7):3974–3994, jul 2021. URL: https://doi.org/10.1021/acs.jctc.1c00167https://pubs.acs.org/doi/10.1021/acs.jctc.1c00167, doi:10.1021/acs.jctc.1c00167.

[274] Yannick J. Franzke, Nils Middendorf, and Florian Weigend. Efficient implementation of one- and two-component analytical energy gradients in exact two-component theory. *J. Chem. Phys.*, 148(10):104110, mar 2018. URL: https://doi.org/10.1063/1.5022153http://aip.scitation.org/doi/10.1063/1.5022153, doi:10.1063/1.5022153.

[275] Yannick J. Franzke, Robert Treß, Tobias M. Pazdera, and Florian Weigend. Error-consistent segmented contracted all-electron relativistic basis sets of double- and triple-zeta quality for NMR shielding constants. *Phys. Chem. Chem. Phys.*, 21(30):16658–16664, 2019. URL: http://dx.doi.org/10.1039/C9CP02382H, doi:10.1039/C9CP02382H.

[276] Yannick J. Franzke and Jason M. Yu. Hyperfine Coupling Constants in Local Exact Two-Component Theory. *J. Chem. Theory Comput.*, 18(1):323–343, jan 2022. URL: https://pubs.acs.org/doi/10.1021/acs.jctc.1c01027, doi:10.1021/acs.jctc.1c01027.

[277] M. J. Frisch, J. A. Pople, and J. S. Binkley. *J. Chem. Phys.*, 80:3265, 1984.

[278] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople. *Gaussian 98, Revision A.8*. Gaussian, Inc., Pittsburgh PA, 1998.

[279] P. Fuentealba, H. Preuss, H. Stoll, and L. von Szentpaly. *Chem. Phys. Lett.*, 89:418–422, 1982.

[280] P. Fuentealba, H. Stoll, L. von Szentpaly, P. Schwerdtfeger, and H. Preuss. *J. Phys. B: At. Mol. Opt. Phys.*, 16:L323, 1983.

[281] P. Fuentealba, L. von Szentpaly, H. Preuss, and H. Stoll. *J. Phys. B: At. Mol. Opt. Phys.*, 18:1287, 1985.

[282] James W Furness, Aaron D Kaplan, Jinliang Ning, John P Perdew, and Jianwei Sun. Accurate and numerically efficient r^{2}SCAN meta-generalized gradient approximation. *J. Phys. Chem. Lett.*, 11(19):8208–8215, 2020. URL: https://doi.org/10.1021/acs.jpclett.0c02405.

[283] D. Ganyushin and F. Neese. *J. Chem. Phys.*, 128:114117, 2008.

[284] D. Ganyushin and F. Neese. *J. Chem. Phys.*, 138:104113, 2013.

[285] M. Garcia-Ratés, U. Becker, and F. Neese. Implicit solvation in domain based pair natural orbital coupled cluster (dlpno-ccsd) theory. *J. Comput. Chem.*, 42(27):1959–1973, 2021. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.26726, doi:10.1002/jcc.26726.

[286] M. Garcia-Ratés and F. Neese. Efficient implementation of the analytical second derivatives of hartree-fock and hybrid dft energies within the framework of the conductor-like polarizable continuum model. *J. Comput. Chem.*, 40(20):1816–1828, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000470923100002, doi:10.1002/jcc.25833.

[287] M. Garcia-Ratés and F. Neese. Effect of the solute cavity on the solvation energy and its derivatives within the framework of the gaussian charge scheme. *J. Comput. Chem.*, 41(9):922–939, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000504862400001, doi:10.1002/jcc.26139.

[288] T. Gatzenmeier, M. Turberg, D. Yepes, Y. W. Xie, F. Neese, G. Bistoni, and B. List. Scalable and highly diastereo- and enantioselective catalytic diels-alder reaction of alpha,beta-Unsaturated methyl esters. *J. Am. Chem. Soc.*, 140(40):12671–12676, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000447354800004, doi:10.1021/jacs.8b07092.

[289] J. Gauss. Molecular properties. In J. Grotendorst, editor, *Modern Methods and Algorithms of Quantum Chemistry*, volume 3, 541–592. John von Neumann Institute for Computing, NIC Series, 2000.

[290] Jürgen Gauss, Kenneth Ruud, and Trygve Helgaker. Perturbation-dependent atomic orbitals for the calculation of spin-rotation constants and rotational g tensors. *The Journal of Chemical Physics*, 105(7):2804–2812, aug 1996. URL: http://aip.scitation.org/doi/10.1063/1.472143, doi:10.1063/1.472143.

[291] R. J. Gdanitz. *Int. J. Quant. Chem.*, 85:281, 2001.

[292] R. J. Gdanitz and R. Ahlrichs. *Chem. Phys. Lett.*, 1988:413, 0143.

[293] Thomas Gerlach, Simon Müller, Andrés González de Castilla, and Irina Smirnova. An open source COSMO-RS implementation and parameterization supporting the efficient implementation of multiple segment descriptors. *Fluid Phase Equil.*, 560:113472, 2022.

[294] Reza Ghafarian Shirazi, Frank Neese, Dimitrios A Pantazis, and Giovanni Bistoni. Physical nature of differential spin-state stabilization of carbenes by hydrogen and halogen bonding: A domain-based pair natural orbital coupled cluster study. *J. Phys. Chem. A*, 123(24):5081–5090, 2019.

[295] Giovanni Ghigo, Björn O. Roos, and Per-Å ke Malmqvist. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chem. Phys. Lett.*, 396:142–149, 09 2004. doi:https://doi.org/10.1016/j.cplett.2004.08.032.

[296] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan. *J. Chem. Phys.*, 128:144117, 2008.

[297] Soumen Ghosh, Achintya Kumar Dutta, Bernardo de Souza, Romain Berraud-Pache, and Róbert Izsák. A new density for transition properties within the similarity transformed equation of motion approach. *Mol. Phys.*, 118(19-20):e1818858, 10 2020. URL: https://doi.org/10.1080/00268976.2020.1818858 (visited on 2021-03-01), doi:10.1080/00268976.2020.1818858.

[298] Soumen Ghosh, Achintya Kumar Dutta, Bernardo de Souza, Romain Berraud-Pache, and Róbert Izsák. A new density for transition properties within the similarity transformed equation of motion approach. *Mol. Phys.*, 0(0):e1818858, 2020. URL: https://doi.org/10.1080/00268976.2020.1818858, doi:10.1080/00268976.2020.1818858.

[299] Andrew T. B. Gilbert, Nicholas A. Besley, and Peter M. W. Gill. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM). *The Journal of Physical Chemistry A*, 112(50):13164–13171, December 2008. Publisher: American Chemical Society. URL: https://doi.org/10.1021/jp801738f, doi:10.1021/jp801738f.

[300] P. M. W. Gill. *Mol. Phys.*, 89:433, 1996.

[301] P. M. W. Gill, B. G. Johnson, and J. A. Pople. *Chem. Phys. Lett.*, 209:506, 1993.

[302] M. K. Gilson and K. K. Irikura. *J. Phys. Chem.*, pages 16304–16317, 2010. doi:10.1021/jp110434s.

[303] A. P. Ginsberg. *J. Am. Chem. Soc.*, 102:111, 1980.

[304] L. Giordano, G. Pacchioni, T. Bredow, and J. F. Sanz. *Surf. Sci.*, 471:21, 2001.

[305] N. Godbout, D. R. Salahub, J. Andzelm, and E. Wimmer. *Can. J. Chem.*, 70:560, 1992.

[306] Stefan Goedecker. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of Chemical Physics*, 120(21):9911–9917, June 2004. Publisher: American Institute of Physics. URL: https://aip.scitation.org/doi/10.1063/1.1724816 (visited on 2022-04-22), doi:10.1063/1.1724816.

[307] L. Goerigk and S. Grimme. *J. Chem. Theory Comput.*, 6:107, 2010.

[308] L. Goerigk and S. Grimme. *J. Chem. Theory Comput.*, 7:291–309, 2011.

[309] L. Goerigk and S. Grimme. *Phys. Chem. Chem. Phys.*, 13:6670, 2011.

[310] L. Moellmann Goerigk and S. Grimme. *Phys. Chem. Chem. Phys.*, 11:4611, 2009.

[311] Lars Goerigk and Stefan Grimme. *J. Chem. Phys.*, 132:184103, 2010.

[312] Lars Goerigk, Andreas Hansen, Christoph Bauer, Stephan Ehrlich, Asim Najibi, and Stefan Grimme. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.*, 19(48):32184–32215, 2017. doi:10.1039/C7CP04913G.

[313] M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, and W. J. Hehre. *J. Am. Chem. Soc.*, 104:2797, 1983.

[314] H. C. Gottschalk, A. Poblotzki, M. Fatima, D. A. Obenchain, C. Perez, J. Antony, A. A. Auer, L. Baptista, D. M. Benoit, G. Bistoni, F. Bohle, R. Dahmani, D. Firaha, S. Grimme, A. Hansen, M. E. Harding, M. Hochlaf, C. Holzer, G. Jansen, W. Klopper, W. A. Kopp, L. C. Kroger, K. Leonhard, M. M. Al-Mogren, H. Mouhib, F. Neese, X. N. Pereira, M. Prakash, I. S. Ulusoy, R. A. Mata, M. A. Suhm, and M. Schnell. The first microsolvation step for furans: New experiments and benchmarking strategies. *J. Chem. Phys.*, 152(16):17, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000531832400001, doi:10.1063/5.0004465.

[315] H. C. Gottschalk, A. Poblotzki, M. A. Suhm, M. M. Al-Mogren, J. Antony, A. A. Auer, L. Baptista, D. M. Benoit, G. Bistoni, F. Bohle, R. Dahmani, D. Firaha, S. Grimme, A. Hansen, M. E. Harding, M. Hochlaf, C. Holzer, G. Jansen, W. Klopper, W. A. Kopp, L. C. Kroger, K. Leonhard, H. Mouhib, F. Neese, M. N. Pereira, I. S. Ulusoy, A. Wuttke, and R. A. Mata. The furan microsolvation blind challenge for quantum chemical methods: First steps. *J. Chem. Phys.*, 148(1):13, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000419394500013, doi:10.1063/1.5009011.

[316] M. Grüning, O. V. Gritsenko, S. J. A. Gisbergen, and E. J. Baerends. *J. Chem. Phys.*, 114:652, 2001.

[317] L Greengard and V Rokhlin. A fast algorithm for particle simulations. *Journal of Computational Physics*, 73(2):325–348, 1987. URL: https://www.sciencedirect.com/science/article/pii/0021999187901409, doi:https://doi.org/10.1016/0021-9991(87)90140-9.

[318] S. Grimme. *J. Chem. Phys.*, 118:9095–9102, 2003.

[319] S. Grimme. *J. Comput. Chem.*, 25:1463, 2004.

[320] S. Grimme. *J. Chem. Phys.*, 124:034108, 2006.

[321] S. Grimme. *J. Comput. Chem.*, 27:1787, 2006.

[322] S. Grimme. *Chem. Eur. J.*, 18:9955–9964, 2012.

[323] S. Grimme. *J. Chem. Phys.*, 138:244104, 2013.

[324] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. *J. Chem. Phys.*, 132:154104, 2010.

[325] S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen. *J. Chem. Phys.*, 143:054107, 2015.

[326] S. Grimme, S. Ehrlich, and L. Goerigk. *J. Comput. Chem.*, 32:1456, 2011.

[327] S. Grimme and A. Hansen. A practicable real-space measure and visualization of static electron-correlation effects. *Angew. Chem. Int. Ed.*, 54:12308–12313, 2015.

[328] S. Grimme and F. Neese. *J. Phys. Chem.*, 127:154116, 2007.

[329] Stefan Grimme. AK Grimme homepage. URL: http://www.thch.uni-bonn.de/tc/grimme.

[330] Stefan Grimme. Exploration of chemical compound, conformer, and reaction space with meta-dynamics simulations based on tight-binding quantum chemical calculations. *J. Chem. Theory Comput.*, 15(5):2847–2862, 2019. doi:10.1021/acs.jctc.9b00143.

[331] Stefan Grimme, Christoph Bannwarth, Sebastian Dohm, Andreas Hansen, Jana Pisarek, Philipp Pracht, Jakob Seibert, and Frank Neese. Fully automated quantum-chemistry-based computation of Spin–Spin-Coupled nuclear magnetic resonance spectra. *Angew. Chem. Int. Ed.*, 56(46):14763–14769, 2017. doi:10.1002/anie.201708266.

[332] Stefan Grimme, Christoph Bannwarth, and Philip Shushkov. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-Block elements (z = 1-86). *J. Chem. Theory Comput.*, 13(5):1989–2009, 2017. doi:10.1021/acs.jctc.7b00118.

[333] Stefan Grimme, Andreas Hansen, Sebastian Ehlert, and Jan-Michael Mewes. r2SCAN-3c: A "Swiss army knife" composite electronic-structure method. *J. Chem. Phys.*, 154(6):064103, 2021. URL: https://doi.org/10.1063/5.0040021, arXiv:https://doi.org/10.1063/5.0040021, doi:10.1063/5.0040021.

[334] Quantum Chemistry Group. TurboMole basis sets. ftp.chemie.uni-karlsruhe.de/pub, since 1988.

[335] M. F. Guest and V. R. Saunders. *Mol. Phys.*, 28:819, 1974.

[336] Sheng Guo, Mark A. Watson, Weifeng Hu, Qiming Sun, and Garnet Kin-Lic Chan. N-Electron Valence State Perturbation Theory Based on a Density Matrix Renormalization Group Reference Function, with Applications to the Chromium Dimer and a Trimer Model of Poly(p-Phenylenevinylene). *J. Chem. Theory Comput.*, 12(4):1583–1591, 04 2016. doi:10.1021/acs.jctc.5b01225.

[337] Y. Guo, W. Li, and S. Li. Improved cluster-in-molecule local correlation approach for electron correlation calculation of large systems. *J. Phys. Chem. A*, 118(39):8996–9004, 2014. doi:https://doi.org/10.1021/jp501976x.

[338] Y. Guo, F. Pavošević, K. Sivalingam, U. Becker, E. F. Valeev, and F. Neese. Sparsemaps—a systematic infrastructure for reduced-scaling electronic structure methods. vi. linear-scaling explicitly correlated n-electron valence state perturbation theory with pair natural orbital. *J. Chem. Phys.*, 2023.

[339] Y. Guo, C. Riplinger, D. G. Liakos, U. Becker, M. Saitow, and F. Neese. Linear scaling perturbative triples correction approximations for open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory dlpno-ccsd(t-0/t). *J. Chem. Phys.*, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000519813300005, doi:10.1063/1.5127550.

[340] Yang Guo, Ute Becker, and Frank Neese. Comparison and combination of "direct" and fragment based local correlation methods: Cluster in molecules and domain based local pair natural orbital perturbation and coupled cluster theories. *J. Chem. Phys.*, 148(12):124117, 2018. doi:10.1063/1.5021898.

[341] Yang Guo, Christoph Riplinger, Ute Becker, Dimitrios G. Liakos, Yury Minenkov, Luigi Cavallo, and Frank Neese. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.*, 148(1):011101, 2018. doi:10.1063/1.5011798.

[342] Yang Guo, Kantharuban Sivalingam, Christian Kollmar, and Frank Neese. Approximations of density matrices in N-electron valence state second-order perturbation theory (NEVPT2). II. The full rank NEVPT2 (FR-NEVPT2) formulation. *J. Chem. Phys.*, 154(21):214113, 06 2021. doi:10.1063/5.0051218.

[343] Yang Guo, Kantharuban Sivalingam, and Frank Neese. Approximations of density matrices in N-electron valence state second-order perturbation theory (NEVPT2). I. Revisiting the NEVPT2 construction. *J. Chem. Phys.*, 154(21):214111, 06 2021. doi:10.1063/5.0051211.

[344] Yang Guo, Kantharuban Sivalingam, Edward F. Valeev, and Frank Neese. SparseMaps—A systematic infrastructure for reduced-scaling electronic structure methods. III. Linear-scaling multireference domain-based pair natural orbital N-electron valence perturbation theory. *J. Chem. Phys.*, 144(9):094111, 03 2016. doi:10.1063/1.4942769.

[345] Yang Guo, Kantharuban Sivalingam, Edward F. Valeev, and Frank Neese. Explicitly correlated N-electron valence state perturbation theory (NEVPT2-F12). *J. Chem. Phys.*, 147(6):064110, 08 2017. doi:10.1063/1.4996560.

[346] M. Häser and R. Ahlrichs. *J. Comput. Chem.*, 10:104, 1989.

[347] U. Häuß ermann, M. Dolg, H. Stoll, and H. Preuss. *Mol. Phys.*, 78:1211–1224, 1993.

[348] M. Hülsen, A. Weigand, and M. Dolg. *Theor. Chem. Acc.*, 122:23, 2009.

[349] Ida-Marie Hø yvik, Branislav Jansik, and Poul Jø rgensen. Trust region minimization of orbital localization functions. *J. Chem. Theory Comput.*, 8(9):3137–3146, 2012. doi:10.1021/ct300473g.

[350] J. Hachmann, W. Cardoen, and G. K.-L. Chan. *J. Chem. Phys.*, 125:144101, 2006.

[351] Diptarka Hait and Martin Head-Gordon. Orbital optimized density functional theory for electronic excited states. *J. Phys. Chem. Lett.*, 12:4517–4529, 5 2021. doi:10.1021/acs.jpclett.1c00744.

[352] S. Haldar, C. Riplinger, B. Demoulin, F. Neese, R. Izsak, and A. K. Dutta. Multilayer approach to the ip-eom-dlpno-ccsd method: theory, implementation, and application. *J. Chem. Theory Comput.*, 15(4):2265–2277, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000464475500015, doi:10.1021/acs.jctc.8b01263.

[353] T. P. Hamilton and P. Pulay. *J. Chem. Phys.*, 84:5728, 1986.

[354] B. Hammer, L. B. Hansen, and J. K. Nø rskov. *Phys. Rev. B*, 59:7413, 1999.

[355] C. Hampel, K. A. Peterson, and H. J. Werner. *Chem. Phys. Lett.*, 190:1, 1992.

[356] C. Hampel and H. J. Werner. *J. Chem. Phys.*, 104:6286, 1996.

[357] Matthias Hanauer and Andreas Köhn. Pilot applications of internally contracted multireference coupled cluster theory, and how to choose the cluster operator properly. *J. Chem. Phys.*, 134(20):204111, 05 2011. doi:10.1063/1.3592786.

[358] N. C. Handy and A. J. Cohen. *Mol. Phys.*, 99:403, 2001.

[359] N. C. Handy, P. J. Knowles, and K. Somasundram. *Theor. Chem. Acc.*, 68:87, 1985.

[360] N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks. *Chem. Phys. Lett.*, 164:185, 1989.

[361] A. Hansen, D. G. Liakos, and F. Neese. *J. Chem. Phys.*, 135:214102, 2011.

[362] D. J. Harding, P. Gruene, M. Haertelt, G. Meijer, A. Fielicke, S. M. Hamilton, W. S. Hopkins, S. R. Mackenzie, S. P. Neville, and T. R. Walsh. Probing the structures of gas-phase rhodium cluster cations by far-infrared spectroscopy. *J. Chem. Phys.*, 133(21):214304, 2010. doi:10.1063/1.3509778.

[363] P. C. Hariharan and J. A. Pople. *Theor. Chim. Acta*, 28:213, 1973.

[364] John E. Harriman. *Theoretical Foundations of Electron Spin Resonance: Physical Chemistry: A Series of Monographs*. Academic Press, 1978. ISBN 978-1483175855.

[365] J. N. Harvey, M. Aschi, H. Schwarz, and W. Koch. *Theor. Chem. Acc.*, 99:95, 1998.

[366] Remco W. A. Havenith, Peter R. Taylor, Celestino Angeli, Renzo Cimiraglia, and Kenneth Ruud. Calibration of the n-electron valence state perturbation theory approach. *J. Chem. Phys.*, 120:4619, 2004. doi:10.1063/1.1645243.

[367] P. J. Hay and W. R. Wadt. *J. Chem. Phys.*, 82:270 & 284 & 299, 1985.

[368] W. M. Haynes, D. R. Lide, and T. J. Bruno. *Handbook of Chemistry and Physics*. CRC Press, 95th edition, 2014. ISBN 978-1-4822-0868-9.

[369] M. Head-Gordon. *Chem. Phys. Lett.*, 372:508, 2003.

[370] M. Head-Gordon and J. A. Pople. *Chem. Phys. Lett.*, 153:503, 1988.

[371] M. Head-Gordon, R. A. Rico, M. Oumi, and T. J. Lee. *Chem. Phys. Lett.*, 219:21–29, 1994.

[372] W. J. Hehre, R. Ditchfield, and J. A. Pople. *J. Chem. Phys.*, 56:2257, 1972.

[373] W. J. Hehre, R. Ditchfield, R. F. Stewart, and J. A. Pople. *J. Chem. Phys.*, 52:2769, 1970.

[374] W. J. Hehre, R. F. Stewart, and J. A. Pople. *J. Chem. Phys.*, 51:2657, 1969.

[375] T. Helgaker, M. Jaszuński, and K. Ruud. *Chem. Rev.*, 99:293, 1999.

[376] T. Helgaker and P. R. Taylor. In D. R. Yarkony, editor, *Modern Electronic Structure Theory*, pages 725. World Scientific, 1995.

[377] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*, chapter 9. John Wiley & Sons, 2013.

[378] Trygve U. Helgaker, Jan Almlöf, Hans Jørgen Aa. Jensen, and Poul Jørgensen. Molecular Hessians for large-scale MCSCF wave functions. *J. Chem. Phys.*, 84(11):6266–6279, 1986. doi:10.1063/1.450771.

[379] Benjamin Helmich-Paris. Benchmarks for electronically excited states with CASSCF methods. *J. Chem. Theory Comput.*, 15(7):4170–4179, 2019. URL: https://doi.org/10.1021/acs.jctc.9b00325, doi:10.1021/acs.jctc.9b00325.

[380] Benjamin Helmich-Paris. CASSCF linear response calculations for large open-shell molecules. *J. Chem. Phys.*, 150(17):174121, 2019. URL: https://doi.org/10.1063/1.5092613, doi:10.1063/1.5092613.

[381] Benjamin Helmich-Paris. A trust-region augmented Hessian implementation for restricted and unrestricted Hartree–Fock and Kohn–Sham methods. *J. Chem. Phys.*, 154(16):164104, 2021. URL: https://doi.org/10.1063/5.0040798, doi:10.1063/5.0040798.

[382] Benjamin Helmich-Paris. A trust-region augmented Hessian implementation for state-specific and state-averaged CASSCF wave functions. *J. Chem. Phys.*, 156(20):204104, 2022. URL: https://doi.org/10.1063/5.0090447, arXiv:https://doi.org/10.1063/5.0090447, doi:10.1063/5.0090447.

[383] Benjamin Helmich-Paris, Bernardo de Souza, Frank Neese, and Róbert Izsák. An improved chain of spheres for exchange algorithm. *J. Chem. Phys.*, 155(10):104109, 2021. doi:10.1063/5.0058766.

[384] G. Henkelman and H. Jónsson. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.*, 113(22):9978–9985, 2000.

[385] G. Henkelman, B.P. Uberuaga, and H. Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.*, 113(22):9901–9904, 2000.

[386] Laura Hernández-Martínez, Eric Brémond, Angel J. Pérez-Jiménez, Emilio San-Fabián, Carlo Adamo, and Juan C. Sancho-García. Nonempirical (double-hybrid) density functionals applied to atomic excitation energies: a systematic basis set investigation. *Int. J. Quantum Chem.*, 120:e26193, 2020. URL: https://onlinelibrary.wiley.com/doi/10.1002/qua.26193, doi:https://doi.org/10.1002/qua.26193.

[387] R. H. Hertwig and W. Koch. *Chem. Phys. Lett.*, 268:345, 1997.

[388] G Herzberg. *Infrared and Raman Spectra*. Van Nostrand Reinhold, 1945.

[389] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. *J. Chem. Theory Comput.*, 4:435, 2008.

[390] Bernd A. Hess, Christel M. Marian, Ulf Wahlgren, and Odd Gropen. *Chem. Phys. Lett.*, 251:365–371, 1996.

[391] J. L. Heully and J.-P. Malrieu. *J. Mol. Struct.: THEOCHEM*, 768:53, 2006.

[392] J. Hillenbrand, M. van Gastel, E. Bill, F. Neese, and A. Furstner. Isolation of a homoleptic non-oxo Mo(V) alkoxide complex: Synthesis, structure, and electronic properties of penta-tert-butoxymolybdenum. *J. Am. Chem. Soc.*, 142(38):16392–16402, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000575684100032, doi:10.1021/jacs.0c07073.

[393] So Hirata and Martin Head-Gordon. Time–Dependent density functional theory within the Tamm–Dancoff approximation. *Chem. Phys. Lett.*, 314:291–299, 1999.

[394] F. L. Hirshfeld. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta*, 44:129–138, 1977.

[395] R. Hockney. *Methods Comp. Phys.*, 9:136–211, 1970.

[396] Manuel Hodecker, Malgorzata Biczysko, Andreas Dreuw, and Vincenzo Barone. Simulation of vacuum uv absorption and electronic circular dichroism spectra of methyl oxirane: the role of vibrational effects. *Journal of Chemical Theory and Computation*, 12(6):2820–2833, 2016. URL: https://doi.org/10.1021/acs.jctc.6b00121, doi:10.1021/acs.jctc.6b00121.

[397] W.-M. Hoe, A. J. Cohen, and N. C. Handy. *Chem. Phys. Lett.*, 341:319, 2001.

[398] T. Holmgaard List, N. Saue and P. Norman. Rotationally averaged linear absorption spectra beyond the electric-dipole approximation. *Mol. Phys.*, 115(1-2):63–74, 2017. doi:10.1080/00268976.2016.1187773.

[399] H. Horn, H. Wei, M. Häser, M. Ehrig, and R. Ahlrichs. *J. Comput. Chem.*, 12:1058, 1991.

[400] W. Hujo and S. Grimme. *J. Chem. Theory Comput.*, 2011. doi:dx.doi.org/10.1021/ct200644w.

[401] W. Hujo and S. Grimme. *Phys. Chem. Chem. Phys.*, 13:13942, 2011.

[402] T. F. Hunter and R. F. Wyatt. Intersystem crossing in anthracene. *Chem. Phys. Lett.*, 6(3):221–224, 1970. doi:10.1016/0009-2614(70)80224-X.

[403] L .M. J. Huntington, M. Krupička, F. Neese, and R. Izsák. Similarity transformed equation of motion coupled-cluster theory based on an unrestricted hartree-fock reference for applications to high-spin open-shell systems. *J. Chem. Phys.*, 147:174104, 2017.

[404] L. M. J. Huntington, A. Hansen, F. Neese, and M. Nooijen. *J. Chem. Phys.*, 136:064101, 2012.

[405] L. M. J. Huntington and M. Nooijen. *J. Chem. Phys.*, 133:184109, 2010.

[406] L. M. J. Huntington and M. Nooijen. *J. Chem. Phys.*, 142:194111, 2015.

[407] L. M. J. Huntington and M. Nooijen. *J. Chem. Theory Comput.*, 12:114, 2016.

[408] Marcella Iannuzzi, Alessandro Laio, and Michele Parrinello. Efficient exploration of reactive potential energy surfaces using Car–Parrinello molecular dynamics. *Phys. Rev. Lett.*, 90(23):238302, 2003. doi:10.1103/PhysRevLett.90.238302.

[409] G. Igel-Mann, H. Stoll, and H. Preuss. *Mol. Phys.*, 65:1321–1328, 1988.

[410] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao. *J. Chem. Phys.*, 115:3540–3544, 2001.

[411] Mark A. Iron and Trevor Janes. Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database. *J. Phys. Chem. A*, 123(17):3761–3781, 05 2019. URL: https://doi.org/10.1021/acs.jpca.9b01546 (visited on 2020-09-15), doi:10.1021/acs.jpca.9b01546.

[412] Kazuhiro Ishida, Keiji Morokuma, and Andrew Komornicki. The intrinsic reaction coordinate. An ab initio calculation for HNC\textrightarrow HCN and H-+CH4\textrightarrow CH4+H-. *J. Chem. Phys.*, 66(5):2153–2156, 1977. doi:10.1063/1.434152.

[413] Naoya Iwahara and Liviu F. Chibotaru. Exchange interaction between j multiplets. *Physical Review B*, 91:174438, May 2015. URL: https://link.aps.org/doi/10.1103/PhysRevB.91.174438, doi:10.1103/PhysRevB.91.174438.

[414] Naoya Iwahara, Liviu Ungur, and Liviu F. Chibotaru. \~\~J-pseudospin states and the crystal field of cubic systems. *Phys. Rev. B*, 98(5):054436, 08 2018. URL: https://link.aps.org/doi/10.1103/PhysRevB.98.054436, doi:10.1103/PhysRevB.98.054436.

[415] R. Izsak and F. Neese. *J. Chem. Phys.*, 135:144105, 2011.

[416] H. Jónsson, G. Mills, and K.W. Jacobsen. *Classical and Quantum Dynamics in Condensed Phase Simulations*. World Scientific Publishing Company, 1998.

[417] Poul Jø rgensen, Hans Jø rgen Aagaard Jensen, and Jeppe Olsen. Linear response calculations for large scale multiconfiguration self-consistent field wave functions. *J. Chem. Phys.*, 89(6):3654–3661, 1988.

[418] K. Jankowski and J. Paldus. Applicability of coupled-pair theories to quasidegenerate electronic states: A model study. *Int. J. Quantum Chem.*, 18(5):1243–1269, 11 1980. doi:10.1002/qua.560180511.

[419] F. Jensen. *Introduction to Computational Chemistry*. Wiley, 1999.

[420] B. Jeziorski and H. J. Monkhorst. *Phys. Rev. A*, 24:1668–1681, 1981.

[421] S. Jiang, D. Maganas, N. Levesanos, E. Ferentinos, S. Haas, K. Thirunavukkuarasu, J. Krzystek, M. Dressel, L. Bogani, F. Neese, and P. Kyritsis. *J. Am. Chem. Soc.*, 137:12923, 2015.

[422] B. G. Johnson, P. M. W. Gill, and J. A. Pople. *J. Chem. Phys.*, 98:5612, 1993.

[423] B. G. Johnson, P. M. W. Gill, and J. A. Pople. *Chem. Phys. Lett.*, 220:377, 1994.

[424] E. R. Johnson and A. D. Becke. *J. Chem. Phys.*, 123:024101, 2005.

[425] E. R. Johnson and A. D. Becke. *J. Chem. Phys.*, 124:174104, 2006.

[426] E. Bright Wilson Jr, J. C. Decius, and Paul C. Cross. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*. Dover Publications, revised ed. edition edition, 03 1980. ISBN 978-0-486-63941-3.

[427] J. Jung, S. T. Loffler, J. Langmann, F. W. Heinemann, E. Bill, G. Bistoni, W. Scherer, M. Atanasov, K. Meyer, and F. Neese. Dispersion forces drive the formation of uranium-alkane adducts. *J. Am. Chem. Soc.*, 142(4):1864–1870, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000510531900033, doi:10.1021/jacs.9b10620.

[428] Julie Jung, Mihail Atanasov, and Frank Neese. Ab Initio Ligand-Field Theory Analysis and Covalency Trends in Actinide and Lanthanide Free Ions and Octahedral Complexes. *Inorg. Chem.*, 56(15):8802–8816, 08 2017. doi:10.1021/acs.inorgchem.7b00642.

[429] P. Jurečka, J. Šponer, J. Černý, and P. Hobza. *Phys. Chem. Chem. Phys.*, 8:1985, 2006.

[430] Johannes Kästner, Hans Martin Senn, Stephan Thiel, Nikolaj Otte, and Walter Thiel. QM/MM free-energy perturbation compared to thermodynamic integration and umbrella sampling: Application to an enzymatic reaction. *J. Chem. Theory Comput.*, 2(2):452–461, 2006. doi:10.1021/ct050252w.

[431] H. J. Köhler and F. Birnstock. *Z. Chem.*, 5:196, 1972.

[432] W. Küchle, M. Dolg, H. Stoll, and H. Preuss. *Mol. Phys.*, 74:1245–1263, 1991.

[433] W. Küchle, M. Dolg, H. Stoll, and H. Preuss. *J. Chem. Phys.*, 100:7535–7542, 1994.

[434] V. Kairys and J.D. Head. Geometry optimization of charged molecules in an external electric field applied to F-$\odot $ H2O and I-$\odot $ H2O. *J. Phys. Chem. A*, 102(8):1365–1370, 1998.

[435] Jaroslaw Kalinowski, Frank Wennmohs, and Frank Neese. Arbitrary angular momentum electron repulsion integrals with graphical processing units: application to the resolution of identity hartree–fock method. *J. Chem. Theory Comput.*, 13(7):3160–3170, 2017.

[436] A. Karton, A Tarnopolsky, J.-F. Lamère, G. C. Schatz, and J. M. L. Martin. Highly accurate first-principles benchmark data sets for the parametrization and validation of density functional and other approximate methods. derivation of a robust, generally applicable, double-hybrid functional for thermochemistry and thermochemical kinetics. *J. Phys. Chem. A*, 112:12868, 2008. URL: https://pubs.acs.org/doi/abs/10.1021/jp801805p, doi:https://doi.org/10.1021/jp801805p.

[437] M. Kaupp, P. v. R. Schleyer, H. Stoll, and H. Preuss. *J. Chem. Phys.*, 94:1360–1366, 1991.

[438] M. Keilwerth, J. Hohenberger, F. W. Heinemann, J. Sutter, A. Scheurer, H. Y. Fang, E. Bill, F. Neese, S. F. Ye, and K. Meyer. A series of iron nitrosyl complexes Fe-NO(6-9) and a fleeting Fe-NO(10) intermediate en route to a metalacyclic iron nitrosoalkane. *J. Am. Chem. Soc.*, 141(43):17217–17235, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000493866300028, doi:10.1021/jacs.9b08053.

[439] R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison. *J. Chem. Phys.*, 96:6769, 1992.

[440] R. A. Kendall and H. A. Früchtl. *Theor. Chem. Acc.*, 97:158, 1997.

[441] Mikaël Kepenekian, Vincent Robert, and Boris Le Guennic. What zeroth-order Hamiltonian for CASPT2 adiabatic energetics of Fe(II)N6 architectures? *J. Chem. Phys.*, 131:114702, 09 2009. doi:https://doi.org/10.1063/1.3211020.

[442] Manoj K. Kesharwani, Brina Brauer, and Jan M. L. Martin. Frequency and Zero-Point Vibrational Energy Scale Factors for Double-Hybrid Density Functionals (and Other Selected Methods): Can Anharmonic Force Fields Be Avoided? *J. Phys. Chem. A*, 119(9):1701–1714, 03 2015. URL: https://doi.org/10.1021/jp508422u (visited on 2020-04-14), doi:10.1021/jp508422u.

[443] Abhishek Khedkar and Michael Roemelt. Active Space Selection Based on Natural Orbital Occupation Numbers from n-Electron Valence Perturbation Theory. *J. Chem. Theory Comput.*, 15:3522–3536, 2019.

[444] H. F. King, R. E. Stanton, H. Kim, R. E. Wyatt, and R. G. Parr. *J. Chem. Phys.*, 47:1936, 1967.

[445] A. Klamt and F. Eckert. COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equil.*, 172:43–72, 2000. doi:https://doi.org/10.1016/S0378-3812(00)00357-5.

[446] A. Klamt and Schüürmann. *J. Chem. Soc. Perkin Trans. 2*, pages 799, 1993.

[447] Andreas Klamt. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.*, 99:2224–2235, 1995.

[448] Andreas Klamt, Jonas Volker, Thorsten Bürger, and John C. W. Lohrenz. Refinement and parametrization of COSMO-RS. *J. Phys. Chem. A*, 102:5074–5085, 1998.

[449] G. Knizia. *J. Chem. Theory Comput.*, 9:4834–4843, 2013.

[450] P. J. Knowles, J. S. Andrews, R. D. Amos, N. C. Handy, and J. A. Pople. *Chem. Phys. Lett.*, 186:130, 1991.

[451] W. Koch and M. C. Holthausen. *A Chemist's Guide to Density Functional Theory*. Wiley-VCH, 2000.

[452] C. Kollmar. *J. Chem. Phys.*, 105:8204, 1996.

[453] C. Kollmar. *Int. J. Quant. Chem.*, 62:617, 1997.

[454] C. Kollmar and A. Hesselmann. *Theor. Chem. Acc.*, 127:311, 2010.

[455] C. Kollmar and F. Neese. *Mol. Phys.*, 108:2449, 2010.

[456] C. Kollmar and F. Neese. *J. Chem. Phys.*, 135:064103, 2011.

[457] C. Kollmar and F. Neese. *J. Chem. Phys.*, 135:084102, 2011.

[458] Christian Kollmar, Kantharuban Sivalingam, Yang Guo, and Frank Neese. An efficient implementation of the NEVPT2 and CASPT2 methods avoiding higher-order density matrices. *J. Chem. Phys.*, 155(23):234104, 12 2021.

[459] Christian Kollmar, Kantharuban Sivalingam, Benjamin Helmich-Paris, Celestino Angeli, and Frank Neese. A perturbation-based super-CI approach for the orbital optimization of a CASSCF wave function. *J. Comput. Chem.*, 40:1463–1470, 2019. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.25801, doi:10.1002/jcc.25801.

[460] Christian Kollmar, Kantharuban Sivalingam, and Frank Neese. An alternative choice of the zeroth-order Hamiltonian in CASPT2 theory. *J. Chem. Phys.*, 152(21):214110, 06 2020. doi:10.1063/5.0010019.

[461] Christian Kollmar, Kantharuban Sivalingam, and Frank Neese. An alternative choice of the zeroth-order Hamiltonian in CASPT2 theory. *J. Chem. Phys.*, 152(21):214110, 06 2020. doi:10.1063/5.0010019.

[462] S. Koseki, M. S. Gordon, M. W. Schmidt, and N. Matsunaga. *J. Phys. Chem.*, 99:12764–12772, 1995.

[463] S. Koseki, M. W. Schmidt, and M. S. Gordon. *J. Phys. Chem.*, 96:10768–10772, 1992.

[464] S. Koseki, M. W. Schmidt, and M. S. Gordon. *J. Phys. Chem. A*, 102:10430–10435, 1998.

[465] S. Kossmann and F. Neese. *Chem. Phys. Lett.*, 481:240–243, 2009.

[466] M. Kotzian, N. Rösch, and M. C. Zerner. *Theor. Chim. Acta*, 81:201, 1992.

[467] Sebastian Kozuch, David Gruzman, and Jan M. L. Martin. DSD-BLYP: A general purpose double hybrid density functional including spin component scaling and dispersion correction. *J. Phys. Chem. C*, 114(48):20801–20808, 2010. doi:10.1021/jp1070852.

[468] Sebastian Kozuch and Jan M. L. Martin. DSD-PBEP86: in search of the best double-hybrid DFT with spin-component scaled MP2 and dispersion corrections. *Phys. Chem. Chem. Phys.*, 13(45):20104–20107, 2011. doi:10.1039/C1CP22592H.

[469] Sebastian Kozuch and Jan M. L. Martin. Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *J. Comput. Chem.*, 34(27):2327–2344, 2013. doi:10.1002/jcc.23391.

[470] M. Krack and A. M. Köster. *J. Chem. Phys.*, 108:3226, 1998.

[471] V. Krewald, F. Neese, and D. A. Pantazis. Implications of structural heterogeneity for the electronic structure of the final oxygen-evolving intermediate in photosystem II. *J. Inorg. Biochem.*, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000488146900034, doi:10.1016/j.jinorgbio.2019.110797.

[472] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople. *J. Chem. Phys.*, 72:650, 1980.

[473] R. Krishnan, M. J. Frisch, and J. A. Pople. *J. Chem. Phys.*, 72:4244, 1980.

[474] M. Krupička, K. Sivalingam, L. Huntington, A. A. Auer, and F. Neese. A toolchain for the automatic generation of computer codes for correlated wavefunction calculations. *J. Comput. Chem.*, 38:1853, 2017.

[475] H. Kruse, L. Goerigk, and S. Grimme. *J. Org. Chem.*, 23:10824, 2012.

[476] H. Kruse and S. Grimme. *J. Chem. Phys.*, 16:136, 2012.

[477] Adam Kubas, Felix Hoffmann, Alexander Heck, Harald Oberhofer, Marcus Elstner, and Jochen Blumberger. Electronic couplings for molecular charge transfer: Benchmarking CDFT, FODFT, and FODFTB against high-level ab initio calculations. *The Journal of Chemical Physics*, 140(10):104105, March 2014. URL: https://doi.org/10.1063/1.4867077 (visited on 2024-06-18), doi:10.1063/1.4867077.

[478] Adam Kubas, Johannes Noak, Annette Trunschke, Robert Schlögl, Frank Neese, and Dimitrios Maganas. A combined experimental and theoretical spectroscopic protocol for determination of the structure of heterogeneous catalysts: developing the information content of the resonance raman spectra of m1 movo x. *Chem. Sci.*, 8(9):6338–6353, 2017.

[479] Adam Kubas, Max Verkamp, Josh Vura-Weis, Frank Neese, and Dimitrios Maganas. Restricted open-shell configuration interaction singles study on m-and l-edge x-ray absorption spectroscopy of solid chemical systems. *J. Chem. Theory Comput.*, 14(8):4320–4334, 2018.

[480] A. Kumar, F. Neese, and E. Valeev. Near-linear scaling explicitly correlated coupled cluster singles and doubles method based on an open-shell domain-based local pair natural orbitals. *Abstr. Pap. Am. Chem. Soc.*, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000525061504143.

[481] A. Kumar, F. Neese, and E. F. Valeev. Explicitly correlated coupled cluster method for accurate treatment of open-shell molecules with hundreds of atoms. *J. Chem. Phys.*, 153(9):17, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000570176400001, doi:10.1063/5.0012753.

[482] Shankar Kumar, John M. Rosenberg, Djamal Bouzida, Robert H. Swendsen, and Peter A. Kollman. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.*, 13(8):1011–1021, 1992. doi:10.1002/jcc.540130812.

[483] R. Kutteh. Rattle recipe for general holonomic constraints: angle and torsion constraints. *CCP5 Newslett.*, 1998.

[484] W. Kutzelnigg, U. Fleischer, and M. Schindler. *The IGLO-Method: Ab Initio Calculation and Interpretation of NMR Chemical Shifts and Magnetic Susceptibilities*. Volume 23. Springer Verlag, 1990.

[485] W. Kutzelnigg and D. Mukherjee. *J. Chem. Phys.*, 107:432, 1997.

[486] D. Laikov and C. Van Wüllen. Software. URL: http://www.ccl.net/cca/software/SOURCES/FORTRAN/Lebedev-Laikov-Grids/.

[487] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.*, 99(20):12562–12566, 2002. doi:10.1073/pnas.202427399.

[488] J. Lang, J. Brabec, M. Saitow, J. Pittner, F. Neese, and O. Demel. Perturbative triples correction to domain-based local pair natural orbital variants of mukherjee's state specific coupled cluster method. *Phys. Chem. Chem. Phys.*, 21(9):5022–5038, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000461722700029, doi:10.1039/c8cp03577f.

[489] Lucas Lang. *Development of New Multistate Multireference Perturbation Theory Methods and Their Application*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2020.

[490] Lucas Lang, Mihail Atanasov, and Frank Neese. Improvement of Ab Initio Ligand Field Theory by Means of Multistate Perturbation Theory. *J. Phys. Chem. A*, 124(5):1025–1037, 02 2020. doi:10.1021/acs.jpca.9b11227.

[491] Lucas Lang and Frank Neese. Spin-dependent properties in the framework of the dynamic correlation dressed complete active space method. *J. Chem. Phys.*, 150:104104, 2019.

[492] Lucas Lang, Enrico Ravera, Giacomo Parigi, Claudio Luchinat, and Frank Neese. Solution of a puzzle: High-level quantum-chemical treatment of pseudocontact chemical shifts confirms classic semiempirical theory. *J. Phys. Chem. Lett.*, 11(20):8735–8744, 2020.

[493] Lucas Lang, Kantharuban Sivalingam, and Frank Neese. The combination of multipartitioning of the Hamiltonian with canonical Van Vleck perturbation theory leads to a Hermitian variant of quasidegenerate N-electron valence perturbation theory. *J. Chem. Phys.*, 152(1):014109, 01 2020. doi:10.1063/1.5133746.

[494] A. W. Lange and J. M. Herbert. *J. Chem. Phys.*, 133:244111, 2010.

[495] W. J. Lauderdale, J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett. *Chem. Phys. Lett.*, 187:21, 1991.

[496] V. I. Lebedev. *Zh. Vychisl. Mat. Fiz.*, 15:48, 1975.

[497] V. I. Lebedev. *Zh. Vychisl. Mat. Fiz.*, 16:293, 1976.

[498] V. I. Lebedev and D. N. Laikov. *Dokl. Math.*, 59:477, 1999.

[499] V. I. Lebedev and A. L. Skorokhodov. *Sov. Phys.-Dokl.*, 45:587, 1992.

[500] M. H. Lechner, A. Papadopoulos, K. Sivalingam, A. A. Auer, A. Koslowski, U. Becker, F. Wennmohs, and F. Neese. Code generation in ORCA: Progress, Efficiency and Tight integration. *Phys. Chem. Chem. Phys.*, 26(21):15205–15220, 2024.

[501] Marvin H. Lechner, Róbert Izsák, Marcel Nooijen, and Frank Neese. A perturbative approach to multireference equation-of-motion coupled cluster. *Mol. Phys.*, pages e1939185, 06 2021. URL: http://doi.org/10.1080/00268976.2021.1939185, doi:10.1080/00268976.2021.1939185.

[502] C. Lee, W. Yang, and R. G. Parr. *Phys. Rev. B*, 37:785, 1988.

[503] K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth. *Phys. Rev. B*, 82:081101, 2010.

[504] T. J. Lee and P. R. Taylor. *Int. J. Quant. Chem. Symp.*, 23:199, 1989.

[505] S. Lehtola. *J. Chem. Phys.*, 152:134108, 2020.

[506] S. Lehtola, C. Steigemann, MJT Oliveira, and MAL Marques. Recent developments in libxc – A comprehensive library of functionals for density functional theory. *SoftwareX*, 7:1–5, 2019. URL: https://doi.org/10.1016/j.softx.2017.11.002.

[507] T. Leininger, A. Berning, A. Nicklass, H. Stoll, H.-J. Werner, and H.-J. Flad. *Chem. Phys.*, 217:19, 1997.

[508] T. Leininger, A. Nicklass, W. Küchle, H. Stoll, M. Dolg, and A. Bergner. *Chem. Phys. Lett.*, 255:274, 1996.

[509] T. Leininger, A. Nicklass, H. Stoll, M. Dolg, and P. Schwerdtfeger. *J. Chem. Phys.*, 105:1052–1059, 1996.

[510] Gianluca Levi, Aleksei V Ivanov, and Hannes Jónsson. Variational density functional calculations of excited states via direct optimization. *J. Chem. Theory Comput.*, 16(11):6968–6982, 2020. doi:10.1021/acs.jctc.0c00597.

[511] Mathieu Lewin. *J. Math. Chem.*, 44:967, 2008.

[512] Tiago Leyser da Costa Gouveia, Dimitrios Maganas, and Frank Neese. Restricted open-shell hartree–fock method for a general configuration state function featuring arbitrarily complex spin-couplings. *The Journal of Physical Chemistry A*, 128(25):5041–5053, 2024. PMID: 38886177. URL: https://doi.org/10.1021/acs.jpca.4c00688, arXiv:https://doi.org/10.1021/acs.jpca.4c00688, doi:10.1021/acs.jpca.4c00688.

[513] H. Li and J. H. Jensen. *Theor. Chem. Acc.*, 107:211, 2002.

[514] S. Li, J. Ma, and Y. Jiang. Linear scaling local correlation approach for solving the coupled cluster equations of large systems. *J. Comput. Chem.*, 23:237–244, 2002. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.10003, doi:https://doi.org/10.1002/jcc.10003.

[515] S. Li, J. Shen, W. Li, and Y. Jiang. An efficient implementation of the "Cluster-in-Molecule" approach for local electron correlation calculations. *J. Chem. Phys.*, 125:074109, 2006. doi:https://doi.org/10.1063/1.2244566.

[516] W. Li, P. Piecuch, J. Gour, and S. Li. Local correlation calculations using standard and renormalized coupled-cluster approaches. *J. Chem. Phys.*, 131:114109, 2009. doi:https://doi.org/10.1063/1.3218842.

[517] Zhendong Li and Wenjian Liu. First-order nonadiabatic coupling matrix elements between excited states: A Lagrangian formulation at the CIS, RPA, TD-HF, and TD-DFT levels. *J. Chem. Phys.*, 141(1):014110, 2014. URL: https://doi.org/10.1063/1.4885817, arXiv:https://doi.org/10.1063/1.4885817, doi:10.1063/1.4885817.

[518] D. G. Liakos, Y. Guo, and F. Neese. Comprehensive benchmark results for the domain based local pair natural orbital coupled cluster method (dlpno-ccsd(t)) for closed- and open-shell systems. *J. Phys. Chem. A*, 124(1):90–100, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000507151000012, doi:10.1021/acs.jpca.9b05734.

[519] D. G. Liakos and F. Neese. *J. Phys. Chem. A*, 116:4801–4816, 2012.

[520] Dimitrios G. Liakos, Andreas Hansen, and Frank Neese. *J. Chem. Theory Comput.*, 7:76, 2011.

[521] Dimitrios G. Liakos, R. Izsák, E. F. Valeev, and Frank Neese. *Mol. Phys.*, 111:2653, 2013.

[522] Dimitrios G. Liakos, Manuel Sparta, Manoj K. Kesharwani, Jan M. L. Martin, and Frank Neese. *J. Chem. Theory Comput.*, 11:1525, 2015.

[523] I. S. Lim, P. Schwerdtfeger, B. Metz, and H. Stoll. *J. Chem. Phys.*, 122:104103, 2005.

[524] I. S. Lim, H. Stoll, and P. Schwerdtfeger. *J. Chem. Phys.*, 124:034107, 2006.

[525] Y.-S. Lin, G.-D. Li, S.-P. Mao, and J.-D. Chai. Long-range corrected hybrid density functionals with improved dispersion corrections. *J. Chem. Theory Comput.*, 9:263–272, 2013. URL: https://pubs.acs.org/doi/abs/10.1021/ct300715s, doi:https://doi.org/10.1021/ct300715s.

[526] Roland Lindh, Anders Bernhardsson, and Martin Schütz. Force-constant weighted redundant coordinates in molecular geometry optimizations. *Chem. Phys. Lett.*, 303(5):567–575, 04 1999. URL: http://www.sciencedirect.com/science/article/pii/S000926149900247X, doi:10.1016/S0009-2614(99)00247-X.

[527] M. E. Lines. Orbital angular momentum in the theory of paramagnetic clusters. *J. Chem. Phys.*, 55:2977, 1971.

[528] N. H. List, T. R. L. Melin, M. van Horn, and T. Saue. Beyond the electric-dipole approximation in simulations of x-ray absorption spectroscopy: Lessons from relativistic theory. *J. Chem. Phys.*, 152(18):184110, 05 2020. doi:10.1063/5.0003103.

[529] Z. Liu, O. Demel, and M. Nooijen. Multireference Equation of Motion Coupled Cluster study of atomic excitation spectra of first-row transition metal atoms Cr, Mn, Fe and Co. *J. Mol. Spectrosc.*, 311:54, 2015.

[530] Z. Liu, L. M. J. Huntington, and M. Nooijen. *Mol. Phys.*, 113:2999, 2015.

[531] Rohini C. Lochan and Martin Head-Gordon. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *J. Chem. Phys.*, 126(16):164101, 04 2007. doi:10.1063/1.2718952.

[532] F. London. *Phys. Radium*, 8:397, 1937.

[533] Derek A. Long. *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*. Wiley, 1 edition edition, 11 2001. ISBN 978-0-471-49028-9.

[534] Q. Lu, F. Neese, and G. Bistoni. Formation of agostic structures driven by london dispersion. *Angew. Chem. Int. Ed.*, 57(17):4760–4764, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000430165700058, doi:10.1002/anie.201801531.

[535] Q. Lu, F. Neese, and G. Bistoni. London dispersion effects in the coordination and activation of alkanes in sigma-complexes: a local energy decomposition study. *Phys. Chem. Chem. Phys.*, 21(22):11569–11577, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000472218500051, doi:10.1039/c9cp01309a.

[536] Qing Lu, Frank Neese, and Giovanni Bistoni. Formation of agostic structures driven by london dispersion. *Angew. Chem. Int. Ed.*, 57(17):4760–4764, 2018.

[537] Qing Lu, Frank Neese, and Giovanni Bistoni. London dispersion effects in the coordination and activation of alkanes in σ-complexes: a local energy decomposition study. *Phys. Chem. Chem. Phys.*, 21(22):11569–11577, 2019.

[538] Dmitry I. Lyakh, Monika Musiał, Victor F. Lotrich, and Rodney J. Bartlett. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chem. Rev.*, 112(1):182–243, 12 2012. doi:10.1021/cr2001417.

[539] Klaus Müller and Leo D. Brown. Location of saddle points and minimum energy paths by a constrained simplex optimization procedure. *Theor. Chem. Acc.*, 53(1):75–93, 1979. doi:10.1007/BF00547608.

[540] Marcel Müller, Andreas Hansen, and Stefan Grimme. Ωb97x-3c: a composite range-separated hybrid dft method with a molecule-optimized polarized valence double-ζ basis set. *J. Chem. Phys.*, 158(1):014103, 2023. URL: https://doi.org/10.1063/5.0133026, arXiv:https://doi.org/10.1063/5.0133026, doi:10.1063/5.0133026.

[541] Marcel Müller, Andreas Hansen, and Stefan Grimme. An atom-in-molecule adaptive polarized valence single-ζ atomic orbital basis for electronic structure calculations. *J. Chem. Phys.*, 159:164108, 2023.

[542] Thomas Müller and Hans Lischka. Simultaneous calculation of Rydberg and valence excited states of formaldehyde. *Theor. Chem. Acc.*, 106:369–378, 2001.

[543] J. Mášik and I. Hubač. *Adv. Quant. Chem.*, 31:75–104, 1998.

[544] Satoshi Maeda, Koichi Ohno, and Keiji Morokuma. Updated Branching Plane for Finding Conical Intersections without Coupling Derivative Vectors. *J. Chem. Theory Comput.*, 6(5):1538–1545, 05 2010. URL: https://doi.org/10.1021/ct1000268 (visited on 2021-01-07), doi:10.1021/ct1000268.

[545] D. Maganas, S. DeBeer, and F. Neese. Pair natural orbital restricted open-shell configuration interaction (PNO-ROCIS) approach for calculating x-ray absorption spectra of large chemical systems. *J. Phys. Chem. A*, 122(5):1215–1227, 2018. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jpca.7b10880, doi:https://doi.org/10.1021/acs.jpca.7b10880.

[546] D. Maganas, J. K. Kowalska, C. Van Stappen, S. DeBeer, and F. Neese. Mechanism of L-2,L-3-edge x-ray magnetic circular dichroism intensity from quantum chemical calculations and experiment-A case study on V-(IV)/V-(III) complexes. *J. Chem. Phys.*, 152(11):15, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000521227700001, doi:10.1063/1.5129029.

[547] D. Maganas, S. Sottini, P. Kyritsis, E. J. J. Groenen, and F. Neese. *Inorg. Chem.*, 50:8741, 2011.

[548] Dimitrios Maganas, Serena DeBeer, and Frank Neese. A restricted open configuration interaction with singles method to calculate valence-to-core resonant x-ray emission spectra: a case study. *Inorg. Chem.*, 56(19):11819–11836, 2017.

[549] Dimitrios Maganas, Joanna K. Kowalska, Marcel Nooijen, Serena DeBeer, and Frank Neese. Comparison of multireference ab initio wavefunction methodologies for X-ray absorption edges: A case study on [Fe(II/III)Cl4]2–/1– molecules. *J. Chem. Phys.*, 150(10):104106, 2019. URL: https://pubs.aip.org/aip/jcp/article-abstract/150/10/104106/1058894/, doi:https://doi.org/10.1063/1.5051613.

[550] U. S. Mahapatra, B. Datta, and D. Mukherjee. *J. Chem. Phys.*, 110:6171–6188, 1999.

[551] Sebastian Mai, Felix Plasser, Mathias Pabst, Frank Neese, Andreas Köhn, and Leticia González. Surface hopping dynamics including intersystem crossing using the algebraic diagrammatic construction method. *J. Chem. Phys.*, 147(18):184109, 2017.

[552] M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. A*, 113:5806–5812, 2009.

[553] E. Maras, O. Trushin, A. Stukowski, T. Ala-Nissila, and H. Jónsson. Global transition path search for dislocation formation in Ge on Si (001). *Comput. Phys. Commun.*, 205:13–21, 2016.

[554] Narbe Mardirossian and Martin Head-Gordon. *Phys. Chem. Chem. Phys.*, 16(21):9904–9924, 2014. URL: http://dx.doi.org/10.1039/C3CP54374A, doi:10.1039/C3CP54374A.

[555] Narbe Mardirossian and Martin Head-Gordon. *J. Chem. Phys.*, 142:074111, 2015. URL: https://aip.scitation.org/doi/10.1063/1.4907719, doi:10.1063/1.4907719.

[556] Narbe Mardirossian and Martin Head-Gordon. *J. Chem. Phys.*, 144:214110, 2016. URL: https://aip.scitation.org/doi/10.1063/1.4952647, doi:10.1063/1.4952647.

[557] Narbe Mardirossian and Martin Head-Gordon. Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ωB97M(2) double hybrid density functional. *J. Chem. Phys.*, 148(24):241736, jun 2018. URL: https://doi.org/10.1063/1.5025226, doi:10.1063/1.5025226.

[558] A. V. Marenich, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. B*, 113:6378, 2009.

[559] A.V. Marenich, G.D. Hawkins, D.A. Liotard, and D.G. Cramer. GESOL - version 2008. URL: https://comp.chem.umn.edu/gesol.

[560] J. M. L. Martin and A. Sundermann. *J. Chem. Phys.*, 114:3408, 2001.

[561] Richard L. Martin. Natural transition orbitals. *J. Chem. Phys.*, 118:4775–4777, 2003.

[562] Glenn J. Martyna, Michael L. Klein, and Mark Tuckerman. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97(4):2635–2643, 1992. doi:10.1063/1.463940.

[563] Glenn J. Martyna, Mark E. Tuckerman, Douglas J. Tobias, and Michael L. Klein. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.*, 87(5):1117–1157, 1996. doi:10.1080/00268979600100761.

[564] J. Mason. *Solid State Nucl. Magn. Res.*, 2:285, 1993.

[565] Rémi Maurice, Roland Bastardis, Coen de Graaf, Nicolas Suaud, Talal Mallah, and Nathalie Guihéry. Universal theoretical approach to extract anisotropic spin hamiltonians. *J. Chem. Theory Comput.*, 5:2977–2984, 2009.

[566] Sergey N. Maximoff and Gustavo E. Scuseria. Nuclear magnetic resonance shielding tensors calculated with kinetic energy density-dependent exchange-correlation functionals. *Chem. Phys. Lett.*, 390(4-6):408–412, 06 2004. URL: http://linkinghub.elsevier.com/retrieve/pii/S0009261404005949, doi:10.1016/j.cplett.2004.04.049.

[567] I. Mayer. *Chem. Phys. Lett.*, 97:270, 1983.

[568] I. Mayer. *Int. J. Quant. Chem.*, 26:151, 1984.

[569] I. Mayer. *Theor. Chim. Acta*, 67:315, 1985.

[570] I. Mayer. In Z. B. Maksić, editor, *Modelling of Structure and Properties of Molecules*. John Wiley and Sons, 1987.

[571] Nicholas J. Mayhall, Krishnan Raghavachari, and Hrant P. Hratchian. ONIOM-based QM:QM electronic embedding method using Löwdin atomic charges: Energies and analytic gradients. *J. Chem. Phys.*, 132(11):114107, 2010. URL: https://doi.org/10.1063/1.3315417, doi:10.1063/1.3315417.

[572] A. D. McLean and G. S. Chandler. *J. Chem. Phys.*, 72:5639, 1980.

[573] A. D. McLean and M. Yoshimine. Theory of Molecular Polarizabilities. *The Journal of Chemical Physics*, 47(6):1927–1935, 09 1967. URL: https://doi.org/10.1063/1.1712220, arXiv:https://pubs.aip.org/aip/jcp/article-pdf/47/6/1927/18852261/1927_1_online.pdf, doi:10.1063/1.1712220.

[574] R. McWeeny. *Mol. Phys.*, 28:1273, 1974.

[575] R. McWeeny. *Methods of Molecular Quantum Mechanics. 2nd Edition*. Academic Press, 1992.

[576] M. Melander, K. Laasonen, and H. Jónsson. Removing external degrees of freedom from transition-state search methods using quaternions. *J. Chem. Theory Comput.*, 11(3):1055–1062, 2015.

[577] M. C. R. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. C. Maia, G. B. Rocha, J. V. Ribeiro, J. E. Stone, F. Neese, K. Schulten, and Z. Luthey-Schulten. NAMD goes quantum: an integrative suite for hybrid simulations. *Nat. Methods*, 15(5):351–+, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000431372700019, doi:10.1038/nmeth.4638.

[578] Adamo Meo, Trouillas and Sancho-Garcia. *J. Chem. Phys.*, 139:164104, 2013.

[579] B. Metz, M. Schweizer, H. Stoll, M. Dolg, and W. Liu. *Theor. Chem. Acc.*, 104:22, 2000.

[580] B. Metz, H. Stoll, and M. Dolg. *J. Chem. Phys.*, 113:2563, 2000.

[581] F. Meyer and F. Neese. Impact of modern spectroscopy in inorganic chemistry. *Inorg. Chem.*, 59(19):13805–13806, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000580381700001, doi:10.1021/acs.inorgchem.0c02755.

[582] Wilfried Meyer. Ionization energies of water from pno-ci calculations. *International Journal of Quantum Chemistry*, 5(S5):341–348, 1971.

[583] Wilfried Meyer. Pno–ci studies of electron correlation effects. i. configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane. *The Journal of Chemical Physics*, 58(3):1017–1035, 1973.

[584] Wilfried Meyer. Configuration Expansion by Means of Pseudonatural Orbitals. In Henry F. Schaefer III, editor, *Methods of Electronic Structure Theory*, pages 413–446. Springer US, 01 1977.

[585] G. Mills, H. Jónsson, and G.K. Schenter. Reversible work transition state theory: application to dissociative adsorption of hydrogen. *Surf. Sci.*, 324(2-3):305–337, 1995.

[586] Yury Minenkov, Giovanni Bistoni, Christoph Riplinger, Alexander A Auer, Frank Neese, and Luigi Cavallo. Pair natural orbital and canonical coupled cluster reaction enthalpies involving light to heavy alkali and alkaline earth metals: the importance of sub-valence correlation. *Phys. Chem. Chem. Phys.*, 19(14):9374–9391, 2017.

[587] A. V. Mitin, G. Hirsch, and R. Buenker. *Chem. Phys. Lett.*, 259:151, 1996.

[588] A. V. Mitin, G. Hirsch, and R. Buenker. *J. Comput. Chem.*, 18:1200, 1997.

[589] Mariusz P. Mitoraj, Artur Michalak, and Tom Ziegler. A combined charge and energy decomposition scheme for bond analysis. *J. Chem. Theory Comput.*, 5(4):962–975, 2009.

[590] B. Mondal, F. Neese, E. Bill, and S. F. Ye. Electronic structure contributions of non-herne oxo-iron(v) complexes to the reactivity. *J. Am. Chem. Soc.*, 140(30):9531–9544, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000440877000033, doi:10.1021/jacs.8b04275.

[591] Marco Montalti, Alberto Credi, Luca Prodi, and M. Teresa Gandolfi. *Handbook of Photochemistry, Third Edition*. CRC Press, 3 edition edition, 02 2006. ISBN 978-0-8247-2377-4.

[592] K. Mori, T. P. M. Goumans, E. van Lenthe, and F. Wang. Predicting phosphorescent lifetimes and zero-field splitting of organometallic complexes with time-dependent density functional theory including spin–orbit coupling. *Phys. Chem. Chem. Phys.*, 16(28):14523–14530, 06 2014. URL: http://pubs.rsc.org/en/content/articlelanding/2014/cp/c3cp55438d, doi:10.1039/C3CP55438D.

[593] A. Moritz, X. Cao, and M. Dolg. *Theor. Chem. Acc.*, 117 & 118:473 & 845, 2007.

[594] A. Moritz and M. Dolg. *Theor. Chem. Acc.*, 121:297, 2008.

[595] D. H. Moseley, S. E. Stavretis, K. Thirunavukkuarasu, M. Ozerov, Y. Q. Cheng, L. L. Daemen, J. Ludwig, Z. G. Lu, D. Smirnov, C. M. Brown, A. Pandey, A. J. Ramirez-Cuesta, A. C. Lamb, M. Atanasov, E. Bill, F. Neese, and Z. L. Xue. Spin-phonon couplings in transition metal complexes with slow magnetic relaxation. *Nature Comm.*, 9:11, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000437101700002, doi:10.1038/s41467-018-04896-0.

[596] D. Mukherjee. *Chem. Phys. Lett.*, 274:561, 1997.

[597] R. P. Muller, J. M. Langlois, M. N. Ringnalda, R. A. Friesner, and W. A. Goddard. *J. Chem. Phys.*, 100:1226, 1994.

[598] R. S. Mulliken. Electronic population analysis on lcao–mo molecular wave functions. i. *J. Chem. Phys.*, 23(10):1833–1840, 1955. URL: https://doi.org/10.1063/1.1740588, doi:10.1063/1.1740588.

[599] R. S. Mulliken. Report on notation for the spectra of polyatomic molecules. *J. Chem. Phys.*, 23(11):1997–2011, 1955. URL: https://doi.org/10.1063/1.1740655, doi:10.1063/1.1740655.

[600] C. W. Murray, N. C. Handy, and G. J. Laming. *Mol. Phys.*, 78:997, 1993.

[601] A. Najibi, M. Casanova-Páez, and L. Goerigk. Analysis of recent BLYP- and PBE-based range-separated double-hybrid density functional approximations for main-group thermochemistry, kinetics, and noncovalent interactions. *J. Phys. Chem. A*, 125:4026–4035, 2021.

[602] A. Najibi and L. Goerigk. *J. Chem. Theory Comput.*, 14:5725, 2018.

[603] A. Najibi and L. Goerigk. *J. Comput. Chem.*, 41:2562–2572, 2020.

[604] Haruyuki Nakano. Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.*, 99(10):7983–7992, 11 1993. doi:doi:10.1063/1.465674.

[605] G. Nave, S. Johansson, R. C. M. Learner, A. P. Thorne, and J. W. Brault. *Astrophys. J., Suppl. Ser.*, 94:221, 1994.

[606] E. F. Neese, F.; Valeev. Revisiting the atomic natural orbital approach for basis sets: robust systematic basis sets for explicitly correlated and conventional correlated ab initio methods. *J. Chem. Theory Comput.*, 7:33–43, 2011.

[607] F. Neese. *Chem. Phys. Lett.*, 325:93, 2000.

[608] F. Neese. *Inorg. Chim. Acta*, 337C:181–192, 2002.

[609] F. Neese. *Chem. Phys. Lett.*, 380:721–728, 2003.

[610] F. Neese. *J. Chem. Phys.*, 122:034107, 2005.

[611] F. Neese. *J. Am. Chem. Soc.*, 128:10213, 2006.

[612] F. Neese. *J. Biol. Inorg. Chem.*, 11:702, 2006.

[613] F. Neese. *J. Chem. Phys.*, 127:164112, 2007.

[614] F. Neese. *Coordin. Chem. Rev.*, 253:526, 2009.

[615] F. Neese, M. Atanasov, G. Bistoni, D. Maganas, and S. F. Ye. Chemistry and quantum mechanics in 2019: Give us insight and numbers. *J. Am. Chem. Soc.*, 141(7):2814–2824, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459642000006, doi:10.1021/jacs.8b13313.

[616] F. Neese, A. Hansen, and D. G. Liakos. *J. Chem. Phys.*, 131:064103, 2009.

[617] F. Neese, A. Hansen, F. Wennmohs, and S. Grimme. *Acc. Chem. Res.*, 42:641, 2009.

[618] F. Neese, D. G. Liakos, and S. F. Ye. *J. Biol. Inorg. Chem.*, 16:821, 2011.

[619] F. Neese, T. Schwabe, and S. Grimme. *J. Chem. Phys.*, 126:124115, 2007.

[620] F. Neese, T. Schwabe, S. Kossmann, B. Schirmer, and S. Grimme. *J. Chem. Theory Comput.*, 5:3060, 2009.

[621] F. Neese and E. I. Solomon. *Inorg. Chem.*, 37:6568–6582, 1998.

[622] F. Neese, F. Wennmohs, and A. Hansen. *J. Chem. Phys.*, 130:114108, 2009.

[623] F. Neese, F. Wennmohs, A. Hansen, and U. Becker. *Chem. Phys.*, 356:98–109, 2009.

[624] Frank Neese. The orca program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2(1):73–78, 2012. doi:http://doi.wiley.com/10.1002/wcms.81.

[625] Frank Neese. High-level spectroscopy, quantum chemistry, and catalysis: not just a passing fad. *Angew. Chem. Int. Ed.*, 56(37):11003–11010, 2017.

[626] Frank Neese. *Quantum Chemistry and EPR Parameters*, pages 1–22. American Cancer Society, 2017. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470034590.emrstm1505, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470034590.emrstm1505, doi:https://doi.org/10.1002/9780470034590.emrstm1505.

[627] Frank Neese. Software update: the orca program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 8(1):e1327, 2018. doi:http://doi.wiley.com/10.1002/wcms.1327.

[628] Frank Neese, Taras Petrenko, Dmitry Ganyushin, and Gottfried Olbrich. Advanced aspects of ab initio theoretical optical spectroscopy of transition metal complexes: Multiplets, spin-orbit coupling and resonance Raman intensities. *Coordin. Chem. Rev.*, 251(3-4):288–327, 02 2007. doi:10.1016/j.ccr.2006.05.019.

[629] Frank Neese, Frank Wennmohs, Ute Becker, and Christoph Riplinger. The orca quantum chemistry program package. *J. Chem. Phys.*, 152(22):224108, 2020. doi:https://aip.scitation.org/doi/10.1063/5.0004608.

[630] J. Neugebauer, M. Reiher, C. Kind, and B. A. Hess. *J. Comput. Chem.*, 23:895–910, 2002.

[631] Johannes Neugebauer, Markus Reiher, Carsten Kind, and Bernd A. Hess. Quantum chemical calculation of vibrational spectra of large molecules—Raman and IR spectra for Buckminsterfullerene. *J. Comput. Chem.*, 23(9):895–910, 07 2002. URL: http://onlinelibrary.wiley.com/doi/10.1002/jcc.10089/abstract, doi:10.1002/jcc.10089.

[632] A. Nicklass, M. Dolg, H. Stoll, and H. Preuss. *J. Chem. Phys.*, 102:8942–8952, 1995.

[633] C. Nieke and J. Reinhold. *Theor. Chim. Acta*, 65:99, 1984.

[634] Hidetaka Nishimura, Kazuo Tanaka, Yasuhiro Morisaki, Yoshiki Chujo, Atsushi Wakamiya, and Yasujiro Murata. Oxygen-bridged diphenylnaphthylamine as a scaffold for full-color circularly polarized luminescent materials. *The Journal of Organic Chemistry*, 82(10):5242–5249, 2017. URL: https://www.sciencedirect.com/science/article/pii/S0022326321002929, doi:https://doi.org/10.1021/acs.joc.7b00511.

[635] J. Nocedal. Updating quasi-Newton matrices with limited storage. *Math. Comput.*, 35(151):773–782, 1980.

[636] L. Noodleman. *J. Chem. Phys.*, 74:5737, 1981.

[637] L. Noodleman and E. R. Davidson. *Chem. Phys.*, 109:131, 1986.

[638] Marcel Nooijen and Rodney J. Bartlett. A new method for excited states: Similarity transformed equation-of-motion coupled-cluster theory. *The Journal of Chemical Physics*, 106(15):6441–6448, 04 1997. URL: https://doi.org/10.1063/1.474000, arXiv:https://pubs.aip.org/aip/jcp/article-pdf/106/15/6441/19166848/6441_1_online.pdf, doi:10.1063/1.474000.

[639] Marcel Nooijen, Ondrej Demel, Dipayan Datta, Liguo Kong, K. R. Shamasundar, V. Lotrich, Lee M. Huntington, and Frank Neese. Communication: Multireference equation of motion coupled cluster: A transform and diagonalize approach to electronic structure. *J. Chem. Phys.*, 140(8):081102, 2014.

[640] D. A. Pantazis, X.-Y. Chen, C. R. Landis, and F. Neese. *J. Chem. Theory Comput.*, 4:908–919, 2008.

[641] D. A. Pantazis and F. Neese. *J. Chem. Theory Comput.*, 5:2229–2238, 2009.

[642] D. A. Pantazis and F. Neese. *J. Chem. Theory Comput.*, 7:677–684, 2011.

[643] D. A. Pantazis and F. Neese. *Theor. Chem. Acc.*, 131:1292, 2012.

[644] D. A. Pantazis, M. Orio, T. Petrenko, S. Zein, E. Bill, W. Lubitz, J. Messinger, and F. Neese. *Chem. Eur. J.*, 15:5108, 2009.

[645] Ewa Papajak and Donald G. Truhlar. Convergent partially augmented basis sets for Post-Hartree-Fock calculations of molecular properties and reaction barrier heights. *J. Chem. Theory Comput.*, 7(1):10–18, 2011. URL: https://doi.org/10.1021/ct1005533, doi:10.1021/ct1005533.

[646] R. G. Parr. *Density Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry. Oxford University Press, 1994. ISBN 978-0-19-509276-9.

[647] H. Partridge. *J. Chem. Phys.*, 87:6643, 1987.

[648] H. Partridge. *J. Chem. Phys.*, 90:1043, 1989.

[649] J. L. Pascual-Ahuir and E. Silla. *J. Comput. Chem.*, 11:1047–1060, 1990.

[650] J. L. Pascual-Ahuir, E. Silla, and I. Tunon. *J. Comput. Chem.*, 12:1077–1088, 1991.

[651] J. L. Pascual-Ahuir, E. Silla, and I. Tunon. *J. Comput. Chem.*, 15:1127–1138, 1994.

[652] Shubhrodeep Pathak, Lucas Lang, and Frank Neese. A dynamic correlation dressed complete active space method: Theory, implementation, and preliminary applications. *J. Chem. Phys.*, 147:234109, 2017.

[653] F. Pavošević, P. Pinski, C. Riplinger, F. Neese, and E.F. Valeev. *J. Chem. Phys.*, 144:144109, 2016.

[654] Fabijan Pavošević, Chong Peng, Peter Pinski, Christoph Riplinger, Frank Neese, and Edward F Valeev. Sparsemaps—a systematic infrastructure for reduced scaling electronic structure methods. v. linear scaling explicitly correlated coupled-cluster method with pair natural orbitals. *J. Chem. Phys.*, 146(17):174108, 2017.

[655] Kasper S Pedersen, Daniel N Woodruff, Saurabh Kumar Singh, Alain Tressaud, Etienne Durand, Mihail Atanasov, Panagiota Perlepe, Katharina Ollefs, Fabrice Wilhelm, Corine Mathonière, and others. [osf6] x-: molecular models for spin-orbit entangled phenomena. *Chem. Eur. J.*, 23(47):11244–11248, 2017.

[656] M. R. Pederson and S. N. Khanna. *Phys. Rev. B*, 60:9566, 1999.

[657] Andrew J. Pell, Guido Pintacuda, and Clare P. Grey. Paramagnetic NMR in solution and the solid state. *Prog. Nucl. Magn. Reson. Spectrosc.*, 2018. doi:10.1016/j.pnmrs.2018.05.001.

[658] Daoling Peng, Nils Middendorf, Florian Weigend, and Markus Reiher. An efficient implementation of two-component relativistic exact-decoupling methods for large molecules. *J. Chem. Phys.*, 138(18):184105, may 2013. URL: http://dx.doi.org/10.1063/1.4803693{]}http://aip.scitation.org/doi/10.1063/1.4803693, doi:10.1063/1.4803693.

[659] Daoling Peng and Markus Reiher. Local relativistic exact decoupling. *J. Chem. Phys.*, 136(24):244108, jun 2012. URL: http://aip.scitation.org/doi/10.1063/1.4729788, doi:10.1063/1.4729788.

[660] J. P. Perdew. *Phys. Rev. B*, 33:8822, 1986.

[661] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.*, 77:3865, 1996.

[662] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.*, 80:891, 1998.

[663] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais. *Phys. Rev. A*, 46:6671, 1992.

[664] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais. *Phys. Rev. A*, 48:4978, 1993.

[665] J. P. Perdew, A. Ruzsinsky, G. I. Csonka, L. A. Constantin, and J. Sun. *Phys. Rev. Lett.*, 103:026403, 2009.

[666] J. P. Perdew, A. Ruzsinsky, G. I. Csonka, L. A. Constantin, and J. Sun. *Phys. Rev. Lett.*, 106:179902, 2011.

[667] J. P. Perdew and Y. Wang. *Phys. Rev. B*, 45:13244, 1992.

[668] J. P. Perdew and Yue Wang. *Phys. Rev. B*, 33:8800, 1986.

[669] J. P. Perdew and Wang Yue. *Phys. Rev. B*, 40:3399, 1986.

[670] K. A. Peterson. *J. Chem. Phys.*, 119:11099, 2003.

[671] K. A. Peterson, D. Figgen, M. Dolg, and H. Stoll. *J. Chem. Phys.*, 126:124101, 2007.

[672] K. A. Peterson, D. Figgen, E. Goll, H. Stoll, and M. Dolg. *J. Chem. Phys.*, 119:11113, 2003.

[673] K. A. Peterson and C. Puzzarini. *Theor. Chem. Acc.*, 114:283, 2005.

[674] K. A. Peterson, B. C. Shepler, D. Figgen, and H. Stoll. *J. Phys. Chem. A*, 110:13877, 2006.

[675] T. Petrenko, S. Kossmann, and F. Neese. *J. Chem. Phys.*, 134:054116, 2011.

[676] Taras Petrenko, Serena DeBeer-George, Núria Aliaga-Alcalde, Eckhard Bill, Bernd Mienert, Yuming Xiao, YiSong Guo, Wolfgang Sturhahn, Stephen P. Cramer, Karl Wieghardt, and Frank Neese. Characterization of a genuine iron(v)-nitrido species by nuclear resonant vibrational spectroscopy coupled to density functional calculations. *J. Am. Chem. Soc.*, 129:11053–11060, 2007.

[677] Taras Petrenko, Simone Kossmann, and Frank Neese. Efficient time-dependent density functional theory approximations for hybrid density functionals: Analytical gradients and parallelization. *J. Chem. Phys.*, 134(5):054116, 02 2011. URL: http://aip.scitation.org/doi/abs/10.1063/1.3533441, doi:10.1063/1.3533441.

[678] Taras Petrenko and Frank Neese. Analysis and prediction of absorption band shapes, fluorescence band shapes, resonance Raman intensities, and excitation profiles using the time-dependent theory of electronic spectroscopy. *J. Chem. Phys.*, 127(16):164319, 10 2007. URL: https://doi.org/10.1063/1.2770706, doi:10.1063/1.2770706.

[679] Taras Petrenko, Wolfgang Sturhahn, and Frank Neese. *Hyperfine Interact.*, 175:165, 2007.

[680] R. A. Pierotti. *Chem. Rev.*, 76:717, 1976.

[681] W. J. Pietro, E. S. Blurock, R. F. Hout, Hehrem W. J., D. J. DeFrees, and R. F. Stewart. *Inorg. Chem.*, 20:3650, 1981.

[682] W. J. Pietro and W. J. Hehre. *J. Comput. Chem.*, 4:241, 1983.

[683] W. J. Pietro, B. A. Levi, W. J. Hehre, and R. F. Stewart. *Inorg. Chem.*, 19:2225, 1980.

[684] P. Pinski, C. Riplinger, E. F. Valeev, and Frank Neese. *J. Chem. Phys.*, 143:034108, 2015.

[685] Peter Pinski and Frank Neese. Communication: Exact analytical derivatives for the domain-based local pair natural orbital MP2 method (DLPNO-MP2). *J. Chem. Phys.*, 148:031101, 2018. doi:10.1063/1.5011204.

[686] Peter Pinski and Frank Neese. Analytical gradient for the domain-based local pair natural orbital second order Møller-Plesset perturbation theory method (DLPNO-MP2). *J. Chem. Phys.*, 150:164102, 2019.

[687] F. Plasser, M. Wormit, and A. Dreuw. *J. Chem. Phys.*, 141:024106, 2014.

[688] Christoph Plett, Marcel Stahn, Markus Bursch, Jan-Michael Mewes, and Stefan Grimme. Improving quantum chemical solvation models by dynamic radii adjustment for continuum solvation method (DRACO). *J. Phys. Chem. Lett.*, 15:2462, 2024.

[689] Patrik Pollak and Florian Weigend. Segmented contracted error-consistent basis sets of double- and triple-ζ valence quality for one- and two-component relativistic all-electron calculations. *J. Chem. Theory Comput.*, 13(8):3696–3705, 2017. URL: https://doi.org/10.1021/acs.jctc.7b00593, doi:10.1021/acs.jctc.7b00593.

[690] Christopher J. Pollock, Mario Ulises Delgado-Jaime, Mihail Atanasov, Frank Neese, and Serena DeBeer. Kβ mainline x-ray emission spectroscopy as an experimental probe of Metal–Ligand covalency. *J. Am. Chem. Soc.*, 136(26):9453–9463, 2014. URL: https://pubmed.ncbi.nlm.nih.gov/24914450/, doi:https://doi.org/10.1021/ja504182n.

[691] J. A. Pople and D. L. Beveridge. *Approximate Molecular Orbital Theory*. McGraw Hill Inc, 1970.

[692] J. A. Pople, D. L. Beveridge, and P. A. Dobosh. *J. Chem. Phys.*, 47:2026, 1967.

[693] J. A. Pople, J. S. Binkley, and R. Seeger. *Int. J. Quant. Chem. Symp.*, 10:1, 1976.

[694] J. A. Pople, P. M. W. Gill, and B. G. Johnson. *Chem. Phys. Lett.*, 199:557, 1992.

[695] J. A. Pople and G. A. Segal. *J. Chem. Phys.*, 43:136, 1965.

[696] J. A. Pople and G. A. Segal. *J. Chem. Phys.*, 44:3289, 1966.

[697] Philipp Pracht, Fabian Bohle, and Stefan Grimme. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Physical Chemistry Chemical Physics*, 22(14):7169–7192, April 2020. Publisher: The Royal Society of Chemistry. URL: https://pubs.rsc.org/en/content/articlelanding/2020/cp/c9cp06869d (visited on 2022-04-22), doi:10.1039/C9CP06869D.

[698] Philipp Pracht, Eike Caldeweyher, Sebastian Ehlert, and Stefan Grimme. A robust non-self-consistent tight-binding quantum chemistry method for large molecules. *ChemRxiv*, pages DOI: 10.26434/chemrxiv.8326202.v1, 2019.

[699] Philipp Pracht and Stefan Grimme. Calculation of absolute molecular entropies and heat capacities made simple. *Chemical Science*, 12(19):6551–6568, May 2021. Publisher: The Royal Society of Chemistry. URL: https://pubs.rsc.org/en/content/articlelanding/2021/sc/d1sc00621e (visited on 2022-08-16), doi:10.1039/D1SC00621E.

[700] Philipp Pracht, Rainer Wilcken, Anikó Udvarhelyi, Stephane Rodde, and Stefan Grimme. High accuracy quantum-chemistry-based calculation and blind prediction of macroscopic pKa values in the context of the SAMPL6 challenge. *J. Comput.-Aided Mol. Des.*, 08 2018. doi:10.1007/s10822-018-0145-7.

[701] P. Pulay. ImprovedSCF convergence acceleration. *J. Comput. Chem.*, 3(4):556–560, 0024. URL: http://doi.wiley.com/10.1002/jcc.540030413 (visited on 2021-02-24), doi:10.1002/jcc.540030413.

[702] P. Pulay. *Chem. Phys. Lett.*, 73:393, 1980.

[703] P. Pulay. *J. Comput. Chem.*, 3:556, 1982.

[704] P. Pulay, S. Saebo, and W. Meyer. *J. Chem. Phys.*, 81:1901, 1984.

[705] C. C. Pye and T. Ziegler. *Theor. Chem. Acc.*, 101:396, 1999.

[706] José M Pérez-Jordá and Weitao Yang. A concise redefinition of the solid spherical harmonics and its use in fast multipole methods. *The Journal of chemical physics*, 104(20):8003–8006, 1996.

[707] M. Römelt, S. Ye, and F. Neese. *Inorg. Chem.*, 48:784, 2009.

[708] E. Ramos-Cordoba, E. Matito, I. Mayer, and P. Salvador. *J. Chem. Theory Comput.*, 8:1270, 2012.

[709] Dmitrij Rappoport. Property-optimized Gaussian basis sets for lanthanides. *J. Chem. Phys.*, 155:124102, 2021. URL: https://doi.org/10.1063/5.0065611, doi:10.1063/5.0065611.

[710] Dmitrij Rappoport and Filipp Furche. *J. Chem. Phys.*, 133:134105, 2010.

[711] V. Rassolov, J. A. Pople, M. Ratner, and T. L. Windus. *J. Chem. Phys.*, accepted 1998.

[712] K. Ray, S. DeBeer-George, E. I. Solomon, K. Wieghardt, and F. Neese. *Chem. Eur. J.*, 13:2783, 2007.

[713] Sarah Reimann, Ulf Ekström, Stella Stopkowicz, Andrew M Teale, Alex Borgoo, and Trygve Helgaker. The importance of current contributions to shielding constants in density-functional theory. *Phys. Chem. Chem. Phys.*, 17(28):18834–18842, 2015. URL: http://pubs.rsc.org/en/content/articlehtml/2015/cp/c5cp02682b, arXiv:26123927, doi:10.1039/C5CP02682B.

[714] Jeffrey R. Reimers. A practical method for the use of curvilinear coordinates in calculations of normal-mode-projected displacements and Duschinsky rotation matrices for large molecules. *J. Chem. Phys.*, 115(20):9103–9109, 11 2001. URL: http://aip.scitation.org/doi/abs/10.1063/1.1412875, doi:10.1063/1.1412875.

[715] Kevin Reiter, Michael Kühn, and Florian Weigend. Vibrational circular dichroism spectra for large molecules and molecules with heavy elements. *The Journal of Chemical Physics*, 146(5):054102, feb 2017. URL: http://aip.scitation.org/doi/10.1063/1.4974897, doi:10.1063/1.4974897.

[716] Marius Retegan, Nicholas Cox, Dimitrios A. Pantazis, and Frank Neese. A First-Principles Approach to the Calculation of the on-Site Zero-Field Splitting in Polynuclear Transition Metal Complexes. *Inorg. Chem.*, 53(21):11785–11793, 11 2014. doi:10.1021/ic502081c.

[717] Young Min Rhee and Martin Head-Gordon. *J. Phys. Chem. A*, 111:5314–5326, 2007.

[718] J. Ribas-Arino and D. Marx. *Chem. Rev.*, 112(10):5412–5487, 2012.

[719] J. Ridley and M. C. Zerner. *Theor. Chim. Acta*, 32:111, 1973.

[720] C. Riplinger and F. Neese. *J. Chem. Phys.*, 138:034106, 2013.

[721] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, and Frank Neese. *J. Chem. Phys.*, 144:024109, 2016.

[722] C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese. *J. Chem. Phys.*, 139:134101, 2013.

[723] Christoph Riplinger, Joseph P. Y. Kao, Gerald M. Rosen, Velavan Kathirvelu, Gareth R. Eaton, Sandra S. Eaton, Andrei Kutateladze, and Frank Neese. *J. Am. Chem. Soc.*, 131:10092, 2009.

[724] T. Risthaus, A. Hansen, and S. Grimme. *Phys. Chem. Chem. Phys.*, 16:14408–14419, 2014.

[725] Melvin B. Robin. *Higher Excited States of Polyatomic Molecules*. Academic Press, 1974. ISBN 978-0-12-589901-7.

[726] M. Roemelt and F. Neese. *J. Phys. Chem. A*, 117:3069–3082, 2013.

[727] J. D. Rolfes, M. van Gastel, and F. Neese. Where is the fluoro wall?: A quantum chemical investigation. *Inorg. Chem.*, 59(2):1556–1565, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000509420100065, doi:10.1021/acs.inorgchem.9b03474.

[728] Julian D. Rolfes, Frank Neese, and Dimitrios A. Pantazis. All-electron scalar relativistic basis sets for the elements Rb–Xe. *J. Comput. Chem.*, 41:1842–1849, 2020. doi:10.1002/jcc.26355.

[729] M Rolik, Z.; Kallay. A general-order local coupled-cluster method based on the cluster-in-molecule approach. *J. Chem. Phys.*, 135:104111, 2011. doi:https://doi.org/10.1063/1.3632085.

[730] C. Romelt, S. F. Ye, E. Bill, T. Weyhermuller, M. van Gastel, and F. Neese. Electronic structure and spin multiplicity of iron tetraphenylporphyrins in their reduced states as determined by a combination of resonance raman spectroscopy and quantum chemistry. *Inorg. Chem.*, 57(4):2141–2148, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000426014800048, doi:10.1021/acs.inorgchem.7b03018.

[731] Björn O. Roos and Kerstin Andersson. Multiconfigurational perturbation theory with level shift — the Cr2 potential revisited. *Chem. Phys. Lett.*, 245:215–223, 10 1995. doi:https://doi.org/10.1016/0009-2614(95)01010-7.

[732] L. E. Roy, J. Hay, and R. L. Martin. *J. Chem. Theory Comput.*, 4:1029, 2008.

[733] E. Ruiz, J. Cano, S. Alvarez, and P. Alemany. *J. Comput. Chem.*, 20:1391, 1999.

[734] Paweł Sał ek, Stinne Hø st, Lea Thø gersen, Poul Jø rgensen, Pekka Manninen, Jeppe Olsen, Branislav Jansík, Simen Reine, Filip Pawł owski, Erik Tellgren, Trygve Helgaker, and Sonia Coriani. Linear-scaling implementation of molecular electronic self-consistent field theory. *J. Chem. Phys.*, 126(11):114110, 2007.

[735] Ali Sadeghi, S. Alireza Ghasemi, Bastian Schaefer, Stephan Mohr, Markus A. Lill, and Stefan Goedecker. Metrics for measuring distances in configuration spaces. *The Journal of Chemical Physics*, 139(18):184118, November 2013. Publisher: American Institute of Physics. URL: https://aip.scitation.org/doi/abs/10.1063/1.4828704 (visited on 2022-04-25), doi:10.1063/1.4828704.

[736] S. Saebo and J. Almlöf. *Chem. Phys. Lett.*, 154:83, 1989.

[737] T. Saito, S. Nishihara, Y. Kataoka, Y. Nakanishi, Y. Kitagawa, T. Kawakami, S. Yamanaka, M. Okumura, and K. Yamaguchi. *J. Phys. Chem. A*, 114:7967, 2010.

[738] T. Saito and W. Thiel. *J. Phys. Chem. A*, 116:10864, 2012.

[739] M. Saitow, U. Becker, C. Riplinger, E. F. Valeev, and F. Neese. *J. Chem. Phys.*, 146:164105, 2017.

[740] M. Saitow, A. K. Dutta, and F. Neese. Accurate ionization potentials, electron affinities and electronegativities of single-walled carbon nanotubes by state-of-the-art local coupled-cluster theory. *Bull. Chem. Soc. Jpn.*, 92(1):170–174, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000455409900003, doi:10.1246/bcsj.20180254.

[741] Masaaki Saitow, Yuki Kurashige, and Takeshi Yanai. Multireference configuration interaction theory using cumulant reconstruction with internal contraction of density matrix renormalization group wave function. *J. Chem. Phys.*, 139:044118, 07 2013. doi:10.1063/1.4816627.

[742] Masaaki Saitow and Frank Neese. Accurate spin-densities based on the domain-based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.*, 149:034104, 2018. doi:10.1063/1.5027114.

[743] C. A. M. Salla, J. T. dos Santos, G. Farias, A. J. Bortoluzi, S. F. Curcio, T. Cazati, R. Izsak, F. Neese, B. de Souza, and I. H. Bechtold. New Boron(III) blue emitters for all-solution processed OLEDs: Molecular design assisted by theoretical modeling. *Eur. J. Inorg. Chem.*, pages 2247–2257, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000471302400001, doi:10.1002/ejic.201900265.

[744] E. A. Salter, G. W. Trucks, and R. J. Bartlett. *J. Chem. Phys.*, 90:1752, 1989.

[745] B. Sandhoefer and F. Neese. *J. Chem. Phys.*, 137:094102, 2012.

[746] Gerald M. Sando and Kenneth G. Spears. Ab Initio Computation of the Duschinsky Mixing of Vibrations and Nonlinear Effects. *J. Phys. Chem. A*, 105(22):5326–5333, 06 2001. URL: http://dx.doi.org/10.1021/jp004230b, doi:10.1021/jp004230b.

[747] Fabrizio Santoro, Roberto Improta, Alessandro Lami, Julien Bloino, and Vincenzo Barone. Effective method to compute Franck-Condon integrals for optical spectra of large molecules in solution. *The Journal of Chemical Physics*, 126(8):084509, 02 2007. URL: https://doi.org/10.1063/1.2437197, arXiv:https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.2437197/15395332/084509_1_online.pdf, doi:10.1063/1.2437197.

[748] D. P. Santry. *J. Am. Chem. Soc.*, 90:3309, 1968.

[749] D. P. Santry and G. A. Segal. *J. Chem. Phys.*, 47:158, 1967.

[750] V. R. Saunders and I. H. Hillier. *Int. J. Quant. Chem.*, VII:699, 1973.

[751] Elvira R. Sayfutyarova and Sharon Hammes-Schiffer. Constructing molecular π-orbital active spaces for multireference calculations of conjugated systems. *J. Chem. Theory Comput.*, 15(3):1679–1689, March 2019. URL: https://doi.org/10.1021/acs.jctc.8b01196, doi:10.1021/acs.jctc.8b01196.

[752] Elvira R. Sayfutyarova, Qiming Sun, Garnet Kin-Lic Chan, and Gerald Knizia. Automated construction of molecular active spaces from atomic valence orbitals. *J. Chem. Theory Comput.*, 13(9):4063–4078, September 2017. URL: https://doi.org/10.1021/acs.jctc.7b00128, doi:10.1021/acs.jctc.7b00128.

[753] A. Schäfer, H. Horn, and R. J. Ahlrichs. *Chem. Phys.*, 97:2571, 1992.

[754] M. Schütz and H. J. Werner. *Chem. Phys. Lett.*, 318:370, 2000.

[755] M. Schütz and H. J. Werner. *J. Chem. Phys.*, 114:661, 2001.

[756] Igor Schapiro, Kantharuban Sivalingam, and Frank Neese. Assessment of n-Electron Valence State Perturbation Theory for Vertical Excitation Energies. *J. Chem. Theory Comput.*, 9(8):3567–3580, 08 2013. doi:10.1021/ct400136y.

[757] Caspar Jonas Schattenberg and Martin Kaupp. Effect of the current dependence of tau-dependent exchange-correlation functionals on nuclear shielding calculations. *J. Chem. Theory Comput.*, 17(3):1469–1479, 2021. URL: https://dx.doi.org/10.1021/acs.jctc.0c01223, doi:10.1021/acs.jctc.0c01223.

[758] F. Schautz, H.-J. Flad, and M. Dolg. *Theor. Chem. Acc.*, 99:231, 1998.

[759] B. Scheibe, C. Pietzonka, O. Mustonen, M. Karppinen, A. J. Karttunen, M. Atanasov, F. Neese, M. Conrad, and F. Kraus. The U2F12 (2-) anion of sr U2F12. *Angew. Chem. Int. Ed.*, 57(11):2914–2918, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000426490700027, doi:10.1002/anie.201800743.

[760] P. Scheurer and W. H. E. Schwarz. Continuous degeneracy of sets of localized orbitals. *Int. J. Quantum Chem.*, 76:428–433, 2000.

[761] B. Schimmelpfennig. AMFI - an atomic mean-field spin-orbit integral program. 1996.

[762] H. B. Schlegel. In K. P. Lawley, editor, *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry, Part I*, volume 67, pages 249. John Wiley and Sons, 1987.

[763] H. B. Schlegel. In D. R. Yarkony, editor, *Modern Electronic Structure Theory*, pages 459. World Scientific, 1995.

[764] H. B. Schlegel. In P. v. R. Schleyer, editor, *Encyclopedia of Computational Chemistry*, pages 1136. John Wiley and Sons, 1998.

[765] Y. L. A. Schmerwitz, V. Ásgeirsson, and H. Jónsson. Improved initialization of optimal path calculations using sequential traversal over the image dependent pair potential surface. *arXiv:2310.04531*, 2023. URL: https://arxiv.org/abs/2310.04531.

[766] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. J. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery. *J. Comput. Chem.*, 14:1347, 1993.

[767] W. Schneider, G. Bistoni, M. Sparta., C. Riplinger, M. Saitow, A. Auer, and F. Neese. Decomposition of intermolecular interaction energies within the local pair natural orbital coupled cluster framework. *J. Chem. Theory Comput.*, 12(10):4778–4792, 2016. doi:10.1021/acs.jctc.6b00523.

[768] C. E. Schulz, R. G. Castillo, D. A. Pantazis, S. DeBeer, and F. Neese. Structure-spectroscopy correlations for intermediate q of soluble methane monooxygenase: Insights from QM/MM calculations. *J. Am. Chem. Soc.*, 143(17):6560–6577, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000648704100032, doi:10.1021/jacs.1c01180.

[769] C. E. Schulz, M. van Gastel, D. A. Pantazis, and F. Neese. Converged structural and spectroscopic properties for refined qm/mm models of azurin. *Inorg. Chem.*, 60(10):7399–7412, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000653539100063, doi:10.1021/acs.inorgchem.1c00640.

[770] C. J. H. Schutte, J. E. Bertie, P. R. Bunker, J. T. Hougen, I. M. Mills, J. K. G. Watson, and B. P. Winnewisser. Notations and conventions in molecular spectroscopy: part 2. symmetry notation (iupac recommendations 1997). *Pure & Appl. Chem.*, 69(8):1641–1649, 1997. URL: https://doi.org/10.1351/pac199769081641, doi:10.1351/pac199769081641.

[771] T. Schwabe and L. Goerigk. *J. Chem. Theory Comput.*, 13:4307, 2017.

[772] T. Schwabe and S. Grimme. *Phys. Chem. Chem. Phys.*, 8:4398, 2006.

[773] P. Schwerdtfeger, M. Dolg, W. H. E. Schwarz, G. A. Bowmaker, and P. D. W. Boyd. *J. Chem. Phys.*, 91:1762–1774, 1989.

[774] Martin Schütz. A new, fast, semi-direct implementation of linear scaling local coupled cluster theory. *Physical Chemistry Chemical Physics*, 4(16):3941–3947, 2002.

[775] Martin Schütz and Frederick R Manby. Linear scaling local coupled cluster theory with density fitting. part i: 4-external integrals. *Physical Chemistry Chemical Physics*, 5(16):3349–3358, 2003.

[776] G. E. Scuseria and H. F. Schaefer III. *Chem. Phys. Lett.*, 142:354, 1987.

[777] G. E. Scuseria, C. L. Janssen, and H. F. Schaefer III. *J. Chem. Phys.*, 89:7382, 1988.

[778] J. Sedlej and I. L. Cooper. *Semi-Emipirical Methods of Quantum Chemistry*. 1985, John Wiley and Sons.

[779] R. Seeger and J. A. Pople. *J. Chem. Phys.*, 66:3045, 1977.

[780] Emmanouil Semidalas and Jan M. L. Martin. Automatic generation of complementary auxiliary basis sets for explicitly correlated methods. *Journal of Computational Chemistry*, 43(25):1690–1700, 2022. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.26970, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.26970, doi:https://doi.org/10.1002/jcc.26970.

[781] Avijit Sen, Bernardo de Souza, Lee M. J. Huntington, Martin Krupička, Frank Neese, and Róbert Izsák. An efficient pair natural orbital based configuration interaction scheme for the calculation of open-shell ionization potentials. *J. Chem. Phys.*, 149(11):114108, 2018.

[782] Robert Send and Filipp Furche. First-order nonadiabatic couplings from time-dependent hybrid density functional response theory: Consistent formalism, implementation, and performance. *J. Chem. Phys.*, 132(4):044107, 01 2010. URL: https://aip.scitation.org/doi/10.1063/1.3292571 (visited on 2020-06-08), doi:10.1063/1.3292571.

[783] Carlos Serpa, Luis G. Arnaut, Sebastião J. Formosinho, and K. Razi Naqvi. Calculation of triplet–triplet energy transfer rates from emission and absorption spectra. The quenching of hemicarcerated triplet biacetyl by aromatic hydrocarbons. *Photochem. Photobiol. Sci.*, 2(5):616–623, 05 2003. doi:10.1039/B300049D.

[784] Rami Shafei, Ai Hamano, Christophe Gourlaouen, Dimitrios Maganas, Keiko Takano, Chantal Daniel, and Frank Neese. Theoretical spectroscopy for unraveling the intensity mechanism of the optical and photoluminescent spectra of chiral re(i) transition metal complexes. *The Journal of Chemical Physics*, 159(8):084102, 2023. URL: https://doi.org/10.1063/5.0153742.

[785] Sason Shaik, Rajeev Ramanan, David Danovich, and Debasish Mandal. Structure and reactivity/selectivity control by oriented-external electric fields. *Chem. Soc. Rev.*, 47:5125–5145, 2018. URL: http://dx.doi.org/10.1039/C8CS00354H, doi:10.1039/C8CS00354H.

[786] Tareq M. Shami, Ayman A. El-Saleh, Mohammed Alswaitti, Qasem Al-Tashi, Mhd Amen Summakieh, and Seyedali Mirjalili. Particle Swarm Optimization: A Comprehensive Survey. *IEEE Access*, 10:10031–10061, 2022. URL: https://ieeexplore.ieee.org/document/9680690 (visited on 2024-06-25), doi:10.1109/ACCESS.2022.3142859.

[787] Honghui Shang and Jinlong Yang. The Moving-Grid Effect in the Harmonic Vibrational Frequency Calculations with Numeric Atom-Centered Orbitals. *The Journal of Physical Chemistry A*, 124(14):2897–2906, April 2020. Publisher: American Chemical Society. URL: https://doi.org/10.1021/acs.jpca.0c01453 (visited on 2023-07-13), doi:10.1021/acs.jpca.0c01453.

[788] S. Sharma and G. K.-L. Chan. *J. Chem. Phys.*, 136:124121, 2012.

[789] I. Shavir and L. T. Redmon. *J. Chem. Phys.*, 73:5711, 1980.

[790] James Shee, Matthias Loipersberger, Adam Rettig, Joonho Lee, and Martin Head-Gordon. Regularized second-order møller–plesset theory: a more accurate alternative to conventional MP2 for noncovalent interactions and transition metal thermochemistry for the same computational cost. *The Journal of Physical Chemistry Letters*, 12(50):12084–12097, December 2021. URL: https://doi.org/10.1021/acs.jpclett.1c03468, doi:10.1021/acs.jpclett.1c03468.

[791] D. Sheppard, R. Terrell, and G. Henkelman. Optimization methods for finding minimum energy paths. *J. Chem. Phys.*, 128(13):134106, 2008.

[792] Toru Shiozaki, Werner Győrffy, Paolo Celani, and Hans-Joachim Werner. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.*, 135:081106–081106–4, 08 2011.

[793] R. G. Shirazi, F. Neese, and D. A. Pantazis. Accurate spin-state energetics for aryl carbenes. *J. Chem. Theory Comput.*, 14(9):4733–4746, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000444792700020, doi:10.1021/acs.jctc.8b00587.

[794] R. G. Shirazi, F. Neese, D. A. Pantazis, and G. Bistoni. Physical nature of differential spin-state stabilization of carbenes by hydrogen and halogen bonding: A domain-based pair natural orbital coupled cluster study. *J. Phys. Chem. A*, 123(24):5081–5090, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000472800600009, doi:10.1021/acs.jpca.9b01051.

[795] R. G. Shirazi, D. A. Pantazis, and F. Neese. Performance of density functional theory and orbital-optimised second-order perturbation theory methods for geometries and singlet-triplet state splittings of aryl-carbenes. *Mol. Phys.*, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000535113400001, doi:10.1080/00268976.2020.1764644.

[796] Per E. M. Siegbahn. Direct configuration interaction with a reference state composed of many reference configurations. *Int. J. Quant. Chem.*, 18(5):1229–1242, 11 1980. doi:10.1002/qua.560180510.

[797] Saurabh Kumar Singh, Mihail Atanasov, and Frank Neese. Challenges in multireference perturbation theory for the calculations of the g-tensor of first-row transition-metal complexes. *J. Chem. Theory Comput.*, 14(9):4662–4677, 2018.

[798] Saurabh Kumar Singh, Julien Eng, Mihail Atanasov, and Frank Neese. Covalency and chemical bonding in transition metal complexes: An ab initio based ligand field perspective. *Coordin. Chem. Rev.*, 344:2–25, 08 2017. doi:10.1016/j.ccr.2017.03.018.

[799] S. Sinnecker and F. Neese. *J. Phys. Chem. A*, 110:12267, 2006.

[800] A. Sirohiwal, F. Neese, and D. A. Pantazis. Microsolvation of the redox-active tyrosine-d in photosystem II: Correlation of energetics with EPR spectroscopy and oxidation-induced proton transfer. *J. Am. Chem. Soc.*, 141(7):3217–3231, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459642000056, doi:10.1021/jacs.8b13123.

[801] A. Sirohiwal, F. Neese, and D. A. Pantazis. Protein matrix control of reaction center excitation in photosystem II. *J. Am. Chem. Soc.*, 142(42):18174–18190, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000580559000041, doi:10.1021/jacs.0c08526.

[802] A. Sirohiwal, F. Neese, and D. A. Pantazis. Chlorophyll excitation energies and structural stability of the cp47 antenna of photosystem ii: a case study in the first-principles simulation of light-harvesting complexes. *Chem. Sci.*, 12(12):4463–4476, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000635768300025, doi:10.1039/d0sc06616h.

[803] A. Sirohiwal, F. Neese, and D. A. Pantazis. How can we predict accurate electrochromic shifts for biochromophores? a case study on the photosynthetic reaction center. *J. Chem. Theory Comput.*, 17(3):1858–1873, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000629135700044, doi:10.1021/acs.jctc.0c01152.

[804] Abhishek Sirohiwal, Romain Berraud-Pache, Frank Neese, Róbert Izsák, and Dimitrios A. Pantazis. Accurate computation of the absorption spectrum of chlorophyll a with pair natural orbital coupled cluster methods. *J. Phys. Chem. B*, 124(40):8761–8771, 2020. URL: https://doi.org/10.1021/acs.jpcb.0c05761, doi:10.1021/acs.jpcb.0c05761.

[805] Kantharuban Sivalingam, Martin Krupička, Alexander A. Auer, and Frank Neese. Comparison of fully internally and strongly contracted multireference configuration interaction procedures. *J. Chem. Phys.*, 145(5):054104, 08 2016. doi:10.1063/1.4959029.

[806] J. C. Slater. *The Quantum Theory of Atoms Molecules and Solids, Vol. 4*. McGraw Hill, New York, 1974.

[807] S. Smidstrup, A. Pedersen, K Stokbro, and H. and Jónsson. Improved initial guess for minimum energy path calculations. *J. Chem. Phys.*, 140(21):214106, 2014.

[808] T. Soda, Y. Kitagawa, T. Onishi, Y. Takano, Y. Shigeta, H. Nagao, Y. Yoshioka, and K. Yamaguchi. *Chem. Phys. Lett.*, 319:223, 2000.

[809] Alessandro Soncini and Willem Van den Heuvel. Communication: Paramagnetic NMR chemical shift in a spin state subject to zero-field splitting. *J. Chem. Phys.*, 138:021103, 2013. doi:doi:10.1063/1.4775809.

[810] L. K. Sorensen, M. Guo, R. Lindh, and M. Lundberg. *J. Chem. Theory Comput.*, 115:174, 2016.

[811] Manuel Sparta, Marius Retegan, Peter Pinski, Christoph Riplinger, Ute Becker, and Frank Neese. Multilevel approaches within the local pair natural orbital framework. *J. Chem. Theory Comput.*, 13(7):3198–3207, 07 2017. URL: https://pubs.acs.org/doi/10.1021/acs.jctc.7b00260, arXiv:28590754, doi:10.1021/acs.jctc.7b00260.

[812] Sebastian Spicher, Christoph Plett, Philipp Pracht, Andreas Hansen, and Stefan Grimme. Automated Molecular Cluster Growing for Explicit Solvation by Efficient Force Field and Tight Binding Methods. *Journal of Chemical Theory and Computation*, 18(5):3174–3189, May 2022. Publisher: American Chemical Society. URL: https://doi.org/10.1021/acs.jctc.2c00239, doi:10.1021/acs.jctc.2c00239.

[813] N. Spiller, V. G. Chilkuri, S. DeBeer, and F. Neese. Sulfur vs. Selenium as bridging ligand in di-iron complexes: A theoretical analysis. *Eur. J. Inorg. Chem.*, 2020(15-16):1525–1538, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000520715700001, doi:10.1002/ejic.202000033.

[814] M. Stahn, S. Ehlert, and S. Grimme. Extended conductor-like polarizable continuum solvation model (cpcm-x) for semiempirical methods. *J. Phys. Chem. A*, 127:7036–7043, 2023. doi:https://doi.org/10.1021/acs.jpca.3c04382.

[815] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew. *J. Chem. Phys.*, 119:12129, 2003.

[816] S. E. Stavretis, Y. Q. Cheng, L. L. Daemen, C. M. Brown, D. H. Moseley, E. Bill, M. Atanasov, A. J. Ramirez-Cuesta, F. Neese, and Z. L. Xue. Probing magnetic excitations in co-ii single-molecule magnets by inelastic neutron scattering. *Eur. J. Inorg. Chem.*, pages 1119–1127, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000459919200008, doi:10.1002/ejic.201801088.

[817] K. K. Stavrev and M. C. Zerner. Spin-averaged hartree–fock procedure for spectroscopic calculations: the absorption spectrum of mn2+ in zns crystals. *Int. J. Quant. Chem.*, 65(5):877–884, 1997. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-461X%281997%2965%3A5%3C877%3A%3AAID-QUA51%3E3.0.CO%3B2-T, doi:https://doi.org/10.1002/(SICI)1097-461X(1997)65:5<877::AID-QUA51>3.0.CO;2-T.

[818] Peter J. Steinbach and Bernard R. Brooks. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.*, 15(7):667–683, 1994. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540150702, doi:10.1002/jcc.540150702.

[819] Johannes Steinmetzer, Stephan Kupfer, and Stefanie Gräfe. Pysisyphus: Exploring potential energy surfaces in ground and excited states. *International Journal of Quantum Chemistry*, 121(3):e26390, 2021. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.26390, doi:10.1002/qua.26390.

[820] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, 98:11623, 1994. URL: https://pubs.acs.org/doi/pdf/10.1021/j100096a001, doi:https://doi.org/10.1021/j100096a001.

[821] J. P. Stewart. *J. Comput. Chem.*, 10:209 & 221, 1989.

[822] R. F. Stewart. *J. Chem. Phys.*, 50:2485, 1969.

[823] H. Stoll, B. Metz, and M. Dolg. *J. Comput. Chem.*, 23:767, 2002.

[824] H. Stoll, P. Schwerdtfeger P. Fuentealba, J. Flad, L. von Szentpaly, and H. Preuss. *J. Chem. Phys.*, 81:2732–2736, 1984.

[825] Georgi L Stoychev, Alexander A Auer, Róbert Izsák, and Frank Neese. Self-consistent field calculation of nuclear magnetic resonance chemical shielding constants using gauge-including atomic orbitals and approximate two-electron integrals. *J. Chem. Theory Comput.*, 14(2):619–637, 2018. URL: http://pubs.acs.org/doi/10.1021/acs.jctc.7b01006, doi:10.1021/acs.jctc.7b01006.

[826] Georgi L. Stoychev, Alexander A. Auer, Jürgen Gauss, and Frank Neese. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *J. Chem. Phys.*, 154(16):164110, 2021. URL: https://aip.scitation.org/doi/10.1063/5.0047125, doi:10.1063/5.0047125.

[827] Georgi L. Stoychev, Alexander A. Auer, and Frank Neese. Automatic generation of auxiliary basis sets. *J. Chem. Theory Comput.*, 13(2):554, 2017. URL: https://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b01041, doi:https://doi.org/10.1021/acs.jctc.6b01041.

[828] Georgi L. Stoychev, Alexander A. Auer, and Frank Neese. Efficient and accurate prediction of nuclear magnetic resonance shielding tensors with double-hybrid density functional theory. *J. Chem. Theory Comput.*, 14(9):4756–4771, 09 2018. URL: http://pubs.acs.org/doi/10.1021/acs.jctc.8b00624, doi:10.1021/acs.jctc.8b00624.

[829] R. E. Stratmann, G. E. Scuseria, and M. J. Frisch. *Chem. Phys. Lett.*, 257:213, 1996.

[830] S. J. Strickler and Robert A. Berg. Relationship between Absorption Intensity and Fluorescence Lifetime of Molecules. *J. Chem. Phys.*, 37(4):814–822, 08 1962. URL: http://aip.scitation.org/doi/abs/10.1063/1.1733166 (visited on 2017-12-26), doi:10.1063/1.1733166.

[831] A. H. Stroud. *Approximate Calculation of Multiple Integrals*. Prentice-Hall, Englewood Cliffs, 1971.

[832] Jianwei Sun, Adrienn Ruzsinszky, and John P. Perdew. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.*, 115(3):036402, 07 2015. doi:10.1103/PhysRevLett.115.036402.

[833] R. Sure and S. Grimme. *J. Comput. Chem.*, 34:1672–1685, 2013.

[834] E. A. Suturina, D. Maganas, E. Bill, M. Atanasov, and F. Neese. *Inorg. Chem.*, 54:9948–9961, 2015.

[835] Elizaveta A Suturina, Joscha Nehrkorn, Joseph M Zadrozny, Junjie Liu, Mihail Atanasov, Thomas Weyhermüller, Dimitrios Maganas, Stephen Hill, Alexander Schnegg, Eckhard Bill, and others. Magneto-structural correlations in pseudotetrahedral forms of the [co (sph) 4] 2–complex probed by magnetometry, mcd spectroscopy, advanced epr techniques, and ab initio electronic structure calculations. *Inorg. Chem.*, 56(5):3102–3118, 2017.

[836] Elizaveta A. Suturina, Dimitrios Maganas, Eckhard Bill, Mihail Atanasov, and Frank Neese. Magneto-Structural Correlations in a Series of Pseudotetrahedral [CoII(XR)4]2– Single Molecule Magnets: An ab Initio Ligand Field Study. *Inorg. Chem.*, 54(20):9948–9961, 10 2015. doi:10.1021/acs.inorgchem.5b01706.

[837] Marcel Swart and F. Matthias Bickelhaupt. Optimization of strong and weak coordinates. *Int. J. Quantum Chem.*, 106(12):2536–2544, 01 2006. doi:10.1002/qua.21049.

[838] William C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. *J. Chem. Phys.*, 76:637–649, 1982.

[839] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 1989. ISBN 978-0-486-69186-2. URL: http://books.google.de/books?id=6mV9gYzEkgIC.

[840] P. G. Szalay and R. J. Bartlett. *Chem. Phys. Lett.*, 214:481, 1993.

[841] K. Takatsuka, T. Fueno, and K. Yamaguchi. *Theor. Chim. Acta*, 48:175, 1978.

[842] Matthias Tamm, Luong Phong Ho, Alexandre Nasr, Peter G Jones, Ahmet Altun, Frank Neese, and Giovanni Bistoni. London dispersion interactions in pnictogen cations [ECl$_2$]$^+$ and [E=E]$^2+$(E= p, as, sb) supported by anionic n-heterocyclic carbenes. *Chem. Eur. J.*, 2018. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/chem.201804714.

[843] Alex Tarnopolsky, Amir Karton, Rotem Sertchook, Dana Vuzman, and Jan M L Martin. Double-hybrid functionals for thermochemical kinetics. *J. Phys. Chem. A*, 112(1):3–8, 2008. URL: https://pubs.acs.org/doi/abs/10.1021/jp710179r, doi:https://doi.org/10.1021/jp710179r.

[844] M. Tarrago, C. Romelt, J. Nehrkorn, A. Schnegg, F. Neese, E. Bill, and S. F. Ye. Experimental and theoretical evidence for an unusual almost triply degenerate electronic ground state of ferrous tetraphenylporphyrin. *Inorg. Chem.*, 60(7):4966–4985, 2021. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000637850300081, doi:10.1021/acs.inorgchem.1c00031.

[845] Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao. *J. Chem. Phys.*, 120:8425–8433, 2004.

[846] G. Te Velde and E. J. Baerends. *J. Comp. Phys.*, 99:84, 1992.

[847] W. Thiel and A. A. Voityuk. *Theor. Chim. Acta*, 81:391, 1992.

[848] M. K. Thomsen, A. Nyvang, J. P. S. Walsh, P. C. Bunting, J. R. Long, F. Neese, M. Atanasov, A. Genoni, and J. Oyergaard. Insights into single-molecule-magnet behavior from the experimental electron density of linear two-coordinate iron complexes. *Inorg. Chem.*, 58(5):3211–3218, 2019. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000460600300035, doi:10.1021/acs.inorgchem.8b03301.

[849] Van Anh Tran and Frank Neese. Double-hybrid density functional theory for g-tensor calculations using gauge including atomic orbitals. *J. Chem. Phys.*, 153(5):054105, 08 2020. URL: https://doi.org/10.1063/5.0013799, doi:10.1063/5.0013799.

[850] O. Treutler and R. J. Ahlrichs. *J. Chem. Phys.*, 102:346, 1994.

[851] G. N. R. Tripathi and Robert H. Schuler. The resonance Raman spectrum of phenoxyl radical. *J. Chem. Phys.*, 81(1):113–121, 07 1984. URL: http://aip.scitation.org/doi/abs/10.1063/1.447373 (visited on 2018-02-07), doi:10.1063/1.447373.

[852] T. N. Truong and E. V. Stefanovich. *Chem. Phys. Lett.*, 240:253–260, 1995.

[853] S.A. Trygubenko and D.J. Wales. A doubly nudged elastic band method for finding transition states. *J. Chem. Phys.*, 120(5):2082–2094, 2004.

[854] Anil Kumar Tummanapelli and Sukumaran Vasudevan. Dissociation constants of weak acids from ab initio molecular dynamics using metadynamics: Influence of the inductive effect and hydrogen bonding on pKa values. *J. Phys. Chem. B*, 118(47):13651–13657, 2014. doi:10.1021/jp5088898.

[855] Anil Kumar Tummanapelli and Sukumaran Vasudevan. Estimating successive pKa values of polyprotic acids from ab initio molecular dynamics using metadynamics: the dissociation of phthalic acid and its isomers. *Phys. Chem. Chem. Phys.*, 17(9):6383–6388, 2015. doi:10.1039/C4CP06000H.

[856] Liviu Ungur. *Ab Initio Methodology for the Investigation of Magnetism in Strongly Anisotropic Complexes*. PhD thesis, KU Leuven, 10 2010.

[857] Liviu Ungur. Introduction to the electronic structure, luminescence, and magnetism of lanthanides. In Pablo Martín-Ramos and Manuela Ramos Silva, editors, *Lanthanide-Based Multifunctional Materials*, Advanced Nanomaterials, pages 1–58. Elsevier, 2018. URL: http://www.sciencedirect.com/science/article/pii/B9780128138403000016, doi:10.1016/B978-0-12-813840-3.00001-6.

[858] Liviu Ungur and Liviu F. Chibotaru. Ab initio crystal field for lanthanides. *Chem. Eur. J.*, 23(15):3708–3718, 2017. URL: https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/chem.201605102, arXiv:https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/chem.201605102, doi:10.1002/chem.201605102.

[859] O. Vahtras, J. Almlöf, and M. W. Feyereisen. *Chem. Phys. Lett.*, 213:514, 1993.

[860] O. Vahtras, B. Minaev, and H. Ågren. Ab initio calculations of electronic g-factors by means of multiconfiguration response theory. *Chem. Phys. Lett.*, 281(1):186–192, 1997. doi:https://doi.org/10.1016/S0009-2614(97)01169-X.

[861] C. Van Alsenoy. *J. Comput. Chem.*, 9:620, 1988.

[862] Willem Van den Heuvel and Alessandro Soncini. NMR chemical shift as analytical derivative of the Helmholtz free energy. *J. Chem. Phys.*, 138:054113, 2013. doi:doi:10.1063/1.4789398.

[863] E. van Lenthe, E. J. Baerends, and J. G. Snijders. *J. Chem. Phys.*, 99(6):4597–4610, 1993.

[864] E. van Lenthe, E. J. Baerends, and J. G. Snijders. *J. Chem. Phys.*, 101:9783–9792, 1994.

[865] C. Van Stappen, D. Maganas, S. DeBeer, E. Bill, and F. Neese. Investigations of the magnetic and spectroscopic properties of V(III) and V(IV) complexes. *Inorg. Chem.*, 57(11):6421–6438, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000434491700027, doi:10.1021/acs.inorgchem.8b00486.

[866] C. van Wüllen. *J. Chem. Phys.*, 109:392–399, 1998.

[867] Libor Veis, Andrej Antalík, Jiri Brabec, Frank Neese, Ors Legeza, and Jiri Pittner. Coupled cluster method with single and double excitations tailored by matrix product state wave functions. *J. Phys. Chem. Lett.*, 7(20):4072–4078, 2016.

[868] L. Verlet. *Phys. Rev.*, 159:98–103, 1967.

[869] Toon Verstraelen, Steven Vandenbrande, Farnaz Heidar-Zadeh, Louis Vanduyfhuys, Veronique Van Speybroeck, Michel Waroquier, and Paul W. Ayers. Minimal basis iterative stockholder: atoms in molecules for force-field development. *J. Chem. Theory Comput.*, 12:3894–3912, 2016. URL: https://doi.org/10.1021/acs.jctc.6b00456, doi:10.1021/acs.jctc.6b00456.

[870] Veacheslav Vieru, Naoya Iwahara, Liviu Ungur, and Liviu F Chibotaru. Giant exchange interaction in mixed lanthanides. *Scientific reports*, 6:24046, 2016.

[871] L. Visscher and K. G. Dyall. *Atom. Data Nucl. Data Tabl.*, 67:207, 1997.

[872] L. von Szentpaly, P. Fuentealba, H. Preuss, and H. Stoll. *Chem. Phys. Lett.*, 93:555–559, 1982.

[873] S. H. Vosko, L. Wilk, and M. Nusair. *Can. J. Phys.*, 58:1200, 1980.

[874] T. Vreven, B. Mennucci, C. O. da Silva, K. Morokuma, and J. Tomasi. *J. Chem. Phys.*, 115:62, 2001.

[875] O. A. Vydrov and T. Van Voorhis. *J. Chem. Phys.*, 133:244103, 2010.

[876] A. J. H. Wachters. *J. Chem. Phys.*, 52:1033, 1970.

[877] David J. Wales and Jonathan P. K. Doye. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, July 1997. Publisher: American Chemical Society. URL: https://doi.org/10.1021/jp970984n (visited on 2022-04-22), doi:10.1021/jp970984n.

[878] K. N. Walzl, C. F. Koerting, and A. Kuppermann. Electron-impact spectroscopy of acetaldehyde. *J. Chem. Phys.*, 87:3796–3803, 1987. doi:10.1063/1.452935.

[879] Y. Wang and M. Dolg. *Theor. Chem. Acc.*, 100:124, 1998.

[880] A. Weigand, X. Cao, J. Yang, and M. Dolg. *Theor. Chem. Acc.*, 126:117–127, 2010.

[881] F. Weigend and M. Häser. *Theor. Chem. Acc.*, 97:331, 1997.

[882] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs. *Chem. Phys. Lett.*, 294:143, 1998.

[883] Florian Weigend and Alexander Baldes. Segmented contracted basis sets for one- and two-component Dirac–Fock effective core potentials. *J. Chem. Phys.*, 133(17):174102, 2010. URL: https://doi.org/10.1063/1.3495681, doi:10.1063/1.3495681.

[884] Florian Weigend, Marco Kattannek, and Reinhart Ahlrichs. Approximated electron repulsion integrals: Cholesky decomposition versus resolution of the identity methods. *J. Chem. Phys.*, 130:164106, 04 2009. doi:https://doi.org/10.1063/1.3116103.

[885] F. Wennmohs and F. Neese. *Chem. Phys.*, 343:217–230, 2008. doi:.

[886] Hans-Joachim Werner and Wilfried Meyer. Pno-ci and pno-cepa studies of electron correlation effects: v. static dipole polarizabilities of small molecules. *Molecular Physics*, 31(3):855–872, 1976.

[887] J. L. Whitten. *J. Chem. Phys.*, 58:4496, 1973. doi:.

[888] K. B. Wiberg. *Tetrahedron*, 24:1083, 1968. doi:.

[889] E. B. Wilson, J. C. Decius, and P. C. Cross. *Molecular Vibrations – the Theory of Infrared and Raman Vibrational Spectra*. Dover Publications, 1955.

[890] L. Wittmann, H. Neugebauer, S. Grimme, and M. Bursch. Dispersion-corrected r2scan based double-hybrid functionals. *J. Chem. Phys.*, 2023.

[891] D. E. Woon and T. H. Dunning Jr. *J. Chem. Phys.*, 98:1358, 1993.

[892] D. E. Woon and T. H. Dunning Jr. *J. Chem. Phys.*, 100:2975, 1994.

[893] Axel Wuttke and Ricardo A Mata. Visualizing dispersion interactions through the use of local orbital spaces. *J. Comput. Chem.*, 38(1):15–23, 2017.

[894] Xin Xu and William A. Goddard, III. *Proc. Nat. Acad. Sci.*, 101:2673, 2004.

[895] Kiyoshi Yagi, Kimihiko Hirao, Tetsuya Taketsugu, Michael W. Schmidt, and Mark S. Gordon. \emph Ab initio vibrational state calculations with a quartic force field: Applications to H2CO, C2H4, CH3OH, CH3CCH, and C6H6. *J. Chem. Phys.*, 121(3):1383–1389, 07 2004. URL: http://aip.scitation.org/doi/10.1063/1.1764501 (visited on 2020-04-14), doi:10.1063/1.1764501.

[896] K. Yamaguchi, F. Jensen, A. Dorigo, and K. N. Houk. *Chem. Phys. Lett.*, 149:537, 1988.

[897] K. Yamaguchi, Y. Takahara, and T. Fueno. In V. H. Smith, editor, *Applied Quantum Chemistry*, pages 155. Wiley, 1986.

[898] K. Yamamoto, J. K. Li, J. A. O. Garber, J. D. Rolfes, G. B. Boursalian, J. C. Borghs, C. Genicot, J. Jacq, M. van Gastel, F. Neese, and T. Ritter. Palladium-catalysed electrophilic aromatic C-H fluorination. *Nature*, 554(7693):511–514, 2018. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000425597400043, doi:10.1038/nature25749.

[899] T. Yanai, D. P. Tew, and N. C. Handy. *Chem. Phys. Lett.*, 393:51–57, 2004.

[900] J. Yang and M. Dolg. *Theor. Chem. Acc.*, 113:212, 2005.

[901] Danny L. Yeager and Poul Jø rgensen. A multiconfigurational time-dependent Hartree–Fock approach. *Chem. Phys. Lett.*, 65:77–80, 1979.

[902] D. Yepes, F. Neese, B. List, and G. Bistoni. Unveiling the delicate balance of steric and dispersion interactions in organocatalysis using high-level computational methods. *J. Am. Chem. Soc.*, 142(7):3613–3625, 2020. URL: \T1\textless{}GotoISI\T1\textgreater{}://WOS:000515214000044, doi:10.1021/jacs.9b13725.

[903] Diana Yepes, Frank Neese, Benjamin List, and Giovanni Bistoni. Unveiling the delicate balance of steric and dispersion interactions in organocatalysis using high-level computational methods. *J. Am. Chem. Soc.*, 142(7):3613–3625, 2020. URL: https://doi.org/10.1021/jacs.9b13725, arXiv:https://doi.org/10.1021/jacs.9b13725, doi:10.1021/jacs.9b13725.

[904] D. M. York and M. Karplus. *J. Phys. Chem. A*, 103:11060–11079, 1999.

[905] Feng Yu. Spin-Component-Scaled Double-Hybrid Density Functionals with Nonlocal van der Waals Correlations for Noncovalent Interactions. *J. Chem. Theory Comput.*, 10(10):4400–4407, 10 2014. URL: https://doi.org/10.1021/ct500642x (visited on 2020-09-15), doi:10.1021/ct500642x.

[906] M. C. Zerner. *Int. J. Quant. Chem.*, 35:567, 1989.

[907] M. C. Zerner. In K. B. Lipkowitz and D. B. Boyd, editors, *Reviews in Computational Chemistry*, volume 2, pages 313. Wiley-VCH, 1990.

[908] M. C. Zerner. In D. R. Salahub and N. Russo, editors, *Metal-Ligand Interactions: From Atoms to Clusters to Surfaces*, pages 101. Kluwer Academic Publishers, 1992.

[909] M. C. Zerner. In D. R. Salahub and N. Russo, editors, *Metal-Ligand Interactions: Structure and Reactivity*, pages 493. Kluwer Academic Publishers, 1992.

[910] M. C. Zerner and M. Hehenberger. *Chem. Phys. Lett.*, 62:550, 1979.

[911] M. C. Zerner, G. H. Loew, R. F. Kirchner, and U. T. Mueller-Westerhoff. *J. Am. Chem. Soc.*, 102:589, 1980.

[912] Dominika Zgid, Debashree Ghosh, Eric Neuscamman, and Garnet Kin-Lic Chan. A study of cumulant approximations to n-electron valence multireference perturbation theory. *J. Chem. Phys.*, 130:194107, 05 2009.

[913] Y. Zhang and W. Yang. *Phys. Rev. Lett.*, 80:890, 1998.

[914] Y. Zhao and D. G. Truhlar. *J. Phys. Chem. A*, 109:5656–5667, 2005.

[915] Y. Zhao and Donald G. Truhlar. *J. Chem. Phys.*, 125:194101, 2006.

[916] Y. Zhao and Donald G. Truhlar. *Theor. Chem. Acc.*, 120:215, 2008.

[917] Jingjing Zheng, Xuefei Xu, and Donald G Truhlar. *Theor. Chem. Acc.*, 128:295–305, 2010.

[918] Y. C. Zheng and J. Almlöf. *Chem. Phys. Lett.*, 214:397, 1993.

[919] Y. C. Zheng and J. Almlöf. *J. Mol. Struct.: THEOCHEM*, 388:277, 1996.

[920] Ting Zhu, Ju Li, Amit Samanta, Hyoung Gyu Kim, and Subra Suresh. Interfacial plasticity governs strain rate sensitivity and ductility in nanostructured metals. *Proc. Nat. Acad. Sci.*, 104(9):3031–3036, 2007. URL: https://www.pnas.org/content/104/9/3031, arXiv:https://www.pnas.org/content/104/9/3031.full.pdf, doi:10.1073/pnas.0611097104.

[921] J. Patrick Zobel, Juan J. Nogueira, and Leticia Gonzalez. The IPEA Dilemma in CASPT2. *Chem. Sci.*, 09 2016.

[922] Gamess program. Gamess user manual.