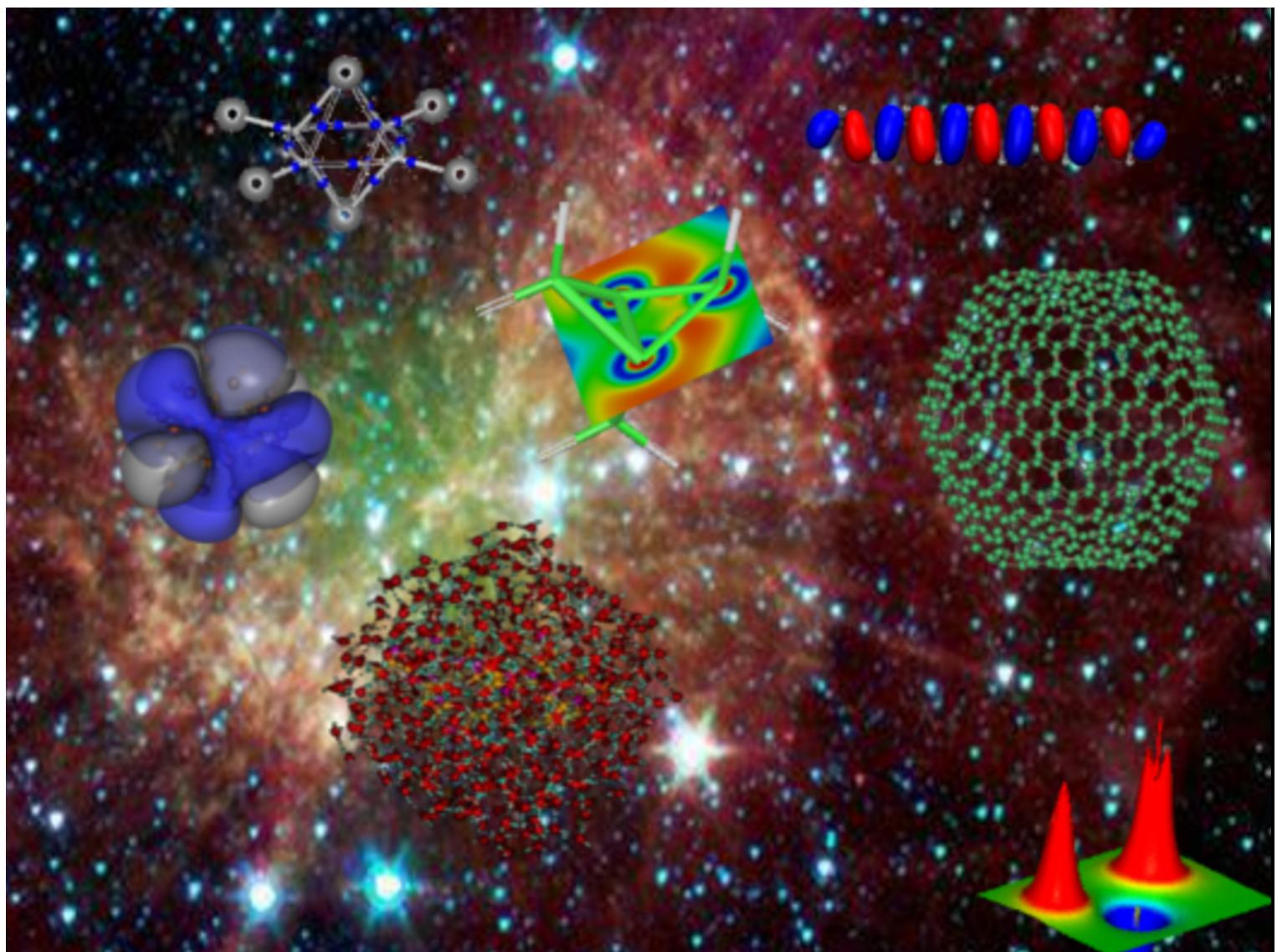


# The deMon2k Users' Guide

## Installation and Reference Manual

<http://www.deMon-software.com>





## Preface<sup>†</sup>

This is the third Users' Guide written for the program deMon2k. It describes version 5.0 and has increased with respect to the last edition for version 3.0 by more than 60 pages! This reflects the substantial increase in the functionality and the number of keywords and their options in the new version of deMon2k.

I would like to take this opportunity to thank all deMon2k developers and users for their corrections and comments to this guide. The community effort to write this guide is well reflected by the author and editor list. The critical feedback of users and developers is essential and very desirable. Please keep it up!

I also thank Cinvestav for allowing the development of deMon2k and in particular for the sabbatical leave in the summer semester of 2017 in which the guide was concluded. Type-setting was done with  $\text{\LaTeX}$ .

Cinvestav, Mexico-City,

*Andreas M. Köster*

January 2018

---

<sup>†</sup>Dedicated to Dr. Annick Goursot who passed away on March 31<sup>st</sup>, 2017.



# **The deMon2k Users' Guide**

## **Installation and Reference Manual**

<http://www.deMon-software.com>

**Version 5.0**  
**February 2018**

Reference: The deMon2k Users' Guide, Version 5.0, Cinvestav, Mexico-City, 2018

Authors: A.M. Köster, G. Geudtner, G.U. Gamboa, A. Alvarez-Ibarra, P. Calaminici, R. Flores-Moreno, A. Goursot, A. de la Lande, D. Mejia-Rodriguez, T. Mineva, L.G.M. Petersson, J.M. Vasquez-Perez, B. Zuniga-Gutierrez

Editors: S.B. Trickey and D.R. Salahub

The procedures and applications presented in this Guide have been included for their instructional value. They have been tested with care but are not guaranteed for any particular purpose. The authors and editors do not offer any warranties or representations, nor do they accept any liabilities with respect to the programs or applications.

The software described in this Guide is furnished under a license and may be used or copied only in accordance with the terms of that license.

Copyright ©2018 by the authors and editors

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission of at least one of the authors or editors.

# Contents

<b>1 Getting Acquainted</b>	<b>1</b>
1.1 The Game of the Name . . . . .	1
1.2 How to Avoid Reading this Manual . . . . .	5
1.3 How to Read this Manual . . . . .	5
1.4 How to Use deMon2k . . . . .	6
<b>2 Getting Started</b>	<b>13</b>
2.1 Before Installing deMon2k . . . . .	13
2.1.1 FORTRAN90 Compiler . . . . .	13
2.1.2 Bourne Again Shell . . . . .	13
2.1.3 MPI . . . . .	13
2.2 How to Install deMon2k . . . . .	13
2.3 How to Compile or Re-compile deMon2k . . . . .	16
2.4 How to Change the Compiler . . . . .	20
2.5 The Database File . . . . .	20
2.5.1 F90 Variable . . . . .	21
2.5.2 FFLAGS Variable . . . . .	22
2.5.3 LFLAGS Variable . . . . .	22
2.5.4 LIBS Variable . . . . .	22
2.5.5 MPIBIT Variable . . . . .	22
2.5.6 EXCLUDE Variable . . . . .	23
2.5.7 Source File Name Variables . . . . .	23
2.6 How to Run deMon2k . . . . .	23
2.6.1 Running deMon2k Directly . . . . .	23
2.6.2 Running the deMon Script . . . . .	24
<b>3 Carrying On</b>	<b>27</b>
3.1 Parameters . . . . .	27
3.2 Files . . . . .	28
3.3 Input Syntax . . . . .	29
3.4 Job Restart . . . . .	31
<b>4 Keywords</b>	<b>32</b>

4.1	Molecular Geometry Input . . . . .	32
4.1.1	Keyword <b>GEOMETRY</b> . . . . .	32
4.1.2	Keywords <b>REACTANT</b> and <b>PRODUCT</b> . . . . .	40
4.1.3	Keyword <b>CONSTANTS</b> . . . . .	42
4.1.4	Keyword <b>CONSTRAINTS</b> . . . . .	44
4.1.5	Keyword <b>VARIABLES</b> . . . . .	46
4.1.6	Keyword <b>SYMMETRY</b> . . . . .	47
4.1.7	Keyword <b>ALIGNMENT</b> . . . . .	48
4.1.8	Keyword <b>PATTERN</b> . . . . .	49
4.2	Methodology Selection . . . . .	51
4.2.1	Keyword <b>VXCTYPE</b> . . . . .	51
4.2.2	Keyword <b>DISPERSION</b> . . . . .	54
4.2.3	Keyword <b>FORCEFIELD</b> . . . . .	56
4.2.4	Keyword <b>QM/MM</b> . . . . .	57
4.2.5	Keyword <b>EFIELD</b> . . . . .	59
4.2.6	Keyword <b>EMBED</b> . . . . .	60
4.3	Atom Specific Input . . . . .	61
4.3.1	Keyword <b>BASIS</b> . . . . .	61
4.3.2	Keyword <b>AUGMENT</b> . . . . .	65
4.3.3	Keyword <b>AUXIS</b> . . . . .	65
4.3.4	Keyword <b>ECPS</b> . . . . .	67
4.3.5	Keyword <b>MCPS</b> . . . . .	72
4.3.6	Keyword <b>GRID</b> . . . . .	75
4.3.7	Keyword <b>FREEZE</b> . . . . .	76
4.4	Electronic State Control . . . . .	78
4.4.1	Keyword <b>MULTIPLICITY</b> . . . . .	78
4.4.2	Keyword <b>CHARGE</b> . . . . .	78
4.4.3	Keyword <b>MOEXCHANGE</b> . . . . .	78
4.4.4	Keyword <b>MOMODIFY</b> . . . . .	80
4.4.5	Keyword <b>FIXMOS</b> . . . . .	81
4.4.6	Keyword <b>SMEAR</b> . . . . .	82
4.4.7	Keyword <b>CONFIGURE</b> . . . . .	84
4.5	SCF Control and Stabilization . . . . .	88

4.5.1 Keyword <b>SCFTYPE</b> . . . . .	88
4.5.2 Keyword <b>ORBITALS</b> . . . . .	93
4.5.3 Keyword <b>PHASE</b> . . . . .	93
4.5.4 Keyword <b>ERIS</b> . . . . .	95
4.5.5 Keyword <b>GUESS</b> . . . . .	99
4.5.6 Keyword <b>MIXING</b> . . . . .	100
4.5.7 Keyword <b>SHIFT</b> . . . . .	100
4.5.8 Keyword <b>DIIS</b> . . . . .	101
4.6 Optimization, Interpolation, and Transition State Search . . . . .	103
4.6.1 Keyword <b>OPTIMIZATION</b> . . . . .	103
4.6.2 Keyword <b>SADDLE</b> . . . . .	105
4.6.3 Keyword <b>SCAN</b> . . . . .	108
4.6.4 Keyword <b>IRC</b> . . . . .	110
4.6.5 Keyword <b>HESSIAN</b> . . . . .	112
4.6.6 Keyword <b>UPDATE</b> . . . . .	113
4.6.7 Keyword <b>STEPTYPE</b> . . . . .	114
4.7 Born-Oppenheimer Molecular Dynamics . . . . .	115
4.7.1 Keyword <b>DYNAMICS</b> . . . . .	115
4.7.2 Keyword <b>TRAJECTORY</b> . . . . .	115
4.7.3 Keyword <b>SIMULATION</b> . . . . .	118
4.7.4 Keyword <b>LPCONSERVE</b> . . . . .	125
4.7.5 Keyword <b>VELOCITIES</b> . . . . .	126
4.7.6 Keyword <b>BATH</b> . . . . .	127
4.7.7 Keyword <b>RDF</b> . . . . .	131
4.8 Molecular Property Control . . . . .	133
4.8.1 Keyword <b>FREQUENCY</b> . . . . .	133
4.8.2 Keyword <b>THERMO</b> . . . . .	134
4.8.3 Keyword <b>DIPOLE</b> . . . . .	135
4.8.4 Keyword <b>POLARIZABILITY</b> . . . . .	135
4.8.5 Keyword <b>EXCITATION</b> . . . . .	140
4.8.6 Keyword <b>NONCOLLINEAR</b> . . . . .	141
4.8.7 Keyword <b>XRAY</b> . . . . .	141
4.8.8 Keyword <b>NMR</b> . . . . .	146

4.8.9 Keyword <b>MAGNETIZABILITY</b> . . . . .	151
4.8.10 Keyword <b>NQR</b> . . . . .	152
4.8.11 Keyword <b>FUKUI</b> . . . . .	155
4.9 Electronic Structure Analysis . . . . .	156
4.9.1 Keyword <b>POPULATION</b> . . . . .	156
4.9.2 Keyword <b>DOS</b> . . . . .	159
4.9.3 Keyword <b>SIGPI</b> . . . . .	159
4.9.4 Keyword <b>LOCALIZATION</b> . . . . .	160
4.10 Visualization and Topological Analysis . . . . .	161
4.10.1 Keyword <b>VISUALIZATION</b> . . . . .	161
4.10.2 Keyword <b>TOPOLOGY</b> . . . . .	162
4.10.3 Keyword <b>PLOT</b> . . . . .	163
4.10.4 Keyword <b>CPSEARCH</b> . . . . .	168
4.10.5 Keyword <b>ISOSURFACE</b> . . . . .	169
4.10.6 Keyword <b>GEOINTERFACE</b> . . . . .	172
4.10.7 Keyword <b>BOX</b> . . . . .	174
4.10.8 Keyword <b>POINTS</b> . . . . .	176
4.11 Performance and Accuracy Settings . . . . .	178
4.11.1 Keyword <b>PARALLEL</b> . . . . .	178
4.11.2 Keyword <b>MATDIA</b> . . . . .	178
4.11.3 Keyword <b>MATINV</b> . . . . .	179
4.11.4 Keyword <b>DAVIDSON</b> . . . . .	181
4.11.5 Keyword <b>WEIGHTING</b> . . . . .	181
4.11.6 Keyword <b>QUADRATURE</b> . . . . .	182
4.11.7 Keyword <b>CFPINTEGRATION</b> . . . . .	182
4.12 Miscellaneous Keywords . . . . .	184
4.12.1 Keyword <b>TITLE</b> . . . . .	184
4.12.2 Keyword <b>PRINT</b> . . . . .	184
<b>A Automatic Generation of Auxiliary Functions</b>	<b>187</b>
<b>B Test Set for deMon2k Installation</b>	<b>189</b>
<b>C MHM Program for deMon2k Trajectory Files</b>	<b>192</b>



## List of Tables

1	Parameter Settings . . . . .	27
2	Physical and Logical Filenames . . . . .	29
3	Special Input Symbols . . . . .	30
4	Internal Coordinate Substitution . . . . .	37
5	Exchange Functionals . . . . .	53
6	Correlation Functionals . . . . .	54
7	Basis Sets . . . . .	62
8	ROKS Parametrizations . . . . .	89
9	SCF Tightening . . . . .	92
10	ERI Calculation . . . . .	96
11	Relationship between the keywords GEOMETRY and OPTIMIZATION . .	104
12	Molecular Fields . . . . .	165
13	Content of the tests directory. . . . .	189

## List of Figures

1	Directory Tree of deMon2k . . . . .	14
2	Z-Matrix Input Orientation . . . . .	35
3	Dihedral Angle Definition . . . . .	36
4	Mixed Input Definition . . . . .	38
5	QM/MM Input Definition . . . . .	57
6	Basis Set Exchange Library . . . . .	64
7	Stuttgart-Dresden ECPs . . . . .	69
8	Los Alamos National Laboratory ECPs . . . . .	70
9	Lovallo-Klobukowski MCPs . . . . .	74
10	Saddle Alignment . . . . .	106
11	BOMD Temperature Profiles . . . . .	129
12	Phase Space Distribution . . . . .	130
13	Elliptic Coordinates . . . . .	158
14	Molecular Field Plots . . . . .	167
15	Critical Points of Molecular Fields . . . . .	169
16	Isodensity Surface and Electrostatic Potential of Insulin . . . . .	171

# 1 Getting Acquainted

## 1.1 The Game of the Name

The deMon program system was developed for density functional theory (DFT) calculations of atoms, molecules, and clusters [1–3]. Its first widely available version appeared in 1992. It was based on the Ph.D. work of A. St-Amant, performed under the supervision of D.R. Salahub at the Université de Montréal [4]. In fact, the name **deMon** stands for "density of **M**ontréal". A key motivation for the deMon program development was to merge the well-established Gaussian quantum chemistry techniques with the just emerging DFT developments for finite systems. For this reason, deMon is based on the linear combination of Gaussian type orbital (LCGTO) method. To maintain efficiency, auxiliary functions were used from the very beginning in deMon. To avoid the calculation of four-center electron repulsion integrals (ERIs), which represents the computational bottleneck of the corresponding LCGTO Hartree-Fock calculations, the variational fitting of the Coulomb potential, as developed by Dunlap, Connolly, and Sabin [5–7], was introduced. A second auxiliary function set was used for the numerical calculation of the exchange-correlation potential as suggested by Sambe and Felton [8]. This second fit was not variational nor was the corresponding energy.

Shortly after its first appearance, the original deMon code was substantially modified for commercialization by BIOSYM Technologies. The beta-release of that version appeared in 1993. It also served as the basis for the deMon-KS [9] series of programs developed in Montréal until 1997. The focus of that program development was mainly upon molecular property calculations with LCGTO-DFT methods. Remarkable progress was achieved for the calculation of nuclear magnetic shielding tensors [10], nuclear quadrupole coupling constants [11], and excited states via time-dependent DFT [12]. A drawback of the first deMon-KS release, deMon-KS1, was its poor self-consistent field (SCF) convergence. At that time it was not obvious if this difficulty was due only to technical problems or if the fitting of the Coulomb and exchange-correlation potential represented a fundamental problem for the SCF convergence. In fact, a formally exact resolution of this issue was published only in 2009 [13]. However, it had already been proven that the energy gradients were not variational due to the least-squares fit of the exchange-correlation potential [14–16].

Triggered by these concerns, the original deMon-KS version was developed further in Montpellier by A. Goursot and in Stockholm by L.G.M. Pettersson. A major achievement

was the implementation of accurate numerical integration algorithms for the calculation of the exchange-correlation matrix elements. In this way the troublesome least-squares fit could be avoided. However, the computational demands associated with this early numerical integration algorithm rendered the approach impractical for routine applications. Nevertheless, the variational nature of the resulting gradients resolved some of the basic problems encountered with deMon-KS1. Furthermore, the calculation of core-level spectra was implemented in the Stockholm version. Originally, the Montpellier and Stockholm efforts were independent from each other. In 1997 the two variants were merged into the deMon-KS3 [17] program version. In cooperation with K. Hermann from the Fritz-Haber Institut in Berlin, the deMon branch from Stockholm was modified further for the calculation of core-level spectra. It is distributed under the name StoBe [18] and is particularly well suited for the calculation of core-level spectra.

Independently from the deMon development, the ALLCHEM [19] project was initiated by the group of A.M. Köster in Hannover in 1995. The aim of that project was to write a well-structured DFT code from scratch for the further study of auxiliary density approaches. For this purpose, general recurrence relations for three-center ERIs that appear in the variational fitting of the Coulomb potential were developed and coded [20]. With this approach, the original limitation of deMon to *d*-type orbital and auxiliary functions was removed in ALLCHEM. Other key developments in the ALLCHEM project were adaptive grid techniques [21] for the numerical integration of the exchange-correlation matrix elements. With these smaller yet reliable grids, variational energy calculations became routinely available. The first ALLCHEM calculations appeared in the literature in 1997. The structured programming of ALLCHEM proved very useful for the development and testing of new auxiliary function DFT approaches and algorithms. On the other hand, ALLCHEM was increasingly being used for standard computational chemistry applications, for which the primitive input structure was error prone and rather cumbersome to use.

Based upon an initiative of D.R. Salahub, A. Goursot, and A.M. Köster, the first deMon developers meeting was held in Ottawa in March 2000. At that meeting, the deMon and ALLCHEM developers agreed to merge their codes in order to keep a Tower of Babel from arising. As a result, a new code was developed. It couples deMon functionality and input utilities with the stable and efficient integral and SCF branch from ALLCHEM. This code, named deMon2k [22], was presented for the first time at the third deMon developers meeting in Geneva in 2002. Over the next three years, the beta version of this code was dis-

tributed and tested within the deMon developers community. The annual meetings of this community at the deMon Developers workshops were used to discuss the code structure and to exchange algorithmic developments.

Since the 6<sup>th</sup> deMon Developers workshop in 2005 (Dresden), deMon2k has been distributed to the public under a license agreement from Cinvestav, Mexico-City. Originally the 2.2.6 release served as the distribution version. It was superseded by the 2.3.1 release. Besides the standard LCGTO-DFT auxiliary function methodology ported from ALLCHEM, the 2.2.6 release contained the first stable implementation of auxiliary density functional theory (ADFT) [23] that enables calculations of large systems with modest computational resources. With the 2.3.1 release, a parallel ADFT SCF module was realized on the basis of MPI [24]. That version and its successors have been distributed to more than 400 research groups around the world. The 2.3.1 release served as our development platform over the intervening four years. Today, ADFT is accepted as a stable and efficient alternative to conventional Kohn-Sham DFT calculations, as evidenced by the increasing number of deMon2k ADFT calculations.

In February 2009, the 9<sup>th</sup> deMon Developers workshop was held in Pune, India. The new 3.0 release of deMon2k was presented there. It is a systematic further development of version 2.3.1. The program is fully parallelized and has been successfully ported to massively parallel architectures. Major improvements are the implementation of an iterative solver for the fitting equations that can handle hundreds of thousands of auxiliary functions [25], of improved functionals [26–29], of auxiliary density perturbation theory (ADPT) [30], of a hierarchical transition state finder [31], and of an extended Born-Oppenheimer molecular dynamics module [32–34]. Based on the demand from the deMon2k user community, we have also extended the molecular property section of the program. The new version contains internal modules for time-dependent density functional theory (TD-DFT) calculations in the framework of ADFT from the group of M.E. Casida [35], nuclear magnetic resonance shielding tensor calculations employing gauge invariant atomic orbital (GIAO) basis sets and nuclear quadrupole resonance calculations as well as polarizability and hyperpolarizability calculations [36,37]. The electronic structure analysis module also has been extended to include both atoms-in-molecules (AIM) [38,39] and natural bond orbital (NBO) [40] analysis. A standardized interface to the newest VU version [41] for scientific visualization of molecular fields is available, too. Routines for calculating x-ray absorption (XAS) [42] and emission (XES) spectra as well as x-ray photoelectron spectroscopy (XPS) were

included.

At the 17<sup>th</sup> deMon Developers workshop in May 2017 in Calgary, Canada, the new 5.0 public release of deMon2k was announced. It breaks with the original focus on DFT methods only and now includes also a Hartree-Fock and an empirical force field module. This permits hybrid DFT calculations and quantum mechanical molecular mechanical (QM/MM) calculations within deMon2k. Despite these extensions only three-center ERIs are needed. This is possible due to a variational fitting of the Fock potential [43] very much in the same spirit as the variational fitting of the Coulomb potential. For the QM/MM calculations in deMon2k [44] the OPLS-AA force field [45] was implemented. An interface with CHARMM is also available [46]. The deMon2k version discussed here is also part of several multiscale packages and approaches [47,48]. To learn more about these packages and other deMon2k utility programs as well as activities of the deMon2k user and developer community please visit <http://www.deMon-software.com>.

From the many changes and extensions of the new deMon2k 5.0 version we would like to mention two more because they alter default settings. The first one is the default setting for the ERIS keyword (see 4.5.4) that was changed from CONVENTIONAL to MULTIPOLE. The second is the default setting of the auxiliary function set by the AUXIS keyword (see 4.3.3) which was changed from A2 to GEN-A2. **Note that due to this change old inputs without explicitly defined auxiliary function sets will generate different results with respect to version 3!** Because of the intensive development of deMon2k we also decided to distribute developer versions to the public. They are usually not fully tested! They can be identified easily by an even version number.

The new deMon2k version is available for just about any computer made today, including parallel architectures. The versions for the different platforms are essentially the same. In particular, there is only one common source for the serial and parallel versions. Thus, no special knowledge about MPI is needed for source code modifications. An input file created according to this Guide should give, within the numerical accuracy of the compile flags used, the same output.

The written word has certain complications not characteristic of the spoken, and the need to distinguish deMon2k from similarly spelled products restricts how you may write it. The proper way to refer to the program in text is to use the unusual capitalization. The new deMon2k version must be cited as:

A.M. Koster, G. Geudtner, A. Alvarez-Ibarra, P. Calaminici, M.E. Casida, J. Carmona-Espindola, V.D. Dominguez, R. Flores-Moreno, G.U. Gamboa, A. Goursot, T. Heine, A. Ipatov, A. de la Lande, F. Janetzko, J.M. del Campo, D. Mejia-Rodriguez, J.U. Reveles, J. Vasquez-Perez, A. Vela, B. Zuniga-Gutierrez and D.R. Salahub, deMon2k, Version 5, The deMon developers, Cinvestav, Mexico-City (2018).

There is a companion to this Guide, *The deMon2k Tutorial*, which is in a comparatively early stage of development at the time of this writing. It gives a practical introduction to the use of deMon2k by example inputs and outputs. Some of them are taken from the Computational Chemistry course given in Cinvestav, Mexico-City, University of Calgary, Canada and a CECAM summer school in Paris. Most of these examples have a rather basic and didactic description. Some others touch quite advanced topics and may also be of interest to very expert users. The tutorial also provides a "Troubleshooting" section in which common problems are discussed and practical solutions are provided. At several points in this Guide we will refer to specific examples in the Tutorial for more practical instructions.

The deMon Authors and Editors do not run a help desk. An FAQ is under development for the website <http://www.deMon-software.com>. However, users who find bugs and non-trivial difficulties should provide a concise, comprehensive description of the problem and suspected diagnosis in an email to [akoster@cinvestav.mx](mailto:akoster@cinvestav.mx) or [geudtner@cinvestav.mx](mailto:geudtner@cinvestav.mx). All mails should include the corresponding deMon2k input and output files.

## 1.2 How to Avoid Reading this Manual

If you are familiar with an older version of deMon2k<sup>1</sup> or you would rather learn by doing, look at the directory `examples` in the deMon tree you have received. It contains sample inputs for the most common applications that are discussed in the deMon2k Tutorial. If you prefer to learn more about the deMon2k input syntax before you use the code, read on.

## 1.3 How to Read this Manual

While the example inputs cover the most common standard applications of a quantum chemistry program like deMon2k, there is much more to learn about input possibilities.

---

<sup>1</sup>Please be aware that some default settings have changed with respect to previous versions!

Eventually, you will want to control specific parts of a calculation or you will need more output information. Then you will need to consult this manual for the necessary keyword information. Each keyword section is self-contained, *i.e.* you can read the section about a keyword without knowing the preceding ones. However, **all** keyword sections assume that you have read Chapters 1 and 3.

For example, suppose you want to print molecular orbitals (MOs). The Table of Contents will lead you to Section 4.12.2, which describes the PRINT keyword. The concept of keywords and options, as well as their notation in this Guide, is described in Chapter 3. It will take you just a moment to construct the input line for printing molecular orbitals if you have already read chapter 3. It could be quite frustrating if you haven't. So it is best to read at least this and the third Chapter now. Here is a brief sketch of the chapter contents:

**Chapter 1** reviews the history of deMon2k and provides a brief background for its use.

**Chapter 2** gives an overview about how to install and run the program using the scripts that come with the deMon2k distribution.

**Chapter 3** explains how to customize the program parameters and input/output files. It also explains the input syntax.

**Chapter 4** gives a detailed description of all input keywords and their options.

**Appendix A** describes the automatic generation of auxiliary fitting function sets.

**Appendix B** lists the test cases that come with the deMon2k distribution.

**Appendix C** describes the input for the multiple histogram method (MHM) program.

## 1.4 How to Use deMon2k

The deMon2k program implements DFT in the Kohn-Sham formulation. It uses the linear combination of Gaussian type orbital (LCGTO) method. In this framework, the Kohn-Sham orbitals  $\psi_i(\mathbf{r})$  are expanded in an atomic orbital basis:

$$\psi_i(\mathbf{r}) = \sum_{\mu} c_{\mu i} \mu(\mathbf{r}) \quad (1.1)$$

Here  $\mu(\mathbf{r})$  denotes an atomic orbital (built from contracted Gaussian basis functions) and  $c_{\mu i}$  the corresponding molecular orbital coefficient. With this expansion, the electronic density is:

$$\rho(\mathbf{r}) = \sum_{\mu, \nu} P_{\mu\nu} \mu(\mathbf{r}) \nu(\mathbf{r}) \quad (1.2)$$

$P_{\mu\nu}$  is an element of the (closed-shell, also called non-spin-polarized in the DFT literature) density matrix defined as:

$$P_{\mu\nu} = 2 \sum_i^{occ} c_{\mu i} c_{\nu i} \quad (1.3)$$

Using the LCGTO expansions for the Kohn-Sham orbitals (1.1) and the electronic density (1.2), the Kohn-Sham self-consistent field (SCF) energy expression [49] can be expressed as:

$$E_{SCF} = \sum_{\mu, \nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu, \nu} \sum_{\sigma, \tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu \| \sigma\tau \rangle + E_{xc}[\rho] \quad (1.4)$$

The total energy is the sum of  $E_{SCF}$  and the nuclear repulsion energy, which can be calculated analytically. In (1.4),  $H_{\mu\nu}$  are elements of the core Hamiltonian matrix. They are built from the kinetic and nuclear attraction energy operators of the electrons and describe the distribution of an independent electron in the nuclear framework. The second term in (1.4) is the Coulomb repulsion energy of the electrons. In the short-hand notation for the four-center electron repulsion integrals (ERIs) the symbol  $\|$  represents the two-electron Coulomb operator and separates functions of electron 1 from those of electron 2. In contrast to Hartree-Fock theory, the calculations of the Coulomb and exchange energies are separate in Kohn-Sham DFT. Calculation of the exchange-correlation energy  $E_{xc}[\rho]$  requires numerical integration. In deMon2k, the  $N^4$  scaling of straight-forward calculation of the Coulomb repulsion energy is avoided by introducing an auxiliary function density [50]. This approximated density  $\tilde{\rho}(\mathbf{r})$  is expanded in primitive Hermite Gaussians  $\bar{k}(\mathbf{r})$  which are centered on the atoms [51,52]:

$$\tilde{\rho}(\mathbf{r}) = \sum_{\bar{k}} x_{\bar{k}} \bar{k}(\mathbf{r}) \quad (1.5)$$

The primitive Hermite Gaussian auxiliary functions are grouped in auxiliary function sets that share the same exponent [53,54]. For this reason, they usually are denoted as  $s$ ,  $p$ ,  $d$  etc. auxiliary function sets. With the LCGTO expansion for  $\rho(\mathbf{r})$  and  $\tilde{\rho}(\mathbf{r})$  we obtain the following approximate SCF energy:

$$E_{SCF} = \sum_{\mu, \nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu, \nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k}, \bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle + E_{xc}[\rho] \quad (1.6)$$

Therefore, only three-center electron repulsion integrals are necessary for the SCF and energy calculation in deMon2k. This represents the density fitting Kohn-Sham method available in deMon2k. It is activated by the keyword VXCTYPE BASIS (see Section 4.2.1 for more details about the VXCTYPE keyword). However, by default (VXCTYPE AUXIS), the approximated density is also used for the calculation of the exchange-correlation energy:

$$E_{SCF} = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle + E_{xc}[\tilde{\rho}] \quad (1.7)$$

This is the auxiliary density functional theory (ADFT) energy expression. For more details on ADFT, see the reviews [55–59]. Typically, the optimized ADFT structure parameters are indistinguishable from their full DFT counterparts even for weakly bound systems (here the use of the GEN-A2\* auxiliary function set is recommended; see Section 4.3.3 and Appendix A). For binding energies, ADFT and Kohn-Sham results typically deviate by less than 1 kcal/mol if GEN-A2\* or larger auxiliary function sets are used. Thus, the differences between ADFT and Kohn-Sham DFT geometries and bond energies are usually in the range of the accuracy of the underlying approximate exchange-correlation functional. Because of the considerable savings in computational time, we suggest to use ADFT for all studies including frequency analysis and property calculations. The VXCTYPE BASIS option Eq. (1.6) should be employed only if direct comparison with four-center DFT calculations is required. It should be noted that the default setting for the auxiliary functions is GEN-A2, independent of which energy expression is used (see Section 4.3.3). For all theoretical models available in deMon2k, VXCTYPE AUXIS results can be used as a restart guess (GUESS RESTART; see Section 4.5.5) for VXCTYPE BASIS calculations.

The most frequently encountered problem in DFT calculations is the failure to achieve SCF convergence. Usually this is caused by the small energy gap between the highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbital. In deMon2k, the DIIS procedure (Section 4.5.8) is activated by default. For a small HOMO-LUMO gap, DIIS may be counterproductive and should be switched off. There are several options available in deMon2k to achieve SCF convergence. Most important are modifications of the choice of the starting GUESS (Section 4.5.5) and the MIXING (Section 4.5.6) of the old and new (auxiliary) densities as well as enlargement of the HOMO-LUMO gap by the level-SHIFT (Section 4.5.7) procedure. If a static level-shift is employed it is advisable to check the orbital energies and occupations at the HOMO-LUMO gap by use of the PRINT keyword (Section 4.12.2). Other relevant keywords to alter or achieve SCF convergence are MOEX-CHANGE (Section 4.4.3), FIXMOS (Section 4.4.5) and SMEAR (Section 4.4.6). For atomic

calculations, the CONFIGURE keyword (Section 4.4.7) should be used in order to ensure SCF convergence.

In deMon2k 5.0 the calculation of Hartree-Fock energies by the variational fitting of the Coulomb and Fock potentials is also available. The corresponding SCF energy has the form [43]:

$$\begin{aligned} E_{SCF} = & \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle - \\ & \frac{1}{4} \sum_{\mu,\nu} \sum_{\sigma,\tau} \sum_{\bar{k},\bar{l}} P_{\mu\nu} P_{\sigma\tau} \langle \mu\sigma \| \bar{k} \rangle \langle \bar{k} \| \bar{l} \rangle^{-1} \langle \bar{l} \| \tau\nu \rangle \end{aligned} \quad (1.8)$$

Note that the same auxiliary function sets are used for the Coulomb and Fock potential fitting. As a result, the approximated Hartree-Fock energy, Eq. (1.8), is self-interaction free. To obtain a computationally efficient methodology the Fock potential fitting is performed with localized molecular orbitals [61]. This yields a computationally efficient and very accurate approximate Hartree-Fock energy expression that only requires three-center ERIs. Deviations with respect to four center ERIs total energies are below 1 kcal/mol if GEN-A2\* auxiliary function sets are used. With this development hybrid functionals such as B3LYP [62,63], PBE0 [64,65] and M06-2X [66] are now available in deMon2k [67].

For QM/MM calculations in deMon2k 5.0 the following energy expression is used [44]:

$$E = E^{QM} + E^{QMMM} + E^{MM} \quad (1.9)$$

The QM energy,  $E^{QM}$ , can be calculated with any of the above discussed SCF energy expressions given in Eqs. (1.6) to (1.8) or corresponding hybrid functional expressions. In all cases the core Hamiltonian matrix elements,  $H_{\mu\nu}$ , are augmented in order to take into account the electrostatic embedding of the QM system by the MM region:

$$H_{\mu\nu} = H_{\mu\nu}^{(0)} - \sum_D^{MM} \langle \mu\nu | \hat{\mathcal{A}}_D(\mathbf{0}) \rangle Q_D \quad (1.10)$$

In Eq. (1.10)  $H_{\mu\nu}^{(0)}$  denotes original core Hamilton matrix elements of the QM system and  $Q_D$  denotes the atomic charges of the MM atoms  $D$ . The general form of the nuclear attraction type operator,  $\hat{\mathcal{A}}_D(\mathbf{0})$ , is given by:

$$\hat{\mathcal{A}}_D(\mathbf{m}) \equiv \left( \frac{\partial}{\partial D_x} \right)^{m_x} \left( \frac{\partial}{\partial D_y} \right)^{m_y} \left( \frac{\partial}{\partial D_z} \right)^{m_z} \frac{1}{| \mathbf{r} - \mathbf{D} |} \quad (1.11)$$

This general definition permits immediately the inclusion of MM atoms with higher point moments. Note that Eq. (1.10) is also used for pure electrostatic embedding [68] with the

EMBED keyword (see 4.2.6). In both cases asymptotic expansions for the long-range nuclear attraction type integrals are implemented in order to improve computational efficiency [69]. Another part of the QM energy in Eq. (1.9) is the MM augmented nuclear repulsion energy,

$$E_{NN} = \sum_{A>B}^{QM} \frac{Z_A Z_B}{|\mathbf{A} - \mathbf{B}|} + \sum_A^{QM} \sum_D^{MM} \frac{Z_A Q_D}{|\mathbf{A} - \mathbf{D}|}, \quad (1.12)$$

which can be calculated analytically from the structure of the QM/MM system. Because the so-defined QM energy contains all quantum mechanical terms plus the electrostatic embedding from the MM region the Kohn-Sham or Hartree-Fock matrix elements can be defined as partial derivatives of this energy with respect to density matrix elements.

The second term in Eq. (1.9) contains the mechanical interaction energy between the QM and MM regions. It is expressed in the form of a Lennard-Jones potential:

$$E^{QMMM} = \sum_A^{QM} \sum_D^{MM} \epsilon_{AD} \left[ \left( \frac{R_{AD}}{|\mathbf{A} - \mathbf{D}|} \right)^{12} - 2 \left( \frac{R_{AD}}{|\mathbf{A} - \mathbf{D}|} \right)^6 \right] \quad (1.13)$$

The  $R_{AD}$  are combinations of the van der Waals radii of QM atom  $A$  and MM atom  $D$ . By default these radii are taken from the MM force field. The parameter  $\epsilon_{AD}$  defines the depth of the Lennard-Jones potential. As for the van der Waals radii it is also taken from the MM force field. Therefore, an MM atom type has to be assigned to each QM atom in the input. This is done with the QM/MM keyword (Section 4.2.4).

The last term in Eq. (1.9) is the MM energy. In deMon2k 5.0 it can contain the following terms:

$$E^{MM} = E^{bond} + E^{bend} + E^{tors} + E^{urey} + E^{vdW} + E^{QQ} \quad (1.14)$$

The first four terms in Eq. (1.14) denote bond stretching, angle bending, dihedral torsion and Urey-Bradley energy terms. Their calculation requires molecular connectivity information that is usually given in the input along with the geometrical definition of the MM atoms under the GEOMETRY keyword (see 4.1.1). As an alternative, the automatic generation of molecular connectivity information on the basis of the distances between MM atoms is also available. The last two terms in Eq. (1.14) represent van der Waals and point-charge interaction energies between the MM atoms. The force fields for MM and QM/MM calculations available in deMon2k are OPLS-AA [45] and AMBER [70]. They are selected by the FORCEFIELD keyword (see 4.2.3) and read from the FFDS (force field dataset) file. For these MM and QM/MM calculations all deMon2k functionalities, such as geometry optimization, transition state finding, molecular dynamics, frequency analysis etc., are avail-

able. Also property calculations for the QM system in QM/MM calculations are possible [71].

Besides the internal MM capability, deMon2k can also be externally interfaced with force fields. To this end a standard interface output for CHARMM [72] can be activated with the QM/MM keyword [44,46,73].

By default, the ERIs are calculated in each SCF cycle (direct SCF) using recurrence relations for near-field ERIs [20,51] and double asymptotic expansions [74] for far-field ERIs. This approach minimizes the random access memory (RAM) demand of deMon2k. If sufficient RAM is available the code performance can be improved by the MIXED option of the ERIS keyword (see 4.5.4). The RAM usage of deMon2k can be monitored by PRINT RAM (see Section 4.12.2 for more details). It also should be noted that, for larger systems, the linear algebra steps in deMon2k may become a bottleneck. With the keywords MATDIA and MATINV (see 4.11.2 and 4.11.3) alternative diagonalizers and matrix inversion techniques can be selected.

Several optimization and transition state search algorithms are implemented in deMon2k. For structure optimization, the default setting is the Levenberg-Marquardt restricted step method in delocalized internal redundant coordinates. This method has excellent convergence behavior and is very robust. However, it requires an iterative back transformation of the coordinates. Thus, to reach tight structure convergence, it may be necessary to switch to Cartesian coordinates at the end of the optimization (see 4.6.1). For ultimate accuracy, this might be combined with a Hessian calculation in each optimization step (UPDATE EXACT; Section 4.6.5). If effective core potentials (ECPs), Section 4.3.4, or model core potentials (MCPs), Section 4.3.5, are used, care must be taken regarding the accuracy of the gradients. Here it may be necessary to tighten the numerical integration threshold with the GRID keyword (see 4.3.6). Usually a FINE grid will be sufficient. The same holds for weak and nonbonded interactions. For the local transition state search, we recommend starting the optimization from a calculated Hessian (see 4.6.5) or restarting it from a frequency analysis (the Hessian from the frequency analysis is then used in the first optimization step). If a SADDLE point interpolation (Section 4.6.2) is to be performed, the starting points must be local minima, *i.e.* reactants and products. All optimizations and interpolations can be restarted with the deMon.new and deMon.mem files. These must be copied into the new input file deMon.inp and the corresponding restart file deMon.rst. The new input file may be modified and extended but the molecular geometry definitions must

be left untouched in order to guarantee a successful restart run.

Born-Oppenheimer molecular dynamics (BOMD) simulations are initialized by the DYNAMICS keyword (see 4.7.1). In these calculations a trajectory file deMon.trj is created which can be large! For compatibility reasons, the trajectory file is written in ASCII (Note that \*\*\*\*\* are used as separations in this file). It should not be modified. The data from the trajectory file can be used to restart BOMD runs or to analyze them (Sections 4.7.2 and 4.7.3). Because BOMD runs may take weeks, we recommend that regular snapshots of the deMon working directory be produced from which restarts are possible. During such a copy the trajectory, deMon.trj, and new input file, deMon.new, must be unchanged. With these files, a restart run is possible just as in the case of structure optimizations, *i.e.* the deMon.new must be copied into the new input file deMon.inp. If requested, the restart file can also be used, e.g. for a restart density (GUESS RESTART; see 4.5.5). However, this is not mandatory.

Usually the default settings of deMon2k are sufficient for standard calculations. However, if extended basis sets are used or higher accuracy is required, it may be necessary to adjust the accuracy and performance settings of the code. This is achieved by the keywords GRID, SCFTYPE and ERIS (see 4.3.6, 4.5.1 and 4.5.4) for the electronic structure calculation, the keyword OPTIMIZATION (see 4.6.1) for the structure optimization and the keywords MATDIA and MATINV (see 4.11.2 and 4.11.3) for the linear algebra parts of the code. The keywords WEIGHTING, QUADRATURE and CFPINTEGRATION control the accuracy settings for the numerical integration (see 4.11.5, 4.11.6 and 4.11.7). The keyword DAVIDSON (4.11.4) controls the iterative diagonalization in time-dependent DFT calculations. In general, modification of the standard settings may alter the performance and accuracy of the code quite substantially. Therefore, such modifications should be tested carefully before being used for production runs.

## 2 Getting Started

### 2.1 Before Installing deMon2k

#### 2.1.1 FORTRAN90 Compiler

To install deMon2k you will need a Unix or Linux operating system running on your machine and a working FORTRAN90 compiler. If you do not have a FORTRAN90 compiler, install it first. You can find different FORTRAN90 compilers on the web. Some are free of charge.

#### 2.1.2 Bourne Again Shell

You will need a Bourne Again Shell (bash) for the installation of deMon2k. The bash is included in almost all current Linux distributions and installed by default. However, if you are running a Unix system like AIX, bash might be missing. If you are not sure if you have the Bourne Again Shell just type

`which bash`

in a console window. If you have bash you should get a message like

`/bin/bash`

If not, you must install bash first. You can check the web for available packages or contact your system administrator.

#### 2.1.3 MPI

The deMon2k code is parallelized using the MPI (Message Passing Interface) routines from the 1.1 standard. The MPI routines are addressed in the deMon2k source by interface subroutines. If you want to run deMon2k in parallel, you need to install the libraries for the MPI 1.1 standard in your system first (they are not included in the deMon2k package).

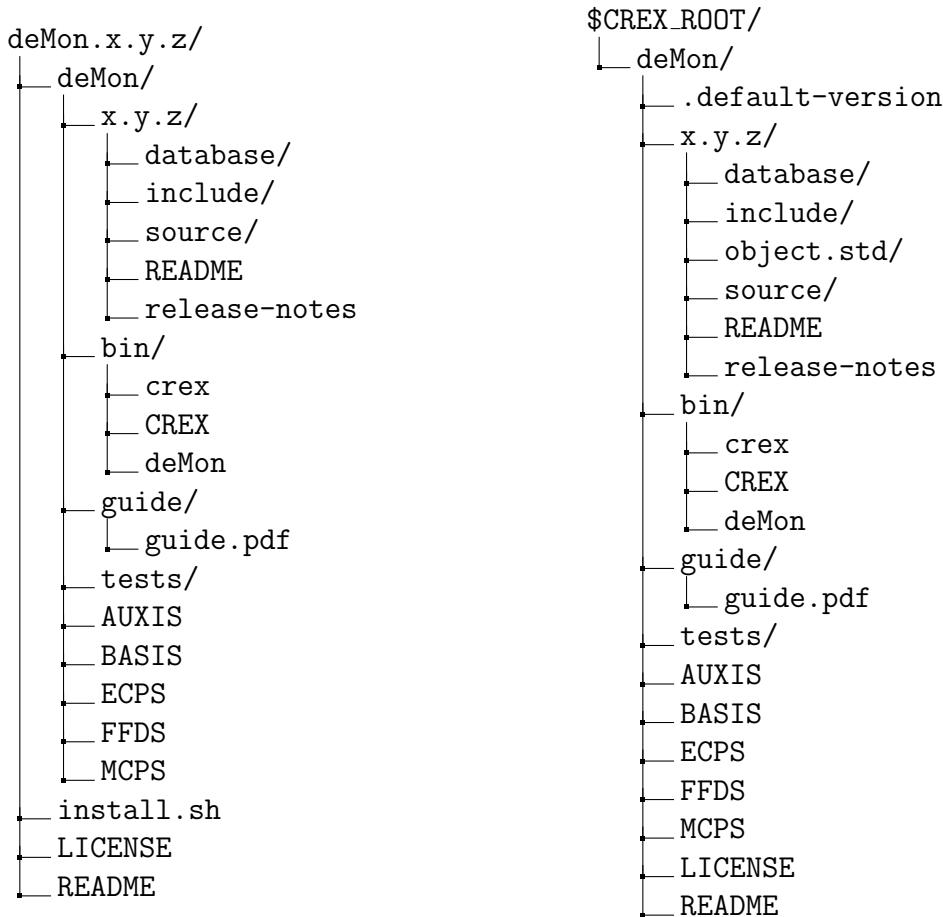
### 2.2 How to Install deMon2k

The deMon2k program is shipped as a packed archive named deMon2k.sh. For the installation, unpack this archive first by typing

`sh ./deMon2k.sh`

This will provide you with an unpacked directory tree `deMon.x.y.z` as depicted in Figure 1, left panel. The numbers x, y and z refer to the version, release and revision number of your deMon2k source.

Figure 1: Directory tree of the unpacked deMon2k version as shipped (left) and of the corresponding deMon2k master version (right).



Inside this directory you will find the deMon2k program tree, the install script, a README file and the LICENSE file (Figure 1, left). If you want to install deMon2k so that it is accessible to all users of a machine you should first become the root. Then run the installation script inside `deMon.x.y.z` by typing

```
./install
```

and follow the instructions of the script. The script will first display the license text which you have to confirm. After this you will be asked for the name of the directory where you want the deMon2k master version to be installed.

Where do you want to install deMon [/usr/local]?

The standard directory for installing deMon2k as root is

\$CREX\_ROOT=/usr/local

To accept this just hit <Enter>, otherwise choose another directory by giving the absolute path e.g. /home/gerald/programs. The script will check if the path exists. If not, it will be created automatically. If you have already installed an older version of deMon2k, the script will ask if you want to update the version. If you choose to do so, the script will install the new version and make it the default. The older versions will not be deleted. If you try to install the same version again, you will be asked, if you want to override the version which is already installed. In that case, the old files will be deleted.

The installation script next tries to identify your operating system and the "platform" you are using. Here "platform" refers to specific hardware and software environments for which compile flags are available in the system files in the deMon2k database directory. Two outcomes are possible:

(A) If your specific environment is unknown to the script, you will have to specify compiler and link flags by yourself. For this purpose, you must edit the file "database.new". This file can be found in the database directory (see left panel of Figure 1). The README file in the database directory or Section 2.5 of this manual will guide you through the customization of the database file for your environment. **We strongly recommend that you start with the database file for the standard compilation.** If you have successfully created a new database file for your specific environment, please email it to [speedy@cinvestav.mx](mailto:speedy@cinvestav.mx) in order that we may include it in the deMon2k distribution.

(B) If your hardware and software environment was recognized by the script, you will be asked for the type of installation.

Choose a compilation type:

1. Standard compilation
2. Debug compilation
3. Optimization compilation
4. Parallel MPI compilation
5. Parallel MPI debug compilation

Select a number:

**We strongly recommend to start with the standard compilation and create other executables after a successful standard compilation.** After your selection of the compilation type, the installation will start. The install script executes the CREX script for the compilation of the deMon2k master version. You will see the message

Compiling with CREX

when CREX takes over control. For more details on the output, see Section 2.3. After the successful installation you will see the message

deMon2k installation successful

You should modify your profile according to the information provided by the installation script at the end of the installation process. If you are updating your deMon2k version, you will be asked if some files like AUXIS, BASIS and so on should be updated also. After the successful installation of the deMon2k master version, a directory tree as shown in the right panel of Figure 1 is created under \$CREX\_ROOT. It must be accessible (read and execute) to all users of deMon2k. If the compilation was unsuccessful, the installation script will halt with the message

deMon compilation failed

In this case you should check the README file in the directory deMon.x.y.z for troubleshooting.

### ADVICE

Different compilers have different options. Sometimes these compile options change within different versions of the same compiler or they are not compatible for different CPUs. Also the installation paths for libraries may change. For these reasons, it is difficult to provide a set of compile options which will work in all environments. Therefore, we strongly recommend that you start the basic installation with a standard compilation. This type of compilation has the highest success rate. The compilation of an optimized version should be done after a successful basic installation. This compilation is described in the next section 2.3. We also recommend that you read Section 2.5 for the specification of compile options.

## 2.3 How to Compile or Re-compile deMon2k

Once deMon2k is installed on your system, you might want to recompile an executable because you need to change some parameters or include some modifications. Or you want to compile an executable of a different type, e.g. an executable for execution in parallel. The

deMon2k program is shipped with two compilation scripts called CREX and crex (CReate EXecutable). CREX is intended for the master version, crex for an individual version.

### CREX

The CREX script can be used to recompile or create a (new) master version of deMon2k. This is done by the installation script also. The "master" version is intended to be an "original" version, *i.e.* a version without any user modifications of the source code. If, during the installation, you have created, for example, a standard version "std" and you want to generate a debug version "dbg", you can do this by running CREX. The master version is usually installed in a directory accessible to all users of a machine. The standard directory is /usr/local. Therefore, you should do this installation as root. CREX expects the following settings:

\$CREX_ROOT=/usr/local	! or another directory where ! deMon2k is/will be installed; ! you have to export this variable
\$CREX_ROOT/deMon/.default-version	! contains the version number ! x.y.z of the version to be ! compiled
\$CREX_ROOT/deMon/x.y.z/source/	! directory for source files
\$CREX_ROOT/deMon/x.y.z/include/	! directory for include files
\$CREX_ROOT/deMon/x.y.z/database/	! directory for database files

The variable \$CREX\_ROOT **must** point to the directory where the master version of deMon2k is installed, if it is not installed in /usr/local. If all required files are there, then type CREX

If you are running on a LINUX system, a menu similar to the following one will appear:

Select compiler (Linux only).

1. Portland Group pgf90
2. Intel Fortran 90
3. xlf (IBM)
4. GNU Fortran 90 (gfortran)

Enter a number:

Choose your compiler and type <Enter> (the menu above is not shown after a successful basic installation). Next, choose which kind of executable is to be created.

Choose a compilation type:

1. Debug compilation
2. Standard compilation
3. Optimization compilation
4. Parallel MPI compilation
5. Parallel debug compilation

Select a number:

After your selection of the compile type, CREX will start the compilation (you will not see this menu in the basic installation because the compile type was chosen earlier). If compilation and linking is successful, CREX will terminate with the message

Executable program deMon.x.y.z.\* created.

Normal termination of CREX.

The executable is written to

\$CREX\_ROOT/x.y.z/object.\*/deMon.x.y.z.\*

where the \* stands for the type of executable that was created, e.g. std if you have generated a standard version. If the compilation fails, the message

Abnormal termination of CREX.

appears. In this case you should check the README file in the directory deMon.x.y.z for troubleshooting.

### **crex**

A local version of deMon2k can be created with the crex script. Thus, each user can create a personal version of deMon2k by adding or modifying source and/or include files without changing the master version. The crex script will compile only the modified files and link them with the unchanged files of the master version. For this reason, the crex script is also well-suited to exchange local developments with other users or between research groups.

The crex script expects the following settings:

```

$CREX_ROOT=/usr/local                                ! or another directory where
                                                       ! the master version of deMon2k
                                                       ! is installed. The default
                                                       ! is /usr/local
$CREX_ROOT/deMon/.default-version                  ! specifies deMon2k master version
$CREX_ROOT/deMon/x.y.z/object.*/                  ! that contains all object files
                                                       ! of the x.y.z.* version. The "*"!
                                                       ! denotes the kind of version,
                                                       ! e.g. "std"
$HOME/demon/                                       ! main directory for the local
                                                       ! version (lower-case characters!)
$HOME/demon/.default-version                      ! specifies local deMon2k version
$HOME/demon/x.y.z/source/                         ! contains the modified source
                                                       ! files that are compiled and
                                                       ! then linked to the x.y.z
                                                       ! master version
$HOME/demon/x.y.z/include/                        ! contains modified include files
                                                       ! (directory must exist,
                                                       ! but can be empty)

```

Type

crex

and follow the instructions. They are similar to those of the CREX script (see above). After successful compilation, the executable is written to

`$HOME/demon/x.y.z/object.*/deMon.x.y.z.*`

If the compilation is unsuccessful, the message

Abnormal termination of crex.

appears. In this case you should check the README file in the directory `deMon.x.y.z` for trouble shooting.

## 2.4 How to Change the Compiler

Suppose, as an example, that for the first compilation the GNU Fortran90 compiler was used and you now want to switch to the Intel® ifort Fortran compiler. To do so you must change the file \$CREX\_ROOT/deMon/.default-version. This file contains only one line similar to

```
x.y.z.gfortran
```

For our particular example it has to be changed to

```
x.y.z.ifdefort
```

Next, you need to identify your system, so execute the command

```
CREX -i
```

You'll get an output like the following

```
Creation of a master executable for deMon:  
Program version: x.y  
Revision      : z  
ifort Fortran Compiler is used  
Platform identified.  
Platform code: Linux-x86_64 (Xeon/Athlon 64-bit architecture)  
CREX system code: Linux-x86_64  
Compiler code: ifort
```

The line CREX system code: is the important one. It points to the new database file. Thus, you have to ensure that in the directory \$CREX\_ROOT/deMon/x.y.z/database the file #.\*.ifdefort exists. Here # stands for the value of the CREX system code, in this case Linux-x86\_64. The \* stands for the type of executable, which can be std, dbg, opt, mpi, pro or dpg. You have to change the compile options in the file #.\*.ifdefort according to your needs. In case such a database file does not exist, it must be created. Before compilation with the new compiler the old object directories must be removed (\rm -rf \$CREX\_ROOT/deMon/x.y.z/object.x)

## 2.5 The Database File

The database files contain the compile and link flags as well as the corresponding options and exceptions in terms of CREX/crex variable assignments. These files are located in the directory \$CREX\_ROOT/deMon/x.y.z/database (see Figure 1). Beware that a lot of

compile options in the database files are for Intel® CPUs. For AMD® CPUs these options must be changed. The name of a database file follows the structure *system.type.comp*. Here *system* stands for the CREX *system* code which describes the operating system you are using. The CREX *system* code can be displayed by the use of the command CREX -i (see section 2.4). The *type* suffix describes the kind of executable you can create with the corresponding database file. Possible *types* are *opt* for an executable that was created with compiler optimization flags, *std* with no optimization during the compilation process, *dbg* for an executable that contains debugging information, *mpi* for an executable which was compiled for parallel execution, *pro* for a parallel executable that contains profiling information and *dpg* for a parallel executable that contains debugging information. The last suffix *comp* identifies the compiler you are using. This name must be consistent with the compiler abbreviation that appears in the file \$CREX\_ROOT/deMon/.default-version. However, it does not have to be identical with the real name of the compiler. A database file template is provided in the directory \$CREX\_ROOT/deMon/x.y.z/database (see Figure 1) by the file database.new. The corresponding README file describes in detail how to customize a deMon2k database file for a new platform.

The content of a database file is divided into assignment lines. Each line is started by a CREX/crex variable or the name of a deMon2k source file. A database assignment line can spawn several physical input lines. To continue an assignment line with the next physical input line, the preceding physical line must be terminated with the backslash “\” symbol. The individual assignment options are separated by blanks. Their syntax is given by the syntax of the compiler options. It is possible to have informative text lines before the first assignment line, *i.e.* at the beginning of the database file. These lines do not have to be marked in any special way.

### 2.5.1 F90 Variable

This CREX/crex variable contains the name of the command to start the FORTRAN compiler. Possible names are *e.g.* gfortran, ifort, pgi, or mpif90.

Some special notes about the compiler for the creation of a parallel executable are important at this point. The compiler mpif90 is often a script/wrapper which calls the actual compiler with predefined compile/link options. The actual compiler depends on the MPI installation. If the MPI installation was done with *e.g.*, the gfortran compiler, then mpif90 will use the gfortran compiler. But it is possible to use, for the compilation of the deMon

source, a compiler which is different from the one used for the MPI installation. In this case the user has to take care about additional include paths, library paths and library names. An example of using the Intel® compiler for the compilation of the deMon2k source code together with MPI libraries created with the gfortran compiler (from OpenMPI) is shown in the database file `Linux-x86_64.mpi.mixed`.

### 2.5.2 FFLAGS Variable

This variable contains the compilation flags. Note that for deMon2k real variable padding, *i.e.* the automatic conversion of 4-byte long REAL variables to 8-byte long REAL variables, is required at the compiler level. Thus, the compiler options for real variable padding must always be present.

### 2.5.3 LFLAGS Variable

This variable contains the flags for the linking step. It should not contain the names of external libraries, *e.g.* the BLAS library. But it should contain the search path for these libraries.

### 2.5.4 LIBS Variable

This variable contains the flags for the names of external libraries. If you are using the Intel® MKL library you can use the Intel® Math Kernel Library Link Line Advisor (<http://software.intel.com/en-us/articles/intel-mkl-link-line-advisor/>) for more information on how you must link these libraries.

### 2.5.5 MPIBIT Variable

This variable is used only in the case of the creation of a parallel executable. It can possess only one out of two possible settings. One MPIBIT setting is 8x4, which means you are going to use MPI libraries which support 8-byte long REAL and 4-byte long INTEGER variables. This is also the default value for the MPIBIT variable. The other setting is 8x8, which means that you are going to use MPI libraries which support 8-byte long REAL and 4-byte long INTEGER variables whereas the deMon2k program itself uses 8-byte long INTEGER. Note that the MPIBIT variable itself does not automatically sets the compilation flags for 8-byte long INTEGER. Instead, the compile option for 8-byte long INTEGER has to

be set additionally with the FFLAGS variable (ifort: -i8; gfortran: -fdefault-integer=8. See also the `Linux-x86_64.mpi.ifort` file in the database directory of the shipped version as an example for the 8x8 setting of the MPIBIT variable.

### 2.5.6 EXCLUDE Variable

This variable contains the names of the source files which are **not** to be compiled because they are part of external libraries. This choice is mostly relevant to the source files for BLAS and LAPACK routines. See the `system.opt.comp` files in the database directory for examples.

### 2.5.7 Source File Name Variables

These variables carry the names <filename> of source files. After the source file name, the compile options for that specific source file are given. These compile options are used only for the compilation of the source file <filename>. They substitute for the global compile flags given by the variable FFLAGS. This flexibility is useful e.g. in case you have to reduce the optimization level for certain source files. There can be several of these lines at the end of the database file, each one for a different source file. In the case of the 8x8 setting for the MPIBIT variable it is mandatory to compile the files `mp_abort.f`, `mp_collect.f`, `mp_end.f`, `mp_get.f`, `mp_send.f`, `mp_send_all.f`, `mp_start.f`, `mp_sum.f`, `mp_sumonone.f` and `mp_sync.f` with 4-byte long INTEGER while 8-byte long INTEGER are used via the FFLAGS variable for the other source files of the deMon2k program.

## 2.6 How to Run deMon2k

You can run deMon2k either directly or with the deMon script. Because the script must be adapted to your environment you will typically start with direct runs and then utilize a script.

### 2.6.1 Running deMon2k Directly

Generate a working directory, e.g.

```
mkdir /tmp/job
```

Copy your input file (`your.inp`) into this directory. Note that the name of the input file in the

working directory **must** be deMon.inp. Also copy or link the AUXIS, BASIS, ECPS, FFDS and MCPS files into this directory. These files are located in the \$CSEX\_ROOT/deMon directory (see Figure 1, right panel). Assuming the default value /usr/local for \$CSEX\_ROOT, you should execute the commands

```
cp your.inp /tmp/job/deMon.inp
cp /usr/local/deMon/AUXIS /tmp/job/AUXIS
cp /usr/local/deMon/BASIS /tmp/job/BASIS
cp /usr/local/deMon/ECPS /tmp/job/ECPS
cp /usr/local/deMon/FFDS /tmp/job/FFDS
cp /usr/local/deMon/MCPS /tmp/job/MCPS
```

Change to the directory

```
cd /tmp/job
```

Start the execution of deMon2k by typing

```
/usr/local/deMon/x.y.z/object.*/deMon.x.y.z.*
```

The \* denotes the kind of deMon2k version, *e.g.* opt, you are using for your calculation. If you want to run the program in the background in order to continue working in the foreground you have to start the executable by typing

```
nohup /usr/local/deMon/x.y.z/object.*/deMon.x.y.z.* &
```

In case of the parallel executable (\* = mpi) the execution has to be started according to the rules of the installed MPI system. In case of MPICH or OpenMPI the execution has to be started by typing

```
mpirun -np N /usr/local/deMon/x.y.z/object.mpi/deMon.x.y.z.mpi
```

Here *N* denotes the number of CPU cores one wants to use. If the calculation is spread over physically different nodes, one has to assure that the working directory exists on each node.

Output from the calculation is in the file deMon.out.

## 2.6.2 Running the deMon Script

By default, a deMon script for an NQS/PBS queuing system is installed in the directory \$CSEX\_ROOT/deMon/bin (Figure 1). This script is used in the research groups at CINVESTAV for working with deMon2k. **It is only an example of what can be done and certainly must be modified for other queuing systems and/or environments.** However, by using this script as a template, modifications for other systems and environments should be

straightforward. As a language reference for bash scripting, we suggest [75]. To run a job with the deMon script just type

```
deMon your.inp
```

As in the foregoing example `your.inp` denotes the input file which must be in the current directory. You are asked to select an execution mode. If you are running NQS/PBS it could look similar to the following (`serial`, `pl16`, `pl32` and `pl12` are batch queues on the `rodio` mainframe at Cinvestav):

Select an execution mode.

1. Interactive job (background)
2. Interactive job (foreground)
3. NQS/PBS batch job in queue `serial` at `rodio`
4. NQS/PBS batch job in queue `pl16` at `rodio`
5. NQS/PBS batch job in queue `pl32` at `rodio`
6. NQS/PBS batch job in queue `pl12` at `rodio`

Enter a number:

After you have selected the execution mode, *i.e.* the queue (and machine) to which the job is submitted, the script checks which deMon2k versions are available. For this purpose the script tries to read first the version number in your home directory in the file `$HOME/deMon/.default-version`. If this file is not available, it will take the version which is specified in the `$CREX_ROOT/deMon/.default-version` file. You can then choose between the available executables, *e.g.*:

The following program versions are available:

1. `/usr/local/deMon/x.y.z/object.opt/deMon.x.y.z.opt`
2. `/usr/local/deMon/x.y.z/object.dbg/deMon.x.y.z.dbg`
3. `/usr/local/deMon/x.y.z/object.dpg/deMon.x.y.z.dpg`
4. `/usr/local/deMon/x.y.z/object.pro/deMon.x.y.z.pro`
5. `/usr/local/deMon/x.y.z/object.mpi/deMon.x.y.z.mpi`
6. `/home/florian/demon/x.y.z/object.opt/deMon.x.y.z.opt`
7. `/home/florian/demon/x.y.z/object.dbg/deMon.x.y.z.dbg`

Enter a number:

After choosing the executable, the calculation is started. The output files will be written to `$HOME/demon/x.y.z/outputs` after the calculation is finished. You must create this directory manually. If this directory doesn't exist, the deMon script will stop the execution with the following error message

```
deMon: The directory /home/florian/demon/x.y.z/outputs does not exist
```

Abnormal termination of deMon.

You can also specify a different output directory. Type

```
deMon -h
```

to get further information on the deMon script.

## 3 Carrying On

### 3.1 Parameters

As shipped, deMon2k has some preset limits for memory and disk space requirements, number of atoms, and number of basis and auxiliary functions. Table 1 summarizes those as-shipped limits.

Table 1: Parameter settings in deMon2k as shipped.

Parameter	Setting	Description
MAXDISK	51200	Maximum disk size for scratch files in Mbytes
MAXRAM	2048	Maximum RAM size for program kernel in Mbytes
PRORAM	256	RAM size reserved for program data in Mbytes
MAXATMM	15000	Maximum number of MM atoms
MAXATOM	1000	Maximum number of QM atoms
MAXAUX	20000	Maximum number of auxiliary functions
MAXAUXSET	10000	Maximum number of auxiliary function sets
MAXAUXSHL	10000	Maximum number of auxiliary function shells
MAXCON	21	Maximum degree of contraction
MAXCUBE	500000	Maximum number of (cube) embedding points
MAXECPGTO	10000	Maximum number of ECP Gaussian functions
MAXECPSHL	3000	Maximum number of ECP shells
MAXGTO	20000	Maximum number of primitive Gaussian functions
MAXLAUX	6	Maximum L quantum number for auxiliary functions
MAXLBAS	6	Maximum L quantum number for basis functions
MAXLECP	5	Maximum L quantum number for ECP functions
MAXLINK	4	Maximum number of molecular mechanics links
MAXLMCP	3	Maximum L quantum number for MCP functions
MAXLSUM	8	Maximum L quantum number for double asymptotic ERIs
MAXMCPGTO	10000	Maximum number of MCP Gaussian functions
MAXMCPSHL	3000	Maximum number of MCP shells
MAXMCPSTO	6000	Maximum number of contracted MCP (STO) orbitals
MAXPAT	10	Maximum number of geometry pattern
MAXSHL	5000	Maximum number of (STO) orbital shells
MAXSTO	10000	Maximum number of contracted (STO) orbitals
MAXTBSTO	10000	Maximum number of tight-binding (STO) orbitals

These limits are specified in the section `user defined parameters` of the `parameter.h` file. If a calculation violates these limits, the program stops and prints `PARAMETER EXCEEDED` statement(s) for the parameters that must be changed. In this case, please change only the first parameter listed as exceeded to the suggested (`USED`) value and

recompile the program (the following ones may be corrupted by array boundary violations). Repeat this procedure until all PARAMETER EXCEEDED statements disappear. We recommend that you **not** increase the MAXRAM value over the physical memory (RAM) available. Doing otherwise may result in large paging overhead during program execution. Remember also that the size for MAXRAM is given in MB/Core, *i.e.* the total amount of RAM assumed available for the program in case of a parallel execution equals MAXRAM  $\times$  number of cores.

In the case of a conventional SCF calculation, disk requirements are dominated by the size of the electron repulsion integral scratch file `ioeri.scr`. In the case of a direct SCF (see 4.5.1) calculation, the files `ioscf.scr`, `iocdf.scr`, and `iogrd.scr` dominate. The files `4GBYTE.H` and `6GBYTE.H` contain alternative parameter settings for larger deMon2k versions with up to 1000 and 2000 atoms, respectively. If one of these settings is to be used, copy the corresponding file to `parameter.h`.

## 3.2 Files

All files in deMon2k are opened with filenames. Throughout this Guide, the name used in the FORTRAN OPEN statement is referred to as the logical filename. These logical filenames are read by the program or created in the working directory.

The logical files are connected to FORTRAN I/O units in the file `fileio.h`. Table 2 provides the logical and physical filenames used in the program.

**From deMon2k 5.0 on the input file `deMon.inp` is modified into a fixed standard structure and, therefore, overwritten during program execution. The same applies for the geometry file `deMon.geo`. Thus, these files should be copied prior to program execution in order to preserve the original input formats.**

Besides these files, special plot files may be created for the visualization of molecular fields, like the electronic density, electrostatic potential, etc., with VU [41]. These files are named according to the molecular field to be visualized. For example, a file `RHO.bin` is created for the visualization of the electronic density. The `deMon.pie` file is the VU control file and refers in this case to the `RHO.bin` file. If the job is executed by a script, care must be taken that these files are copied from the working directory to the output path and that the filenames are consistent with the ones in the `deMon.pie` file. These entries can also be modified manually in order to use the same VU control file for different molecular fields.

Table 2: Physical and logical filenames in deMon. The I/O units are given in parentheses.

Physical Filename	Logical Filename	Description	Type
AUXIS	AUX (1)	Auxiliary function file	ASCII
BASIS	BAS (2)	Basis set file	ASCII
ECPS	ECP (3)	ECP file	ASCII
MCPS	MCP (4)	MCP file	ASCII
FFDS	FFDS (5)	Force field dataset file	ASCII
deMon.inp	INP (6)	Input file	ASCII
deMon.out	OUT (7)	Output file	ASCII
deMon.geo	GEO (8)	Molecular geometry file	ASCII
deMon.new	NEW (9)	New input file for restart	ASCII
deMon.trj	TRJ (10)	Trajectory file	ASCII
deMon.mol <sup>a</sup>	MOL (11)	MOLDEN input file	ASCII
deMon.mkl <sup>a</sup>	MOL (11)	MOLEKEL input file	ASCII
deMon.wfn <sup>a</sup>	MOL (11)	XAIM input file	ASCII
deMon.qmm	QMM (12)	QM/MM interface file	ASCII
deMon.xry	XRY (13)	XAS/XES analysis file	ASCII
deMon.nbo	NBO (14)	NBO 5.G input file	ASCII
deMon.pie	PIE (15)	VU control file	ASCII
deMon.mem	MEM (16)	MEMORY file	BINARY
deMon.rst	RST (17)	RESTART file	BINARY
deMon.cub	CUB (18)	Embedding file	ASCII
deMon.lat	LAT (19)	Plot lattice file	ASCII / BINARY
deMon.edu	EDU (20)	Input file for reactant	ASCII
deMon.pro	PRO (21)	Input file for product	ASCII

<sup>a</sup>The I/O unit 11 is connected either to the MOLDEN, MOLEKEL or XAIM input file depending on the VISUALIZATION (4.10.1) keyword. For the XYZ output the MOLDEN file (deMon.mol) is used in all cases.

### 3.3 Input Syntax

The input of deMon2k is easy and mnemonic and at the same time offers high flexibility. The input file contains keywords, options and keyword bodies. A keyword together with its options, and the keyword body forms a keyword block. The ordering of these blocks is free. The keyword body starts in a new input line directly after the keyword line(s), and must have the appropriate structure given in Chapter 4. All keywords, with the exception of GEOMETRY or its equivalent, REACTANT and PRODUCT, have associated default values which are used if the keyword is not explicitly specified in the input. Thus, the simplest deMon2k input contains only the coordinates of the atoms in the molecule to be calculated in the keyword body of GEOMETRY. deMon2k uses special symbols to make the keyword

input flexible and efficient. Table 3 summarizes the special symbols that are allowed in the job input file deMon.inp.

Table 3: Special symbols of the deMon input.

Symbol	Description
&	Continuation of a keyword line
#	First character of a comment line
=	Assignment of numerical values to options
>	Mathematical greater than symbol
<	Mathematical smaller than symbol
^	Mathematical power symbol
-	Combination of keyword options
Space	Separation of keyword and options
,	Separation of different options (optional)
END	End of a keyword data block

Lines in the job input file deMon.inp are restricted to 80 characters. The input is not case sensitive. A keyword line can be continued by the & symbol at the very end of the line. At most, five continuation lines are permitted. The lines in the keyword body cannot be continued and, therefore, are limited to 80 characters. **The TITLE line also cannot be continued and the length of the title string is restricted to 60 characters.**

The symbol # indicates a comment line. Comment lines can occur only between complete keyword blocks including the keyword body, if it exists (Placing comments inside a keyword line or body results in unpredictable read operations). The end of a keyword block can be indicated by an END statement, by a comment symbol # in a new input line, or simply by a new keyword line.

Keyword options must be given directly after the keyword, separated by space(s). If options are not present, the default settings for them are used. Numerical values are assigned by the = sign to keyword options (e.g. SCFTYPE Tol = <Real>, where <Real> indicates a real number). The – sign is reserved for combining keyword options (e.g. VXCTYPE B88-LYP) where that is allowed. The other mathematical symbols in Table 3 are used according to their mathematical meaning. They can be surrounded by space(s).

## 3.4 Job Restart

For several applications it is convenient to restart a deMon2k job. This is done with the RESTART option of the keywords GUESS, HESSIAN, FREQUENCY and TRAJECTORY (see 4.5.5, 4.6.5, 4.8.1 and 4.7.2, respectively). During all runs deMon2k generates a memory file `deMon.mem` and a new input file `deMon.new`. The memory file contains the electronic structure information of the current job in the form of the molecular orbital coefficients as well as the corresponding occupation numbers and eigenvalues. Optionally, it may also contain the Hessian matrix. The file is in binary format and, thus, not human-readable. On the other hand, the new input file `deMon.new` is an ASCII file that contains an updated copy of the original input file `deMon.inp`. To perform a job restart, these two files must be copied. The memory file `deMon.mem` must be copied to the restart file `deMon.rst` and the new input file `deMon.new` to the input file `deMon.inp`. The input file may be modified. For the restart of a Born-Oppenheimer molecular dynamics simulation, the deMon2k trajectory file `deMon.trj` must also be available.

## 4 Keywords

The description of the keywords is grouped as follows according to their functionality: geometry input, methodology selection, atom specific inputs (basis sets, auxiliary functions etc.), electronic state control, self-consistent field (SCF) control, optimization and transition state search, Born-Oppenheimer molecular dynamics, molecular properties, electronic structure analysis, visualization/topological analysis, performance/accuracy settings, and miscellaneous. Keyword options which are mutually exclusive, *i.e.* are incompatible with one another, are listed on one line separated by / signs. If none of the options is specified in the input, the underlined one is used by default. If more than one of these options is given, the last one listed will override the preceding ones. Non-exclusive options are separated from exclusive ones by an extra blank line in the description block. The boldface printed part (first five letters) of the keywords and options are mandatory for the correct input interpretation by deMon2k. Therefore, **OPTIMIZATION**, **OPTIMIZE**, and **OPTIM** are all allowed input forms for the keyword OPTIMIZATION.

### 4.1 Molecular Geometry Input

#### 4.1.1 Keyword GEOMETRY

This keyword is mandatory! It specifies the molecular geometry.

Options:

##### CARTESIAN / ZMATRIX / MIXED

**CARTESIAN**      The molecular structure is given in Cartesian coordinates.

**ZMATRIX**      The molecular structure is defined by a Z-matrix. The form **Z-MATRIX** is equivalent to **ZMATRIX**.

**MIXED**      The first atoms of the molecular structure are defined by Cartesian coordinates and the following ones are defined by a Z-matrix.

##### ANGSTROM / BOHR

**ANGSTROM**      Coordinates or bond distances are given in Ångströms.

**BOHR**      Coordinates or bond distances are given in atomic units.

Description:

The geometry is read in the keyword body of GEOMETRY in free format, one line for each atom. In the case of a CARTESIAN input the atomic symbol (*e.g.* H), which may carry an

identification number (*e.g.* H1), and the x, y and z coordinates of each atom of the system must be specified. As an example, the geometry of H<sub>2</sub>O may be specified as:

```
GEOMETRY CARTESIAN ANGSTROM
O  0.00  0.00  0.00
H  0.76  0.00  0.52
H -0.76  0.00  0.52
```

The Cartesian coordinates of an individual atom, defined by the atomic symbol (*e.g.* H2), or an atom group, defined by the element symbol (*e.g.* H), can be frozen during the geometry optimization in redundant coordinates. To keep all three degrees of freedom of an atom fixed the corresponding atomic symbol or element symbol has to be specified with the string XYZ in the keyword body of CONSTANTS. The string X will freeze only the x coordinate of the corresponding atom, the string XY fixes the x and y coordinates, etc. In the following example, the oxygen atom is kept frozen and the movements of all hydrogen atoms are restricted to the xz-plane (y coordinate frozen) during the optimization.

```
GEOMETRY CARTESIAN ANGSTROM
O  0.00  0.00  0.00
H  0.76  0.00  0.52
H -0.76  0.00  0.52
#
CONSTANTS
O XYZ
H Y
```

As usual, the comment line (#) is given only for clarity. The effective nuclear charge (in atomic units), which is identical to the number of electrons counted for the atom, and the nuclear mass (in atomic mass units; amu) can be specified by an integer and real number after the coordinates, respectively. Therefore, the following input for a water-like system consists of a frozen oxygen atom, a hydrogen atom, H1, fixed in the xz-plane with the effective nuclear charge 0, *i.e.* a ghost atom that can be used to define basis functions at a point in space, and another hydrogen atom, H2, with the nuclear charge 1 and the nuclear mass 2.014, *i.e.* a deuterium atom.

```
GEOMETRY CARTESIAN ANGSTROM
O  0.00  0.00  0.00
H1 0.76  0.00  0.52  0
H2 -0.76  0.00  0.52  1  2.014
#
CONSTANTS
O XYZ
H1 Y
```

If quantities are not specified, default values are used (*e.g.* the O atom in the example above has a nuclear charge of 8 and a mass of 15.999 amu).

Negative integer numbers after the atomic coordinates indicate an MM atom. The absolute value of this integer number specifies the atom type in the selected force field. These atom types can be found in the FFDS file. Thus, the geometry part of an MM water input can have the form:

```
GEOMETRY CARTESIAN ANGSTROM
O   0.00  0.00  0.00  -76
H1  0.76  0.00  0.52  -77
H2 -0.76  0.00  0.52  -77
```

Here 76 and 77 are the atom types of the MM oxygen and hydrogen atoms, respectively. Note that the connectivity between the atoms is generated automatically. The following input shows the geometry definition of a QM/MM water dimer, the first one being QM and the second one being MM.

```
GEOMETRY CARTESIAN ANGSTROM
O    1.159094   0.359472   -1.228183      8   15.999400
H    0.165274   0.461227   -1.247218      1   1.007940
H    1.357609   -0.118689   -2.058595      1   1.007940
O   -1.496793   0.620303   -1.289943     -76
H   -2.123457   0.067624   -0.732608     -77
H   -1.996365   1.430082   -1.611984     -77
```

For the QM atoms the nuclear charges and masses are also explicitly defined in this input. Again the connectivity of the MM atoms is generated automatically.

Even though it is usually advantageous to sort a QM/MM geometry definition according to the QM and MM atoms, this is not mandatory for a deMon2k input. In the following input QM and MM atoms are mixed among themselves. Note also the extended element symbols. This naming must be chosen such that the first two characters remain specific for the element. Thus, Hm and Hq used below are possible choices for hydrogen whereas He or Ho would denote helium or holmium, respectively!

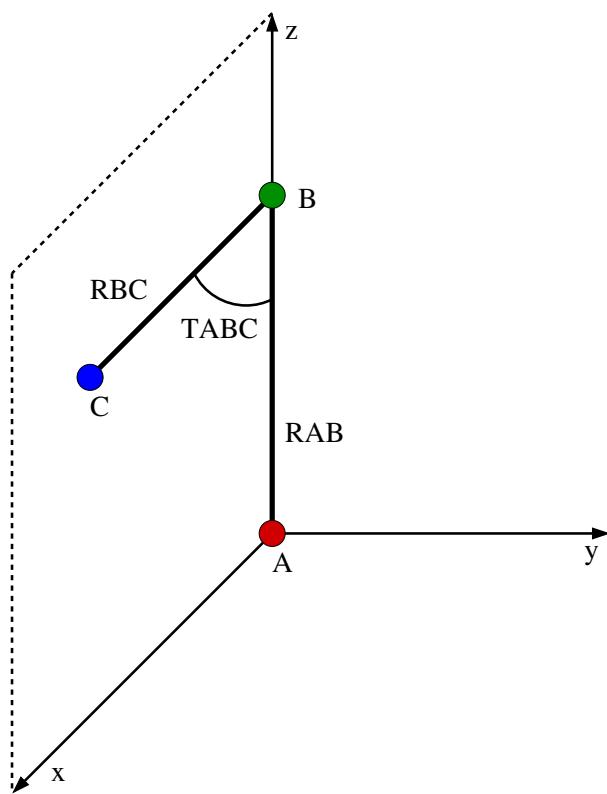
```
GEOMETRY CARTESIAN ANGSTROM
Hm1b  1.262389   -1.937796   1.222724     -77   6
Hq1a  0.217713   0.505507   -1.369657
Hm1a  1.124357   -0.596134   0.419304     -77   6
Hq1b  1.503324   -0.129952   -1.999029
Oq2  -1.628323   0.600651   -1.413703
Om1  0.867185   -1.043515   1.275905     -76   1   3
Hq2a -1.885969   0.153697   -0.557057
Om2  -1.948708   -0.776469   1.058175     -76   10  12
Hq2b -2.021673   1.495719   -1.360149
Hm2a -0.978710   -0.948944   1.231841     -77   8
Oq1  1.187632   0.332759   -1.195914
Hm2b -2.264749   -0.314968   1.861890     -77   8
```

In this input the integer numbers after the negative atom type specifications of the MM atoms denote their connectivity. Thus, hydrogen atom Hm1b and Hm1a are connected to

atom 6 which is the MM oxygen atom  $\text{O}_{\text{m}1}$ . For this atom two connectivity integer numbers are given, 1 and 3, that indicate the connectivities to atom 1,  $\text{H}_{\text{m}1\text{b}}$ , and atom 3,  $\text{H}_{\text{m}1\text{a}}$ .

If the geometry of the system is given by a Z-matrix, using the option ZMATRIX, each line in the keyword body of GEOMETRY describes the connectivity of the atom, which again is defined by its atomic symbol. The first atom (A) of the Z-matrix defines the origin of the input (see Figure 2) and therefore has no connectivity information.

Figure 2: Orientation of the first three atoms defined in the Z-matrix.



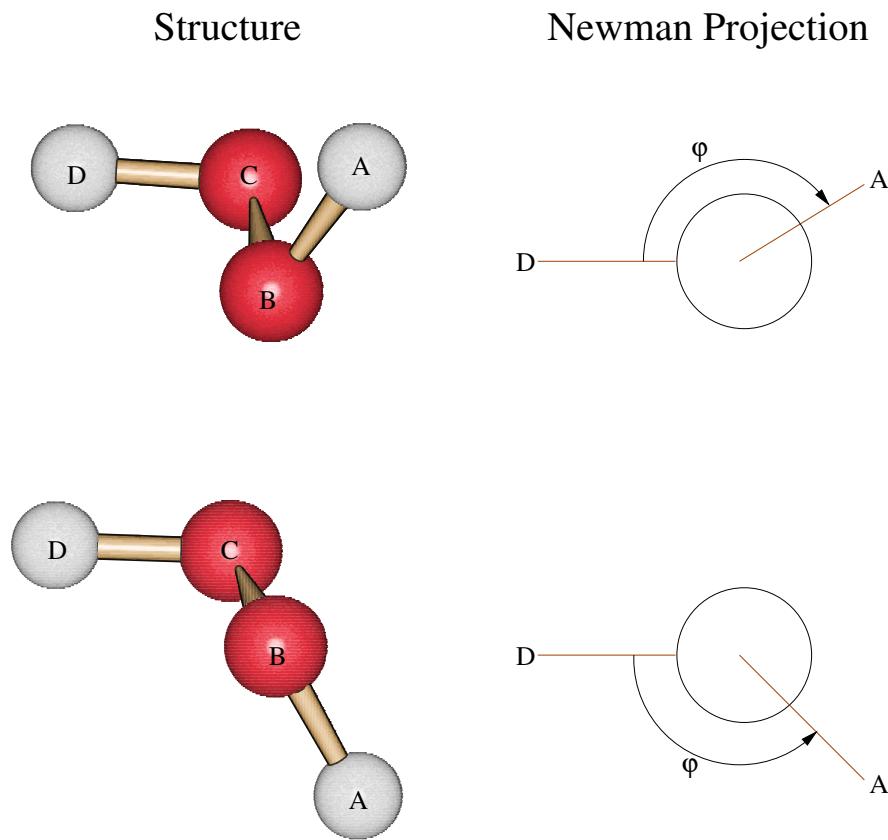
The second atom (B) lies on the z-axis of the input coordinate system and is connected to the first atom at a distance RAB. The third atom (C) lies in the xz-plane and connects to the two previously defined atoms at a distance RBC and an angle TABC. Each of the following atoms is defined by an input line in free format containing the information:

LABEL	NB	RAB	NC	TABC	ND	PABCD
-------	----	-----	----	------	----	-------

Here LABEL is the atomic label of atom A. NB, NC, and ND are the Z-matrix line numbers or the atomic labels of the atoms B, C and D with respect to which atom A is defined by a length (RAB), angle (TABC), and dihedral angle (PABCD). These are the so-called internal

coordinates of the system. Their values are given by RAB (in Ångström or Bohr), TABC, and PABCD (in degrees), respectively. The maximum length of these numbers in the Z-matrix is limited to ten characters. The dihedral angle is defined by the angle between the planes spanned by the atoms {A, B, C} and {B, C, D}, respectively. The sign of the dihedral angle is defined according to the Newman projection shown in Figure 3. If the projection angle is oriented clockwise then the dihedral angle is positive.

Figure 3: Definition of positive (top) and negative (bottom) dihedral angles in deMon2k.



Instead of giving the values for RAB, TABC, and PABCD directly in the Z-matrix, they may also be represented by symbolic strings with no more than eight characters each. These strings must be assigned to CONSTANT or VARIABLE values (see 4.1.3 and 4.1.5) in the input, as the following example of the C<sub>2</sub>H<sub>4</sub> input demonstrates:

```
GEOMETRY ZMATRIX ANGSTROM
C1
C2  C1  rCC
H1  C1  rCH  C2  aCCH
H2  C1  rCH  C2  aCCH  H1  dHCCH
```

```

H3  C2  rCH  C1  aCCH  H2  dHCCH
H4  C2  rCH  C1  aCCH  H1  dHCCH
#
VARIABLES
rCC  1.3390
rCH  1.0850
aCCH 121.09
#
CONSTANT
dHCCH 180.0

```

The internal coordinates listed under VARIABLES are changed during geometry optimization, whereas the coordinates listed under CONSTANTS are kept frozen. The RAB, TABC, and PABCD values that are directly defined in the Z-matrix are always interpreted as constants in internal or internal redundant structure optimizations.

If one uses internal coordinates, dummy atoms are often useful. The atomic symbol X has been reserved for them (note that the symbol XX is reserved for automatically generated dummy atoms; it is not to be processed as an input symbol). The following example shows the use of a dummy atom in the definition of HCN.

```

GEOMETRY ZMATRIX ANGSTROM
H
C 1 R1
X 2 1.0 1 90.0
N 2 R2 3 A1 1 180.0
#
VARIABLES
R1 1.089
R3 1.166
A1 90.00

```

Here the dummy atom is used to avoid a 180° angle which might cause problems in a geometry optimization. Please note that the structure parameters which are defined inside the Z-matrix are kept constant during structure optimization. Example 2.2 on page 4 of the deMon2k tutorial shows the use of a dummy atom for the definition of a ring system.

Table 4: Substitution strings for the internal coordinate substitution in the Z-matrix.

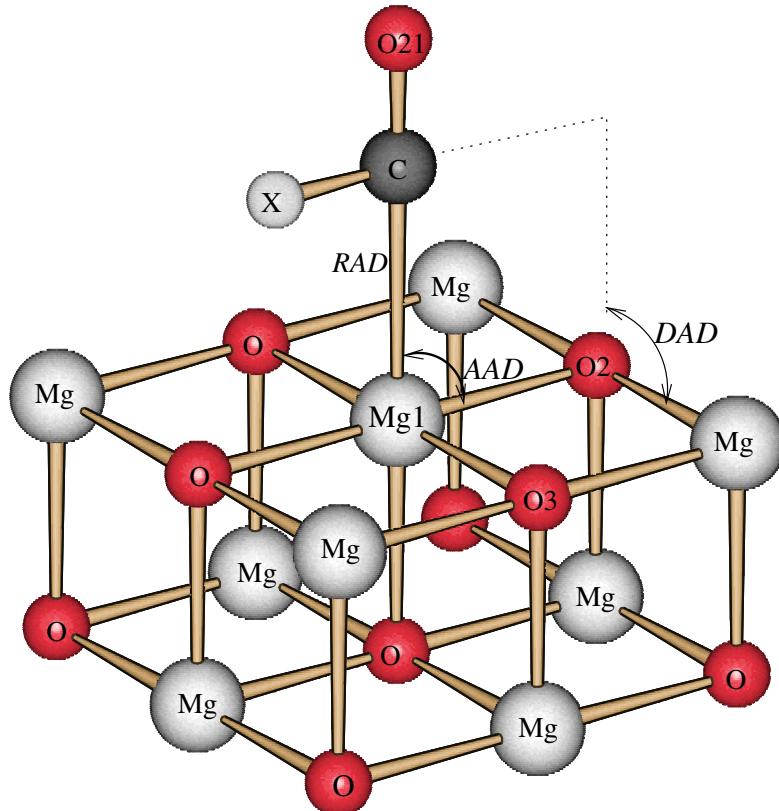
Subst. String	Internal Coordinates	Description
RAD	RAB TABC PABCD	No substitution.
RRD	RAB RAC PABCD	Length AC substitutes for angle ABC.
RAA	RAB TABC TABD	Angle ABD substitutes for dihedral ABCD.
RRA	RAB RAC TABD	Combination of RRD and RAA.
RRR	RAB RAC RAD	Definition over lengths only.

To facilitate the input of the internal coordinates, substitutions of angles and dihedral angles are possible in the Z-matrix input. The substitutions are indicated by strings after the

connectivity information in each Z-matrix line. The possible substitution strings are listed in Table 4. Thus, the connectivity information in a Z-matrix line can be extended by the substitution string, the nuclear charge, and the atomic mass, respectively. This ordering is necessary! However, defaults can be omitted. That is, the definition of a substitution string and an atomic mass will work and the default nuclear charge will be assigned to the given atom. The examples 2.3 and 2.4 on page 5 and page 6 of the tutorial show some common applications of Z-matrix inputs.

If the geometry of the system is given by a MIXED input, the first few atoms are defined by Cartesian coordinates. At least the first three atoms, including dummy atoms, must be defined in this way. Subsequent atoms can then be defined by a Z-matrix line as described above. The reference atoms in that Z-matrix line can be defined either by Cartesian or internal coordinates. In either case, the coordinate system orientation is defined by the first three atoms. The following example, depicted in Figure 4, shows a MIXED input for a CO molecule adsorbed on a MgO cluster.

Figure 4: Coordinate definition of the example MIXED input.



```

GEOMETRY MIXED ANGSTROM
Mg1    0.000  0.000  0.000
O2     0.000  2.105  0.000
O3     2.105  0.000  0.000
O4     0.000 -2.105  0.000
O5    -2.105  0.000  0.000
Mg6    2.105  2.105  0.000
Mg7    2.105 -2.105  0.000
Mg8   -2.105 -2.105  0.000
Mg9   -2.105  2.105  0.000
O10    0.000  0.000 -2.105
Mg11   0.000  2.105 -2.105
Mg12   2.105  0.000 -2.105
Mg13   0.000 -2.105 -2.105
Mg14  -2.105  0.000 -2.105
O15    2.105  2.105 -2.105
O16    2.105 -2.105 -2.105
O17   -2.105 -2.105 -2.105
O18  -2.105  2.105 -2.105
C19    Mg1    RAD    O2      AAD    O3      DAD
X20    C19    1.0    Mg1    90.0    O2      180.0
O21    C19    RCO    X20    90.0    Mg1    180.0
#
CONSTANTS
Mg    XYZ
O2    XYZ
O3    XYZ
O4    XYZ
O5    XYZ
O10   XYZ
O15   XYZ
O16   XYZ
O17   XYZ
O18   XYZ
#
VARIABLES
RAD   2.0
AAD   90.0
DAD   90.0
RCO   1.4

```

This example calls for the free optimization of a CO molecule, including internal relaxation, on a frozen MgO cluster surface. Thus, the MgO cluster is kept constant by freezing its Cartesian coordinates (see 4.1.3). The carbon atom of the CO molecule is defined with respect to the Mg1, O2, and O3 atoms of the cluster (see Figure 4) through the internal coordinates RAD, AAD, and DAD. These coordinates are free for optimization because they are declared as variables (see 4.1.5). The same holds for the CO bond length that is defined by the internal coordinate RCO. The dummy atom X20 must be used to avoid an ill-defined dihedral angle if the O21, C, Mg1 arrangement becomes linear.

In deMon2k Z-MATRIX and MIXED geometry definitions are also available for MM and QM/MM calculations. Note that the dummy atom line in the following Z-matrix definition for an MM input has to be taken into account for the MM connectivity numbering.

```

GEOMETRY Z-MATRIX ANGSTROM
O                               -76   15.999400 2 3
H     1  R1                   -77   1.007940 1
H     1  R2       2  A1      RAD  -77   1.007940 1
H     1  R3       2  A2      3  D1      RAD  -77   1.007940 6
X1    4  RCON    1  ACON    2  DCON    RAD   0    0.000000
O     4  R4       5  A3      1  D2      RAD  -76   15.999400 4 7
H     6  R5       4  A4      5  D3      RAD  -77   1.007940 6
#
VARIABLES
R1      1.00275988
R2      1.00278094
R3      1.72377877
R4      1.01622488
R5      0.99833459
A1      108.809982
A2      122.743681
A3      86.814590
A4      108.264348
D1      -152.130713
D2      181.980680
D3      33.280793
#
CONSTANTS
RCON    2.00000000
ACON    90.000000
DCON    -138.000000

```

#### 4.1.2 Keywords REACTANT and PRODUCT

These two keywords must appear together (cannot be used separately, but can be used in either order). They define the reactant and product structures for a saddle-point (see 4.6.2) interpolation of a transition state structure.

Options:

##### CARTESIAN / ZMATRIX / MIXED

- |                  |   |
|------------------|---|
| <b>CARTESIAN</b> | The molecular structure is given in Cartesian coordinates.  |
| <b>ZMATRIX</b>   | The molecular structure is defined by a Z-matrix. Equivalent to the option <b>ZMATRIX</b> is the option <b>Z-MATRIX</b> .         |
| <b>MIXED</b>     | The first atoms of the molecular structure are defined by Cartesian coordinates and the following ones are defined by a Z-matrix. |

##### ANGSTROM / BOHR

- |                 |  |
|-----------------|--|
| <b>ANGSTROM</b> | Coordinates or bond distances are given in Ångströms.    |
| <b>BOHR</b>     | Coordinates or bond distances are given in atomic units. |

Description:

The reactant and product geometries for a saddle-point interpolation are read in the key-

word bodies of REACTANT and PRODUCT, respectively. Both structures must be local minima on the potential energy surface. If the molecular structures are defined by a Z-matrix, two different VARIABLE definitions are allowed but CONSTANTS must be the same for both REACTANT and PRODUCT . In the following input example, the first variable definition is assigned to the product whereas the second one is assigned to the reactant.

```

PRODUCT Z-MATRIX ANGSTROM
C                               6   12.011000
C     1   R1                   6   12.011000
O     1   R2                   8   15.999400
H     3   R3                   1   1.007940
H     2   R4                   1   1.007940
H     1   R5                   1   1.007940
H     2   R6                   1   1.007940
#
VARIABLES
R1      1.33753412
R2      1.35406571
R3      .98316245
R4      1.09714328
R5      1.10195029
R6      1.10140201
A1      126.727334
A2      108.484045
A3      120.171938
A4      122.350600
A5      122.089435
D1      .026953
D2      180.007587
D3      -180.000000
D4      .000000
#
#
#
REACTANT Z-MATRIX ANGSTROM
C                               6   12.011000
O     1   R1                   8   15.999400
C     1   R2                   6   12.011000
H     3   R3                   1   1.007940
H     3   R4                   1   1.007940
H     3   R5                   1   1.007940
H     1   R6                   1   1.007940
#
VARIABLES
R1      1.21898657
R2      1.48370403
R3      1.10453256
R4      1.11075723
R5      1.11077688
R6      1.13169337
A1      124.467029
A2      111.060189
A3      109.154845
A4      109.123964
A5      120.433172
D1      -.070961
D2      -122.316147
D3      122.144371
D4      179.992103

```

Constant (frozen) structure parameters must be common to both reactants and products. Thus, only one set of constant definitions is permitted. The following input is used for a saddle-point interpolation in which the positions of C1 and C2 remain unchanged.

```

PRODUCT
C1      .000000      .000000      .000000
C2      .000000      .000000     1.508857
H      1.028689      .000000    -.410316
H     -1.015045     -.024889     1.940572
H     -.529072     -.891070    -.389824
H     -.517448     .901443    -.385027
H     .544687     .857413     1.954166
H     .676484     -.872804     1.873463
SI     -.129351     -2.427077     2.452048
H     -1.252212     -2.428853     1.388199
H     -1.024260     -1.779782     3.537837
#
#
#
REACTANT
C1      .000000      .000000      .000000
C2      .000000      .000000     1.508857
H      1.035894      .000000    -.397290
SI     .882110     -1.504909     2.206066
H     -.504846     -.904544    -.397302
H     -.520431     .886537    -.415328
H     .492408     .912128     1.916629
H     -1.036466     .015343     1.916608
H     .883975     -1.510440     3.722866
H     .196804     -2.766324     1.714760
H     2.318323     -1.522790     1.716988
#
CONSTANTS
C XYZ

```

Please note that the Cartesian coordinates of these two atoms have to be the same in the PRODUCT and REACTANT definitions. In fact, these two structures were optimized with the same constraint. Examples 2.6 and 2.7 on page 7 and 8 in the tutorial, show typical inputs for saddle point interpolations.

#### 4.1.3 Keyword CONSTANTS

This keyword assigns constant Cartesian and internal structure parameters.

Options: None

Description:

The body of CONSTANTS lists all atoms and internal coordinates which are kept frozen during the geometry optimization. The atomic or element symbol with the string XYZ is used to freeze the atom or group of atoms in redundant coordinates. The string X will only freeze the x coordinate of the atom, the string XY the x and y coordinates, etc.

For internal coordinate inputs, freezing of a particular degree of freedom is done by giving the symbolic coordinate string of the Z-matrix and its assigned (frozen) value (real number). The maximum length of the value string is 15 characters. Please note that structure parameters that are explicitly defined inside the Z-matrix are always treated as CONSTANTS in optimizations. In the case of a MIXED input, these two syntax forms can be mixed (see Figure 4 and the corresponding input example), for example in the form

```
Constants
H   XYZ
A   120.0
```

Here, the first line freezes all hydrogen atoms at their positions. In the second line the internal coordinate A is frozen to 120. To avoid inconsistencies in the definition of mixed constraints, it is advisable to freeze three (linearly independent) atoms of the Cartesian part of the mixed input. These atoms may also be dummy atoms. In this way, a global reference coordinate system is defined (see example 2.5 on page 6 of the tutorial). A similar strategy should be used for mixed constraints on internal coordinates, as the following input example shows.

```
Geometry Z-Matrix
X1
Fe  X1  1.0
X2  Fe  0.5    X1  90.0
X3  Fe  0.5    X1  90.0    X2  180.0
X4  Fe  rFeCp  X1  90.0    X3  180.0
X5  Fe  rFeCp  X1  90.0    X2  180.0
C1  X4  rCX    Fe  90.0    X1  0.0
C2  X4  rCX    C1  72.0    Fe  90.0
C3  X4  rCX    C2  72.0    Fe  90.0
C4  X4  rCX    C3  72.0    Fe  90.0
C5  X4  rCX    C4  72.0    Fe  90.0
C6  X5  rCX    Fe  90.0    X1  0.0
C7  X5  rCX    C6  72.0    Fe  90.0
C8  X5  rCX    C7  72.0    Fe  90.0
C9  X5  rCX    C8  72.0    Fe  90.0
C10 X5  rCX   C9  72.0    Fe  90.0
H1  C1  rCH    X4  alpha   Fe  180.0
H2  C2  rCH    X4  alpha   Fe  180.0
H3  C3  rCH    X4  alpha   Fe  180.0
H4  C4  rCH    X4  alpha   Fe  180.0
H5  C5  rCH    X4  alpha   Fe  180.0
H6  C6  rCH    X5  alpha   Fe  180.0
H7  C7  rCH    X5  alpha   Fe  180.0
H8  C8  rCH    X5  alpha   Fe  180.0
H9  C9  rCH    X5  alpha   Fe  180.0
H10 C10 rCH   X5  alpha   Fe  180.0
#
VARIABLES
rFeCP  1.19723
rCH    1.09414
alpha  162.12553
#
```

```
CONSTANTS
rCX 1.21719
X1 XYZ
X2 XYZ
X3 XYZ
Fe XYZ
```

Furthermore it is advisable, though not mandatory, that mixed constraints be defined to match the mixed input, *i.e.* Cartesian constraints for atoms that are defined by Cartesian coordinates and internal constraints for the bonds and angles that are defined in the Z-matrix part of a mixed input. Examples 2.8 and 2.9 on page 9 and 10 of the tutorial demonstrate the use of mixed constraints in combination with equivalent and constant coordinate definitions.

#### 4.1.4 Keyword CONSTRAINTS

This keyword specifies molecular dynamics constraints.

Options: None

Description:

Molecular dynamics constraints are introduced force neutral along interatomic connection lines with the Rattle [76] and Shake [77] algorithms. They can be defined in terms of internal coordinates, *i.e.* bond lengths, bond angles and dihedral angles, in the keyword body of the CONSTRAINTS keyword. In the following example the bond length between atoms 1 and 3 (C-H2), the bond angle between atoms 1, 2 and 7 (C-H1-0) and the dihedral angle between atoms 1, 3, 7 and 8 (C-H2-O-H) are constrained during a BOMD simulation of CH<sub>4</sub> and H<sub>2</sub>O.

```
DYNAMICS STEP=0.5 MAX=200 INT=1
CONSTRAINTS
 1 3
 1 2 7
 1 3 7 8
PRINT CONSTRAINTS
BATH NOSE NHC=3 T=300
VELOCITIES RANDOM LP=0
LPCONSERVE FORCES
VISUALIZE MOLDEN MD
#
GEOMETRY CARTESIAN ANGSTROMS
C    0.000000    0.000000    0.000000
H1   0.000000    0.000000    1.092841
H2   1.033412    0.000000   -0.357505
H3   -0.517308   0.895483   -0.357014
H4   -0.518495   -0.891704   -0.361515
H    0.521017    0.588124   -2.579748
O    0.928735    0.380831   -3.434937
H    0.920449   -0.588239   -3.473172
```

As this example shows, each constraint is defined by a set of integer numbers in a line of the CONSTRAINTS keyword body. Alternatively, it is possible to constrain an entire set of atoms as shown in the following example.

```
DYNAMICS STEP=0.5 MAX=200 INT=1
CONSTRAINTS
1 - 5
PRINT CONSTRAINTS
BATH NOSE NHC=3 T=300
VELOCITIES RANDOM LP=0
LPCONSERVE FORCES
VISUALIZE MOLDEN MD
#
GEOMETRY CARTESIAN ANGSTROMS
C    0.000000    0.000000    0.000000
H1   0.000000    0.000000    1.092841
H2   1.033412    0.000000   -0.357505
H3   -0.517308   0.895483   -0.357014
H4   -0.518495   -0.891704   -0.361515
H    0.521017    0.588124   -2.579748
O    0.928735    0.380831   -3.434937
H    0.920449   -0.588239   -3.473172
```

Here the first five atoms (the CH<sub>4</sub> molecule) are constrained during the BOMD simulation.

Note that dummy atoms must be taken into account for the atom numbering in constraints as the following example shows.

```
DYNAMICS STEP=1.0 MAX=100 R=6.0 INT=1
PRINT CONSTRAINTS
BASIS
C (STO-3G)
H (DZVP)
N (STO-3G)
AUXIS
C (A2)
H (A2)
N (GEN-A2)
MULTIPLICITY 1
VXCTYPE PBE
TRAJECTORY RESTART
BATH NOSE T=500 NHC=3 FREQ=200
VELOCITIES RANDOM LP=0 T=500
LPCONSERVE ON
GUESS RESTART
VISUALIZATION MOLDEN MD
#
# Z-Matrix coordinates of MD step 100
#
GEOMETRY Z-MATRIX BOHR
C                               6    12.011000
H    1    RCH                  1    1.007940
X    1    RCX      2    A1      RAD    0    0.000000
N    1    RCN      3    A2      2    DX1      RAD    7    14.006740
#
VARIABLES
RCH      2.06518100
RCX      1.89096462
RCN      2.27049634
```

```

A1          89.538122
A2          89.876160
DX1         163.715357
#
CONSTRAINTS
  1      2
  1      4

```

In this input the C-H and C≡N bond lengths in HCN are fixed during the BOMD. The CONSTRAINTS keyword can also be used for MM MD simulations. The following example shows an MM MD input for (H<sub>2</sub>O)<sub>4</sub> in which the first water molecule is kept frozen.

```

LPCONSERVE
CONSTRAINTS
 1 - 3
VISUALIZATION MD
PRINT CONSTRAINTS
DISPERSION S6=1.0
BATH NOSE NHC=3 T=300
VELOCITIES RANDOM LP=0
FORCEFIELD FF=OPLS-AA
DYNAMICS STEP=0.5 MAX=1000 INT=1
#
GEOMETRY MIXED ANGSTROM
O1   1.187632    0.332759    -1.195914    -76 2 3
H2   0.217713    0.505507    -1.369657    -77 1
H3   1.503324    -0.129952    -1.999029    -77 1
O4   -1.628323   0.600651    -1.413703    -76
H5   -1.885969   0.153697    -0.557057    -77
H6   -2.021673   1.495719    -1.360149    -77
O7   0.867185    -1.043515    1.275905    -76 8 9
H8   1.124357    -0.596134    0.419304    -77 7
H9   1.262389    -1.937796    1.222724    -77 7
O10  O7  R1  O4  A1  O1  D1  -76 11 12
H11  O10 R2  O4  A2  O1  D2  -77 10
H12  O10 R3  H11 A3  O7  D3  -77 10
#
Variables
R1  3.0
R2  1.0
R3  1.0
A1  90.0
A2  10.0
A3  99.0
D1  0.0
D2  45.0
D3  90.0

```

The extension of the CONSTRAINTS keyword to QM/MM molecular dynamics is straightforward. It can be applied to QM and MM atoms simultaneously.

#### 4.1.5 Keyword **VARIABLES**

This keyword assigns variable internal coordinates.

Options: None

Description:

The body of VARIABLES collects all symbolic coordinate strings of the Z-matrix, along with their initial values, which are to be optimized (see 4.1.1 for examples). The maximum length for the variable value assigned under VARIABLES is 15 characters. **The maximum length for symbolic strings in the Z-matrix that denote variables or constants is 8 characters.**

#### 4.1.6 Keyword SYMMETRY

This keyword controls the symmetry analysis of the molecular structure.

Options:

**NONE / ON / OFF**

**NONE** No symmetry transformations are performed.

**OFF** The symmetry analysis is switched off.

**ON** A symmetry analysis is performed.

Description:

With the option ON of the SYMMETRY keyword a symmetry analysis is performed. In this case the point group of the molecule is detected automatically and the molecular INPUT ORIENTATION is transformed into the STANDARD ORIENTATION defined by the center of mass and the principal axes of the inertia tensor. All further calculations are done in this orientation. This is the only way to obtain symmetry adapted tensor components e.g. for electrostatic moments (see 4.8.3) and polarizabilities (see 4.8.4). For atomic calculations, the ON option of the SYMMETRY keyword enforces the zeroing of those off-diagonal Kohn-Sham elements which connect (atomic) orbitals of different spherical symmetry.

With the default option NONE of the SYMMETRY keyword no symmetry transformation is performed but the point group of the molecule is detected. The INPUT ORIENTATION is used for the SCF calculation. However, if a frequency analysis is requested (see 4.8.1), the normal modes are always printed in STANDARD ORIENTATION. Also a Bader population analysis (see 4.9.1) is always performed in STANDARD ORIENTATION. The option OFF of the SYMMETRY keyword skips the symmetry analysis module and assigns a C<sub>1</sub> point group for the input structure. This setting is advisable for systems with more than 1,000 atoms.

#### 4.1.7 Keyword ALIGNMENT

This keyword allows the control of the alignment procedure used either for the SADDLE interpolation method, or for pattern-geometry alignment (see 4.1.8). If no saddle interpolation (see 4.6.2) or similarity analysis (see 4.7.3) is requested, this keyword performs the alignment between the geometry and the pattern coordinates specified in the input.

Options:

<b>ENANTIOMER</b>	Include enantiomers for the pattern-geometry alignment.
<b>EXCLUDE</b>	Exclude the labels specified in the next line.
<b>OFF</b>	Disable the alignment procedure (enabled by default).
<b>CONNECT</b>	Try to preserve connectivity during the alignment procedure (always enabled for SADDLE interpolation).
<b>UNIFORM</b>	Use non-weighted coordinates instead of mass-weighted coordinates.
<b>CUBE=&lt;Integer&gt;</b>	Maximum grid density. The default is 6.
<b>TOL=&lt;Integer&gt;</b>	Convergence threshold. The default is 50.

Description:

The automatic alignment of REACTANT and PRODUCT is enabled by default when performing a SADDLE interpolation for a transition state search. In this case molecular connectivity is preserved and it cannot be disabled unless the automatic alignment is disabled. Disabling of the automatic alignment means that the user has to define the input coordinates of the reactant and product in exactly the same order. The pattern-geometry alignment in an MD comparison analysis is enabled by default, molecular connectivity is not preserved by default but it can be enabled with the keyword option CONNECT. Inclusion of enantiomers of patterns can be enabled with the keyword option ENANTIOMER.

In all cases you can disable the automatic alignment of specific atoms using the EXCLUDE option or completely turn off the alignment procedure with the OFF option. The TOL option specifies the convergence threshold [78], a higher value is better to find the optimum alignment but it requires a denser grid and will take more time to complete. The CUBE option specifies the maximum grid density, each level makes the grid 8 times denser. A higher value is needed only when the superposition procedure does not converge. In such cases a warning message will be printed to the output file. The following is an example input for the ALIGNMENT keyword. Note the definition of two structures.

```
Alignment ENANTIOMER
#
# Definition of reference structure
```

```

#
GEOMETRY
C      0.00000    0.00000    0.00000
H     -1.94688   -0.00000    0.89636
C     -0.69473    1.05820    2.81736
H      0.66851    0.48204   -0.72556
O     -0.86210    1.65475    1.61542
H     -1.63798    1.38692   -0.25767
C     -1.19879    0.71251    0.49953
H     -0.47414    1.84223    3.57858
H      0.31112   -0.94053    0.47087
O     -0.77370   -0.14111    3.05464
PATTERN
O      0.129600   1.190498   -0.168647
C      1.134846   0.276024    0.115270
O      0.649460   -1.029686    0.045734
H     -1.273979   -1.662177    0.548734
H     -1.031468   -1.191061   -1.188865
H     -1.283657    0.657175    1.285269
C     -0.787840   -0.942833   -0.135600
C     -1.096353    0.518732    0.200367
H     -1.916909    0.952610   -0.396583
H      2.069202   0.455344   -0.442920

```

#### 4.1.8 Keyword PATTERN

This keyword is used to specify the geometries to be used as patterns in a (MD) similarity analysis.

Options:

<Name>      Pattern name with max. 60 characters.

Description:

It is possible to concatenate up to 10 PATTERN geometries in a single calculation. There should be no blank lines, comments or any keyword between the PATTERN keyword line and the pattern geometry definition in the keyword body. It is possible to specify a pattern name in the PATTERN keyword line to identify that pattern in the output as the following example shows.

```

Alignment ENANTIMER
#
# Definition of reference structure
#
GEOMETRY
NA      0.000000    0.000000    0.000000
NA      0.000000    0.000000    3.393525
NA      3.268461    0.000000    1.479245
NA      0.888324   -3.056532    4.774686
NA     -1.629822   -4.455197    2.686041
NA     -2.583769   -1.943555    4.967887
NA     -2.989937   -0.984704    1.738553
NA      1.366162   -2.955290    1.350093

```

---

NA	-1.635500	-3.218847	-0.544082
NA	-5.055099	-3.819702	3.167361
PATTERN	min2		
NA	0.000000	0.000000	0.000000
NA	0.000000	0.000000	3.575759
NA	3.549454	0.000000	-0.439483
NA	2.938549	1.851826	2.596886
NA	2.827134	-1.667030	2.498461
NA	-0.441082	3.023305	1.834449
NA	1.766140	3.022677	-0.663340
NA	-0.477096	-3.183526	1.786813
NA	1.714285	-3.182376	-0.693793
NA	2.282350	5.408371	2.016496
PATTERN	min3		
NA	0.000000	0.000000	0.000000
NA	0.000000	0.000000	3.566787
NA	3.067758	0.000000	1.276279
NA	0.794195	-3.055461	1.660894
NA	-2.609997	-1.801127	1.799058
NA	0.193436	-3.141545	5.116727
NA	-1.732165	-5.064679	2.834014
NA	-4.954470	-3.629680	3.753795
NA	-2.916452	-4.658506	6.253583
NA	-2.747943	-1.156385	5.271665
PATTERN	min4		
NA	0.000000	0.000000	0.000000
NA	0.000000	0.000000	3.492390
NA	3.139007	0.000000	1.618174
NA	3.285473	0.000000	-1.910498
NA	0.146554	0.000080	-3.529251
NA	-3.034789	0.000000	-1.727631
NA	-3.032251	0.000000	1.762791
NA	3.285875	-0.000019	-5.402903
NA	6.317725	-0.000215	-3.673600
NA	6.320538	-0.000196	-0.183041

## 4.2 Methodology Selection

### 4.2.1 Keyword VXCTYPE

This keyword specifies the choice of exchange-correlation functional. A keyword option controls the selection of the density (Kohn-Sham or auxiliary) to be used in calculating the exchange-correlation energy and matrix elements of the exchange-correlation potential.

Options:

**VWN** / **PZ81** / **PW92** / **PW86** / **BLYP** / **OLYP** / **PW91** / **PW91SSF** / **PBE** / **PBESSF** /  
**PBESOL** / **KT1** / **KT2** / **KT3** / **CAP** / **SO11** / **N12** / **GAM** / **VS98** / **PKZB** / **TPSS** / **M06L**  
/ **M11L** / **MN12** / **B3LYP** / **BH&H** / **PBE0** / **M062X** / **M06HF** / **M06** / **XALPHA** / **FOCK** /  
**NONE**

<b>VWN</b>	Dirac exchange with local VWN correlation [79,88].
<b>PZ81</b>	Dirac exchange with local PZ81 correlation [79,89].
<b>PW92</b>	Dirac exchange with local PW92 correlation [79,90].
<b>PW86</b>	PW86 GGA exchange with P86 GGA correlation [80,91].
<b>BLYP</b>	B88 GGA exchange with LYP GGA correlation [81,93–95].
<b>OLYP</b>	HC01 GGA exchange with LYP GGA correlation [86,93–95].
<b>PW91</b>	PW91 GGA exchange and correlation [82].
<b>PW91SSF</b>	PW91 with full spin scaling function [82].
<b>PBE</b>	PBE GGA exchange and correlation [83].
<b>PBESSF</b>	PBE with full spin scaling function [83].
<b>PBESOL</b>	PBE GGA exchange and correlation for solids [87,87].
<b>KT1/2/3</b>	KT1-3 GGA exchange and correlation for NMR shieldings [100,101].
<b>SO11</b>	SO11 GGA exchange and correlation [102].
<b>N12</b>	N12 GGA exchange and correlation [103].
<b>GAM</b>	GAM GGA exchange and correlation [104].
<b>CAP</b>	CAP GGA exchange and correlation [105].
<b>VS98</b>	VS98 meta-GGA exchange and correlation [106].
<b>PKZB</b>	PKZB meta-GGA exchange and correlation [107].
<b>TPSS</b>	TPSS meta-GGA exchange and correlation [108,109].
<b>M06L</b>	M06L meta-GGA exchange and correlation [110].
<b>M11L</b>	M11L meta-GGA exchange and correlation [111].
<b>MN12</b>	MN12 meta-GGA exchange and correlation [112].
<b>B3LYP</b>	B3LYP GGA hybrid exchange and correlation [62,63].

---

<b>BH&amp;H</b>	Becke half-and-half GGA hybrid exchange and correlation [113].
<b>PBE0</b>	PBE0 GGA hybrid exchange and correlation [64,65].
<b>M062X</b>	M062X meta-GGA hybrid exchange and correlation [66].
<b>M06HF</b>	M06HF meta-GGA hybrid exchange and correlation [66].
<b>M06</b>	M06 meta-GGA hybrid exchange and correlation [66].
<b>XALPHA</b>	$X_\alpha$ calculation. The default $\alpha$ value is 0.75. A user defined $\alpha$ value can be selected with the X = <Real> option [114].
<b>FOCK</b>	Variational fitted Fock exchange [115].
<b>NONE</b>	No exchange-correlation functional is used.

### **AUXIS / BASIS**

<b>AUXIS</b>	The auxiliary function density is used for the calculation of the exchange-correlation energy and matrix elements of the potential. This is the default.
<b>BASIS</b>	The Kohn-Sham density is used for the calculation of the exchange-correlation energy and matrix elements of the potential.

#### Description:

The options listed above represent the most common combinations of exchange and correlation functionals. In addition to these combinations, the exchange and correlation functionals listed in Table 5 and 6, respectively, can be combined by the user.

The syntax for user-defined combinations of functionals is <Exchange> - <Correlation> (e.g. B88 - P86). With the option NONE, the exchange or correlation functional (or both) can be omitted. Distinct from the literature, the PW86 exchange functional is implemented in deMon2k with a cutoff and the local contribution of the P86 correlation functional is calculated with the VWN functional. The BH&H hybrid functional energy is defined in deMon2k as [113]:

$$E^{BH\&H} = \frac{1}{2}E_x^{LDA} + \frac{1}{2}\delta E_x^{B88} + \frac{1}{2}E_x^{HF} + E_c^{LYP} \quad (4.1)$$

Also the M06 correlation functional implementation differs slightly from the original proposed form [66] by modifying the self-correlation correction term inside the VS98 [106] contribution and the M06 part, as suggested by Gräfenstein et al. [116]. All other functionals are implemented according to the literature cited. Where possible, second derivatives of the density were eliminated by integration by parts [117].

The AUXIS and BASIS options specify the density that is used for the calculation of the exchange-correlation energy and matrix elements of the exchange-correlation potential.

Table 5: Exchange functionals available in deMon2k. The option acronym is given in bold.

<b>DIRAC</b>	Local Dirac exchange [79].
<b>FOCK</b>	Variational fitted Fock exchange [115].
<b>PW86</b>	Perdew and Wang GGA exchange from 1986 [80].
<b>B88</b>	Becke GGA exchange from 1988 [81].
<b>PW91</b>	Perdew and Wang GGA exchange from 1991 [82].
<b>ECMV</b>	Engel, Chevary, McDonald, Vosko GGA exchange [96,97].
<b>EV93</b>	Engel and Vosko GGA exchange [98].
<b>LB94</b>	Leeuwen and Baerends model exchange potential with correct asymptotic behavior [99].
<b>PBE96</b>	Original Perdew, Burke and Ernzerhof GGA exchange from 1996 [83].
<b>PBE98</b>	PBE GGA exchange modified by Y. Zhang and W. Yang [84].
<b>PBE99</b>	PBE GGA exchange modified by B. Hammer et al. [85].
<b>PBESOL</b>	PBE GGA exchange for solids [87].
<b>OPTX</b>	Handy and Cohen GGA exchange [86].
<b>KT1</b>	Keal and Tozer GGA exchange for NMR shielding calculations, formula 1 [100].
<b>KT2</b>	Keal and Tozer GGA exchange for NMR shielding calculations, formula 2 [100].
<b>KT3</b>	Keal and Tozer GGA exchange with improved energetics [101].
<b>VMT</b>	Vela, Medel, Trickey GGA exchange [29].
<b>VMTSOL</b>	Vela, Medel, Trickey GGA exchange for solids [29].
<b>S011</b>	Peverati et al. GGA exchange from 2011 [102].
<b>N12</b>	Peverati and Truhlar GGA exchange from 2012 [103].
<b>CAP</b>	Carmona-Espíndola, Gázquez, Vela, Trickey GGA exchange with correct asymptotic potential [105].
<b>GAM</b>	Yu et al. GGA exchange from 2015 [104].
<b>VS98</b>	Voorhis and Scuseria meta-GGA exchange [106].
<b>PKZB</b>	Perdew, Kurth, Zupan, Blaha meta-GGA exchange [107].
<b>TPSS</b>	Tao, Perdew, Staroverov, Scuseria meta-GGA exchange [108,109].
<b>M06L</b>	M06 meta-GGA exchange [66].
<b>M11L</b>	Peverati and Truhlar meta GGA exchange from 2011 [111].
<b>MN12</b>	Peverati and Truhlar meta GGA exchange from 2012 [112].
<b>PBE0</b>	PBE0 hybrid exchange [64,65].

By default, the auxiliary function density is used for both calculations, i.e. the auxiliary density functional theory approach is employed. The option BASIS invokes the Kohn-Sham methodology. Note that this choice may slow down the calculation significantly. For meta-GGAs the option BASIS **must** be specified. For Fock exchange the AUXIS and BASIS options are meaningless because it is always calculated by the variational fitting of the Fock potential. For more details and recommendations see Section 1.4, “How to Use deMon2k”. For the PW91 and PBE functionals, two different spin scaling functions are implemented

in deMon2k. By default a numerically more stable cutoff function is used. By invoking the extension SSF (options PW91SSF and PBESSF), the originally published form of the spin-scaling function is selected. This choice may change the orbital energies considerably, but for the total energies the effect is usually negligible.

Table 6: Correlation functionals available in deMon2k. The option acronym is given in bold.

<b>VWN</b>	Local Vosko, Wilk and Nusair correlation [88].
<b>PVWN</b>	Pade interpolated VWN correlation [88].
<b>RVWN</b>	RPA interpolated VWN correlation [88].
<b>PZ81</b>	Local Perdew and Zunger correlation [89].
<b>PW92</b>	Local Perdew and Wang correlation from 1992 [90].
<b>KT2</b>	Keal and Tozer local correlation for NMR shielding calculations [100].
<b>P86</b>	Perdew GGA correlation from 1986 with VWN local correlation [92].
<b>PZ86</b>	Perdew GGA correlation from 1986 [91].
<b>LYP</b>	Lee, Yang and Parr GGA correlation from 1988 [93–95].
<b>PW91</b>	Perdew and Wang GGA correlation from 1991 [82].
<b>PW91SSSF</b>	PW91 with full spin scaling function.
<b>PBE</b>	Perdew, Burke and Ernzerhof GGA correlation from 1996 [83].
<b>PBESSF</b>	PBE with full spin scaling function.
<b>PBESOL</b>	PBE GGA correlation for solids [87].
<b>KT3</b>	Keal and Tozer GGA exchange with improved energetics [101].
<b>S011</b>	Peverati et al. GGA correlation from 2011 [102].
<b>N12</b>	Peverati and Truhlar GGA correlation from 2012 [103].
<b>GAM</b>	Yu et al. GGA correlation from 2015 [104].
<b>VS98</b>	Voorhis and Scuseria meta-GGA correlation [106].
<b>PKZB</b>	Perdew, Kurth, Zupan, Blaha meta-GGA correlation [107].
<b>TPSS</b>	Tao, Perdew, Staroverov, Scuseria meta-GGA correlation [108,109].
<b>M11L</b>	Peverati and Truhlar meta GGA correlation from 2011 [111].
<b>MN12</b>	Peverati and Truhlar meta GGA correlation from 2012 [112].

#### 4.2.2 Keyword DISERSION

This keyword adds an empirical dispersion term to the electronic energy and its derivatives.

The correction is limited to the dipolar  $C_6$  term.

Options:

##### ON / OFF

- |             |  |
|-------------|--|
| <b>ON</b>   | An empirical dispersion term is added.               |
| <b>OFF</b>  | No empirical dispersion term is added.               |
| <b>READ</b> | The $C_6$ coefficients are read from the input file. |

**S6=<Real>** Global scaling factor. The default value is 1.0.

Description:

The empirical dispersion energy available in deMon2k is given by [26]:

$$E_{disp} = -S_6 \sum_A \sum_{B>A} f(R_{AB}) \frac{C_6^{AB}}{R_{AB}^6} \quad (4.2)$$

Here  $R_{AB}$  denotes the distance between atoms  $A$  and  $B$ , while  $S_6$  is a global scaling factor that adapts the empirical dispersion formula to the specified exchange functional [121]. The default value of  $S_6 = 1$  should be used in combination with the revised PBE exchange functionals PBE98 and PBE99 (see Table 5). Their use in combination with the LYP correlation functional is recommended. The function  $f(R_{AB})$  is a damping term that switches off the dispersion correction at short range. It is chosen as a Fermi type function

$$f(R_{AB}) = \frac{1}{1 + e^{-\alpha \left( \frac{R_{AB}}{R_o} - 1 \right)}}, \quad (4.3)$$

with the empirical parameter  $\alpha$  set to 23. The symbol  $R_o$  denotes the sum of the van der Waals radii [122] of the two atoms  $A$  and  $B$ . The diatomic  $C_6$  coefficients in (4.2) are calculated from the atomic  $C_6$  coefficients by the formula

$$C_6^{AB} = \frac{2 C_6^A C_6^B}{C_6^A + C_6^B}. \quad (4.4)$$

For the elements H, C, N and O, atomic  $C_6$  coefficients from Wu and Yang [123] are available within deMon2k. Distinct from the original suggestion, however, they are averaged over all possible hybridization states of the atom. For all other elements the  $C_6$  coefficients from the universal force field [124] are used in deMon2k. The user can provide  $C_6$  coefficients in the input file with the READ option of the DISPERSION keyword. User defined  $C_6$  coefficients will override the default values. Different  $C_6$  coefficients can be assigned to individual atoms by the atomic symbol (e.g. F1) or to atom groups by the element symbol (e.g. F) in the keyword body of DISPERSION:

```
DISPERSION READ
F1 10.74
```

Here the  $C_6$  dispersion coefficient must be given in atomic units. Atoms for which  $C_6$  coefficients are set to zero do not contribute to the deMon2k dispersion energy.

### 4.2.3 Keyword FORCEFIELD

This keyword is used for Molecular Mechanics (MM) or Quantum Mechanics/Molecular Mechanics (QM/MM) calculations, performed inside the deMon2k program.

Options:

**FF=<String>** Force field name as given in the FFDS file.

Description:

In MM calculations the potential energy surface of a system is calculated by an empirical force field. The necessary force field parameters are provided in the FFDS file. The FFDS file format follows TINKER [125]. This file can be changed and modified by the user in order to provide alternative force field parameters to deMon2k. In all cases, the first processed entry in this file is the force field name that must be equal to the name given by the FORCEFIELD keyword in the deMon2k input. By default, the **FFDS** file contains parameters for the OPLS-AA force field published by Jorgensen and co-workers [45,126,127]. Alternatively, a FFDS file for the AMBER force field is also delivered with deMon2k. This file gathers parameters from the different AMBERxx force fields (xx=94, 96, 98, 99, 99sb) [70,128–130].

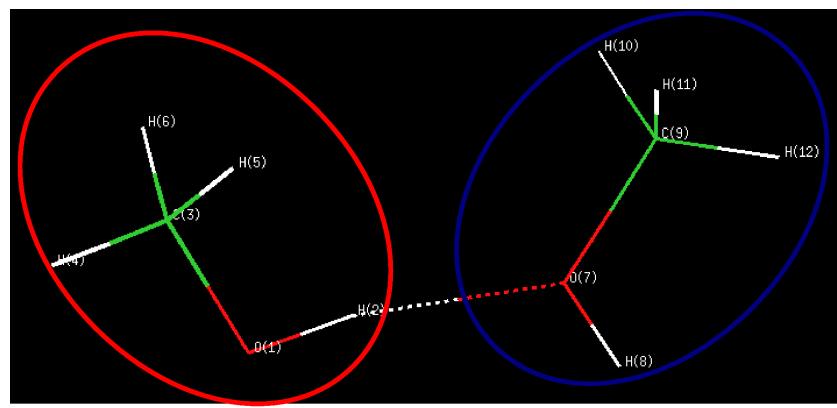
As already described in the GEOMETRY keyword (Section 4.1.1) MM atoms are specified by their atom type number which is given as a negative (-) integer number after the corresponding atomic coordinates or Z-matrix connectivities in the GEOMETRY keyword body. The following shows a single-point energy MM deMon2k input for a water molecule. The negative integer numbers, -76 and -77, after the atomic coordinates specify the atom types for the MM oxygen and hydrogen atoms, respectively. According to the default FFDS file of deMon2k this input defines an SPC water molecule.

```
TITLE SPC WATER
#
FORCEFIELD FF=OPLS-AA
#
GEOMETRY CARTESIAN ANGSTROM
O    -1.628323    0.600651   -1.413703    -76
H    -1.885969    0.153697   -0.557057    -77
H    -2.021673    1.495719   -1.360149    -77
```

Note that the atom connectivities are not specified and are, therefore, automatically generated. Alternatively, connectivities can be provided by the user as described in the GEOMETRY keyword (see 4.1.1).

The QM/MM approach combines the strengths of QM accuracy and MM speed, allowing the study of chemical processes in solution, proteins and other large systems. For these calculations within deMon2k the FORCEFIELD **and** QM/MM keywords must be specified

Figure 5: Definition of QM (red) and MM (blue) regions in a QM/MM input.



and a QM/MM geometry input must be provided. The following shows a QM/MM input for two methanol molecules (Figure 5), the first one is QM (red border in Figure 5) and the second one is MM (blue border in Figure 5). Again the MM connectivity is generated automatically.

```

TITLE QM/MM CH3OH Dimer
#
FORCEFIELD FF=OPLS-AA
QM/MM
O 96
H 97
C 99
HC 98
#
GEOMETRY CARTESIAN ANGSTROM
O   1.234808    0.310286   -1.235364
H   0.282891    0.511411   -1.454987
C   1.839149   -0.324975   -2.392312
HC  2.892167   -0.547951   -2.127264
HC  1.326470   -1.277517   -2.664724
HC  1.833392    0.354508   -3.275284
O7  -1.571368    0.671857   -1.491851   -96
H8  -1.869011    0.181864   -0.675330   -97
C9  -2.487673    0.359200   -2.572448   -99
H10 -2.204433    0.996007   -3.434393   -98
H11 -2.423630   -0.711536   -2.879602   -98
H12 -3.540363    0.591861   -2.291386   -98

```

The QM/MM keyword body is mandatory for a QM/MM calculation in deMon2k and will be explained in detail in Section 4.2.4.

#### 4.2.4 Keyword QM/MM

This keyword defines the mechanical interaction between the QM and the MM region in a QM/MM calculation. To this end atom types of the force field are assigned to the QM atoms.

Options:

- CHARMM**      The CHARMM-deMon2k interface is activated.  
**NOPOLES**     Disables the asymptotic expansion of QM/MM embedding integrals.

Description:

The mechanical interaction between the QM and MM regions is expressed by a Lennard-Jones potential, Eq. (4.13) in Section 1.4. Therefore, Lennard-Jones parameters must be assigned to the QM atoms. They are taken from the force field. To do so, atom types are assigned to the QM atoms with the QM/MM keyword. These atom types serve only for the assignment of van der Waals radii and potential depths. They are defined in the keyword body of the QM/MM keyword as shown in the following input example. Note that they are assigned by positive integer numbers.

```
TITLE QM WATER AND MM WATER
#
FORCEFIELD FF=OPLS-AA
QM/MM
O    76
H    77
#
GEOMETRY CARTESIAN ANGSTROM
O    1.187632    0.332759    -1.195914
H    0.217713    0.505507    -1.369657
H    1.503324    -0.129952   -1.999029
O    -1.628323   0.600651    -1.413703    -76
H    -1.885969   0.153697    -0.557057    -77
H    -2.021673   1.495719    -1.360149    -77
```

The QM atoms can be addressed in the QM/MM keyword body by a specific atomic symbol or by atom groups using the element symbol.

With the CHARMM option of the QM/MM keyword the deMon2k CHARMM interface is activated. In such calculations CHARMM drives the QM/MM calculations and deMon2k calculates only energies and gradients for a given, electrostatically embedded, QM system. The following deMon2k input shows again a QM/MM water dimer, now for a CHARMM-deMon2k calculation.

```
TITLE CHARMM QM WATER AND MM WATER
#
VXCTYPE PBE
QM/MM CHARMM
MATINV ANALYTICAL
#
EMBED READ
      0.88590000    0.36292000    0.37595000    -0.83400000
      1.25000000    0.75598000    1.17374000    0.41700000
      0.28602000    1.03645000    0.04525000    0.41700000
GEOMETRY
O          2.3563900000    -1.4195200000    -1.1274700000
```

H	1.8724500000	-0.7537800000	-0.6025600000
H	1.7965500000	-2.1894500000	-1.0546000000

Here the MM water is given in the keyword body of the EMBED keyword by coordinates and point charges. The QM water geometry is given as usual in the keyword body of the GEOMETRY keyword.

Processing this input with deMon2k will create, besides the usual output files, a QM/MM interface file, deMon.qmm, that has the following format:

```

TOTAL ENERGY          =      -76.3904859540375

QMFORCES
-0.115663E-01    -0.150402E-01    0.179275E-02
 0.716778E-02    -0.110224E-01    -0.794221E-02
 0.111375E-01    0.189759E-01    -0.250127E-03

EMBEDFORCES
-0.904640E-02    0.126830E-01    0.978137E-02
 0.650241E-04    -0.258697E-02    -0.295494E-02
 0.203853E-02    -0.304493E-02    -0.311790E-03

FINAL INPUT ORIENTATION
NO. ATOM      X           Y           Z           Z-ATOM   MASS
 1  O         2.356390   -1.419520   -1.127470     8       15.999
 2  H         1.872450   -0.753780   -0.602560     1       1.008
 3  H         1.796550   -2.189450   -1.054600     1       1.008

```

This file contains the energy of the QM system,  $E^{QM}$  in Eq. (4.9) of Section 1.4, and the corresponding forces on the QM (QMFORCES) and MM (EMBEDFORCES) atoms. At the end of the file the final input orientation of the QM system is also given.

The option NOPOLES of the QM/MM keyword disables the asymptotic expansion of the electrostatic embedding integrals, Eq. (4.10), in Section 1.4. For large MM regions this can slow down the calculation and, therefore, should only be used for testing and validation.

#### 4.2.5 Keyword EFIELD

With this keyword an external homogeneous electric field can be specified.

Options:

<Real>        Three real numbers that give the x, y and z components of the external electric field in atomic units.

Description:

The external electric field will be added to the core Hamilton operator by multiplying with the corresponding dipole moment matrices. The SCF is then converged with this modified core Hamilton operator. The electric field components must be given in atomic units by real numbers.

---

```
EFIELD 0.0 0.0 -0.01
```

#### 4.2.6 Keyword EMBED

This keyword specifies the embedding of the molecule in a point-charge array.

Options:

##### **FILE / READ**

- |             |  |
|-------------|--|
| <b>FILE</b> | The embedding point coordinates and charges are read from the deMon.cub file. This is the default. |
| <b>READ</b> | The embedding point coordinates and charges are read from the input file.                          |

Description:

The coordinates and charges are given in free format as real numbers either in the embedding file deMon.cub (see Table 2) or in the keyword body of EMBED. The Cartesian coordinates of the embedding points refer to the input orientation of the molecule as defined in the input file. If a Z-matrix input is used, the embedding point coordinates are defined with respect to the Z-matrix input orientation (see Figure 2). The point charge coordinate units, Ångström or Bohr, correspond to the unit system selected for the input coordinates (see keyword GEOMETRY in 4.1.1). Optionally, an element symbol or a radius (real number) may be specified after the point charge coordinates and the charge values. This information will be used for drawing the bonds in a VU output of the embedding region [41]. The following input specifies embedding in a (distorted) octahedron of positive point charges:

```
EMBED READ
 5.0  0.0  0.0  1.0  H
 -5.0  0.0  0.0  1.0  H
 0.0  5.0  0.0  1.0  H
 0.0 -5.0  0.0  1.0  H
 0.0  0.0  2.0  1.0  1.5
 0.0  0.0 -2.0  1.0  1.5
```

The first line is necessary only if the embedding points are defined within the input file. In this example, the hydrogen atom radius is assigned to each of the first four embedding points. For the other two, the radius is explicitly defined to be 1.5 Ångström or Bohr depending on the unit system specification in the GEOMETRY keyword (see Section 4.1.1).

## 4.3 Atom Specific Input

### 4.3.1 Keyword BASIS

This keyword specifies the basis set.

Options:

<Basis>      The basis set string <Basis> defines the global basis set. If absent, the DZVP basis set is used by default.

Description:

The global basis set (DZVP) is used for all atoms not defined explicitly in the BASIS keyword body. In the BASIS keyword body, basis sets can be assigned to individual atoms by a specific atomic symbol (e.g. H4) or to atom groups by the element symbol (e.g. H). The assignment of the basis set by atomic symbols, element symbols, and global basis set definition has the hierarchy (highest to lowest):

<atomic symbol>

<element symbol>

<global basis set>

Thus any basis set definition for an atom can be overridden by the explicit assignment of the basis set using the atomic symbol. For example, for C<sub>2</sub>H<sub>4</sub> (see input in 4.1.1), the following basis set definition

```
Basis (DZVP)
C    (TZVP)
C1   (STO-3G)
```

assigns an STO-3G basis to atom C1, a TZVP basis to the other carbon and the DZVP basis to all other atoms. Instead of using basis set strings, the basis set of an atom may also be specified by the Huzinaga notation given in the BASIS file. In that notation, the foregoing basis set definition would read

```
Basis (DZVP)
C    (7111/4111/1*)
C1   (33/3)
```

The file BASIS contains the basis sets listed in Table 7. Other basis sets can be obtained from the Extensible Computational Chemistry Environment Basis Set Database [131] at <https://bse.pnl.gov/bse/portal> by choosing the deMon2k basis set format.

Table 7: Basis sets available in the deMon2k BASIS file.

Basis Set	Elements	Description
DZV	H, C	LDA double $\zeta$ basis set.
DZV-GGA	H, C	GGA double $\zeta$ basis set.
DZVP	H-Xe	LDA double $\zeta$ polarization <sup>a</sup> basis set [132].
DZVP-GGA	H-Xe	GGA double $\zeta$ polarization <sup>a</sup> basis set [133].
DZVP2	Be-F, Al-Ar, Sc-Zn	Modified DZVP basis set.
TZVP	H, Li, C-F, Si-Cl	LDA triple $\zeta$ polarization basis set.
TZVP-GGA	Sc-Cu	GGA triple $\zeta$ polarization basis set.
TZVP-FIP1	H, C-F, Na, P, S, Cl, Cu	TZVP with field-induced polarization [134,135].
TZVP-FIP2	H, C-F, Na, P, S, Cl, Cu	for $\alpha$ , $\beta$ (FIP1) and $\gamma$ (FIP2) calculations.
EPR-III	H-F	EPR basis set [136].
IGLO-II	H, Li, B-F, Si	NMR basis set [137].
IGLO-III	H, B-F, Si, Cr, Fe	NMR basis set [137].
aug-PCJ-X	H, B-F, Si, P	J-Coupling basis set [138].
STO-3G	H-Ar	HF single $\zeta$ basis set [139].
6-31G**	H-Ar	HF double $\zeta$ polarization basis set [140].
6-311G**	H-Ar	MP2 triple $\zeta$ polarization basis set [141].
SAD	H, C-F	Sadlej FIP basis set [142].
LIC	H-Ne	Lie-Clementi basis set [143].
WACHTERS	Sc-Cu	Wachters basis set without $f$ functions [144].
DZ-ANO	H-Zn	Double $\zeta$ ANO basis set from Roos [145].
def2-TZVPP	H-Ar	Enlarged triple $\zeta$ polarization basis set [146]
aug-cc-pVXZ	H, Li-F, Na, Al, Cl, Cr, Zn, Mo, Ru, Au	Augmented correlation consistent basis [147].
ECP SD	See Figure 7, top	Valence basis for SD ECPs [148].
RECP SD	See Figure 7, middle	Valence basis for SD RECPs [148].
QECP SD	See Figure 7, bottom	Valence basis for SD QECPs [148].
ECP LANL2DZ	See Figure 8, top	Valence basis for LANL ECPs [131].
QECP LANL2DZ	See Figure 8, bottom	Valence basis for LANL QECPs [131].
ECP HW	K-Cu	Hay-Wadt basis for LANL ECPs [149–151].
QECP HW	Rb-Ag, Cs-La, Hf-Au	Hay-Wadt basis for LANL QECPs [149–151].
MCP LK	See Figure 9, top	Valence basis for LK MCPs [152,153].
RMCP LK	See Figure 9, bottom	Valence basis for LK RMCPs [152,153].
XAS-I	Li-F	XAS augmentation basis for first row [154].
XAS-II	Na-Cl	XAS augmentation basis for second row [154].

<sup>a</sup>Modified for Li and Na [155].

Instead of reading the basis set from the BASIS file, the user can define the basis directly in the input file according to the format:

```

Basis
SYMBOL Read
N   L   K
  EXPONENT COEFFICIENT
  :
  EXPONENT COEFFICIENT

```

Here SYMBOL can be an element (e.g. OXYGEN) or atomic symbol (e.g. O), N and L are

the principal and angular momentum quantum numbers of this shell and K is the degree of contraction. A shell collects all contracted orbitals of the same angular momentum quantum number, such as  $p_x$ ,  $p_y$ ,  $p_z$  or  $d_{xx}$ ,  $d_{xy}$ ,  $d_{xz}$ ,  $d_{yy}$ ,  $d_{yz}$  and  $d_{zz}$ . The contracted orbitals,  $\mu(\mathbf{r})$ , are linear combinations of (atom-centered) Gaussian type orbitals (LCGTO), which are called the primitive orbitals,  $g(\mathbf{r})$ :

$$\mu(\mathbf{r}) = \sum_{k=1}^K d_{\mu k} g_k(\mathbf{r}) \quad (4.5)$$

$$g_k(\mathbf{r}) = (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} e^{-\alpha_k(\mathbf{r}-\mathbf{A})^2} \quad (4.6)$$

The exponents  $\alpha_k$  and contraction coefficients  $d_{\mu k}$  are listed in free format under the shell definition line, one line for each primitive orbital (EXONENT and COEFFICIENT). Here is an example for a user-defined basis set input for the oxygen molecule:

```
Multiplicity 3
SCFTYPE ROKS
VXCType Basis BLYP
Mixing 0.4
PRINT GTO CGTO AUXIS
Geometry Z-Matrix
0
0 0 r
#
Variables
r 1.207
BASIS
0 Read
1 0 2
49.98097100 0.4301280000
8.896588000 0.6789140000
2 0 2
1.945237000 0.4947200000E-01
0.493363000 0.9637820000
2 1 2
1.945237000 0.5115410000
0.493363000 0.6128200000
AUXIS (GEN-A2)
```

In this example the basis set was taken from <https://bse.pnl.gov/bse/portal> using the deMon2k format (see Figure 6). The first line in this format (here 3) must be deleted in the user-defined basis set input. This line defines the number of contractions and is only needed for the basis set definition in the file BASIS. The basis set in this example contains a  $1s$ ,  $2s$  and  $2p$  shell. Each shell has a contraction degree of two. The  $2s$  and  $2p$  shells share a common set of exponents. However, in deMon2k both those shells have to be listed independently, as shown here.

The specification of ECP and MCP valence basis sets will also trigger the use of the corresponding effective or model core potential if the ECP and MCP keywords are not specified.

Figure 6: STO-2G basis set for oxygen, taken from <https://bse.pnl.gov/bse/portal>

```

#  STO-2G  EMSL  Basis Set Exchange Library 7/23/15 10:51 AM
# Elements          References
# -----
# H - He: W.J. Hehre, R.F. Stewart and J.A. Pople, J. Chem. Phys. 2657 (1969).
# Li - Ne:
# Na - Ar: W.J. Hehre, R. Ditchfield, R.F. Stewart, J.A. Pople, J. Chem. Phys.
# K - Kr: 52, 2769 (1970).
# 

# This basis set uses cartesian components
O-OXYGEN O (STO-2G)
#
# Primary Functions: (4s,2p) -> [2s,1p]
# Augmenting Functions: (4s,2p) ->[2s,1p]
3
1    0    2
  49.9809710      0.4301280
  8.8965880       0.6789140
2    0    2
  1.9452370      0.0494720
  0.4933630       0.9637820
2    1    2
  1.9452370      0.5115410
  0.4933630       0.6128200

```

Thus, their specification usually is **not necessary**. An exception is the Hay-Wadt valence basis sets (ECP|HW and QECP|HW), which represent alternative choices to the corresponding LANL (Los Alamos National Laboratory) double  $\zeta$  valence basis sets. If the Hay-Wadt basis sets are used, the ECP must be specified explicitly with the ECPS keyword. This is a special case of the more general capability of defining basis sets and ECPs or MCPs independently by the BASIS and ECPS or MCPS keywords. **Be aware that you can mix any ECPs/MCPs with any basis set, including all-electron ones!** Thus, care must be taken if the keywords BASIS and ECPS/MCPS are used together in the deMon2k input. For less experienced users, we recommend using the PRINT keyword with the BASIS options ECPS and MCPS (see example 2.11 on page 15 of the tutorial) in order to obtain full information about the basis sets and ECPs/MCPs actually utilized. If the ECP or MCP valence basis sets are obtained from external resources like <https://bse.pnl.gov/bse/portal>, the principal quantum number indexing may be wrong. As a result the tight-binding start

density cannot be generated correctly. As a quick fix, we suggest switching to a CORE start density by GUESS CORE (see 4.5.5).

### 4.3.2 Keyword AUGMENT

This keyword specifies the basis set augmentation in XAS calculations.

Options:

<Augment>      The basis set augmentation string <Augment> defines the global augmentation basis set. By default, no augmentation is performed.

As for the orbital basis set options, the XAS (see 4.8.7) augmentation basis set can be assigned to individual atoms by the atomic symbol or to atom groups by the element symbol in the keyword block of AUGMENT. The syntax and hierarchy of these assignments is analogous to the orbital basis set assignments (see 4.3.1). Thus, the following input sequence

```
AUGMENT (XAS-I)
S          (XAS-II)
```

requests the XAS-I augmentation basis set for all atoms, except sulfur, for which the XAS-II augmentation basis set is used. These are the two augmentation basis sets available in the deMon2k basis set file. XAS-I is typically used for first-row atoms whereas XAS-II is somewhat more extended and, therefore, appropriate for heavier atoms. Note that for reliable XAS spectra the augmentation basis is needed, at least on the core-excited atom. Otherwise the sampling of the continuum and Rydberg states will be too coarse. Augmentation basis sets can also be defined directly in the input file using the format:

```
Augment
SYMBOL Read
N   L   K
    EXPONENT COEFFICIENT
    :
    EXPONENT COEFFICIENT
```

The syntax is the same as for the reading of basis sets.

### 4.3.3 Keyword AUXIS

This keyword specifies the auxiliary function set.

Options:

<Auxis> The auxiliary function set string <Auxis> defines the global auxiliary function set. If this option is absent, the GEN-A2 auxiliary function set is used by default.

Description:

Similar to the orbital basis set options, the auxiliary function sets can also be assigned to individual atoms by the atomic symbol or to atom groups by the element symbol in the keyword block of AUXIS. The syntax and hierarchy of these assignments is analogous to the orbital basis set assignments (see 4.3.1). For C<sub>2</sub>H<sub>4</sub> (see input in 4.1.1) the following auxiliary function definition

```
AUXIS (A2)
C      (GEN-A2*)
C1     (GEN-A2)
```

assigns a GEN-A2 auxiliary function set to atom C1, a GEN-A2\* auxiliary function set to the other carbon and the A2 auxiliary function set to all other atoms. The AUXIS file of deMon contains only the A2 auxiliary function set. The other auxiliary function sets used above (GEN-A2 and GEN-A2\*) are generated automatically [133,156] according to the procedure described in Appendix A. The GEN-An+ ( $n = 2, 3, 4$ ) auxiliary function sets are particular for the variational fitting of Fock exchange. They have a set structure as the corresponding GEN-An\* auxiliary function sets but without the angular momentum index augmentation to *f* and *g* Hermite Gaussians during the SCF iterations. Instead, this augmentation is only performed post-SCF for the final Fock energy calculation. Thus, the computational demand for the GEN-An+ auxiliary function set during the SCF is close to that of GEN-An, whereas the final (non-variational) energy is close to GEN-An\*. Just as with the orbital basis sets, auxiliary function sets can also be specified directly in the input file using the format:

```
Auxis
SYMBOL Read
LMAX EXPONENT
:
:
LMAX EXPONENT
```

Here LMAX denotes the maximum angular momentum quantum number of the auxiliary function subset and EXPONENT the exponent which is shared [53,54] by all functions in that subset. Thus an auxiliary function subset with LMAX = 2 contains ten functions, namely one *s*, three *p* and six (Cartesian) *d* functions. In deMon2k, these functions are primitive Hermite Gaussians [51,157] of the form (without normalization):

$$\bar{a}(\mathbf{r}) = \left( \frac{\partial}{\partial A_x} \right)^{\bar{a}_x} \left( \frac{\partial}{\partial A_y} \right)^{\bar{a}_y} \left( \frac{\partial}{\partial A_z} \right)^{\bar{a}_z} e^{-\beta(\mathbf{r}-\mathbf{A})^2} \quad (4.7)$$

In the following we extend the previous example of a user-defined basis set input for an oxygen molecule with a user-defined auxiliary function set input:

```
Multiplicity 3
SCFTYPE ROKS
VxcType Basis BLYP
Mixing 0.4
PRINT GTO CGTO AUXIS
Geometry Z-Matrix
0
0 0 r
#
Variables
r 1.207
BASIS
0 Read
1 0 2
49.98097100      0.4301280000
8.896588000      0.6789140000
2 0 2
1.945237000     0.4947200000E-01
0.493363000     0.9637820000
2 1 2
1.945237000     0.5115410000
0.493363000     0.6128200000
AUXIS
0 Read
0 100.0
0 25.0
2 5.0
2 1.0
```

The auxiliary function set in this example consists of two *s* and two *d* sets. Together these comprise 22 auxiliary functions (1+3+6 from each *d* set and 1 from each *s* set).

#### 4.3.4 Keyword ECPS

This keyword specifies the effective core potentials (ECPS).

Options:

<ECP>      The ECP string <ECP> defines the global effective core potentials. If absent, an all-electron calculation is assumed by default.

Usage of the keyword ECPS is very similar to that for the keyword BASIS (Section 4.3.1). Different ECPs can be assigned to individual atoms by the atomic symbol (e.g. Au1) or to atom groups by the element symbol (e.g. Au). The global ECP is used for all atoms which are not specified explicitly. **If the keyword ECPS is not specified for an atom but the substring "ECP" is present in the basis set definition of the keyword BASIS, then an ECP of the same name will be assigned automatically to the atom.** The assignment of

the ECP by atomic symbols, element symbols, and global ECPS definition possesses the hierarchy (highest to lowest):

```
<atomic symbol>
<element symbol>
<global ECPS>
<basis set name>
```

Thus, any ECPS definition for an atom can be overridden by the explicit assignment of the ECP using the atomic symbol. For example the following ECPS definition

```
ECPS (ECP|SD)
Au   (ECP19|SD)
Au1  (ECP1|SD)
```

assigns the one-electron ECP denoted by (ECP|SD) to the atom Au1, the 19-electron ECP denoted by (ECP19|SD) to all other gold atoms, and the globally defined (ECP|SD) ECP to any other atom in the input. The file ECPS contains the ECPs from Stuttgart-Dresden [148] and from Los Alamos National Laboratory [131] in the deMon2k format [158]. Figures 7 and 8 give an overview of the effective core potentials available in the ECPS file of deMon2k [159]. All default ECPs that are obtained by the specification of the ECP acronym alone, without an explicit valence electron number, e.g. (ECP|SD), (QECP|SD), (RECP|SD), (ECP|LANL2DZ) and (QECP|LANL2DZ), refer to the ground state configuration of the corresponding element. If an ECP for an excited state is selected, e.g. for lanthanides or actinides, the tight-binding guess generation may fail. In such cases please switch to a core start density by using GUESS CORE (see 4.5.5 for further details).

**Specifying only the keyword ECPS and not specifying a corresponding valence basis set leads to a situation in which the default all-electron DZVP basis set is used in combination with the defined effective core potential. Therefore, it is handy to define the effective core potential and the corresponding basis set at the same time with the keyword BASIS.** Then the keyword ECPS should be used only to define an effective core potential different from the one automatically defined by the specified basis set. In the following example,

```
BASIS (ECP|SD)
Au   (ECP19|SD)
Au1  (ECP1|SD)
ECPS
Cu   (RECP|SD)
```

Figure 7: Stuttgart-Dresden ECPs (top), RECPs (middle) and QECPs (bottom) that are available in the ECPS file of deMon2k.

H	(ECP SD)																		He
Li	Be																		
Na	Mg																		
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr		
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe		
Cs	Ba	La*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn		
Fr	Ra	Ac*																	

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu				
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr				

H	(RECP SD)																		He
Li	Be																		
Na	Mg																		
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr		
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe		
Cs	Ba	La*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn		
Fr	Ra	Ac*																	

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu				
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr				

H	(QECP SD)																		He
Li	Be																		
Na	Mg																		
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr		
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe		
Cs	Ba	La*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn		
Fr	Ra	Ac*																	

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu				
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr				

Figure 8: Los Alamos National Laboratory ECPs (top) and QECPs (bottom) that are available in the ECPS file of deMon2k.

H	(ECP LANL2DZ)														He
Li	Be														
Na	Mg														
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	C	N	O
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te
Cs	Ba	La*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po
Fr	Ra	Ac*													At Rn

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	He
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	

H	(QECP LANL2DZ)														He
Li	Be														
Na	Mg														
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se Br Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te I Xe
Cs	Ba	La*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po At Rn
Fr	Ra	Ac*													

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	He
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	

the relativistic (RECP|SD) is used for copper atoms in combination with the corresponding non-relativistic valence basis set. For more elaborate inputs, it is important to note that explicit effective core potential definitions from the ECPS keyword will override implicit effective core potentials specified from the BASIS keyword. Take this input sequence as an example.

```
BASIS (ECP|SD)
Au (ECP19|SD)
Au1 (ECP1|SD)
ECPS (RECP|SD)
```

Here the effective core potential assigned by the keyword ECPS will override **all** definitions of the effective core potential by the BASIS keyword. Also, the effective core potentials can be specified directly in the input file using the format

```
ECPS
SYMBOL Read ELECTRONS
```

---

N	L	K
	EXONENT	COEFFICIENT
:	:	
	EXONENT	COEFFICIENT

where SYMBOL can be an element or atomic symbol, ELECTRONS is an integer number specifying the number of valence electrons, N denotes the radial power of the operator, L the angular momentum of the effective potential, and K the contraction degree. The exponents and contraction coefficients are listed in free format under the ECP block definition line, one line for each Gaussian (EXONENT and COEFFICIENT). The following example shows a user-defined ECP input for gold in the  $\text{Au}(\text{H}_2\text{O})_2^+$  cation:

```

CHARGE 1
BASIS (DZVP)
Au (AUG-CC-PVTZ-PP)
ECPS
Au Read 19
 0 0 2
 13.2051000000    426.7098400000
   6.6025500000    35.9388240000
  0 1 2
 10.4520200000    261.1610230000
   5.2260100000    26.6262840000
  0 2 2
   7.8511000000    124.7568310000
   3.9255500000    15.7722600000
  0 3 2
   4.7898000000    30.5684750000
   2.3949100000    5.1837740000
  0 4 1
   1.0000000000    0.0000000000
AUXIS (GEN-A2*)
PRINT BASIS ECPS
#
GEOMETRY CARTESIAN ANGSTROM
AU      0.000000  -0.004153    0.000033    79    196.966540
O       2.088157   0.052069   -0.041878     8    15.999400
O      -2.088163   0.052040    0.041551     8    15.999400
H       2.507511  -0.813975   -0.228354     1    1.007940
H      -2.507532  -0.814141    0.227289     1    1.007940
H       2.490319   0.393274    0.784571     1    1.007940
H      -2.490152   0.393861   -0.784705     1    1.007940

```

Calculation of the basis set superposition error (BSSE) using the counterpoise correction [160] when ECP atoms are involved requires special consideration because of the ECP projectors. If the ECP-carrying atom represents a ghost atom in the BSSE calculation, then the effective core potential must be disabled but the valence basis set must still be defined and used. The following input sequence describes this situation for the gold atom in  $\text{AuF}$ .

```

ECPS
Au (NONE)

```

```
BASIS (DZVP)
Au (RECP|SD)
#
GUESS CORE
#
GEOMETRY ZMATRIX
Au          0
F   Au 1.936660
```

Thus, the result of this calculation is the BSSE corrected energy for the F atom in AuF. Please note the use of the CORE start density for this atomic calculation. On the other hand, if the BSSE for the ECP-carrying atom is to be calculated, no modification of the effective core potential should be made. The following input can be used to calculate the BSSE corrected energy of the Au atom in AuF:

```
BASIS (DZVP)
Au (RECP|SD)
#
GUESS CORE
#
GEOMETRY ZMATRIX
Au
F   Au 1.936660 0
```

Example 2.12 on page 15 of the deMon2k Tutorial shows an input for BSSE calculations involving more than one ECP center. For accurate energy derivatives in ECP calculations, extended grids are often necessary. Therefore, we recommend the use of GRID FINE or equivalent GRID options (see 4.3.6) if ECPs are specified. This is particularly important for heavier elements.

### 4.3.5 Keyword MCPS

This keyword specifies the use of model core potentials (MCPS). Its syntax is very similar to that for the ECPS keyword.

Options:

<MCP> MCP string <MCP> defines the global model core potentials. If absent, an all-electron calculation is assumed by default.

Different MCPs can be assigned to individual atoms by the atomic symbol (e.g. Au1) or to atom groups by the element symbol (e.g. Au). The global MCP is used for all atoms for which MCPs are not specified explicitly. Simultaneous use of ECPs and MCPs for the same atomic center is forbidden. **If MCPS is not specified for an atom but the substring "MCP" is present in the basis set definition of the keyword BASIS, then an MCP of**

**the same name will be assigned automatically to that atom.** The assignment of the MCP by atomic symbols, element symbols, and global MCPS definition follows the usual deMon2k hierarchy (highest to lowest):

```
<atomic symbol>
<element symbol>
<global MCPS>
<basis set name>
```

Thus, as in the case of ECPS, any MCPS definition for an atom can be overridden by the explicit assignment of the MCP using the atomic symbol. For example, in a system containing two copper atoms, e.g. linked to an organic ligand, the following MCPS definition,

```
MCPS (MCP|LK)
Cu (MCP17|LK)
Cu1 (RMCP17|LK)
```

assigns the 17-electron quasi-relativistic MCP from Lovallo and Klobukowski [152,153] denoted by (RMCP17|LK) to copper atom Cu1, the 17-electron non-relativistic MCP by the same authors, denoted by (MCP17|LK), to other copper atoms, and the globally defined (MCP|LK) non-relativistic MCP to any atom of other types that is present in the ligand. The file MCPS contains the MCPs from Lovallo and Klobukowski [152,153] in the deMon2k format. Figure 9 shows the elements for which MCPS are available in the MCPS file.

**Specifying only the keyword MCPS and not specifying the corresponding valence basis set leads to a situation in which the default all-electron DZVP basis set is used in combination with the specified model core potentials. Therefore it is handy to define the model core potential and the corresponding valence basis set at the same time with the keyword BASIS.** The keyword MCPS should then be used only to define a model core potential different from the one automatically invoked by the specified basis set. In the following example

```
BASIS (MCP|LK)
MCPS
Cu (MCP17|LK)
Cu1 (RMCP17|LK)
```

the explicit specification of the model core potential by the MCPS keyword overrides the global model core potential invoked with the keyword BASIS. The model core potentials can also be specified in the input file using the format

```
MCPS
SYMBOL Read ELECTRONS
```

Figure 9: Lovallo-Klobukowski MCPs (top) and RMCPs (bottom) that are available in the MCPS file of deMon2k.

H	(MCP LK)												He				
Li	Be																
Na	Mg																
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	B	C	N	O	F	Ne
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	Al	Si	P	S	Cl	Ar
Cs	Ba	La*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Ga	Ge	As	Se	Br	Kr
Fr	Ra	Ac*										In	Sn	Sb	Te	I	Xe
												Tl	Pb	Bi	Po	At	Rn

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

H	(RMCP LK)												He				
Li	Be																
Na	Mg																
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	B	C	N	O	F	Ne
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	Al	Si	P	S	Cl	Ar
Cs	Ba	La*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Ga	Ge	As	Se	Br	Kr
Fr	Ra	Ac*										In	Sn	Sb	Te	I	Xe
												Tl	Pb	Bi	Po	At	Rn

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

```

LS NLS
N K
EXONENT COEFFICIENT
:
EXONENT COEFFICIENT
N K F
EXONENT COEFFICIENT
:
EXONENT COEFFICIENT

```

Here SYMBOL can be an element or atomic symbol, ELECTRONS is an integer number specifying the number of valence electrons, LS is the number of local shells in the model potential, and NLS is the number of non-local (core) shells. All local shells are loaded first. For each local shell, N represents the radial power and K the degree of contraction. For non-local shells, N represents the principal quantum number, K is the degree of contraction, and F is the energy weight-factor of the non-local projector of this shell. The exponents and contraction coefficients are listed in free format under the MCP block definition line, one line

for each Gaussian (EXPOENT and COEFFICIENT).

#### 4.3.6 Keyword GRID

This keyword specifies the grid for the numerical integration of the exchange-correlation energy and potential.

Options:

##### ADAPTIVE / FIXED

<b>ADAPTIVE</b>	An adaptive grid is used for the integration. This is the default.
<b>FIXED</b>	A fixed grid is used for the integration.

##### MEDIUM / COARSE / FINE / REFERENCE

<b>MEDIUM</b>	Medium grid accuracy is requested. This is the default.
<b>COARSE</b>	Coarse grid accuracy is requested.
<b>FINE</b>	Fine grid accuracy is requested.
<b>REFERENCE</b>	A (fixed) reference grid is requested.
<b>TOL=&lt;Real&gt;</b>	Adaptive grid tolerance.
<b>RAD=&lt;Integer&gt;</b>	Number of radial shells for adaptive grid.

##### SCF / GUESS

<b>SCF</b>	The converged SCF density is used for the adaptive grid generation. This is the default.
<b>GUESS</b>	The start density is used for the adaptive grid generation.
<b>DIRECT</b>	The grid coordinates and weights are rebuilt in each SCF cycle. This is the default for parallel runs.

Description:

By default, deMon2k uses an adaptive grid [21] with a tolerance of  $10^{-5}$  (MEDIUM) for the numerical integration of the exchange-correlation energy and matrix elements of the associated potential. This tolerance indicates the accuracy of the numerical integration of the diagonal elements of the exchange-correlation potential matrix. With this setting, the converged SCF energy accuracy is usually better than  $100 \mu\text{Hartree}$  [118]. The adaptive grid COARSE and FINE options refer to grid tolerances of  $10^{-4}$  and  $10^{-6}$ , respectively. Thus, the stability of the numerical integration can be easily checked by choosing different grid tolerances. The COARSE adaptive grid should not be used for final energy or property cal-

culations. With the TOL option, a user-defined grid tolerance can be specified. Of course, this option is only applicable for an adaptive grid. The same holds for the RAD option that can be used to define the number of radial shells for an adaptive grid. Its upper limit is 500 specified by the MAXRAD option in the `parameter.h` file. The SCF and GUESS options specify the density used for the exchange-correlation potential which generates the adaptive grid. By default, the converged SCF density is used. This requires an extra SCF for the initial energy calculation. If the option GUESS is specified, the adaptive grid is built using the start density for the calculation of the grid generating exchange-correlation potential. Fixed grids are also available in deMon2k. They can be selected by the option FIXED in combination with the options COARSE, MEDIUM, and FINE. MEDIUM is the default for fixed grids. With the options COARSE, MEDIUM, and FINE, the pruned grids (50,194)p, (75,302)p and (99,590)p are selected. In this notation the first number refers to the radial shells and the second to the number of Lebedev grid points [119] on each of these shells. The "p" stands for pruned [120], which indicates that, for some radial shells, smaller Lebedev grids have been used. With the REFERENCE option, the fixed unpruned (200,1202) reference grid is specified. Due to its size, its application should be restricted to small systems.

The DIRECT option is used to avoid the I/O of the grid point coordinates and weights. With this option, the grid is built on the fly in each SCF step. This results in a small additional computational overhead. For serial runs, the I/O of the grid data is usually not significant. However, this behavior is architecture dependent. In parallel runs, the grid I/O becomes a bottleneck. Therefore, the DIRECT option is the default setting in such calculations.

#### 4.3.7 Keyword FREEZE

This keyword specifies a frozen auxiliary function fit.

Options:

**NONE / CUSP / CORE / VALENCE**

<b>NONE</b>	No freezing of auxiliary function sets is performed.
<b>CUSP</b>	Auxiliary function sets with exponents above 1000 are frozen.
<b>CORE</b>	Auxiliary function sets with exponents above 100 are frozen.
<b>VALENCE</b>	Auxiliary function sets with exponents above 1 are frozen.
<b>TOL=&lt;Real&gt;</b>	User-defined tolerance for auxiliary function freezing.

This keyword freezes the coefficients of auxiliary function sets with exponents above the

given threshold to their free (spherical) atom values. This technique usually improves SCF convergence but at the cost of possibly unreliable results. Thus, careful testing is required.

## 4.4 Electronic State Control

### 4.4.1 Keyword MULTPLICITY

This keyword specifies the multiplicity of the system.

Options:

<Integer>      Multiplicity of the system.

Description:

The default multiplicities are 1 for closed and 2 for open shell systems. The program will check automatically if the specified multiplicity is allowed for the given molecular system considering its charge.

### 4.4.2 Keyword CHARGE

This keyword specifies the net charge of the system.

Options:

<Integer>      Charge of the system.

Description:

The default net charge is 0. Beware that net negative charges (anions) can lead to unphysical behavior with simple approximate density functionals [161].

### 4.4.3 Keyword MOEXCHANGE

This keyword alters the molecular orbital ordering in the start or restart density.

Options:

<Integer1>      Number of  $\alpha$  molecular orbital exchanges.

<Integer2>      Number of  $\beta$  molecular orbital exchanges.

Description:

The indices of the molecular orbitals to be exchanged are read in pairs of integers in the keyword body of MOEXCHANGE, one line for each pair. First all  $\alpha$  exchanges (or MO exchanges in case of RKS and ROKS) are read, and then, in the case of an unrestricted calculation (see UKS in 4.5.1)  $\beta$  exchanges are read. A maximum of 10 MO exchanges in each spin manifold is permitted. The following N<sub>2</sub> input shows how the start density, which here is the tight-binding density, can be altered with MOEXCHANGE.

```

Multiplicity 3
SCFTYPE ROKS Max=100
VxcType Basis BLYP
Mixing 0.4
Guess TB
MOExchange 1
7 9
PRINT MOs = 6-9
FIXMOS SCF=1 Fixed
Geometry Z-Matrix
N
N N R
#
Variables
R 1.0975
Basis (STO-3G)
End

```

In this example orbitals 7 and 9 are exchanged. In the output (see below), the orbitals 6 to 9 of the tight-binding start density are printed. This output is triggered by the PRINT keyword with the option "MOs = 6-9" (see 4.12.2). The exchange of orbital 7 and 9 results in an unoccupied orbital (9) with lower energy than that of the two highest occupied orbitals which are degenerate. Here is the output of the tight-binding orbitals before the SCF iteration:

#### TIGHT-BINDING MO COEFFICIENTS OF CYCLE 1

				6	7	8	9
				-.3043	.2841	.2841	-.1589
				2.0000	1.0000	1.0000	.0000
1	1	N	1s	.0000	.0000	.0000	.0413
2	1	N	2s	.0000	.0000	.0000	-.3246
3	1	N	2py	.5622	-.8336	-.0006	.0000
4	1	N	2pz	.0000	.0000	.0000	.6345
5	1	N	2px	.2728	-.0006	.8336	.0000
6	2	N	1s	.0000	.0000	.0000	.0413
7	2	N	2s	.0000	.0000	.0000	-.3246
8	2	N	2py	.5622	.8336	.0006	.0000
9	2	N	2pz	.0000	.0000	.0000	-.6345
10	2	N	2px	.2728	.0006	-.8336	.0000

Clearly, this is an excited state configuration which will not be preserved during the SCF iterations. With the keyword FIXMOS (see 4.4.5) this configuration can be fixed during the SCF iteration. With the options "SCF=1" and "Fixed" the configuration of the first SCF cycle (here the above displayed tight-binding MOs) is fixed and retained through the SCF. The converged molecular orbitals are also printed:

#### MO COEFFICIENTS OF CYCLE 8

6

7

8

9

				-.3091	.0647	.0647	-.3023
				2.0000	1.0000	1.0000	.0000
1	1	N	1s	.0000	.0000	.0000	.0741
2	1	N	2s	.0000	.0000	.0000	-.4167
3	1	N	2py	.6205	.8129	.1848	.0000
4	1	N	2pz	.0000	.0000	.0000	.5995
5	1	N	2px	-.0736	-.1848	.8129	.0000
6	2	N	1s	.0000	.0000	.0000	.0741
7	2	N	2s	.0000	.0000	.0000	-.4167
8	2	N	2py	.6205	-.8129	-.1848	.0000
9	2	N	2pz	.0000	.0000	.0000	-.5995
10	2	N	2px	-.0736	.1848	-.8129	.0000

Note the orbital energy of the lowest unoccupied MO.

#### 4.4.4 Keyword MOMODIFY

This keyword modifies the molecular orbital occupation.

Options:

- <Integer1> Number of  $\alpha$  molecular orbital occupation modifications.
- <Integer2> Number of  $\beta$  molecular orbital occupation modifications.

Description:

The modified molecular orbital occupations are read in the keyword body of MOMODIFY, one line for each modification. Such a modification consists of an integer number that specifies the molecular orbital to be modified and a real number for the molecular orbital occupation. First all  $\alpha$  MO modifications are read, and then, in the case of an unrestricted or spin-restricted open-shell calculation (see 4.5.1)  $\beta$  MO modifications are read. At most 10 molecular orbitals of each spin manifold can be modified. The following He atom input modifies the  $\beta$  occupation of orbital 1 to 0.75.

```
PRINT MOS
SCFTYPE UKS
MOMODIFY 0 1
1 0.75
GEOMETRY CARTESIAN ANGSTROM
He 0.0 0.0 0.0
```

Note that the MULTIPLICITY keyword (see 4.4.1) refers to the unmodified system, *i.e.* without fractional MO occupation. The fractional MO occupation can be best monitored by printing the molecular orbitals. We find for the above He atom input after SCF convergence:

```
ALPHA MO COEFFICIENTS OF CYCLE 6
```

				<sup>1</sup> -0.7814	<sup>2</sup> 0.3711
				1.0000	0.0000
1 2	1 1	He He	1s 2s	0.8294 0.2402	-0.9999 1.2768

## BETA MO COEFFICIENTS OF CYCLE 6

			<sup>1</sup> -0.7342	<sup>2</sup> 0.3936	
			0.7500	0.0000	
1 2	1 1	He He	1s 2s	0.8089 0.2661	-1.0166 1.2716

As this output shows the 0.75 occupation of the  $\beta$  MO breaks the energy degeneracy of the  $\alpha$  and  $\beta$  MOs. Also the MO coefficients are different. To obtain total energies that match with integer occupied calculations, e.g. for the cation or anion, the UKS SCFTYPE (see 4.5.1) option must be selected!

#### 4.4.5 Keyword FIXMOS

This keyword fixes a molecular orbital configuration by projection during all or part of the SCF.

Options:

**FIXED / ITERATIVE**

**FIXED**              The molecular orbital configuration at the beginning of the projection is used as reference configuration.

**ITERATIVE**          The molecular orbital configuration from the preceding SCF cycle is used as reference configuration.

**SCF=<Integer>** SCF cycle at which the projection procedure is turned on. By default the projection starts with the first SCF cycle.

**TOL=<Real>** SCF convergence tolerance after which the projection procedure is turned on.

Description:

The projection of the molecular orbital configuration is based on a maximum overlap criterion  $\Omega_{ij}$  of the form

$$\Omega_{ij} = \sum_{\mu,\nu} c_{\mu i}^{\text{new}} S_{\mu\nu} c_{\nu j}^{\text{old}} \quad (4.8)$$

where  $\{c_{\mu i}^{\text{new}}\}$  are the new molecular orbital coefficients and  $\{c_{\nu j}^{\text{old}}\}$  are the ones for the reference configuration. The example input in 4.4.3 shows the application of FIXMOS. The converged MO coefficients and energies clearly show the convergence to an excited state.

#### 4.4.6 Keyword SMEAR

This keyword specifies a certain kind of fractional occupation at the Fermi level. Note two things. This smearing is not necessarily the same as Fermi-level smearing in other codes. The variational principle may be violated if SMEAR is used.

Options:

<Real>	Energy range $\Delta E$ [in a.u.] around the HOMO energy in which orbitals are fractionally occupied.
<b>UNIFORM</b>	Specifies uniform fractional occupation within $\Delta E$ [in a.u.] around the HOMO level.

Description:

The SMEAR keyword affects only the molecular orbitals within the specified energy interval  $[E_{\text{HOMO}} - \Delta E/2, E_{\text{HOMO}} + \Delta E/2]$ . Therefore, the  $\Delta E$  value should be selected based on the orbital energy spectrum (see MOS option of PRINT in 4.12.2). The smearing is done by inverse proportionality to the energy interval of a given orbital energy to the reference energy  $E_{\text{HOMO}} - \Delta E/2$ . The closer the MO energy is to this reference energy, the larger will its occupation number be set. To enforce uniform occupation of the orbitals within the  $\Delta E$  interval, the option UNIFORM should be used. In either case, the converged fractional orbital occupation is used in any further step of the calculation (optimization, frequencies, properties etc.). The use of SMEAR for the symmetry-adapted calculation of the NO doublet is shown in the following example:

```

DIIS on
Guess TB
Smear 0.1
Print M0s=7-10
Geometry Z-Matrix
N
O N R
Variables
R 1.188

```

In this particular example the energy obtained is variational. The converged  $\alpha$  and  $\beta$  MO coefficients for orbitals 7 through 10 are:

```
ALPHA MO COEFFICIENTS OF CYCLE 10
```

				7	8	9	10
				-.4318	-.1778	-.1778	.1876
				1.0000	.5000	.5000	.0000
1	1	N	1s	.1121	.0000	.0000	.0817
2	1	N	2s	-.2273	.0000	.0000	-.1440
3	1	N	3s	-.4724	.0000	.0000	-1.5219
4	1	N	2py	.0000	.0034	.6623	.0000
5	1	N	2pz	.5063	.0000	.0000	-.3326
6	1	N	2px	.0000	.6623	-.0034	.0000
7	1	N	3py	.0000	.0021	.4037	.0000
8	1	N	3pz	.0927	.0000	.0000	-1.4349
9	1	N	3px	.0000	.4037	-.0021	.0000
10	1	N	3d-2	.0000	.0000	.0000	.0000
11	1	N	3d-1	.0000	.0000	-.0074	.0000
12	1	N	3d+0	.0323	.0000	.0000	.0304
13	1	N	3d+1	.0000	-.0074	.0000	.0000
14	1	N	3d+2	.0000	.0000	.0000	.0000
15	2	0	1s	-.0259	.0000	.0000	.1201
16	2	0	2s	.0483	.0000	.0000	-.2451
17	2	0	3s	-.0984	.0000	.0000	1.5780
18	2	0	2py	.0000	-.0028	-.5499	.0000
19	2	0	2pz	-.4739	.0000	.0000	-.2873
20	2	0	2px	.0000	-.5499	.0028	.0000
21	2	0	3py	.0000	-.0017	-.3420	.0000
22	2	0	3pz	-.1745	.0000	.0000	-.9839
23	2	0	3px	.0000	-.3420	.0017	.0000
24	2	0	3d-2	.0000	.0000	.0000	.0000
25	2	0	3d-1	.0000	-.0001	-.0117	.0000
26	2	0	3d+0	.0195	.0000	.0000	-.0004
27	2	0	3d+1	.0000	-.0117	.0001	.0000
28	2	0	3d+2	.0000	.0000	.0000	.0000

## BETA MO COEFFICIENTS OF CYCLE 10

				7	8	9	10
				-.4072	-.1391	-.1391	.1995
				1.0000	.0000	.0000	.0000
1	1	N	1s	-.1196	.0000	.0000	.0811
2	1	N	2s	.2374	.0000	.0000	-.1270
3	1	N	3s	.5017	.0000	.0000	-1.5760
4	1	N	2py	.0000	.1054	.6454	.0000
5	1	N	2pz	-.4988	.0000	.0000	-.3042
6	1	N	2px	.0000	.6454	-.1054	.0000
7	1	N	3py	.0000	.0714	.4376	.0000
8	1	N	3pz	-.0992	.0000	.0000	-1.4761
9	1	N	3px	.0000	.4376	-.0714	.0000
10	1	N	3d-2	.0000	.0000	.0000	.0000
11	1	N	3d-1	.0000	-.0008	-.0048	.0000
12	1	N	3d+0	-.0281	.0000	.0000	.0221
13	1	N	3d+1	.0000	-.0048	.0008	.0000
14	1	N	3d+2	.0000	.0000	.0000	.0000
15	2	0	1s	.0201	.0000	.0000	.1206

16	2	0	2s	-.0323	.0000	.0000	-.2389
17	2	0	3s	.0819	.0000	.0000	1.6249
18	2	0	2py	.0000	-.0849	-.5200	.0000
19	2	0	2pz	.4565	.0000	.0000	-.2597
20	2	0	2px	.0000	-.5200	.0849	.0000
21	2	0	3py	.0000	-.0570	-.3493	.0000
22	2	0	3pz	.1750	.0000	.0000	-1.0090
23	2	0	3px	.0000	-.3493	.0570	.0000
24	2	0	3d-2	.0000	.0000	.0000	.0000
25	2	0	3d-1	.0000	-.0022	-.0136	.0000
26	2	0	3d+0	-.0185	.0000	.0000	.0048
27	2	0	3d+1	.0000	-.0136	.0022	.0000
28	2	0	3d+2	.0000	.0000	.0000	.0000

Note the symmetry-adapted occupation of the  $\alpha$  MOs 8 and 9.

#### 4.4.7 Keyword CONFIGURE

This keyword controls the atomic configuration. Therefore it is applicable **only** to calculations for single atoms.

Options:

**MAX=<Integer>** Maximum number of SCF cycles for which the configuration is used.  
By default it is used for all SCF cycles.

**OCCUPY** The (fractional) occupation of the specified atomic configuration is given explicitly in the subsequent input.

Description:

The orbital configuration is defined in the first line of the keyword body of CONFIGURE. In the case of OKS and ROKS calculations (see 4.5.1), two lines, the first for the  $\alpha$ -orbital configuration and the second for the  $\beta$ -orbital configuration, are required. If spherical orbitals (see 4.5.2) are to be used, one integer for  $s$ , three for  $p$ , five for  $d$ , etc. must be specified. Empty shells (see 4.3.1 for the definition of a shell) can be omitted. In deMon2k spherical orbitals are defined over real spherical harmonic Gaussians [162] as (without normalization):

$$\phi_{lm}(\mathbf{r}) = r^l e^{-\alpha r^2} S_l^m(\vartheta, \phi) \quad (4.9)$$

The  $S_l^m$  denote real spherical harmonics that are constructed from complex spherical harmonics by [163]:

$$S_l^m = \frac{1}{\sqrt{2}} (Y_l^m + Y_l^{-m}) \quad (4.10)$$

$$S_l^{-m} = -\frac{i}{\sqrt{2}} (Y_l^m - Y_l^{-m}) \quad (4.11)$$

It should be noted that the  $S_l^m$  are not eigenfunctions of the  $\hat{l}_z$  operator and, therefore,  $m$  is not a good quantum number for these orbitals.

The ordering of the integer numbers in the configuration line follows the  $l$  (shell) and  $m$  index of the real spherical harmonics:

Shell	$s$	$p$			$d$			...		
$l$	0		1			2		...		
$m$	0	-1	0	1	-2	-1	0	1	2	...
Orbital	$s$	$p_y$	$p_z$	$p_x$	$d_{xy}$	$d_{yz}$	$d_{z^2}$	$d_{xz}$	$d_{x^2-y^2}$	...

As an example the open-shell (OKS or ROKS) triplet ground state configuration of the carbon atom,  $1s^2 2s^2 2p^2$ , can be defined in the keyword block of CONFIGURE as:

```
2 1 1 0
2 0 0 0
```

The first line defines the  $\alpha$ -orbital configuration with two electrons in the most stable  $s$ -orbitals ( $1s$  and  $2s$ ) and one electron each in the most stable  $p_y$  ( $m = -1$ ) and  $p_z$  ( $m = 0$ )  $\alpha$ -orbital ( $2p_y$  and  $2p_z$ ). The second line assigns two electrons to the two most stable  $\beta$   $s$ -orbitals. Unspecified shells (here  $d$  and higher) are not considered. If Cartesian orbitals are used, only the number of electrons in the  $s$ ,  $p$ ,  $d$ , etc. shells needs to be specified. Thus, for Cartesian orbitals the foregoing configuration definition changes to:

```
2 2
2 0
```

Because of the  $p$ -orbital degeneracy, the chosen Cartesian configuration has to be stabilized during the SCF procedure (see 4.4.5 or 4.5.7). The calculation of the triplet carbon ground state with spherical and Cartesian orbitals using CONFIGURE is described by the examples 2.16 and 2.17 in the deMon2k Tutorial, respectively.

To access excited atomic states or to use fractional occupation numbers, the option OC-CUPY must be used. In this case, an explicit definition of the orbital occupation is expected after the orbital configuration line(s). In the case of spherical atomic orbitals an input line describing the occupation of each real spherical harmonic orbital type is needed. The occupation numbers can be given as real or integer values. To describe the excited triplet state of carbon,  $1s^2 2s^1 2p^2 3s^1$ , for example, the following keyword body of CONFIGURE, using the option OCCUPY, is needed.

```
3 1 1 0
2 0 0 0
1 0 1
1
1
1 1
```

Here, the 3 in the first configuration line indicates that the three lowest  $s$ -type  $\alpha$ -orbitals will be occupied (note that this number does not represent the number of electrons in these orbitals). The 1 1 0 in the first line indicates that the  $p_y$  and  $p_z$   $\alpha$ -orbitals will be occupied (we assume the default spherical orbitals with the order given above). The second line gives the  $\beta$  configuration. Here, only the two lowest  $\beta s$  orbitals will be occupied. Occupation patterns begin in the third line, which, for the three lowest  $s$ -type  $\alpha$  orbitals in this example is ( 1 0 1 ). This scheme has the lowest  $\alpha s$ -orbital occupied with 1 electron, the next one empty, and the third one again occupied with 1 electron. Therefore, a hole in the  $\alpha s$ -orbital occupation is produced. The next two occupation lines assign 1 electron to the  $p_y$  and 1 to the  $p_z$   $\alpha$ -orbital. The last occupation has the two lowest  $\beta s$ -orbitals occupied with 1 electron each. Please also note that zero (0) entries in the configuration line(s) do not have corresponding occupation lines. Example 2.18 on page 24 of the deMon2k Tutorial describes the calculation of the excited  $1s^2 2s^1 2p^2 3s^1$  carbon triplet state.

As already mentioned, the OCCUPY option may also be used to generate fractional occupations, e.g. for the calculation of spherical atoms. In the case of the triplet carbon ground state, the following CONFIGURE keyword body produces a spherical atom.

```
2 1 1 1
2 0 0 0
1 1
0.6666
0.6666
0.6666
1 1
```

The first configuration line describes the occupation of two  $\alpha s$ -orbitals and of all three  $p$ -orbitals. According to the specified occupation (third line), the two  $s$ -orbitals are occupied by one electron each. The three  $\alpha p$ -orbitals are occupied uniformly by 2/3 of an electron each (0.6666 in lines 4, 5, and 6). Finally, the  $\beta s$ -orbital occupation is given by the last line. (See Example 2.19 on page 25 of the tutorial for the discussion of the corresponding output).

Because the Kohn-Sham method is a single-determinant approach, atomic states are approximated by a single configuration (see however [164] for a multi-determinantal approach). This can be done as in spatially unrestricted Hartree-Fock calculations [165]. For  $s$  and  $p$  occupations, all possible configurations yield the correct spatial symmetry. However, for  $d$  occupations this is not the case and care must be taken with the occupation scheme. Correct spatial symmetry uniquely defines the orbital occupancies for  $d^2$ ,  $d^3$ ,  $d^5$ ,  $d^7$  and  $d^8$ . The other  $d$  occupations are selected to maximize the absolute angular momen-

tum (this choice is arbitrary because  $m$  is no longer a good quantum number). Thus, the following  $d$  configurations must be used:

$$\begin{aligned}d^1(^2D) &: (d_{xy})^1 \\d^2(^3F) &: (d_{z^2})^1 (d_{x^2-y^2})^1 \\d^3(^4F) &: (d_{xy})^1 (d_{xz})^1 (d_{yz})^1 \\d^4(^5D) &: (d_{z^2})^1 (d_{xy})^1 (d_{xz})^1 (d_{yz})^1 \\d^5(^6S) &: (d_{z^2})^1 (d_{x^2-y^2})^1 (d_{xy})^1 (d_{xz})^1 (d_{yz})^1 \\d^6(^5D) &: (d_{z^2})^1 (d_{x^2-y^2})^1 (d_{xy})^2 (d_{xz})^1 (d_{yz})^1 \\d^7(^4F) &: (d_{z^2})^2 (d_{x^2-y^2})^2 (d_{xy})^1 (d_{xz})^1 (d_{yz})^1 \\d^8(^3F) &: (d_{z^2})^1 (d_{x^2-y^2})^1 (d_{xy})^2 (d_{xz})^2 (d_{yz})^2 \\d^9(^2D) &: (d_{z^2})^2 (d_{x^2-y^2})^2 (d_{xy})^1 (d_{xz})^2 (d_{yz})^2\end{aligned}$$

## 4.5 SCF Control and Stabilization

### 4.5.1 Keyword SCFTYPE

With this keyword the SCF method is selected and SCF convergence control is specified.

Options:

#### **RKS / UKS / ROKS / CUKS**

**RKS** The restricted Kohn-Sham method will be used. This is the default for closed-shell systems.

**UKS** The unrestricted Kohn-Sham method will be used. This is the default for open-shell systems.

**ROKS** The spin-restricted open-shell Kohn-Sham method will be used.

<ROKS option> The ROKS option string defines the ROKS parametrization according to Table 8. By default the Guest & Saunders parametrization [174] is used.

**CUKS** The constrained unrestricted Kohn-Sham method will be used.

**MAX=<Integer>** Maximum number of SCF iterations. Default is 100.

**TOL=<Real>** MinMax SCF energy convergence criterion. Default is  $10^{-5}$  a.u.

**CDF=<Real>** Auxiliary density convergence criterion. Default is  $5 \cdot 10^{-4}$  a.u.

**NOTIGHTEN** The SCF convergence criteria will not be adjusted during an optimization, frequency analysis or property calculation.

Description:

The implementations of the Kohn-Sham SCF methods are similar to the corresponding Hartree-Fock methods [166–170] and have similar advantages and disadvantages. For historical reasons the three-center ERI Hartree-Fock methods are also selected by the RKS, UKS, ROKS and CUKS options of the SCFTYPE keyword. In the case of UKS calculations, spin-contamination may appear. In deMon2k, the approximate spin contamination [171, 172] is calculated and printed just before the converged UKS SCF energy. This printed value has proven to be a good guide for evaluation of spin-contamination in UKS calculations. The ROKS implementation [173] in deMon2k follows the original Roothaan work for Hartree-Fock [169]. Six different parametrizations for the mixing factors of the diagonal

blocks of the ROKS Kohn-Sham matrix,

$$\begin{pmatrix} \mathbf{K}^2 & \mathbf{K}^\beta & \frac{\mathbf{K}^\alpha + \mathbf{K}^\beta}{2} \\ \mathbf{K}^\beta & \mathbf{K}^1 & \mathbf{K}^\alpha \\ \frac{\mathbf{K}^\alpha + \mathbf{K}^\beta}{2} & \mathbf{K}^\alpha & \mathbf{K}^0 \end{pmatrix}, \quad (4.12)$$

where  $\mathbf{K}^\alpha$  and  $\mathbf{K}^\beta$  are the usual  $\alpha$  and  $\beta$  Kohn-Sham matrices in MO representation and  $\mathbf{K}^2$ ,  $\mathbf{K}^1$  and  $\mathbf{K}^0$  are the diagonal blocks for double, single and unoccupied MOs, are implemented in deMon2k. The corresponding mixing equations are given by:

$$\begin{aligned} \mathbf{K}^2 &= A_2 \mathbf{K}^\alpha + B_2 \mathbf{K}^\beta \\ \mathbf{K}^1 &= A_1 \mathbf{K}^\alpha + B_1 \mathbf{K}^\beta \\ \mathbf{K}^0 &= A_0 \mathbf{K}^\alpha + B_0 \mathbf{K}^\beta \end{aligned}$$

The mixing factors of the different parametrizations [175] along with their SCFTYPE keyword options are given in Table 8.

Table 8: Mixing factors and SCFTYPE options of the ROKS parametrizations available in deMon2k.

Parametrization	$A_0$	$B_0$	$A_1$	$B_1$	$A_2$	$B_2$	SCFTYPE option
Guest & Saunders [174]	1/2	1/2	1/2	1/2	1/2	1/2	<b>GUEST</b> (default)
Rootaan [169]	3/2	-1/2	1/2	1/2	-1/2	3/2	<b>ROOTHAAN</b>
Davidson [176]	1	0	1	0	1/2	1/2	<b>DAVIDSON</b>
Binkley et al. [170]	0	1	1	0	1/2	1/2	<b>BINKLEY</b>
McWeeny & Diercksen [177]	2/3	1/3	1/3	1/3	1/3	2/3	<b>MCWEENY</b>
Faegri & Manne [178]	1/2	1/2	1	0	1/2	1/2	<b>FAEGRI</b>
Plakhutin et al. [175]	1	0	1	0	0	1	<b>PLAKHUTIN</b>

Note that the ROKS orbital energies vary with respect to these parametrizations and, therefore, are not uniquely defined. In particular, the Aufbau principle for the doubly and singly occupied molecular orbital energy manifolds can be violated [179,180]. Moreover, ROKS orbital energies include additional constraints and cannot be compared with RKS or UKS orbital energies. Specifically, the Slater-Janak theorem [181,182] does not hold for ROKS calculations. For this reason the block-diagonalized ROKS  $\alpha$  and  $\beta$  MO energies and coefficients [183], also named semi-canonical spin orbitals, are printed after the ROKS energy as shown in the following example output for a B3LYP/STO-3G/GEN-A2\* calculation of the triplet O<sub>2</sub> ground state.

\*\*\* SCF CONVERGED \*\*\*

MO COEFFICIENTS OF CYCLE 7

				7	8	9	10
				-0.3794	-0.0601	-0.0601	0.4267
				2.0000	1.0000	1.0000	0.0000
1	1	0	1s	0.0000	0.0000	0.0000	-0.0917
2	1	0	2s	0.0000	0.0000	0.0000	0.5671
3	1	0	2py	0.6576	0.0822	0.7637	0.0000
4	1	0	2pz	0.0000	0.0000	0.0000	0.9484
5	1	0	2px	0.0374	0.7637	-0.0822	0.0000
6	2	0	1s	0.0000	0.0000	0.0000	0.0917
7	2	0	2s	0.0000	0.0000	0.0000	-0.5671
8	2	0	2py	0.6576	-0.0822	-0.7637	0.0000
9	2	0	2pz	0.0000	0.0000	0.0000	0.9484
10	2	0	2px	0.0374	-0.7637	0.0822	0.0000

RANDOMIZED SCF GRID GENERATED IN 3 CYCLES

REFERENCE VALUE OF S\*\*2 FOR PURE SPIN STATE S(S+1): 2.0000

S\*\*2 BEFORE SPIN PROJECTION: 2.0000  
 S\*\*2 AFTER SPIN PROJECTION: 2.0000

ELECTRONIC CORE ENERGY	=	-260.328412563
ELECTRONIC COULOMB ENERGY	=	101.099784250
ELECTRONIC HARTREE ENERGY	=	-159.228628313
EXCHANGE ENERGY	=	-16.350046036
CORRELATION ENERGY	=	-0.733801619
EXCHANGE-CORRELATION ENERGY	=	-17.083847655
ELECTRONIC SCF ENERGY	=	-176.312475968
NUCLEAR-REPULSION ENERGY	=	28.059106307
TOTAL ENERGY	=	-148.253369660

BLOCK DIAGONAL ALPHA MO COEFFICIENTS

				7	8	9	10
				-0.4028	-0.1618	-0.1618	0.4020
				1.0000	1.0000	1.0000	0.0000
1	1	0	1s	0.0729	0.0000	0.0000	-0.0917
2	1	0	2s	-0.3516	0.0000	0.0000	0.5671
3	1	0	2py	0.0000	-0.7148	-0.2813	0.0000
4	1	0	2pz	0.6083	0.0000	0.0000	0.9484
5	1	0	2px	0.0000	0.2813	-0.7148	0.0000

6	2	0	1s	0.0729	0.0000	0.0000	0.0917
7	2	0	2s	-0.3516	0.0000	0.0000	-0.5671
8	2	0	2py	0.0000	0.7148	0.2813	0.0000
9	2	0	2pz	-0.6083	0.0000	0.0000	0.9484
10	2	0	2px	0.0000	-0.2813	0.7148	0.0000

## BLOCK DIAGONAL BETA MO COEFFICIENTS

				7	8	9	10
				-0.3239	0.0415	0.0415	0.4515
				1.0000	0.0000	0.0000	0.0000
1	1	0	1s	0.0000	0.0000	0.0000	-0.0917
2	1	0	2s	0.0000	0.0000	0.0000	0.5671
3	1	0	2py	-0.3948	-0.7637	-0.0825	0.0000
4	1	0	2pz	0.0000	0.0000	0.0000	0.9484
5	1	0	2px	0.5272	0.0825	-0.7637	0.0000
6	2	0	1s	0.0000	0.0000	0.0000	0.0917
7	2	0	2s	0.0000	0.0000	0.0000	-0.5671
8	2	0	2py	-0.3948	0.7637	0.0825	0.0000
9	2	0	2pz	0.0000	0.0000	0.0000	0.9484
10	2	0	2px	0.5272	-0.0825	0.7637	0.0000

The printing of the MO energies and coefficients is activated with PRINT MOS=7-10 (see Section 4.12.2). The block diagonal  $\alpha$  and  $\beta$  MO energies and coefficients can be directly compared with their UKS counterparts. In deMon2k they are also used for ROKS perturbation theory calculations. As an alternative to ROKS the constrained unrestricted Kohn-Sham (CUKS) method can be used for spin-projection [184] where the semi-canonical orbitals are obtained directly. The convergence behavior of ROKS and CUKS calculations is generally different and it is advisable to switch among them to test convergence in problematic cases.

As a consequence of the variational fitting of the density [5,6], deMon2k can exploit a MinMax SCF procedure [185]. In a variational MinMax procedure, there is no strict convergence from above. Therefore, it is possible to obtain energies below the converged energy during the SCF iterations in deMon2k. Note that the printed SCF ERROR for an SCF cycle is the difference between upper and lower MinMax energy bounds and, thus, a direct measurement of how far the current SCF step is away from the convergence point.

The maximum number of SCF iterations is specified with the MAX option. With MAX=0 an "energy only" calculation with the molecular orbital coefficients from the restart file can be performed. No SCF iteration is done! The MAX=0 option automatically triggers GUESS

RESTART (see 4.5.5) and, therefore, fails if no adequate restart file `deMon.rst` exists. The ordering and, thus, the occupation of the molecular orbitals in the restart file can be altered with the MOEXCHANGE keyword (see Section 4.4.3).

The SCF energy convergence criterion can be defined by the user with the TOL option. Such a user-defined SCF convergence criterion is valid for the first single-point SCF calculation. During a geometry optimization, however, the convergence criterion is automatically tightened according to the residual forces (see Table 9). However, if the user-defined convergence criterion is smaller than the automatic-tightening value, the user-defined value is used instead. If self-consistent perturbation calculations are performed, the automatically requested SCF energy convergence criterion is  $10^{-6}$  Hartree. It too may be overridden by a smaller value with the TOL option.

Table 9: Tightening of the SCF convergence criteria during structure optimization.

RMS Gradient [a.u.]	Tightened SCFTYPE Option	
	TOL[a.u.]	CDF
$\leq 0.02$	$1.0 \cdot 10^{-5}$	$1.0 \cdot 10^{-3}$
$\leq 0.01$	$5.0 \cdot 10^{-6}$	$5.0 \cdot 10^{-4}$
$\leq 0.005$	$1.0 \cdot 10^{-7}$	$1.0 \cdot 10^{-4}$
$\leq 0.001$	$7.5 \cdot 10^{-8}$	$7.5 \cdot 10^{-5}$
$\leq 7.5 \cdot 10^{-4}$	$5.0 \cdot 10^{-8}$	$5.0 \cdot 10^{-5}$
$\leq 5.0 \cdot 10^{-4}$	$2.5 \cdot 10^{-8}$	$2.5 \cdot 10^{-5}$
$\leq 1.0 \cdot 10^{-4}$	$1.0 \cdot 10^{-8}$	$1.0 \cdot 10^{-5}$
$\leq 7.5 \cdot 10^{-5}$	$7.5 \cdot 10^{-9}$	$7.5 \cdot 10^{-6}$
$\leq 5.0 \cdot 10^{-5}$	$5.0 \cdot 10^{-9}$	$5.0 \cdot 10^{-6}$
$\leq 2.5 \cdot 10^{-5}$	$2.5 \cdot 10^{-9}$	$2.5 \cdot 10^{-6}$
$\leq 1.0 \cdot 10^{-5}$	$1.0 \cdot 10^{-9}$	$1.0 \cdot 10^{-6}$
$\leq 7.5 \cdot 10^{-6}$	$7.5 \cdot 10^{-10}$	$7.5 \cdot 10^{-7}$
$\leq 5.0 \cdot 10^{-6}$	$5.0 \cdot 10^{-10}$	$5.0 \cdot 10^{-7}$
$\leq 2.5 \cdot 10^{-6}$	$2.5 \cdot 10^{-10}$	$2.5 \cdot 10^{-7}$
$\leq 1.0 \cdot 10^{-6}$	$1.0 \cdot 10^{-10}$	$1.0 \cdot 10^{-7}$
$\leq 7.5 \cdot 10^{-7}$	$1.0 \cdot 10^{-10}$	$7.5 \cdot 10^{-8}$
$\leq 5.0 \cdot 10^{-7}$	$1.0 \cdot 10^{-10}$	$5.0 \cdot 10^{-8}$
$\leq 2.5 \cdot 10^{-7}$	$1.0 \cdot 10^{-10}$	$2.5 \cdot 10^{-8}$

The auxiliary density convergence criterion can be defined by the user with the CDF option. As with the user-defined energy convergence criterion, the user-defined auxiliary density convergence criterion holds for the first single-point SCF calculation and overrides the automatically determined CDF values (see Table 9) during the optimization if the user-defined

value is smaller. If self-consistent perturbation calculations are performed, the default auxiliary density convergence criterion is tightened to  $5 \cdot 10^{-5}$ . Again, this value can be overridden by a smaller user-defined value with the CDF option. At SCF convergence both energy and auxiliary density convergence criteria are satisfied.

The option NOTIGHTEN disables the automatic tightening of the SCF convergence criteria according to the root mean square (RMS) gradient (as shown in Table 9). Thus, the SCF convergence tolerance is relaxed during geometry optimization and, therefore, SCF convergence failures are less likely. However, accuracy of the gradients is compromised and the structural optimization may not converge. Good practice, therefore, is to use NOTIGHTEN only at the beginning of the optimization or in combination with carefully tested user-defined TOL and CDF values. The NOTIGHTEN option also disables tightening of SCF convergence for perturbation calculations.

#### 4.5.2 Keyword ORBITALS

This keyword controls the atomic orbital choice.

Options:

##### SPHERICAL / CARTESIAN

**SPHERICAL**      Spherical atomic orbitals ( $5d, 7f$ ) are used. This is the default.

**CARTESIAN**      Cartesian atomic orbitals ( $6d, 10f$ ) are used.

Description:

In deMon2k the spherical (see 4.4.7 for the definition of real spherical harmonic Gaussians) and Cartesian atomic orbitals have the general form (without normalization):

$$\begin{aligned}\phi_{lm}(\mathbf{r}) &= r^l e^{-\alpha r^2} S_l^m(\vartheta, \phi) \\ \phi_{ijk}(\mathbf{r}) &= x^i y^j z^k e^{-\alpha r^2}\end{aligned}$$

Because spherical orbitals have no linear dependencies, they are the recommended choice for most applications. See 4.4.7 regarding CARTESIAN orbitals in the case of an atom.

#### 4.5.3 Keyword PHASE

This keyword permits the explicit definition of molecular orbital phases.

Options:

<Integer> Molecular orbital for which an explicit phase convention is defined in the input.

Description:

The molecular orbital (MO) phases are undefined in standard quantum chemical calculations. Therefore, the phase may differ for different diagonalization routines or computational architectures. The keyword PHASE permits the definition of a user-defined phase convention. This is useful if molecular orbitals are plotted for different geometries, *e.g.* in correlation diagrams, along reaction coordinates, or for a movie from a molecular dynamics run. If the keyword PHASE is specified without any options, then the MO phase is chosen such that the largest molecular orbital coefficient is positive. This choice usually preserves the MO phases for small geometry changes. However, it is not fail-safe. If the molecular orbital phase must be preserved over large geometry changes, an explicit phase convention should be specified with the PHASE keyword combined with some integer numbers. The integer numbers specify the molecular orbitals for which an explicit phase convention is given in the input. The reference MO coefficients are then given in free format, one after the other, respecting the 80 character length limit for each input line, in the body of the keyword PHASE as the following example shows:

```
PHASE 21 22
-0.03127 +0.15629 -0.31825 -0.04906 -0.02041 +0.03267 -0.00927 -0.02741
+0.02988 +0.00172 +0.00354 -0.00638 +0.00096 +0.00580 +0.00753 +0.00063
+0.00101 +0.00121 -0.03098 +0.15345 -0.32858 -0.06885 +0.01440 -0.03169
+0.02076 +0.01737 -0.03349 +0.01821 -0.00553 +0.00307 +0.00218 +0.00501
+0.00624 +0.00003 +0.00351 -0.00073 -0.01639 +0.08148 -0.16645 -0.09271
-0.02134 +0.00160 +0.02254 -0.03551 +0.00426 +0.03639 -0.01114 +0.00271
+0.00982 +0.00644 +0.00203 +0.00302 +0.00077 -0.00037 -0.02184 +0.10920
-0.21773 -0.10390 +0.02340 -0.00197 -0.03176 +0.03411 -0.00520 -0.04247
+0.00826 -0.00167 -0.00962 +0.00639 +0.00263 +0.00267 +0.00082 -0.00181
-0.00645 +0.03270 -0.06593 -0.03674 +0.04837 -0.00504 -0.05537 +0.12202
-0.02485 -0.12622 +0.03715 -0.00640 -0.03966 -0.00016 -0.00982 +0.00232
+0.00760 -0.00687 +0.00398 -0.02055 +0.03840 +0.01576 +0.04400 -0.00549
-0.05095 +0.11616 -0.02418 -0.12269 +0.02478 -0.00534 -0.02588 -0.00075
+0.01033 -0.00326 -0.00795 +0.01127 -0.03146 +0.15476 -0.35558 -0.21744
-0.00997 -0.00249 +0.01294 +0.00384 -0.00194 -0.00387 -0.00141 +0.00248
-0.00098 -0.00335 +0.00100 -0.00271 +0.00256 +0.00016 +0.02919 -0.14448
+0.32441 +0.19459 -0.00417 -0.00267 +0.01116 +0.01717 -0.00444 -0.01650
+0.00471 -0.00559 -0.00013 +0.00552 -0.00060 +0.00393 -0.00208 -0.00028
```

The foregoing input defines explicit phase conventions for MOs 21 and 22. After the diagonalization of the Kohn-Sham matrix, the phases of the new MOs 21 and 22 are obtained by projection with the reference MO coefficients given in the preceding definition.

#### 4.5.4 Keyword ERIS

This keyword controls the calculation method for the three-center electron repulsion integrals (ERIs).

Options:

**MULTIPOLE / CONVENTIONAL / DIRECT / MIXED**

**MULTIPOLE** All ERIs are recalculated at each SCF iteration. For long-range ERIs the double asymptotic expansion is used. This is the default.

**CONVENTIONAL** All ERIs are calculated at the beginning of the SCF procedure and stored.

**DIRECT** All ERIs are recalculated at each SCF iteration.

**MIXED** Short-range ERIs are calculated at the beginning of the SCF procedure and stored in RAM. Long-range ERIs are recalculated in each SCF iteration employing the double asymptotic expansion.

**RAM=<Real>** Specifies the RAM per core [in MB] usable in the calculation.

**TOL=<Real>** Threshold for ERI screening.

Description:

The ERIS default option MULTIPOLE for the three-center ERI calculations is a compromise between computational performance and memory (RAM) demand [187]. Because no ERIs are stored the full RAM size is available for SCF matrices. At the same time the double asymptotic expansion for the long-range ERIs [74] improves the computational performance. The ERIS option DIRECT [186] is similar to the MULTIPOLE option. The only difference is that the double asymptotic expansion for the long-range ERIs is disabled. Thus, all ERIs are recalculated by recurrence relations [51] twice in each SCF iteration. As a result, the DIRECT option is always computationally more demanding than the MULTIPOLE option. It should only be used for testing and benchmarking.

Whereas the MULTIPOLE option is the method of choice for calculations that are memory bound ( $\geq 25000$  basis functions), the MIXED option can be computationally beneficial for smaller systems, particularly in parallel runs with many SCF iterations. With this option the short-range ERIs are calculated only once before the SCF procedure and stored in RAM. The long-range ERIs are calculated in each SCF iteration employing the double asymptotic expansion. In parallel runs the ERI storage is distributed over all cores that have free RAM space, i.e. that are not involved in the allocation of SCF matrices. In Table

10 average timings per SCF cycle for PBE/DZVP/GEN-A2 calculations of n-alkane chains with the DIRECT, MULTIPOLE and MIXED option of the ERIS keyword are shown. Also listed are the number of basis functions,  $N_{Basis}$ , the number of auxiliary functions,  $N_{Auxis}$ , and the number of SCF cycles until convergence is reached. Note that the number of SCF cycles was the same for all options and that the converged energies were identical to  $10^{-7}$  a.u. or better. All calculations were performed in parallel on a single compute node with 2 octo-core Intel Xeon E5-2650v2 CPUs @ 2.60GHz with a total of 64 GB RAM.

Table 10: Average time per SCF cycle [sec] for PBE/DZVP/GEN-A2 calculations of n-alkanes using different options of the ERIS keyword.

n-Alkane	$N_{Basis}$	$N_{Auxis}$	SCF cycles	DIRECT	MULTIPOLE	MIXED
C <sub>100</sub> H <sub>202</sub>	2510	4208	15	19	13	11
C <sub>150</sub> H <sub>302</sub>	3760	6308	16	57	42	39
C <sub>200</sub> H <sub>402</sub>	5010	8408	16	122	97	89
C <sub>250</sub> H <sub>502</sub>	6260	10508	39	182	137	136
C <sub>300</sub> H <sub>602</sub>	7510	12608	39	294	229	226

Table 10 shows that the ERIS option choice makes a noticeable difference for smaller systems where the linear algebra tasks are still not dominant. For such systems the MIXED option is advisable, particularly for Born-Oppenheimer molecular dynamics simulations (see Section 4.7).

To monitor the RAM space for the ERIS MIXED option a RAM allocation table can be printed with PRINT RAM (see also 4.12.2). The following example shows such a table for a C<sub>24</sub>H<sub>50</sub> calculation with the aug-cc-pVQZ basis (5270 basis functions) and GEN-A2 auxiliary function set employing 32 cores with 4 GB RAM each.

```
*** ERI STATISTIC ***
Est. Integrated ERIs: 8486773340
Est. Asymptotic ERIS: 96798170
```

```
*** RAM Allocation ***
Program part          Size in MBytes
SCF Kernel           905.121
ERI Kernel            4.416
DAE Kernel            1.152
FIT Kernel             0.613
LAG Kernel            225.340
```

Max RAM	4096.000
Max SHM	65536.000

Integrated ERIs: Incore on 31 CPUs  
 Asymptotic ERIs: Direct SCF method

\*\*\* Incore ERI Storage \*\*\*

#CPU	#ERIs	Sizes [MBytes]		
		ERI	Vector	Max. Size
0	277791345	SCF	kernel allocation	
1	273321035		2119.380	3614.660
2	270188345		2085.274	3614.660
3	270513580		2061.373	3614.660
4	270640015		2063.855	3614.660
5	271920380		2064.819	3614.660
6	273630545		2074.588	3614.660
7	278321900		2087.635	3614.660
8	277542740		2123.428	3614.660
9	279195440		2117.483	3614.660
10	278907830		2130.092	3614.660
11	274740355		2127.898	3614.660
12	276984955		2096.103	3614.660
13	276904470		2113.228	3614.660
14	274954730		2112.613	3614.660
15	273873285		2097.738	3614.660
16	274457520		2089.487	3614.660
17	272614335		2093.945	3614.660
18	270426255		2079.882	3614.660
19	270650360		2063.189	3614.660
20	270687890		2064.898	3614.660
21	270468465		2065.185	3614.660
22	271087330		2063.511	3614.660
23	273499175		2068.232	3614.660
24	274764620		2086.633	3614.660
25	273490215		2096.288	3614.660
26	271100105		2086.565	3614.660
27	274784945		2068.330	3614.660
28	272456980		2096.443	3614.660
29	272598035		2078.682	3614.660
30	274256160		2079.758	3614.660
31			2092.408	3614.660

The ERI STATISTIC lists the estimated number of ERIs calculated by recurrence relations (Est. Integrated ERIs) and by the double asymptotic expansion (Est. Asymptotic ERIS). These numbers are estimates because screening due to density matrix elements or fitting coefficients is not included. The following RAM Allocation table lists the RAM sizes required for individual calculation tasks. These are self-consistent field (SCF) iteration, near-field ERI recurrence relation (ERI), double asymptotic expansion (DAE), density fitting (FIT) and linear algebra (LAG) operations. The following two lines, Max RAM and Max SHM, print the maximum RAM per CPU (in this case set by the MAXRAM parameter; see Table 1) and the maximum shared memory size available to the SCF matrices. Note that

the maximum shared memory size for the SCF matrices is 16 times the maximum RAM per CPU because there are 16 CPUs on each board of the cluster here used. The following output is specific to ERIS MIXED. It states that the near-field ERIs (Integrated ERIs) are held in-core on 31 CPUs and that the double-asymptotically-expanded ERIs (Asymptotic ERIs) are calculated according to the direct SCF method, i.e. they are recalculated twice (once for the Kohn-Sham matrix and another time for the Coulomb vector) in each SCF iteration. As this table shows, no ERIs are stored on CPU 0 because its RAM is used for the storage of the SCF matrices. On the other 31 CPUs a little bit more than 2 GB are used for ERI storage. With the ERIS MIXED option the computational time for ERI calculation can be reduced to below 10% of the total computational time [188]. Thus it is advisable to explore if the ERIS MIXED option can be used for an application at hand. Note that in the case of a serial run with only near-field ERIs the ERIS MIXED option is equivalent to an in-core SCF.

The CONVENTIONAL option of the ERIS keyword calculates the three-center ERIs before the SCF procedure and, if possible, stores them in RAM. This so-called in-core method is fast as long as all integrals fit into the RAM. The available RAM size (as distinct from system RAM size) is set by deMon2k with the MAXRAM parameter (see Table 1) or the RAM option of the ERIS keyword. If the RAM space is not sufficient, deMon2k will write all ERIs to the scratch file `ioeri.scr`. As a result, the ERIs must be read from disk at each SCF step. For larger systems this disk I/O becomes the bottleneck of the calculation. Note that the printing of the ERIs in the `deMon.out` file enabled by PRINT ERIS (see 4.12.2) requires the ERIS CONVENTIONAL option. The same holds for PRINT DEBUG which includes PRINT ERIS.

With the RAM option of the ERIS keyword the allocatable RAM size per core can be defined in the deMon input file. This overrides the MAXRAM definition in the `parameter.h` file. Note that a RAM size definition larger than the available physical memory will result in large paging overhead during program execution.

Screening of the ERIs can be controlled with the TOL option. The threshold  $\tau$  is calculated as:

$$\tau = \frac{\text{TOL}}{\text{Number of Electrons}} \quad (4.13)$$

ERIs with an orbital overlap smaller than  $\tau$  are not calculated (screened). The threshold  $\tau$  also enters into the double asymptotic expansion radii for the MULTIPOLE method. The density screening threshold for the numerical integration of the exchange-correlation

potential is also given by  $\tau$ . The default settings for TOL are  $10^{-14}$  and  $10^{-10}$  a.u. for the CONVENTIONAL and DIRECT/MIXED/MULTIPOLE method, respectively. Beware that aggressive screening of the ERIs can harm SCF convergence.

#### 4.5.5 Keyword GUESS

This keyword specifies the SCF starting density.

Options:

**TB / CORE / FERMI / RESTART / ONLY / PROJECTED**

<b>TB</b>	A tight-binding SCF starting density is calculated. This is the default.
<b>CORE</b>	The SCF starting density is obtained from diagonalization of the core Hamiltonian.
<b>HBOMD</b>	The CORE start density is used in each BOMD step.
<b>FERMI</b>	The starting density is obtained by quenching a fractionally occupied SCF solution to integer occupation numbers.
<b>RESTART</b>	The starting density is read from the restart file deMon.rst.
<b>PROJECTED</b>	A projected starting density is requested.
<b>ONLY</b>	The program stops after the generation of the starting density.

Description:

A good choice of the starting density can be crucial for the SCF convergence. In most cases, the tight-binding starting density is the recommended choice. For metal clusters, the core starting density may sometimes be advantageous. In Born-Oppenheimer molecular dynamics (BOMD) simulations the converged SCF density of the previous MD step is used by default as start density for the next MD step. With the GUESS option HBOMD a core start density is requested for each BOMD step. This decouples the SCF solutions of the individual BOMD steps. Note that this choice will increase computational demand. For the definition of the core Hamiltonian recall Section 1.4.

If the molecular orbitals of the start density (for the printing of the MOs see Section 4.12.2) exhibit only a very small HOMO-LUMO gap, the FERMI option may be used to obtain a better starting density. It should be noted that this option implies an SCF calculation and, therefore, is much more time consuming than the other starting density options. If a FERMI starting density is used, the OMA option of the keyword MIXING (see 4.5.6) should not be applied! The starting density can also be read from the restart file deMon.rst of a previous run in form of MO coefficients. Note that the restart file always contains canonical MO

coefficients. With the keyword MOEXCHANGE (see 4.4.3), the molecular orbital ordering of the starting density can be altered. With the option PROJECTED MO coefficients from a different (usually smaller) basis set can be read from the restart file and used as start density. This can be a good strategy to accelerate SCF convergence for very large basis sets.

The option ONLY stops the program after the starting density is generated and written to the restart file. This option is recommended if a FERMI starting density was requested or the starting density needs to be altered after inspection. The starting density can be printed with the MAX=1 and MOS option of the PRINT keyword (4.12.2).

#### 4.5.6 Keyword MIXING

This keyword controls the charge density mixing.

Options:

- + <Real> Fixed charge density mixing parameter between 0 and 1.
- <Real> Dynamic charge density mixing parameter. The default is -0.3.
- OMA** The optimal mixing algorithm is activated.

Description:

With the charge density mixing parameter, Hartree damping [189] (also called Pratt or linear mixing) of the charge density fitting coefficients is performed. The smaller the mixing parameter, the larger is the damping. For some systems, a very small mixing parameter (0.1) may be necessary. If a dynamic mixing parameter is chosen, the parameter is reduced during the SCF iterations if the MinMax SCF error increases. With the OMA option, the optimal mixing algorithm from Cancés [190] is activated. This option is recommended if the dynamical mixing fails or is slow to converge. It should be noted that use of OMA adds considerable overhead and that the method exhibits bad convergence for the case of small HOMO-LUMO gaps.

#### 4.5.7 Keyword SHIFT

This keyword activates the level-shift procedure.

Options:

- + <Real> Fixed level-shift value [in a.u.].
- <Real> Dynamic level-shift value [in a.u.].

**MIN=<Real>** Minimum value for dynamical level-shift [in a.u.].

Description:

The level-shift procedure enables enlargement of the HOMO-LUMO gap during the SCF iterations [191]. This stabilizes the initial SCF configuration preferentially. Because of the small HOMO-LUMO gap common in DFT calculations, this procedure can be very valuable for improving SCF convergence. Beware, however, that with a fixed level-shift, incorrect convergence to a spurious excited state is possible. When using SHIFT therefore, it is imperative to check the converged orbital occupations with the MOS option of the PRINT keyword (see 4.12.2). The problem can be avoided if the dynamical level-shift procedure [155] is used. In that case, the shift value is adapted to the MinMax SCF error. For large errors, which usually occur at the beginning of the SCF procedure, the full shift value is used. As the error decreases, the shift value is also decreased (see Example 2.20 on page 27 of the tutorial). A small shift value may remain at the end of the SCF procedure. In fact a minimum shift value can be requested with the MIN option. Small remaining shift values at the end of the SCF are usually of no concern. The dynamical level-shift procedure has proven very valuable for the SCF convergence of transition-metal systems [192]. In the case of ROKS calculations, the level-shift procedure applies twice, once between the doubly and singly occupied orbitals and again between the singly-occupied and the unoccupied orbitals. In both cases, the specified level-shift value is used.

If the SHIFT keyword is used, the level-shift procedure is switched on for all calculation steps (optimization, frequencies, properties, etc.). In the case that the level-shift procedure is used only to prepare a start density, a single-point SCF must be performed with the SHIFT keyword and the resulting density then has to be read in as the restart density (see 4.5.5) in the subsequent calculation without the SHIFT keyword.

#### 4.5.8 Keyword DIIS

This keyword activates the DIIS procedure.

Options:

**ON / OFF**

**ON** The DIIS procedure is switched on. This is the default.

**OFF** The DIIS procedure is switched off.

**TOL=<Real>** The DIIS procedure is switched on after the SCF energy error is smaller than <Real>.

Description:

By default the DIIS (direct inversion in the iterative subspace) procedure [193] is switched on if the energy error is less than 0.01 a.u. in an SCF step. For most cases, this results in a considerable speed-up of the SCF convergence. However, for some systems this threshold is too large and a SCF convergence failure will occur. In such cases, initiation of the DIIS procedure should be manipulated with the TOL option. **If the HOMO-LUMO gap of a system is very small, DIIS can be counterproductive for the SCF convergence.** In those cases the DIIS procedure can be disabled with the OFF option. It should be noted that in deMon2k different DIIS algorithms are invoked for the AUXIS and BASIS options specified with the VXCTYPE keyword (see 4.2.1). Both algorithms are based on the energy gradient with respect to the charge-density fitting coefficients. In the case of the AUXIS option, only charge-density fitting coefficients are used in the DIIS step. Therefore, no extra I/O is necessary. In the case of the BASIS option, density matrices are also used in the DIIS step and, therefore, added I/O overhead may occur. The same holds for Fock exchange.

## 4.6 Optimization, Interpolation, and Transition State Search

### 4.6.1 Keyword OPTIMIZATION

This keyword controls the local geometry optimization and transition state search.

Options:

#### **REDUNDANT / INTERNAL / CARTESIAN**

**REDUNDANT** The optimization is performed in delocalized redundant internal coordinates. This is the default.

**INTERNAL** The optimization is performed in the defined internal coordinates (Z-matrix input).

**CARTESIAN** The optimization is performed in Cartesian coordinates.

**MAX=<Integer>** Maximum number of optimization steps. Default is 50.

**TOL=<Real>** Optimization convergence criterion for RMS gradient. Default is  $3 \cdot 10^{-4}$  atomic units.

**STEP=<Real>** Maximum step size in optimization. Default is 0.3 atomic units.

**TS** Activate local transition state search.

**MOD=<Integer>** Hessian eigenmode to be followed in the transition state search. Default is 1, the mode with the lowest (most negative) frequency.

Description:

By default, deMon2k uses delocalized internal redundant coordinates for the geometry optimization or transition state search [194,195]. For full geometry optimizations, this is also the recommended method. Because of the possible linear dependencies of the internal coordinates defined in the Z-matrix, optimization in these coordinates with OPTIMIZATION INTERNAL may become problematic even for small systems ( $\leq 100$  atoms). With the REDUNDANT option and a Cartesian input, deMon2k constructs a Z-matrix from the Cartesian input and a set of linear combinations of internal coordinates (delocalized coordinates) which avoids linear dependencies. This option should be used if an optimization gets stuck and has to be restarted. With the REDUNDANT option and an internal Z-matrix input, the optimization is also performed in delocalized internal redundant coordinates. However, now the user-defined internal coordinates of the Z-matrix are included in the delocalized internal redundant coordinates. At each optimization step these delocalized internal redundant

coordinates are transformed back to the coordinates of the Z-matrix input. Therefore the combination of a Z-matrix input and the option REDUNDANT can lead to different results than the combination of a Z-matrix input and the option INTERNAL. In particular, equivalent internal coordinates could be broken during a redundant optimization. To enforce coordinate equivalences the option INTERNAL must be used for the optimization. Even though the default optimization in redundant coordinates is usually most efficient some exceptions exist. For very tight optimizations ( $TOL \leq 10^{-6}$  a.u.) the iterative back transformation may hamper convergence of the optimization. Another problem can arise from the automatic Z-matrix construction for redundant coordinates. This construction is prone to linear dependencies in systems that consist of many individual molecules, e.g. solvent clusters. In such cases it is advisable to optimize in Cartesian coordinates by using OPTIMIZATION CARTESIAN. Table 11 shows the relationships between the types of input and types of optimization options. Each cell in that Table contains two values. The upper value shows the coordinate system used for an optimization step, while the lower value shows the coordinate system in the deMon.new file which contains the optimized geometry.

Table 11: Relationship between different options for the keywords GEOMETRY and OPTIMIZATION. Upper values denote the coordinate system for the optimization and lower values the geometry definition in the file deMon.new.

GEOMETRY option	OPTIMIZATION option		
	REDUNDANT	INTERNAL	CARTESIAN
Z-MATRIX	Redundant	Z-Matrix	Cartesian
	Z-Matrix	Z-Matrix	Cartesian
CARTESIAN	Redundant	<i>Impossible</i>	Cartesian
	Redundant		Cartesian
MIXED	Redundant	Redundant	Cartesian
	Redundant	Redundant	Cartesian

With the MAX and TOL options, the maximum number of optimization steps and the optimization convergence criterion can be specified. Convergence of the optimization is based on the remaining maximum and root mean square (RMS) forces. The TOL option specifies the RMS force convergence criterion. The other convergence criteria are calculated from

this value. Note that the SCF convergence criteria are also tightened according to Table 9. For very tight optimizations ( $\text{TOL} \leq 10^{-6}$  a.u.) it might also be necessary to use finer grids for the numerical integration of the exchange-correlation contributions (see 4.3.6). The maximum step length (in atomic units) used in the optimization can be specified with the STEP option.

The option TS activates a local transition state search based on the eigenvector-following of the Hessian matrix. The success of this approach depends crucially upon the starting structure and quality of the starting Hessian. Therefore, use of a calculated start Hessian (see 4.6.5) is recommended for the transition state search. By default, deMon2k employs an uphill trust region method [196,197] for the local transition state search. As an alternative (P)-RFO steps [198,199] can be chosen by selecting the RFO STEPTYPE (see 4.6.7). Both methods guarantee that the (right) Hessian structure (one negative eigenvalue) is preserved. To avoid reversion to a positive-definite Hessian, the Powell update [200] is used by default for the transition state search.

If a mode other than the lowest Hessian eigenmode is to be followed, the option MOD is used to select the desired eigenvector (use the PRINT keyword 4.12.2 with the DE2 option to print the Hessian eigenvectors). On subsequent steps, this mode is selected by the largest overlap with the eigenvector followed in the previous cycle.

#### 4.6.2 Keyword SADDLE

This keyword invokes a saddle-point interpolation of a transition state neighborhood. This keyword can only be used in combination with the keywords REACTANT and PRODUCT (see 4.1.2). They define the reactant and product structures for the saddle-point interpolation.

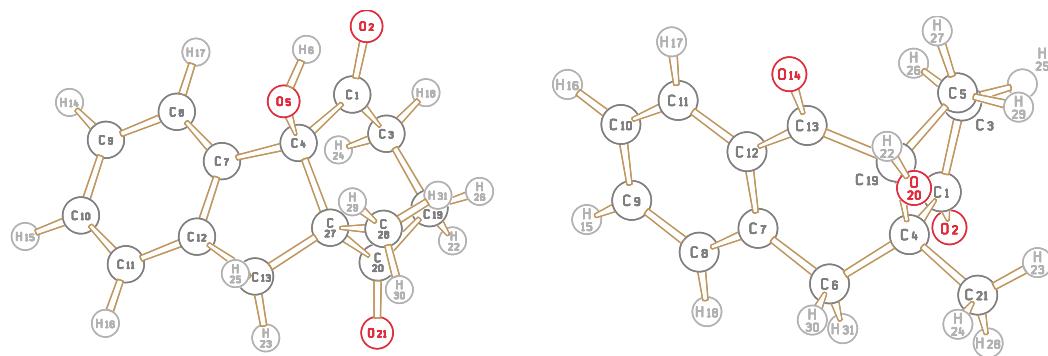
Options:

<b>TS</b>	A local transition state search is performed directly after the successful saddle-point interpolation.
<b>INT=&lt;Real&gt;</b>	Reduction of interpolation distance in each saddle step. By default a 5% reduction (INT=0.05) is used.
<b>MAX=&lt;Integer&gt;</b>	Maximum number of saddle-point interpolation steps. Default is 100.
<b>TOL=&lt;Real&gt;</b>	Interpolation distance tolerance [a.u.] for saddle-point convergence. Default is 0.1 atomic units.

Description:

The saddle-point interpolation is used to find starting structures for a local transition state search. This is particularly useful if chemical intuition fails to provide such structures. In any case, the exact transition state structure must be optimized by a local transition state search as described in 4.6.1. **The geometry from a saddle-point interpolation is only a starting structure!** The subsequent local transition state search can be requested directly with the TS option of the SADDLE keyword. In that case, the saddle-point interpolation and local transition state search are performed in one run. For simple transition states this so-called hierarchical transition state search [31] is most convenient. For more complicated transition states it is advisable that the output of the saddle-point interpolation be analyzed first. Because the saddle-point interpolation is based on a modified restricted step method [31], transition state starting structures are always found. However, it is not guaranteed that the interpolation follows only one minimum energy path. If switches occur, the transition state starting structure may lead to another reaction coordinate than the one that connects the input structures given by the REACTANT and PRODUCT keywords. Thus, it is always advisable to calculate the intrinsic reaction coordinate (see 4.6.4) after a transition state structure is optimized. To avoid switching of reaction coordinates a proper alignment of reactant and product structures is important. By default, the alignment of these structures is performed before the saddle interpolation. Note that numbering of atoms enforces alignment of these atoms to each other in the reactant and product structure. This can further stabilize a selected reaction coordinate. Figure 10 shows MOLDEN (see 4.10.1) outputs with atom numbering of reactant and product for a concerted intramolecular rearrangement reaction.

Figure 10: Atom numbering for saddle alignment to follow a concerted intramolecular rearrangement reaction coordinate.



The relevant atoms for the concerted rearrangement are C(4), O(5), H(6), C(1) and O(2) in the reactant and C(13), O(14), C(19), O(20) and H(22) in the product structure. To ensure that the alignment algorithm maps these atoms correctly onto each other such that the saddle interpolation follows the concerted reaction path, they are numbered accordingly in the structure definition of the reactant and product.

```

VXCTYPE PBE
SCFTYPE MAX=500
SADDLE INT=2.0 MAX=200
#
# Optimized reactant coordinates
#
REACTANT CARTESIAN ANGSTROM
C1 -1.146304 1.580047 0.304830 6 12.011000
O1 -1.474470 2.704811 -0.094272 8 15.999400
C -1.478037 1.068974 1.680762 6 12.011000
C2 -0.388141 0.637819 -0.631850 6 12.011000
O2 -0.381260 1.189947 -1.947188 8 15.999400
H1 -0.758490 2.097150 -1.833935 1 1.007940
C 1.048313 0.421392 -0.145352 6 12.011000
C 1.939655 1.366263 0.372306 6 12.011000
C 3.241408 0.952931 0.686023 6 12.011000
C 3.637166 -0.375842 0.466195 6 12.011000
C 2.745608 -1.314937 -0.071382 6 12.011000
C 1.442073 -0.907573 -0.371235 6 12.011000
C 0.308225 -1.703520 -0.973142 6 12.011000
H 3.960249 1.673741 1.103430 1 1.007940
H 4.662742 -0.682994 0.718950 1 1.007940
H 3.070566 -2.354154 -0.245424 1 1.007940
H 1.634675 2.414219 0.528049 1 1.007940
H -2.066621 1.836068 2.219392 1 1.007940
C -2.257501 -0.257740 1.553539 6 12.011000
C -1.511818 -1.286833 0.722868 6 12.011000
O -1.427834 -2.455869 1.104029 8 15.999400
H -2.469807 -0.701939 2.544632 1 1.007940
H 0.203258 -2.717154 -0.544618 1 1.007940
H -0.531868 0.883086 2.232117 1 1.007940
H 0.440534 -1.793379 -2.070772 1 1.007940
H -3.231445 -0.055345 1.053604 1 1.007940
C -0.945955 -0.847999 -0.649669 6 12.011000
C -2.078732 -1.028854 -1.680712 6 12.011000
H -1.724896 -0.679675 -2.668085 1 1.007940
H -2.366493 -2.098307 -1.735541 1 1.007940
H -2.979948 -0.434441 -1.415635 1 1.007940
#
# Optimized product coordinates
#
PRODUCT CARTESIAN ANGSTROM
C 1.364437 1.346285 0.835797 6 12.011000
O 1.215160 2.555667 0.991251 8 15.999400
C 1.651005 0.332453 1.938566 6 12.011000
C 1.302791 0.639095 -0.533580 6 12.011000
C 2.023969 -0.959913 1.201220 6 12.011000
C 0.050517 1.093638 -1.304525 6 12.011000
C -1.228417 0.650885 -0.640377 6 12.011000
C -2.383053 1.445681 -0.648124 6 12.011000
C -3.566686 0.978715 -0.067793 6 12.011000
C -3.629589 -0.297452 0.514493 6 12.011000

```

C	-2.497068	-1.106654	0.514108	6	12.011000
C	-1.294672	-0.632954	-0.049150	6	12.011000
C2	-0.102654	-1.498701	-0.041826	6	12.011000
O2	-0.177817	-2.735059	0.077486	8	15.999400
H	-4.458601	1.622475	-0.067957	1	1.007940
H	-4.568591	-0.652453	0.965774	1	1.007940
H	-2.504895	-2.117874	0.948295	1	1.007940
H	-2.346458	2.449322	-1.103029	1	1.007940
C1	1.307134	-0.877664	-0.186071	6	12.011000
O1	2.007450	-1.631628	-1.182543	8	15.999400
C	2.569593	1.040706	-1.327857	6	12.011000
H1	1.710225	-2.561139	-1.054706	1	1.007940
H	3.500977	0.773070	-0.785542	1	1.007940
H	2.580433	0.504636	-2.295926	1	1.007940
H	2.425802	0.713660	2.629881	1	1.007940
H	0.714441	0.225449	2.526799	1	1.007940
H	1.738154	-1.883766	1.738902	1	1.007940
H	2.558691	2.136998	-1.488046	1	1.007940
H	3.112220	-0.998838	0.991781	1	1.007940
H	0.097062	0.665875	-2.331098	1	1.007940
H	0.069499	2.196290	-1.397068	1	1.007940

This input ensures that the first atom of the reactant, C(1) in Figure 10, which is named C1 in the reactant structure definition for the saddle interpolation is aligned to the nineteenth carbon atom of the product, C(19) in Figure 10, which is also named C1, now, however, in the product structure. The same holds for the other numbered atoms in the above SADDLE input. Also note that the interpolation distance reduction is reduced to 2% by setting INT=2.0 in the SADDLE keyword line. This stabilizes further the saddle interpolation along the desired reaction coordinate. Of course, this reduction induces an increase in the number of saddle-point interpolation steps and, thus, an increase in the computational effort. The MAX and TOL options can be used to change the maximum number of saddle interpolation steps and the tolerance for the interpolation convergence.

#### 4.6.3 Keyword SCAN

This keyword is used to perform a scan of the potential energy surface along a Cartesian or internal coordinate.

Options:

##### ADIABATIC / VERTICAL

- |                  |  |
|------------------|--|
| <b>ADIABATIC</b> | The start structure for geometry optimization is given by the optimized structure of the previous scan step. |
| <b>VERTICAL</b>  | The start structure for geometry optimization is given by the initial structure.                             |

**END=<Real>** End value for scanning coordinate. This specification is mandatory.

**STEP=<Integer>** Number of scan steps. This specification is mandatory.

Description:

Note that SCAN requires both END and STEP to be specified, even though they have the syntax of options. If a prospective reaction coordinate has been identified, a potential energy surface scan ("linear transit") can be used as a computationally efficient approach for the localization of transition-state starting structures. The options ADIABATIC and VERTICAL of the SCAN keyword are used to select starting structures for a geometry optimization (if requested) along the scanning coordinate. If the keyword OPTIMIZATION (see 4.6.1) is specified, all geometric variables except the scan coordinate are optimized. In the following example, the dihedral angle DX1 is successively reduced from 180.0 to 0.0 degrees in 9 steps. At each step, the variables R1 and R2 are optimized. Their starting values are taken from the previous scan step.

```
OPTIMIZATION
SCAN DX1 END=0.0 STEP=9
#
GEOMETRY Z-MATRIX
C
N 1 R1
X 2 RCON 1 ACON
H 2 R2 3 A1 1 DX1
#
VARIABLES
R1 1.18
R2 1.01
DX1 180.0
#
CONSTANTS
RCON 1.0
ACON 90.0
A1 90.0
```

If the keyword OPTIMIZATION had been omitted in the example, the variables R1 and R2 would be kept constant during the potential energy scan. If a Cartesian coordinate is scanned the syntax of the SCAN keyword changes to:

```
SCAN H4 END=(0.9,-0.4,-0.3) STEP=10
```

Here H4 denotes the atom that is moved during the potential energy scan. The END option now defines the final Cartesian coordinates of this atom. These coordinates are kept constant during the optimizations in the individual scan steps. Example 2.21 on page 31 of the tutorial discusses the input and output of a Cartesian potential energy surface scan in more detail.

#### 4.6.4 Keyword IRC

This keyword activates the intrinsic reaction coordinate (IRC) calculation. The keyword HESSIAN (see 4.6.5) is mandatory for this keyword.

Options:

##### **FORWARD / REVERSE**

- |                |  |
|----------------|--|
| <b>FORWARD</b> | The IRC is calculated in the forward direction of the transition state eigenvector of the Hessian. |
| <b>REVERSE</b> | The IRC is calculated in the reverse direction of the transition state eigenvector of the Hessian. |

##### **MASS / NOMASS**

- |               |   |
|---------------|---|
| <b>MASS</b>   | The IRC is calculated in mass-weighted coordinates.       |
| <b>NOMASS</b> | The IRC is calculated without mass-weighting coordinates. |

- |                             |  |
|-----------------------------|--|
| <b>RESTART</b>              | The IRC calculation is restarted. Previous steps are read from the file deMon.trj.                       |
| <b>EXTEND</b>               | An already-calculated IRC branch is extended.  |
| <b>MAX=&lt;Integer&gt;</b>  | Maximum number of IRC steps. Default is 50.  |
| <b>STEP=&lt;Real&gt;</b>    | Maximum radius size for IRC optimization. Default is $0.1 \text{ (amu}^{1/2}\text{)} \cdot \text{\AA}$ . |
| <b>TOL=&lt;Real&gt;</b>     | Convergence criterion for IRC end-point gradients. Default is $3 \times 10^{-4}$ atomic units.           |
| <b>EVIB=&lt;Real&gt;</b>    | Energy threshold for transition state kick-off. Default is $10^{-3}$ Hartree.                            |
| <b>READ=&lt;Integer&gt;</b> | Number of IRC pivot points for which coordinates are read from the file deMon.trj.                       |

Description:

Calculation of the intrinsic reaction coordinate (IRC) yields the connectivity among reactant, transition state, and product on the potential energy surface [201]. Only with this connectivity may a reaction be broken down into its elementary steps. IRC calculations are also useful to discover unexpected intermediates that are often overlooked in the transition state search. In deMon2k, IRC calculations are performed with the Gonzalez-Schlegel algorithm [202,203]. Each IRC step possesses two stages. First is the generation of a pivot point. Second is constrained optimization on a hypersphere around this point. With the option STEP, the (maximum) radius of this hypersphere is defined. The IRC path is divided into two branches, one for the forward direction of the transition state eigenvector (which is the

eigenvector that corresponds to the imaginary frequency of the transition state), and the other for the reverse direction of this eigenvector [204]. Thus, the IRC calculation always starts at the transition state structure and needs the Hessian matrix for this structure. **As a result the keyword HESSIAN is mandatory for an IRC input.** It is convenient to use the restart Hessian from the frequency analysis of the optimized transition state structure. To do so the `deMon.mem` file from the transition state frequency analysis must be copied to the `deMon.rst` file for the IRC calculation and HESSIAN RESTART must be specified in the IRC input file. Note that the structure definition for such an IRC calculation must coincide with the one in the `deMon.new` file from the transition state frequency analysis. The IRC keyword options FORWARD and REVERSE are used to specify IRC calculations along the forward and reverse direction of the Hessian eigenvector that corresponds to the imaginary frequency of the transition state (please note that eigenvector directions may change on account of variations in diagonalization routines and computer architectures). The two branches can be merged with the option EXTEND. Assuming that an IRC FORWARD calculation was already performed (see example 2.22 on page 31 of the tutorial), the following input extends the IRC path with the REVERSE branch:

```
IRC REVERSE EXTEND
```

As a result, the trajectory and molden files `deMon.trj` and `deMon.mol` contain both branches and, thus, the full reaction coordinate. Coordinates of a specified pivot point along the IRC path can be loaded with the READ option of the IRC keyword. In this way, properties or molecular fields can be calculated or plotted along the IRC path.

By default, the IRC path is calculated in mass-weighted coordinates. The mass-weighting is disabled with the option NOMASS. With the MAX option of the IRC keyword, the maximum number of IRC steps can be specified. The RESTART option permits the restart of an IRC calculation, e.g. if not enough IRC steps were requested in a previous run. The IRC calculation restarts with the last successfully calculated pivot point. As always, the restart input can be found in the `deMon.new` file. For the IRC restart, the file `deMon.mem`, must be renamed to `deMon.rst`, and the file `deMon.trj` must also be available. The TOL option specifies the convergence criterion for the RMS gradient at the IRC path end points. The convergence of the constrained optimizations in each IRC step can be modified with the OPTIMIZATION keyword. Cartesian constants can be used in an IRC calculation. With the EVIB option, the energy decrease for the transition state kick-off (that is, the energy decrement to initiate the search) along the transition state eigenvector is specified. This energy value is given in Hartree.

#### 4.6.5 Keyword HESSIAN

This keyword specifies the choice of the starting Hessian for geometry optimization or transition state search.

Options:

**BAKER / FISHER / UNITY / INTERNAL / LINDH / CALCULATE / RESTART / READ**

<b>BAKER</b>	Starting Hessian constructed from Baker force weights. This is the default for the internal optimization.
<b>FISHER</b>	Starting Hessian constructed from Fisher force weights. This is the default for the redundant optimization.
<b>UNITY</b>	The unit matrix is used as starting Hessian. This is the default for the Cartesian optimization.
<b>INTERNAL</b>	Starting Hessian from the unit matrix of an internal Hessian constructed from an auxiliary Z-matrix.
<b>LINDH</b>	Starting Hessian constructed from Lindh force weights.
<b>CALCULATE</b>	Starting Hessian is calculated from numerical second derivatives.
<b>RESTART</b>	Starting Hessian is read from the restart file deMon.rst.
<b>READ</b>	Starting Hessian is read from the input file.

**PLAIN / DEMON**

<b>PLAIN</b>	The Hessian is read from the input file in free format row by row. Only valid in combination with the READ option.
<b>DEMON</b>	The Hessian is read from the input file in the deMon2k output format. Only valid in combination with the READ option.
<b>VIB=&lt;Real&gt;</b>	Scaling factor for the numerical step size during the calculation of second derivatives. Only valid in combination with the CALCULATE option.

Description:

With the options BAKER, FISHER, and LINDH, a starting Hessian in primitive redundant coordinates is generated, then transformed into the coordinate system used for the optimization. For the BAKER option, the diagonal elements for the primitive bonds, bends, and dihedral angles are set to 0.5, 0.2 and 0.1, respectively [194]. For the FISHER option, the diagonal elements are calculated according to the empirical rules of Fisher et al. [205]. This is the default for the starting Hessian of the default redundant optimization. With the

option LINDH, the diagonal elements are calculated according to the empirical rules of Lindh et al. [206]. For a Cartesian optimization, a unit Cartesian starting Hessian is the default option. If other options are chosen for Cartesian optimizations, an auxiliary Z-matrix is constructed from the Cartesian input, and either a unit (option INTERNAL) or an empirical (options BAKER, FISHER, or LINDH) internal Hessian matrix (options BAKER, FISHER, or LINDH) is calculated. Then, this internal Hessian is transformed into the Cartesian starting Hessian.

A unit matrix starting Hessian in the space of the primitive redundant coordinates (for redundant optimizations), in the space of the internal Z-matrix coordinates (for internal optimizations), or in the space of the Cartesian coordinates is specified with the option UNITY. The option CALCULATE triggers calculation of the starting Hessian from numerical second derivatives. This is the most accurate but also the most time-consuming method. It is recommended for the transition state search. With the options RESTART and READ, the starting Hessian is read from the restart or input file respectively. The (updated) Hessian matrix is written to the restart file at the end of the optimization and also after it has been calculated using the HESSIAN or FREQUENCY keyword (see 4.8.1). With the RESTART option, that Hessian can be recovered. If the READ option is specified, the elements of the Hessian matrix must be given either in free format as real values (see example 2.23 on page 32 of the tutorial) or in the deMon2k output format (see example 2.24 on page 33 of the tutorial) in the keyword body of the HESSIAN keyword. This way the reading of the binary restart file can be bypassed. The Hessian input format is selected by the PLAIN or DEMON option. The option VIB is valid only if the option CALCULATE is also specified. For more information on the VIB option, see the keyword FREQUENCY (4.8.1).

#### 4.6.6 Keyword UPDATE

This keyword controls the Hessian update in quasi-Newton methods during geometry optimization or transition state search.

Options:

**BFGS / POWELL / BERNY / DFP / MSP / SR1 / EXACT**

<b>BFGS</b>	Broyden, Fletcher, Goldfarb, Shanno (BFGS) update [207–210]. This is the default for the geometry optimization.
<b>POWELL</b>	Powell update [200]. This is the default for the transition state search.
<b>BERNY</b>	Berny update from H.B. Schlegel [211].

---

<b>DFP</b>	Davidon, Fletcher, Powell update [212,213].
<b>MSP</b>	Murtagh-Sargent-Powell update [197].
<b>SR1</b>	Symmetric rank-one update [214].
<b>EXACT</b>	No update is performed, the Hessian matrix is always calculated anew.
<b>VIB=&lt;Real&gt;</b>	Scaling factor for the numerical step size during the calculation of second derivatives. Only valid in combination with the EXACT option.

Description:

For the optimization in redundant and Cartesian coordinates, the BFGS update is recommended. For optimization in internal coordinates, the BERNY update may be preferred. For the transition state search, only the POWELL or MSP updates are suitable. The option VIB is valid only if the option EXACT is also specified. For more information on this option see the keyword FREQUENCY (4.8.1).

#### 4.6.7 Keyword STEPTYPE

With this keyword the geometry optimization step is selected.

Options:

##### LEVENBERG / RFO / WALK / DESCENT

<b>LEVENBERG</b>	Levenberg-Marquardt.
<b>RFO</b>	Rational function optimization.
<b>WALK</b>	Potential energy surface walking.
<b>DESCENT</b>	Steepest descent.

Description:

By default deMon2k uses Levenberg-Marquardt steps [215,216] with robust root finding [217] in the structure optimization and the transition state search by the uphill trust region method [196,197]. Alternatively, RFO [198] and WALK [218] steps are available for restricted step method minimizations [219]. For the transition state search the (P)-RFO step can be selected as alternative to the default Levenberg-Marquardt step by specifying the RFO option.

## 4.7 Born-Oppenheimer Molecular Dynamics

### 4.7.1 Keyword DYNAMICS

This keyword activates Born-Oppenheimer molecular dynamics (BOMD) simulations.

Options:

**STEP=<Real>** The time length of a BOMD step in fs. The default value is 1 fs.

**MAX=<Integer>** The number of BOMD steps in the simulation. This value **must** be specified.

**INT=<Integer>** Step interval to update deMon.out file. The default value is 10.

**R=<Real>** The radius of an Anderson containment sphere.

Description:

Note that DYNAMICS requires MAX to be specified, even though it has the syntax of an option. A BOMD step in deMon2k consists of solving the time-independent electronic Schrödinger equation to get the forces on the nuclei, followed by a classical velocity Verlet [220,221] propagation of the nuclei according to those forces. For each BOMD step, the step time, molecular energies, atomic coordinates, and velocities are written to the trajectory file deMon.trj. The BOMD time step and the number of those steps are defined by the options STEP and MAX, respectively. The INT option controls the updating of the deMon.out and deMon.mol files. It specifies the step interval after which data are written to those two files. By default, the deMon.out file records the time and the instantaneous temperature as well as the instantaneous kinetic, potential, and total energies. It also records averages of temperature and (total) energy. The deMon.mol file records only energies (see 4.7.3 for energy selection) and atomic coordinates for visualization of the BOMD propagation with MOLDEN (see 4.10.1). Please note that the number of propagation steps in the animation is limited by your specific MOLDEN implementation. The R option of the DYNAMICS keyword activates an Anderson containment sphere [221]. It bounces particles back into its center under energy and (linear) momentum conservation. The radius of the containment sphere is defined in Ångström or Bohr according to the unit setting in the GEOMETRY keyword (see 4.1.1).

### 4.7.2 Keyword TRAJECTORY

This keyword controls BOMD trajectory operations. It also triggers additional output to the deMon.trj file.

Options:

<b>RESTART</b>	A BOMD trajectory is restarted.
<b>READ=&lt;Integer&gt;</b>	Read specified trajectory step from deMon.trj file.
<b>PART=&lt;Integer1&gt;-&lt;Integer2&gt;</b>	The trajectory part from step <Integer1> through <Integer2> is loaded from the deMon.trj file.
<b>CUT&gt; &lt;Integer&gt;</b>	Trajectory steps above <Integer> are cut and removed from the deMon.trj file.
<b>CUT&lt; &lt;Integer&gt;</b>	Trajectory steps below <Integer> are cut and removed from the deMon.trj file.
<b>PLOT=&lt;Integer1&gt;-&lt;Integer2&gt;</b>	A deMon.mol file for the trajectory part from step <Integer1> through <Integer2> is created.
<b>INT=&lt;Integer&gt;</b>	Interval for trajectory plotting with the PLOT option.
<b>FORCES</b>	Atomic forces are added to the trajectory file deMon.trj.
<b>MOS=&lt;Integer1&gt;-&lt;Integer2&gt;</b>	MO <Integer1> through MO <Integer2> energies are added to the trajectory file deMon.trj.

Description:

The data of deMon2k BOMD runs are stored in ASCII format in the file deMon.trj. This permits transferability between different computational architectures. A BOMD run can be restarted with the option RESTART of the TRAJECTORY keyword in combination with the DYNAMICS keyword (see 4.7.1). In this case the old trajectory is read from the deMon.trj file and augmented during the run. As with all deMon2k restarts, a restart input for a BOMD run is written into the deMon.new file. Copying this file to deMon.inp is an easy way to set up a BOMD restart. The restart geometry is read from the input file whereas the restart velocities are read from the corresponding trajectory file deMon.trj. A BOMD restart has the following minimal syntax:

```
TRAJECTORY RESTART
DYNAMICS MAX=10000
```

This can be augmented by other BOMD relevant keywords such as BATH (see 4.7.6), LP-CONSERVE (see 4.7.4), etc. With the READ option of the TRAJECTORY keyword specific trajectory steps can be read from the deMon.trj file. In the following example input, trajectory step 61 is read from deMon.trj and electrostatic moments (see 4.8.3 for the keyword DIPOLE) are calculated for this geometry.

```
DIPOLE
TRAJECTORY READ=61
#
GEOMETRY CARTESIAN ANGSTROM
NA      .000000      .000000     -.666740
LI      .000000      .000000     2.208355
```

Please note that the specified input geometry is overwritten by the geometry read from the trajectory file. The corresponding output notes this in the geometry specification.

\*\*\* GEOMETRY \*\*\*

TRAJECTORY STEP 61 ORIENTATION IN ANGSTROM

NO.	ATOM	X	Y	Z	Z-ATOM	MASS
1	NA	.000000	.000000	-.701224	11	22.990
2	LI	.000000	.000000	2.322574	3	6.941

Instead of reading only one step from the trajectory, the option PART for the TRAJECTORY keyword reads step sequences. Typically this option is used in combination with the SIMULATION keyword (see 4.7.3) for a trajectory analysis or the calculation of properties along the BOMD trajectory. With the CUT option trajectory parts, such as equilibration steps, can be removed from the deMon.trj file. The PLOT option has the same syntax as the PART option. However, it does not load trajectory steps but instead generates MOLDEN plot outputs in the file deMon.mol for the specified trajectory part. The step interval between the trajectory snapshots can be defined with the INT option. The following input example will generate 11 MOLDEN snapshots between the trajectory steps 30 and 80 with a 5-step interval.

TRAJECTORY PLOT=30-80 INT=5

Additional trajectory outputs are triggered by the options FORCES and MOS. The option FORCES causes Cartesian atomic forces for each BOMD step to be written to the trajectory file. With the option MOS, the molecular orbital energies of the specified MO range are written to the deMon.trj file. Please note that trajectory files are usually demanding in terms of storage.

In order to calculate thermodynamic partition functions it is convenient to join together several trajectory files by the multiple histogram method (MHM) [222,223]. This permits the generation of nanosecond statistics with ADFT BOMD simulations, e.g. for the study of melting transitions in finite systems [224]. The input syntax of the external MHM program [225] and its output is described in Appendix C.

### 4.7.3 Keyword **SIMULATION**

This keyword controls property evaluations along BOMD trajectories.

Options:

**ANALYZE / CALCULATE** (must be specified)

**ANALYZE** A BOMD (property) calculation is analyzed.

**CALCULATE** A property calculation along a BOMD trajectory is requested.

### **DIPOLE / POLARIZABILITY / NMR / MAGNETIZABILITY**

**DIPOLE** Dipole moments are calculated along a BOMD trajectory.

**POLARIZABILITY** Polarizabilities are calculated along a BOMD trajectory.

**NMR=<Integer>** NMR shieldings are calculated along a BOMD trajectory. The <Integer> denotes the atom for which the magnetic shielding data are printed.

**MAGNETIZABILITY** Magnetizabilities are calculated along a BOMD trajectory.

### **MOMENTA / PHASESPACE**

**MOMENTA** Linear and angular molecular momenta are calculated along an MD trajectory.

**PHASESPACE** The reduced coordinate and momentum are calculated along an MD trajectory. This option is valid only for dimers.

### **RDF / SIMILARITY / LINDEMANN / MEANDIS / PROLATE**

**RDF=A1-A2** Calculate the radial distribution function between atom A1 and atom A2.

**SIMILARITY** Calculate the similarity index along an MD trajectory with respect to pattern geometries.

**LINDEMANN** Calculate the Lindemann parameter for an MD trajectory.

**MEANDIS** Calculate the mean interatomic distance along an MD trajectory.

**PROLATE** Calculate the prolate deformation parameter along an MD trajectory.

### **LENGTH / ANGLE / DIHEDRAL**

---

<b>LENGTH</b>	Calculate the specified bond lengths along an MD trajectory.
<b>ANGLE</b>	Calculate the specified angles along an MD trajectory.
<b>DIHEDRAL</b>	Calculate the specified dihedral angles along an MD trajectory.

**E=STANDARD / E=KINETIC / E=POTENTIAL / E=TOTAL / E=SYSTEM**

<b>E=STANDARD</b>	Standard output for the energy analysis of a trajectory file.
<b>E=KINETIC</b>	The kinetic energy is averaged in the energy analysis of a trajectory file.
<b>E=POTENTIAL</b>	The potential energy is averaged in the energy analysis of a trajectory file.
<b>E=TOTAL</b>	The total energy is averaged in the energy analysis of a trajectory file.
<b>E=SYSTEM</b>	The system energy is averaged in the energy analysis of a trajectory file.

**INT=<Integer>** Step interval for property analysis or calculation. Default is 1.

## Description:

In deMon2k, molecular dynamics property calculations proceed in two steps. First, a trajectory file, `deMon.trj`, is created by a BOMD run as described in 4.7.1. In the second step, properties are analyzed or calculated along the stored trajectory coordinates according to the options ANALYZE or CALCULATE of the keyword SIMULATION. It is important to note that the BOMD run and the property calculation are completely independent. Thus, different basis sets, functionals etc. can be used in the property calculation. When properties are calculated, the trajectory file `deMon.trj` is modified in order to store property values. Thus, it is important to note that the geometries and property data in the `deMon.trj` file are results of **two** independent calculations. Once the property values are stored in the trajectory file, they can be analyzed with the ANALYZE option. The SIMULATION options DIPOLE, POLARIZABILITY, NMR and MAGNETIZABILITY specify the property to be analyzed or calculated. A BOMD dipole calculation is triggered by the following input line:

SIMULATION CALCULATE DIPOLE

The corresponding output has the form:

TIME [FS]	X[A.U.]	Y[A.U.]	Z[A.U.]	MU  [A.U.]	<MU> [A.U.]	<MU^2> [A.U.]
101.0	.089	.000	.835	.840	.840	.000
102.0	.092	.000	.839	.844	.842	.002

103.0	.093	.000	.842	.848	.844	.003
104.0	.091	.000	.846	.851	.846	.004
105.0	.088	.000	.850	.854	.847	.005
106.0	.083	.000	.852	.856	.849	.006
107.0	.076	.000	.855	.858	.850	.006
108.0	.068	.000	.871	.874	.853	.010
109.0	.060	.000	.864	.866	.855	.010
110.0	.051	.000	.868	.870	.856	.010

For each time step, the instantaneous dipole moment components and the corresponding absolute dipole moment are listed, together with the average value and the standard deviation. The BOMD polarizability [33] output is similar and discussed in more detail in example 2.25 on page 34 of the tutorial.

A BOMD property calculation can be specified further by the corresponding property keyword. Note that NMR shieldings along BOMD trajectories are only printed for the atom specified by the NMR option of the SIMULATION keyword, even though all shieldings are calculated by default. To obtain the NMR shielding for a specific atom use SIMULATION ANALYZE NMR=<Integer>, where <Integer> denotes the number of the desired atom in the GEOMETRY definition. The following input example,

```
TRAJECTORY RESTART PART=10-100 INT=1
SIMULATION ANALYZE NMR=3 INT=1
GEOMETRY CARTESIAN
O    0.0000    0.0000    0.1173  17.0
H    0.0000    0.7572   -0.4692
H    0.0000   -0.7572   -0.4692
```

generates an NMR shielding output for atom 3, here the second hydrogen of the water molecule, from trajectory step 10 to 100. Note the TRAJECTORY RESTART option that is used to address only a part of the trajectory file. Besides the detailed NMR shielding information (tensor diagonal elements, instantaneous shielding and averaged shielding) for the specified atom, the corresponding output lists also the NMR shielding statistics for all atoms.

```
*****
*** MOLECULAR DYNAMICS TRAJECTORY ANALYSIS ***
*****
```

TIME [FS]	XX	YY	ZZ	SIGMA	<SIGMA>
10.0	24.73	40.58	31.27	32.19	32.19
11.0	24.01	39.42	29.88	31.11	31.65
12.0	23.23	37.98	28.70	29.97	31.09
13.0	22.70	36.99	28.05	29.25	30.63
14.0	22.51	36.65	27.96	29.04	30.31
15.0	22.70	37.04	28.41	29.38	30.16

16.0	23.27	38.09	29.36	30.24	30.17
17.0	24.03	39.36	30.52	31.31	30.31
18.0	24.52	40.04	31.28	31.95	30.49
19.0	24.36	39.48	31.18	31.68	30.61
20.0	23.70	38.01	30.44	30.72	30.62
:	:	:	:	:	:
90.0	23.49	37.26	30.88	30.54	30.43
91.0	23.05	36.30	30.36	29.90	30.43
92.0	22.85	35.89	30.14	29.63	30.42
93.0	22.94	36.10	30.29	29.78	30.41
94.0	23.39	37.12	30.86	30.46	30.41
95.0	24.27	39.07	31.74	31.69	30.42
96.0	25.00	40.90	32.05	32.65	30.45
97.0	24.71	40.92	30.93	32.19	30.47
98.0	23.72	39.41	29.20	30.77	30.47
99.0	23.02	38.12	28.13	29.76	30.46
100.0	22.75	37.57	27.79	29.37	30.45

\*\*\* NMR SHIELDINGS STATISTIC \*\*\*

CHEMICAL SHIELDING FOR ATOM 1 ( 0 )  
AVERAGE SHIELDING [ppm] = 332.87  
STD DEV SHIELDING [ppm] = 13.06

CHEMICAL SHIELDING FOR ATOM 2 ( H )  
AVERAGE SHIELDING [ppm] = 30.40  
STD DEV SHIELDING [ppm] = 0.81

CHEMICAL SHIELDING FOR ATOM 3 ( H )  
AVERAGE SHIELDING [ppm] = 30.45  
STD DEV SHIELDING [ppm] = 1.02

The atoms for which NMR shieldings should be calculated along a BOMD trajectory can be selected with the READ option of the NMR keyword, see 4.8.8. Also the nuclear spin-rotation constant [226] can be calculated along BOMD trajectories as shown by the following input example.

```

DISPERSION
BASIS (AUG-CC-PVDZ)
AUXIS (GEN-A2)
SCFTYPE MAX=1000 TOL=0.100E-04
VXCTYPE OPTX-PBE AUXIS
TRAJECTORY RESTART PART=10000-12000 INT=20
SHIFT -0.2
NMR SPINROT
SIMULATION CALCULATE NMR=1 INT=20
#
# Cartesian coordinates of MD step 410000
#
GEOMETRY CARTESIAN ANGSTROM
F   -66.916797   19.630729  -13.088069    9  18.998000
H   -67.101518   18.720636  -12.997091    1  2.014000
C    54.166066  -15.936748   9.958016    6 12.0000000
C    53.957601  -15.619981  11.173715    6 12.0000000
H    54.766661  -16.089846   9.024499    1  1.008000
H    53.312450  -15.623649  12.049412    1  1.008000

```

Magnetizabilities [228] are calculated along a BOMD trajectory by the following input line:

## SIMULATION CALCULATE MAGNETIZABILITY

Similarly to the spin-rotation constant, the rotational g-tensor [229] can be calculated along BOMD trajectories by specifying MAGNETIZABILITY GTENSOR in the the input.

The options MOMENTA and PHASESPACE are generic to MD simulations and, therefore, have no corresponding property keyword. While linear and angular momentum analyses are general, the PHASESPACE analysis can be performed only for diatomic molecules. In general, MOMENTA and PHASESPACE results probe the MD sampling in detail and are recommended if the calculation of thermodynamic properties from the MD trajectory is the objective. Plots from PHASESPACE outputs are shown in Figure 12.

The RDF, SIMILARITY, LINDEMANN, MEANDIS and PROLATE options are intended to be used in combination with the ANALYSE option to perform the respective analysis on an MD trajectory. The RDF option enables a radial distribution function calculation. It requires two string arguments to specify the atomic pair defining the radial distribution function. According to this specification the RDF for a specific atom pair or for all pairs of an element combination are calculated. Specific RDF options can be set by the RDF keyword (see 4.7.7) as in the following example.

```
RDF MAX=2.0 WIDTH=0.01
SIMULATION ANALYSE RDF=C-H INT=10
#
GEOMETRY
C      1.251600    -0.743355    -0.192056
O      2.419080     0.016064     0.186570
C     -0.012609     0.001665     0.300080
H      3.194749    -0.353196    -0.223934
O     -0.044331     1.355726    -0.228495
H      0.628856     1.905515     0.161926
C     -1.248651    -0.749384    -0.241064
O     -2.455867    -0.060507     0.201783
H     -2.338769     0.832194    -0.188720
H     -0.073137    -0.089445     1.394002
H     -1.226536    -1.787913     0.146590
H      1.132009    -1.855775     0.371678
H      1.127880    -0.963356    -1.265091
H     -1.042307    -0.734311    -1.348578
```

This input generates an RDF for all C-H distances below 2 Ångström of a glycerol BOMD. The SIMILARITY option requires the definition of at least one PATTERN geometry. This option triggers the calculation of the normalized RMS distance between the MD and PATTERN geometries which can be used as an index to classify isomers or calculate isomer lifetimes. By default, pattern and MD geometries are aligned (see 4.1.7) before anything is calculated. The automatic pattern-geometry alignment can be modified with the ALIGNMENT keyword if desired. The following input performs a similarity analysis, excluding hydrogen atoms, of glycerol conformers along a BOMD trajectory.

```

ALIGN ENANTIOMER EXCLUDE
H
SIMULATION ANALYSE SIMILARITY INT=1
#
# Pattern definitions
#
PATTERN GLYA
C      0.000000    0.000000    0.000000
O      0.000000    0.000000    1.462899
C      1.474005    0.000000   -0.440301
H     -0.893691   -0.260594    1.766994
O      2.162480   -1.176562    0.084086
H      1.941117   -1.214952    1.038194
C      1.632514   -0.059827   -1.970356
O      3.025324   -0.048486   -2.363928
H      3.464920   -0.717132   -1.794492
H      1.971460    0.922883   -0.056060
H      1.170832    0.835048   -2.435026
H     -0.510825    0.910396   -0.391119
H     -0.502224   -0.913371   -0.392432
H     1.121602   -0.977513   -2.350343
PATTERN GLYB
C      0.000000    0.000000    0.000000
O      0.000000    0.000000    1.447257
C      1.448191    0.000000   -0.553309
H      0.683368    0.660325    1.699576
O      2.197151    1.114528    0.025822
H      1.975852    1.886268   -0.546110
C      1.505454    0.056976   -2.092055
O      1.020939    1.380454   -2.494249
H      1.337270    1.554323   -3.403051
H     -0.524157   -0.927069   -0.313648
H      2.557147   -0.088302   -2.425107
H      0.859665   -0.741313   -2.524473
H     -0.548895    0.882612   -0.405640
H     1.966016   -0.916375   -0.200050
GEOMETRY
C      1.251600   -0.743355   -0.192056
C      2.419080    0.016064    0.186570
C     -0.012609    0.001665    0.300080
O      3.194749   -0.353196   -0.223934
O     -0.044331    1.355726   -0.228495
O      0.628856    1.905515    0.161926
H     -1.248651   -0.749384   -0.241064
H     -2.455867   -0.060507    0.201783
H     -2.338769    0.832194   -0.188720
H     -0.073137   -0.089445    1.394002
H     -1.226536   -1.787913    0.146590
H     1.132009   -1.855775    0.371678
H     1.127880   -0.963356   -1.265091
H     -1.042307   -0.734311   -1.348578

```

With the Lindemann option the root-mean-square bond length fluctuation [230],

$$\delta = \frac{2}{N(N-1)} \sum_{A < B} \frac{\sqrt{\langle R_{AB}^2 \rangle - \langle R_{AB} \rangle^2}}{\langle R_{AB} \rangle} ,$$

with  $N$  and  $R_{AB}$  being the number of atoms and interatomic distances, respectively, are

calculated along a trajectory. Similarly, the mean interatomic distance [231],

$$l_{miad} = \frac{1}{N(N-1)} \sum_{A < B} |R_{AB}| ,$$

is calculated with the MEANDIS option. It measures the average "extension" of a structure in a molecular dynamics simulation. The prolate deformation coefficient [232],

$$\varepsilon_{pro} = \frac{2Q_1}{Q_2 + Q_3} ,$$

with  $Q_1 \geq Q_2 \geq Q_3$  being the eigenvalues of the geometrical quadrupole tensor over all atoms with elements,

$$Q_{ij} = \sum_A R_{Ai} R_{Aj} ; i, j = x, y, z ,$$

are calculated along a trajectory with the PROLATE option.

The LENGTH, ANGLE and DIHEDRAL options require the definition of the atoms involved, in the next input lines. Two numbers are required to define each bond LENGTH, three numbers to define each ANGLE and four numbers to define each DIHEDRAL angle. Up to ten parameters can be defined in this way, in consecutive lines after the keyword, as the following input example shows.

```
SIMULATION ANALYSE DIHEDRAL INT=10
2 1 3 7
2 1 3 5
8 7 3 1
8 7 3 5
1 5 3 8
8 3 5 1
4 5 6 7
7 6 5 4
3 1 10 2
12 11 10 9
#
GEOMETRY
C      1.251600    -0.743355    -0.192056
O      2.419080     0.016064    0.186570
C     -0.012609     0.001665    0.300080
H      3.194749    -0.353196   -0.223934
O     -0.044331    1.355726   -0.228495
H      0.628856    1.905515    0.161926
C     -1.248651    -0.749384   -0.241064
O     -2.455867    -0.060507    0.201783
H     -2.338769     0.832194   -0.188720
H     -0.073137    -0.089445    1.394002
H     -1.226536    -1.787913    0.146590
H     1.132009    -1.855775    0.371678
H     1.127880    -0.963356   -1.265091
H    -1.042307    -0.734311   -1.348578
```

The standard energy output of a deMon BOMD run has the following form:

TIME [FS]	T [K]	EKIN	EPOT	ETOT	$\langle T \rangle$ [K]	$\langle ETOT \rangle$
10.0	103.5	0.00098	-648.62892	-648.62793	103.0	-648.62793
20.0	100.3	0.00095	-648.62889	-648.62794	102.7	-648.62793
30.0	91.4	0.00087	-648.62885	-648.62799	100.2	-648.62795
40.0	84.9	0.00081	-648.62882	-648.62801	96.8	-648.62796
50.0	84.2	0.00080	-648.62874	-648.62794	94.2	-648.62797
60.0	107.4	0.00102	-648.62869	-648.62767	94.3	-648.62794
70.0	185.8	0.00176	-648.62858	-648.62681	102.2	-648.62784
80.0	252.2	0.00240	-648.62831	-648.62592	118.8	-648.62763
90.0	154.0	0.00146	-648.62799	-648.62653	128.2	-648.62747
100.0	56.3	0.00053	-648.62772	-648.62719	124.2	-648.62742

For each time step it lists the instantaneous temperature, kinetic, potential, and total energy as well as the average temperature and total energy. With the INT option of the SIMULATION keyword, the step interval for the energy output can be modified. Energy statistics along an existing trajectory file can be obtained with:

SIMULATION ANALYZE E=STANDARD

If E=STANDARD is substituted by E=KINETIC, E=POTENTIAL, E=TOTAL, or E=SYSTEM, the averages of the kinetic, potential, total or system<sup>2</sup> energies are listed. In the case of E=KINETIC the output takes the form:

TIME [FS]	EKIN	EPOT	ESYS	ETOT	$\langle EKIN \rangle$
10.0	0.00098	-648.62892	-648.62793	-648.62793	0.0009836
20.0	0.00095	-648.62889	-648.62791	-648.62794	0.0009680
30.0	0.00087	-648.62885	-648.62795	-648.62799	0.0009348
40.0	0.00081	-648.62882	-648.62798	-648.62801	0.0009027
50.0	0.00080	-648.62874	-648.62788	-648.62794	0.0008821
60.0	0.00102	-648.62869	-648.62765	-648.62767	0.0009051
70.0	0.00176	-648.62858	-648.62747	-648.62681	0.0010279
80.0	0.00240	-648.62831	-648.62741	-648.62592	0.0011989
90.0	0.00146	-648.62799	-648.62729	-648.62653	0.0012282
100.0	0.00053	-648.62772	-648.62710	-648.62719	0.0011588

After the trajectory output, energy and temperature statistics are printed. As for the DYNAMICS keyword, the SIMULATION keyword can be combined with the TRAJECTORY keyword in order to analyze or calculate properties for parts of a trajectory.

#### 4.7.4 Keyword LPCONSERVE

This keyword enables enforcement of linear and angular momentum conservation during a BOMD run.

Options:

**ON / OFF**

<sup>2</sup>The system energy refers to the extended system energy of the Nosé type thermostats.

---

<b>ON</b>	Momentum conservation is enabled. This is the default.
<b>OFF</b>	Momentum conservation is disabled.
<b>TOL=&lt;Real&gt;</b>	Tolerance for momentum conservation. The default is $10^{-10}$ atomic units.
<b>FORCES</b>	Momentum conservation is applied on particle forces.
<b>VELOCITIES</b>	Momentum conservation is applied on particle velocities.

Description:

In nature the linear and angular momentum of an isolated system are conserved. In molecular dynamics simulations, these conservation laws can be violated through numerical noise. The keyword LPCONSERVE enforces momentum conservation by adding compensating forces to the system [233]. These compensating forces modify the calculated forces in a least-squares sense to conserve momentum. Note that momentum conservation cannot be guaranteed if the system is in contact with an external thermostat.

Option FORCES act on particle forces only, this is sufficient to ensure momentum conservation if a microcanonical simulation is performed, but if a canonical simulation is requested, it is advisable to enable the modification on the velocities as well, with the option VELOCITIES. The option TOL specifies the tolerance for the requested linear and angular momentum conservation.

#### 4.7.5 Keyword VELOCITIES

This keyword specifies the initial nuclear velocities and, therefore, the initial temperature of a BOMD simulation.

Options:

##### RANDOM / ZERO / READ

<b>RANDOM</b>	Random velocities from a uniform distribution.
<b>ZERO</b>	Initial velocities are set to zero.
<b>READ</b>	Initial velocities are read from the input file.

<b>P=0</b>	Adjust the initial velocities to zero linear momentum.
<b>LP=0</b>	Adjust the initial velocities to zero linear and angular momentum.
<b>T=&lt;Real&gt;</b>	Specifies the temperature [K] to which the initial velocities are scaled. Default is 300 K.

Description:

The positions and velocities of the nuclei must be initialized at the beginning of a BOMD simulation. Although the nuclear positions are defined by the GEOMETRY keyword (see 4.1.1), the nuclear velocities are initialized by the VELOCITIES keyword. By default, random velocities are assigned to the nuclei. Under the option ZERO, the nuclei are frozen at the beginning of the BOMD simulation. The nuclear velocities can also be read from the input file deMon.inp by use of the read option of the VELOCITIES keyword. The individual atomic velocities are then read from the keyword body of the VELOCITIES keyword. The syntax and hierarchy of these assignments is similar to those for basis set assignment (see 4.3.1). Thus, initial nuclear velocities can be assigned to individual atoms by the atomic symbol (*e.g.* H4) or to atom groups by the element symbol (*e.g.* H). The following example assigns nuclear velocities to the atoms in a water molecule:

```
VELOCITIES READ
O  0.00  0.00  0.00
H1 0.10  0.05  0.00
H2 0.20 -0.07  0.00
```

Depending upon the unit definition (Å or Bohr) in the GEOMETRY keyword (see 4.1.1), the velocity unit is either Å/fs or Bohr/fs. The READ option can be combined with the RANDOM or ZERO option which then acts only on the atom velocities that are not explicitly defined in the keyword body of VELOCITIES.

The options P=0 and LP=0 of the VELOCITIES keyword are used to adjust the initial velocities to zero linear or zero total linear and angular momentum. Please note that these options will modify the read-in velocities. The temperature of the initial nuclear velocities can be specified according to the equipartition theorem via the option T.

#### 4.7.6 Keyword BATH

This keyword specifies the temperature bath for canonical BOMD simulations.

Options:

**NONE / SCALING / BERENDSEN / HOOVER / NOSE**

- |                  |   |
|------------------|---|
| <b>NONE</b>      | No temperature bath. This is the default. |
| <b>SCALING</b>   | Velocity scaling thermostat [234].        |
| <b>BERENDSEN</b> | Berendsen thermostat [235].               |
| <b>HOOVER</b>    | Isokinetic Hoover thermostat [236].       |
| <b>NOSE</b>      | Nosé-Hoover type thermostat [237–239].    |

---

<b>T=&lt;Real&gt;</b>	Temperature [K] of BOMD simulation. Default is 300 K.
<b>TAU=&lt;Real&gt;</b>	Berendsen parameter $\tau$ in ps. Default is 0.5 ps.
<b>INT=&lt;Integer&gt;</b>	Interval for velocity scaling. Default is 1.
<b>FREQ=&lt;Real&gt;</b>	Nosé coupling parameter. Default is 1500.0 cm <sup>-1</sup> .
<b>NHC=&lt;Integer&gt;</b>	Number of Nosé-Hoover chain thermostats. Default is 3.

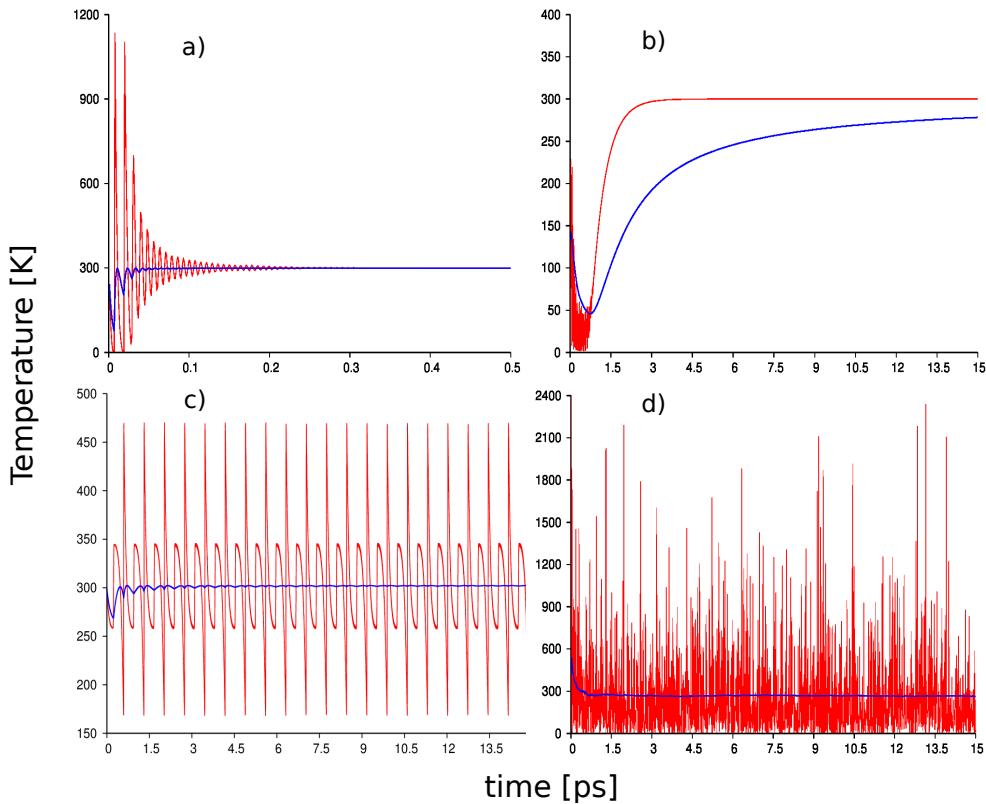
Description:

In experiments, it is often easier to control temperature than energy. Thus, many experiments are performed under conditions more appropriately described by the canonical ( $N$ ,  $V$ ,  $T$ ) ensemble. On the other hand, molecular dynamics simulations are most easily performed in the microcanonical ( $N$ ,  $V$ ,  $E$ ) ensemble. While this poses no difficulty in treating macroscopic systems, it is of critical importance in the case of microscopic systems because the transformation between different ensembles is no longer possible. Therefore, we need to perform canonical-ensemble BOMD simulations if a direct comparison to experiments under temperature control is the objective. Doing so requires a modification of the equations of motion for the nuclei [233]. In particular, thermostats must be introduced to control the temperature during the simulation. All the thermostats implemented in deMon2k work on the basis of the equipartition principle. In certain applications, e.g. the calculation of the temperature dependence of the molecular heat capacity, the thermostat must control the temperature and must also ensure a canonical phase space sampling.

The option SCALING of the BATH keyword activates a simple velocity scaling [234]. The interval for the velocity scaling can be defined with the option INT. By default, the scaling is performed in each step. Figure 11a shows the instantaneous (red) and average (blue) temperature curves for this thermostat in a BOMD run of Cu<sub>2</sub>. The thermostat temperature was set to 300 K.

As this figure shows, velocity scaling is very efficient to adjust both instantaneous and average temperatures to the thermostat value. The freezing of the instantaneous temperature indicates that the phase space distribution is not canonical. Indeed, analysis of the phase space distribution (see 4.7.3) shows a pattern similar to a microcanonical simulation rather than a canonical one [34]. The option BERENDSEN of the BATH keyword selects the BERENDSEN thermostat [235]. This thermostat couples the physical system, namely the molecule, to an external heat bath that is fixed at the desired temperature. The coupling between the molecule and the heat bath can be modified with the TAU option of the BATH keyword. Because of this coupling, the instantaneous temperature can fluctuate

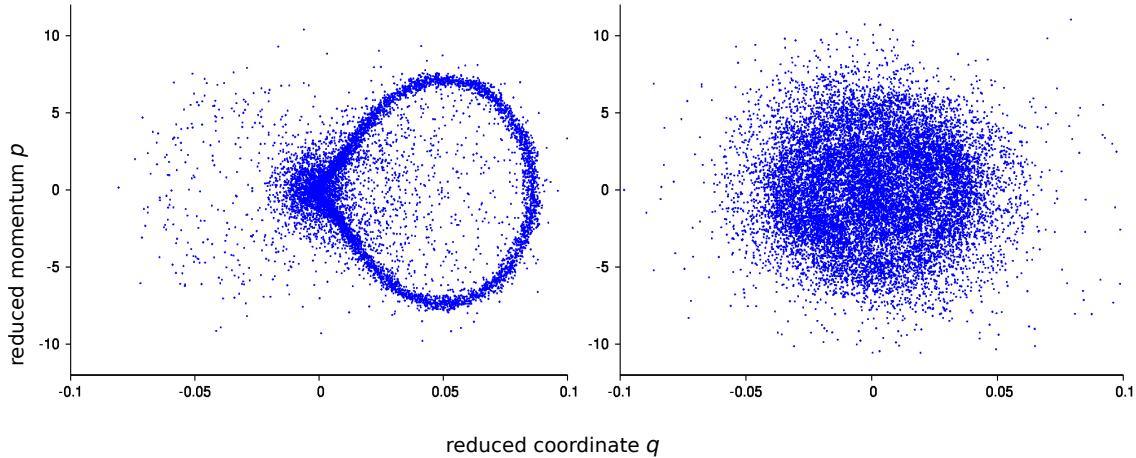
Figure 11: Temperature profiles of  $\text{Cu}_2$  BOMD simulations with (a) velocity scaling, (b) Berendsen, (c) Hoover, and (d) Nosé-Hoover thermostats. The red curves are instantaneous temperatures and the blue ones average temperatures. The thermostat temperature is set to 300 K in all runs.



around the heat bath temperature. Figure 11b shows the instantaneous (red) and average (blue) temperature curves of the Berendsen thermostat in the  $\text{Cu}_2$  simulation. In contrast to the velocity scaling, the fluctuations in the instantaneous temperature remain over the full trajectory length. However, the phase space distribution of the Berendsen thermostat is similar to the one from velocity scaling [34]. Thus, the Berendsen thermostat cannot be recommended for canonical BOMD simulations of molecules. It is, however, well suited for simulated annealing due to its exponential convergence behavior. The option HOOVER of the BATH keyword specifies a thermostat for the so-called constant kinetic temperature dynamics [236]. This method generates correct configurational properties in the canonical ensemble. The instantaneous (red) and average (blue) temperature curves for this thermostat are depicted in Figure 11c, again for the  $\text{Cu}_2$  simulation. The repetitive instantaneous

temperature fluctuations during the full simulation are characteristic of this thermostat. As a result, it causes only a few different points in phase space to be visited during the full simulation [34].

Figure 12: Phase space distributions of  $\text{N}_2$  BOMD simulations at 300 K with the Nosé-Hoover (left) and Nosé-Hoover chain (right) thermostats.



The option NOSE of the keyword BATH invokes a Nosé-Hoover (chain) thermostat [237–239]. The chain length is specified by the NHC option. By default, three chain links (thermostats) are used. With NHC=1 a pure Nosé-Hoover thermostat is specified. The instantaneous (red) and average (blue) temperature curves for the default Nosé-Hoover chain thermostat with three links are depicted in Figure 11d, again for the  $\text{Cu}_2$  simulation. This figure shows that the instantaneous temperature exhibits stochastic fluctuations during the full simulation. Despite these fluctuations, the average temperature of 300 K is reached rather quickly and is stable along the trajectory. The coupling between the Nosé-Hoover thermostat and the real system (molecule), as well as between the thermostats in the chain, is controlled by the FREQ option of the BATH keyword. The larger the FREQ value is, the weaker the coupling is. Because the performance of the Nosé-Hoover (chain) thermostat is sensitive to this coupling frequency, test calculations are recommended. This is particularly important for sensitive properties such as temperature-dependent molecular energies and entropies. The phase space distribution for a canonical  $\text{N}_2$  BOMD simulation at 300 K employing the Nosé-Hoover (NHC=1) and Nosé-Hoover chain thermostat with the default value of NHC=3 are depicted in Figure 12. The plotted reduced momenta and coordinates were obtained from a SIMULATION PHASESPACE analysis of the trajectory file.

As this figure shows, the Nosé-Hoover chain thermostat yields an almost perfect canonical phase space distribution for this system (see reference [233] for a detailed discussion). For this reason the Nosé-Hoover chain thermostat is the method of choice for the calculation of molecular properties in canonical ensembles, e.g. the calculation of the temperature dependence of molecular heat capacities.

#### 4.7.7 Keyword RDF

This keyword controls the radial distribution function calculation.

Options:

**MAX=<Real>** Maximum distance for radial distribution function calculation. The default is 10.0 Å.

**WIDTH=<Real>** Width of bins in the grid for radial distribution. The default is 0.01 Å.

Description:

The MAX option specifies how far the radial distribution function ranges out in space. With the WIDTH option the extent of a bin on the radial grid is specified. Note that the ratio MAX/WIDTH defines the number of bins on the grid. Units are automatically chosen from the GEOMETRY block specification. The following example input calculates the RDF for C-O in glycerol until 10 Å with a resolution of 0.05 Å.

```
RDF MAX=10.0 WIDTH=0.05
SIMULATION ANALYSE RDF=C-O INT=10
#
GEOMETRY
C      1.251600    -0.743355    -0.192056
O      2.419080     0.016064    0.186570
C     -0.012609     0.001665    0.300080
H      3.194749    -0.353196   -0.223934
O     -0.044331     1.355726   -0.228495
H      0.628856     1.905515    0.161926
C     -1.248651    -0.749384   -0.241064
O     -2.455867    -0.060507    0.201783
H     -2.338769     0.832194   -0.188720
H     -0.073137    -0.089445    1.394002
H     -1.226536    -1.787913    0.146590
H      1.132009    -1.855775    0.371678
H      1.127880    -0.963356   -1.265091
H     -1.042307    -0.734311   -1.348578
```

The corresponding output is given below.

```
*** RADIAL DISTRIBUTION [1/A^3] BETWEEN ATOM C AND ATOM O ***
RADIAL EXTENSION =      10.0000 ANGSTROM
RADIAL BIN WIDTH =      0.0500 ANGSTROM
```

BIN	COUNTS	DISTANCE	RDF
1	0	0.0250	0.0000
2	0	0.0750	0.0000
3	0	0.1250	0.0000
4	0	0.1750	0.0000
5	0	0.2250	0.0000
:	:	:	:
27	0	1.3250	0.0000
28	8	1.3750	0.0673
29	93	1.4250	0.7288
30	164	1.4750	1.1996
31	34	1.5250	0.2327
32	1	1.5750	0.0064
33	0	1.6250	0.0000
:	:	:	:
195	0	9.7250	0.0000
196	0	9.7750	0.0000
197	0	9.8250	0.0000
198	0	9.8750	0.0000
199	0	9.9250	0.0000
200	0	9.9750	0.0000

## 4.8 Molecular Property Control

### 4.8.1 Keyword FREQUENCY

This keyword activates the frequency analysis.

Options:

<b>FIXED</b>	A partial frequency analysis is requested.
<b>RAMAN</b>	The Raman intensities are calculated.
<b>RESTART</b>	The frequency analysis is restarted from the deMon.rst file.
<b>VIB=&lt;Real&gt;</b>	Scaling factor for the numerical step size.

Description:

By default, the infra-red intensities of the vibrations are calculated. With the option RAMAN, the Raman intensities (in atomic units) and the depolarization ratio [240] are calculated also. This is considerably more time-consuming than a standard frequency analysis. The RESTART option permits the restart of a frequency analysis. The Hessian matrix elements already calculated are read from the restart file deMon.rst and the analysis is continued. The numerical step size can be scaled with the VIB option. By default, the step size is 0.005 a.u. for all theoretical models. With VIB=2 or VIB=0.2, this step size can be scaled to 0.01 a.u. or 0.001 a.u., respectively. Note that numerical frequency calculations can be sensitive to the SCF convergence. To improve the SCF convergence in the finite-step SCF calculations the options TOL and CDF of the SCFTYPE keyword (see 4.5.1) can be used. The FIXED option specifies a partial frequency analysis, that is, calculation of frequencies involving only the motion of certain atoms. The element or atomic symbols of the atoms that are to be treated in the frequency analysis must be specified in the keyword body of the FREQUENCY keyword. The following example input shows a frequency analysis for the CO in the CuCO<sup>+</sup> cation.

```

CHARGE +1
FREQUENCIES FIXED
C  O
#
GEOMETRY Z-MATRIX ANGSTROM
Cu
C    Cu   R1
O    C    R2   Cu   180.0
#
VARIABLES
R1  1.779932
R2  1.136288

```

The calculated CO stretching frequency of 2250 cm<sup>-1</sup> is in excellent agreement with the result from an unconstrained frequency analysis (2251 cm<sup>-1</sup>).

#### 4.8.2 Keyword THERMO

This keyword activates the calculation of thermal corrections to thermodynamic functions.

It can be used **only** in combination with the FREQUENCY keyword (see Section 4.8.1).

Options:

<b>MAX=&lt;Real&gt;</b>	Maximum temperature. The default is 2000 K.
<b>MIN=&lt;Real&gt;</b>	Minimum temperature. The default is 100 K.
<b>INT=&lt;Real&gt;</b>	Temperature interval. The default is 100 K.
<b>VIBONLY</b>	Restrict calculation to vibrational contributions plus zero point energy.

Description:

The thermal corrections to thermodynamic functions are calculated in the approximation of the ideal polyatomic gas [241]. For diatomic molecules, a ro-vibronic correction term is included. Electronic contributions are neglected. The temperature range can be specified with the options MAX and MIN. To avoid numerical instabilities, the minimum possible temperature is fixed to 5 K. The step size of the temperature interval is defined by the option INT. The output contains the heat capacities  $c_p$ , the entropy correction  $\Delta S$ , noted by S, the enthalpy correction  $\Delta H$ , noted by H, and the internal energy correction  $\Delta U$ , noted by U, for each temperature step. The natural logarithm of the partition function  $z = z_{trans}z_{rot}z_{vib}$  is also reported. Note that zero point contributions are not included. These are given elsewhere in the output. Thus,  $\Delta H$  and  $\Delta U$  go to zero for  $T \rightarrow 0K$  as the following output for CO<sub>2</sub> shows.

##### THERMODYNAMIC FUNCTIONS (EXCLUDING ZERO-POINT ENERGY)

T[K] Cp[J/(K\*MOL)] S[J/(K\*MOL)] H[KJ/MOL] U[KJ/MOL] ln(z)

5.0	29.10	97.76	0.15	0.10	63.01
100.0	29.23	184.95	2.91	2.08	73.50
200.0	32.68	206.03	5.97	4.31	75.94
300.0	37.58	220.23	9.49	7.00	77.44
400.0	41.59	231.61	13.46	10.13	78.57
500.0	44.76	241.24	17.78	13.62	79.49
600.0	47.34	249.64	22.39	17.40	80.29

The VIBONLY option restricts the energetic contributions to the internal degrees of freedom (vibrational contributions). Thermodynamic functions calculated with this option enabled include the vibrational zero point energy as the following CO<sub>2</sub> output shows.

##### VIBRATIONAL THERMODYNAMIC FUNCTIONS (INCLUDING ZERO-POINT ENERGY)

T [K]	Cp [J/(K*MOL)]	S [J/(K*MOL)]	H [KJ/MOL]	U [KJ/MOL]	ln (z)
5.0	8.31	0.00	30.27	30.27	-728.24
100.0	8.45	0.02	30.28	30.28	-36.41
200.0	11.89	0.92	30.43	30.43	-18.19
300.0	16.79	3.32	31.03	31.03	-12.04
400.0	20.80	6.33	32.09	32.09	-8.89
500.0	23.97	9.47	33.50	33.50	-6.92
600.0	26.56	12.56	35.20	35.20	-5.55

#### 4.8.3 Keyword DIPOLE

This keyword activates the calculation of molecular electrostatic moments.

Options: None

Description:

The DIPOLE keyword specifies calculation of the electrostatic dipole, quadrupole, and octupole moments. If higher moments are required, the parameter MAXMOM in the parameter.h file must be reset. The electrostatic moments are calculated from the orbital density (not the AUXIS density; see Section 1.4) independent of whether the energy calculation was specified as AUXIS or BASIS. If the electrostatic moments are to be calculated in the STANDARD ORIENTATION (see Section 4.1.6 for the definition), a single-point calculation with the symmetry keyword activated (SYMMETRY ON) must be performed. Depending on the unit selection in the GEOMETRY keyword (4.1.1), the electrostatic moments are calculated either in esu (for ANGSTROM) or atomic units (for BOHR). Traceless moments are given in the last column of the moment output.

#### 4.8.4 Keyword POLARIZABILITY

This keyword activates the calculation of polarizabilities.

Options:

##### ALPHA / BETA / GAMMA

**ALPHA**      The polarizability tensor  $\alpha$  is calculated. This is the default.

**BETA**      The first hyperpolarizability  $\beta$  is calculated.

**GAMMA**      The second hyperpolarizability  $\gamma$  is calculated.

##### ANALYTICAL / NIACPKS / NUMERICAL

**ANALYTICAL**      The polarizability is calculated analytically. This is the default.

<b>NIACPKS</b>	The polarizability is calculated with the non-iterative coupled-perturbed Kohn-Sham method.
<b>NUMERICAL</b>	The polarizability is calculated numerically with a finite-field method.
<b><u>DD / DQ</u></b>	
<b>DD</b>	Dipole-dipole polarizabilities are calculated. This is the default.
<b>DQ</b>	Dipole-quadrupole polarizabilities are calculated. Only valid in combination with the option ANALYTICAL.
<b><u>SHG / EOPE / OR</u></b>	
<b>SHG</b>	Calculate second harmonic generation. Only valid in combination with the options ANALYTICAL and W=<Real>.
<b>EOPE</b>	Calculate electro-optical Pockels effect. Only valid in combination with the options ANALYTICAL and W=<Real>.
<b>OR</b>	Calculate optical rectification. Only valid in combination with the options ANALYTICAL and W=<Real>.
<b>W=&lt;Real&gt;</b>	Dynamical polarizability frequency (or wavelength) that can be given in a.u., eV, cm <sup>-1</sup> , and nm by the unit abbreviations AU, EV, CM, and NM, respectively. By default, the field frequency is given in atomic units. The form <b>W1=&lt;Real&gt;</b> is an equivalent input format.
<b>W2=&lt;Real&gt;</b>	Second dynamical hyperpolarizability frequency. Input format as for <b>W=&lt;Real&gt;</b> .
<b>EFISH</b>	Activate the EFISH orientation.
<b>FFS=&lt;Real&gt;</b>	Finite field strength for the polarizability calculation with the option NUMERICAL.
<b>TOL=&lt;Real&gt;</b>	Threshold for numerical exchange-correlation kernel derivative pruning. Default value is 1.0.
<b>FALDA</b>	The LDA kernel will be used in the response equation system. Note that this setting will apply to all response calculations in a run.
<b>GALDA</b>	The LDA kernel derivative will be used for the hyperpolarizability response equation system.
<b>FDKERNEL</b>	A finite difference kernel will be used in the response equation system. Note that this setting will apply to all response calculations in a run.

Description:

In deMon2k polarizabilities can be calculated either analytically by auxiliary density perturbation theory (ADPT) [30,158], semi-numerically by the non-iterative coupled-perturbed Kohn-Sham (NIA-CPKS) method [37] or numerically by the finite-field method [36]. The method is selected with the options ANALYTICAL, NIACPKS or NUMERICAL. All three methods are available for RKS, UKS and ROKS. The mean polarizability  $\bar{\alpha}$  and polarizability anisotropy  $|\Delta\alpha|^2$  are calculated in the principal axis system of the polarizability tensor as:

$$\bar{\alpha} = \frac{1}{3} (\alpha_{xx} + \alpha_{yy} + \alpha_{zz}) \quad (4.14)$$

$$|\Delta\alpha|^2 = \frac{1}{2} [(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{xx} - \alpha_{zz})^2 + (\alpha_{yy} - \alpha_{zz})^2] \quad (4.15)$$

Dipole-quadrupole polarizabilities [242] can be calculated with the option ANALYTICAL or NIACPKS employing the DQ option of the POLARIZABILITY keyword. First hyperpolarizabilities,  $\beta$  can be calculated with the options ANALYTICAL or NUMERICAL specifying BETA in the POLARIZABILITY keyword line.

```

Dipoles
Grid Fine
Symmetry On
AUXIS (GEN-A2*)
BASIS (TZVP-FIP1)
VXCTYP AUXIS BLYP
POLARIZABILITY BETA ANALYTICAL EFISH
#
Geometry Z-Matrix
C1
 01   C1   LC01
Constants
LC01   1.1281

```

In analytic first hyperpolarizability calculations the full  $\beta$  tensor is calculated as the following output shows.

MOLECULE ORIENTATION FOR POLARIZABILITIES IN ANGSTROM							
NO.	ATOM	X	Y	Z	Z-ATOM	MASS	TYPE
1	C1	0.000000	0.000000	-0.644365	6	12.011	QM
2	O1	0.000000	0.000000	0.483735	8	15.999	QM

\*\*\* HYPERPOLARIZABILITY \*\*\*

BETA TENSOR COMPONENTS XIJ

	<sup>1</sup>	<sup>2</sup>	<sup>3</sup>
1	0.118052	0.097283	5.999359
2	0.097283	0.004124	-0.040942
3	5.999359	-0.040942	-0.133288

## BETA TENSOR COMPONENTS YIJ

	<sup>1</sup>	<sup>2</sup>	<sup>3</sup>
1	0.097283	0.004124	-0.040942
2	0.004124	-0.053564	5.749705
3	-0.040942	5.749705	-0.182612

## BETA TENSOR COMPONENTS ZIJ

	<sup>1</sup>	<sup>2</sup>	<sup>3</sup>
1	5.999359	-0.040942	-0.133288
2	-0.040942	5.749705	-0.182612
3	-0.133288	-0.182612	26.180178

## HYPERPOLARIZABILITY TENSOR NORMS [A.U.]

DIPOLAR BETA NORM : 29.38  
 QUADRUPOLAR BETA NORM : 5.44  
 TOTAL BETA TENSOR NORM: 29.88

AVERAGE FIRST HYPERPOLARIZABILITY: 22.76

In numerical first hyperpolarizability calculations only selected components of the  $\beta$  tensor are calculated. Therefore, the numerical  $\beta$  calculations are always performed in the so-called EFISH (Electric Field Induced Second Harmonic generation) orientation. In it, the  $z$ -axis of the molecule is oriented along the permanent dipole moment of the system. The same holds for numerical second hyperpolarizability,  $\gamma$ , calculations. In EFISH orientation, the mean first and second hyperpolarizabilities are defined as:

$$\bar{\beta} = \frac{3}{5} \sum_i \beta_{iiz} \quad (4.16)$$

$$\bar{\gamma} = \frac{1}{5} \sum_{i,j} \gamma_{iijj} \quad (4.17)$$

Thus, the AVERAGE FIRST HYPERPOLARIZABILITY of a numerical  $\beta$  calculation,

## MOLECULE ORIENTATION FOR POLARIZABILITIES IN ANGSTROM

NO.	ATOM	X	Y	Z	Z-ATOM	MASS	TYPE
1	C1	0.000000	0.000000	-0.644365	6	12.011	QM
2	O1	0.000000	0.000000	0.483735	8	15.999	QM

\*\*\* FIRST HYPERPOLARIZABILITY \*\*\*

## INDUCED DIPOLE MOMENTS (IN A.U.)

ELECTRIC FIELD		X	Y	Z	SCF CYCLES
Fz(-0.0100)		0.000000	-0.000000	-0.087547	12
Fz(-0.0100)	Fz(-0.0100)	0.000000	0.000000	-0.244794	13
Fy(-0.0100)		0.000000	-0.123112	0.070726	14
Fy(-0.0100)	Fy(-0.0100)	-0.000000	-0.247706	0.071423	15
Fx(-0.0100)		-0.123112	-0.000000	0.070726	14
Fx(-0.0100)	Fx(-0.0100)	-0.247706	-0.000000	0.071423	15
Fx(+0.0100)		0.123112	0.000000	0.070726	14
Fx(+0.0100)	Fx(+0.0100)	0.247706	-0.000000	0.071423	15
Fy(+0.0100)		-0.000000	0.123112	0.070726	14
Fy(+0.0100)	Fy(+0.0100)	-0.000000	0.247706	0.071423	15
Fz(+0.0100)		0.000000	0.000000	0.230419	11
Fz(+0.0100)	Fz(+0.0100)	0.000000	0.000000	0.395391	12

## FIRST HYPERPOLARIZABILITY [A.U.]

BETA(XXX) = 0.00021	BETA(XYY) = -0.00000	BETA(XZZ) = 0.00000
BETA(YXX) = 0.00000	BETA(YYY) = -0.00000	BETA(YZZ) = 0.00000
BETA(ZXX) = 4.64573	BETA(ZYY) = 4.64570	BETA(ZZZ) = 25.74926

AVERAGE FIRST HYPERPOLARIZABILITY: 21.02

always refers to the EFISH orientation. As a consequence, analytical and numerical  $\bar{\beta}$  results can only be directly compared by using this orientation in both calculations!

Dynamical polarizabilities [243] and hyperpolarizabilities [244] can only be calculated with the analytical ADPT approach, *i.e.* with the option ANALYTICAL. In these calculations the option W=<Real> (or its alternative form W1=<Real>) specifies the frequency (or wavelength) of the (first) external field. It can be given in a.u., eV, cm<sup>-1</sup>, or nm by using the corresponding unit abbreviation. By default W is set to zero. Note that for polarizabilities imaginary frequencies are also possible [245]. They are defined by negative real numbers in the W=<Real> option of the POLARIZABILITY keyword. If for dynamical first hyperpolarizabilities only one frequency is specified either the second harmonic generation (SHG), the electro-optical Pockels effect (EOPE) or the optical rectification (OR) is calculated. These non-linear optical properties are characterized by the following frequency combinations for the dynamical  $\beta$  tensor elements  $\beta_{xyz}(-\omega_3; \omega_1, \omega_2)$  [246]:

SHG	$\omega_1=\pm\omega$	$\omega_2=\pm\omega$	$\omega_3=\pm2\omega$
EOPE	$\omega_1=0$	$\omega_2=\pm\omega$	$\omega_3=\pm\omega$
OR	$\omega_1=\pm\omega$	$\omega_2=\mp\omega$	$\omega_3=0$

By default the SHG is calculated as with the following input line.

```
POLARIZABILITY BETA ANALYTICAL EFISH W=0.199 au
```

To calculate the EOPE use:

```
POLARIZABILITY BETA ANALYTICAL EFISH W=0.199 au EOPE
```

With the W2=<Real> option a second frequency for dynamical  $\beta$  calculations can be explicitly specified in the input.

The default finite-field strength for the numerical polarizability calculations is 0.01 atomic units. With the option FFS, that field strength can be modified. Please note that the calculated polarizabilities are sensitive to both the finite field strength and the SCF convergence. With the FALDA and GALDA options the use of the LDA kernel and kernel derivative is requested, respectively, independent of the functional specified with the VXCTYPE keyword (see 4.2.1). With the FDKERNEL option a finite-difference kernel calculation [37] is requested. This permits static and dynamic ADPT polarizability calculations for all LDA and GGA functionals for which a potential is implemented. By default the FDKERNEL option is only activated for functionals for which no analytic ADPT kernel is available [247]. Note that the FALDA, GALDA and FDKERNEL settings apply to all response calculations in a deMon2k run.

#### 4.8.5 Keyword EXCITATION

This keyword specifies a time-dependent DFT (TDDFT) calculation.

Options:

##### **DAVIDSON / RS / DSYEV / D&C / JACOBI**

**DAVIDSON** Iterative Davidson diagonalization of the TDDFT matrix.

**RS** EISPACK Householder diagonalization of the TDDFT matrix. This is the default.

**DSYEV** LAPACK Householder diagonalization of the TDDFT matrix.

**D&C** LAPACK divide and conquer diagonalization of the TDDFT matrix.

**JACOBI** Jacobi diagonalization of the TDDFT matrix.

**TDA** Activation of Tamm-Dancoff approximation.

Description:

Excitation energies and oscillator strengths are calculated using the formulation of Ref. [12, 35,248]. If full diagonalization is used, *i.e.*, the option DAVIDSON is not used, then the oscillator strength sums are exact for the orbital basis set. In particular, this means that the dipole polarizability calculated as an oscillator strength sum is an analytic derivative value

which should agree with the dipole polarizability found by the finite difference or other method. Note that if the Tamm-Dancoff approximation is used, the oscillator strength sums are no longer exact. The Tamm-Dancoff approximation [249,250] can actually improve the quality of computed excitation energies when the molecular geometry is far from the equilibrium geometry.

#### 4.8.6 Keyword NONCOLLINEAR

This keyword activates non-collinear spin-flip in TDDFT calculations.

Options: None

Description:

This keyword specifies the calculation of excitation energies and oscillator strengths using non-collinear spin-flip TDDFT [251]. It is valid only in combination with the keyword EXCITATION.

#### 4.8.7 Keyword XRAY

This keyword activates an x-ray absorption or emission spectrum calculation.

Options:

##### XAS / XES

**XAS** Request an x-ray absorption spectrum calculation. This is the default.

**XES** Request an x-ray emission spectrum calculation.

**ALPHA=<MO1>-<MO2>** Specifies  $\alpha$  molecular orbital range from integer <MO1> to <MO2> for XES.

**BETA=<MO1>-<MO2>** Specifies  $\beta$  molecular orbital range from integer <MO1> to <MO2> for XES.

**TOL=<Real>** Cutoff tolerance for (extended) XAS basis set orthogonalization. Default is  $10^{-6}$ .

With the XAS option the calculation of x-ray absorption spectra are requested. By default, the core-hole is specified as a fractionally-occupied molecular orbital with the MOMODIFY (see 4.4.4) keyword. Usually a half-core-hole (transition potential) is used for the spectrum calculation [42,252]. The following input yields a  $1s$  oxygen XAS for water:

```

MOMODIFY 1 0
1 0.5
PRINT MOS
XRAY XAS
SCFTYPE UKS
VXCTYPE BLYP
AUXIS (GEN-A2*)
BASIS (IGLO-III)
#
AUGMENT
O (XAS-I)
#
GEOMETRY Z-MATRIX ANGSTROM
O                               8    15.999400
H   1   R                      1    1.007940
H   1   R   2   A               RAD   1    1.007940
#
VARIABLES
R      0.97344553
A      104.736377

```

Note the half-core-hole definition with MOMODIFY and the augmentation of the basis set (see 4.3.2) for the XAS calculation. **For reliable spectra the augmentation basis is needed at least on the core-excited atom. Otherwise the sampling of the continuum and Rydberg states will be too coarse.** The half-core-hole can be verified from the  $\alpha$  MO printing,

ALPHA MO COEFFICIENTS OF CYCLE 9						
			1	2	3	4
			-19.8771	-1.1772	-0.7189	-0.5834
			0.5000	1.0000	1.0000	1.0000
1	1	0	1s	0.1649	-0.0372	0.0000
2	1	0	2s	0.3151	-0.0818	0.0000
3	1	0	3s	0.4328	-0.1554	0.0000
4	1	0	4s	0.2189	-0.0896	0.0000
5	1	0	5s	0.0141	0.3718	0.0000
6	1	0	6s	-0.0051	0.4611	0.0000
7	1	0	7s	-0.0015	0.1289	-0.0001
:	:	:	:	:	:	:
36	2	H	1s	0.0000	0.0449	0.0688
37	2	H	2s	0.0004	0.0876	0.1587
38	2	H	3s	0.0019	0.0285	0.1328
39	2	H	4s	-0.0002	0.0008	0.0120
:	:	:	:	:	:	:
46	3	H	1s	0.0000	0.0449	-0.0688
47	3	H	2s	0.0004	0.0876	-0.1587
48	3	H	3s	0.0019	0.0285	-0.1327
49	3	H	4s	-0.0002	0.0008	-0.0119
:	:	:	:	:	:	,

after SCF convergence. The XAS output, employing the basis set augmentation, is given in the property output section of deMon.out as:

TRANSITION			TRANSITION MOMENTS		
NO.	E [eV]	STRENGTH	X	Y	Z
1	536.10	0.0059	0.0167	0.0000	0.0129
2	538.09	0.0132	0.0193	0.0000	-0.0251
3	539.25	0.0016	-0.0087	0.0003	-0.0067
4	539.30	0.0021	0.0002	0.0126	0.0001
5	539.69	0.0007	0.0056	0.0000	0.0044
6	539.99	0.0011	0.0056	0.0000	-0.0073
7	540.22	0.0000	0.0007	0.0000	0.0005
8	540.25	0.0002	0.0002	-0.0038	0.0002
9	540.25	0.0002	0.0032	0.0002	0.0025
10	540.28	0.0003	-0.0003	0.0044	-0.0002
11	540.28	0.0002	0.0030	0.0005	0.0022
12	540.28	0.0001	0.0013	0.0000	-0.0019
13	540.30	0.0000	0.0000	0.0000	0.0000
:	:	:	:	:	:
:	:	:	:	:	:

The transitions are ordered according to their transition energies. The transition strength [a.u.] is calculated as:

$$S_{hi} = \frac{2}{3} (\varepsilon_i - \varepsilon_h) \langle \psi_h | \mathbf{r} | \psi_i \rangle^2 \quad (4.18)$$

The transition moment components are given in the last three columns of the above output. This information is also written to the file deMon.xxy for further processing with the utility program xray2k. This program convolutes the spectrum and produces output for plotting the results. To improve the absolute energies of the XAS spectrum and place it on an absolute energy scale, it is recommended to calculate explicitly the first core-excited state total energy. For such a calculation we remove a full 1s electron, but only after the occupations are already generated. The excited electron must thus be included as an extra charge (-1) in the molecule. For the water molecule this is the corresponding input:

```

CHARGE -1
MOMODIFY 1 0
1 0.0
SCFTYPE UKS
VXCTYPE BLYP
AUXIS (GEN-A2*)
BASIS (IGLO-III)
#
GEOMETRY Z-MATRIX ANGSTROM
O                      8      15.999400
H   1   R              1      1.007940
H   1   R   2   A       RAD    1      1.007940
#
VARIABLES
R          0.97344553
A          104.736377

```

This yields the following MO occupations after SCF convergence:

ALPHA MO COEFFICIENTS OF CYCLE 9				1	2	...	6	7
				-20.1462	-1.1230	...	-0.1018	-0.0396
				0.0000	1.0000	...	1.0000	0.0000
1	1	0	1s	0.1660	-0.0396	...	-0.0143	0.0000
2	1	0	2s	0.3174	-0.0866	...	-0.0320	0.0000
3	1	0	3s	0.4291	-0.1699	...	-0.0616	0.0000
4	1	0	4s	0.2214	-0.0798	...	-0.0349	0.0000
5	1	0	5s	0.0102	0.4080	...	0.1954	0.0000
6	1	0	6s	-0.0020	0.4585	...	0.1882	-0.0001
7	1	0	7s	0.0000	0.1112	...	0.9772	0.0000
:	:	:	:	:	:	:	:	:
36	2	H	1s	0.0000	0.0422	...	-0.0339	0.0388
37	2	H	2s	0.0001	0.0807	...	-0.1164	0.1102
38	2	H	3s	0.0003	0.0281	...	-0.3905	0.2107
39	2	H	4s	-0.0001	0.0006	...	-0.6048	1.3031
:	:	:	:	:	:	:	:	:
46	3	H	1s	0.0000	0.0422	...	-0.0339	-0.0388
47	3	H	2s	0.0001	0.0807	...	-0.1164	-0.1102
48	3	H	3s	0.0003	0.0281	...	-0.3903	-0.2108
49	3	H	4s	-0.0001	0.0006	...	-0.6053	-1.3029
:	:	:	:	:	:	:	:	:

The energy difference between this first excited core-state and the ground state is used to shift the origin of the transition potential spectrum, *i.e.* the first excitation of the above TRANSITION output. A further correction can be introduced by calculating the core ionization potential, here of the O<sub>1s</sub> orbital in water, and compare with results from experimental x-ray photoelectron spectra (XPS). This offset also corrects for deficiencies in the functional [252,253]. The input for the XPS calculation has the form:

```
MOMODIFY 1 0
1 0.0
SCFTYPE UKS
VXCTYPE BLYP
AUXIS (GEN-A2*)
BASIS (IGLO-III)
#
GEOMETRY Z-MATRIX ANGSTROM
O                               8   15.999400
H   1   R                      1   1.007940
H   1   R   2   A               RAD  1   1.007940
#
VARIABLES
R      0.97344553
A      104.736377
```

If core-levels are close in energy, like in a water cluster, ECPs (see 4.3.4) can be used to have the relevant core-level uniquely defined. The following shows a water pentamer input for an XAS calculation using this technique:

```
MOMODIFY 1 0
1 0.5
```

```

XRAY XAS
BASIS (TZVP)
01 (IGLO-III)
02 (RECP6|SD)
SCFTYPE UKS MAX=20 TOL=0.100E-04
VXCTYPE PBE
AUGMENTATION
01 (XAS-I)
ERIS DIRECT TOL=1.0E-8
#
# Pentamer geometry
#
GEOMETRY CARTESIAN ANGSTROM
01    0.000000    0.000000    0.000000
02    2.264218    1.270427   -1.048492
02   -0.236014    0.380416    2.763979
02   -2.323588    1.242303   -0.947429
02   -0.113699   -2.602651   -1.026294
H   -0.081064   -1.025238   -0.277800
H   -0.036255    0.134954    1.127393
H   -0.909319   -3.125997   -0.746222
H    0.695861   -3.190749   -0.803055
H    0.617200    0.800843    3.188623
H   -0.994919    0.919704    3.110267
H   -3.037164    0.667781   -0.585638
H   -1.475886    0.786984   -0.720841
H    1.514773    0.710601   -0.705452
H    3.110268    0.796835   -0.730599

```

The isolated core-level of 01 can be easily identified (first MO) in the corresponding MO output:

#### ALPHA MO COEFFICIENTS OF CYCLE 10

				1	2	3	4	...
				-19.7860	-1.0776	-1.0646	-1.0562	...
				0.5000	1.0000	1.0000	1.0000	...
1	1	0	1s	0.1647	-0.0327	0.0133	0.0133	...
2	1	0	2s	0.3162	-0.0718	0.0293	0.0292	...
3	1	0	3s	0.4323	-0.1375	0.0560	0.0558	...
4	1	0	4s	0.2183	-0.0774	0.0319	0.0319	...
5	1	0	5s	0.0146	0.3176	-0.1315	-0.1317	...
6	1	0	6s	-0.0061	0.4356	-0.1739	-0.1716	...
7	1	0	7s	0.0065	0.0788	-0.0343	-0.0351	...
:	:	:	:	:	:	:	:	...
:	:	:	:	:	:	:	:	...

With the XES option the calculation of x-ray emission spectra is requested. The core levels are specified with the ALPHA and BETA options of the XRAY keyword. Since the ground state is used we can compute spectra for all relevant core-levels in one calculation so a range of orbitals is specified. This range can include only one orbital as shown in the following input example for the first  $\alpha$  MO XES of water:

```

SCFTYPE UKS
VXCTYPE PBE BASIS

```

```

XRAY XES ALPHA=1-1
BASIS (IGLO-III)
H Read
 1   0   3
    14.9588900000      0.0349460000
    2.2563909000      0.2347270000
    0.5112084000      0.8137573000
 2   0   1
    0.1219492000      1.0000000000
 3   0   1
    0.0360000000      1.0000000000
 2   1   1
    1.1000000000      1.0000000000
AUXIS (GEN-A4*)
#
GEOMETRY CARTESIAN ANGSTROM
O      0.00000000  0.00000000  0.20476407
H      0.66694680  0.00000000 -0.39124903
H     -0.66694680  0.00000000 -0.39124903

```

Here the relevant O 1s level is the lowest in energy, but this depends on the system. The specification is by way of orbital number so that in more complicated cases a ground state calculation should be done first to find the relevant orbital(s). The corresponding XES output is given in the property output section of deMon.out and reads for this example as:

XES CALCULATION FOR MO #: 1

IONIZATION POTENTIAL = 510.08 EV

TRANSITION		TRANSITION MOMENTS			
NO.	E [eV]	STRENGTH	X	Y	Z
2	483.78	0.4311	0.0000	0.0000	-0.0152
3	496.73	5.3353	0.0513	0.0000	0.0000
4	500.29	5.5736	0.0000	0.0000	0.0519
5	502.75	7.2973	0.0000	0.0589	0.0000

This information is also written to the file deMon.xry for further processing with the utility program xray2k. The emission energies are obtained from orbital energy differences and lack effects of relaxation (most important for the inner-shell). An improved absolute energy can be obtained by computing the difference between the IP for the core and for the HOMO and use that to place the transition with highest energy. This assumes that relaxation effects in the valence are similar, which is not necessarily the case [254].

#### 4.8.8 Keyword NMR

This keyword activates the calculation of closed-shell NMR shielding tensors.

Options:

<b>NICS</b>	Calculation of nuclear-independent chemical shifts.
-------------	---

**READ** Specifies that a list of nuclei for shielding tensor calculations is read from the keyword body of NMR.

**SPINROT** Calculation of the Nuclear Spin-Rotation Constants (NSRC).

#### Description:

The magnetic shielding tensors can be calculated either within the AUXIS or BASIS GIAO formalism [255]. One or the other formalism is activated by the AUXIS or BASIS options of the VXCTYPE keyword. The GIAO scheme implemented thereby can be combined with ECPs and MCPs on neighboring centers in order to include scalar relativistic effects in the magnetic shielding tensor calculation. The default output only shows the calculated isotropic magnetic shielding [ppm] of each nucleus in the system:

$$\sigma_{iso}(\vec{P}) = \frac{1}{3} [\sigma_{xx}(\vec{P}) + \sigma_{yy}(\vec{P}) + \sigma_{zz}(\vec{P})] \quad (4.19)$$

To print the full shielding tensor, PRINT NMR must be used (see 4.12.2). The user can select individual atoms by the atomic symbol (e.g. H4) or atom groups by the element symbols (e.g. H) for which magnetic shielding tensors will be calculated with the READ option. In the following example, only the shielding tensors of the hydrogen atoms in NH<sub>3</sub> are calculated.

```
NMR READ
H
SYMMETRY ON
BASIS (IGLO-III)
VXCTYPE BASIS PBE
#
GEOMETRY Z-MATRIX ANGSTROM
N
H1      1   rNH
H2      1   rNH      2   aHNH
H3      1   rNH      3   aHNH      2   wHHNH
#
VARIABLES
rNH     1.01700
aHNH    107.80000
wHHNH   120.00000
```

The NICS option of the NMR keyword [256] activates shielding tensor calculations at the position of dummy atoms as specified in the keyword body of the GEOMETRY keyword. The following input example enables NICS calculations at two different positions, specified by X1 and X2, in an Al<sub>4</sub><sup>2-</sup> cluster.

```
NMR NICS READ
X1 X2
CHARGE -2
PRINT NMR
```

```

BASIS (DZVP)
VXCTYPE BLYP
AUXIS (GEN-A2)
SCFTYP TOL=1.E-8 CDF=1.E-5
#
GEOMETRY Z-MATRIX ANGSTROM
X1
Al      1  AlX2
Al      1  AlX3      2  AlXA13
Al      1  AlX4      2  AlXA14      3  DIH4
Al      1  AlX5      3  AlXA15      2  DIH5
X2      1  1.25      2  90.0      3  90.0
#
VARIABLES
AlX2    1.831000
AlX3    1.831000
AlX4    1.831000
AlX5    1.831000
AlXA13  90.000000
AlXA14  90.000000
AlXA15  90.000000
DIH4    180.000000
DIH5    180.000000

```

The simulation of the magnetic shieldings along a BOMD trajectory is calculated in two steps as discussed in Section 4.7.3. A BOMD magnetic shielding calculation is activated by the option NMR=<Integer> of the simulation keyword (see 4.7.3). Here <Integer> specifies the number of the atom for which the trajectory data are printed. The user can select the nuclei by activating the option READ of the NMR keyword (see 4.8.8), otherwise the magnetic shielding of all nuclei will be calculated along the BOMD trajectory. In the following example, the magnetic shielding of the carbon nucleus of methanol along a BOMD trajectory is calculated as:

```

TITLE METHANOL MD SIMULATION
#
NMR READ
C
BASIS (AUG-CC-PVDZ)
VXCTYPE AUXIS PW91
TRAJECTORY RESTART PART=10000-110000
SIMULATE CALCULATE NMR=1 INT=25
#
GEOMETRY Z-MATRIX
C
H   1  HC2
H   1  HC3      2  HCH3
H   1  HC4      3  HCH4      2  DIH4
O   1  OC5      3  OCH5      2  DIH5
H   5  HO6      1  HOC6      3  DIH6
VARIABLES
HC2    1.098100
HC3    1.098100
HCH3   109.120
HC4    1.098100
HCH4   109.120
DIH4   120.000

```

```

OC5      1.429200
OCH5    109.471
DIH5     240.000
HO6      0.975100
HOC6    107.690
DIH6     180.000
END

```

At the end of the corresponding output, the average and the standard deviation of the isotropic magnetic shieldings of all selected nuclei are printed.

The SPINROT option activates the computation of the NSRC for all magnetically active nuclei in the molecule [257]. For the NSRC, it is necessary to indicate the isotopic composition of the molecule which can be specified by the nuclear masses in the GEOMETRY definition. The NSR tensor is evaluated in the principal axes of the tensor of inertia. The output gives the values of the diagonal elements of the NSR tensor,  $C_{xx}$ ,  $C_{yy}$  and  $C_{zz}$  which are associated with the moments of inertia arranged in the order  $I_{xx} \leq I_{yy} \leq I_{zz}$ . The components are labeled A, B and C. Also the isotropic and anisotropic NSRC are given in the output. The isotropic NSRC is calculated as:

$$C_{iso} = \frac{1}{3} [C_{xx} + C_{yy} + C_{zz}] \quad (4.20)$$

The corresponding anisotropic NSRC is calculated from the eigenvalues of the traceless NSR tensor  $e_{xx} \leq e_{yy} \leq e_{zz}$  as:

$$C_{ani} = e_{zz} - \frac{1}{2} (e_{xx} + e_{yy}) \quad (4.21)$$

The NSRC can also be calculated along a BOMD trajectory. The example given below shows the input for such a calculation of the acetylene-HF complex, the H<sup>12</sup>C-<sup>12</sup>C-DF isotopologue. Note that the SIMULATION keyword options for the NSRC calculation are the same as for the magnetic shielding calculation. The NSRC calculation is triggered by the NMR SPINROT keyword line.

```

DISPERSION
BASIS (AUG-CC-PVDZ)
AUXIS (GEN-A2)
SCFTYPE MAX=1000 TOL=0.100E-04
VXCTYPE OPTX-PBE AUXIS
TRAJECTORY RESTART PART=10000-12000 INT=20
NMR SPINROT
SIMULATION CALCULATE NMR=1 INT=20
#
# Cartesian coordinates of MD step 410000
#
GEOMETRY CARTESIAN ANGSTROM
F   -66.916797   19.630729  -13.088069           18.998

```

H	-67.101518	18.720636	-12.997091	2.014
C	54.166066	-15.936748	9.958016	12.000
C	53.957601	-15.619981	11.173715	12.000
H	54.766661	-16.089846	9.024499	1.008
H	53.312450	-15.623649	12.049412	1.008

In the output,

```

NSR CONSTANTS [kHz] FOR ATOM 1

TIME [FS]      XX       YY       ZZ       C_K      <C_K>
5000.0        1.82     2.33     2.06     2.07     2.07
5010.0        2.76     1.91     1.87     2.18     2.12
5020.0        2.42     1.60     1.71     1.91     2.05
5030.0        1.63     1.77     1.90     1.77     1.98
5040.0        0.61     2.30     2.33     1.75     1.93
5050.0        0.04     2.87     2.83     1.91     1.93
5060.0        1.23     3.17     3.30     2.57     2.02
5070.0        2.32     2.91     3.31     2.85     2.12
5080.0        1.71     2.69     3.23     2.54     2.17
:           :         :         :         :         :
*** NMR SHIELDINGS STATISTIC ***

CHEMICAL SHIELDING FOR ATOM 1 ( F )
AVERAGE SHIELDING [ppm] =    380.82
STD DEV SHIELDING [ppm] =    20.27

CHEMICAL SHIELDING FOR ATOM 2 ( H )
AVERAGE SHIELDING [ppm] =    28.38
STD DEV SHIELDING [ppm] =    1.53
:           :         :         :         :
*** NSR CONSTANTS STATISTIC ***

NSR CONSTANT FOR ATOM 1 ( F )

AVERAGE NSR XX-CONSTANT [kHz] =    3.80
STD DEV NSR XX-CONSTANT [kHz] =    3.63
AVERAGE NSR YY-CONSTANT [kHz] =    2.91
STD DEV NSR YY-CONSTANT [kHz] =    0.99
AVERAGE NSR ZZ-CONSTANT [kHz] =    3.15
STD DEV NSR ZZ-CONSTANT [kHz] =    1.34

NSR CONSTANT FOR ATOM 2 ( H )

AVERAGE NSR XX-CONSTANT [kHz] =   -0.12
STD DEV NSR XX-CONSTANT [kHz] =    0.12
AVERAGE NSR YY-CONSTANT [kHz] =   -0.06
STD DEV NSR YY-CONSTANT [kHz] =    0.01
AVERAGE NSR ZZ-CONSTANT [kHz] =   -0.06
STD DEV NSR ZZ-CONSTANT [kHz] =    0.01
:           :         :         :         :

```

the instantaneous NSR tensor diagonal elements and isotropic value along with its average are listed for the selected NMR atom in the SIMULATION keyword line. The NSR constant

statistics at the end of the output are given for all atoms. Note also the NMR shielding statistics. Thus, NSR constants and NMR shieldings are always calculated together.

#### 4.8.9 Keyword MAGNETIZABILITY

This keyword activates the calculation of closed-shell magnetizabilities.

Options:

**GTENSOR** Calculation of the rotational g-tensor.

Description:

The magnetizability tensor can be calculated either with AUXIS or BASIS GIAO formalism [255]. One or the other formalism is activated by the AUXIS or BASIS options of the VX-CTYPE keyword. The default output prints the magnetizability tensor in atomic units and the isotropic and anisotropic magnetizability in both atomic units and SI units [ $10^{-30}$  JT $^{-2}$ ].

The isotropic magnetizability is calculated as:

$$\xi_{iso} = \frac{1}{3} [\xi_{xx} + \xi_{yy} + \xi_{zz}] \quad (4.22)$$

The corresponding anisotropic magnetizability is calculated in the principal axis system of the magnetizability tensor as:

$$|\Delta\xi|^2 = \frac{1}{2} [(\xi_{xx} - \xi_{yy})^2 + (\xi_{xx} - \xi_{zz})^2 + (\xi_{yy} - \xi_{zz})^2] \quad (4.23)$$

To print the diamagnetic and paramagnetic magnetizability tensors the input line PRINT MAG must be used (see 4.12.2). The magnetizability along a BOMD trajectory is calculated in two steps as discussed in Section (4.7.3). A BOMD magnetizability calculation is activated by the option MAGNETIZABILITY of the SIMULATION keyword (see 4.7.3). In the following example, the magnetizability of ethane along a BOMD trajectory is calculated:

```
TITLE ETHANE - MAGNETIZABILITY MD SIMULATION
#
BASIS (AUG-CC-PVDZ)
VXCTYPE AUXIS PW91
TRAJECTORY RESTART PART=10000-110000
SIMULATE CALCULATE MAG INT=25
#
GEOMETRY Z-MATRIX ANGSTROM
C                               6      12.011000
H    1   R1                   1      1.007940
H    1   R2     2   A1        RAD    1      1.007940
H    1   R3     2   A2     3   D1        RAD    1      1.007940
C    1   R4     2   A3     3   D2        RAD    6      12.011000
H    5   R5     1   A4     2   D3        RAD    1      1.007940
```

```

H   5   R6    1   A5    2   D4      RAD    1       1.007940
H   5   R7    1   A6    2   D5      RAD    1       1.007940
#
VARIABLES
R1          1.10761992
R2          1.10778711
R3          1.10781183
R4          1.51627843
R5          1.10773895
R6          1.10766207
R7          1.10782491
A1          107.574370
A2          107.515075
A3          111.387913
A4          111.403485
A5          111.398744
A6          111.311745
D1          -115.468944
D2          122.337112
D3          59.891940
D4          180.000000
D5          -60.021828
#
END

```

The GTENSOR option of the magnetizability keyword activates the calculation of the rotational g-tensor [227], a dimensionless quantity, whose calculation requires the magnetizability tensor. In the default output, aside from the  $3 \times 3$  rotational g-tensor calculated in the principal axis system of the tensor of inertia, its isotropic and anisotropic values are also printed:

$$g_{iso} = \frac{1}{3} [g_{xx} + g_{yy} + g_{zz}] \quad (4.24)$$

and

$$\Delta g = g_{zz} + \frac{1}{2} [g_{xx} + g_{yy}] \quad (4.25)$$

The diagonal elements of the g-tensor are ordered according to  $g_{xx} \geq g_{yy} \geq g_{zz}$ . Together with the magnetizability the rotational g-tensor can also be calculated along a BOMD trajectory. It is activated by adding the following line to the above example of the magnetizability MD simulation:

```
MAGNETIZABILITY GTENSOR
```

#### 4.8.10 Keyword NQR

This keyword activates the calculation of the nuclear quadrupole splitting as measured in NQR and NMR spectroscopy.

Options:

**BARN / FM^2**

**BARN** The values of the nuclear quadrupole moments are given in barns ( $10^{-24} \text{ cm}^2$ ). This is the default.

**FM^2** The values of the nuclear quadrupole moments are given in  $\text{fm}^2$ .

**READ** Specifies that nuclear quadrupole moments and nuclear spins are read from the keyword body of NQR.

Description:

The basic outputs of the NQR keyword are the traceless electric field gradient (EFG) components at the positions of the nuclei and the corresponding anisotropy parameters  $\eta$  defined as:

$$\eta = \frac{V_{xx} - V_{yy}}{V_{zz}} \quad (4.26)$$

with

$$| V_{zz} | > | V_{yy} | \geq | V_{xx} | \quad (4.27)$$

If nuclear quadrupole moments for specific nuclei are supplied by the READ option of the NQR keyword, then the nuclear quadrupole coupling constants (in MHz) are calculated [258,259]. In the following input, the nuclear quadrupole moment of Cl is given by the assignment line (CL Q = -7.97) in the body of the NQR keyword. In this particular case, the nuclear quadrupole moment is given in  $\text{fm}^2$ .

```
NQR READ fm^2
CL Q = -7.97
#
GEOMETRY Z-MATRIX
C1
C2 C1 RCC
C1 C1 RCC1 C2 ACCC1
H1 C1 RCH1 C2 ACCH1 C1 180.0
H2 C2 RCH2 C1 ACCH2 C1 180.0
H3 C2 RCH3 C1 ACCH3 C1 0.0
#
VARIABLES
RCC 1.332
RCC1 1.726
RCH1 1.079
RCH2 1.078
RCH3 1.090
ACCC1 122.3
ACCH1 123.8
ACCH2 119.5
ACCH3 121.0
```

The components of the nuclear quadrupole coupling constants for the Cl atom are printed in the output as:

TRACELESS EFG COMPONENTS FOR ATOM 3 ( CL )

XX COUPLING CONSTANT: -29.794 MHz  
 YY COUPLING CONSTANT: -39.685 MHz  
 ZZ COUPLING CONSTANT: 69.479 MHz  
 EFG ANISOTROPY ETA : 0.142

If the nuclear spin is also given in the input, as in the following example,

NQR READ  
 Cl Q = -0.0797 S = 1.5

then the energy levels of the nuclear quadrupole splitting [11] are also calculated. To print the corresponding NQR spin Hamiltonian matrix, use PRINT NQR in the input. The additional output reads as follows:

PRINCIPAL AXES OF EFG TENSOR

	1	2	3
1	0.00000	0.52718	0.84976
2	1.00000	-0.00002	0.00001
3	0.00003	0.84976	-0.52718

NQR SPIN HAMILTONIAN MATRIX

	1	2	3	4
1	17.36967	0.00000	1.42758	0.00000
2	0.00000	-17.36967	0.00000	1.42758
3	1.42758	0.00000	-17.36967	0.00000
4	0.00000	1.42758	0.00000	17.36967

EIGENVECTORS OF NQR SPIN HAMILTONIAN

	1	2	3	4
1	0.00000	-0.04099	0.99916	0.00000
2	0.99916	0.00000	0.00000	0.04099
3	0.00000	0.99916	0.04099	0.00000
4	-0.04099	0.00000	0.00000	0.99916

NQR ENERGY LEVELS

ENERGY LEVEL 1 = -0.1743E+02 MHz  
 ENERGY LEVEL 2 = -0.1743E+02 MHz  
 ENERGY LEVEL 3 = 0.1743E+02 MHz  
 ENERGY LEVEL 4 = 0.1743E+02 MHz

#### 4.8.11 Keyword FUKUI

This keyword is used to enable reactivity analysis based on the Fukui function.

Options:

**FALDA** The LDA kernel will be used in the response equation system. Note that this setting will apply to all response calculations in a run.

**FDKERNEL** A finite difference kernel will be used in the response equation system. Note that this setting will apply to all response calculations in a run.

Description:

Upon activation both Fukui functions, electrophilic and nucleophilic, are evaluated. It is assumed that electrons are removed or added to frontier orbitals. Note that the Fukui function is computed analytically using the method reported by Flores-Moreno et al. [260,261]. In case of a degeneracy of the frontier orbitals an average frontier density is considered.

## 4.9 Electronic Structure Analysis

### 4.9.1 Keyword POPULATION

This keyword selects a population analysis.

Options:

**MULLIKEN / LOEWIDIN / BADER / NBO / BECKE / HIRSHFELD / VORONOI**

<b>MULLIKEN</b>	A Mulliken population analysis is performed. This is the default.
<b>LOEWIDIN</b>	A Löwdin population analysis is performed.
<b>BADER</b>	A Bader population analysis is performed.
<b>NBO</b>	Create input file for an NBO analysis.
<b>BECKE</b>	Real space is partitioned into overlapping (fuzzy) Becke cells.
<b>HIRSHFELD</b>	Real space is partitioned into overlapping Hirshfeld cells.
<b>VORONOI</b>	Real space is partitioned into non-overlapping Voronoi cells.

### SCF / ITERATIVE / DEFORMATION / FULL

<b>SCF</b>	The SCF density is used for the Hirshfeld weights. Only meaningful for the HIRSHFELD option.
<b>ITERATIVE</b>	The Hirshfeld weights are refined iteratively. Only meaningful for the HIRSHFELD option.
<b>FRACTIONAL</b>	Same as the ITERATIVE option but separates $\alpha$ and $\beta$ spin densities. Only meaningful for the HIRSHFELD option.
<b>DEFORMATION</b>	The deformation density is analyzed. Only meaningful for the BECKE, HIRSHFELD and VORONOI options.
<b>FULL</b>	More information is printed.

Description:

For all population analyses, deMon2k calculates atomic charges and, in the case of open-shell systems, atomic spin populations. From the orbital-based population analysis of Mulliken [262] or Löwdin [263], the bond order [264] or valence matrix [267] is calculated as well (closed-shell systems only) with the option FULL. In the case of the Bader population analysis [38], the critical points of the electron density and the molecular graph are also calculated. These quantities can be visualized with VU [41] using the deMon.pie file. The option BADER together with the option FULL prints out extra information about the dipole and quadrupole moments of atoms and non-nuclear attractors as well as informa-

tion about their kinetic energies. For other options related to the Bader population analysis see (4.10.2). The input file for the NBO population analysis is created for use with the NBO program [268]. The options MULLIKEN, LOEWdin, BADER, NBO, BECKE, HIRSHFELD and VORONOI are mutually exclusive.

With the options BECKE, HIRSHFELD or VORONOI a topological atom-in-molecule (AIM) analysis of the electronic density is performed. The electronic density,  $\rho(\mathbf{r})$  is integrated numerically on a fixed grid of points. By selecting the BASIS or AUXIS option of the TOPOLOGY keyword (see 4.10.2), the user triggers the analysis of the Kohn-Sham density (BASIS) or of the auxiliary density (AUXIS). By default the Kohn-Sham density is used for all topological analyses. The quality of the integration grid is also set with the TOPOLOGY keyword using the options COARSE, MEDIUM, FINE or REFERENCE. The DEFORMATION option of the POPULATION keyword triggers the analysis of the deformation density,  $\Delta\rho(\mathbf{r})$ , defined as:

$$\Delta\rho(\mathbf{r}) = \rho^{SCF}(\mathbf{r}) - \rho^{Pro}(\mathbf{r}) \quad (4.28)$$

In Eq. (4.28)  $\rho^{SCF}(\mathbf{r})$  is the relaxed electronic density produced by the SCF and  $\rho^{Pro}(\mathbf{r})$  is the promolecular density obtained by the superposition of spherically averaged atomic densities. According to the TOPOLOGY options BASIS or AUXIS these are either the corresponding Kohn-Sham (default) or auxiliary densities, respectively. Having defined the electronic density to analyze, the number of electrons,  $N_A$ , is calculated for each atom  $A$  according to:

$$N_A = \sum_i^{\text{Grid}} \rho(\mathbf{r}_i) \omega_q(\mathbf{r}_i) \omega_A(\mathbf{r}_i) \quad (4.29)$$

The summation runs over all the grid points. The quadrature weights  $\omega_q$  are calculated according to the Euler-McLaurin (radial) and Lebedev (angular) quadrature schemes. The atomic weights  $\omega_A$  determine how the electron density is distributed over the atoms. Their mathematical definition is different for each topological population scheme. In particular, the following atomic weight definitions are implemented for the VORONOI, BECKE and HIRSHFELD options.

Voronoi:

$$\omega_A(\mathbf{r}_i) = \begin{cases} 1 & \forall \mathbf{r}_i \text{ most close to atom } A \\ 0 & \text{else} \end{cases} \quad (4.30)$$

Becke:

$$\omega_A(\mathbf{r}_i) = \frac{P_A(\mathbf{r}_i)}{\sum_B P_A(\mathbf{r}_i)} \quad (4.31)$$

Hirshfeld:

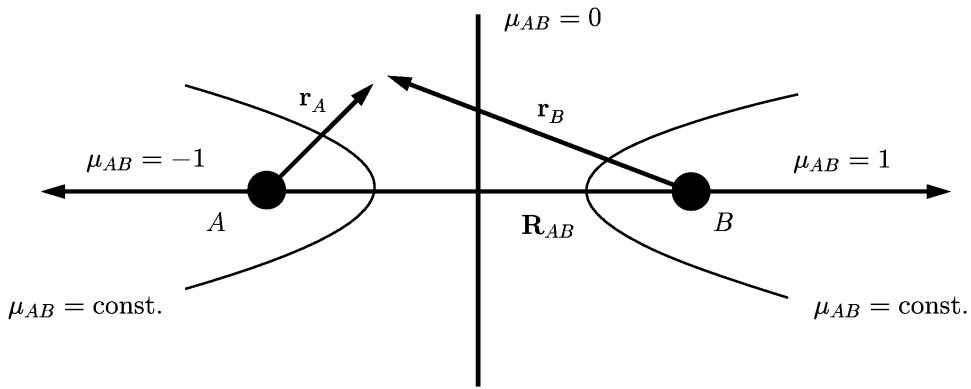
$$\omega_A(\mathbf{r}_i) = \frac{\tilde{\rho}_A^{\text{ref}}(\mathbf{r})}{\sum_B \tilde{\rho}_B^{\text{ref}}(\mathbf{r})} \quad (4.32)$$

The cell function  $P_A(\mathbf{r}_i)$  of the BECKE weights is defined by a "soft" step function in terms of the elliptic coordinates  $\mu_{AB}$  defined as [269]:

$$\mu_{AB} = \frac{\mathbf{r}_A - \mathbf{r}_B}{\mathbf{R}_{AB}} \quad (4.33)$$

The geometrical construction of these coordinates is depicted in Figure 13.

Figure 13: Elliptic coordinate definition used for the cell function of BECKE atomic weights.



As can be seen from Eq. (4.32) the atomic Hirshfeld weights are always calculated from the auxiliary densities, independent of whether the Kohn-Sham or auxiliary density is partitioned. For the reference densities in Eq. (4.32) 4 different options are available. By default spherically averaged neutral atom densities are used to calculate the atomic Hirshfeld weights. This is the standard Hirshfeld scheme from the literature [265]. With the SCF option of the POPULATION keyword  $\rho_A^{\text{ref}}$  is calculated from the atomic SCF auxiliary densities. With the ITERATIVE option  $\rho_A^{\text{ref}}$  of each atom is calculated from a spherically averaged ion with the current charge of the atom (see Eq. 19 of Ref. [266]). Note that this approach is computationally more involved than the standard Hirshfeld scheme because it requires repetitive integration of the electronic density over the grid. The tolerance criteria and the maximum number of iterations are set by the TOPOLOGY keyword options TOL and MAX, respectively. The iterative Hirshfeld scheme produces higher partial charges in absolute value than the standard one and is also applicable to ionic molecules. Moreover the arbitrariness in the definition of the reference densities of the standard Hirshfeld approach is

removed with the iterative version. In open-shell molecules either the density or the spin densities can be refined iteratively according to scheme that was just discussed. This selection is triggered by the options ITERATIVE or FRACTIONAL, respectively. Therefore, the input lines,

```
POPULATION HIRSHFELD ITERATIVE
```

and

```
POPULATION HIRSHFELD FRACTIONAL
```

yield the same iterative Hirshfeld charges for closed-shell molecules but different charges for open-shell molecules.

#### 4.9.2 Keyword DOS

This keyword specifies the printing of the squares of the elements of the orthonormalized MO eigenvectors which are essential ingredients for generating a density of states.

Options:

**MOS=<Integer1>-<Integer2>** Specification of orbital range for DOS calculation, first MO through last MO.

Description:

Use of the DOS keyword causes Löwdin-orthonormalized MO eigenvectors to be printed. As printed, each element is the square of the corresponding eigenvector component, *i.e.* the sum of the values of the printed eigenvector is unity. Thus, each element reflects the proportional contribution of an AO to the partial density represented by that one MO. A density of states (DOS) diagram can be created from these values. Printing of the vectors can be limited with the MOS option. The input MOS = <First> - <Last> specifies that only the MOs in the range <First> through <Last> are printed. Here <First> and <Last> refer to the (integer) indices of the molecular orbitals. For example, to print the vectors related to orbitals 10 through 25, the following input can be used:

```
DOS MOS=10-25
```

#### 4.9.3 Keyword SIGPI

With this keyword a  $\sigma - \pi$  energy separation will be performed.

Options: None

### Description:

The  $\sigma - \pi$  energy separation will be performed as described in [270]. The molecule must be planar to perform this calculation.

#### 4.9.4 Keyword LOCALIZATION

This keyword localizes user-defined  $\pi$ -groups in extended  $\pi$ -systems.

Options: None

### Description:

Each line in the keyword body defines a  $\pi$ -group. During the SCF the  $\pi$  interaction between these groups is suppressed. Thus, the energy difference with respect to the unconstrained calculation represents the resonance energy. The definition of the  $\pi$ -group consists of one or more atomic labels. These labels must be defined in the keyword body of GEOMETRY (see 4.1.1). For example, assume the following geometry definition for benzene:

GEOMETRY			
C1	0.000000	0.000000	0.000000
C2	0.000000	0.000000	1.394129
C3	1.207359	0.000000	-0.697056
C4	1.207341	0.000000	2.091222
C5	2.414698	0.000000	0.000038
C6	2.414694	0.000000	1.394163
H	-0.952713	0.000000	1.944171
H	-0.952718	0.000000	-0.550037
H	3.367421	0.000000	-0.549983
H	3.367408	0.000000	1.944198
H	1.207334	0.000000	3.191314
H	1.207389	0.000000	-1.797150

Then the following LOCALIZATION input defines three localized ethylene  $\pi$ -groups:

```
LOCALIZATION
C1 C2
C4 C6
C3 C5
```

Thus, the interactions between these  $\pi$ -groups are suppressed. The  $\sigma$ -orbitals are not affected. The LOCALIZATION keyword is applicable only to planar systems. However, the system orientation is arbitrary. **The LOCALIZATION and SYMMETRY keywords are incompatible and, therefore, cannot be used together in the input.**

## 4.10 Visualization and Topological Analysis

The keywords PLOT, ISOSURFACE, and GEOSURFACE of this section are mutually exclusive and **cannot** be combined. The same holds for the keywords BOX and POINTS.

### 4.10.1 Keyword VISUALIZATION

This keyword activates interfaces to the visualization programs MOLDEN [271], MOLEKEL [272], and XAIM [273].

Options:

#### **MOLDEN / MOLEKEL / WFN**

<b>MOLDEN</b>	The MOLDEN interface is activated. The MOLDEN input is written to the file <code>deMon.mol</code> . This is the default.
<b>MOLEKEL</b>	The MOLEKEL interface is activated. The MOLEKEL input is written to the file <code>deMon.mkl</code> .
<b>WFn</b>	The XAIM interface is activated. A WFn input is written to the file <code>deMon.wfn</code> .

#### **XYZ / OPT / MD / FULL / OFF**

<b>XYZ</b>	Only the molecular structures and energies of each optimization step are written to the <code>deMon.mol</code> file. This is the default if the options MOLDEN, MOLEKEL, or WFN are omitted.
<b>OPT</b>	The molecular structures as well as the forces and step sizes of each optimization step are written to the <code>deMon.mol</code> file. This option triggers the MOLDEN interface.
<b>MD</b>	The structures and energies of MD steps defined by the step interval option INT of the DYNAMICS keyword (4.7.1) are written to the <code>deMon.mol</code> file.
<b>FULL</b>	A full MOLDEN, MOLEKEL, or WFN input is written. This is the default if the option MOLDEN, MOLEKEL, or WFN is set.
<b>OFF</b>	The visualization interface is switched off.

#### **E=POTENTIAL / E=KINETIC / E=TOTAL / E=SYSTEM**

<b>E=POTENTIAL</b>	The potential energy is written to the <code>deMon.mol</code> file if the MD option is activated. This is the default.
--------------------	--

---

<b>E=KINETIC</b>	The kinetic energy is written to the deMon.mol file if the MD option is activated.
<b>E=TOTAL</b>	The total energy is written to the deMon.mol file if the MD option is activated.
<b>E=SYSTEM</b>	The system energy is written to the deMon.mol file if the MD option is activated. Only meaningful with the Nose-Hoover thermostat.

Description:

The options MOLDEN, MOLEKEL, and WFN are mutually exclusive. They activate the writing of the interface files deMon.mol, deMon.mkl, and deMon.wfn. With the options XYZ, OPT, MD, and FULL the output data are selected. By default, a MOLDEN interface file that contains only structural data and energies (XYZ option) is written. The option OFF of the VISUALIZATION keyword suppresses this output. Note that for large systems, the FULL option of the VISUALIZATION keyword produces very large interface files. We recommend use of the VU interface for the visualization of molecular fields of large systems (see PLOT and ISOSURFACE keywords below). In case of a frequency analysis the FULL option is activated automatically in order to animate the normal modes. With the MD option, a deMon.mol file is generated that contains energies and structures along the calculated trajectory. The step size for these snapshots is set by the INT option of the DYNAMICS keyword (see 4.7.1). The MD energy is selected by the E=POTENTIAL, E=KINETIC, E=TOTAL, and E=SYSTEM options. The deMon.wfn file contains molecular orbital information in terms of the primitive basis. This file format is used by several external programs for the topological analysis of molecular fields. For the deMon2k internal topological analysis modules see the keywords TOPOLOGY (4.10.2) and CPSEARCH (4.10.4).

#### 4.10.2 Keyword TOPOLOGY

This keyword controls parameters for the Bader population analysis; see Section (4.9.1).

Options:

**MEDIUM / COARSE / FINE / REFERENCE**

<b>MEDIUM</b>	Fixed, medium-grid accuracy is specified for a topological population analysis. This is the default.
<b>COARSE</b>	Fixed coarse-grid accuracy is specified for a topological population analysis.

**FINE** Fixed fine-grid accuracy is specified for a topological population analysis.

**REFERENCE** A reference grid accuracy is specified for a topological population analysis.

### **BASIS / AUXIS**

**BASIS** The Kohn-Sham density is used for the topological population analysis. This is the default.

**AUXIS** The auxiliary function density is used for the topological population analysis.

**DIRECT** The grid coordinates and weights are built on-the-fly for the topological population analysis. This is the default for parallel runs.

**MAX=<Integer>** Maximum number of iterations for attractor search or iterative Hirshfeld analysis. Default is 999.

**TOL=<Real>** Iterative Hirshfeld charge convergence criterion. Default is  $5 \cdot 10^{-4}$  atomic units.

Description:

This keyword specifies the grid and the density for topological population analyses (see 4.9.1). For more details about fixed grid please consult the keyword GRID (4.3.6). The TOPOLOGY keyword also allows changing the number of iterations and the convergence criterion for the iterative Hirshfeld analysis. In case of a Bader analysis the MAX option sets the maximum number of steps in the attractor search [39]. If the POPULATION keyword is absent the TOPOLOGY keyword triggers a Bader analysis by default.

### **4.10.3 Keyword PLOT**

This keyword controls the calculation and plotting of molecular fields. See the keywords BOX (4.10.7) and POINTS (4.10.8) for definitions of the plot support.

Options:

**<FIELD>** Molecular field specification. The available field acronyms are given in Table 12. This specification is mandatory, even though the syntax is that of an option!

### **BASIS / AUXIS**

**BASIS** The Kohn-Sham density is used for the construction of the plot function. This is the default.

**AUXIS** The auxiliary function density is used for the construction of the plot function. This option is incompatible with the READ option.

### **BINARY / ASCII / TABLE**

**BINARY** A binary file FIELD.bin is written in VU file format. It contains the coordinates and plot function values. The VU control file deMon.pie is written too. This is the default.

**ASCII** An ASCII file FIELD.asc is written. It contains the coordinates and plot function values.

**TABLE** A table with coordinates and plot function values is written to the output file deMon.out.

**READ** Specifies that an orbital list is read in the keyword body of PLOT. This keyword is incompatible with the EFIDS field.

**LMO** Localize molecular orbitals by Cholesky decomposition of the density matrix [274] before molecular fields are calculated. Only meaningful with the BASIS option.

**FBLMO** Localize molecular orbitals by Foster-Boys method [275] before molecular fields are calculated. Only meaningful with the BASIS option.

**PMLMO** Localize molecular orbitals by Pipek-Mezey method [276] before molecular fields are calculated. Only meaningful with the BASIS option.

**DD** The deformation density is used for the molecular field calculation.

#### Description:

Definition of the molecular field is mandatory for the PLOT keyword, even though the syntax is that of an option. The available molecular fields are listed in Table 12. If the BOX and POINTS keywords are omitted, the default BOX setting is used for the plot. Only one molecular field at a time can be calculated. If several fields are to be calculated, restart calculations with SCFTYPE MAX=0 and GUESS RESTART (see 4.5.1 and 4.5.5) may be done. In any case, no more than 20 molecular field entries at a time can be calculated. A field along with its first and second derivatives counts as 10 entries (1 + 3 + 6).

Table 12: Molecular fields available in deMon2k

Acronym	Description
<b>PSI</b>	Molecular orbitals. By default, the five highest energy occupied and five lowest energy unoccupied orbitals are calculated.
<b>D1PSI</b>	Molecular orbitals and first derivatives. By default, the HOMO and LUMO are calculated.
<b>D2PSI</b>	Molecular orbitals, first, and second derivatives. By default the HOMO is calculated.
<b>FUKUI</b>	Fukui function matrices. Only valid with the keyword FUKUI.
<b>RHO</b>	Electron density. By default, the density for the occupied orbitals is calculated.
<b>D1RHO</b>	Electron density and first spatial derivatives. By default, the density and its derivatives for the occupied molecular orbitals are calculated.
<b>D2RHO</b>	Electron density, first, and second spatial derivatives. By default, the density and its derivatives for the occupied orbitals are calculated.
<b>PRHO</b>	Perturbed densities in $x$ , $y$ , and $z$ electric field directions.
<b>SPIN</b>	Spin density. By default, the spin density for the occupied molecular orbitals is calculated.
<b>D1SPIN</b>	Spin density and first spatial derivatives. By default, the spin density and its derivatives for the occupied molecular orbitals are calculated.
<b>D2SPIN</b>	Spin density, first, and second spatial derivatives. By default, the spin density and its derivatives for the occupied orbitals are calculated.
<b>LAP</b>	Electron density Laplacian. By default, the Laplacian of the occupied molecular orbitals is calculated.
<b>ESP</b>	Molecular electrostatic potential. By default, the MESP of the occupied molecular orbitals is calculated. A minus sign in front indicates that only the electronic part is calculated.
<b>D1ESP</b>	Molecular electrostatic potential and its first derivatives. By default, the MEP and its derivatives of the occupied orbitals are calculated.
<b>D2ESP</b>	Molecular electrostatic potential and its first and second derivatives. By default, the MEP and its derivatives of the occupied orbitals are calculated.
<b>EFI</b>	Molecular electric field. By default, the electric field of the occupied molecular orbitals is calculated.
<b>D1EFI</b>	Molecular electric field and its first derivatives. By default, the electric field and its derivatives of the occupied orbitals are calculated.
<b>EFG</b>	Molecular electric field gradients. Identical to D1EFI.
<b>EFIDS</b>	Anion surface generation [277].
<b>ELF</b>	Electron localization function [278,279].
<b>ITELF</b>	Information theory derived electron localization function [280].
<b>IMF</b>	Induced magnetic field and NICS.
<b>SGGA</b>	Dimensionless density gradient $s$ as defined in the PBE functional.

The BASIS and AUXIS options specify the density (Kohn-Sham or auxiliary function den-

sity, respectively) used for the calculation of the molecular field. The AUXIS option is incompatible with the orbital specification by the READ option. Use of the auxiliary function density for the calculation of the molecular field will yield a considerable speedup. However, the field values may show deviations of as much as 10% compared to the Kohn-Sham density results. Most certainly these deviations depend on the chosen auxiliary function set (see 4.3.3). In general, the field topology obtained with the auxiliary function density is qualitatively correct.

The PLOT keyword produces two additional output files, the binary file FIELD.bin and the VU control file deMon.pie. With these files, a VU [41] session can be started. With the ASCII option, only the file FIELD.asc is created. The coordinates and molecular field values are listed in this file. With the TABLE option, the output is created but written to the deMon.out output file.

A set of molecular orbitals can be specified with the READ option for the calculation of molecular fields. However, this option cannot be used for the EFIDS field. In the case of the PSI and D1PSI fields, the READ option specifies the calculated molecular orbitals directly. Again, for PSI no more than 20 orbitals at a time can be calculated and, for D1PSI, no more than 5. If orbitals are selected in the case of other fields, then those fields are constructed from the sum of the selected orbitals. For example the following input specifies the calculation of the density using the molecular orbitals 1, 2, and 4:

```
PLOT RHO READ
1 2 4
```

Virtual orbitals can be included. They will be occupied by one electron. In the case of an unrestricted calculation (UKS option of keyword SCFTYPE; see 4.5.1),  $\alpha$  and  $\beta$ -orbitals can be selected, each set in a separate input line. In the following example, the  $\alpha$ -orbitals 1, 2, 3, 4, 5, 6, 7, and the  $\beta$ -orbitals 1, 2, 3, 4, 5, 6 are selected for the calculation of the molecular electrostatic potential:

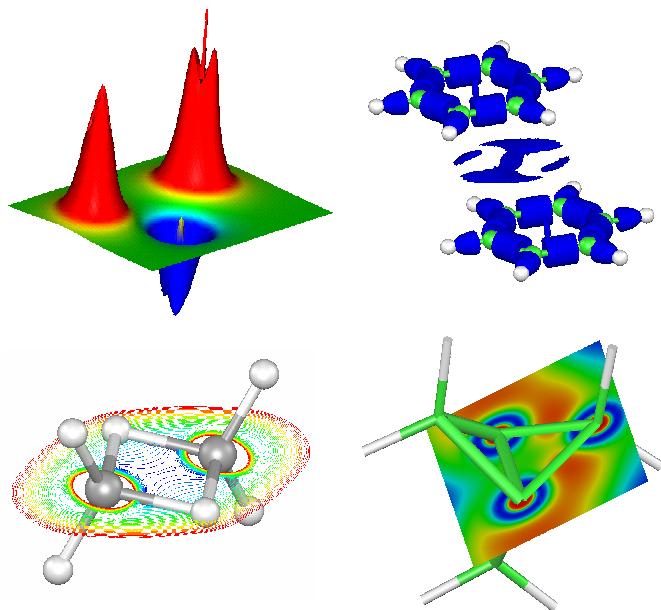
```
PLOT ESP READ
1 2 3 4 5 6 7
1 2 3 4 5 6
```

If only the  $\beta$ -orbitals are to be selected, the first line in the preceding example must have a 0 entry, as follows:

```
PLOT ESP READ
0
1 2 3 4 5 6
```

With the options LMO, FBLMO, and PMLMO the localization of molecular orbitals is requested. These localized orbitals can either be plotted (in combination with the PSI field option) or used for the calculation of molecular fields. Again, fields can be calculated with subsets of localized orbitals. For the electrostatic potential and its derivatives, it is also possible to plot only the electronic part by using minus signs in front of the corresponding field acronyms, *i.e.* -ESP, -D1ESP, and -D2ESP. Figure 14 depicts some molecular field plots obtained with the PLOT keyword and rendered with VU [41].

Figure 14: Plots of spin density for Cr<sub>3</sub>, (top left), dimensionless exchange gradient for two C<sub>6</sub>H<sub>6</sub> (top right) and for B<sub>2</sub>H<sub>6</sub>, (bottom left), and IT-ELF for C<sub>5</sub>H<sub>6</sub> (bottom right).



The DD option of the PLOT keyword triggers the use of deformation densities calculated according to Eq. (4.28). The deformation density itself is plotted by:

```
PLOT RHO DD
```

With the IMF PLOT option the NICS, the induced magnetic field vectors, and the absolute induced magnetic field values at each point of the grid are calculated. The external magnetic field can be specified with the READ option:

```
PLOT IMF READ
1.0 0.0 0.0
#
Geometry
C1    0.000000    1.396400    0.000000
```

C2	0.000000	0.698200	1.209318
C3	0.000000	-0.698200	1.209318
C4	0.000000	-1.396400	0.000000
C5	0.000000	-0.698200	-1.209318
C6	0.000000	0.698200	-1.209318
H1	0.000000	2.479500	0.000000
H2	0.000000	1.239750	2.147310
H3	0.000000	-1.239750	2.147310
H4	0.000000	-2.479500	0.000000
H5	0.000000	-1.239750	-2.147310
H6	0.000000	1.239750	-2.147310

In this example the external field is oriented along the *x*-axis which is perpendicular to the benzene molecule defined in the *y,z*-plane. By default the external magnetic field is oriented along the *z*-axis.

#### 4.10.4 Keyword CPSEARCH

This keyword activates the critical point search of scalar molecular fields. For the definition of the search region, see the keywords BOX (4.10.7) and POINTS (4.10.8).

Options:

**<FIELD>** Scalar molecular field specification. The available field acronyms are RHO and ESP. See Table 12 for the acronym meanings. This specification is mandatory, even though the syntax is that of an option!

#### BASIS / AUXIS

**BASIS** The Kohn-Sham density is used for the calculation of the scalar molecular field. This is the default.

**AUXIS** The auxiliary function density is used for the calculation of the scalar molecular field. Only available in combination with the ESP option.

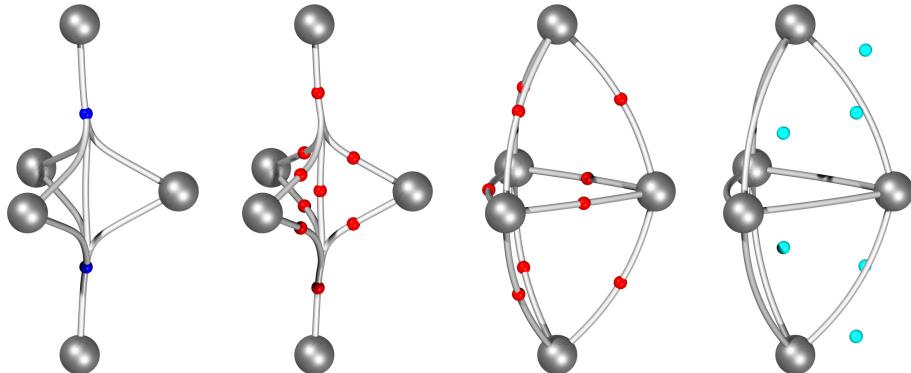
Description:

The critical points of the density (RHO) and molecular electrostatic potential (ESP) can be searched with this option. By default, the critical point search is performed in a marching cube-like style [281]. In it, the molecule is embedded in a rectilinear box (BOX LARGE; see Section 4.10.7), which in turn is divided into subboxes. Each subbox is searched for a critical point [282]. The box size and shape can be manipulated by the BOX keyword.

As an alternative, start points for the critical point search can be supplied by the POINTS keyword (see Section 4.10.8). With the POINTS option, POLYGON start points between atom pairs and triples are generated automatically. This option is recommended for the

critical point search of the density. In addition to the critical point search, the molecular connectivity over (3,-1) critical points is generated [38]. Figure 15 depicts critical point search results for  $\text{Li}_5$ . The two left-hand graphs show the non-nuclear attractors and bond critical points of the electron density in this system. The molecular graph from the density is also depicted to guide the eye. The figures were rendered with VU [41]. The right-hand two graphs in Figure 15 depict the bond critical points and minima of the molecular electrostatic potential (MESP). Again the molecular graph, however now derived from the molecular electrostatic potential, is also depicted.

Figure 15: Non-nuclear attractors (left) and bond critical points (middle left) of the density in  $\text{Li}_5$ . For comparison, the bond critical points of the MESP (middle right) and MESP minima (right) are also depicted.



#### 4.10.5 Keyword ISOSURFACE

This keyword controls the calculation and plotting of molecular field isosurfaces. Only one isosurface at a time can be generated. For the generation of multiple isosurfaces, the restart specification for the keyword PLOT (Section 4.10.3) can be applied. See the keyword BOX (4.10.7) for the definition of the isosurface boundary.

Options:

- |                      |   |
|----------------------|---|
| <b>&lt;FIELD&gt;</b> | Molecular field specification. The available field acronyms are PSI, RHO, SPIN, LAP, ESP, EFIDS, ELF, ITEL, SGGA, and PRHO in form of PRHOX, PRHOY or PRHOZ. See Table 12 for the acronym meanings. This specification is mandatory, even though the syntax is that of an option! |
|----------------------|---|

**BASIS / AUXIS**

- BASIS**      The Kohn-Sham density is used for the construction of the isosurface.  
                 This is the default.
- AUXIS**      The auxiliary function density is used for the construction of the iso-  
                 surface. This option is incompatible with the READ option.

**LINEAR / BILINEAR / LOGARITHMIC**

- LINEAR**      Linear interpolation scheme for the isosurface construction. This is the  
                 default.
- BILINEAR**      Bilinear interpolation scheme for the isosurface construction.
- LOGARITHMIC**    Logarithmic interpolation scheme for the isosurface construction.

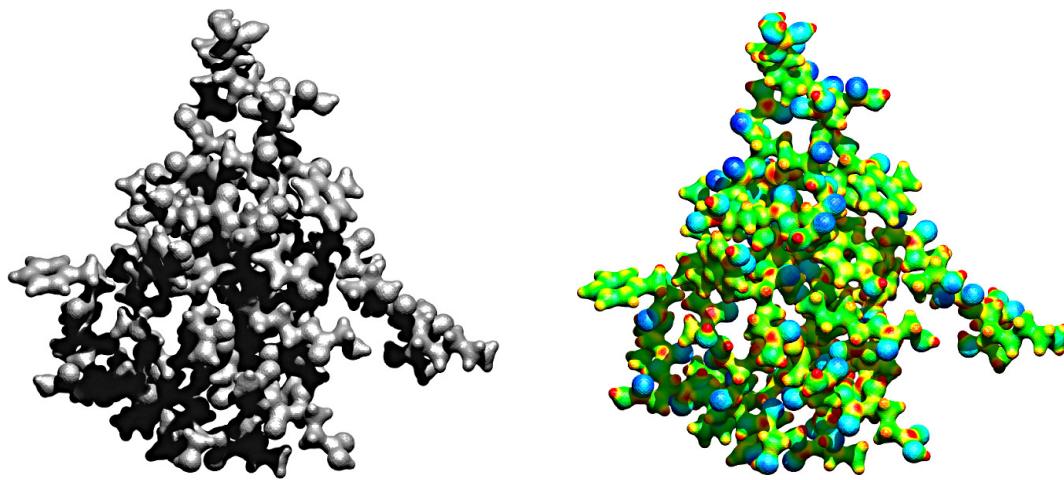
**BINARY / ASCII / TABLE**

- BINARY**      The isosurface is written to the binary file LAT.bin using the VU file  
                 format. The VU control file deMon.pie is written, too. This is the de-  
                 fault.
- ASCII**      An ascii output of the isosurface is written in the file deMon.lat.
- TABLE**      A function table of isosurface coordinates is written in the output file  
                 deMon.out.
- READ**      Specifies that an orbital list is read in the body of ISOSURFACE.
- ISO=<Real>** Isosurface value. This specification is mandatory!
- TOL=<Real>** Tolerance for data reduction.

Description:

The options AUXIS, BASIS, and READ of the ISOSURFACE keyword are identical in function to the corresponding options of the PLOT keyword (4.10.3). The BINARY option specifies that the isosurface coordinates and connectivities are written to the binary file LAT.bin. With the VU control file deMon.pie that is also generated, the isosurface grid can be visualized in VU [41]. It is important to note that the LAT.bin file can be used as an input file for the plotting of a molecular field on the isosurface! For this purpose, that file has to be in the same directory as the deMon.inp file and the POINTS keyword (see Section 4.10.8) must be used to define the isosurface in LAT.bin as plot support. Figure 16 shows the density isosurface of human insulin and the molecular electrostatic potential plotted on top of this isosurface.

Figure 16: Density isosurface of insulin, left, and the molecular electrostatic potential plotted on top of this isosurface, right.



With the ASCII option, the ascii file deMon.lat is generated. Here is an example of the data structure of this file:

```
RHO ( .100000E+00 A.U. ) ISOSURFACE COORDINATES IN ANGSTROM
```

```
NUMBER OF VERTICES: 1814
```

```
NUMBER OF FACETS: 3624
```

```
VOLUME: .163076E+04
```

```
AREA: .176664E+03
```

X	Y	Z
.312196E+00	-.729474E+00	.105260E+01
.243158E+00	-.765796E+00	.105260E+01
.243158E+00	-.729474E+00	.917119E+00
:	:	:

#### CONNECTIVITY

1	2	3
4	5	6
6	7	8
:	:	:

The first line is the file header including the field information (RHO), the isovalue (0.1 a.u.), and the units (ANGSTROM) used for the coordinates, area, and volume data. The NUMBER OF VERTICES value corresponds to the number of interpolation points for the isosurface. The NUMBER OF FACETS is the number of surface elements of the isosurface. The volume (here in  $\text{\AA}^3$ ) and the surface area (here in  $\text{\AA}^2$ ) then appear. Note that the volume includes all volume elements with field values below the given threshold. Therefore, the volume of a 0.1 a.u. isosurface of a field with positive and negative values includes all

volume elements with positive field values less than 0.1 a.u. **and** all volume elements with negative field values. The next block lists the coordinates of the interpolation points (here in Å). The last block, entitled CONNECTIVITY, describes the triangulation of the interpolation points. With the option TABLE, the same data as with the ASCII option are written into the file deMon.out.

The options LINEAR, BILINEAR, and LOGARITHMIC specify the interpolation scheme for the construction of the isosurface by use of the marching tetrahedron algorithm [281]. The value of the isosurface is specified with the ISO option. A value in vertical bar delimiters (e.g.,  $|0.1|$ ) indicates an absolute isosurface value. It can be used to generate two isosurfaces simultaneously with the same positive and negative threshold (e.g., for molecular orbitals). If the thresholds for the two isosurfaces should be different (e.g., for molecular electrostatic potentials), the two values can be specified as:

`ISO = 0.1/-0.05`

As noted above, the ISO specification has no default setting and, therefore, is mandatory, despite its syntax as an option! With the TOL option, the amount of data reduction is specified. Allowed values, from  $10^{-3}$  to 0.5, specify the severity of the data reduction. A value of  $10^{-3}$  indicates no data reduction at all. This is the default. With increasing TOL values, the data reduction increases, thus decreasing the memory requirements and, at the same time, the resolution.

Perturbed density isosurfaces are calculated with respect to the external electric field direction. Thus, the options PRHOX, PRHOY, and PRHOZ generate isosurfaces of the perturbed density originating from an external electric field applied along the  $x$ -axis,  $y$ -axis, or  $z$ -axis, respectively.

#### 4.10.6 Keyword GEOSURFACE

This keyword controls the calculation and plotting of geometrical surfaces such as spheres or ellipsoids. See the keyword BOX (4.10.7) for the definition of the geosurface boundary. Options:

##### **SPHERE / ELLIPSOID**

**SPHERE**            A sphere of radius  $a$  around the origin  $\vec{r}_o$  is constructed.

**ELLIPSOID**        An ellipsoid with semi-axes  $a$ ,  $b$ , and  $c$  around the origin  $\vec{r}_o$  is constructed.

##### **BINARY / ASCII / TABLE**

<b>BINARY</b>	A binary output of the geometrical surface is written in the file LAT.bin using the VU file format. The VU control file deMon.pie is also written. This is the default.
<b>ASCII</b>	An ascii output of the geometrical surface is written in the file deMon.lat.
<b>TABLE</b>	A function table of the geometrical surface coordinates is written in the output file deMon.out.
<b>TOL=&lt;Real&gt;</b>	Tolerance for data reduction.

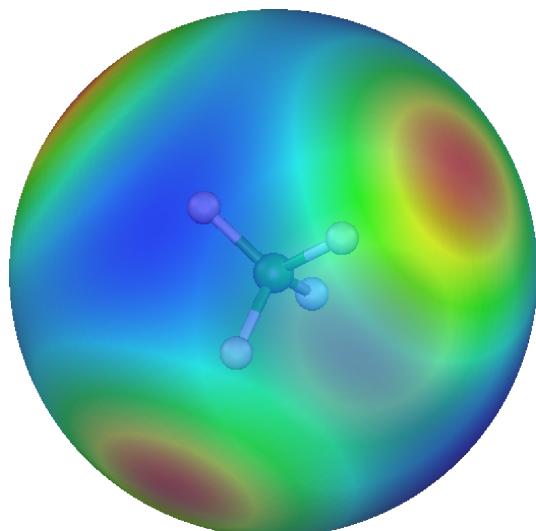
Description:

The options BINARY, ASCII, TABLE, and TOL are identical in definition to the corresponding options of the ISOSURFACE keyword. Despite the syntax, specification of SPHERE or ELLIPSOID is mandatory for the GEOSURFACE keyword. The origin and radius or semi-axes of the sphere or ellipsoid must be defined in the keyword body of GEOSURFACE. For a sphere with the origin at  $\vec{r}_o = (1, 3, 4)$  and a radius of 7 the input has the form:

```
GEOSURFACE SPHERE
1.0 3.0 4.0
7.0
```

The units are defined by the unit definition in the GEOMETRY keyword (see Section 4.1.1). The following input is used to generate the sphere, shown just below on the right, on which the molecular electrostatic potential is plotted:

```
GEOSURFACE SPHERE
0.0 0.0 0.0
3.0
#
BOX READ
 100 100 100
 -6.0 -6.0 -6.0
  6.0 -6.0 -6.0
 -6.0  6.0 -6.0
 -6.0 -6.0  6.0
#
Geometry Z-Matrix
C
H 1 RCH
H 1 RCH 2 AHCH
H 1 RCH 2 AHCH 3 AHCH RAA
H 1 RCH 2 AHCH 3 -AHCH RAA
VARIABLES
RCH    1.094
CONSTANTS
AHCH   109.47122063449
```



Note the explicit box definition that must enclose the sphere. For an ellipsoid with semi-axes 7.0, 1.0, and 0.5 in the current unit system and centered at  $\vec{r}_o = (1, 3, 4)$ , the following input form must be used:

```
GEOSURFACE ELLIPSOID
1.0 3.0 4.0
7.0 1.0 0.5
```

For the option ELLIPSOID the definition of all three semi-axes is mandatory.

#### 4.10.7 Keyword BOX

This keyword defines a box around a molecule for plotting (4.10.3), critical point search (4.10.4), and iso- and geosurface generation (4.10.5 and 4.10.6).

Options:

##### STANDARD / SMALL / LARGE

- |                 |  |
|-----------------|--|
| <b>STANDARD</b> | A box enclosing all defined atoms with an extension of twice their covalent radii. This is the default.                                    |
| <b>SMALL</b>    | A box enclosing all defined atoms with an extension of 1.2 times their covalent radii.   |
| <b>LARGE</b>    | A box enclosing all defined atoms with an extension of four times their covalent radii. This is the default for the critical point search. |

##### MEDIUM / COARSE / FINE

- |               |  |
|---------------|--|
| <b>MEDIUM</b> | The mesh spacing of the box is 0.3 Bohr. In the case of the critical point search it is 0.5 Bohr. This is the default. |
| <b>COARSE</b> | The mesh spacing of the box is 0.9 Bohr. In the case of the critical point search it is 1.0 Bohr.                      |
| <b>FINE</b>   | The mesh spacing of the box is 0.1 Bohr.   |
| <b>READ</b>   | The box definition is read from the keyword body.  |

Description:

The box size is specified with the options STANDARD, SMALL, and LARGE. The box is always created in STANDARD ORIENTATION (see 4.1.6) and, therefore, is symmetry-adapted to the molecule. The options MEDIUM, COARSE, and FINE define the mesh spacing within the box. The FINE mesh spacing will generate huge data files and thus is appropriate only for small systems. Via the READ option, the box can be explicitly defined

in the keyword body of BOX. If only a part of the molecule should be considered, the box can be defined simply by the atoms that should be included. In this case, the STANDARD, SMALL, and LARGE, as well as the MEDIUM, COARSE, and FINE options remain active. For a SMALL box with FINE resolution containing the H1 and O atoms of the water molecule, the input has the form:

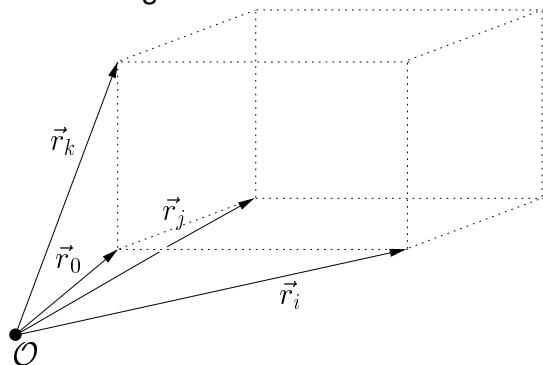
```
BOX SMALL FINE READ
O H1
```

Note that all atoms are defined in one input line. Instead of the explicit definition of each atom, elements can be used for the box definition, too. Therefore, a SMALL box with MEDIUM resolution containing only the hydrogen atoms of water can be defined as:

```
BOX SMALL MEDIUM READ
H
```

Alternatively, the plot box can be defined by vectors in the keyword body of BOX. In this case the options STANDARD, SMALL, and LARGE, as well as MEDIUM, COARSE, and FINE lose their meanings. The following format is used for the box definition via vectors:

```
BOX READ
NI NJ NK
X0 Y0 Z0
XI YI ZI
XJ YJ ZJ
XK YK ZK
```



The  $\vec{r}_0 = (x_0, y_0, z_0)$  position vector defines the origin of the box. The other three vectors  $\vec{r}_i = (x_i, y_i, z_i)$ ,  $\vec{r}_j = (x_j, y_j, z_j)$  and  $\vec{r}_k = (x_k, y_k, z_k)$  span the box as sketched above. NI, NJ and NK specify the number of points along the directions  $\vec{r}_i - \vec{r}_0$ ,  $\vec{r}_j - \vec{r}_0$  and  $\vec{r}_k - \vec{r}_0$ , respectively. Please note that these three vectors,  $\vec{r}_i - \vec{r}_0$ ,  $\vec{r}_j - \vec{r}_0$  and  $\vec{r}_k - \vec{r}_0$ , must be perpendicular to each other. In the case of a plane (two-dimensional box), definitions of NK, XK, YK, and ZK are not needed. For a line (one-dimensional box), only NI,  $\vec{r}_0$ , and  $\vec{r}_i$  must be defined. The explicit definition of a two-dimensional plot box is given in the following example:

```
PLOT LAPLACIAN
VXCTYP AUXIS VWN
#
GEOMETRY ANGSTROM
```

```
N 0.0 0.0 0.0
N 0.0 0.0 1.114553
#
BOX
200 200
0.0 -2.0 -2.0
0.0 2.0 -2.0
0.0 -2.0 3.114553
```

The definition of a one-dimensional plot box is described in Example 2.26 on page 35 of the deMon2k Tutorial.

#### 4.10.8 Keyword POINTS

Points for the plotting of molecular fields (4.10.3) and starting points for the critical point search (4.10.4) can be specified with this keyword.

Options:

##### BINARY / ASCII / POLYGON / READ

- |                |  |
|----------------|--|
| <b>BINARY</b>  | The plot point coordinates and connectivities are read from the external binary file LAT.bin. This is the default. |
| <b>ASCII</b>   | The plot point coordinates and connectivities are read from the external ascii file LAT.asc.                       |
| <b>POLYGON</b> | Starting points for the critical point search are generated by the polygon algorithm.                              |
| <b>READ</b>    | The plot point coordinates are read from the file deMon.inp.   |

Description:

The option BINARY causes coordinates and connectivities to be read from the file LAT.bin. Usually this file has been produced by a previous run which calculated isosurfaces or geometrical surfaces (see Sections 4.10.5 and 4.10.6). In this way, molecular fields can be plotted on isosurfaces. Example 2.27 on page 35 of the deMon2k tutorial describes the plotting of the electrostatic potential on the isodensity surface of benzene. The LAT.bin file and the corresponding plot output can be visualized with VU [41]. With the ASCII option, plot point coordinates, and optionally, connectivities, can be read from the external ascii file LAT.asc. The file format is:

```
1814
.312196E+00    -.729474E+00     .105260E+01
```

```
.243158E+00   -.765796E+00    .105260E+01
.243158E+00   -.729474E+00    .917119E+00
:
:
1      2      3
4      5      6
6      7      8
:
:
```

The first integer (here 1814) denotes the number of points. For each point, a line is then given with its coordinates ( $x$ ,  $y$  and  $z$ ). The integer triples after the coordinate specifications are the connectivities. This part of the input is optional. The LAT.asc file can be generated easily from the deMon.lat file (see 4.10.5) and used directly for visualization with VU. With the POLYGON option, starting points for the critical point search are generated automatically. In the case of the electronic density, this algorithm usually generates a sufficient set of starting points to find all critical points. **This is not the case for the critical points of other molecular fields.** The READ option can be used to read point coordinates from the input file deMon.inp. The coordinates are given in free format in the keyword body of POINTS, one input line for each point.

## 4.11 Performance and Accuracy Settings

### 4.11.1 Keyword PARALLEL

The parallel processing mode can be changed with this keyword.

Options:

#### HXYZ / HSCF

<b>HXYZ</b>	Parallelization of frequency analysis over coordinates. This is the default.
<b>HSCF</b>	Parallelization of frequency analysis over the SCF.
<b>GRAM</b>	Store inverse Coulomb matrix in RAM.

Description:

In a parallel single-point calculation, the work is distributed inside the SCF. That is, the integral calculation and linear algebra are parallelized and therefore the various CPUs all work on the same SCF calculation. We call this parallelization over the SCF. For the calculation of second derivatives (see 4.8.1 and 4.6.5), a different parallel assignment is made. Here each CPU works alone on a distinct SCF cycle. We call this parallelization over the coordinates. This is the default for the numerical second derivative calculation. But such parallelization also means that each process uses a complete set of matrices and files, just as a calculation on a single CPU would. As a result, performance problems can occur on some computational architectures, mainly because of swapping and/or intensive I/O. In those cases, use of the option HSCF will change the parallelization scheme to the one which is used for a single-point SCF calculation, *i.e.* all CPUs work on the same SCF. That approach is less efficient and should be used only if necessary.

The GRAM option of the PARALLEL keyword triggers the in-core storage of the inverse Coulomb matrix [25] in frequency calculations. This increases the RAM demand of the calculation but reduces I/O operations. We recommend this option for parallel runs of smaller systems with less than 500 atoms. This option has no effect when the option HSCF is also active.

### 4.11.2 Keyword MATDIA

One specifies the matrix diagonalization technique with this keyword.

Options:

#### DSYEV / RS / D&C / JACOBI

---

<b>DSYEV</b>	LAPACK Householder diagonalization. This is the default in serial runs except for atomic calculations.
<b>RS</b>	EISPACK Householder diagonalization. This is the default in parallel runs except for atomic calculations.
<b>D&amp;C</b>	LAPACK divide and conquer diagonalization.
<b>JACOBI</b>	Jacobi diagonalization. This is the default for atomic calculations.

Description:

The fastest diagonalizer available in deMon2k is the divide and conquer diagonalization from LAPACK [283]. However, it may fail in special situations. Fail-safe diagonalizers in deMon2k are based on the Householder algorithm. The DSYEV diagonalizer is considerably faster than the RS diagonalizer. However, DSYEV is also more memory intensive. Therefore, it should be used only if enough RAM space is available. The JACOBI diagonalizer is much slower but produces very pure eigenvectors. It is used by default for atomic calculations.

#### 4.11.3 Keyword MATINV

The matrix inversion technique for the Coulomb matrix in the variational density fitting is selected with this keyword.

Options:

##### NUMERICAL / ANALYTICAL

<b>NUMERICAL</b>	Iterative calculation of the fitting coefficients is specified. This is the default.
<b>ANALYTICAL</b>	Analytical calculation of the fitting coefficients is specified.

##### SVD / LUD / ONE

<b>SVD</b>	Use singular value decomposition for the calculation of the inverse of the Coulomb matrix. This is the default.
<b>LUD</b>	Use Cholesky decomposition for the calculation of the inverse of the Coulomb matrix.
<b>ONE</b>	Use the identity matrix as initial approximation to the inverse of the Coulomb matrix. Only meaningful in combination with the NUMERICAL option.

---

<b>DIAGONAL</b>	Performs the calculation of the fitting coefficients in the diagonal basis.
<b>TOL=&lt;Real&gt;</b>	Tolerance for the singular value decomposition. The default is $10^{-6}$ .
<b>MAX=&lt;Integer&gt;</b>	Maximum number of iteration cycles for the iterative calculation of the fitting coefficients. The default is 5000.

Description:

Calculation of the Coulomb and, in the case of ADFT, exchange-correlation fitting coefficients requires the solution of linear equation systems with the dimension of the auxiliary function set. For very large systems this task can represent a computational bottleneck. With the options NUMERICAL and ANALYTICAL of the MATINV keyword, numerical or analytical algorithms [25,192] for the solution of the fitting equation systems are requested. Whereas analytic solutions must be initialized at each geometry, numerical solutions need only be initialized at the first geometry [284]. The actual algorithm is selected with the options SVD, LUD, or ONE. The SVD and LUD options work with both numerical and analytical algorithms. The option SVD specifies singular value decomposition of the Coulomb matrix. This involves the diagonalization of that matrix, either for the first geometry only, if used in combination with NUMERICAL, or for each geometry if used in combination with ANALYTICAL. SVD is particularly recommended if large auxiliary function sets such as GEN-A2\*, GEN-A3\*, or GEN-A4\* are used. With the option LUD, the inverse Coulomb matrix is calculated via Cholesky decomposition [285]. This approach avoids matrix diagonalization and is, therefore, well-suited for large systems with medium-sized auxiliary function sets such as A2 or GEN-A2. The option ONE is meaningful only for the numerical approach. Here the inverse Coulomb matrix is initialized as a unit matrix and then updated during the first SCF. This is the least memory-demanding approach available in deMon2k. The options DIAGONAL and TOL of the keyword MATINV are meaningful only in combination with the SVD option. In the DIAGONAL SVD approach, eigenvalues and eigenvectors are quenched [13] and the equation systems are always solved in this diagonal representation. This is the numerically most stable approach. Thus, we recommend the use of,

MATINV SVD DIAGONAL ANALYTICAL

if significantly large negative MinMax errors hamper SCF convergence. The SVD threshold can be modified with the TOL option. The default value of  $10^{-6}$  has proven to be robust for most applications. Reducing this value usually will complicate SCF convergence. **Note that calculations with different SVD thresholds are not compatible with each other!** The option MAX of the MATINV keyword specifies the maximum number of iterative fitting

cycles and is, therefore, only meaningful in combination with the option NUMERICAL. The convergence of the numerical fitting can be printed by using PRINT FIT (see 4.12.2) in the input.

#### 4.11.4 Keyword DAVIDSON

This keyword controls the DAVIDSON diagonalization during a TDDFT calculation. This keyword has no effect if the keyword EXCITATION is not active.

Options:

- BAS=**<Integer> The maximum size of the basis to be used in the iterative process, *i.e.*, the size of the matrix to be diagonalized. The default value is 30.
- EIG=**<Integer> The number of the lowest excited states to be used in the iterative process. The default value is 10.
- MAX=**<Integer> The maximum number of iterations to be used in the iterative process. The default value is 500.
- TOL=**<Real> The convergence criterion. The default value is  $10^{-5}$ .

Description:

In deMon2k the block Davidson algorithm [286] is used for finding the lowest eigenvalues and eigenvectors of the large TDDFT matrix.

#### 4.11.5 Keyword WEIGHTING

This keyword specifies the weight function for the grid partitioning.

Options:

##### SCREENED / BECKE

- SCREENED** A screened weight function is used. This is the default.
- BECKE** The original Becke weight function is used. This option also activates the calculation of weight derivatives.

Description:

The BECKE weight function [287] introduces a cubic scaling into the grid generation [288]. By introducing a new piece-wise defined weight function [118], near linear scaling for the grid generation can be obtained. This weight function is specified with the option SCREENED.

#### 4.11.6 Keyword QUADRATURE

With this keyword the three-dimensional numerical quadrature scheme can be modified.

Options:

##### **GAUSS / EULER / MURA**

<b>GAUSS</b>	Gauss-Chebyshev radial quadrature is selected.
<b>EULER</b>	Euler-MacLaurin radial quadrature is selected.
<b>MURA</b>	Mura-Knowles radial quadrature [289] is selected.

##### **RANDOM / NORANDOM**

<b>RANDOM</b>	Random rotation of the Lebedev grids is activated.
<b>NORANDOM</b>	Random rotation of the Lebedev grids is deactivated.

<b>REFERENCE</b>	A reference grid is generated.
------------------	--------------------------------

Description:

With the options GAUSS, EULER, and MURA, the radial quadrature for the numerical integration [21,118] is selected. By default, the Gauss-Chebyshev radial quadrature is used for adaptive grids, while the Euler-MacLaurin radial quadrature is used for fixed grids. The options RANDOM and NORANDOM activate or deactivate the random rotations of the angular Lebedev grids. By default, random rotations are activated except for electrostatic property calculations. The option REFERENCE of the QUADRATURE keyword activates weight derivatives and deactivates grid screening. Thus, it also switches to BECKE weights (see 4.11.5). For the case of adaptive grids, it triggers adaptive grid generation for each atom in the molecule. As a result, this option might slow down the grid generation in larger systems.

#### 4.11.7 Keyword CFPINTEGRATION

Accuracy setting for the numerical radial integration in ECP and MCP integrals.

Options:

##### **MEDIUM / COARSE / FINE / REFERENCE**

<b>MEDIUM</b>	Numerical radial integration with 99 quadrature points. This is the default.
<b>COARSE</b>	Numerical radial integration with 67 quadrature points.
<b>FINE</b>	Numerical radial integration with 131 quadrature points.

**REFERENCE** Numerical radial integration with 195 quadrature points.

Description:

The keyword CFPINTEGRATION specifies the number of quadrature points for the numerical integration of local ECP and MCP integrals [158,159]. In the case of ECPs, it also defines the convergence threshold for the numerical radial integration of the semi-local integrals. These thresholds are  $10^{-7}$ ,  $10^{-8}$ ,  $10^{-9}$  and  $10^{-10}$  for the COARSE, MEDIUM, FINE, and REFERENCE option, respectively.

## 4.12 Miscellaneous Keywords

### 4.12.1 Keyword TITLE

With this keyword the job title is specified.

Options: None

The job title is limited to 60 characters and the title line **cannot** be continued.

### 4.12.2 Keyword PRINT

Optional printing can be activated with this keyword.

Options:

<b>AUXIS</b>	Print auxiliary function table.
<b>BASIS</b>	Print molecular orbitals in long format.
<b>C6</b>	Print C <sub>6</sub> coefficients.
<b>CGTO</b>	Print GTO contraction table.
<b>CIS</b>	Print CIS three-center electron repulsion integral batching.
<b>CON</b>	Print MD constraints.
<b>COORD</b>	Print primitive internal coordinates.
<b>CP</b>	Print critical points during Bader population analysis.
<b>DE2</b>	Print Hessian matrix.
<b>DEBUG</b>	Generate debug output.
<b>ECP</b>	Print ECP table.
<b>EMBED</b>	Print embedding charges.
<b>ERIS</b>	Print three-center electron repulsion integrals.
<b>FIT</b>	Print preconditioned conjugate gradient cycles.
<b>FUKUI</b>	Print Fukui matrices.
<b>G</b>	Print auxiliary function Coulomb matrix.
<b>GRID</b>	Generate full grid output.
<b>GTO</b>	Print primitive GTO table.
<b>IRC</b>	Print verbose IRC output.
<b>KS</b>	Print Kohn-Sham and core Hamiltonian matrix.
<b>MAX=&lt;Integer&gt;</b>	Maximum number of SCF cycles with print output.
<b>MCP</b>	Print MCP table.
<b>MD</b>	Print verbose MD simulation output.
<b>MOE</b>	Print molecular orbital energies and occupations.

<b>MOS</b>	Print molecular orbital energies, occupations and coefficients.
<b>NMR</b>	Print full NMR output.
<b>NOGEO</b>	Suppress geometry output.
<b>MAG</b>	Print full magnetizability output.
<b>NIA</b>	Print NIA CPKS cycles.
<b>NQR</b>	Print full NQR output.
<b>OPT</b>	Generate full optimization output.
<b>ORTHO</b>	Print orthogonalization matrix $S^{-1/2}$ .
<b>P</b>	Print density matrix.
<b>POPAN</b>	Generate full population analysis output.
<b>R</b>	Print response matrix.
<b>RAM</b>	Print RAM allocation table.
<b>S</b>	Print overlap matrix.
<b>SADDLE</b>	Print verbose saddle-point interpolation output.
<b>SCF</b>	Print SCF cycles which are usually omitted during, for example, optimization and frequency analysis.
<b>SYMMETRY</b>	Generate full symmetry output.
<b>T</b>	Print kinetic energy matrix.
<b>TB</b>	Print tight-binding matrices.
<b>TD</b>	Print full TDDFT output.
<b>VERBOSE</b>	Print verbose information regarding the SVD procedure.
<b>XCE</b>	Generate full exchange-correlation energy output.
<b>XCV</b>	Generate full exchange-correlation potential output.
<b>XRAY</b>	Generate full x-ray spectroscopy output.

#### Description:

The MAX option specifies the number of SCF cycles with active print output. All matrices that are changing during the SCF (e.g. P, KS, MOS, etc.) can be printed for each cycle. The following example activates the printing of the density matrix in the first 10 SCF cycles:

```
PRINT MAX=10 P
```

The printing of molecular orbitals can be limited with the MOS option. The input MOS = <First> - <Last> specifies that only the MOs in the range from <First> through <Last> are printed. Here, <First> and <Last> refer to the (integer) indices of the molecular orbitals. With the MOE option, printing of the molecular orbital coefficients can be suppressed. This is recommended for larger basis sets if only MO energies are of interest. To

print the orbital energies, occupation numbers, and coefficients of the orbitals 10 through 25 the following input can be used:

```
PRINT MOS=10-25
```

If only orbital energies and occupation numbers are of interest the input can be modified to:

```
PRINT MOE=10-25
```

In general, several print options such as MD, OPT, etc. will produce rather large outputs. In particular, the DEBUG option enables the printing of all matrices! Of course such an output is huge, even for small systems. Note also that the DEBUG option and some other PRINT options are only available in serial runs.

## A Automatic Generation of Auxiliary Functions

Automatically generated auxiliary function sets are provided in deMon2k by means of the auxiliary function specification GEN- $A_n$ , with  $n = 1, 2, 3, 4$ , and GEN- $A_n^*$ , with  $n = 2, 3, 4$  [133,156]. The GEN-A1 set possesses only  $s$  auxiliary functions and is usually used only for debugging. The GEN- $A_n$  sets consist of  $s$ ,  $p$ , and  $d$  Hermite Gaussian functions. In addition, the GEN- $A_n^*$  also have  $f$  and  $g$  Hermite Gaussians. Because the auxiliary functions are used to fit the electron density they are grouped in  $s$ ,  $spd$ , and  $spdfg$  sets. The exponents are shared within each of these sets [53,54]. Therefore, the auxiliary function notation (3,2,2) describes 3  $s$  sets with a total of 3 functions, 2  $spd$  sets with a total of 20 functions and 2  $spdfg$  sets with a total of 70 functions (see also 4.3.3). The range of exponents of all automatically generated auxiliary functions is determined by the smallest,  $\beta_{\min}$ , and largest,  $\beta_{\max}$ , primitive Gaussian exponent of the specified orbital basis set. Therefore, the GEN- $A_n$  and GEN- $A_n^*$  auxiliary function sets differ for different orbital basis sets. The number of exponents  $N$  (auxiliary function sets) is given by:

$$N = \text{Int} \left( \frac{\ln(\beta_{\max}/\beta_{\min})}{\ln(6-n)} + 0.5 \right) \quad (\text{A.1})$$

Here  $n$  is 1, 2, 3, or 4 according to the index of the selected GEN- $A_n$  or GEN- $A_n^*$  set. The exponents are generated in almost even-tempered form (see details below) and split into  $s$ ,  $spd$  and, if a GEN- $A_n^*$  set is specified,  $spdfg$  sets. The tightest (largest) exponents are assigned to the  $s$  sets, followed by the  $spd$ , and, if specified,  $spdfg$  sets. The basic exponent from which the generation starts is defined as:

$$\beta_o = 2 \beta_{\min} (6-n)^{(N-1)} \quad (\text{A.2})$$

In case ECPs or MCPs are used, equation (A.2) changes to:

$$\beta_o = 2 \beta_{\min} (6-n)^N \quad (\text{A.3})$$

Furthermore, only  $spd$  and  $spdfg$  sets are generated on ECP and MCP centers. From  $\beta_o$  the two tightest  $s$ , or  $spd$  set exponents,  $\beta_1$  and  $\beta_2$  are generated according to the formulas:

$$\beta_1 = \left( 1 + \frac{n}{12 - 2n} \right) \beta_o \quad (\text{A.4})$$

$$\beta_2 = \frac{\beta_o}{6-n} \quad (\text{A.5})$$

The other *s* or *spd* set exponents are generated according to the even-tempered progression:

$$\beta_{i+1} = \frac{\beta_i}{6 - n} \quad (\text{A.6})$$

The  $\beta_o$  exponent of the subsequent *spd* sets is also generated according to the progression (A.6). Based on this  $\beta_o$  exponent, the exponents of the first two *spd* sets are calculated with the formulas (A.4) and (A.5). The subsequent *spd* set exponents are calculated again according to the even-tempered progression (A.6). In the same way, the *spdfg* set exponents are calculated. In the case of *3d* elements, an extra diffuse *s* auxiliary function set is added.

## B Test Set for deMon2k Installation

A test set is provided with the distribution for the validation of a deMon2k installation. These tests are located in the directory `$CREX_ROOT/deMon/tests`. For each test, a subdirectory with the name `test<No>.d` exists. Here `<No>` refers to the number (index) of the test. Each subdirectory contains an input file (`test<No>.inp`) as well as a MOLDEN file (`TEST<No>.mol`), a new input file (`TEST<No>.new`), and an output file that was obtained by a deMon2k run of the original input. Within the usual accuracy limits (variation among machines and compilers), these output files should be reproduced by any deMon2k installation. The following table lists the tests with a short description of the test type and the input molecule used for the test. This information is also available in updated form in the `README` file of the `tests` directory. Please, in communications regarding bugs, always refer to the test number given below and append input and output files of your run.

Table 13: Content of the `tests` directory.

No.	Test type	Input molecule
1	Input syntax	<chem>H2O</chem>
2	Primitive coordinates	<chem>C2H2</chem>
3	Cartesian dummy atom	<chem>C2H2</chem>
4	Add dummy atom to Z-Matrix	<chem>HCN</chem>
5	Coordinate substitution RRA	<chem>H2O2</chem>
6	Redundant optimization with RRA	<chem>H2O2</chem>
7	Coordinate substitution RAA	<chem>H2O2</chem>
8	Redundant optimization with RAA	<chem>H2O2</chem>
9	Generation of Z-Matrix	<chem>C6H2</chem>
10	Symmetry adapted input	<chem>C6H6</chem>
11	Symmetry analyzer	<chem>C8H8</chem> (Cubane)
14	Mixed input	<chem>CO@MgO</chem>
16	Basis set substitution	<chem>HCN</chem>
17	User defined basis set	<chem>O2</chem>
18	Charged molecule	<chem>[B12H12]2-</chem>
19	Exchange of molecular orbitals	<chem>N2</chem>
20	Projection of molecular orbitals	<chem>[Cu(H2O)]2+</chem>
21	Fractional occupation	<chem>Cu3</chem>
27	Spherical (5d) orbitals	<chem>Fe(C5H5)2</chem>
28	Cartesian (6d) orbitals	<chem>Fe(C5H5)2</chem>
29	CONVENTIONAL ERI calculation	<chem>C24H50</chem>
30	DIRECT ERI calculation	<chem>C24H50</chem>
31	MULTIPOLE ERI calculation	<chem>C24H50</chem>
34	OMA mixing	<chem>CrF6</chem>

Table 13: Content of the tests directory (continuation).

No.	Test type	Input molecule
35	Dynamical level SHIFT	<chem>Cr2P5(CpMe5)2</chem>
75	GRID/DIRECT/PRINT calculation	<chem>CH3COCH3</chem>
77	GRID/DIRECT/VWN calculation	<chem>C6H5NHCOCH3</chem>
78	EMBEDDING calculation	<chem>NaCl@324</chem>
80	Semilocal ECP integral derivatives	<chem>Au4</chem>
81	ECP, BASIS and AUXIS pointers	<chem>AuH</chem>
82	ECP frequency analysis	<chem>Au4</chem>
83	Numerical integration of local ECP integrals	<chem>Fe(CO)5</chem>
85	MCP gradients	<chem>Cu2</chem>
86	FREEZE CORE with MULTIPOLE ERIS	<chem>C24H50</chem>
94	Linear dependencies in basis set	<chem>H2O</chem>
95	Full population analysis	<chem>H2O</chem>
96	OLYP/BASIS exchange-correlation	<chem>NiH</chem>
97	OLYP/AUXIS exchange-correlation	<chem>NiH</chem>
98	Linear basis set dependencies	<chem>H2O</chem>
100	Optimization with empirical dispersion term	<chem>(CH3-CH2-CH3)2</chem>
103	Numerical polarizabilities with AUXIS	<chem>CO</chem>
117	VERBOSE molecular dynamics	<chem>Na2</chem>
121	Generate MD trajectory	<chem>Na2</chem>
150	Unrelaxed potential energy surface scan	<chem>HCN</chem>
151	Adiabatic potential energy surface scan	<chem>HCN</chem>
161	$\pi$ localization with $\sigma$ - $\pi$ energy separation	<chem>C6H6</chem>
165	Analytical dynamical polarizability	<chem>NH2</chem>
242	Saddle interpolation	<chem>HCN</chem>
243	Hierarchical TS search	<chem>HCN</chem>
257	Chemical shifts of all centers; standard output	<chem>CH4</chem>
258	Chemical shifts with tensor output in $C_s$ symmetry	<chem>NH3</chem>
259	Chemical shifts of selected centers; tensor output	<chem>H2O</chem>
270	Polarizability with analytic VWN kernel	<chem>H2O</chem>
271	Polarizability with numerical VWN kernel	<chem>H2O</chem>
272	Polarizability with finite field (VWN)	<chem>H2O</chem>
283	Bader analysis with non-nuclear attractors	<chem>NH3</chem>
284	Bader analysis with very close CPs	<chem>C2H2</chem>
285	Bader analysis with full output	<chem>C4H8</chem>
291	Bondpath generation with non-nuclear attractors	<chem>Li5</chem>
293	Dynamic closed shell VWN polarizability	<chem>H2O</chem>
331	Forward IRC with calculated Hessian	<chem>HCN \rightarrow CNH</chem>
443	BOMD with Nose-Hoover chain Thermostat	<chem>H2O</chem>
455	Second row 1s XRAY calculation	<chem>H2S</chem>
464	Fukui reactivity analysis with Loewdin	<chem>C2H2</chem>
488	GGA Magnetizability with BASIS/GRID DIRECT	<chem>CH3NH2</chem>
507	XAS S1s alpha with S and O augmentation	<chem>SO2</chem>
514	XES calculation with one MO	<chem>H2O</chem>

Table 13: Content of the tests directory (continuation).

No.	Test type	Input molecule
522	Optimization and vibrational thermo corrections	CH <sub>4</sub>
536	Alignment with 5 pattern and enantiomers	Na <sub>10</sub>
548	Iterative Hirshfeld population analysis	LiH
609	CHARMM QM/MM step of RNA fragment	RNA
612	BOMD group constraints	Glycerol/Water
648	Average bond length analysis of BOMD trajectory	Glycerol
680	MM Cartesian optimization, freq and thermo	(H <sub>2</sub> O) <sub>4</sub>
703	QM/MM Cartesian opt. with QM vdW	Glycin@H <sub>2</sub> O
714	QM/MM Internal Z-Matrix optimization	(H <sub>2</sub> O) <sub>4</sub>
720	QM/MM opt. of mixed input with Cart. const.	(H <sub>2</sub> O) <sub>4</sub>
806	Nuclear spin-rotation constants	SO <sub>2</sub>
826	Foster-Boys MO localization	3-CH <sub>3</sub> -C <sub>5</sub> H <sub>11</sub>
834	FOCK/OPT+FREQ/ERIS DIRECT/GEN-A2*	CH <sub>4</sub>
849	Fock optimization with generated Z-Matrix	C <sub>6</sub> H <sub>2</sub>
853	M06 optimization	CH <sub>3</sub> COCH <sub>3</sub>
861	PBE0/OPT+FREQ+ALPHA/GEN-A2*	CH <sub>4</sub>
874	B3LYP XAS transition state for O1s alpha	H <sub>2</sub> O

## C MHM Program for deMon2k Trajectory Files

This appendix describes the compilation and input syntax of the "ttbin" implementation of the multiple histogram method (MHM) version 2.5. This program is intended to analyze the trajectory files created by the deMon2k software and will not work with trajectory files created by other programs. The program can be obtained directly from its author, Dr. José Manuel Vásquez Pérez (email: josemanuel\_vasquez@uaeh.edu.mx).

### Compilation

Change to the directory where the README.pdf file is located and locate the `ttmake.sh` script. Ensure that the gfortran compiler is installed and that the script is executable. To compile just type:

```
./ttmake.sh
```

The script will compile the source files producing the object files and the binary `ttbin` inside the newly created `object` directory. By default the script uses the gfortran compiler but you can switch to any FORTRAN compiler which supports the FORTRAN 90 standard by changing it.

### Input File

The input file should indicate at least the number of atoms in the molecule or cluster and the file names of the trajectories to analyze. The "#" character may be used at the beginning of a line to comment it. The list of possible keywords in the input is given below:

**ATOMS <Integer>**

Number of atoms in the system. This keyword is mandatory, if omitted an error is indicated.

**BINS <Integer>**

This keyword sets the number of bins into which the energy interval is divided for the histogram construction. If the keyword is omitted the default is 1000.

**TNODES <Integer>**

This keyword sets the number of nodes into which the temperature interval is divided for the calculation of thermodynamic functions. If the keyword is omitted the default is 100.

**TOLERANCE <Real>**

This keyword sets the convergence tolerance for the MHM self consistent procedure. If the keyword is omitted the default is  $10^{-9}$ . The default value should be fine for most cases. You can tighten it if you suspect that convergence is not reached.

**EGROUND <Real>**

This keyword sets a reference energy such that all trajectory energies will be offset relative to it. This value should be set equal to the ground state energy of the system to be physically meaningful. If the ground state energy is unknown the special value LOWEST may be used, which sets this reference equal to the lowest energy found along the trajectory. The selection of this value will shift the absolute values of internal and free energies calculated by the program. If the keyword is omitted the default is LOWEST.

**QUANTUM <Logical>**

This keyword enables quantum corrections for the partition function. If enabled the FREQS block must also be defined, otherwise an error is indicated. If the keyword is omitted its default value is FALSE.

**CUMULATIVE <Logical>**

This keyword enables an alternative integration method based on cumulative histograms for the MHM procedure. If the keyword is omitted its default value is FALSE.

**READ <String>**

This keyword specifies the intended energy to read from the trajectory. If omitted the default is the total energy. The possible values for this keyword are:

TOTAL	Read the total energy (default)
POTENTIAL	Read only the potential energy
KINETIC	Read only the kinetic energy

**FILES**

This keyword opens a block to define the filenames of the trajectory files to analyze. After each filename you may indicate the first and last steps of each trajectory to be read. If

you do not indicate the last step to read it is set to the last step in the trajectory. If you do not indicate neither the first nor the last step, they are set to the first and last step in the trajectory. The block must be closed with the END keyword. For example:

```
FILES
<filename> [<integer>] [<integer>]
<filename> [<integer>] [<integer>]
<filename> [<integer>] [<integer>]
END
```

### FREQS

This keyword opens a block to define the list of normal mode frequencies (in  $\text{cm}^{-1}$ ), as calculated in a normal mode analysis of the system, to recover the correct quantum zero temperature limit of the partition function and its derivatives. The block must be closed with the END keyword. For example:

```
FREQS
<frequency>
<frequency>
<frequency>
END
```

The following is a simple input example as found in the `example` directory of the MHM distribution:

```
ATOMS 2
EGROUND -108.662165418
QUANTUM TRUE
#BINS 1000
#TNODES 100
#READ TOTAL
#CUMULATIVE FALSE
#TOLERANCE 1.0E-9

FILES
N2_0500.trj      5001
N2_1500.trj      5001
N2_2500.trj      5001
END

FREQS
2361.8
END
```

Besides this input file the specified trajectory files along with the expected output files are also in the `example` directory of the MHM distribution.

## Execution

Ensure that the `ttbin` binary is executable and run it in the same directory where the trajectory and input files are located by typing

```
./ttbin
```

The program will print the following messages during its execution:

Reading trajectory file...

Printed as trajectory files are read.

Solving self consistent equations...

Printed when the self consistent MHM procedure starts.

Calculating thermodynamic functions...

Printed when the partition function derivatives calculation starts.

Additionally, several warnings and errors may be printed.

A script named `ttplot.sh` is also included in the MHM distribution. It will plot the free energy, entropy, internal energy and heat capacity curves. Make sure that it is executable and run it in the same directory where the output files are located by typing:

```
./ttplot.sh [<real> <real>]
```

It will create a gnuplot script named `makeplots.gnu` and, if gnuplot is installed, will create the postscript files `f.ps`, `u.ps`, `cv.ps` and `s.ps` that contain plots of the free energy, internal energy, heat capacity and entropy, respectively. By default the whole temperature range of the trajectories is plotted, but the script accepts two optional arguments to set the lower and upper temperatures of the plots.

## Output Files

A successful execution of the program creates the following output files:

`output`

This file contains the settings and parameters used for the construction of the histograms by the multiple histogram method.

**trj-data**

This file contains statistics for each trajectory file. Internal energies and heat capacities are directly calculated from statistical fluctuations of the energy.

**can-data**

This file contains canonical properties as a function of temperature. It lists the free energy,  $F$ , entropy,  $S$ , internal energy,  $U$  and heat capacity,  $c_v$  as a function of temperature.

**his-data**

This file contains the energy histograms.

## References

- [1] P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- [2] W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [3] R.M. Dreizler, E.K.U. Gross, *Density Functional Theory* (Springer-Verlag, Berlin, 1990).
- [4] A. St-Amant, *deMon, un programme LCGTO-MCP-DF, et une etude théorique de l'adsorption de l'hydrogène sur les monomères et dimères de Ni, Rh, et Pd*, Ph.D. Thesis, Université de Montréal (1992).
- [5] B.I. Dunlap, J.W.D. Connolly, J.R. Sabin, *J. Chem. Phys.* **71**, 4993 (1979).
- [6] J.W. Mintmire, B.I. Dunlap, *Phys. Rev. A* **25**, 88 (1982).
- [7] J.W. Mintmire, J.R. Sabin, S.B. Trickey, *Phys. Rev. B* **26**, 1743 (1982).
- [8] H. Sambe, R.H. Felton, *J. Chem. Phys.* **62**, 1122 (1975).
- [9] M.E. Casida, C. Daul, A. Goursot, A.M. Köster, L.G.M. Pettersson, E. Proynov, A. St-Amant, D.R. Salahub, H. Duarte, N. Godbout, J. Guan, C. Jamorski, M. Leboeuf, V.G. Malkin, O.L. Malkina, F. Sim, A. Vela, *deMon-KS Version 3.4*, deMon Software, Montréal (1996).
- [10] V.G. Malkin, O.L. Malkina, M.E. Casida, D.R. Salahub, *J. Am. Chem. Soc.* **116**, 5898 (1994).
- [11] A.M. Köster, P. Calaminici, N. Russo, *Phys. Rev. A* **53**, 3865 (1996).
- [12] C. Jamorski, M.E. Casida, D.R. Salahub, *J. Chem. Phys.* **104**, 5134 (1996).
- [13] A.M. Köster, J.M. del Campo, F. Janetzko, B. Zúñiga-Gutiérrez, *J. Chem. Phys.* **130**, 114106 (2009).
- [14] L. Versluis, T. Ziegler, *J. Chem. Phys.* **88**, 322 (1988).
- [15] J. Andzelm, E. Wimmer, *Physica B* **172**, 307 (1991).
- [16] J. Guan, M.E. Casida, A.M. Köster, D.R. Salahub, *Phys. Rev. B* **52**, 2184 (1995).

- [17] M.E. Casida, C. Daul, A. Goursot, A.M. Köster, L.G.M. Pettersson, E. Proynov, A. St-Amant, D.R. Salahub, H. Duarte, N. Godbout, J. Guan, K. Hermann, C. Jamorski, M. Leboeuf, V.G. Malkin, O.L. Malkina, M. Nyberg, L. Pedocchi, F. Sim, L. Triguero, A. Vela, deMon Software (2001).
- [18] StoBe-deMon version 2.1, K. Hermann and L.G.M. Pettersson, M.E. Casida, C. Daul, A. Goursot, A.M. Köster, E. Proynov, A. St-Amant, and D.R. Salahub. Contributing authors: V. Caravetta, H. Duarte, C. Friedrich, N. Godbout, J. Guan, C. Jamorski, M. Leboeuf, M. Leetmaa, M. Nyberg, L. Pedocchi, F. Sim, L. Triguero, and A. Vela, StoBe Software (2005).
- [19] A.M. Köster, M. Krack, M. Leboeuf, B. Zimmermann, ALLCHEM, Universität Hannover (1998).
- [20] A.M. Köster, *J. Chem. Phys.* **104**, 4114 (1996).
- [21] M. Krack, A.M. Köster, *J. Chem. Phys.* **108**, 3226 (1998).
- [22] A.M. Koster, G. Geudtner, P. Calaminici, M.E. Casida, V.D. Dominguez, R. Flores-Moreno, G.U. Gamboa, A. Goursot, T. Heine, A. Ipatov, F. Janetzko, J.M. del Campo, J.U. Reveles, A. Vela, B. Zuniga-Gutierrez and D.R. Salahub, deMon2k, Version 3, The deMon developers, Cinvestav, Mexico-City (2011).  
See also <http://www.deMon-software.com>.
- [23] A.M. Köster, J.U. Reveles, J.M. del Campo, *J. Chem. Phys.* **121**, 3417 (2004).
- [24] G. Geudtner, F. Janetzko, A.M. Köster, A. Vela, P. Calaminici, *J. Comput. Chem.* **27**, 483 (2006).
- [25] V.D. Dominguez-Soria, G. Geudtner, J.L. Morales, P. Calaminici, A.M. Köster, *J. Chem. Phys.* **131**, 124102 (2009).
- [26] A. Goursot, T. Mineva, R. Kevorkyants, D. Talbi, *J. Chem. Theor. Comput.* **3**, 755 (2007).
- [27] P. Jurečka, J. Černý, P. Hobza, D.R. Salahub, *J. Comput. Chem.* **28**, 555 (2007).
- [28] Y. Zhang, A. Vela, D.R. Salahub, *Theor. Chem. Acc.* **118**, 693 (2007).

- [29] A. Vela, V. Medel, S.B. Trickey, J. Chem. Phys. **130**, 244103 (2009).
- [30] R. Flores-Moreno, A.M. Köster, J. Chem. Phys. **128**, 134105 (2008).
- [31] J.M. del Campo, A.M. Köster, J. Chem. Phys. **129**, 024107 (2008).
- [32] S. Krishnamurty, T. Heine, A. Goursot, J. Phys. Chem. B **107**, 5728 (2003).
- [33] G.U. Gamboa, P. Calaminici, G. Geudtner, A.M. Köster, J. Phys. Chem. A **112**, 11969 (2008).
- [34] G.U. Gamboa, J.M. Vásquez-Pérez, P. Calaminici, A.M. Köster, Int. J. Quant. Chem. **110**, 2172 (2010).
- [35] A. Ipatov, A. Fouqueau, C. Perez del Valle, F. Cordova, M.E. Casida, A.M. Köster, A. Vela, J. Mol. Structure: THEOCHEM **762**, 179 (2006).
- [36] P. Calaminici, K. Jug, A.M. Köster, J. Chem. Phys. **109**, 7756 (1998).
- [37] S. Shedge, J. Carmona-Espíndola, S. Pal, A.M. Köster, J. Phys. Chem. **114**, 2357 (2010).
- [38] R.F.W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, New York, 1990).
- [39] J.I. Rodríguez, A.M. Köster, P.W. Ayers, A. Santos-Valle, A. Vela, G. Merino, J. Comput. Chem. **30**, 1082 (2009).
- [40] F. Weinhold, C.R. Landis, *Valency and Bonding: A Natural Bonding Orbital Donor-Acceptor Perspective* (Cambridge University Press, Cambridge, 2003).
- [41] See <http://www.invisu.ca/>.
- [42] L. Triguero, L.G.M. Pettersson, H. Ågren, Phys. Rev. B **58**, 8097 (1998).
- [43] D. Mejía-Rodríguez, A.M. Köster, J. Chem. Phys. **141**, 124114 (2014).
- [44] D.R. Salahub, S.Y. Noskov, B. Lev, R. Zhang, V. Ngo, A. Goursot, P. Calaminici, A.M. Köster, A. Alvarez-Ibarra, D. Mejía-Rodríguez, J. Řezáč, F. Cailliez, A. de la Lande, Molecules **20**, 4780 (2015).

- [45] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives *J. Am. Chem. Soc.* **118**, 11225 (1996).
- [46] B. Lev, R. Zhang, A. de la Lande, D.R. Salahub, S.Y. Noskov, *J. Comput. Chem.* **31**, 1015 (2010).
- [47] O. Bertran, S.B. Trickey, J. Torras, *J. Comput. Chem.* **31**, 2669 (2010).
- [48] Cuby. Available online: <http://cuby.molecular.cz> (accessed 25 February 2015).
- [49] R.G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- [50] E.J. Baerends, D.E. Ellis, P. Ros, *Chem. Phys.* **2**, 41 (1973).
- [51] A.M. Köster, *J. Chem. Phys.* **118**, 9943 (2003).
- [52] S.B. Trickey, J.A. Alford, J.C. Boettger in *Computational Materials Science and Computational Chemistry*, Vol. 15 (Elsevier, Amsterdam, 2004).
- [53] J. Andzelm, E. Radzio, D.R. Salahub, *J. Comput. Chem.* **6**, 520 (1985).
- [54] J. Andzelm, N. Russo, D.R. Salahub, *J. Chem. Phys.* **87**, 6562 (1987).
- [55] A.M. Köster, A. Goursot, D.R. Salahub in *Comprehensive Coordination Chemistry-II: From Biology to Nanotechnology*, Vol. 1, Chapter 2.57, pp. 681, Editors: J. McCleverty, T.J. Meyer, B. Lever (Elsevier, Amsterdam, 2003).
- [56] D.R. Salahub, J. Weber, A. Goursot, A.M. Köster, A. Vela in *Theory and Applications of the Computational Chemistry: The First 40 Years*, Editors: C.E. Dykstra, G. Frenking, K.S. Kim, G. Scuseria (Elsevier, Amsterdam, 2005).
- [57] F. Janetzko, A. Goursot, T. Mineva, P. Calaminici, R. Flores-Moreno, A.M. Köster, D.R. Salahub in *Nanoclusters: A Bridge across Disciplines*, Editors: P. Jena, A. Castleman, Jr. (Elsevier, Amsterdam, 2010).
- [58] D.R. Salahub, P. Calaminici, G.U. Gamboa, A.M. Köster, J.M. Vásquez-Pérez, in *Metallic Systems: A Quantum Chemist's Perspective*, Editors: T.C. Allison, O. Coskuner, C.A. González (CRC Press, Boca Raton, 2011).

- [59] P. Calaminici, A. Alvarez-Ibarra, D. Cruz-Olvera, V.D. Dominguez-Soria, R. Flores-Moreno, G.U. Gamboa, G. Geudtner, A. Goursot, D. Mejía-Rodríguez, D.R. Salahub, B. Zuniga-Gutierrez, A.M. Köster, in *Handbook of Computational Chemistry*, Second Edition, Editor: J. Leszczynski (Springer, Berlin, Heidelberg, 2016)
- [60] D. Mejía-Rodríguez, *Low-Order Scaling Methods for Auxiliary Density Functional Theory*, Ph.D. Thesis, Cinvestav, Mexico-City (2015).
- [61] R. Polly, H.-J. Werner, F.R. Manby, P.J. Knowles, Mol. Phys. **102**, 2311 (2004).
- [62] A.D. Becke, J. Chem. Phys. **98**, 5648 (1993).
- [63] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M. Frisch, J. Chem. Phys. **98**, 11623 (1994).
- [64] J.P. Perdew, M. Ernzerhof, K. Burke, J. Chem. Phys. **105**, 9982 (1996).
- [65] C. Adamo, V. Barone, J. Chem. Phys. **110**, 6158 (1999).
- [66] Y. Zhao, D.G. Truhlar, Theor. Chem. Acc. **120**, 215 (2007).
- [67] D. Mejía-Rodríguez, X. Huang, J.M. del Campo, A.M. Köster, Adv. Quantum Chem. **71**, 41 (2015).
- [68] C. Quintanar, R. Caballero, A.M. Köster, Int. J. Quantum Chem. **96**, 483 (2004).
- [69] A. Alvarez-Ibarra, A.M. Köster, R. Zhang, D.R. Salahub, J. Chem. Theor. Comput. **8**, 4232 (2012).
- [70] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
- [71] T. Mineva, Y. Tsoneva, R. Kevorkyants, A. Goursot, Can. J. Chem. **91**, 529 (2013).
- [72] H.L. Woodcock, M. Hodošček, A.T.B. Gilbert, P.M.W. Gill, H.F. Schaefer, B.R. Brooks, J. Comput. Chem. **28**, 1485 (2007).
- [73] B. Lev, B. Roux, S.Y. Noskov, J. Chem. Theor. Comput. **9**, 4165 (2013).
- [74] A. Alvarez-Ibarra A.M. Köster, J. Chem. Phys. **139**, 024102 (2013).

- [75] E. Siever, A. Weber, S. Figgins, R. Love, A. Robbins, *Linux in a Nutshell* (O'Reilly Verlag GmbH & Co. KG, Köln, Germany).
- [76] H.C. Andersen, J. Comput. Phys. **52**, 24 (1983).
- [77] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
- [78] J.M. Vásquez-Pérez, G.U. Gamboa, A.M. Köster, P. Calaminici, J. Chem. Phys. **131**, 124126 (2009).
- [79] P.A.M. Dirac, Proc. Cambr. Phil. Soc. **26**, 376 (1930).
- [80] J.P. Perdew, Y. Wang, Phys. Rev. B **33**, 8800 (1986); **34**, 7406(E) (1986).
- [81] A.D. Becke, Phys. Rev. A **38**, 3098 (1988).
- [82] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
- [83] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); *ibid.* **78**, 1396(E) (1997).
- [84] Y. Zhang, W. Yang, Phys. Rev. Lett. **80**, 890 (1999).
- [85] B. Hammer, L.B. Hansen, J.K. Norskov, Phys. Rev. B **59**, 7413 (1999).
- [86] N.C. Handy, A.J. Cohen, Mol. Phys. **99**, 403 (2001).
- [87] J.P. Perdew, A. Ruzsinszky, G.I. Csonka, O.A. Vydrov, G.E. Scuseria, L.A. Constantin, X. Zhou, K. Burke, Phys. Rev. Lett. **100**, 136406 (2008).
- [88] S.H. Vosko, L. Wilk, M. Nusair, Can. J. Phys. **58**, 1200 (1980).
- [89] J.P. Perdew, A. Zunger, Phys. Rev. B **23**, 5048 (1981).
- [90] J.P. Perdew, Y. Wang, Phys. Rev. B **45**, 13244 (1992).
- [91] J.P. Perdew, Phys. Rev. B **33**, 8822 (1986).
- [92] B. Zimmermann, *Struktur und Stabilität kleiner Kupfercluster*, Ph.D. Thesis, Universität Hannover (1999).

- [93] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B **37**, 785 (1988).
- [94] R. Colle, D. Salvetti, Theor. Chim. Acta **37**, 329 (1975).
- [95] R. Colle, D. Salvetti, J. Chem. Phys. **79**, 1404 (1983).
- [96] E. Engel, J.A. Chevary, L.D. Macdonald, S.H. Vosko, Z. Phys. D **23**, 7 (1992).
- [97] E. Engel, S. Keller, R.M. Dreizler, Phys. Rev. A **53**, 1367 (1996)
- [98] E. Engel, S.H. Vosko, Phys. Rev. B **47**, 13164 (1993)
- [99] R. Leeuwen, E.J. Baereends, Phy. Rev. A. **49**, 2421 (1994).
- [100] T.W. Keal, D.J. Tozer, J. Chem. Phys. **119**, 3015 (2003).
- [101] T.W. Keal, D.J. Tozer, J. Chem. Phys. **121**, 5654 (2004).
- [102] R. Peverati, Y. Zhao, D.G. Truhlar, J. Phys. Chem. Lett. **2**, 1991 (2011).
- [103] R. Peverati, D.G. Truhlar, J. Chem. Theor. Comput. **8**, 2310 (2012).
- [104] H.S. Yu, W. Zhang, P. Verma, Z. He, D.G. Truhlar, Phys. Chem. Chem. Phys. **17**, 12146 (2015).
- [105] J. Carmona-Espíndola, J.L. Gázquez, A. Vela, S.B. Trickey, J. Chem. Phys. **142**, 054105 (2015).
- [106] T. van Voorhis, G.E. Scuseria, J. Chem. Phys. **109**, 400 (1998); *ibid.* **129**, 219901(E) (2008).
- [107] J.P. Perdew, S. Kurth, A. Zupan, P. Blaha, Phys. Rev. Lett. **82**, 2544 (1999).
- [108] J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
- [109] J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, J. Chem. Phys. **120**, 6898 (2004).
- [110] Y. Zhao, D.G. Truhlar, J. Chem. Phys. **125**, 194101 (2006).
- [111] R. Peverati, D.G. Truhlar, J. Phys. Chem. Lett. **3**, 117 (2012).
- [112] R. Peverati, D.G. Truhlar, Phys. Chem. Chem. Phys. **14**, 13171 (2012).

- [113] A.D. Becke J. Chem. Phys. **98**, 1372 (1993).
- [114] J.C. Slater, Phys. Rev. **81**, 385 (1951).
- [115] D. Mejíia-Rodríguez, A.M. Köster, J. Chem. Phys. **141**, 124114 (2014).
- [116] J. Gräfenstein, D. Izotov, D. Cremer, J. Chem. Phys. **127**, 214103 (2007).
- [117] B. Miehlich, A. Savin, H. Stoll, H. Preuss, Chem. Phys. Lett. **157**, 200 (1989).
- [118] A.M. Köster, R. Flores-Moreno, J.U. Reveles, J. Chem. Phys. **121**, 681 (2004).
- [119] V.I. Lebedev, Russian Acad. Sci. Dokl. Math. **50**, 283 (1995).
- [120] P.M.W. Gill, B.G. Johnson, J.A. Pople, Chem. Phys. Lett. **209**, 506 (1993).
- [121] S. Grimme, J. Comput. Chem. **25**, 1463 (2004).
- [122] A. Bondi, J. Phys. Chem. **68**, 441 (1964).
- [123] Q. Wu, W. Yang, J. Chem. Phys. **116**, 515 (2002).
- [124] A.K. Rappé, C.J. Casewit, K.S. Colwell, W.A. Goddard III, W.M. Skiff, J. Am. Chem. Soc **114**, 10024 (1992)
- [125] P. Ren, J.W. Ponder, J. Phys. Chem. B **114**, 2549 (1993).  
See also <https://dasher.wustl.edu/tinker>.
- [126] W.L. Jorgensen, N.A. McDonald, J. Phys. Chem. B **102**, 8049 (1998).
- [127] K.P. Jensen, W.L. Jorgensen, J. Chem. Theory Comput. **2**, 1499 (2006).
- [128] T.E. Cheatham, P. Cieplak, P.A. Kollman, J. Biomol. Struct. Dyn. **16**, 845 (1999).
- [129] J.W. Ponder, D.A. Case, Adv. Prot. Chem. **66**, 27 (2003).
- [130] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Proteins **65**, 712 (2006).
- [131] D.J. Feller, J. Comput. Chem. **17**, 1571 (1996); K.L. Schuchardt, B.T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, T.L. Windus, J. Chem. Inf. Model. **47**, 1045 (2007).  
See also: <https://bse.pnl.gov/bse/portal>

- [132] N. Godbout, D.R. Salahub, J. Andzelm, E. Wimmer, Can. J. Phys. **70**, 560, (1992).
- [133] P. Calaminici, F. Janetzko, A.M. Köster, R. Mejia-Olvera, B. Zúñiga-Gutiérrez, J. Chem. Phys. **126**, 044108 (2007).
- [134] J. Guan, P. Duffy, J.T. Carter, D.P. Chong, K.C. Casida, M.E. Casida, M. Wrinn, J. Chem. Phys. **98**, 4753 (1993).
- [135] D.P. Chong (private communication).
- [136] N. Rega, M. Cossi, V. Barone, J. Chem. Phys. **105**, 11060 (1996).
- [137] S. Huzinaga, J. Chem. Phys. **42**, 1293 (1965).
- [138] F. Jensen, J. Chem. Theory Comput. **2**, 1360 (2006).
- [139] W.J. Hehre, R.F. Stewart, J.A. Pople, J. Chem. Phys. **51**, 2657 (1969); W.J. Hehre, R. Ditchfield, R.F. Stewart, J.A. Pople, J. Chem. Phys. **52**, 2769 (1970).
- [140] P.C. Hariharan, J.A. Pople, Theor. Chim. Acta **28**, 213 (1973); M.M. Franci, W.J. Pietro, W.J. Hehre, J.S. Binkley, M.S. Gordon, D.J. DeFress, J.A. Pople, J. Chem. Phys. **77**, 3654 (1982).
- [141] R. Krishnan, J.S. Binkley, R. Seeger, J.A. Pople, J. Chem. Phys. **72**, 650 (1980).
- [142] A.J. Sadlej, Collection Czech. Chem. Commun. **53**, 1995 (1988).
- [143] G.C. Lie, E. Clementi, J. Chem. Phys. **60**, 1275 (1974).
- [144] A.J.H. Wachters, J. Chem. Phys. **52**, 1033 (1970).
- [145] P.-O. Widmark, B.J. Persson, B.O. Roos, Theor. Chim. Acta **79**, 419 (1991); R. Pou-Amérigo, M. Merchán, I. Nebot-Gil, P.-O. Widmark, B.O. Roos, Theor. Chim. Acta **92**, 149 (1995).
- [146] F. Weigend, R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).
- [147] T.H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
- [148] <http://www.theochem.uni-stuttgart.de>.
- [149] P.J. Hay, W.R. Wadt, J. Chem. Phys. **82**, 270 (1985).

- [150] P.J. Hay, W.R. Wadt, J. Chem. Phys. **82**, 284 (1985).
- [151] P.J. Hay, W.R. Wadt, J. Chem. Phys. **82**, 299 (1985).
- [152] C.C. Lovallo, M. Klobukowski, Int. J. Quant. Chem. **90**, 1099 (2002).
- [153] C.C. Lovallo, M. Klobukowski, J. Comput. Chem. **24**, 1009 (2003).
- [154] L.G.M. Pettersson, private communication.
- [155] A.M. Köster, *Entwicklung einer LCGTO-Dichtefunktionalmethode mit Hilfsfunktionen*, Habilitation Thesis, Universität Hannover (1998).
- [156] P. Calaminici, R. Flores-Moreno, A.M. Köster, Comput. Lett. **1**, 164 (2005).
- [157] J.C. Boettger, S.B. Trickey, Phys. Rev. B **53**, 3007 (1996).
- [158] R. Flores-Moreno, *Analytical Derivatives in LCGTO-DFT Pseudo-Potential Methods with Auxiliary Functions*, Ph.D. Thesis, Cinvestav, Mexico-City (2006).
- [159] R. Flores-Moreno, R.J. Álvarez-Méndez, A. Vela, A.M. Köster, J. Comput. Chem. **27**, 1009, (2006).
- [160] S.F. Boys, F. Bernardi, Mol. Phys. **19**, 553 (1970).
- [161] N. Rösch, S.B. Trickey, J. Chem. Phys. **106**, 8940 (1997).
- [162] H.B. Schlegel, M.J. Frisch, Int. J. Quant. Chem. **54**, 83 (1995).
- [163] H.-J. Glaeske, J. Reinhold, P. Volkmer, *Quantenchemie*, Band 5, Eds. W. Haberditzl, M. Scholz, L. Zülicke (Dr. Alfred Hüthig Verlag, Heidelberg, 1987).
- [164] C. Daul, Int. J. Quant. Chem. **52**, 867 (1994).
- [165] P.J. Hay, J. Chem. Phys. **66**, 4377 (1977).
- [166] C.C.J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
- [167] G.G. Hall, Proc. Roy. Soc. Ser. A **205**, 541 (1951).
- [168] J.A. Pople, R.K. Nesbet, J. Chem. Phys. **22**, 571 (1954).
- [169] C.C.J. Roothaan, Rev. Mod. Phys. **32**, 179 (1960).

- [170] J.S. Binkley, J.A. Pople, P.A. Dobosh, Mol. Phys. **28**, 1423 (1974).
- [171] T. Amos, L.C. Snyder, J. Chem. Phys. **41**, 1773 (1964).
- [172] A.D. Bacon, M.C. Zerner, Theor. Chim. Acta **53**, 21 (1979).
- [173] T.V. Russo, R.L. Martin, P.J. Hay, J. Chem. Phys. **101**, 7729 (1994).
- [174] M.F. Guest, V.R. Saunders, Mol. Phys. **28**, 819 (1974).
- [175] B.N. Plakhutin, E.V. Gorelik, N.N. Breslavskaya, J. Chem. Phys. **125**, 204110 (2006).
- [176] E.R. Davidson, Chem. Phys. Lett. **21**, 565 (1973).
- [177] R. McWeeny, G. Diercksen, J. Chem. Phys. **49**, 4852 (1968).
- [178] K. Faegri, R. Manne, Mol. Phys. **31**, 1037 (1976).
- [179] B.N. Plakhutin, E.R. Davidson, J. Phys. Chem. A **113**, 12386 (2009).
- [180] K.R. Glaesemann, M.W. Schmidt, J. Phys. Chem. A **114**, 8772 (2010).
- [181] J.C. Slater, J.H. Wood, Int. J. Quant. Chem. S **4**, 3 (1971).
- [182] J.F. Janak, Phys. Rev. B **18**, 7165 (1978).
- [183] P.J. Knowles, J.S. Andrews, R.D. Amos, N.C. Handy, J.A. Pople, Chem. Phys. Lett. **186**, 130 (1991).
- [184] T. Tsuchimochi, G.E. Scuseria, J. Chem. Phys. **134**, 064101 (2011).
- [185] A.M. Köster, J.M. del Campo, F. Janetzko, B. Zúñiga-Gutiérrez, J. Chem. Phys. **130**, 114106 (2009).
- [186] J. Almlöf, K. Faegri, Jr., K. Korsell, J. Comput. Chem. **3**, 385 (1982).
- [187] A. Álvarez Ibarra, *Asymptotic Expansion of Molecular Integrals in Self-Consistent Auxiliary Density Functional Methods*, Ph.D. Thesis Cinvestav, Mexico-City (2013).
- [188] A. Alvarez-Ibarra A.M. Köster, Mol. Phys. **113**, 3128 (2015).
- [189] D.R. Hartree, *The Calculation of Atomic Structures* (Wiley, New York, 1957).

- [190] E. Cancés, J. Chem. Phys. **114**, 10616 (2001).
- [191] V.R. Saunders, I.H. Hillier, Int. J. Quant. Chem. **8**, 699 (1973).
- [192] A.M. Köster, P. Calaminici, Z. Gómez, U. Reveles in *Reviews of Modern Quantum Chemistry, A Celebration of the Contribution of Robert G. Parr*, Ed. K. Sen (World Scientific Publishing Co., Singapore, 2002).
- [193] P. Pulay, J. Comput. Chem. **3**, 556 (1982).
- [194] J. Baker, A. Kessi, B. Delley, J. Chem. Phys. **105**, 192 (1996).
- [195] J.U. Reveles, *Geometry Optimization in LCGTO-DFT Methods with Auxiliary Functions*, Ph.D. Thesis, Cinvestav, Mexico-City (2004).
- [196] P. Culot, G. Dive, V.H. Nguyen, J.M. Ghuysen, Theor. Chim. Acta **82**, 189 (1992).
- [197] J.M. Bofill, J. Comput. Chem. **15**, 1 (1994).
- [198] A. Banerjee, N. Adams, J. Simons, R. Shepard, J. Phys. Chem. **89**, 52 (1985).
- [199] J. Baker, J. Comput. Chem. **7**, 385 (1986).
- [200] M.J.D. Powell, Math. Prog. **1**, 26 (1971).
- [201] K. Fukui, Acc. Chem. Res. **14**, 363 (1981).
- [202] C. Gonzalez, H.B. Schlegel, J. Chem. Phys. **90**, 2154 (1989).
- [203] C. Gonzalez, H.B. Schlegel, J. Chem. Phys. **95**, 5853 (1991).
- [204] J.M. del Campo, *Exploring Chemical Reactivity with Auxiliary Density Functional Theory*, Ph.D. Thesis, Cinvestav, Mexico-City (2008).
- [205] T.H. Fisher, J. Almlöf, J. Phys. Chem. **96**, 9768 (1992).
- [206] R. Lindh, A. Bernhardsson, G. Karlström, P. Malmqvist, Chem. Phys. Lett. **241**, 423 (1995).
- [207] C.G. Broyden, J. Inst. Maths. Applns. **6**, 76 (1970); *ibid.* **6**, 222 (1970).
- [208] R. Fletcher, Computer J. **13**, 317 (1970).

- [209] D. Goldfarb, Maths. Comp. **24**, 23 (1970).
- [210] D.F. Shanno, Maths. Comp. **24**, 647 (1970).
- [211] H.B. Schlegel, J. Comput. Chem. **3**, 214 (1982).
- [212] W.C. Davidon, AEC Res. & Dev. Report ANL-5990 (1959).
- [213] R. Fletcher, M.J.D. Powell, Computer J. **6**, 163 (1963).
- [214] J. Nocedal, S.J. Wright, *Numerical Optimization* (Springer, New York, 1999).
- [215] K. Levenberg, Quart. Appl. Math. **2**, 164 (1944).
- [216] D.W. Marquardt, SIAM J. **11**, 431 (1963).
- [217] J.M. del Campo, A.M. Köster, Croatica Chemica Acta **82**, A029 (2009).
- [218] J. Nichols, H. Taylor, P. Schmidt, J. Simons, J. Chem. Phys. **92**, 340 (1990).
- [219] R. Fletcher, *Practical Methods of Optimization*, Second Edition (Wiley, New York, 1987).
- [220] L. Verlet, Phys. Rev. **159**, 98 (1967).
- [221] W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson, J. Chem. Phys. **76**, 637 (1982).
- [222] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
- [223] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
- [224] J.M. Vásquez-Pérez, G.U. Gamboa, D. Mejía-Rodríguez, A. Alvarez-Ibarra, G. Geudtner, P. Calaminici, A.M. Köster, J. Phys. Chem. Lett. **6**, 4646 (2015).
- [225] J.M. Vásquez-Pérez, *Propiedades Termodinámicas de Sistemas Finitos a partir de Dinámica Molecular de Born-Oppenheimer*, Ph.D. Thesis Cinvestav, Mexico-City (2011).
- [226] B. Zuniga-Gutierrez, M. Camacho Gonzalez, P. Simon Bastida, A. Bendana Castillo, P. Calaminici, A.M. Köster, J. Chem. Phys. **143**, 104103 (2015).

- [227] B. Zuniga-Gutierrez, M. Camacho-Gonzalez, P. Simon-Bastida, A. Bendana-Castillo, P. Calaminici, A.M. Köster, J. Phys. Chem. A, **119**, 1469 (2015).
- [228] B. Zuniga-Gutierrez, G. Geudtner, A.M. Köster, J. Chem. Phys. **137**, 094113 (2012)
- [229] B. Zuniga-Gutierrez, M. Camacho-Gonzalez, P. Simon-Bastida, A. Bendana-Castillo, P. Calaminici, A.M. Köster, J. Phys. Chem. A, **119**, 1469 (2015).
- [230] T.L. Beck, J.D. Doll, D.L. Freeman, J. Chem. Phys. **90**, 5651 (1989).
- [231] S. Kümmel, J. Akola, M. Manninen, Phys. Rev. Lett. **84**, 3827 (2000).
- [232] K. Joshi, D.G. Kanhere, S.A. Blundell, Phys. Rev. B **67**, 235413 (2003).
- [233] G.U. Gamboa Martínez, *Born-Oppenheimer Molecular Dynamics with Auxiliary Density Functional Theory*, Ph.D. Thesis Cinvestav, Mexico-City (2011).
- [234] L.V. Woodcock, Chem. Phys. Lett. **10**, 257 (1971).
- [235] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, J. Chem. Phys. **81**, 3684 (1984).
- [236] D.J. Evans, W.G. Hoover, B.H. Failor, B. Moran, A.J.C. Ladd, Phys. Rev. B **28**, 1016 (1983).
- [237] S. Nosé, J. Chem. Phys. **81**, 511 (1984).
- [238] W.G. Hoover, Phys. Rev. A **31**, 1695 (1985).
- [239] G.J. Martyna, M.L. Klein, M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).
- [240] W.F. Murphy, W. Holzer, H.J. Bernstein, Appl. Spectrosc. **23**, 211 (1969).
- [241] K. Jug, F. Janetzko, A.M. Köster, J. Chem. Phys. **114**, 5472 (2001).
- [242] S.V. Shedge, S. Pal, A.M. Köster, Chem. Phys. Lett. **552**, 146 (2012).
- [243] J. Carmona-Espíndola, R. Flores-Moreno, A.M. Köster, J. Chem. Phys. **133**, 084102 (2010).
- [244] J. Carmona-Espíndola, R. Flores-Moreno, A.M. Köster, Int. J. Quantum Chem. **112**, 3461 (2012).

- [245] A. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford, 2013).
- [246] J. Carmona-Espíndola, *Time-Dependent Auxiliary Density Perturbation Theory: Method, Implementation and Applications* Ph.D. Thesis, Cinvestav, Mexico-City (2012).
- [247] B. Zuniga-Gutierrez, A.M. Köster, Mol. Phys. **114**, 1026 (2016).
- [248] M.E. Casida, A. Ipatov, F. Cordova, in *Time-Dependent Density-Functional Theory*, edited by M.A.L. Marques, C. Ullrich, F. Nogueira, A. Rubio, E.K.U. Gross, Lecture Notes in Physics (Springer, Berlin, 2006).
- [249] S. Hirata, M. Head-Gordon, Chem. Phys. Lett. **314**, 291 (1999).
- [250] M.E. Casida, F. Gutierrez, J. Guan, F.-X. Gadea, D. Salahub, J.-P. Daudey, J. Chem. Phys. **113**, 7062 (2000).
- [251] F. Wang, T. Ziegler, J. Chem. Phys. **121**, 12191 (2004).
- [252] M. Leetmaa, M.P. Ljungberg, A. Lyubartsev, A. Nilsson, L.G.M. Pettersson, J. Electron Spec. Rel. Phen. **177**, 135 (2010).
- [253] O. Takahashi, L.G.M. Pettersson, J. Chem. Phys. **121**, 10339 (2004).
- [254] J.-L. Pascual, L.G.M. Pettersson, H. Ågren, Phys. Rev. B **56**, 7716 (1997).
- [255] B. Zúñiga-Gutiérrez, G. Geudtner, A.M. Köster, J. Chem. Phys. **134**, 124108 (2011).
- [256] B.A. Zúñiga Gutiérrez, *Magnetic Shielding, Spin-Spin Coupling Constants and Magnetizability Tensors Calculated with Auxiliary Density Perturbation Theory*, Ph.D. Thesis, Cinvestav, Mexico-City (2011).
- [257] B. Zúñiga-Gutiérrez, M. Camacho Gonzalez, P. Simon Bastida, A. Bendana Castillo, P. Calaminici, A.M. Köster, J. Chem. Phys. **143**, 104103 (2015).
- [258] G. de Luca, N. Russo, A.M. Köster, P. Calaminici, K. Jug, Mol. Phys. **97**, 347 (1999).
- [259] E. Dib, T. Mineva, P. Gaveau, B. Alonso, Eur. Phys. J. Special Topics **224**, 1769 (2015).

- [260] R. Flores-Moreno, J. Melin, J.V. Ortiz, G. Merino, *J. Chem. Phys.* **128**, 134105 (2008).
- [261] R. Flores-Moreno, *J. Chem. Theory & Comput.* **6**, 48 (2010).
- [262] R.S. Mulliken, *J. Chem. Phys.* **23**, 1833 (1955).
- [263] P.O. Löwdin, *J. Chem. Phys.* **18**, 365 (1950).
- [264] I. Mayer, *Chem. Phys. Lett.* **97**, 270 (1983).
- [265] F.L. Hirshfeld, *Theor. Chim. Acta* **44**, 129 (1977).
- [266] P. Bultinck, C. Van Alsenoy, P.W. Ayers, R. Carbó-Dorca, *J. Chem. Phys.* **126**, 144111 (2007).
- [267] M.S. Gopinathan, K. Jug, *Theor. Chim. Acta* **63**, 497 (1983).
- [268] E.D. Glendening, J.K. Badenhoop, A.E. Reed, J.E. Carpenter, J.A. Bohmann, C.M. Morales, F. Weinhold, NBO 5.G  
<http://www.chem.wisc.edu/~nbo5>.
- [269] A. de la Lande, C. Clavagura, A.M. Köster, *J. Mol. Modeling* **23**, 99 (2017).
- [270] G. Geudtner, Z. Gómez-Sandoval, F. Janetzko, P. Calaminici, *Theor. Chem. Acc.* **144**, 137 (2005).
- [271] G. Schaftenaar, Molden3.2, CAOS/CAMM Center, University of Nijmegen (1997).
- [272] S. Portmann, P.F. Fluekiger, CSCS/ETHZ and CSCS/UNI Geneva (2002).
- [273] <http://argo.urv.es/XAIM>.
- [274] F. Aquilante, T.B. Pedersen, A.S. de Merás, H. Koch, *J. Chem. Phys.* **125**, 174101 (2006).
- [275] J.M. Foster, S.F. Boys, *Rev. Mod. Phys.* **32**, 300 (1960).
- [276] J. Pipek, P.G. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).
- [277] R.K. Pathak, S.R. Gadre, *J. Chem. Phys.* **93**, 1770 (1990).
- [278] A.D. Becke, K. Edgecomb, *J. Chem. Phys.* **92**, 5397 (1990).

- [279] B. Silvi, A. Savin, *Nature* **371**, 683 (1994).
- [280] R.F. Nalewajski, A.M. Köster, S. Escalante, *J. Phys. Chem. A* **109**, 10038 (2005).
- [281] W.E. Lorensen, H.E. Cline, *Comp. Graph.* **21**, 163 (1987).
- [282] M. Leboeuf, A.M. Köster, K. Jug, D.R. Salahub, *J. Chem. Phys.* **111**, 4893 (1999).
- [283] LAPACK driver routine (version 3.0) –  
Univ. of Tennessee, Univ. of California Berkeley, NAG Ltd.,  
Courant Institute, Argonne National Lab, and Rice University, 1999.
- [284] J.L. Morales, J. Nocedal, *SIAM J. Optim.* **10**, 1079 (2000).
- [285] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN*, Second Edition (Cambridge University Press, Cambridge, 1992).
- [286] C.W. Murray, S.C. Racine, E.R. Davidson, *J. Comput. Phys.* **103**, 382 (1991).
- [287] A.D. Becke, *J. Chem. Phys.* **88**, 2547 (1987).
- [288] R.E. Stratmann, G.E. Scuseria, M.J. Frisch, *Chem. Phys. Lett.* **257**, 213 (1996).
- [289] M.E. Mura, P.J. Knowles, *J. Chem. Phys.* **104**, 9848 (1996).