CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO

PROGRAMA DE NANOCIENCIAS Y NANOTECNOLOGÍA

Ajuste iterativo de la densidad para simulaciones ab initio en nanoescala

T E S I S

que presenta

Jesús Naín Pedroza Montero

Para obtener el grado de

DOCTOR EN CIENCIAS

EN NANOCIENCIAS Y NANOTECNOLOGÍA

Directores de la Tesis:
Dra. Patrizia Calaminici
Dr. Andreas M. Köster

Ciudad de México **Octubre 2020**

CENTER FOR RESEARCH AND ADVANCED STUDIES OF THE NATIONAL POLYTECHNIC INSTITUTE

ZACATENCO CAMPUS

NANOSCIENCIE AND NANOTECHNOLOGY PROGRAM

Iterative density fitting for ab initio nanoscale simulations

T H E S I S

submitted by

Jesús Naín Pedroza Montero

In order to obtain the

DOCTOR IN SCIENCE

IN NANOSCIENCE AND NANOTECNOLOGY

Advisors:
Dr. Patrizia Calaminici
Dr. Andreas M. Köster

Mexico City **October 2020**

Acknowledgments

I gratefully acknowledge Dr. Patrizia Calaminici and Dr. Andreas Köster for their invaluable guidance and support through my Ph.D. path. Our scientific discussions and your support helped me a lot in the development of this work. It is worth to mention that your advices have made an enormous change in the way I look at science. In order words, you have made me a good scientist.

I would like to thank Dr. Gerald Geutdner for his patience and the way he teaches FORTRAN programming. It was a great opportunity to work with you in one of my manuscripts. Besides work, also I appreciate the good talks and barbecue times where we had a beer and a glass of wine, respectively.

I also want to thank Dr. Alberto Vela for the good talks in the Theoretical Chemistry area. Despite I could not attend any of your courses because they were too early, I know you are an excellent physicist and an excellent chemist.

I happily want to thank my friends who became my first mentors inside the Theoretical Chemistry group: Rogelio Delgado and Daniel Mejía, who were the first ones to introduce me to the deMon2k coding and still help me when I have doubts.

A special acknowledgement to my friend Antonio Delesma. Many moments and discussions together do not fit into this page. I really appreciate you friendship.

In general, I gratefully acknowledge the whole Theoretical Chemistry group members and friends: Ángel Albavera, Adrián Martínez, Ignacio Hernández, Luis López-Sosa, Juan Diego Samaniego, Carlos Galindo, Raúl Quintero, Esaú Martínez, Carlos Zepactonal and Jorge Nochebuena. Also, to some friends that were part of the group but we still keep in touch as: Anabel

Galindo, Aurelio Álvarez, Lizandra Barrios, Luis Cota, Sandra Mejía, Sara Pérez. Your friendship is really appreciated.

I gratefully acknowledge the evaluation commission members: Dr. Mary Rosales-Hoz, Dr. Omar Solorza-Feria, Dr. Joel Ireta and Dr. Lars G. M. Pettersson. Many thanks for the very good suggestions and corrections to my Ph.D. thesis.

Thanks to the Nanoscience and Nanotechnology Ph.D. program who accepted me as a student 5 years ago.

The CONACyT financial fellowship No. 245798 is also acknowledged.

Contents

List of Figures

List of Tables

Related Work

The following publications contain complementary information and additional results obtained during my Ph.D. project:

- **J.N. Pedroza-Montero**, F.A. Delesma, R.I Delgado-Venegas, P. Calaminici, A.M. Köster, Static and Dynamic Polarizabilities of Oligothiophenes, *Theor. Chem. Acc.* **135**, 230, (2016).

- **J.N. Pedroza-Montero**, J.L. Morales, G. Geudtner, A. Álvarez-Ibarra, P. Calaminici, A.M. Köster, Variational Density Fitting with a Krylov Subspace Method, *J. Chem. Theory Comput.* **16**, 2965, (2020).

- **J.N. Pedroza-Montero**, F.A. Delesma, J.L. Morales, P. Calaminici, A.M. Köster, Variational Fitting of Fock Exchange Potential with Modified Cholesky Decomposition *J. Chem. Phys. just accepted*.

The following manuscripts with results from this thesis are currently in preparation:

- B. Zúñiga-Gutiérrez, L.G. Cota, **J.N. Pedroza-Montero**, A.M. Köster, Nuclear Spin-Spin Coupling Constants from Auxiliary Density Functional Theory, *to be submitted*.

- F.A. Delesma, **J.N. Pedroza-Montero**, R.I. Delgado-Venegas, J.M. del Campo, P. Calaminici, A.M. Köster, Auxiliary Density Perturbation Theory for Fock Exchangeand Hybrid Functionals, *in preparation*.

- **J.N. Pedroza-Montero**, P. Calaminici, A.M. Köster, First-principle Polarizabilities of Nanosystems, *in preparation*.

Parts of this Ph.D. thesis were presented as oral or poster contributions at the following national

and international scientific conferences:

- Poster: **J.N. Pedroza-Montero**, J.L. Morales, P. Calaminici, A.M. Köster,
  Simple, Consistent and Efficient Density Fitting with the Eirola-Nevanlinna Algorithm
  Nano-Cinves 2017, México, City.

- Talk: **J.N. Pedroza-Montero**, J.L. Morales, P. Calaminici, A.M. Köster,
  Simple, Consistent and Efficient Density Fitting with the Eirola-Nevanlinna Algorithm
  17^{th} deMon Developers Workshop
  May, 2017, Calgary, Canada.

- Talk: **J.N. Pedroza-Montero**, J.L. Morales, P. Calaminici, A.M. Köster,
  Variational Density Fitting with a Krylov Subspace Method
  18^{th} deMon Developers Workshop
  June, 2018, Guadalajara, Jalisco.

- Talk: **J.N. Pedroza-Montero**, J.L. Morales, P. Calaminici, A.M. Köster,
  Iterative Density Fitting with Krylov Methods
  $15th$ ICCMSE Symposium in Honor of Dennis Salahub
  May, 2019, Rhodes, Greece.

- Talk: **J.N. Pedroza-Montero**, J.L. Morales, P. Calaminici, A.M. Köster,
  Variational Density Fitting with a Krylov Subspace Method
  XXVII International Materials Research Congress
  August, 2019, Cancún, Quintana Roo.

List of Acronyms and Symbols

ADFT	Auxiliary density functional theory
AO	Atomic orbital
BOMD	Born-Oppenheimer molecular dynamics
deMon2k	Computational chemistry code: density of Montreal 2000
DFT	Density functional theory
DZVP	Double-zeta valence polarization
EN	Eirola-Nevanlinna algorithm
ERI	Electron repulsion integral
GEN-A2	Automatic generated auxiliary function set of A2 quality
GGA	Generalized gradient approximation
GTO	Gaussian-type orbitals
HGTF	Hermite Gaussian-type function
LCAO	Linear combination of atomic orbitals
LCGTO	Linear combination of Gaussian type orbitals
LDA	Local density approximation
MAE	Mean absolute error
MINRES	Minimal residual algorithm
MM	Molecular mechanics
MO	Molecular orbital
QM	Quantum mechanics
QM/MM	Quantum mechanics / Molecular mechanics
RAM	Random access memory
SCF	Self-consistent field
SCP	Self-consistent perturbation
TED	Truncated eigenvalue decomposition

XC	Exchange-correlation
$\mathbf{a}$	Contracted Gaussian-type orbital at nucleus **A**
A, B, C	Nuclei indices
$\mathbf{A}$	Spatial coordinate vector of nucleus **A**
$\hat{\mathcal{A}}_C(\bar{\mathbf{c}})$	Generalized nuclear attraction operator
$c_{\mu i}$	Molecular orbital coefficient
$\bar{\mathbf{c}}$	Hermite Gaussian auxiliary function at nucleus **C**
$\hat{\mathcal{D}}_C$	Differentiation operator with respect to nucleus $\vec{C}$
δ_{ij}	Kronecker delta
d_k	Contraction coefficient
E	Total energy
E_0	Ground state energy
E_H	Coulombic electron repulsion energy
E_{xc}	Exchange-correlation energy
$\mathcal{E}_2$	Error of the variational fitting of the Coulomb potential
F_n	Boys function of order n
$G_{\bar{k}\bar{l}}$	Coulomb matrix element
$\mathbf{G}$	Coulomb matrix
$\hat{H}$	Hamilton operator
$H_{\mu\nu}$	Core Hamilton matrix element
$\hat{H}_{el}$	Electronic Hamilton operator
i, j	Electron or orbital index
$\mathbf{J}$	Coulomb vector
$K_{\mu\nu}$	Kohn-Sham matrix element
$\mathbf{K}$	Kohn-Sham matrix
$\mathcal{L}$	Lagrange function
λ_{ij}	Lagrange multiplier
∇^2	Laplacian operator
∇	Nabla operator
M_A	Mass of nucleus
N	Number of electrons
N_{aux}	Number of auxiliary functions
N_{bas}	Number of basis functions
$P_{\mu\nu}$	Density matrix element
$\mathbf{r}_i$	Spatial coordiante vector of electron i

$S_{\mu\nu}$	Overlap matrix element
$\hat{T}_{el}$	Electronic kinetic energy operator
$\hat{T}_N$	Nuclear kinectic energy operator
μ, ν	Atomic orbitals
v	External potential
$\mathbf{R}$	Response matrix
$\mathbf{S}$	Overlap matrix
v_{xc}	Exchange-correlation potential
$x_{\tilde{c}}$	Exchange-correlation fitting coefficient
$\mathbf{x}$	Coulomb coefficient vector
$\mathbf{z}$	Exchange-correlation coefficient vector
Z_A	Charge of nucleus A

Abstract

In this work, we present the implementation of a variational density fitting methodology that uses iterative linear algebra for solving the associated system of linear equations within the ADFT framework. It is well known that most difficulties with this system arise from the fact that the coefficient matrix is in general ill-conditioned and, due to finite precision round-off errors, it may not be positive definite. The dimensionality, given by the number of auxiliary functions, also poses a challenge in terms of memory and time demand since the coefficient matrix is dense. The methodology presented is based on a preconditioned Krylov subspace method called MINRES able to deal with indefinite ill-conditioned equation systems. To assess its potential, it has been combined with double asymptotic electron repulsion integral expansions as implemented in the deMon2k package. A numerical study on a set of problems with up to 130,000 auxiliary functions shows its effectiveness to alleviate the above mentioned problematic. A comparison with the default methodology used in deMon2k based on a truncated eigenvalue decomposition of the coefficient matrix indicates that the proposed method exhibits excellent robustness and scalability when implemented in a parallel setting. Moreover, the implementation of the MINRES algorithm has been adapted to response property calculations such as polarizabilities, hyperpolarizabilities and nuclear spin-spin coupling constants. With this development the linear algebra bottleneck associated to the density fitting is eliminated. It is important to note that this bottleneck was twofold, namely in terms of computational as well as RAM demand. Both problems were overcome with the here presented adaptation and implementation of the Krylov subspace method MINRES. Illustrative examples of nanosystems and microbiological relevant structures have been used to test the MINRES implementation.

Resumen

En este trabajo, presentamos la implementación de una metodología de ajuste de densidad variacional que utiliza un procedimiento iterativo de álgebra lineal para resolver los sistemas de ecuaciones lineales asociados dentro del marco de la teoría del funcional auxiliar. Es bien sabido que la mayoría de las dificultades con estos sistemas surgen del hecho de que la matriz de coeficientes está en general mal condicionada y, debido a errores de redondeo de precisión finita, puede no ser definida positiva. La dimensionalidad, dada por el número de funciones auxiliares, también plantea un desafío en términos de demanda de memoria y tiempo ya que la matriz de coeficientes es densa. La metodología presentada se basa en un método de Krylov precondicionado llamado MINRES capaz de lidiar con sistemas de ecuaciones indefinidas y mal condicionados. Para evaluar su potencial, se ha combinado con expansion asintótica doble para las integrales de repulsión electrónica dentro del código deMon2k. Un estudio numérico sobre un conjunto de problemas con hasta 130,000 funciones auxiliares muestra su eficacia para mitigar la problemática antes mencionada. Una comparación con la metodología predeterminada utilizada en deMon2k basada en una descomposición truncada de valores propios de la matriz de coeficientes, indica que el método propuesto exhibe una excelente robustez y escalabilidad cuando se implementa en un entorno computacional paralelo. Además, la implementación del algoritmo MINRES se ha adaptado a los cálculos de propiedades de respuesta como polarizabilidades, hiperpolarizabilidades y constantes de acoplamiento espín-espín nuclear. Con este desarrollo se elimina el cuello de botella del álgebra lineal asociado al ajuste de la densidad. Es importante señalar que este cuello de botella era doble, a saber, en términos de demanda computacional y RAM. Ambos problemas fueron resuletos con la adaptación e implementación aquí presentada del algorithmo MINRES. Ejemplos ilustrativos de nanosistemas y estructuras microbiológicas relevantes fueron utilizadas para probar la implementación de MINRES.

1

Introduction and Objectives

Contents

1.1 Introduction

With the rise of the variational fitting of the Coulomb[1–10] and Fock[11–16] potential in combination with localized resolution-of-the identity post Hartree-Fock approaches[17–23], the calculation of four-center electron repulsion integrals (ERIs) might become obsolete in ab initio quantum chemistry methods based on the linear combination of Gaussian type orbitals (LCGTO) approximation. However, such a fundamental change in ab initio LCGTO approaches requires numerically stable implementations of the algorithms associated with the variational fitting. Moreover, these algorithms must also offer attractive perspectives for growing system sizes. Although, there is little doubt that for currently accessible system sizes three-center ERI implementations outperform their four-center counterparts, it is not yet clear that a numerically stable variational fitting approach will outperform four-center ERI based implementations of any system size. The main concern is the linear algebra associated with the variational fitting of the Coulomb potential. In its original formulation[3], a linear equation system with dimension equal to the number of auxiliary functions in the molecule has to be solved; the corresponding coefficient matrix, herein called Coulomb matrix, is supposed to be positive definite. However, due to finite precision arithmetic errors, it can become slightly indefinite and ill-conditioned. Auxiliary function sets that contain functions with small exponents and high angular momentum index seem the cause of this difficulty.

At this point it is important to note that the variational fitting of the Fock potential is less problematic due to its intrinsic local nature. As a result, localized molecular orbitals (MOs) are often used for the variational fitting of the Fock potential[13–16]. In principle, this also opens the possibility of a variational fitting of the Coulomb potential with localized MOs. The downside of such an approach is its increased computational work demand because a fitting equation system has to be solved for each occupied MO. In particular, the scalability of the parallelization of the variational fitting of the Coulomb potential might be jeopardized. As a computational attractive alternative ad-hoc partitioning approaches for the fitting of the Coulomb potential have been suggested in the literature[24–30]. Unfortunately, the variational nature of the Coulomb potential fitting is not guaranteed in these approaches. If this nature is of concern, e.g. for the calculation of exchange-correlation contributions with the auxiliary density[31–34], computing the solution of the equation system, either for the full molecule or for each occupied orbital, is unavoidable. To circumvent the numerical and computational difficulties associated with the solution of the fitting equation system for the full system, a BFGS quasi-Newton approach[35] was adapted for the variational fitting of the Coulomb potential in our groups a few years ago[36]. Unfortunately, this approach fails if the Coulomb matrix is not positive definite. As already mentioned this is

for larger auxiliary function sets more the rule than the exception. Therefore, an initial truncated eigenvalue decomposition (TED) of the Coulomb matrix was applied prior to the iterative quasi-Newton calculation of the density fitting coefficients. Even though, the TED has to be computed only once, it represents a critical computational bottleneck when calculating systems with tens of thousands of auxiliary functions. In order to overcome this computational bottleneck, we suggest in this work the use of MINRES[37,38] for the iterative solution of the fitting equation systems. MINRES is a Krylov subspace based method designed for solving symmetric indefinite equation systems. Thus, the positive definiteness of the Coulomb matrix is not required. Experience in different fields shows that Krylov methods are effective, robust and efficient; these conditions are essential to any numerical algorithm used inside self-consistent field (SCF) calculations.

The main purpose of this work is to show that the MINRES algorithm, parallelized in a suitable form, leads to a low memory and low scaling (only matrix-vector operations are needed) implementation for the iterative density fitting. In particular, the reduced memory demand of MINRES in combination with its improved performance due to newly developed asymptotic integral expansions for two-center ERIs opens up new possibilities for first-principles nanoscale calculations. Illustrative examples for energy calculations, structure optimizations and response property calculations of large molecules with nanometric extensions are presented.

The thesis is organized as follows. In Chapter 2 the revision of quantum chemistry fundamentals with a particular focus on the Kohn-Sham formalism is presented. In the following chapter we derive the auxiliary density functional theory formalism as a very efficient alternative to the Kohn-Sham methodology. In Chapter 4 the use of iterative solvers for the density fitting equation systems is reviewed. Special attention is given to Krylov subspace methods. The details of our MINRES adaption to self-consistent field (SCF) and perturbation theory calculations are outlined and the corresponding implementation details are discussed. The validation of our MINRES implementations are shown in Chapter 5 accompanied by relevant benchmark calculations. Finally, Chapter 6 presents the conclusions and perspectives of this thesis.

1.2 Objectives

The main objective of this work is to contribute to the development of a scalable all-electron first principles methodology for the calculation of large molecular systems with hundreds of thousands of auxiliary functions and tens of thousands of basis functions. With double-zeta valence plus polarization (DZVP) basis sets and automatically generated GEN-A2* auxiliary function sets this will allow first-principles calculations of nanosystems with around 10,000 atoms. In deMon2k, the main bottlenecks in such calculations are in the moment the linear algebra tasks arising from the variational fitting of the Coulomb potential. In order to eliminate these bottlenecks, the following specific objectives are addressed:

- Implementation of the MINRES iterative algorithm inside deMon2k for solving linear equation systems.
- Development and implementation of the double asymptotic expansion for the two-center ERIs appearing in the matrix-vector multiplication of MINRES.
- Implementation of MINRES for SCF energy calculations in deMon2k.
- Implementation of MINRES for all available response property calculations such as Fukui functions, polarizabilities, hyperpolarizabilities and nuclear spin-spin coupling constants.
- Development and implementation of an efficient and scalable parallel version of MINRES for variational density fitting.

These objectives are accompanied by corresponding validation and benchmark calculations for the systems specified in Table 1.1 along with their average sizes.

Table 1.1: Average diameters and lengths in nm of the studied systems in this thesis.

Fullerenes	Diameter	Water clusters	Diameter	Zeolites	Length	Amylose units	Length	DNA units	Length	Oligothiophene	Length
C_{20}	0.3	$(H_2O)_{50}$	1.3	MFI-1	1.3	8	2.5	4	1.1	100 units	35.0
C_{60}	0.7	$(H_2O)_{100}$	1.6	MFI-2	2.5	16	5.0	8	2.4		
C_{180}	1.2	$(H_2O)_{200}$	2.0	MFI-3	3.8	32	9.8	16	5.1		
C_{240}	1.4	$(H_2O)_{300}$	2.4	MFI-4	5.0	48	14.3				
C_{540}	2.2	$(H_2O)_{400}$	2.7			64	19.8				
C_{720}	2.6	$(H_2O)_{500}$	3.0								
C_{960}	3.0										

2

Quantum Chemistry Fundamentals

Contents

2.1 The Schrödinger Equation

One of the cornerstones of quantum physics is the Schrödinger equation, which describes the evolution of a system in terms of quantum mechanics. This was a challenge proposed by Debye to Schrödinger in 1925[39], in order to develop a wave equation that could describe matter. In 1926[40], this led to the most general form of the famous equation, namely, the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t) = \hat{H} \Psi(\mathbf{x}, t), \tag{2.1}$$

where Ψ is a function that completely describes the state of a system and is called the wave function $\Psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n,t)$, and the $\mathbf{x}_i$, $i = 1, ..., n$, are the spatial coordinates of all n particles, that constitute the system. In Eq. (2.1), t is the time variable and $\hat{H}$ is the Hamilton operator.

The classical analogy to Eq. (2.1) is the Hamiltonian in Hamilton mechanics, which predicts how a classical system propagates in the future given its current configuration. If the Hamilton operator does not explicitly depend on t and the system does not change with time, it is said that the system is in a stationary state. Hence, for stationary solutions, the time-dependent wave function reduces to a time-independent wave function, $\Psi(\mathbf{x}_1, \mathbf{x}_2,..., \mathbf{x}_n)$, that satisfies the equation

$$\hat{H}\Psi(\mathbf{x}) = E\Psi(\mathbf{x}). \tag{2.2}$$

For an isolated system with N electrons and M nuclei the explicit form of the non-relativistic Hamilton operator in atomic units is given by

$$\begin{aligned} \hat{H} = -\frac{1}{2}\sum_{i}^{N} \nabla_i^2 - \sum_{A}^{M} \frac{\nabla_A^2}{2M_A} - \sum_{A}^{M}\sum_{i}^{N} \frac{Z_A}{|\mathbf{r}_i - \mathbf{A}|} + \\ \sum_{i}^{N}\sum_{j>i}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{A}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{|\mathbf{A} - \mathbf{B}|}, \end{aligned} \tag{2.3}$$

where $\mathbf{r}$, $\mathbf{A}$ and $\mathbf{B}$ denote the spatial coordinates of the electrons and nuclei, respectively. The atomic mass of nucleus A is M_A and its charge Z_A. The terms on the right hand side of Eq. (2.3) are the following: The first two terms are the kinetic energies of the N electrons and M nuclei, respectively. The last three terms are the potential part of the Hamilton operator in form of electrostatic particle-particle interactions. The electrostatic attraction between the N electrons and the M nuclei, the electrostatic repulsion between the electrons, and the electrostatic repulsion

between the nuclei[41]. Therefore, Eq. (2.3) can also be expressed as

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn}. \tag{2.4}$$

The term $\hat{V}_{ne}$ couples electronic and nuclear motions, making the solution of Eq. (2.2) more complicated. The so-called Born-Oppenheimer approximation[42] states that due to the mass difference (more than 3 orders of magnitude) between the electrons and nuclei, the latter can be spatially fixed. Semi-classically, this means that the nuclear movement can be neglected on the timescale of electronic transitions. In such a semi-classical framework, the electrons follow the nuclear motion adiabatically without relaxation time[43]. Born and Oppenheimer showed that an effective separation of electronic and nuclear motions can be performed without affecting the accuracy of the solution for many cases[44]. Using the adiabatic approximation[45,46], the following quasi-separable ansatz, restricting ourselves for simplicity to only one electronic state, can be used for Eq. (2.2)

$$\Psi(\mathbf{x}, \mathbf{R}) = \Psi(\mathbf{x}; \mathbf{R}) \cdot \Theta(\mathbf{R}). \tag{2.5}$$

In Eq. (2.5) the electronic wave function, $\Psi(\mathbf{x}; \mathbf{R})$, has a parametric dependency from the nuclear coordinates and Θ is the nuclear wave function. Applying the corresponding electronic and nuclear operators, $\hat{T}_e$, and $\hat{T}_n$, respectively, to this wave function yields

$$\hat{T}_e \Psi(\mathbf{x}, \mathbf{R}) = \hat{T}_e \Psi(\mathbf{x}; \mathbf{R}) \Theta(\mathbf{R}) = \Theta(\mathbf{R}) \hat{T}_e \Psi(\mathbf{x}; \mathbf{R}), \tag{2.6}$$

$$\hat{T}_n \Psi(\mathbf{x}, \mathbf{R}) = \hat{T}_n \Psi(\mathbf{x}; \mathbf{R}) \Theta(\mathbf{R}). \tag{2.7}$$

In the asymptotic limit $1/M_A \rightarrow 0 \; \forall \; M_A \rightarrow \infty$ the nuclear configuration can be assumed frozen for the solution of the electronic problem. Therefore, Eq. (2.7) can be expressed as

$$\hat{T}_n \Psi(\mathbf{x}; \mathbf{R}) = -\sum_A^M \frac{1}{2M_A} \Big[\Psi(\mathbf{x}; \mathbf{R}) \nabla_A^2 \Theta(\mathbf{R}) + \Theta(\mathbf{R}) \cancelto{\hat{C}_n}{\nabla_A^2 \Psi(\mathbf{x}; \mathbf{R})} + 2 \cancelto{\hat{C}_n}{\nabla_A \Psi(\mathbf{x}; \mathbf{R})} \cdot \nabla_A \Theta(\mathbf{R}) \Big] \quad ,$$

where $\hat{C}_n$ denotes the adiabatic correction operator that absorbs, after the electronic expectation values are calculated, the second and third term. Because the adiabatic correction is most often

negligible[47] Eq. (2.7) simplifies in the Born-Oppenheimer approximation to

$$\hat{T}_n \Psi(\mathbf{x};\mathbf{R}) = -\Psi(\mathbf{x};\mathbf{R}) \sum_{A}^{M} \frac{1}{2M_A} \nabla_A^2 \Theta(\mathbf{R}). \tag{2.8}$$

Now substituting Eq. (2.5) into Eq. (2.2) and using approximation (2.8) yields

$$\frac{[\hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne} + \hat{V}_{nn}]\Psi(\mathbf{x};\mathbf{R})}{\Psi(\mathbf{x};\mathbf{R})} + \frac{\hat{T}_n \Theta(\mathbf{R})}{\Theta(\mathbf{R})} = E. \tag{2.9}$$

Thus, the first term of the left hand side can only be a function of $\mathbf{R}$:

$$\frac{[\hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne} + \hat{V}_{nn}]\Psi(\mathbf{x};\mathbf{R})}{\Psi(\mathbf{x};\mathbf{R})} = \mathcal{E}(\mathbf{R}). \tag{2.10}$$

Rearrangement of Eq. (2.10) yields the electronic Schrödinger equation:

$$\begin{aligned} [\hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne} + \hat{V}_{nn}]\Psi(\mathbf{x};\mathbf{R}) &= \mathcal{E}(\mathbf{R})\Psi(\mathbf{x};\mathbf{R}) \\ \hat{H}_e \Psi(\mathbf{x};\mathbf{R}) &= \mathcal{E}(\mathbf{R})\Psi(\mathbf{x};\mathbf{R}). \end{aligned} \tag{2.11}$$

Note that the nuclear repulsion energy, $\hat{V}_{nn}$, was incorporated into the electronic Hamilton operator because it contributes to the potential energy for the nuclei. Technically, the nuclear repulsion energy is calculated analytically by classical electrostatic repulsion between nuclear point charges. The solutions of the electronic Schrödinger equation form a set of orthogonalized eigenfunctions, $\Psi(\mathbf{x};\mathbf{R})$, with corresponding eigenvalues $\mathcal{E}(\mathbf{R})$. For each solution there will be a corresponding nuclear Schrödinger equation. For the electronic ground state, the nuclear Schrödinger equation is given by:

$$\begin{aligned} [\hat{T}_n + \mathcal{E}_0(\mathbf{R})]\Theta(\mathbf{R}) &= E\,\Theta(\mathbf{R}) \\ \hat{H}_n \Theta(\mathbf{R}) &= E\,\Theta(\mathbf{R}), \end{aligned} \tag{2.12}$$

where $\mathcal{E}_0(\mathbf{R})$ is the ground state electronic energy for the (fixed) nuclear configuration $\mathbf{R}$, i.e. a point of the corresponding ground state potential energy surface (PES).

The usual procedure to solve Eq. (2.12) is to solve first Eq. (2.11), then substituting the electronic energy into the nuclear Schrödinger equation and solve it. According to Eqs. (2.11) and (2.12) the solution of the electronic Schrödinger equation represents the potential in which the

nuclear wave function can be propagated. This is the basis for the so-called Born-Oppenheimer molecular dynamics (BOMD) approaches that have become very popular over recent years[48]. In BOMD, the motions of the nuclei on the PES are usually solved using Newtonian mechanics in order to permit on-the-fly BOMD simulations. Note that each electronic state will give rise to an own PES, and that crossing between two surfaces is not allowed by the adiabatic theorem[45]. Because our focus is on the solution of the electronic Schrödinger equation, we drop the electronic index *e* of the electronic Hamilton operator as well as the parametric dependency of the electronic wave function on **R**.

Even with the Born-Oppenheimer approximation, the Schrödinger equation cannot be solved analytically for a system with two or more electrons in general. Thus, only approximate solutions are available for most chemical systems of interest. Several approaches have been developed for this purpose[49,50], the most relevant being the ones based on the Rayleigh-Ritz variational method[51,52]. The variational principle in ab initio quantum mechanics states that any approximate wave function will always have an energy expectation value that is above, the ground state energy[53,54]

$$\mathcal{E}[\Psi_t] \geq \mathcal{E}[\Psi_0]. \tag{2.13}$$

In Eq. (2.13), Ψ_t and Ψ_0 denote trial and ground state wave functions, respectively. The variational method allows to approximate a solution to the Schrödinger equation for many-electron systems, nevertheless the task is rather complicated. For an *N* electron system, Ψ depends on *3N* spatial coordinates. Thus, even for very simple molecules the number of variables becomes large. Therefore, a more pragmatic methodology for solving this problem is needed.

2.2 Density Functional Theory

Due to the difficulties for obtaining solutions of the electronic Schrödinger equation in many-electron systems, Thomas[55] and Fermi[56] suggested the use of the electronic density $\rho(\mathbf{r})$ rather than the wave function, $\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n)$, to calculate atomic properties. Further work by Dirac[57] as well as Wigner and Seitz[58,59] improved the model by introducing a local expression for the exchange potential. Although these methods do not yield results with chemical accuracy[60], Slater introduced the idea of approximating Fock exchange in the Hartree-Fock method by an average local potential based on the free-electron gas model[61]. The result was an exchange potential expressed solely in terms of $\rho(\mathbf{r})$. To improve the quality of this approximation an adjustable, semiempirical parameter α was introduced which led to the $X\alpha$ methodology of Slater and Johnson[62]. This method raised the hope that an approximated methodology based

on $\rho(\mathbf{r})$ instead of $\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n)$ could indeed be accurate enough to calculate molecules. The main advantage of such a methodology arises from the fact that the electronic density $\rho(\mathbf{r})$ depends only on 3 instead of *3N* spatial variables. It can be expressed as a measurable observable only depending on spatial coordinates

$$\rho(\mathbf{r}) = N \int ... \int \Psi^*(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N)\ \Psi(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r_N})\, d\mathbf{r_2} \ldots d\mathbf{r_N}. \tag{2.14}$$

For the sake of clarity in our discussion we omit spin dependency. Strictly speaking $\rho(\mathbf{r})$ is a probability density, but calling it the electron density is common practice. It represents the probability that one electron is within the volume element $d\mathbf{r}$ whereas the other *N*-1 electrons are in arbitrary positions. Clearly, $\rho(\mathbf{r})$ is a non-negative function of only three spatial variables which vanishes at infinity and integrates to the total number of electrons[41]

$$\begin{aligned} &\rho(\mathbf{r}) \geq 0\ \forall\, \mathbf{r}, \\ &\lim_{\mathbf{r}\to\infty} \rho(\mathbf{r}) = 0, \\ &\int \rho(\mathbf{r})\ d\mathbf{r} = N. \end{aligned} \tag{2.15}$$

A rigorous formulation for an *ab initio* theory based solely on $\rho(\mathbf{r})$ was first given by Hohenberg and Kohn in 1964, giving rise to the well known Density Functional Theory (DFT)[63]. The Hohenberg and Kohn formulation is based on two theorems:

First Hohenberg-Kohn theorem: *The external potential $v(\mathbf{r})$ is a unique functional of the electron density $\rho(\mathbf{r})$, apart from a trivial additive constant.*

To proof the first Hohenberg-Kohn theorem we assume that there exist two external potentials $v(\mathbf{r})$ and $v'(\mathbf{r})$ differing by more than a constant but which both give rise to the same electron density $\rho(\mathbf{r})$. These two potentials are part of two Hamilton operators which only differ in the external potential, $\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{v}$ and $\hat{H}' = \hat{T} + \hat{V}_{ee} + \hat{v}'$. Assuming non-degenerated ground states, the two Hamilton operators $\hat{H}$ and $\hat{H}'$ yield two different ground state wave functions, $\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n)$ and $\Psi'(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n)$, and corresponding ground state energies, $E_0 \neq E_0'$, respectively. Thus, $\Psi(\mathbf{r})$ and $\Psi'(\mathbf{r})$ are different, and we can use $\Psi'(\mathbf{r})$ as trial wave function for $\hat{H}$[64]. By virtue of the variational principle, we find in Dirac's notation[65]

$$E_0 = \langle\Psi|\hat{H}|\Psi\rangle < \langle\Psi'|\hat{H}|\Psi'\rangle = \langle\Psi'|\hat{H}' + \hat{v} - \hat{v}'|\Psi'\rangle = \langle\Psi'|\hat{H}'|\Psi'\rangle + \langle\Psi'|\hat{v} - \hat{v}'|\Psi'\rangle. \tag{2.16}$$

Similarly, taking $\Psi(\mathbf{r})$ as the trial wave function for $\hat{H}'$ yields

$$E_0' = \langle \Psi' | \hat{H}' | \Psi' \rangle < \langle \Psi | \hat{H}' | \Psi \rangle = \langle \Psi | \hat{H} + \hat{v}' - \hat{v} | \Psi \rangle = \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{v}' - \hat{v} | \Psi \rangle. \quad (2.17)$$

Eqs. (2.16) and (2.17) can be rewritten as

$$E_0 < E_0' + \int \Big[v(\mathbf{r}) - v'(\mathbf{r}) \Big] \, \rho(\mathbf{r}) \, d\mathbf{r}, \quad (2.18)$$

and

$$E_0' < E_0 + \int \Big[v'(\mathbf{r}) - v(\mathbf{r}) \Big] \, \rho(\mathbf{r}) \, d\mathbf{r}. \quad (2.19)$$

Adding Eqs. (2.18) and (2.19) yields

$$E_0 + E_0' < E_0' + E_0. \quad (2.20)$$

This proves by *reductio ad absurdum* that it is impossible to have two different $v(\mathbf{r})$ potentials that yield the same ground state electron density, or, in other words, that the ground state density is uniquely determined by the external potential $v(\mathbf{r})$ for a non-degenerated quantum mechanical system. The following mapping can be defined based on the first Hohenberg-Kohn theorem

$$\rho(\mathbf{r}) \leftrightarrow N, v(\mathbf{r}) \leftrightarrow \hat{H} \mapsto \Psi[\rho] \mapsto E[\rho]. \quad (2.21)$$

The consequence of the one-to-one correspondence between the electronic density $\rho(\mathbf{r})$ and the external potential $v(\mathbf{r})$ is that the wave function and energy are functionals of the density. Since the wave function is a functional of the ground state density, the expectation value of any observable $\hat{O}$ is also a unique functional of the density

$$\langle \hat{O} \rangle = \langle \Psi[\rho(\mathbf{r})] \mid \hat{O} \mid \Psi[\rho(\mathbf{r})] = O[\rho(\mathbf{r})]. \quad (2.22)$$

Hence all ground state properties for a non-degenerated system are entirely determined by the ground state electronic density. Among these observables is the ground state energy, the expectation value of the Hamilton operator, which is of great importance. Therefore, the ground-state wave function and the corresponding energy related to a potential $v(\mathbf{r})$ can be expressed as a

functional of $\rho(\mathbf{r})$

$$E[\rho(\mathbf{r})] = \langle \Psi[\rho(\mathbf{r})] | \hat{H} \Psi[\rho(\mathbf{r})] \rangle = T_e[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + \int v(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}. \tag{2.23}$$

To obtain a more convenient form of Eq. (2.23), we introduce the universal Hohenberg-Kohn functional $F[\rho(\mathbf{r})]$ as[63]

$$F[\rho(\mathbf{r})] = T_e[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})]. \tag{2.24}$$

Thus, the energy can be written in the form

$$E[\rho(\mathbf{r})] = F[\rho(\mathbf{r})] + \int \rho(\mathbf{r})\ v(\mathbf{r})\ d\mathbf{r}. \tag{2.25}$$

The term universal arises because $F[\rho(\mathbf{r})]$ does not depend on the external potential $v(\mathbf{r})$. It is the same for atoms, molecules and solids as well as model systems and, therefore, it is a universal functional of $\rho(\mathbf{r})$[66].

Based on the first Hohenberg-Kohn theorem, the variational energy principle can be used to obtain the ground state density as proven by the following theorem:

Second Hohenberg-Kohn theorem: *The ground state density $\rho(\mathbf{r})$ can be determined from the ground state energy functional $E[\rho(\mathbf{r})]$ via the variational energy principle by variation of the density only.*

The ground state energy E_0 which is a functional of the density is given as

$$E_0 = E[\rho_0(\mathbf{r})] = \langle \Psi[\rho_0(\mathbf{r})] \mid \hat{H} \mid \Psi[\rho_0(\mathbf{r})] \rangle. \tag{2.26}$$

For a trial density $\rho_t(\mathbf{r})$, such that $\int \rho_t(\mathbf{r}) d\mathbf{r} = N$ and $\rho_t(\mathbf{r}) \geq 0$, it follows from the first Hohenberg-Kohn theorem that this trial density determines the corresponding external potential $v_t(\mathbf{r})$ and via the resulting Hamilton operator the trial wave function $\Psi_t[\rho_t(\mathbf{r})]$. From the variational energy principle follows

$$E[\rho_t(\mathbf{r})] = \langle \Psi_t[\rho_t(\mathbf{r})] \mid \hat{H} \mid \Psi_t[\rho_t(\mathbf{r})] \rangle \geq E[\rho_0(\mathbf{r})] = E_0 \tag{2.27}$$

Thus, the ground state energy, E_0, and density, $\rho_0(\mathbf{r})$, can be obtained by minimization of the functional $E[\rho(\mathbf{r})]$ of Eq. (2.23) for arbitrary variations of the density. Unfortunately the univer-

sal Hohenberg-Kohn functional $F[\rho(\mathbf{r})]$ remains to be unknown. Although the Hohenberg-Kohn theorem establishes the fact that the wave function and energy are functionals of the ground state electronic density, it does not, however, prescribe the explicit dependencies of Ψ and E from $\rho(\mathbf{r})$.

2.3 Kohn-Sham Methodology

As can be seen from the proof of the first Hohenberg-Kohn theorem, the functional $F[\rho]$ is independent of the external potential $v(\mathbf{r})$ and holds for any number of particles. For this reason, it is called the universal Hohenberg-Kohn functional. If the exact universal functional $F[\rho]$ was known, DFT would be an exact formulation. However, accurate implementations of DFT are far from easy to achieve because of the unfortunate fact that $F[\rho]$ is hard to approximate in a close form[66]. From its appearance in the Hohenberg-Kohn theorem, the universal functional $F[\rho]$ is defined as the expectation value of the Hamilton operator without the external potential. For this reason, $F[\rho]$ is given by the sum of the electronic kinetic energy and electron-electron interactions

$$F[\rho] = T[\rho] + V_{ee}[\rho] = T[\rho] + J[\rho] + V_{ee}^{nc}[\rho], \tag{2.28}$$

where $T[\rho]$ is the kinetic energy functional and $V_{ee}[\rho]$ is a functional that accounts for the electron-electron interaction energy. $V_{ee}[\rho]$ can be split into $J[\rho]$, the classical Coulomb interaction, and $V_{ee}^{nc}[\rho]$, that collects all non-classical electron-electron interaction, known as exchange-correlation energy. Among these functionals only $J[\rho]$ is known, while the explicit form of the other two contributions remains unknown.

A practical solution for the determination of the kinetic energy was provided by Kohn and Sham[67]. They proposed a fictitious system of non-interacting particles that generates the same density as the corresponding system of interacting particles. Their very clever idea was to realize that, if we are not able to accurately determine the kinetic energy through an explicit functional, we should be a bit less ambitious and concentrate on computing as much as we can of the true kinetic energy exactly. We then have to deal with the remainder in an approximate manner. This is based on decomposing $T[\rho]$ into a part that represents the kinetic energy of a non-interacting system of electrons, $T_s[\rho]$, and a remainder, $T_c[\rho]$,

$$T[\rho] = T_s[\rho] + T_c[\rho]. \tag{2.29}$$

The subscripts *s* and *c* stand for *single-particle* and *correlation*, respectively[68]. In the non-

interacting system, the total kinetic energy $T_s[\rho]$ is just the sum of the single-particle kinetic energies

$$\begin{aligned} T_s[\rho] &= \langle \Psi[\rho] | \hat{T} | \Psi[\rho] \rangle \\ &= \sum_{i}^{occ} \langle \psi_i | \hat{T} | \psi_i \rangle, \end{aligned} \tag{2.30}$$

where $\Psi[\rho]$ is the Slater determinant[69] forming the density $\rho(\mathbf{r})$, and $\psi_i(\mathbf{r})$ are the single-particle orbitals of the non-interacting system. The energy functional can be rewritten using Eqs. (2.28) and (2.29) as

$$\begin{aligned} E[\rho] &= T_s[\rho] + \int v(\mathbf{r})\ \rho(\mathbf{r})\ d\mathbf{r} + J[\rho] + V_{ee}^{nc}[\rho] + T_c[\rho] \\ &= T_s[\rho] + V[\rho] + J[\rho] + E_{xc}[\rho], \end{aligned} \tag{2.31}$$

where

$$V[\rho] = \int v(\mathbf{r})\ \rho(\mathbf{r})\ d\mathbf{r}, \tag{2.32}$$

and $E_{xc}[\rho]$ is the *exchange-correlation functional* that contains $T_c[\rho]$ and $V_{ee}^{nc}[\rho]$. Eq. (2.31) is formally exact but $E_{xc}[\rho]$ remains unknown. Since $T_s[\rho]$ is not an explicit functional of $\rho(\mathbf{r})$, Eq. (2.31) cannot be directly minimized. However, Kohn and Sham suggested a scheme where the minimization is carried out in an indirect form. To this end, Kohn and Sham related the minimization condition for a fully interacting system with that of a non-interacting system. For the fully interacting system, the minimization condition is given by

$$\begin{aligned} \frac{\delta E[\rho]}{\delta \rho(\mathbf{r})} &= \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta V[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \\ &= \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} + v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) \equiv 0. \end{aligned} \tag{2.33}$$

The term $\frac{\delta V[\rho]}{\delta \rho(\mathbf{r})}$ yields the external potential $v(\mathbf{r})$. The term $\frac{\delta J[\rho]}{\delta \rho(\mathbf{r})}$ the Hartree (Coulomb) potential, $v_H(\mathbf{r})$, and, once an explicit form for $E_{xc}[\rho]$ is chosen, the term $\frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}$ yields the exchange-correlation potential, $v_{xc}(\mathbf{r})$.

Consider now a system of non-interacting particles moving in a potential $v_s(\mathbf{r})$. The minimization condition is simply

$$\frac{\delta E_s[\rho_s]}{\delta \rho_s(\mathbf{r})} = \frac{\delta T_s[\rho_s]}{\delta \rho_s(\mathbf{r})} + \frac{\delta V_s[\rho_s]}{\delta \rho_s(\mathbf{r})}$$
$$= \frac{\delta T_s[\rho_s]}{\delta \rho_s(\mathbf{r})} + v_s(\mathbf{r}) \equiv 0, \tag{2.34}$$

since there are no Hartree and exchange-correlation terms in the absence of interactions. Comparing Eq. (2.33) with Eq. (2.34) we find that both minimization have the same solution, $\rho_s(\mathbf{r}) \equiv \rho(\mathbf{r})$, if $v_s(\mathbf{r})$ is chosen to be

$$\frac{\delta V_s[\rho]}{\delta \rho(\mathbf{r})} = \frac{\delta V[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}, \tag{2.35}$$

i.e.

$$v_s(\mathbf{r}) \equiv v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}). \tag{2.36}$$

Consequently, one can calculate the density of the interacting system in the external potential $v(\mathbf{r})$ by solving the equation of a non-interacting system in the potential $v_s(\mathbf{r})$. To this end, Kohn and Sham represented the non-interacting wave function as a single Slater determinant and obtained $T_s[\rho]$ through the Kohn-Sham orbitals $\psi_i(\mathbf{r})$ in the same way as in Eq. (2.30). Solving the set of single-particle Schrödinger like equations

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \, \psi_i(\mathbf{r}), \tag{2.37}$$

yields orbitals that produce, by choosing $v_s(\mathbf{r})$ according to Eq. (2.36), the density $\rho(\mathbf{r})$ of the original system

$$\rho(\mathbf{r}) = \rho_s(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2. \tag{2.38}$$

Eq. (2.37) are the celebrated Kohn-Sham equations. Since both $v_H(\mathbf{r})$ and $v_{xc}(\mathbf{r})$ depend on $\rho(\mathbf{r})$, which depends on the $\psi_i(\mathbf{r})$, which in turn depend on $v_s(\mathbf{r})$, the problem of solving the Kohn-Sham equation is a nonlinear one.

As described above, the Kohn-Sham method eliminates the unknown kinetic energy functional by introducing orbitals of a fictitious non-interacting reference system. However, the exchange-correlation energy functional is still unknown. In fact, the quality of any DFT calculation using

the Kohn-Sham method is determined mainly by the approximation used for the evaluation of $E_{xc}[\rho]$. Different types of approximations for $E_{xc}[\rho]$ have been used, e.g. the Local Density Approximation (LDA) in which the Dirac exchange[57] is combined with a fit to the homogeneous electron gas correlation, like the one proposed by Vosko, Wilk and Nusair (VWN)[70]. More sophisticated approaches include the Generalized Gradient Approximations (GGA)[71,72] like the Becke, Lee, Yang and Parr (BLYP)[73–76] and Perdew, Burke and Ernzerhof (PBE)[77] functionals or hybrid functionals, which include the exact exchange energy[78] like B3LYP. The development of these, and many others, highly accurate density functional approximations have resulted in an exponentially growing attention to DFT, to the point where Kohn-Sham DFT has become the standard tool in electronic structure theory calculations[54].

3

Density Fitting Methodologies

Contents

3.1 LCGTO Approximation

The Kohn-Sham single particle equations, Eq. (2.37), given in the previous section, represent a complicated system of coupled integro-differential equations (the kinetic energy operator is a differential operator, while the Coulomb contribution is expressed through an integral operator). Therefore, it is necessary to find a computationally efficient way of solving these equations. In principle, a purely numerical approach to solve these equations is possible and a few benchmark calculations for atoms and small molecules using such a technique are available[79]. However, numerical procedures are much too demanding for routine applications and other techniques are required. Almost all applications of Kohn-Sham DFT to finite systems make use of the linear combination of atomic orbitals (LCAO) expansion of the Kohn-Sham molecular orbitals, a scheme introduced by Mulliken and adapted by Roothaan. In LCAO the molecular orbitals (MOs), $\psi_i(\mathbf{r})$, are expanded as a linear combination of atomic orbitals

$$\psi_i(\mathbf{r}) = \sum_{\mu} c_{\mu i}\, \mu(\mathbf{r}). \tag{3.1}$$

Within the LCAO approximation, the Kohn-Sham equations, can be obtained by varying the LCAO Kohn-Sham energy expression subject to the MO orthonormality constraint

$$\langle \psi_i | \psi_j \rangle = \delta_{ij}, \tag{3.2}$$

where δ_{ij} is the Kronecker delta. In Eq. (3.1), $\mu(\mathbf{r})$ represents an atomic orbital or, more general, a basis function, and $c_{\mu i}$ is a molecular orbital coefficient. In deMon2k[48,80–82] the basis functions are atom-centered (contracted) Gaussian type orbitals (GTO), hence the working ansatz for deMon2k is known as linear combination of Gaussian type orbitals (LCGTO) approximation. An unnormalized Cartesian GTO is given by[83]

$$\mu(\mathbf{r}) = (x - A_x)^{a_x}(y - A_y)^{a_y}(z - A_z)^{a_z} \sum_{k}^{K} d_k e^{\zeta_k (\mathbf{r}-\mathbf{A})^2}. \tag{3.3}$$

A basis function is completely defined by its atomic center $\mathbf{A}$, its angular momentum vector $\mathbf{a} = (a_x, a_y, a_z)$, the degree of contraction K, the contraction coefficients d_k and the orbital exponents ζ_k. All these parameters remain constant for a given geometry. Thus only the molecular orbital coefficients are the variational parameters for which Eq. (2.31) is minimized. Using the explicit form for all the known terms in Eq. (2.31) and assuming a closed-shell system (the

extension to the open-shell formalism is straightforward[84–86]) yields

$$E = -\sum_{i}^{occ}\langle\psi_i|\nabla^2\psi_i\rangle - \sum_{A}\int \frac{Z_A}{|\mathbf{r}-\mathbf{A}|}\,\rho(\mathbf{r})\,\mathbf{dr} + \frac{1}{2}\iint \frac{\rho(\mathbf{r_1})\rho(\mathbf{r_2})}{|\mathbf{r_1}-\mathbf{r_2}|}\,d\mathbf{r_1}d\mathbf{r_2} + E_{xc}[\rho], \quad (3.4)$$

where the upper index *occ* refers to all doubly occupied spatial orbitals in the closed-shell system. The corresponding LCGTO expansion of the electronic density is given by

$$\rho(\mathbf{r}) = 2\sum_{i}^{occ}|\psi_i(\mathbf{r})|^2 = 2\sum_{i}^{occ}\sum_{\mu,\nu} c_{\mu i}c_{\nu i}\mu(\mathbf{r})\nu(\mathbf{r}) = \sum_{\mu,\nu} P_{\mu\nu}\mu(\mathbf{r})\nu(\mathbf{r}). \quad (3.5)$$

Here $P_{\mu\nu}$ is an element of the closed-shell density matrix defined as

$$P_{\mu\nu} = 2\sum_{i}^{occ} c_{\mu i}c_{\nu i}. \quad (3.6)$$

Using the above expansion for the Kohn-Sham orbitals, Eq. (3.1), and the density, Eq. (3.5), the Kohn-Sham energy expression, Eq. (3.4), can be rewritten as

$$\begin{aligned} E = &-\frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu}\langle\mu|\nabla^2\nu\rangle - \sum_{\mu,\nu}\sum_{A} P_{\mu\nu}\left\langle\mu\left|\frac{Z_A}{|\mathbf{r}-\mathbf{A}|}\right|\nu\right\rangle + \\ &\frac{1}{2}\sum_{\mu,\nu}\sum_{\sigma,\tau} P_{\mu\nu}P_{\sigma\tau}\iint \frac{\mu(\mathbf{r_1})\nu(\mathbf{r_1})\sigma(\mathbf{r_2})\tau(\mathbf{r_2})}{|\mathbf{r_1}-\mathbf{r_2}|}d\mathbf{r_1}d\mathbf{r_2} + E_{xc}[\rho]. \end{aligned} \quad (3.7)$$

The first two-terms of Eq. (3.7) represent the one-electron energy, often named the core energy. The third term is the electronic repulsion energy, hence the integrals appearing in it are named electron repulsion integrals (ERIs). To ease notation, we introduce the core Hamilton matrix, $\mathbf{H}$, with elements

$$H_{\mu\nu} = -\frac{1}{2}\langle\mu|\nabla^2\nu\rangle - \sum_{A}\left\langle\mu\left|\frac{Z_A}{|\mathbf{r}-\mathbf{A}|}\right|\nu\right\rangle, \quad (3.8)$$

and a shorthand notation for the electron repulsion integrals

$$\langle\mu\nu||\sigma\tau\rangle = \iint \frac{\mu(\mathbf{r_1})\nu(\mathbf{r_1})\sigma(\mathbf{r_2})\tau(\mathbf{r_2})}{|\mathbf{r_1}-\mathbf{r_2}|}d\mathbf{r_1}d\mathbf{r_2}. \quad (3.9)$$

In this ERI notation the double vertical bar represents the two-electron operator $1/|\mathbf{r_1}-\mathbf{r_2}|$. It also separates the functions that depend on the electronic coordinate $\mathbf{r_1}$ (in the bra), from those

that depend on the electronic coordinate $\mathbf{r_2}$ (in the ket). Analog notations will be used for other types of ERIs throughout the text. With these shorthand notations we can rewrite Eq. (3.7) as

$$E = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu || \sigma\tau \rangle + E_{xc}[\rho]. \tag{3.10}$$

To derive the Kohn-Sham equations, we minimize the energy expression in Eq. (3.10) with respect to the molecular orbitals coefficients under the constraint of MO orthonormality, Eq. (3.2), which we can write in the form

$$\langle \psi_i | \psi_j \rangle = \sum_{\mu,\nu} c_{\mu i} S_{\mu\nu} c_{\nu j} = \delta_{ij}. \tag{3.11}$$

Imposing these constraints to the LCGTO energy expression leads to the Lagrange function

$$L = E - 2 \sum_{i,j}^{all} \lambda_{ij} \left(\sum_{\mu,\nu} c_{\mu i} \, S_{\mu\nu} \, c_{\nu j} - \delta_{ij} \right). \tag{3.12}$$

The variation of the Lagrange function,

$$\begin{aligned} \frac{\partial L}{\partial c_{\mu i}} &= \frac{\partial E}{\partial c_{\mu i}} - 4 \sum_{j} \sum_{\nu} S_{\mu\nu} c_{\nu j} \lambda_{ji} \\ &= 4 \sum_{\nu} \left(H_{\mu\nu} + \sum_{\sigma,\tau} P_{\sigma\tau} \langle \mu\nu \| \sigma\tau \rangle + \langle \mu | v_{xc}[\rho] | \nu \rangle \right) c_{\nu i} - 4 \sum_{j}^{all} \sum_{\nu} S_{\mu\nu} c_{\nu j} \lambda_{ji}, \end{aligned} \tag{3.13}$$

must vanish at a stationary point. To obtain Eq. (3.13) the variation of $E_{xc}[\rho]$ was performed using the chain rule

$$\frac{\partial E_{xc}[\rho]}{\partial c_{\mu i}} = \int \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \frac{\partial \rho(\mathbf{r})}{\partial c_{\mu i}} d\mathbf{r} = 4 \sum_{\nu} c_{\nu i} \int v_{xc}[\rho; \mathbf{r}] \, \mu(\mathbf{r}) \, \nu(\mathbf{r}) \, d\mathbf{r}. \tag{3.14}$$

At this point it is convenient to define the Kohn-Sham matrix $\mathbf{K}$. This matrix represents the variation of the energy with respect to density matrix elements and its elements are given by

$$K_{\mu\nu} \equiv \frac{\partial E}{\partial P_{\mu\nu}} = H_{\mu\nu} + \sum_{\sigma,\tau} P_{\sigma\tau} \, \langle \sigma\tau \| \mu\nu \rangle + \langle \mu | v_{xc}[\rho] | \nu \rangle. \tag{3.15}$$

Substitution of the Eq. (3.15) into Eq. (3.13) under the minimization condition

$$\frac{\partial L}{\partial c_{\mu i}} = 0 \quad \forall\, c_{\mu i}, \tag{3.16}$$

yields

$$\sum_{\nu} K_{\mu\nu} c_{\nu i} = \sum_{\nu} \sum_{j} S_{\mu\nu} c_{\nu j} \lambda_{ji} \quad \forall\, i. \tag{3.17}$$

Eq. (3.17) is a generalized eigenvalue equation[87]. There exists one equation of the form of Eq. (3.17) for each molecular orbital. Collecting all equations into a single matrix equation yields

$$\mathbf{Kc} = \mathbf{Sc}\boldsymbol{\lambda} \tag{3.18}$$

This set of equations has the same form as the famous Roothaan-Hall equations[88,89] appearing in Hartree-Fock theory. Here, $\mathbf{c} = (\mathbf{c_1}, \mathbf{c_2}, ..., \mathbf{c_{occ}}, ..., \mathbf{c_{all}})$, is a squared matrix composed from all occupied, *occ*, and unoccupied, *uno*, molecular orbital coefficient vectors. It is important to note that Eq. (3.18) is a nonlinear generalized eigenvalue equation because the Kohn-Sham matrix depends on, at least, the occupied subspace of $\mathbf{c}$. Therefore, Eq. (3.18) has to be solved iteratively starting from an educated guess for $\mathbf{c}$. It is important to note that $\mathbf{K}$ is invariant under separate orthogonal transformations of the occupied and unoccupied subspaces of $\mathbf{c}$. This property follows from the fact that the electronic density (and hence, also the density matrix) is invariant to such orthogonal transformations

$$\begin{aligned} \mathbf{P} &= 2\, \mathbf{c_{occ}} \mathbf{c_{occ}^T} \\ &= 2\, \mathbf{c_{occ}} \mathbf{U_{occ}} \mathbf{U_{occ}^T} \mathbf{c_{occ}^T}. \end{aligned} \tag{3.19}$$

Thus, the rotations between the occupied and the unoccupied subspaces are responsible for the change of $\mathbf{K}$ between two iterations of the self-consistent field (SCF) procedure. It follows that, at convergence, the Lagrange multiplier matrix $\boldsymbol{\varepsilon}$ must have, in general, the following block diagonal form:

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\lambda_{occ}} & 0 \\ 0 & \boldsymbol{\lambda_{uno}} \end{pmatrix}. \tag{3.20}$$

Therefore, we can choose to work in the molecular orbital representation $\mathbf{cU}$ where $\mathbf{U}$ is a block-diagonal orthogonal matrix

$$\mathbf{U} = \begin{pmatrix} \mathbf{U_{occ}} & 0 \\ 0 & \mathbf{U_{uno}} \end{pmatrix}, \tag{3.21}$$

such that $\boldsymbol{\varepsilon} = \mathbf{U^T \lambda U}$ is a strictly diagonal matrix. Such molecular orbital representation is called *canonical*. From now on we will assume that $\mathbf{c}$ are the canonical MO coefficients and, therefore, $\boldsymbol{\varepsilon}$ is a diagonal matrix collecting the corresponding MO energies.

In order to determine the computational demand for the described LCGTO Kohn-Sham methodology, we now analyze the formal scaling of the individual terms in the energy expression, Eq. (3.10), and Kohn-Sham matrix, Eq. (3.15), calculation. The computation of the core Hamilton matrix $\mathbf{H}$ scales formally quadratic with the number of basis functions, N_{bas}. The same scaling is observed for the computation of the overlap matrix $\mathbf{S}$. Both matrices, $\mathbf{H}$ and $\mathbf{S}$, remain constant during the whole SCF procedure and, therefore, are computed only once and stored. The Coulomb term introduces a formal N_{bas}^4 scaling into the energy calculation. The exchange-correlation contribution has a formal $N_{bas}^2 \times G$ scaling, whereas G is the number of grid points necessary for the numerical integration needed to compute the exchange-correlation contribution. From the above discussion follows immediately that the calculation of the Coulomb repulsion energy represents the computationally most demanding task in Eqs. (3.10) and (3.15). In the next section, we show how the introduction of the variational fitting of the Coulomb potential reduces the formal scaling of this term using auxiliary functions.

3.2 Variational Fitting of the Coulomb Potential

A very popular technique to reduce the formal scaling of computing the Coulomb potential is the so-called variational fitting of the Coulomb potential. This technique was introduced by Dunlap *et al*[2–5], inspired by a former work of Sambe and Felton[90]. It was originally introduced in the deMon-KS[91] and DGauss[92] programs more than 30 years ago. This formulation is equivalent to the application of the so-called resolution of the identity (RI)[6,93] for the Coulomb integrals used in other programs like Gaussian and Turbomole. A more extensive discussion of the influence of the variational fitting technique on electronic structure calculations can be found in the literature[8,94]. The variational approximation of the Coulomb potential, as implemented in deMon2k, is based on the minimization of the following positive semi-definite second-order

fitting error

$$\mathcal{E}_2 = \frac{1}{2} \iint \frac{[\rho(\mathbf{r}_1) - \tilde{\rho}(\mathbf{r}_1)][\rho(\mathbf{r}_2) - \tilde{\rho}(\mathbf{r}_2)]}{|\mathbf{r}_1 - \mathbf{r}_2|} \, d\mathbf{r}_1 d\mathbf{r}_2 \geq 0. \tag{3.22}$$

In deMon2k the approximated density, $\tilde{\rho}(\mathbf{r})$, is expanded as a linear combination of atom centered primitive Hermite-Gaussian auxiliary functions[95,96], $\bar{k}(\mathbf{r})$,

$$\tilde{\rho}(\mathbf{r}) = \sum_{\bar{k}} x_{\bar{k}} \bar{k}(\mathbf{r}). \tag{3.23}$$

An unnormalized primitive Hermite-Gaussian auxiliary function, $\bar{k}(\mathbf{r})$, centered on atom A with exponent $\zeta_{\bar{k}}$ has the following form

$$\bar{k}(\mathbf{r}) = \left(\frac{\partial}{\partial A_x}\right)^{\bar{k}_x} \left(\frac{\partial}{\partial A_y}\right)^{\bar{k}_y} \left(\frac{\partial}{\partial A_z}\right)^{\bar{k}_z} e^{-\zeta_{\bar{k}}(\mathbf{r}-\mathbf{A})^2}. \tag{3.24}$$

As for the GTO basis functions, all parameters appearing in Eq. (3.24) remain constant during an electronic structure calculation unless the geometry of the molecule is changed. In deMon2k the auxiliary functions are grouped into *s*, *spd*, *spdfg* and *spdfghi* sets sharing the same exponent within each set[97,98]. In the automatic generation of auxiliary functions[99], indicated by the abbreviation GEN in deMon2k, the exponent range for the auxiliary function sets is determined by the smallest, ζ_{min}, and largest, ζ_{max}, primitive Gaussian exponent of the basis set used. Specially developed integral recurrence relations[95,100] ensure maximum performance in the analytic molecular integral calculations with these auxiliary function sets. Expanding $\rho(\mathbf{r})$ and $\tilde{\rho}(\mathbf{r})$ in Eq. (3.22) yields

$$\mathcal{E}_2 = \frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu \| \sigma\tau \rangle - \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} + \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} \langle \bar{k} \| \bar{l} \rangle x_{\bar{l}}. \tag{3.25}$$

Since $\mathcal{E}_2$ is positive semi-definite[101] the following inequality holds

$$\frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu \| \sigma\tau \rangle \geq \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} \langle \bar{k} \| \bar{l} \rangle x_{\bar{l}}. \tag{3.26}$$

Note that the equality holds only when $\rho(\mathbf{r})$ equals $\tilde{\rho}(\mathbf{r})$. Thus, the true Coulomb repulsion energy provides an upper bound to any approximated one. In this context it is worth to point out a common misconception in the literature regarding variationally density fitting. It exists several works in which auxiliary function sets are "optimized" in order to reproduce the total energy

corresponding to calculations without fitting[102–104]. However, such an approach is as useful as optimizing basis sets for total rather than relative energies. Instead, variational density fitting approaches should be understood as own methodologies and their quality should be judged according to their accuracy in the calculation of relative energies, i.e. atomization energies, as it is common for basis sets. With the inequality from Eq. (3.26) we can write a new variational energy expression that incorporates the variational fitting of the Coulomb potential

$$E = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} \langle \bar{k} \| \bar{l} \rangle x_{\bar{l}} + E_{xc}[\rho]. \tag{3.27}$$

For short we call this the Kohn-Sham DFT approach. The BASIS[105] option of the VXCTYPE keyword in deMon2k triggers the use of this energy expression. Minimizing $\mathcal{E}_2$ with respect to $x_{\bar{m}}$ yields

$$\frac{\partial \mathcal{E}_2}{\partial x_{\bar{m}}} = -\sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{m} \rangle + \sum_{\bar{l}} x_{\bar{l}} \langle \bar{l} \| \bar{m} \rangle \equiv 0 \quad \forall\, \bar{m}. \tag{3.28}$$

The linear equation system in Eq. (3.28) can be written in short form as

$$\mathbf{G}\mathbf{x} = \mathbf{J} \tag{3.29}$$

where $\mathbf{G}$ is the Coulomb matrix defined as

$$\mathbf{G} = \begin{pmatrix} \langle \bar{1} \| \bar{1} \rangle & \langle \bar{1} \| \bar{2} \rangle & \dots & \langle \bar{1} \| \bar{m} \rangle \\ \langle \bar{2} \| \bar{1} \rangle & \langle \bar{2} \| \bar{2} \rangle & \dots & \langle \bar{2} \| \bar{m} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \bar{m} \| \bar{1} \rangle & \langle \bar{m} \| \bar{2} \rangle & \dots & \langle \bar{m} \| \bar{m} \rangle \end{pmatrix}, \tag{3.30}$$

and $\mathbf{J}$ is the Coulomb vector given by

$$\mathbf{J} = \begin{pmatrix} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{1} \rangle \\ \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{2} \rangle \\ \vdots \\ \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{m} \rangle \end{pmatrix}. \tag{3.31}$$

The fitting coefficients are collected in the vector $\mathbf{x}$. A straightforward solution of Eq. (3.29) is obtained by the inversion of the Coulomb matrix $\mathbf{G}$

$$\mathbf{x} = \mathbf{G}^{-1}\mathbf{J}. \tag{3.32}$$

However, the $\mathbf{G}$ matrix, although formally positive definite, can become in finite precision number models numerically indefinite with small negative eigenvalues. Such ill-conditioning arises commonly due to large auxiliary function sets. Normalization of the auxiliary functions with respect to the Coulomb norm

$$\langle \bar{k} || \bar{k} \rangle = 1 \quad \forall\, \bar{k}, \tag{3.33}$$

improves slightly the numerical stability of $\mathbf{G}$, nevertheless, it can still be ill-conditioned. To overcome this numerical instability a truncated eigenvalue decomposition (TED) of $\mathbf{G}$ is performed in deMon2k. Once the positive definiteness of $\mathbf{G}$ is ensured, Eq. (3.29) is solved by a numerically robust quasi-Newton method[36]. Although, the TED is performed only one time at the beginning of a calculation the involved diagonalization of the $\mathbf{G}$ matrix can become a critical computational bottleneck if more than hundred thousand auxiliary functions are used.

Once the fitting equation system, Eq. (3.29) has been solved for a particular density, the Kohn-Sham matrix, $\mathbf{K}$, can be obtained by varying Eq. (3.27) with respect to density matrix elements

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu || \bar{k} \rangle \, x_{\bar{k}} + \langle \mu | v_{xc}[\rho] | \nu \rangle. \tag{3.34}$$

Thus, the four-center ERIs, have been substituted by three and two-center ERIs in the expressions for the energy, Eq. (3.27), and Kohn-Sham matrix elements, Eq. (3.34). Because the Kohn-Sham density changes with each SCF step, the density fitting must be performed in each SCF step. With the Kohn-Sham and auxiliary density at hand the Kohn-Sham matrix can be calculated. Diagonalization of the Kohn-Sham matrix yields new molecular orbital coefficients from which a new density matrix and in turn a new Coulomb vector can be calculated. The new auxiliary density is then obtained by solving the fitting equation system, Eq. (3.29). This procedure is repeated in each SCF cycle until energy convergence is achieved. Due to the variational fitting of the Coulomb potential, the computational bottleneck for medium sized systems shifts to the numberical integration of the exhange-correlation energy and potential.

3.3 Auxiliary Density Functional Theory

The use of auxiliary functions for the calculation of the exchange-correlation potential has a long history in DFT methods[90,106]. In programs like deMon-KS[91] or DGAUSS[92] the exchange-correlation potential is expanded in Cartesian Gaussian functions as proposed by Sambe and Felton[90]. The expansion coefficients are obtained by a least squares fit on a small grid. A serious drawback of this approximation is that neither the fit nor the energy expression are variational and, therefore, reliable forces (and higher order derivatives) cannot be obtained[107,108]. As an alternative to this approach the direct use of the auxiliary density $\tilde{\rho}(\mathbf{r})$, obtained from the variational fitting of the Coulomb potential, for the calculation of the exchange-correlation energy and potentials[31–33,109–111] has been suggested. The resulting approximation was named Auxiliary Density Functional Theory (ADFT) and is triggered by the AUXIS[105] option of the VXCTYPE keyword in deMon2k[112]. If the auxiliary density is used for the evaluation of the exchange-correlation potential it is desirable that it is positive definite, $\tilde{\rho}(\mathbf{r}) \geq 0$, and integrates to the number of electrons of the system, $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$. The normalization to the number of electrons can be included as a constraint in the fitting equations. However, even without this constraint the number of electrons is conserved to high accuracy. The introduction of the positive semi-definiteness for $\tilde{\rho}(\mathbf{r})$ is less straightforward. Fortunately, regions where $\tilde{\rho}(\mathbf{r}) < 0$ are rather small and usually occur when $\rho(\mathbf{r}) \approx 0$[113]. Therefore, screening of these points does not impact the accuracy of the methodology[32]. The energy expression in ADFT takes the form

$$E = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} \langle \bar{k} \| \bar{l} \rangle x_{\bar{l}} + E_{xc}[\tilde{\rho}]. \tag{3.35}$$

The ADFT approach is the default method for calculating the exchange-correlation contributions in deMon2k. The variation of this energy expression with respect to density matrix elements yields the corresponding Kohn-Sham elements

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} + \frac{\partial E_{xc}[\tilde{\rho}]}{\partial P_{\mu\nu}} \tag{3.36}$$

The last term of Eq. (3.36) can be evaluated in a similar way as in Eq. (3.14)

$$\frac{\partial E_{xc}[\tilde{\rho}]}{\partial P_{\mu\nu}} = \int \frac{\delta E_{xc}[\tilde{\rho}(\mathbf{r})]}{\delta \tilde{\rho}(\mathbf{r})} \frac{\partial \tilde{\rho}(\mathbf{r})}{\partial P_{\mu\nu}} \mathbf{dr} = \sum_{\bar{k}} \frac{\partial x_{\bar{k}}}{\partial P_{\mu\nu}} \int v_{xc}[\tilde{\rho}; \mathbf{r}] \, \bar{k}(\mathbf{r}) \, d\mathbf{r}, \tag{3.37}$$

with

$$v_{xc}[\tilde{\rho};\mathbf{r}] \equiv \frac{\delta E_{xc}[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r})}. \tag{3.38}$$

The derivatives of the Coulomb fitting coefficients are obtained using Eqs. (3.31) and (3.32) to yield

$$\frac{\partial x_{\bar{k}}}{\partial P_{\mu\nu}} = \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}} \langle \bar{l} \| \mu\nu \rangle. \tag{3.39}$$

Inserting Eq. (3.39) into Eq. (3.37) yields

$$\frac{\partial E_{xc}[\tilde{\rho}]}{\partial P_{\mu\nu}} = \sum_{\bar{k},\bar{l}} \langle \mu\nu \| \bar{k} \rangle G^{-1}_{\bar{k}\bar{1}} \langle \bar{l} | v_{xc}[\tilde{\rho}] \rangle. \tag{3.40}$$

To simplify the notation, we now introduce the exchange-correlation fitting coefficient vector $\mathbf{z}$ with the following elements

$$z_{\bar{k}} = \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}} \langle \bar{l} | v_{xc}[\tilde{\rho}] \rangle. \tag{3.41}$$

The expression for the Kohn-Sham matrix elements can, thus, be rewritten in the form

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle (x_{\bar{k}} + z_{\bar{k}}). \tag{3.42}$$

It is important to note that $\mathbf{z}$ is spin-dependent and accounts for the difference between the α and β Kohn-Sham matrices in open-shell calculations. For the actual calculation of the exchange-correlation fitting coefficient, $z_{\bar{k}}$, we reformulate Eq. (3.41) as an inhomogeneous equation system of the form

$$\mathbf{G}\mathbf{z} = \mathbf{L}, \tag{3.43}$$

with

$$\mathbf{L} = \begin{pmatrix} \langle v_{xc}[\tilde{\rho}] | \bar{1} \rangle \\ \langle v_{xc}[\tilde{\rho}] | \bar{2} \rangle \\ \vdots \\ \langle v_{xc}[\tilde{\rho}] | \bar{m} \rangle \end{pmatrix}. \tag{3.44}$$

Here $\mathbf{L}$ is the exchange-correlation vector and $\mathbf{z}$ collects the exchange-correlation coefficients. In deMon2k a preconditioned conjugate gradient iterative solver for Eq. (3.43) was proposed by Domínguez-Soria[36]. This conjugate gradient algorithm uses the quasi-Newton updated Coulomb matrix for the iterative solution of the Coulomb fitting equation system, Eq. (3.28), as preconditioner. In order to keep the approach variational, $\tilde{\rho}(\mathbf{r})$ must be taken unaltered

from the solution of Eq. (3.32) to calculate $v_{xc}[\tilde{\rho}]$. However, this is not mandatory for the calculation of the Coulomb contribution. Therefore, two sets of Coulomb fitting coefficients are generally available in a deMon2k calculation, one set is obtained directly from Eq. (3.32) and is used for the evaluation of the exchange-correlation potential in order to keep the calculation variational whereas the other set results from SCF convergence and acceleration techniques, such as auxiliary density mixing[112] and DIIS[112,114,115], and is used for the calculation of the approximated two-electron Coulomb energy and the corresponding potential.

4

Variational Density Fitting with Krylov Subspace Methods

Contents

4.1 Methods for solving linear equation systems in deMon2k

4.1.1 Iterative methods for solving linear equation systems

Due to the reduced scaling of the analytical and numerical molecular integral calculation in ADFT, the linear algebra operations associated with matrix multiplications and diagonalizations become computational bottlenecks already for medium sized systems with only a few thousand basis functions. In this respect, it is important to note that density fitting introduces additional linear algebra steps, namely, for the solution of the inhomogeneous fitting equation systems, which contribute substantially to this bottleneck. In particular, the TED of the Coulomb matrix requires the diagonalization of a dense matrix that is one-order of magnitude larger than the corresponding Kohn-Sham matrix in the SCF. Although, this diagonalization has to be performed only once in the current standard deMon2k fitting algorithm, it represents for larger systems a notable bottleneck and renders calculations of auxiliary function set sizes above 100,000 impossible on many common available computer architectures. To circumvent this bottleneck partitioning schemes for the variational density fitting have been proposed in the literature[24,25,27,29,116]. However, such approaches compromise the inequality relation of Eq. (3.26) and, thus, the variational nature of the fit. In this work, we present a new iterative approach for the variational density fitting. It is based on MINRES[37,38], a Krylov subspace method that only requires matrix-vector multiplications. The matrix, here the Coulomb matrix $\mathbf{G}$, is calculated on-the-fly using standard recurrence relations and newly developed double asymptotic expansions for two-center electron repulsion integrals (ERIs). This greatly reduces the RAM demand of the algorithm and, therefore, permits density fitting with hundreds of thousands of auxiliary functions. Special attention is given to the preconditioning and stopping criteria of MINRES for the density fitting.

We will use MINRES to solve Eqs. (3.29) and (3.43), given by

$$\mathbf{G}\mathbf{x} = \mathbf{J}, \tag{4.1}$$

and

$$\mathbf{G}\mathbf{z} = \mathbf{L}, \tag{4.2}$$

where $\mathbf{G}$ is a dense and large coefficient matrix, called the Coulomb matrix, $\mathbf{J}$ is the right hand side Coulomb vector and $\mathbf{x}$ is the solution vector which contains the Coulomb fitting coefficients. Similar, in Eq. (4.2) $\mathbf{L}$ denotes the exchange-correlation vector that is obtained by

numerical integration and the $\mathbf{z}$ vector collects the exchange-correlation fitting coefficients. It is important to mention that this equation is solved only if the AUXIS option for the VXCTYPE keyword is triggered. In case of open-shell calculations, both equation systems have to be solved twice, one time for the α spin manifold and another time for the β spin manifold. Thus, we have Coulomb coefficient vectors $\mathbf{x}^{\alpha}$ and $\mathbf{x}^{\beta}$ and corresponding exchange-correlation vectors $\mathbf{z}^{\alpha}$ and $\mathbf{z}^{\beta}$ in deMon2k open-shell calculations. Although both equation systems are equally important for ADFT SCF calculations, we will focus here our discussion for clarity of presentation only on the Coulomb fitting equation system, Eq. (4.1). Because both equation systems have the same coefficients matrix $\mathbf{G}$ the here developed insight can be straightforward transfered to the solution of Eq. (4.2).

The methods for computing the solution $\mathbf{x}$ in Eq. (4.1) fall into two main categories: Direct methods and iterative methods. The direct methods consist in obtaining the inverse matrix $\mathbf{G}^{-1}$ by means of a diagonalization, LU decomposition or any other appropriate approach and calculate the solution as

$$\mathbf{x} = \mathbf{G}^{-1}\mathbf{J}. \tag{4.3}$$

Obviously, this approach assumes that $\mathbf{G}^{-1}$ exists within the underlying finite precision number model. As already discussed, this is by no means always the case for density fitting equation systems. Therefore, a proper conditioning of $\mathbf{G}$, e.g. by TED, is usually needed. The standard approach in deMon2k to solve Eq. (4.1) consists of an initial TED of the $\mathbf{G}$ matrix in order to ensure the existence of its inverse, $\mathbf{G}^{-1}$. This step involves the diagonalization of $\mathbf{G}$. The so conditioned $\mathbf{G}$ matrix is positive definite and properly conditioned for the used finite precision number model. Its inverse, either in full or diagonal form, is used as initial guess for the proposed quasi-Newton method by Domínguez-Soria *et al.*[36] previously developed in our research groups. During this iterative procedure an inverse BFGS update[117–120] of $\mathbf{G}^{-1}$ is performed. This ensures the positive definiteness of $\mathbf{G}$ and adapts $\mathbf{G}^{-1}$ to changed molecular structure parameters. This algorithm relies on the fact that the solution of the Eq. (4.1) is equivalent to the minimization of the quadratic function, $F(\mathbf{x})$, defined as

$$F(\mathbf{x}) = \frac{1}{2}\sum_{\bar{k},\bar{l}} G_{\bar{k}\bar{l}} x_{\bar{k}} x_{\bar{l}} - \sum_{\bar{k}} J_{\bar{k}} x_{\bar{k}}. \tag{4.4}$$

To find a minimum of $F(\mathbf{x})$ we expand it as

$$F(\mathbf{x}+\Delta\mathbf{x}) = \frac{1}{2}\sum_{\bar{k},\bar{l}} G_{\bar{k}\bar{l}}(x_{\bar{k}}+\Delta x_{\bar{k}})(x_{\bar{l}}+\Delta x_{\bar{l}}) - \sum_{\bar{k}} J_{\bar{k}}(x_{\bar{k}}+\Delta x_{\bar{k}}). \tag{4.5}$$

The variation of $F(\mathbf{x}+\Delta\mathbf{x})$ with respect to the step $\Delta\mathbf{x}$ yields

$$\frac{\partial F(\mathbf{x}+\Delta\mathbf{x})}{\partial \Delta x_{\bar{k}}} = \sum_{\bar{l}} G_{\bar{k}\bar{l}}(x_{\bar{l}}+\Delta x_{\bar{l}}) - J_{\bar{k}} \equiv 0 \quad \forall\ \bar{k}. \tag{4.6}$$

The corresponding second derivative,

$$\frac{\partial^2 F(\mathbf{x}+\Delta\mathbf{x})}{\partial \Delta x_{\bar{k}} \partial \Delta x_{\bar{l}}} = G_{\bar{k}\bar{l}}, \tag{4.7}$$

ensures the minimization of $F(\mathbf{x}+\Delta\mathbf{x})$ by Eq. (4.6) due to the positive definite nature of $\mathbf{G}$. For the step direction then follows

$$\sum_{\bar{l}} G_{\bar{k}\bar{l}}\Delta x_{\bar{l}} = -r_{\bar{k}} \Rightarrow \Delta x_{\bar{l}} = -\sum_{\bar{k}} G^{-1}_{\bar{l}\bar{k}} r_{\bar{k}}. \tag{4.8}$$

Here we introduce the residual vector, $\mathbf{r}$, with elements

$$r_{\bar{k}} = \sum_{\bar{l}} G_{\bar{k}\bar{l}} x_{\bar{l}} - J_{\bar{k}}. \tag{4.9}$$

The step formula, Eq. (4.8), shows that the residual takes the role of the gradient and the Coulomb matrix represents the Hessian in the minimization of $F(\Delta\mathbf{x})$. In quasi-Newton methods, the Hessian, here denoted by $\mathbf{G}$, is approximated by a symmetric positive definite matrix, which is updated at each iterative step[121]. The update to the inverse of the $\mathbf{G}$ matrix, denoted here by $\mathbf{G}^{-1}_{upd}$, is performed by vector-vector and matrix-vector multiplications and, therefore, possess a formal quadratic scaling. This update acts directly on the $\mathbf{G}^{-1}$ originally obtained from the TED as follows

$$\mathbf{G}^{-1}_{upd} = \mathbf{G}^{-1}_{old} + \left(1 + \frac{\boldsymbol{\gamma}^T \mathbf{G}^{-1}_{old}\boldsymbol{\gamma}}{\boldsymbol{\delta}^T\boldsymbol{\gamma}}\right)\frac{\boldsymbol{\delta}\boldsymbol{\delta}^T}{\boldsymbol{\delta}^T\boldsymbol{\delta}} - \frac{\boldsymbol{\delta}\boldsymbol{\gamma}^T\mathbf{G}^{-1}_{old} + \mathbf{G}^{-1}_{old}\boldsymbol{\gamma}\boldsymbol{\delta}^T}{\boldsymbol{\delta}^T\boldsymbol{\gamma}}, \tag{4.10}$$

with $\boldsymbol{\delta} = \mathbf{x} - \mathbf{x}^{old}$ and $\boldsymbol{\gamma} = \mathbf{r} - \mathbf{r}^{old}$. Here the superscript "old" denotes quantities from the

previous iteration cycles. With the approximated Hessian matrix at hand the next step direction is calculated by

$$\Delta\mathbf{x} = -\mathbf{G}_{upd}^{-1}\mathbf{r}. \tag{4.11}$$

After the step direction has been calculated a line search along this direction is performed. Because $F(\mathbf{x})$ is a convex quadratic function the line search parameter α can be computed analytically. Thus, we find as "new" fitting coefficient vector

$$x_{\bar{k}}^{new} = x_{\bar{k}} + \alpha\Delta x_{\bar{k}} \quad \text{with} \quad \min_{\alpha>0} F(\mathbf{x}+\Delta\mathbf{x}). \tag{4.12}$$

For the calculation of α we insert Eq. (4.12) into Eq. (4.4) and derive with respect to the α parameter

$$\frac{dF(\mathbf{x}+\alpha\Delta\mathbf{x})}{d\alpha} = \sum_{\bar{k},\bar{l}} G_{\bar{k}\bar{l}}\Delta x_{\bar{k}} x_{\bar{l}} + \alpha\sum_{\bar{k},\bar{l}} G_{\bar{k}\bar{l}}\Delta x_{\bar{k}}\Delta x_{\bar{l}} - \sum_{\bar{k}} J_{\bar{k}}\Delta x_{\bar{k}} \equiv 0. \tag{4.13}$$

Thus, it follows

$$\begin{aligned}\alpha\sum_{\bar{k},\bar{l}} G_{\bar{k}\bar{l}}\Delta x_{\bar{k}}\Delta x_{\bar{l}} &= \sum_{\bar{k}}\left(J_{\bar{k}} - \sum_{\bar{l}} G_{\bar{k}\bar{l}}x_{\bar{l}}\right)\Delta x_{\bar{k}} \\ &= -\sum_{\bar{k}} r_{\bar{k}}\Delta x_{\bar{k}},\end{aligned} \tag{4.14}$$

and further

$$\alpha = -\frac{\sum_{\bar{k}} r_{\bar{k}}\Delta x_{\bar{k}}}{\sum_{\bar{k},\bar{l}} G_{\bar{k}\bar{l}}\Delta x_{\bar{k}}\Delta x_{\bar{l}}}. \tag{4.15}$$

With the determination of the α parameter a new set of Coulomb fitting coefficients, $\mathbf{x}$, is generated according to Eq. (4.12). With these coefficients, a new residual is calculated and based on its norm, the iteration is stopped by a convergence criterium or continued with a new iteration step. For the sake of clarity and simplicity we will denote the updated inverse matrix, $\mathbf{G}_{upd}^{-1}$, as $\mathbf{G}^{-1}$ for the exchange-correlation coefficients calculation discussion. Once the quasi-Newton iterations for solving Eq. (4.1) are converged, the resulting updated $\mathbf{G}^{-1}$ matrix can be used as preconditioner for a conjugate gradient (CG) algorithm[35] to solve Eq. (4.2).

The CG algorithm is an iterative Krylov subspace method for a linear equation system with a symmetric, positive definite coefficients matrix, here $\mathbf{G}$. Similar to the just derived quasi-

Newton approach the CG algorithm relies on the fact that the solution of Eq. (4.2) is equivalent to the minimization of the quadratic function

$$F(\mathbf{z}) = \frac{1}{2}\mathbf{z}^T\mathbf{G}\mathbf{z} - \mathbf{L}^T\mathbf{z}. \tag{4.16}$$

This particular form of Krylov subspace methods is also called quadratic programming. To proceed, we now calculate the gradient of $F(\mathbf{z})$ which has the form

$$\nabla F(\mathbf{z}) = \mathbf{G}\mathbf{z} - \mathbf{L} = \mathbf{r}. \tag{4.17}$$

As in Eq. (4.9), we can identify this gradient with the residual of the corresponding equation system, here given by Eq. (4.2). The Hessian matrix calculated as

$$\nabla \otimes \nabla^T F(\mathbf{z}) = \mathbf{G}. \tag{4.18}$$

is again the symmetric, positive definite coefficient matrix. This guarantees the existence of a unique minimum for the quadratic function, which in turn translates to a unique bounded solution of the equation system. Therefore, this minimum can be searched by the minimization along conjugated gradient directions. To this end, a one-dimensional analytic line search along these directions is performed. Thus, the new exchange-correlation coefficients can be obtained as

$$z_{\bar{k}}^{new} = z_{\bar{k}} + \alpha\Delta z_{\bar{k}}. \tag{4.19}$$

The line search parameter α is obtained analogously as in Eq. (4.15) now given by

$$\alpha = -\frac{\sum\limits_{\bar{k}} r_{\bar{k}}\Delta z_{\bar{k}}}{\sum\limits_{\bar{k},\bar{l}} G_{\bar{k}\bar{l}}\Delta z_{\bar{k}}\Delta z_{\bar{l}}}, \tag{4.20}$$

Different to the calculation of the Coulomb fitting coefficients, the step direction, $\Delta\mathbf{z}$, is obtained by a CG rather than a quasi-Newton step as

$$\Delta z_{\bar{k}} = -s_{\bar{k}} + \beta\Delta z_{\bar{k}}^{old}, \tag{4.21}$$

where $s_{\bar{k}}$ is an element of the preconditioned (with the updated inverse $\mathbf{G}$ matrix) residual given

by

$$s_{\bar{k}} = \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}} r_{\bar{l}}. \tag{4.22}$$

In Eq. (4.22), $r_{\bar{l}}$ is an element of the residual of Eq. (4.2) calculated as

$$r_{\bar{l}} = \sum_{\bar{k}} G_{\bar{l}\bar{k}} z_{\bar{k}} - L_{\bar{l}}. \tag{4.23}$$

Finally, the conjugate gradient parameter β is calculated in terms of the residual vectors as

$$\beta = -\frac{\sum\limits_{\bar{k}} r_{\bar{k}} s_{\bar{k}}}{\sum\limits_{\bar{k}} r^{old}_{\bar{k}} s^{old}_{\bar{k}}}. \tag{4.24}$$

This value is set to zero at the first iteration step. Thus, the preconditioned equation system is solved solely in terms of the residuals of Eq. (4.2) and the updated inverse $\mathbf{G}$ matrix obtained originally from the TED and subsequently updated by the quasi-Newton iterations for the calculation of the Coulomb fitting coefficients. Therefore, no explicit matrix transformations are necessary. Because the TED of the $\mathbf{G}$ matrix is central to the current standard density fitting approach in deMon2k we analyze it in more detail in the next subsection.

4.1.2 Truncated eigenvalue decomposition

The truncated eigenvalue decomposition (TED) of the $\mathbf{G}$ matrix is based on

$$\mathbf{G} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T \tag{4.25}$$

where $\mathbf{Q}$ is an orthogonal matrix that collects the eigenvectors of $\mathbf{G}$, $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_{N_{aux}}$, and $\mathbf{\Lambda}$ is diagonal whose entries are the eigenvalues of $\mathbf{G}$, $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{N_{aux}}$. With these eigenvalues the conditioning of the G matrix can be analyzed by a quantity called the condition number, $\kappa(\mathbf{G})$, which is defined in numerical analysis as

$$\kappa(\mathbf{G}) = \frac{|\lambda_{\max}(\mathbf{G})|}{|\lambda_{\min}(\mathbf{G})|}, \tag{4.26}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ represent the largest and smallest absolute eigenvalues of $\mathbf{G}$, respectively. As Eq. (4.26) shows $\kappa(\mathbf{G})$ measures the range of the eigenvalue spectrum of the matrix $\mathbf{G}$. In

particular, small absolute eigenvalues yield large $\kappa(\mathbf{G})$. Such large condition numbers usually indicate that the matrix inversion is numerically unstable. Furthermore, the condition number is also used in sensitivity analysis as a measurement for the stability of an equation system solution. A linear equation system with a coefficient matrix that has a small condition number is said to be well-conditioned, whereas one with large condition number is called ill-conditioned. The condition number is a specific property of the problem. In the here studied systems, the $\mathbf{G}$ matrix is large, dense and usually ill-conditioned. To analyze the effect of TED on the density fitting, we assume that $\mathbf{G}$ is nonsingular. Then, it is easy to show that the solution of Eq. (4.3) can be expressed in terms of the eigenvectors, $\mathbf{q}_\ell$, of Eq. (4.25) as follows

$$\mathbf{x} = \sum_{\ell=1}^{N_{aux}} c_\ell \mathbf{q}_\ell, \quad c_\ell = \frac{\mathbf{q}_\ell^T \mathbf{J}}{\lambda_\ell}, \quad \ell = 1, \ldots, N_{aux}. \tag{4.27}$$

This expansion shows that potentially large coefficients c_ℓ, due to small eigenvalues, may contaminate the solution with numerical noise. In fact, $||\mathbf{x}||^2 = \sum_{\ell=1}^{N} c_\ell^2$ can be arbitrarily large. To avoid this contamination the summation in Eq. (4.27) might be started at some index $\ell = t + 1$ given by a threshold ϵ (default setting in deMon2k is 10^{-6}) on the magnitude of the eigenvalues of $\mathbf{G}$. This is the underlying idea of the TED currently used in deMon2k. As an example, Figure 4.1 depicts the eigenvalue decomposition of the $\mathbf{G}$ matrix for the C_{60} fullerene employing the PBE functional[77], DZVP basis set[122] and GEN-A2* auxiliary function set[99]. As this figure shows, choosing ϵ could be problematic from the physical point of view, more specifically the meaning of the inequality in Eq. (3.22) and, thus, the variational nature of the fitting, could be lost.

In particular, a relatively large value of ϵ could give a poor representation of $\mathcal{E}_2$, Eq. (3.22), whereas a too small value of ϵ could produce artificial fluctuations that violate the nonnegativity property of Eq. (3.22). In Figure 4.1, the black lines show the numerically stable behaviour of some eigenvalues whereas the red lines show the most problematic ones ($|c_\ell| > 50$) in terms of the coefficients, c_ℓ. Furthermore, from the fact that ϵ for our case was the deMon2k default value of 10^{-6} it can be seen from Figure 4.1 that the problematic eigenvalues are eliminated alongside with some reasonable well behaved ones. Therefore, the TED may appreciably change the chosen auxiliary function set.

Besides these algorithmic problems of TED, the necessary diagonalization of the $\mathbf{G}$ matrix introduces also a severe computational bottleneck due to the cubic scaling of this linear al-

gebra step. In fact, for large systems with auxiliary function dimensions above $100,000$ this step becomes impossible on many current hardware architectures and, therefore, renders such calculations impossible.

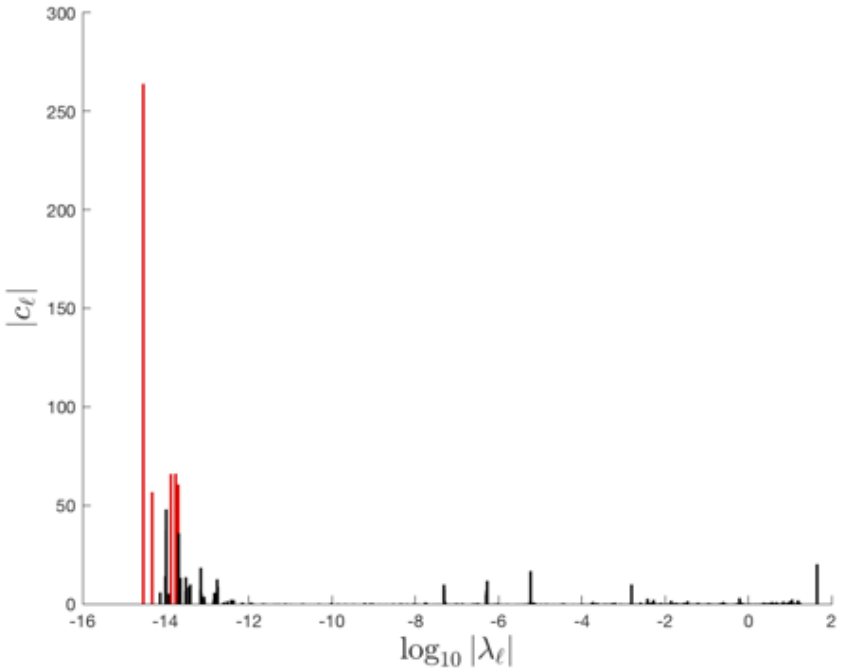

Figure 4.1: Eigenvalue decomposition of $\mathbf{G}$. Absolute values of c_ℓ coefficients in Eq. (4.27) as a function of $\log_{10}|\lambda_\ell|$ for the discussed C_{60} fullerene. The red lines show the highest fluctuations in the systems due to smallest eigenvalues

4.1.3 Krylov subspace methods

After we have identified the TED of the $\mathbf{G}$ matrix as the critical computational bottleneck for large scale variational density fitting, we now turn to alternative approaches. To this end, we will now explore Krylov methods[123] that use matrix-vector product subspaces for solving linear equation systems. Because the solution is found by a residual minimization in this subspace the conditioning of the coefficient matrix is less critical in these methods. The Krylov space is formed as

$$\mathcal{K}(\mathbf{G}, \mathbf{r}^{(0)}, k) = \text{span}\{\mathbf{r}^{(0)}, \mathbf{G}\mathbf{r}^{(0)}, \ldots, \mathbf{G}^{k-1}\mathbf{r}^{(0)}\}, \tag{4.28}$$

where k denotes the coefficient matrix power and $\mathbf{r}$ is the residual vector defined as

$$\mathbf{r} = \mathbf{G}\mathbf{x} - \mathbf{J}. \tag{4.29}$$

The Krylov method avoids matrix-matrix operations. Instead, it works with the resulting vectors of the linear subspace spanned by the images of the residual $\mathbf{r}$ under the powers of the $\mathbf{G}$ matrix. A serious problem of Krylov methods is linear dependency due to the power iteration in Eq. (4.28). To overcome this problem, methods that rely on Krylov subspaces frequently involve an orthogonalization process such as Lanczos[124] (for symmetric cases) or Arnoldi iterations[125] (for nonsymetric cases). Common Krylov subspace methods are Conjugate Gradient[126] (CG), GMRES[127] (generalized minimal residual), BiCGSTAB[128] (biconjugate gradient stabilized), MINRES[37,38] and EN[129] (Eirola-Nevanlinna algorithm), to name a few. Since these methods rely on the Krylov basis, it is evident that the method converges at most in n iterations where n is the equation system size. However, in the presence of rounding errors this statement does not hold. Moreover, in practice n can be very large, and the iterative process reaches sufficient accuracy already far earlier. In general, the convergence of Krylov methods is straightforward dependent on the condition number of the associated coefficient matrix. Thus, a large condition number can result in very slow convergence even with a good preconditioner. Nevertheless, Krylov methods have been successfully applied to many different fields. However, for density fitting they were so far not used.

4.2 Density fitting with the MINRES algorithm

Although, the Coulomb matrix, $\mathbf{G}$, is formally positive definite it becomes in practical applications indefinite due to finite precision round-off errors as already discussed. Therefore, we choose here the MINRES[130] algorithm which is a Krylov subspace based method designed for solving symmetric indefinite equation systems. As a result, the positive definiteness of the Coulomb matrix is no longer required which eliminates the need for the initial TED of $\mathbf{G}$. MINRES is based on the minimization of the residual norm at each iteration over the affine space $\mathbf{x}^{(0)} + \mathcal{K}(\mathbf{G}, \mathbf{r}^{(0)}, k)$

$$\min_{\mathbf{x} \in \mathbf{x}^{(0)} + \mathcal{K}(\mathbf{G}, \mathbf{r}^{(0)}, k)} || \mathbf{G}\mathbf{x} - \mathbf{J} ||, \tag{4.30}$$

where $\mathbf{x}^{(0)}$ is an initial approximation, $\mathbf{r}^{(0)} = \mathbf{G}\mathbf{x}^{(0)} - \mathbf{J}$ is the corresponding residual; k stands for the iteration counter and $||\cdot||$ represents the Euclidean norm. Finally, $\mathcal{K}(\mathbf{G}, \mathbf{r}^{(0)}, k)$, the Krylov subspace of dimension k associated with $\mathbf{G}$, $\mathbf{J}$ and $\mathbf{x}^{(0)}$, is formally defined according to Eq. (4.28) as

$$\mathcal{K}(\mathbf{G}, \mathbf{r}^{(0)}, k) = \text{span}\{\mathbf{r}^{(0)}, \mathbf{G}\mathbf{r}^{(0)}, \ldots, \mathbf{G}^{k-1}\mathbf{r}^{(0)}\}.$$

The Lanczos process[124] provides a more practical representation of $\mathcal{K}(\mathbf{G}, \mathbf{r}^{(0)}, k)$ in terms of an orthonormal basis. At MINRES iteration k we have

$$\mathbf{G}\mathbf{V}^{(k)} = \mathbf{V}^{(k+1)}\mathbf{T}^{(k)}, \tag{4.31}$$

where $\mathbf{T}^{(k)}$ is a $(k+1) \times k$ tridiagonal matrix defined by the α and β values obtained from the Lanczos procedure shown in Scheme 4.1.

$$\mathbf{T}^{(k)} = \begin{pmatrix} \alpha_1 & \beta_2 & 0 & 0 & \cdots & 0 \\ \beta_2 & \alpha_2 & \beta_3 & 0 & \cdots & 0 \\ 0 & \beta_3 & \alpha_3 & \beta_4 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & \beta_{k-1} & \alpha_{k-1} & \beta_k \\ 0 & & & & \beta_k & \alpha_k \\ \hline 0 & & \cdots & & 0 & \beta_{k+1} \end{pmatrix} \tag{4.32}$$

Scheme 4.1: Algorithm to generate a new Lanczos vector.

Lanczos step

Initialize: $\mathbf{v}^{(0)} = \mathbf{0}, \mathbf{v}^{(1)} = \mathbf{J}/||\mathbf{J}||, \beta_1 = ||\mathbf{J}||$

For $k = 1, 2, \cdots$ **do**

$\alpha_k = \mathbf{v}^{(k)^T}\mathbf{G}\mathbf{v}^{(k)}$

$\mathbf{v}^{(k+1)} \leftarrow \mathbf{G}\mathbf{v}^{(k)} - \alpha_k\mathbf{v}^{(k)} - \beta_k\mathbf{v}^{(k-1)}$

$\beta_{k+1} = ||\mathbf{v}^{(k+1)}||$

if $\beta_{k+1} = 0$ **Exit**

$\mathbf{v}^{(k+1)} \leftarrow \mathbf{v}^{(k+1)}/\beta_{k+1}$

End For

Superscripts in parentheses denote MINRES iterations. For the calculation of the α_k and β_k elements in $\mathbf{T}^{(k)}$ we assume, according to the Lanczos process[124], that the Coulomb matrix $\mathbf{G}$ can be brought into tridiagonal form by the following orthogonal transformation

$$\mathbf{T} = \mathbf{V}^T \mathbf{G} \mathbf{V}. \tag{4.33}$$

Because $\mathbf{V}$ is an orthogonal transformation matrix it provides an orthogonal basis, i.e. its column vectors are orthonormal to each other and span a corresponding vector space. To obtain the α_k and β_k elements of $\mathbf{T}$ we solve $\mathbf{GV} = \mathbf{VT}$. Taking the k^{th} column, which is equivalent to the k^{th} dimension of the Krylov subspace and, thus, the k^{th} MINRES iteration, of $\mathbf{GV}$ yields

$$[\mathbf{GV}]_{ik} = \sum_j G_{ij} V_{jk} = \sum_j V_{ij} T_{jk} \qquad \forall\, i. \tag{4.34}$$

By applying the tridiagonal form of $\mathbf{T}$ follows

$$[\mathbf{GV}]_{ik} = V_{i(k-1)} T_{(k-1)k} + V_{ik} T_{kk} + V_{i(k+1)} T_{(k+1)k} \qquad \forall\, i. \tag{4.35}$$

Identifying the $\mathbf{T}$ elements in Eq. (4.35) with the α and β entries in $\mathbf{T}^{(k)}$, i.e. setting $T_{(k-1)k} = \beta_k$, $T_{kk} = \alpha_k$ and $T_{(k+1)k} = \beta_{k+1}$ we obtain

$$[\mathbf{GV}]_{ik} = V_{i(k-1)} \beta_k + V_{ik} \alpha_k + V_{i(k+1)} \beta_{k+1} \qquad \forall\, i. \tag{4.36}$$

During the MINRES iterations this equation can be expressed by the following vector equation

$$\mathbf{G}\mathbf{v}^{(k)} = \beta_k \mathbf{v}^{(k-1)} + \alpha_k \mathbf{v}^{(k)} + \beta_{k+1} \mathbf{v}^{(k+1)}. \tag{4.37}$$

The here appearing $\mathbf{v}^{(k)}$ vectors are the orthonormal column vectors of the $\mathbf{V}$ matrix in Eq. (4.33) that are successively collected during the MINRES iterations. Reformulation of Eq. (4.37) yields as recurrence relation for the successive calculation of the $\mathbf{v}^{(k)}$ vectors as

$$\beta_{k+1} \mathbf{v}^{(k+1)} = (\mathbf{G} - \alpha_k \mathbf{E}) \mathbf{v}^{(k)} - \beta_k \mathbf{v}^{(k-1)}. \tag{4.38}$$

To determine α_k we multiply Eq. (4.37) from the left with $\mathbf{v}^{(k)^T}$ and explore the orthonormality between the $\mathbf{v}^{(k)}$ vectors as

$$\mathbf{v}^{(k)^T}\mathbf{G}\mathbf{v}^{(k)} = \beta_k \underbrace{\mathbf{v}^{(k)^T}\mathbf{v}^{(k-1)}}_{\to 0} + \alpha_k \underbrace{\mathbf{v}^{(k)^T}\mathbf{v}^{(k)}}_{\to 1} + \beta_{k+1} \underbrace{\mathbf{v}^{(k)^T}\mathbf{v}^{(k+1)}}_{\to 0} = \alpha_k. \tag{4.39}$$

The iterative construction of the Krylov subspace starts with setting $\mathbf{x}^{(0)} = \mathbf{0}$, $\beta_1 = ||\mathbf{r}^{(0)}|| = ||\mathbf{J}||$, $\mathbf{v}^{(0)} = \mathbf{0}$ and $\beta_1\mathbf{v}^{(1)} = \mathbf{r}^{(0)} = \mathbf{J}$. The associated LanczosStep to each MINRES iteration k yields α_k and β_{k+1} according to Eqs. (4.39) and (4.38). Hence α_k gives the component of $\mathbf{G}\mathbf{v}^{(k)}$ along $\mathbf{v}^{(k)}$ and β_{k+1} normalizes $\mathbf{v}^{(k+1)}$. Thus, we have $\mathbf{T}^{(k)}$, as given in Eq. (4.32), and $\mathbf{V}^{(k)} = (\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, ..., \mathbf{v}^{(k)})$ at hand. The orthonormal column vector in $\mathbf{V}^{(k)}$ span $\mathcal{K}(\mathbf{G}, \mathbf{r}^{(0)}, k)$, i.e.

$$\text{span}\{\mathbf{r}^{(0)}, \mathbf{G}\mathbf{r}^{(0)}, \ldots, \mathbf{G}^{k-1}\mathbf{r}^{(0)}\} = \text{span}\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)} \ldots, \mathbf{v}^{(k)}\}. \tag{4.40}$$

Note that in exact arithmetic the Krylov subspace construction finishes when $\beta_{k+1} = 0$. A very detailed algorithmic description of MINRES[37] and LanczosStep are given in appendix A. For further explanation and example, of the LanczosStep see appendix B.

The new approximation for $\mathbf{x}$ is written as

$$\mathbf{x} = \mathbf{x}^{(0)} + \mathbf{V}^{(k)}\mathbf{y}; \tag{4.41}$$

substituting the above expression in Eq. (4.30) and considering $\mathbf{x}^{(0)} = \mathbf{0}$ gives the condition to obtain $\mathbf{y}$, which consists in minimizing

$$\min_{\mathbf{y}} ||\mathbf{G}\mathbf{x}^{(0)} + \mathbf{G}\mathbf{V}^{(k)}\mathbf{y} - \mathbf{J}|| = \min_{\mathbf{y}} ||\mathbf{G}\mathbf{V}^{(k)}\mathbf{y} - \mathbf{J}|| \tag{4.42}$$

over all $\mathbf{y} \in \mathbb{R}^k$. To proceed we now substitute $\mathbf{G}\mathbf{V}^{(k)}$ according to Eq. (4.31) by $\mathbf{V}^{(k+1)}\mathbf{T}^{(k)}$ and expand $\mathbf{J}$ in terms of the first orthogonal vector $\mathbf{v}^{(1)}$ in $\mathbf{V}^{(k+1)}$ as

$$\mathbf{J} = \beta_1\mathbf{v}^{(1)}. \tag{4.43}$$

This yields

$$\min_{\mathbf{y}} ||\mathbf{V}^{(k+1)}\mathbf{T}^{(k)}\mathbf{y} - \beta_1\mathbf{v}^{(1)}|| = \min_{\mathbf{y}} ||\mathbf{V}^{(k+1)}(\mathbf{T}^{(k)}\mathbf{y} - \beta_1\mathbf{e}_1)||. \tag{4.44}$$

Thus, the final form of the least-square minimization problem in each MINRES iteration is given by

$$\min_{\mathbf{y}} ||\mathbf{T}^{(k)}\mathbf{y} - \beta_1 \mathbf{e}_1||, \tag{4.45}$$

with $\mathbf{e}_1$ being the first canonical vector in $\mathbb{R}^{k+1}$, which is $\mathbf{e}_1 = (1, 0, 0, ..., 0)^T$. This is an over-constrained system of $k+1$ equations for k unknowns. Because of the tridiagonal structure of the matrix $\mathbf{T}^{(k)}$, the minimization problem of Eq. (4.45) is equivalent to the solution of the inhomogeneous equation system (see also proposition 6.9 in [131])

$$\mathbf{R}^{(k)}\mathbf{y} = \beta_1 \mathbf{Q}^{(k)^T}\mathbf{e}_1, \tag{4.46}$$

where $\mathbf{Q}^{(k)}$ and $\mathbf{R}^{(k)}$ are the $(k+1)\times(k+1)$ orthogonal and $(k+1)\times k$ upper triangular matrix from the QR decomposition of $\mathbf{T}^{(k)}$, respectively (see appendix C for an illustrative introduction to QR decomposition).

To proceed we express $\mathbf{Q}^{(k)^T}$ as a successive product of Givens rotation[124] matrices

$$\mathbf{Q}^{(k)^T} = \mathbf{\Omega}^{(1)}\mathbf{\Omega}^{(2)}\mathbf{\Omega}^{(3)}\cdots\mathbf{\Omega}^{(k)}, \tag{4.47}$$

where $\mathbf{\Omega}^{(k)}$ is given by

$$\mathbf{\Omega}^{(k)} = \begin{bmatrix} \mathbf{E}^{(k-1)} & 0 & 0 \\ 0 & c_k & s_k \\ 0 & -s_k & c_k \end{bmatrix}. \tag{4.48}$$

With this structure of $\mathbf{Q}^{(k)}$ we find for the QR decomposition of $\mathbf{T}^{(k)}$

$$\mathbf{T}^{(k)} = \mathbf{Q}^{(k)}\mathbf{R}^{(k)} \Rightarrow \mathbf{R}^{(k)} = \mathbf{Q}^{(k)^T}\mathbf{T}^{(k)}. \tag{4.49}$$

Enforcing $\mathbf{R}^{(k)}$ to be upper triangular yields

$$\begin{aligned}
\mathbf{R}^{(k)} = \begin{pmatrix} \mathbf{R} \\ \mathbf{0} \end{pmatrix} &= \mathbf{Q}^{(k)^T} \left(\begin{array}{ccc} \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{T}^{(k-1)}$}} & 0 \\ & & \beta_k \\ \hline 0 & \beta_k & \alpha_k \\ 0 & 0 & \beta_{k+1} \end{array}\right) \\
&= \mathbf{\Omega}^{(1)}\mathbf{\Omega}^{(2)}\mathbf{\Omega}^{(3)}\cdots\mathbf{\Omega}^{(k)} \left(\begin{array}{ccc} \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{T}^{(k-1)}$}} & 0 \\ & & \beta_k \\ \hline 0 & \beta_k & \alpha_k \\ 0 & 0 & \beta_{k+1} \end{array}\right) \\
&= \mathbf{\Omega}^{(1)}\mathbf{\Omega}^{(2)}\mathbf{\Omega}^{(3)}\cdots \begin{pmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{E}^{(k-1)}$}} & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & c_k & s_k \\ 0 & 0 & -s_k & c_k \end{pmatrix} \left(\begin{array}{ccc} \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{T}^{(k-1)}$}} & 0 \\ & & \beta_k \\ \hline 0 & \beta_k & \alpha_k \\ 0 & 0 & \beta_{k+1} \end{array}\right).
\end{aligned} \tag{4.50}$$

Therefore, we obtain as condition for the annihilation of the β_{k+1} entry in $\mathbf{T}^{(k)}$

$$-s_k\alpha_k + c_k\beta_{k+1} = 0. \tag{4.51}$$

Together with the corresponding normalization,

$$s_k^2 + c_k^2 = 1, \tag{4.52}$$

follows

$$s_k = \frac{\beta_{k+1}}{\sqrt{\alpha_k^2 + \beta_{k+1}^2}}, \qquad \wedge \qquad c_k = \frac{\alpha_k}{\sqrt{\alpha_k^2 + \beta_{k+1}^2}}. \tag{4.53}$$

In this way, the $k{+}1$ β entries below the diagonal of $\mathbf{T}^{(k)}$ are annihilated by the k rotations with the $\mathbf{\Omega}$ matrices forming $\mathbf{Q}^{(k)^T}$. Note that the $\mathbf{\Omega}$ matrix dimensions formally change during the MINRES iterations according to the dimension changes in $\mathbf{T}^{(k)}$. However, in the practical implementation this is of no concern because only the s_k and c_k are calculated in each MINRES step. For a mathematical consistent QR decomposition with successively applied rotations see appendix D.

Additional MINRES stopping criteria in deMon2k

The intrinsic stopping criteria of MINRES is given in the LanczosStep when $\beta_{k+1} = 0$. It is important to mention that if β_{k+1} becomes close to zero, there is no need to perform another Givens rotation. Hence, the associated angles from two successive rotations will be below a certain tolerance, TOL. This is another indicator that the MINRES algorithm has converged.

In deMon2k we also permit MINRES convergence directly through the residual. To this end, it is important to note that solving Eq. (4.1) in the context of the SCF methodology amounts for solving a sequence of equation systems

$$\mathbf{G}\mathbf{x} = \mathbf{J}^{(i)}, \quad i = 1, \ldots, I \tag{4.54}$$

in which the right hand sides $\mathbf{J}^{(i)}$ change slowly between consecutive SCF iterations. MINRES makes use of the null vector as initial approximation; thus, we reformulate Eq. (4.54) as the equivalent sequence

$$\mathbf{x}^{(0)} = \mathbf{0}, \quad \mathbf{G}\mathbf{d} = \mathbf{J}^{(i)} - \mathbf{G}\mathbf{x}^{(i-1)}, \quad \mathbf{x}^{(i)} = \mathbf{x}^{(i-1)} + \mathbf{d}, \quad i = 1, \ldots, I. \tag{4.55}$$

Each system is approximately solved subject to satisfy at least one of the following conditions in terms of the residual norm approximation

$$||\mathbf{r}^{(k)}|| \leq ||\mathbf{G}||\,||\mathbf{d}^{(k)}||\,\mathrm{TOL}, \qquad ||\mathbf{r}^{(k)}|| \leq \frac{1}{\sqrt{N}}\,\mathrm{TOL}, \tag{4.56}$$

where $\mathrm{TOL} = 1.0 \times 10^{-7}$ is the default value. In case of an open-shell calculation two residual norms are calculated, $||\mathbf{r}_{\alpha}^{(\mathbf{k})}||$ and $||\mathbf{r}_{\beta}^{(\mathbf{k})}||$, and their sum must fulfill the same stopping criteria of Eq. (4.56). It is important to remark that the TOL value for MINRES is fixed through the whole SCF procedure. With these settings, well converged SCF solutions can be obtained over a wide range of the potential energy surface.

4.3 Preconditioning of the MINRES algorithm

In the context of density fitting with large auxiliary function sets, $\mathbf{G}$ is usually ill-conditioned (typical condition numbers range from 10^9 to 10^{16}). This typically results in slow convergence of the MINRES algorithm because the number of iterations taken by MINRES to achieve con-

vergence is closely related to the number of distinct eigenvalues of the coefficient matrix. Although in theory, Krylov methods are guaranteed to reach convergence, their convergence rate is often so slow that stagnation is observed. To overcome this problem Krylov subspace methods make use of a preconditioner also referred to as *accelerators*. Because the convergence of Krylov methods is directly proportional to the number of distinct eigenvalues of the coefficient matrix, here the Coulomb matrix $\mathbf{G}$, preconditioning clusters these eigenvalues. Thus, the perfect preconditioner is $\mathbf{G}^{-1}$ which clusters all eigenvalues of $\mathbf{G}$ to 1. Preconditioning a Krylov subspace method transforms the original problem of Eq. (4.1) into a presumably better conditioned one. It is worth to mention that the development of preconditioners is a large research area[132–135] and a variety of methods have been proposed such as relaxation-type methods, incomplete LU decompositions and approximate inverse methods. Moreover, the preconditioning of Krylov methods is very specific to the problem. Therefore, no general type of preconditioning exists. Most popular preconditioners are nonsingular matrices $\mathbf{M}$ that approximate $\mathbf{G}$ reasonably well; the rationale being that the product $\mathbf{M^{-1}G}$ will have a more favourable eigenvalue distribution than $\mathbf{G}$. Then Eq. (4.1) is replaced by the equivalent system

$$\mathbf{M}^{-1}\mathbf{G}\mathbf{x} = \mathbf{M}^{-1}\mathbf{J}. \tag{4.57}$$

Obviously, the perfect preconditioner is $\mathbf{M} = \mathbf{G}$ since the eigenvalues of $\mathbf{M^{-1}G}$ are all equal to 1, and any iterative method applied to $\mathbf{M^{-1}G}$ will converge in one iteration.

Different to the TED approach in deMon2k the (approximate) inversion of the $\mathbf{G}$ matrix is not an option for the MINRES preconditioner due to the associated CPU time and RAM demand. Therefore, a more pragmatic approach for the preconditioning of MINRES is used. It builds on the atomic block structure of the $\mathbf{G}$ matrix. First, a block diagonal matrix $\mathbf{B}$ is formed from the atomic blocks of $\mathbf{G}$. Second, since $\mathbf{B}$ may contain indefinite blocks due to the linear dependency in large auxiliary function sets, it is replaced by its associated modified Cholesky decomposition[136]

$$\mathbf{M} = \mathbf{P}^T(\mathbf{B} + \Delta\mathbf{B})\mathbf{P} = \mathbf{L}\mathbf{D}\mathbf{L}^T. \tag{4.58}$$

In Eq. (4.58) $\mathbf{P}$ is a permutation matrix, $\mathbf{L}$ is unit lower triangular and $\mathbf{D}$ is block diagonal with diagonal blocks of dimension 1 or 2. The matrix $\Delta\mathbf{B}$ is a small perturbation that makes $\mathbf{M}$ sufficiently positive definite[137–142] and reasonably well conditioned. Therefore, if atomic blocks of $\mathbf{B}$ become indefinite, $||\Delta\mathbf{B}||$ must not be larger than

$$\min\{||\Delta\mathbf{B}|| : \mathbf{B} + \Delta\mathbf{B} \text{ is positive definite.}\} \tag{4.59}$$

The algorithm computes (not explicitly) the $\Delta\mathbf{B}$ matrix to ensure that if $\lambda_{min}(\mathbf{B}) < \delta$ then $\lambda_{min}(\mathbf{B} + \Delta\mathbf{B}) \approx \delta$, where δ is a real and small positive value.

In general the algorithm works as follows. First a symmetric indefinite decomposition is performed using the bounded Bunch-Kaufman (BBK) pivoting for the block $\mathbf{B}$ as

$$\mathbf{P}^T\mathbf{B}\mathbf{P} = \mathbf{L}\tilde{\mathbf{D}}\mathbf{L}^T, \tag{4.60}$$

where the right hand side of Eq. (4.60) has the following form

$$\mathbf{L}\tilde{\mathbf{D}}\mathbf{L}^T = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ L_{21} & 1 & 0 & \cdots \\ L_{31} & L_{32} & 1 & \ddots \end{pmatrix} \begin{pmatrix} \tilde{D}_1 & 0 & 0 & \cdots \\ 0 & \tilde{D}_2 & 0 & \cdots \\ 0 & 0 & \tilde{D}_3 & \ddots \end{pmatrix} \begin{pmatrix} 1 & L_{21} & L_{31} & \cdots \\ 0 & 1 & L_{32} & \cdots \\ 0 & 0 & 1 & \ddots \end{pmatrix}. \tag{4.61}$$

The following recursive relations apply for $\tilde{\mathbf{D}}$ and $\mathbf{L}$

$$\tilde{D}_i = B_{ii} - \sum_{j=1}^{i-1} L_{ij}^2 \tilde{D}_j, \tag{4.62}$$

$$L_{ij} = \frac{1}{\tilde{D}_j}\left(B_{ij} - \sum_{k=1}^{j-1} L_{ik}L_{jk}\tilde{D}_k\right) \forall\ i > j. \tag{4.63}$$

When a pivot, $\tilde{D}_i$ in Eq. (4.62) is small ($\sim 10^{-9}$), the BBK pivoting strategy is used in order to overcome this instability and continue with the symmetric indefinite decomposition. Inside deMon2k, this procedure is carried out by the DSYTRF_RK and DSYTRS_3 LAPACK[143] library routines. Next, the algorithm computes $\mathbf{D} = \tilde{\mathbf{D}} + \Delta\tilde{\mathbf{D}}$ block by block, where $\tilde{\mathbf{D}}$ corresponds to the block diagonal matrix of the decomposition in Eq. (4.60), and $\Delta\tilde{\mathbf{D}}$ is a small perturbation of the indefinite block diagonal matrix, $\tilde{\mathbf{D}}$, obtained as the minimum Frobenius norm such that $\lambda_{min}(\tilde{\mathbf{D}} + \Delta\tilde{\mathbf{D}}) \geq \delta$ using the Higham theorem[136].

Scheme 4.2: Modified Cholesky decomposition

Compute the symmetric indefinite decomposition $\mathbf{P}^T\mathbf{B}\mathbf{P} = \mathbf{L}\mathbf{D}\mathbf{L}^T$ using the BBK pivoting strategy

For each block $\tilde{D}$ of **D do**

  if $\tilde{D}$ is a 1×1 block, $\tilde{d}$, **then**

    if $\tilde{d} < \delta$ **then** $\tilde{d} = \delta$

  else if $\tilde{D}$ is a 2×2 block, **then** compute its eigenvalue decomposition $\mathbf{\Lambda} = \mathbf{U}^T\tilde{\mathbf{D}}\mathbf{U}$,

    where $\Lambda = \text{diag}(\lambda_1, \lambda_2)$

    if $\lambda_1 < \delta$ **then** $\lambda_1 = \delta$

    if $\lambda_2 < \delta$ **then** $\lambda_2 = \delta$

    Build the new block as $\tilde{\mathbf{D}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$

  end if

end for

As Scheme 4.2 shows the perturbation matrix $\Delta\mathbf{B}$ is not explicitly calculated. Instead, the diagonal blocks of $\mathbf{D}$ with dimension 1 or 2 from the indefinite decomposition of $\mathbf{B}$ are directly modified according to Scheme 4.2. In our deMon2k implementation the threshold δ is set to 10^{-9}. Here, the updated $\tilde{\mathbf{D}}$ blocks are the key to success to fulfill the condition in Eq. (4.59). It is worth to mention, that the modified Cholesky decomposition has been recently applied to second order electron propagator calculations employing density fitting[144] and the variational fitting of Fock exchange[145]. Therefore, MINRES is applied to the preconditioned system in Eq. (4.57), and the matrix-vector product $\mathbf{w} = \mathbf{M}^{-1}\mathbf{G}\mathbf{v}^{(k)}$ is computed in two steps

$$\mathbf{v} = \mathbf{G}\mathbf{v}^{(k)}, \qquad \mathbf{L}\mathbf{D}\mathbf{L}^T\mathbf{w} = \mathbf{v}. \tag{4.64}$$

Thus, the product $\mathbf{M}^{-1}\mathbf{G}$ is never computed or stored; the second step in Eq. (4.64) involves only forward and backward substitutions. The forward substitution for an atomic block of dimension m is carried out as follows

$$\mathbf{L}\mathbf{D}\mathbf{y} = \mathbf{v}. \tag{4.65}$$

Eq. (4.65) is given in element-wise form as

$$\begin{array}{ccccccccc} \ell'_{11}y_1 & & & & & & & = & v_1 \\ \ell'_{21}y_1 & + & \ell'_{22}y_2 & & & & & = & v_2 \\ \vdots & & \vdots & & \ddots & & & & \vdots \\ \ell'_{m1}y_1 & + & \ell'_{m2}y_2 & + & \cdots & + & \ell'_{mm}y_m & = & v_m \end{array} \tag{4.66}$$

where the matrix elements of $\mathbf{L}'$ are the $\mathbf{L}$ matrix elements scaled (on-the-fly) by the diagonal elements of $\mathbf{D}$.
Thus, the auxiliary vector $\mathbf{y}$ is straightforward obtained using the forward substitution formulas

$$\begin{aligned} y_1 &= \frac{v_1}{\ell'_{11}}, \\ y_2 &= \frac{v_2 - \ell'_{21}x_1}{\ell'_{22}}, \\ &\vdots \\ y_m &= \frac{v_m - \sum_{i=1}^{m-1} \ell'_{mi}x_i}{\ell'_{mm}}. \end{aligned} \tag{4.67}$$

Because $\mathbf{L}^T$ is upper triangular (transpose of $\mathbf{L}$), $\mathbf{w}$ can be obtained by the backward substitution of

$$\mathbf{L}^T\mathbf{w} = \mathbf{y}, \tag{4.68}$$

in the following form

$$\begin{aligned} w_m &= \frac{y_m}{\ell_{mm}} \\ w_{m-1} &= \frac{y_{m-1} - \ell_{(m-1)m}y_m}{\ell_{(m-1)(m-1)}} \\ &\vdots \\ w_1 &= \frac{y_1 - \sum_{j=2}^{m} \ell_{1j}y_j}{\ell_{11}}. \end{aligned} \tag{4.69}$$

In order to demonstrate the effect of preconditioning on the eigenvalue clustering we present in Figure 4.2 the eigenvalue distribution of $\mathbf{G}$ and $\mathbf{M^{-1}G}$ for C_{60} employing the PBE/DZVP/GEN-A2* methodology. In this example $\mathbf{G}$ has dimension 5580×5580 and its eigenvalue distribution is shown in the upper part of Figure 4.2. By preconditioning with an atomic block diagonal inverse, $\mathbf{M}^{-1}$, calculated by the modified Cholesky decomposition, the lower eigenvalue spectrum in Figure 4.2 is obtained. An obvious discretization of the eigenvalue spectrum above 10 is observed. Thus, the number of distinct eigenvalues is reduced, which in turn will accelerate the convergence of MINRES.

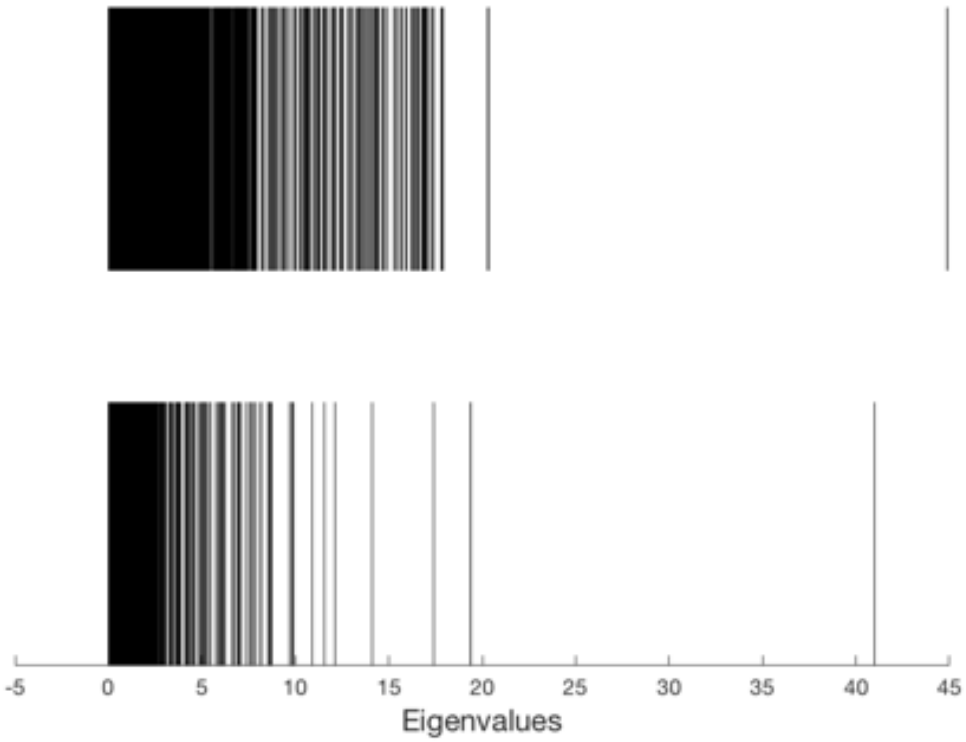

Figure 4.2: Eigenvalues of $\mathbf{G}$ (top) and $\mathbf{M^{-1}G}$ (bottom) for the C_{60} fullerene.

In order to illustrate the effect of the preconditioner on the MINRES performance, Figure 4.3 shows the number of MINRES iterations for the initial SCF cycle in deMon2k without and with preconditioning for the C_{60} fullerene. In this example $\mathbf{G}$ has a condition number of 4.25×10^{16} and 14 negative eigenvalues; $\mathbf{J}$ corresponds to the first right hand side in the SCF cycle, i.e. the Coulomb vector built from the tight-binding start density[146]. The initial approximation was $\mathbf{x}^{(0)} = \mathbf{0}$. The iteration was stopped when $||\mathbf{r}^{(k)}|| \leq \mathrm{TOL}\,||\mathbf{J}||$ with TOL = 10^{-7}. From Figure 4.3 it is straightforward to see how MINRES converges through a plateau (pink dashed line) when the preconditioner is not used. On the other hand, preconditioned MINRES (pMINRES) converges faster because the plateau extension is smaller and reaches the stopping criterium after around 200 iterations.

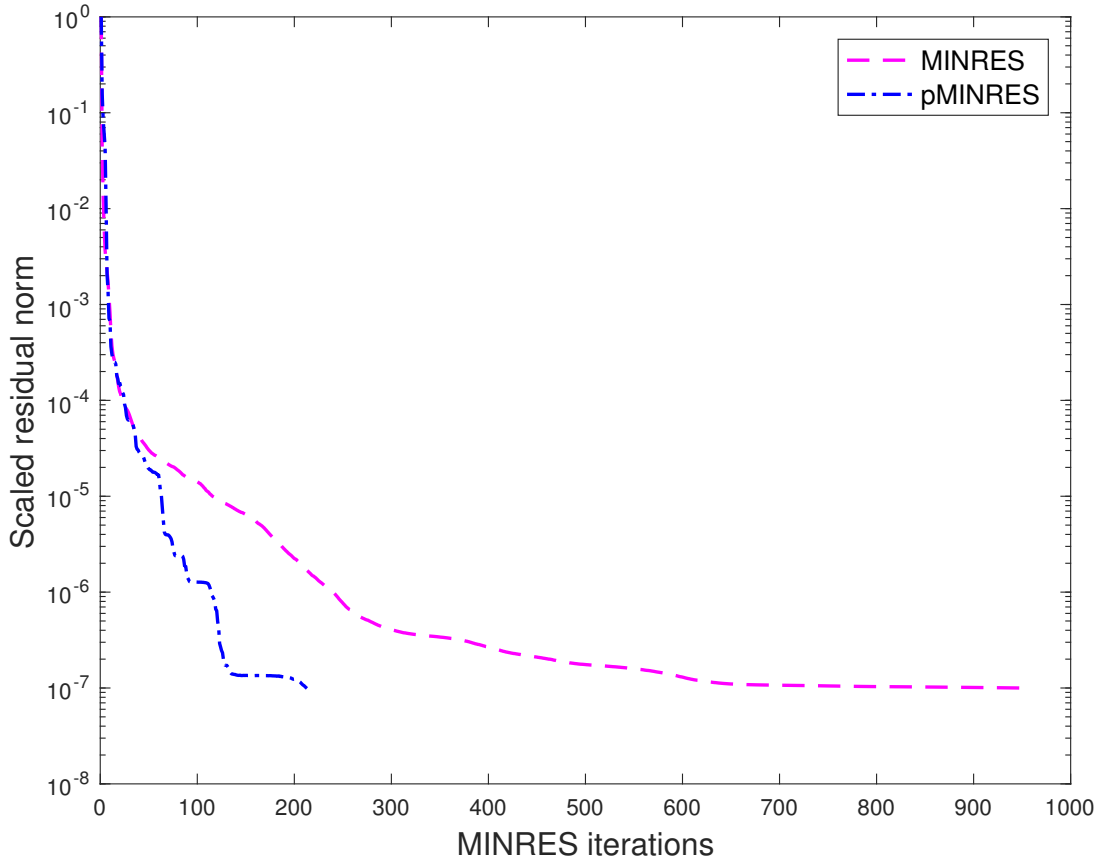

Figure 4.3: Norm of the scaled residual $||\mathbf{r}^{(k)}||/||\mathbf{J}||$ for the C_{60} fullerene as a function of the number of MINRES iterations. See text for further details.

Although the modified Cholesky (ModChol) decomposition preconditioner served well for most of our benchmark calculations, we found some systems where MINRES significantly altered the SCF convergence. Therefore, we developed a second type of preconditioning based on the TED of the atomic blocks of **G**. The underlying idea is that TED eliminates numerical instabilities of ill-conditioned matrices. Thus, a TED based inversion of atomic blocks of **G** should be particularly beneficial if these blocks are ill-conditioned. To validate this TED based preconditioner, we took the same system from the example of Figure 4.3 and we tested the TED preconditioner. In Figure 4.4, the norm of the scaled residuals without using any preconditioner (dashed pink line) compared with both preconditioners (blue line for ModChol and red line for TED) are depicted versus the number of MINRES iterations.

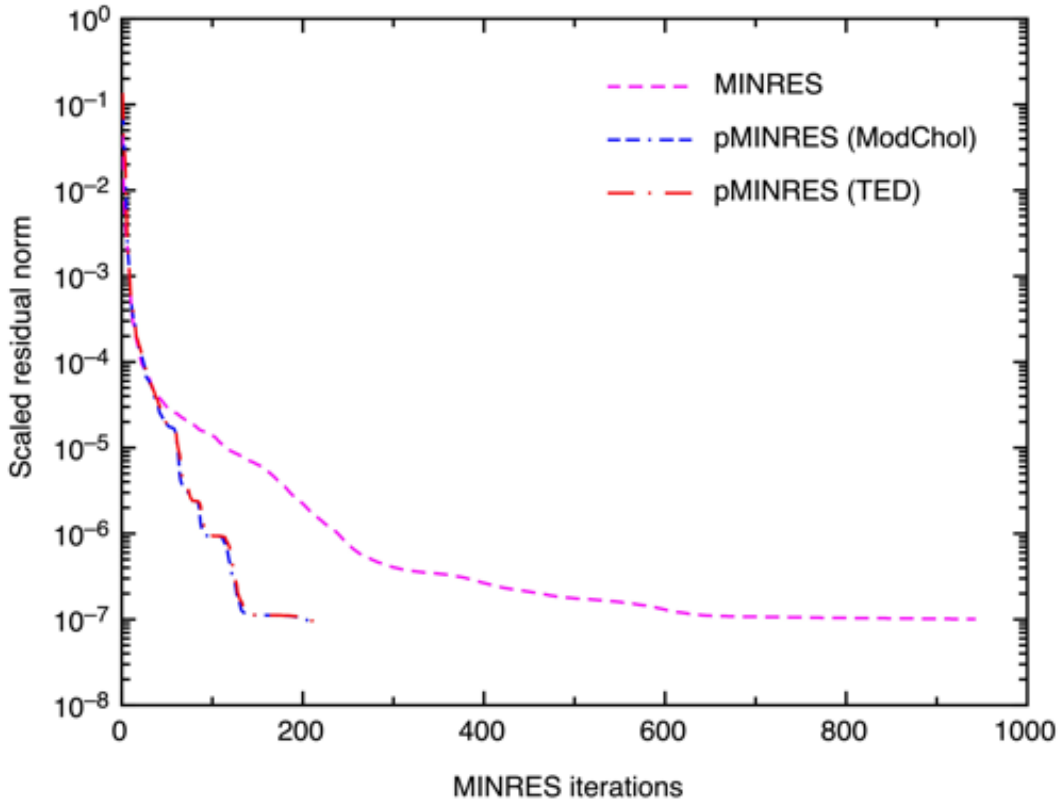

Figure 4.4: Norm of the scaled residual $||\mathbf{r}^{(k)}||/||\mathbf{J}||$ for the C_{60} fullerene as a function of the number of MINRES iterations for different preconditioners. See text for further details.

From Figure 4.4 it is obvious to see that for this basis set and auxiliary function set combination, the effect of the preconditioners is identical. Nevertheless, we experienced that for molecular assemblies, such as water clusters (see also Chapter 5), the number of SCF cycles was increased by a factor of 3 using the modified Cholesky decomposition preconditioner compared to current standard density fitting implementation in deMon2k. It is worth to mention that these systems present a linear dependency of truncated Coulomb matrix eigenvalues, N_{TE}, that arises from the molecular assembly of the structure. Hence, the same calculations were performed using the TED preconditioner inside the MINRES algorithm. With the TED preconditioner in MINRES the SCF convergence was reached in the same number of cycles as with the deMon2k standard density fitting implementation based on the explicit diagonalization of the Coulomb matrix. This underlines once again the importance of the preconditioner for MINRES.

To compare the TED and ModChol preconditioning of MINRES under extreme conditions, we

performed single point energy calculations for a benzene molecule employing the PBE/DZVP/ GEN-A4* level of theory. The here used GEN-A4* auxiliary function set is prone to linear dependencies, even in atomic blocks, due to the dense spacing of auxiliary functions. Figure 4.5 compares the MINRES convergence in the first SCF cycle of this benzene calculation with the two preconditioners, ModChol and TED. As Figure 4.5 shows MINRES with the ModChol preconditioner fails to converge in the first SCF cycle. The scaled residual norm stays almost constant at 10^7 during the MINRES iterations. On the other hand, if the TED preconditioner is used, the convergence for the first SCF cycle is achieved in less than 100 iterations. Thus, TED eliminates from the preconditioner spurious eigenvalues and eigenvectors which can hamper the convergence of MINRES completely. Due to the superior performance of the TED preconditioner in terms of MINRES and SCF convergence, we implemented TED preconditioning as the default for MINRES in deMon2k.

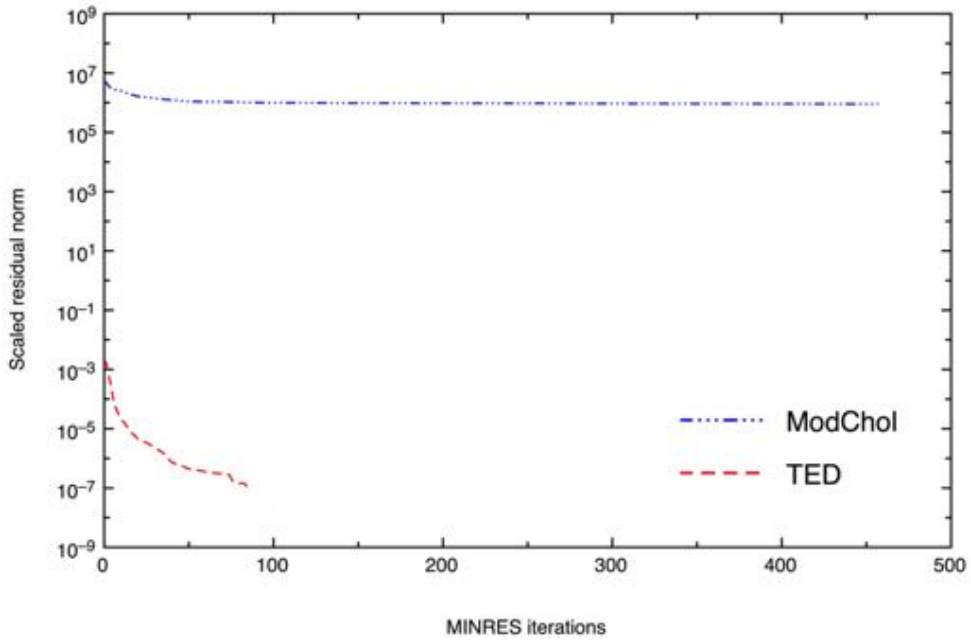

Figure 4.5: Norm of the scaled residual $||\mathbf{r}^{(k)}||/||\mathbf{J}||$ for the benzene molecule as a function of the number of MINRES iterations for different preconditioners. See text for further details.

Because the computational work per iteration for preconditioned and standard MINRES is essentially the same (one product of $\mathbf{G}$ with a vector), preconditioning yields substantial CPU time savings at a very small cost in memory.

4.4 Double asymptotic expansion of two-center ERIs inside MINRES

From a computational point of view the matrix-vector multiplication in each MINRES iteration represents the critical bottleneck of the algorithm. This is particularly true for our quantum chemical MINRES implementation because we must recalculate the **G** matrix on-the-fly in each MINRES iteration in order to keep memory use to a minimum. Because **G** is dense, the solution of Eqs. (4.1) and (4.2) with MINRES scale formally quadratic with respect to the number of auxiliary functions. The quadratic scaling is dominated by the on-the-fly two-center ERI calculation for the **G** matrix elements. To overcome this performance bottleneck we extended the double asymptotic ERI expansion[147] to the two-center ERIs that form the **G** matrix elements. For the variational fitting of the Coulomb potential in deMon2k atom centered primitive Hermite Gaussians, denoted by Latin letters with a bar, are used as auxiliary functions. They are grouped together in sets with common exponents[99]. With these functions an element of the Coulomb matrix is given as

$$G_{\bar{c}\bar{d}} = \langle \bar{c} || \bar{d} \rangle = \iint \frac{\bar{c}(\mathbf{r})\bar{d}(\mathbf{r}\,')}{|\mathbf{r} - \mathbf{r}\,'|} d\mathbf{r}\, d\mathbf{r}\,'. \tag{4.70}$$

This two-electron two-center ERI can be written as a one-electron integral by

$$\langle \bar{c} || \bar{d} \rangle = \int \bar{c}(\mathbf{r}) \phi_{\bar{d}}(\mathbf{r}) d\mathbf{r}. \tag{4.71}$$

The here introduced electrostatic potential, $\phi_{\bar{d}}(\mathbf{r})$, of the primitive Hermite Gaussian auxiliary function $\bar{d}(\mathbf{r})$ is defined as

$$\phi_{\bar{d}}(\mathbf{r}) = \int \frac{\bar{d}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}\,'|} d\mathbf{r}\,'. \tag{4.72}$$

Based on the asymptotic expansion of the Boys function, we find as asymptotic expansion for this potential[95,147]

$$\phi_{\bar{d}}(\mathbf{r}) \sim \left(\frac{\pi}{\zeta_{\bar{d}}}\right)^{3/2} \left(\frac{\partial}{\partial D_x}\right)^{\bar{d}_x} \left(\frac{\partial}{\partial D_y}\right)^{\bar{d}_y} \left(\frac{\partial}{\partial D_z}\right)^{\bar{d}_z} \frac{1}{|\mathbf{r} - \mathbf{D}|}. \tag{4.73}$$

Using this asymptotic expansion and the operator definition

$$\hat{\mathcal{A}}_D(\bar{d}) = \left(\frac{\partial}{\partial D_x}\right)^{\bar{d}_x} \left(\frac{\partial}{\partial D_y}\right)^{\bar{d}_y} \left(\frac{\partial}{\partial D_z}\right)^{\bar{d}_z} \frac{1}{|\mathbf{r} - \mathbf{D}|},$$

we obtain as asymptotic expansion for the Coulomb matrix elements defined in Eq. (4.70) and (4.71)

$$\langle \bar{c}||\bar{d}\rangle \sim \left(\frac{\pi}{\zeta_D}\right)^{3/2} \langle \bar{c}|\hat{\mathcal{A}}_D(\bar{d})\rangle. \tag{4.74}$$

Thus, the first asymptotic expansion of the $\langle \bar{c}||\bar{d}\rangle$ ERI yields a nuclear attraction like integral at nucleus D where the asymptotically expanded auxiliary function potential, $\phi_{\bar{d}}(\mathbf{r})$, is centered. Because $\bar{c}(\mathbf{r})$ is a primitive Hermite Gaussian function the nuclear attraction integral in Eq. (4.74) can be rewritten as

$$\begin{aligned}\langle \bar{c}|\hat{\mathcal{A}}_D(\bar{d})\rangle &= \widehat{\mathcal{D}}_C(\bar{c})\widehat{\mathcal{D}}_D(\bar{d})\langle \bar{s}|\hat{\mathcal{A}}_D(\mathbf{0})\rangle, \\ &= \widehat{\mathcal{D}}_C(\bar{c})\widehat{\mathcal{D}}_D(\bar{d})\frac{2\pi}{\zeta_{\bar{c}}}F_0\left[\zeta_{\bar{c}}(\mathbf{C}-\mathbf{D})^2\right].\end{aligned} \tag{4.75}$$

Here F_0 is the zeroth-order Boys function[83], and, $\widehat{\mathcal{D}}_C$ and $\widehat{\mathcal{D}}_D$ denote the operators

$$\widehat{\mathcal{D}}_C(\bar{c}) \equiv \left(\frac{\partial}{\partial C_x}\right)^{\bar{c}_x}\left(\frac{\partial}{\partial C_y}\right)^{\bar{c}_y}\left(\frac{\partial}{\partial C_z}\right)^{\bar{c}_z},$$

$$\widehat{\mathcal{D}}_D(\bar{d}) \equiv \left(\frac{\partial}{\partial D_x}\right)^{\bar{d}_x}\left(\frac{\partial}{\partial D_y}\right)^{\bar{d}_y}\left(\frac{\partial}{\partial D_z}\right)^{\bar{d}_z}.$$

Applying once again the asymptotic expansion of the Boys function in Eq. (4.75) yields as double asymptotic expansion for the two-electron two-center ERIs of the Coulomb matrix

$$\begin{aligned}\langle \bar{c}||\bar{d}\rangle &\sim \left(\frac{\pi}{\zeta_{\bar{c}}}\right)^{3/2}\left(\frac{\pi}{\zeta_{\bar{d}}}\right)^{3/2}\widehat{\mathcal{D}}_C(\bar{c})\widehat{\mathcal{D}}_D(\bar{d})\frac{1}{|\mathbf{C}-\mathbf{D}|} \\ &= (-1)^{\bar{d}}\left(\frac{\pi}{\zeta_{\bar{c}}}\right)^{3/2}\left(\frac{\pi}{\zeta_{\bar{d}}}\right)^{3/2}\widehat{\mathcal{D}}_C(\bar{c}+\bar{d})\frac{1}{|\mathbf{C}-\mathbf{D}|} \\ &= (-1)^{\bar{c}}\left(\frac{\pi}{\zeta_{\bar{c}}}\right)^{3/2}\left(\frac{\pi}{\zeta_{\bar{d}}}\right)^{3/2}\widehat{\mathcal{D}}_D(\bar{c}+\bar{d})\frac{1}{|\mathbf{C}-\mathbf{D}|}.\end{aligned} \tag{4.76}$$

Within the double asymptotic expansion of $\langle \bar{c}||\bar{d}\rangle$ the auxiliary function shells of atoms can be collected together. As a result, the corresponding matrix-vector multiplication scales with the number of atoms rather than auxiliary functions.

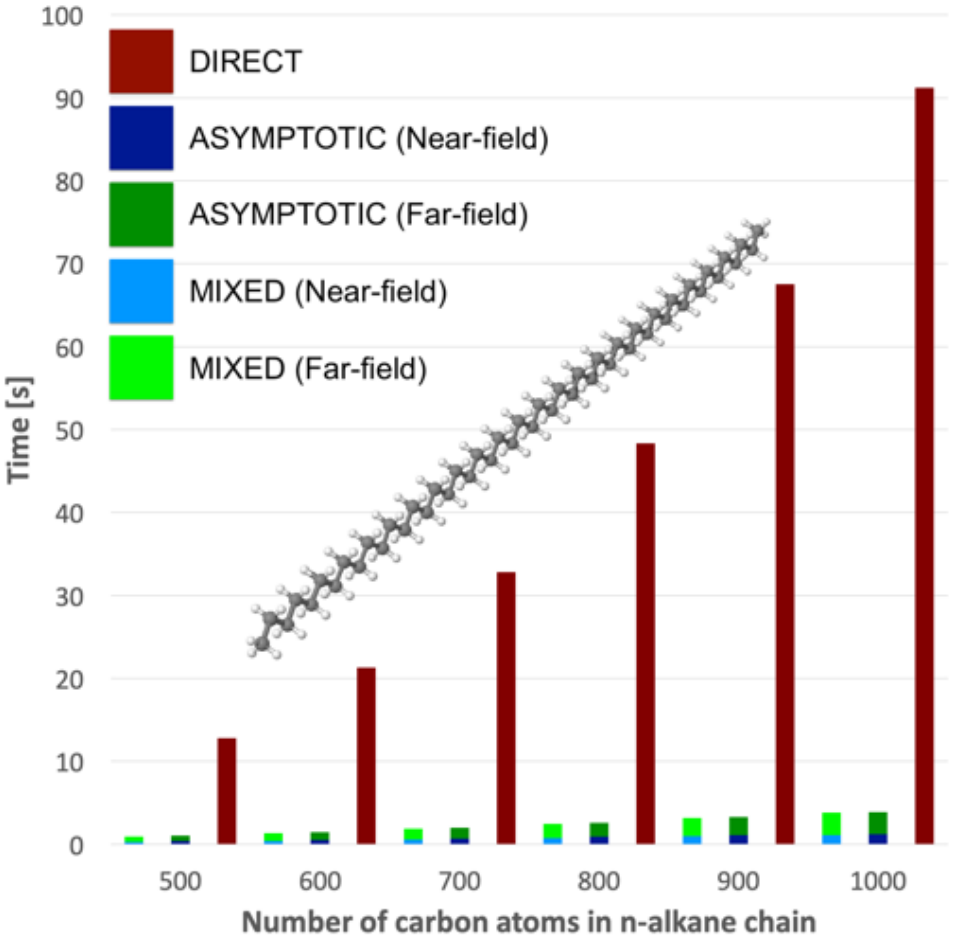

Figure 4.6: Timings for 1 MINRES matrix-vector multiplication using the DIRECT (red bars), ASYMPTOTIC (dark green and blue bars) and MIXED (light green and blue bars) approaches for the calculation of two-center ERIs.

With the double asymptotic expansion of the two-center ERIs, the matrix-vector product in MINRES can be divided into near-field and far-field multiplications. The latter scales with the number of atoms rather than auxiliary functions. Thus, a significant performance improvement can be expected. To this end, we benchmarked the MINRES matrix-vector multiplications for a set of linear n-alkane chains with 500 to 1000 carbon atoms employing the PBE/DZVP/GEN-A2* methodology. The timings for 24 core runs of three different matrix-vector multiplications approaches are depicted in Figure 4.6. The DIRECT approach (red bars in Figure 4.6) refers to MINRES matrix-vector multiplications with on-the-fly calculated **G** matrix elements. For the largest system, $C_{1000}H_{2002}$, with more than 130,000 auxiliary functions a single matrix-vector multiplication takes around 90 seconds with the DIRECT approach. This time can be

substantially reduced with the ASYMPTOTIC approach for the matrix-vector multiplication in MINRES. Here, the near-field ERIs (dark blue bar in Figure 4.6) are calculated on-the-fly with standard recurrence relations, whereas the far-field ERIs (dark green bar in Figure 4.6) are calculated with the double asymptotic ERI expansion. As a result, the time for a matrix-vector multiplication in MINRES is reduced to less than 5 seconds. It is worthy of note that most of this time is spent for the calculation of the far-field two center ERIs (dark green bar in Figure 4.6) which is a marked difference to the corresponding calculation of three-center ERIs. This also explains why the MIXED (light green and blue bars in Figure 4.6) approach, in which the near-field two-center ERIs are stored in RAM rather than being recalculated on-the-fly, shows only little improvement over the ASYMPTOTIC one for the MINRES matrix-vector multiplications.

4.5 Perturbation theory with MINRES

Many molecular properties are obtained as (higher order) perturbations of the molecular energy with respect to corresponding perturbation parameters. The associated perturbation problem has usually to be solved iteratively due to the large size of the equation systems involved. In Kohn-Sham DFT the coupled perturbed Kohn-Sham (CPKS)[148–151] method is used for this purpose. Recently, there have been many efforts to reduce the computational complexity of CPKS methods by several authors[152–161]. A computational attractive alternative to CPKS is auxiliary density perturbation theory (ADPT). ADPT is based on self-consistent perturbation (SCF) theory[162] applied to ADFT. It has proven to be robust and efficient for the calculation of response properties such as polarizabilities[163,164], Fukui functions[165,166], vibrational frequencies[167] and nuclear spin-spin coupling constants[168]. The basic idea of ADPT is to develop the molecular response through the auxiliary density instead of the orbital density. As a consequence, the response of the density matrix is substituted by the response of the Coulomb and exchange-correlation fitting coefficients.

In the original ADPT implementation in deMon2k[169] the response matrix $\mathbf{R}$ was built and inverted in order to solve the ADPT equation system. The computational bottleneck of this approach is not the $\mathcal{O}(N_{aux}^3)$ matrix inversion but rather the $\mathcal{O}(N^4)$ building of $\mathbf{R}$. To avoid this bottleneck, one may choose to solve the ADPT equations by an iterative procedure similar to the MINRES algorithm used for solving the variational Coulomb fitting equation systems. The iterative procedure chosen for this purpose was the Eirola-Nevanlinna (EN) algorithm[129] because it is well suited for large nonsymmetrical indefinite linear equation systems. To proceed we first derive the ADPT equation systems and discuss their solution with the EN algorithm.

Throughout the following discussion, the perturbed quantities are denoted by λ superscripts enclosed in parentheses; small Greek letters, μ and ν, denote atomic orbitals and small Latin letters with a bar, $\bar{k}$ and $\bar{l}$, denote auxiliary functions. According to McWeeny's[162] self-consistent perturbation (SCP) theory an element of the dynamic first-order perturbed density matrix is given by

$$P_{\mu\nu}^{(\lambda)}(\omega) = 2\sum_{i}^{occ}\sum_{a}^{uno}\frac{\mathcal{K}_{ai}^{(\lambda)}(\omega)}{\epsilon_i - \epsilon_a - \omega}c_{\mu a}c_{\nu i} + \sum_{i}^{occ}\sum_{a}^{uno}\frac{\mathcal{K}_{ia}^{(\lambda)}(\omega)}{\epsilon_i - \epsilon_a + \omega}c_{\mu i}c_{\nu a}, \tag{4.77}$$

where the ϵ_i correspond to the occupied MO energies and the ϵ_a to the unoccupied MO energies. The $\boldsymbol{\mathcal{K}}^{(\lambda)}(\omega)$ is the molecular orbital representation of the perturbed Kohn-Sham matrix whose elements are given by

$$\mathcal{K}_{ia}^{(\lambda)}(\omega) = \sum_{\mu,\nu} c_{\mu i}\, K_{\mu\nu}^{(\lambda)}(\omega)\, c_{\nu a}. \tag{4.78}$$

The perturbed ADFT Kohn-Sham matrix elements are given by

$$K_{\mu\nu}^{(\lambda)}(\omega) = H_{\mu\nu}^{(\lambda)} + \sum_{\bar{k}}\langle\mu\nu||\bar{k}\rangle\left(x_{\bar{k}}^{(\lambda)}(\omega) + z_{\bar{k}}^{(\lambda)}(\omega)\right). \tag{4.79}$$

The here appearing perturbed core Hamilton matrix elements, $H_{\mu\nu}^{(\lambda)}$, depend on the particular perturbation being studied. The perturbed fitting coefficients, $\mathbf{x}^{(\lambda)}(\omega)$, and the perturbed exchange-correlation fitting coefficients, $\mathbf{z}^{(\lambda)}(\omega)$, must be computed in order to obtain $\mathbf{P}^{(\lambda)}(\omega)$. To this end, ADPT takes advantage of the perturbed Coulomb fitting equation systems which, for perturbation-independent basis and auxiliary functions, takes the form

$$\sum_{\bar{l}} G_{\bar{k}\bar{l}}\, x_{\bar{l}}^{(\lambda)}(\omega) = \sum_{\mu,\nu}\langle\bar{k}||\mu\nu\rangle P_{\mu\nu}^{(\lambda)}(\omega). \tag{4.80}$$

Inserting Eq. (4.77) into Eq. (4.80) yields

$$\begin{aligned}\sum_{\bar{l}} G_{\bar{k}\bar{l}}x_{\bar{l}}^{(\lambda)}(\omega) = {} & 2\sum_{\mu,\nu}\sum_{i}^{occ}\sum_{a}^{uno}\frac{\mathcal{K}_{ai}^{(\lambda)}(\omega)}{\omega_{ia} - \omega}c_{\mu a}c_{\nu i}\langle\mu\nu\|\bar{k}\rangle + \\ & 2\sum_{\mu,\nu}\sum_{i}^{occ}\sum_{a}^{uno}\frac{\mathcal{K}_{ia}^{(\lambda)}(\omega)}{\omega_{ia} + \omega}c_{\mu i}c_{\nu a}\langle\mu\nu\|\bar{k}\rangle.\end{aligned} \tag{4.81}$$

Here we have introduced $\omega_{ia} = \epsilon_i - \epsilon_a$. The right hand side of Eq. (4.81) can be further simplified by performing the sum over the AOs and using the symmetry of the ERIs and the perturbed Kohn-Sham matrix

$$\sum_{\bar{l}} G_{\bar{k}\bar{l}} x_{\bar{l}}^{(\lambda)}(\omega) = 4 \sum_{i}^{occ} \sum_{a}^{uno} \mathcal{K}_{ia}^{(\lambda)}(\omega) \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \langle ia||\bar{k}\rangle. \tag{4.82}$$

Note that this simplification holds only for LDA and GGA[170]. To proceed further, we now expand $\mathcal{K}_{ia}^{(\lambda)}(\omega)$ according to Eqs. (4.78) and (4.79)

$$\begin{aligned}\sum_{\bar{l}} G_{\bar{k}\bar{l}} x_{\bar{l}}^{(\lambda)}(\omega) = & 4 \sum_{i}^{occ} \sum_{a}^{uno} \mathcal{H}_{ia}^{(\lambda)} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \langle ia||\bar{k}\rangle + \\ & 4 \sum_{\bar{l}} \sum_{i}^{occ} \sum_{a}^{uno} \langle \bar{l}||ia\rangle \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \langle ia||\bar{k}\rangle \left[x_{\bar{l}}^{(\lambda)}(\omega) + z_{\bar{l}}^{(\lambda)}(\omega) \right],\end{aligned} \tag{4.83}$$

where the $\mathcal{H}_{ia}^{(\lambda)}$ are elements of the perturbed core Hamilton matrix in MO representation

$$\mathcal{H}_{ia}^{(\lambda)} = \sum_{\mu,\nu} c_{\mu i} H_{\mu\nu}^{(\lambda)} c_{\nu a}. \tag{4.84}$$

Collecting all terms that depend on the perturbed density fitting coefficients on the left side transforms Eq. (4.83) into

$$\sum_{\bar{l}} G_{\bar{k}\bar{l}} x_{\bar{l}}^{(\lambda)}(\omega) - 4 \sum_{\bar{l}} A_{\bar{k}\bar{l}}(\omega) \left[x_{\bar{l}}^{(\lambda)}(\omega) + z_{\bar{l}}^{(\lambda)}(\omega) \right] = 4\, b_{\bar{k}}^{(\lambda)}(\omega), \tag{4.85}$$

where

$$A_{\bar{k}\bar{l}}(\omega) = \sum_{i}^{occ} \sum_{a}^{uno} \langle \bar{k}||ia\rangle \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \langle ia||\bar{l}\rangle, \tag{4.86}$$

is an element of the Coulomb response matrix, $\mathbf{A}(\omega)$, and

$$b_{\bar{k}}^{(\lambda)}(\omega) = \sum_{i}^{occ} \sum_{a}^{uno} \mathcal{H}_{ia}^{(\lambda)} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \langle ia||\bar{k}\rangle, \tag{4.87}$$

is an element of the perturbation vector $\mathbf{b}^{(\lambda)}(\omega)$. The perturbed exchange-correlation fitting

coefficients in Eq. (4.85) can be expressed as

$$z_{\bar{l}}^{(\lambda)}(\omega) = \sum_{\bar{k},\bar{m}} G_{\bar{l}\bar{k}}^{-1} F_{\bar{k}\bar{m}} x_{\bar{m}}^{(\lambda)} \tag{4.88}$$

with the ADPT kernel matrix defined by

$$F_{\bar{k}\bar{m}} = \langle \bar{k} | f_{xc}[\tilde{\rho}] | \bar{m} \rangle. \tag{4.89}$$

Here f_{xc} is the exchange-correlation kernel that corresponds to the functional derivative of the exchange-correlation potential $v_{xc}[\tilde{\rho}; \mathbf{r}]$. Finally, inserting Eqs. (4.86), (4.87) and (4.88) into (4.85) yields the following linear equation system in matrix form

$$\mathbf{R}(\omega)\mathbf{x}^{(\lambda)}(\omega) = \mathbf{b}^{(\lambda)}(\omega), \tag{4.90}$$

where $\mathbf{R}(\omega)$ is the response matrix given by

$$\mathbf{R}(\omega) = \tfrac{1}{4}\mathbf{G} - \mathbf{A}(\omega)(\mathbf{E} + \mathbf{G}^{-1}\mathbf{F}). \tag{4.91}$$

In order to solve the equation system in Eq. (4.90) a version of the Eirola-Nevanlinna (EN2)[129] algorithm has been implemented in deMon2k[163]. This algorithm solves iteratively Eq. (4.90). Afterwards, the obtained Coulomb perturbed fitting coefficients vector, $\mathbf{x}^{(\lambda)}$, is used to obtain the exchange-correlation fitting coefficients vector, $\mathbf{z}^{(\lambda)}$, Eq. (4.88), simply by matrix-vector products. The EN2 algorithm is also a Krylov subspace method. Its pseudocode is shown in Figure 4.7.

To avoid the explicit calculation of the response matrix a direct variant of the EN2 algorithm is implemented in deMon2k. This method is triggered by the DIRECT[105] option of the ADPTYPE keyword. Hence, it avoids the explicit calculation and storage of the **A** and **F** matrices. Instead, it directly calculates the actions of these matrices on trial vectors. Taking as example the multiplication of the Coulomb response matrix, Eq. (4.86), with a trial vector of perturbed

Coulomb fitting coefficients yields

$$a_{\bar{k}}(\omega) \equiv \sum_{\bar{l}} A_{\bar{k}\bar{l}} x_{\bar{l}}^{(\lambda)} = \sum_{\bar{l}} \sum_{i}^{occ} \sum_{a}^{uno} \langle \bar{k}||ia\rangle \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \langle ia||\bar{l}\rangle x_{\bar{l}}^{(\lambda)}$$
$$= \sum_{i}^{occ} \sum_{a}^{uno} \sum_{\mu,\nu} \sum_{\sigma,\tau} \sum_{\bar{l}} \langle \bar{k}||\mu\nu\rangle c_{\mu i} c_{\nu a} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} c_{\sigma i} c_{\tau a} \langle \sigma\tau||\bar{l}\rangle x_{\bar{l}}^{(\lambda)}. \tag{4.92}$$

1: **procedure** EN2
2: Initialization: $\mathbf{x}_0^{(\lambda)}$, $\mathbf{H}_0$ arbitrary, $\mathbf{r}_0 = \mathbf{b}^{(\lambda)} - \mathbf{R}\mathbf{x}_0^{(\lambda)}$
3: **for** $j = 0$, n **do**
4: **if** $\|\mathbf{r}_j\|_\infty \leq \tau$ **then**
5: Exit
6: **end if**
7: $\alpha_i = \mathbf{c}_i^T(\mathbf{r}_j - \mathbf{R}\mathbf{H}_0\mathbf{r}_j)$ for $i = 0, \ldots, j-1$
8: $\boldsymbol{\eta}_j = \mathbf{H}_0\mathbf{r}_j + \sum_{i=0}^{j-1} \alpha_i \mathbf{u}_i$
9: $\boldsymbol{\xi}_j = \mathbf{r}_j - \mathbf{R}\mathbf{H}_0\mathbf{r}_j - \sum_{i=0}^{j-1} \alpha_i \mathbf{c}_i$
10: $\beta_i = \mathbf{c}_i^T \mathbf{R}\mathbf{H}_0\boldsymbol{\xi}_j$ for $i = 0, \ldots, j-1$
11: $\mathbf{u}_j = \tau(\mathbf{H}_0\boldsymbol{\xi}_j - \sum_{i=0}^{j-1} \beta_i \mathbf{u}_i)$
12: $\mathbf{c}_j = \tau(\mathbf{R}\mathbf{H}_0\boldsymbol{\xi}_j - \sum_{i=0}^{j-1} \beta_i \mathbf{c}_i)$, where τ is such that $\|\mathbf{c}_j\|_2 = 1$
13: $\mathbf{x}_{j+1}^{(\lambda)} = \mathbf{x}_j^{(\lambda)} + \boldsymbol{\eta}_j + \mathbf{u}_j \mathbf{c}_j^T \boldsymbol{\xi}_j$
14: $\mathbf{r}_{j+1} = \boldsymbol{\xi}_j - \mathbf{c}_j \mathbf{c}_j^T \boldsymbol{\xi}_j$
15: **end for**
16: **end procedure**

Figure 4.7: Pseudocode for the Eirola-Nevanlinna iterative algorithm (EN2) for solving nonsymmetric linear equation systems.

In Eq. (4.92) the expansion of the molecular three-center ERIs into their atomic counterparts avoids the $\mathcal{O}(N^4)$ AO to MO ERI transformation. As a result, a low-order scaling algorithm for currently accessible system sizes, i.e. molecules with less than 100,000 auxiliary functions, is

obtained. Whereas the direct EN2 variant is the method of choice for large molecules its conventional counterpart triggered by ADPTYPE CONVENTIONAL[105], can be computationally superior for small systems with less than 50,000 auxiliary functions. In this system size range the unfavourable scaling of the **A** matrix calculation can be compensated by the fact that this matrix has only to be computed one time, independent of how many perturbations have to be calculated. Thus, the conventional EN2 algorithm becomes particularly advantageous for the calculation of many perturbations for a fixed molecular structure like for second analytic ADFT energy derivatives and nuclear spin-spin coupling constants, to name a few.

As a Krylov subspace method, the EN2 also uses a preconditioner ($\mathbf{H}_0$, in Figure 4.7) in order to accelerate its convergence. Instead of inverting the **R** matrix, which would be the perfect preconditioner for this equation system, $\mathbf{G}^{-1}$ is used as preconditioner. With this implementation, only two actions (matrix-vector products) of the **R** matrix are needed (Figure 4.7 lines 7 and 9, for the first action and lines 10 and 12 for the second one). This algorithm does not update the inverse matrix explicitly as the BFGS algorithm, instead only the actions of those updates are carried through the vectors **u** and **c** (lines 8-14 in Figure 4.7). Furthermore, to avoid memory overflow these **u** and **c** vectors are kept and discarded after m steps of the algorithm. In deMon2k $m = 15$ is used. Due to its low memory demand and the good convergence performance the EN2 is the method of choice for the solution of the ADPT equation system. Nevertheless, the current implementation still requires explicitly $\mathbf{G}^{-1}$, which limits treatable systems sizes on many HPC architectures to around 100,000 auxiliary functions. To demonstrate how this limitation can be lifted by the use of MINRES within the EN2 algorithm, we now discuss the action of **A** on a trial vector of perturbed Coulomb fitting coefficients, Eq. (4.92), step by step. To this end, we first rewrite Eq. (4.92) as

$$a_{\bar{k}}(\omega) = \sum_{\mu,\nu} \langle \bar{k} || \mu\nu \rangle \sum_{i}^{occ} \sum_{a}^{uno} c_{\mu i} c_{\nu a} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \sum_{\sigma,\tau} c_{\sigma i} c_{\tau a} \sum_{\bar{l}} \langle \sigma\tau || \bar{l} \rangle x_{\bar{l}}^{(\lambda)}. \tag{4.93}$$

To proceed, we define the matrix **Q** with elements

$$Q_{\sigma\tau} = \sum_{\bar{l}} \langle \sigma\tau || \bar{l} \rangle x_{\bar{l}}^{(\lambda)}. \tag{4.94}$$

The sum in Eq. (4.94) resembles the one appearing in the Kohn-Sham matrix construction, Eq. (3.34), for the direct SCF procedure. Therefore, integral screening and double asymptotic ERI

expansions can be used to obtain an asymptotic $\mathcal{O}(N_{bas})$ scaling for the calculation of the **Q** matrix elements. Substituting Eq. (4.94) into Eq. (4.93) yields

$$a_{\bar{k}}(\omega) = \sum_{\mu,\nu} \langle \bar{k}||\mu\nu\rangle \sum_{i}^{occ} \sum_{a}^{uno} c_{\mu i} c_{\nu a} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \sum_{\sigma,\tau} c_{\sigma i} Q_{\sigma\tau} c_{\tau a}. \tag{4.95}$$

The transformation of **Q** into MO representation yields

$$Q_{ia} = \sum_{\sigma,\tau} c_{\sigma i} Q_{\sigma\tau} c_{\tau a}. \tag{4.96}$$

Although these matrix multiplications introduce a cubic scaling step, it is in practical applications most often not noticeable. The reason lies in its implementation with optimized BLAS routines that perform close to peak performance and, thus, hide these steps[171]. Then, the resulting Q_{ia} elements are scaled with the corresponding MO energy differences

$$Q'_{ia}(\omega) = \frac{Q_{ia} w_{ia}}{w_{ia}^2 - w^2}. \tag{4.97}$$

Substituting Eq. (4.97) into Eq. (4.93) gives the following expression

$$a_{\bar{k}}(\omega) = \sum_{\mu,\nu} \langle \bar{k}||\mu\nu\rangle \sum_{i}^{occ} \sum_{a}^{uno} c_{\mu i} Q'_{ia}(\omega) c_{\nu a}. \tag{4.98}$$

Next, the $Q'_{ia}(\omega)$ matrix elements are transformed into AO representation

$$T_{\mu\nu}(\omega) = \sum_{i}^{occ} \sum_{a}^{uno} c_{\mu i} Q'_{ia}(\omega) c_{\nu a}. \tag{4.99}$$

For this step the same comments as for Eq. (4.96) hold. By inserting Eq. (4.99) into Eq. (4.98), it follows that

$$a_{\bar{k}}(\omega) = \sum_{\mu,\nu} \langle \bar{k}||\mu\nu\rangle T_{\mu\nu}(\omega). \tag{4.100}$$

This equation is algebraically identical to the one for the calculation of the Coulomb vector **J** according to Eq. (3.31). Integral screening and double asymptotic expansions for three-center ERIs can be used to reduce the scaling from $\mathcal{O}(N_{bas}^3)$ to $\mathcal{O}(N_{bas})$. As this discussion shows the direct action of $\mathbf{A}(\omega)$ on a test vector can be calculated in a nearly linear scaling approach

without the explicit use of $\mathbf{G}^{-1}$.

Similarly, the third term in Eq. (4.91) can be expanded as

$$a'_{\bar{k}}(\omega) = \sum_{\bar{m},\bar{n}} A_{\bar{k}\bar{m}}(\omega) G^{-1}_{\bar{m}\bar{n}} \sum_{\bar{l}} \langle \bar{n}|f_{xc}[\tilde{\rho}]|\bar{l}\rangle x^{(\lambda)}_{\bar{l}}. \tag{4.101}$$

For the efficient calculation of this term a vector $\mathbf{q}$ is defined with elements

$$q_{\bar{n}} = \sum_{\bar{l}} \langle \bar{n}|f_{xc}[\tilde{\rho}]|\bar{l}\rangle x^{(\lambda)}_{\bar{l}} = \langle \bar{n}|f_{xc}[\tilde{\rho}]|\tilde{\sigma}\rangle, \tag{4.102}$$

where $\tilde{\sigma}$ is given by

$$\tilde{\sigma}(\mathbf{r}) = \sum_{\bar{l}} \bar{l}(\mathbf{r})\, x^{(\lambda)}_{\bar{l}}. \tag{4.103}$$

This step involves the computation of a vector with the size of the auxiliary function set over the grid. Thus, the formal scaling is expected to be $\mathcal{O}(N_{aux} \times N_{grid})$. This scaling can be further reduced by grid screening techniques. Introducing the $\mathbf{q}$ vector into Eq. (4.101) yields

$$a'_{\bar{k}}(\omega) = \sum_{\bar{m},\bar{n}} A_{\bar{k}\bar{m}} G^{-1}_{\bar{m}\bar{n}} q_{\bar{n}}. \tag{4.104}$$

The next step is the multiplication of the $\mathbf{q}$ vector with the inverse of the Coulomb matrix, i.e. $\mathbf{q}' = \mathbf{G}^{-1}\mathbf{q}$. In the current ADPT implementation in deMon2k the $\mathbf{G}^{-1}$ matrix needed for this operation is read from disk were it was stored in the corresponding TED step. However, if MINRES is used in the SCF the TED step is avoided and $\mathbf{G}^{-1}$ is at this point of the calculation not available. Therefore, we reformulate the multiplication of $\mathbf{q}$ with $\mathbf{G}^{-1}$ in form of the following linear equation system

$$\mathbf{G}\mathbf{q}' = \mathbf{q}, \tag{4.105}$$

which we solve with the here developed MINRES algorithm. Once $\mathbf{q}'$ is at hand Eq. (4.101) can be simplified to

$$a'_{\bar{k}}(\omega) = \sum_{\bar{l}} A_{\bar{k}\bar{l}}(\omega) q'_{\bar{l}}. \tag{4.106}$$

The multiplication of the $\mathbf{A}(\omega)$ matrix with the $\mathbf{q}'$ vector is calculated according to the previously described action algorithm for the Coulomb response matrix.

Finally, the first term of Eq. (4.91) yields in the action algorithm the multiplication of the

Coulomb matrix $\mathbf{G}$ with a trial vector of perturbed Coulomb fitting coefficients. This matrix-vector product is always calculated on-the-fly and it can be performed in the three different forms already discussed in Section 4.4. In conclusion, the here obtained action algorithm for the response matrix $\mathbf{R}$ on a trial vector of Coulomb fitting coefficients is nearly linear scaling and free of $\mathbf{G}^{-1}$ operations.

Therefore, the only $\mathbf{G}^{-1}$ operations left are the preconditioning steps in the EN2 algorithm. They are given by the $\mathbf{H}_0 = \mathbf{G}^{-1}$ multiplications in line 7 and 10 of the EN2 pseudocode in Figure 4.8. Those are

$$\mathbf{r}' = \mathbf{H}_0\mathbf{r} \rightarrow \mathbf{r}' = \mathbf{G}^{-1}\mathbf{r}, \tag{4.107}$$

and

$$\xi' = \mathbf{H}_0\xi \rightarrow \xi' = \mathbf{G}^{-1}\xi. \tag{4.108}$$

In the new EN2 + MINRES (see Figure 4.8) implementation these matrix multiplications are reformulated into the linear equation systems

$$\mathbf{G}\mathbf{r}' = \mathbf{r} \tag{4.109}$$

and

$$\mathbf{G}\xi' = \xi \tag{4.110}$$

that are solved by the here developed MINRES algorithm (see Figure 4.8).

Thus, MINRES is called three times, two times for preconditioning, Eqs. (4.109) and (4.110), as shown in Figure 4.8 and one time to solve iteratively Eq. (4.88) as

$$\mathbf{G}\mathbf{z}^{(\lambda)} = \mathbf{F}\mathbf{x}^{(\lambda)}. \tag{4.111}$$

This concludes our description of the new EN2 + MINRES algorithm that is free of $\mathbf{G}^{-1}$ operations. With this thesis the new EN2 + MINRES algorithm was also implemented in deMon2k for the calculation of polarizabilities and hyperpolarizabilities, Fukui functions and nuclear spin-spin coupling constants. Illustrative examples of its application are shown in the next chapter.

1: **procedure** EN2 + MINRES

2: Initialization: $\mathbf{x}_0^{(\lambda)}$, $\mathbf{H}_0$ arbitrary, $\mathbf{r}_0 = \mathbf{b}^{(\lambda)} - \mathbf{R}\mathbf{x}_0^{(\lambda)}$

3: **for** $j = 0, n$ **do**

4: **if** $\|\mathbf{r}_j\|_\infty \leq \tau$ **then**

5: Exit

6: **end if**

7: $\alpha_i = \mathbf{c}_i^T(\mathbf{r}_j - \mathbf{R}\mathbf{H}_0\mathbf{r}_j)$ for $i = 0, \ldots, j-1$

8: $\boldsymbol{\eta}_j = \mathbf{H}_0\mathbf{r}_j + \sum_{i=0}^{j-1} \alpha_i \mathbf{u}_i$

9: $\boldsymbol{\xi}_j = \mathbf{r}_j - \mathbf{R}\mathbf{H}_0\mathbf{r}_j - \sum_{i=0}^{j-1} \alpha_i \mathbf{c}_i$

10: $\beta_i = \mathbf{c}_i^T \mathbf{R}\mathbf{H}_0\boldsymbol{\xi}_j$ for $i = 0, \ldots, j-1$

11: $\mathbf{u}_j = \tau(\mathbf{H}_0\boldsymbol{\xi}_j - \sum_{i=0}^{j-1} \beta_i \mathbf{u}_i)$

12: $\mathbf{c}_j = \tau(\mathbf{R}\mathbf{H}_0\boldsymbol{\xi}_j - \sum_{i=0}^{j-1} \beta_i \mathbf{c}_i)$, where τ is such that $\|\mathbf{c}_j\|_2 = 1$

13: $\mathbf{x}_{j+1}^{(\lambda)} = \mathbf{x}_j^{(\lambda)} + \boldsymbol{\eta}_j + \mathbf{u}_j\mathbf{c}_j^T\boldsymbol{\xi}_j$

14: $\mathbf{r}_{j+1} = \boldsymbol{\xi}_j - \mathbf{c}_j\mathbf{c}_j^T\boldsymbol{\xi}_j$

15: **end for**

16: **end procedure**

Figure 4.8: Pseudocode for the Eirola-Nevanlinna EN2 + MINRES iterative algorithm for solving nonsymmetric linear equation systems.

5

Validations, Benchmarks and Applications

Contents

5.1 Energy calculations

To assess the performance of the new MINRES fitting approach, we performed single point energy calculations of systems with systematically increasing sizes[130] using three benchmark sets. The first benchmark set consists of 7 fullerene cages ranging from 20 to 960 carbon atoms, the second of 6 water clusters containing 50 to 500 water molecules and the third set consists of saturated MFI zeolites containing 1 to 4 unit cells. The calculations were conducted with the iterative MINRES algorithm using three options: DIRECT, ASYMPTOTIC and MIXED. The MINRES timings are compared with the default TED based approach[35,36] currently implemented in deMon2k. The methodology used for these benchmark calculations was PBE/DZVP/GEN-A2*. The calculations of the fullerene and water clusters were performed using 24 Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20GHz cores with 4 GB of RAM per core. For the zeolite calculations 96 cores with the same specification were used. The results of these single point energy calculations will be presented and discussed in the following three subsections.

5.1.1 Fullerenes

Table 5.1 lists for the fullerene benchmark set the total density fitting and single point SCF timings (in seconds) for the TED as well as the MINRES DIRECT, ASYMPTOTIC and MIXED density fitting approaches.

Table 5.1: Number of auxiliary functions, N_{aux}, CPU timings for the density fitting and single point SCF energy calculations using the listed density fitting approaches as well as the number of truncated eigenvalues in TED, N_{TE}, for the systems in the fullerene benchmark set. All calculations were performed with the PBE/DZVP/GEN-A2* methodology on a 24 core board.

System	N_{aux}	Total density fitting timings [s]				Total SCF timings [s]				N_{TE}
		TED	DIRECT	ASYMPTOTIC	MIXED	TED	DIRECT	ASYMPTOTIC	MIXED	
C_{20}	1860	1.7	14.8	29.19	20.4	8.2	21.0	35.5	26.7	163
C_{60}	5580	61.6	43.7	53.7	17.7	116.3	97.2	115.7	71.2	590
C_{180}	16740	1573.1	566.5	506.8	134.9	2185.8	1170.7	1109.9	738.7	2212
C_{240}	22320	3694.4	1032.9	556.5	206.2	4687.6	2011.4	1533.9	1187.0	2794
C_{540}	50220	43150.6	6885.6	2586.0	1450.2	47492.3	11699.7	7401.5	6275.7	6625
C_{720}	66960	103841.7	13823.7	4148.1	2740.6	112594.3	22647.7	12965.5	11062.4	8796
C_{960}	89280	255725.4	26665.1	7386.6	5665.9	274767.2	44616.4	25239.4	23634.8	12024

The total fitting timings are plotted in Figure 5.1 versus the number of auxiliary functions, N_{aux},

given in the second column of Table 5.1. As Table 5.1 shows the MINRES density fitting CPU times are smaller than the corresponding TED timings from C_{60} on. Such a crossover is expected because of the different formal scalings of TED and MINRES with respect to system size. Whereas TED shows a cubic scaling (red dots in Figure 5.1) due to the one-time diagonalization of the Coulomb matrix **G**, MINRES possesses a formal quadratic scaling (blue stars in Figure 5.1) due to the matrix-vector multiplications. As Table 5.1 shows this crossover appears for the here presented MINRES implementation at C_{60} in the fullerene benchmark set.

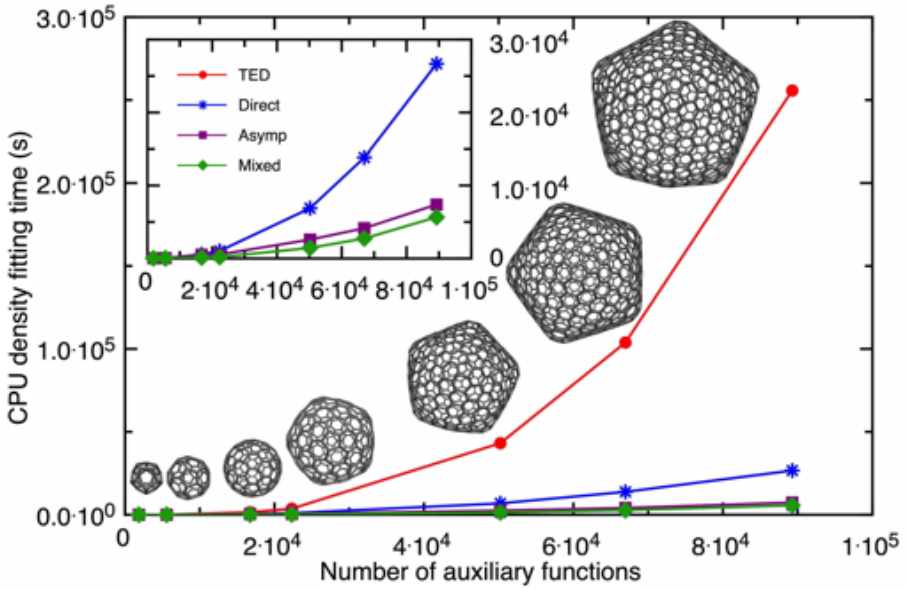

Figure 5.1: CPU time for density fitting as a function of the number of auxiliary functions for the fullerenes. inset depicts different MINRES approaches. To guide the eye the individual data points are connected by lines. See text for further details.

The quadratic scaling of MINRES can be further reduced with the ASYMPTOTIC (purple squares in the inset of Figure 5.1) and MIXED (green diamonds in the inset of Figure 5.1) MINRES versions. For the largest system, C_{960}, the density fitting time for the ASYMPTOTIC and MIXED MINRES is roughly 3 and 5 times faster than for the DIRECT MINRES as Table 5.1 shows. In comparison to the standard TED implementation speed up factors of more than 10, 35 and 45 are obtained with the DIRECT, ASYMPTOTIC and MIXED MINRES versions, respectively. This dramatic reduction in the density fitting timings also affects significantly the single point SCF timings as Table 5.1 shows. The ASYMPTOTIC and MIXED single point SCF times for C_{960} are by more than one order of magnitude smaller than the corresponding TED timing. Moreover, the density fitting with the MIXED MINRES approach requires roughly

25% of the total single point SCF timing whereas the corresponding TED fraction is over 90%. An interesting detail in the density fitting timings of Table 5.1 is found for C_{20} and C_{60}. Here MINRES DIRECT is slightly faster than MINRES ASYMPTOTIC. Although the differences are only 15 and 10 s for C_{20} and C_{60}. This behaviour is not expected at first glance. The reason lies in the nonexistence of far-field ERIs in C_{20} and C_{60}. Thus, the observed difference reflects the overhead due to the near-field vs. far-field ERI decision process in the ASYMPTOTIC (and MIXED) approach that is not needed for MINRES DIRECT. Finally, it is also important to mention that the converged TED and MINRES total energies differ slightly according to the number of truncated eigenvalues, N_{TE} in Table 5.1, in TED. For the here studied fullerenes this difference is a few μHartree for the smaller systems and increases to around 1 mHartree for the largest one (see Table E.1 in appendix E for energy differences and number of SCF cycles).

The enormous improvement of the computational performance employing the iterative MINRES algorithm for solving the fitting equation system is also rooted in its good parallelization. To analyze this in more detail we have studied the scaling of the MINRES density fitting for the C_{540} fullerene with respect to the number of cores in parallel runs. To this end we performed single point PBE/DZVP/GEN-A2* energy calculations for this fullerene using 12, 24, 48 and 96 Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20GHz cores with 4 GB of RAM per core. The obtained scaling of the reciprocal CPU time with respect to the number of cores is depicted in Figure 5.2 (blue stars) and compared with the corresponding ideal scaling (red dots). As this figure shows, the monotonically increasing trend of the MINRES acceleration by parallelization is close to the ideal scaling (red dots).

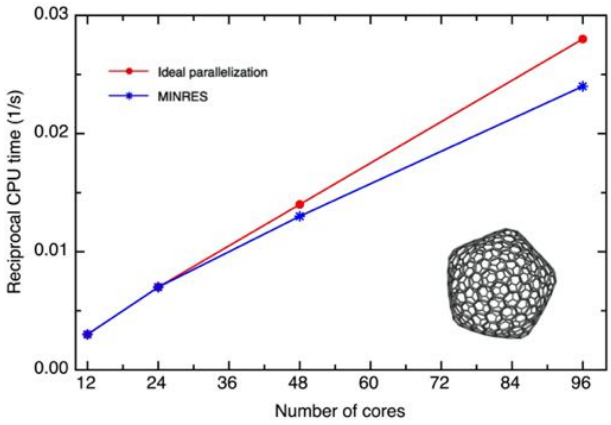

Figure 5.2: Reciprocal CPU time as a function of the number of cores for the C_{540} fullerene. The data points are connected to guide the eye.

This graph shows the good parallelization for the implemented MINRES algorithm due to its matrix-vector product based algorithmic structure.

5.1.2 Water clusters

To study the performance of MINRES density fitting for assemblies of molecules we run benchmark calculations on water clusters. Thus, our second test set consists of 6 water clusters containing 50 to 500 water molecules. Again we performed single point energy calculations with the already discussed methodology, i.e. PBE/DZVP/GEN-A2*. Table 5.2 lists for this water cluster benchmark set the total density fitting and single point SCF timings (in seconds) for the TED as well as the MINRES DIRECT, ASYMPTOTIC and MIXED approaches.

Table 5.2: Number of auxiliary functions, N_{aux}, CPU timings for the density fitting and single point SCF energy calculations using the listed density fitting approaches as well as the number of truncated eigenvalues in TED, N_{TE}, for the water cluster benchmark set. All calculations were performed with the PBE/DZVP/GEN-A2* methodology on a 24 core board.

System	N_{aux}	Total density fitting timings [s]				Total SCF timings [s]				N_{TE}
		TED	DIRECT	ASYMPTOTIC	MIXED	TED	DIRECT	ASYMPTOTIC	MIXED	
$(H_2O)_{50}$	6850	127.8	187.9	144.6	54.2	200.7	260.8	217.0	126.1	100
$(H_2O)_{100}$	13700	959.0	957.6	480.7	203.1	1227.3	1240.6	830.9	484.2	200
$(H_2O)_{200}$	27400	7340.7	4048.6	1426.0	1003.1	8694.9	6695.8	4112.9	3640.2	401
$(H_2O)_{300}$	41100	24651.1	10081.6	3187.9	2519.3	28194.8	17676.3	10667.0	9648.0	601
$(H_2O)_{400}$	54800	59771.5	19634.9	5196.4	5025.8	67029.2	35611.9	22079.7	21422.6	803
$(H_2O)_{500}$	68500	118926.1	32414.4	8266.8	8260.9	132543.7	63514.7	37311.7	36637.5	1002

The fitting timings are plotted in Figure 5.3 with respect to the number of auxiliary functions, N_{aux}, given in the second column of Table 5.2. As Table 5.2 shows the MINRES MIXED density fitting CPU times are always smaller than the corresponding TED timings. The reason is the larger number of auxiliary functions in the smallest water cluster, $(H_2O)_{50}$, compared to the smallest fullerene, C_{20}. Thus, the crossover between the TED and MINRES MIXED density fitting is already passed by the smallest water cluster in Table 5.2.

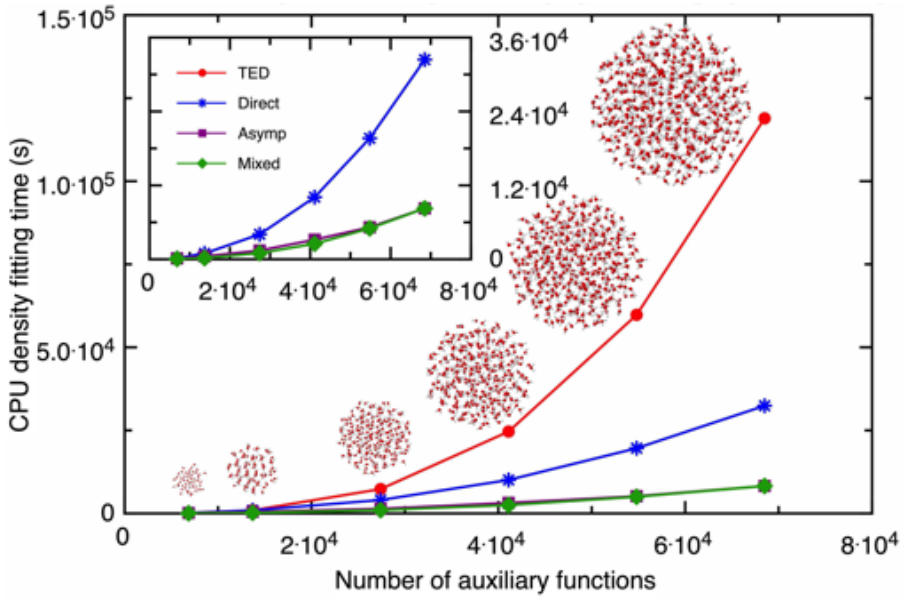

Figure 5.3: CPU time for density fitting as a function of the number of auxiliary functions for the water clusters. inset depicts different MINRES approaches. To guide the eye the individual data points are connected by lines. See text for further details.

The TED timings (red dots in Figure 5.3) show again the expected cubic scaling. The density fitting CPU times using MINRES DIRECT given in column 4 of Table 5.2 are dominated by the matrix-vector product. Their scaling of $N_{aux}^{2.2}$ with respect to the number of auxiliary functions is best observed from the inset of Figure 5.3 (blue stars). Also, the scaling reduction by the ASYMPTOTIC (purple squares) and MIXED (green diamonds) MINRES versions is clearly visible from this figure inset. The difference between these two versions is for the water clusters smaller than for the fullerenes. In fact, for the largest water cluster this difference becomes negligible. Again the speed-up of the density fitting by MINRES accelerates the corresponding single point SCF calculations, albeit with a more moderate factor of around 4 for the largest water cluster, $(H_2O)_{500}$. Because all water clusters in Table 5.2 possess far-field ERIs the ASYMPTOTIC and MIXED MINRES fittings are always faster than their DIRECT counterpart.

The comparison of the fullerenes and water cluster results (Table 5.1 and 5.2) reveals some interesting details. Although, the number of auxiliary functions of C_{20} is considerably smaller than for the $(H_2O)_{50}$ cluster, C_{20} has more truncated eigenvalues. This shows that the altering of the auxiliary function set by TED is system dependent and in this particular case very different for a molecule in comparison to an assembly of molecules. The assembly structure of the water

clusters is also the reason for the almost perfect correlation between N_{TE} and cluster size in Table 5.2. This indicates that the system type has a profound impact on the structure of the **G** matrix. Despite these very different **G** matrix structures, MINRES reliably reduces the scaling of the density fitting (see Figures 5.1 and 5.3) and, therefore, accelerates single point SCF calculations significantly.

The here obtained results for molecular assemblies are of particular interest for the future development of a truly linear scaling quantum chemistry methodology on the basis of ADFT. By eliminating the linear algebra bottlenecks associated to density fitting only the transformation and diagonalization of the Kohn-Sham matrix remain as higher order scaling steps in Roothaan-Hall ADFT SCF iterations. To overcome this bottleneck the pseudodiagonalization of the Kohn-Sham matrix can be used. In this approach SCF convergence is reached by annihilation of the occupied-virtual blocks of the Kohn-Sham matrix in MO representation[172]. To make the method linear scaling MO localization must be enforced. Thus, molecular assemblies with their natural local electronic structure are ideal test systems for such an approach. In fact, the development of an ADFT SCF based on MINRES density fitting and Kohn-Sham matrix pseudodiagonalization is a straightforward extension of the work developed in this thesis.

5.1.3 Zeolite structures

Besides its superior computational performance MINRES also reduces substantially the memory demand for the density fitting. To study this in more detail we performed single point energy calculations of hydrogen saturated MFI zeolite cutouts consisting of 1 to 4 unit cells[130]. The total density fitting timings for the TED and MINRES alongside with the memory demand per core and the number of truncated eigenvalues, N_{TE}, are listed in Table 5.3.

Table 5.3: Number of auxiliary functions, N_{aux}, CPU timings and RAM demand per core for the listed density fitting approaches as well as the number of truncated eigenvalues in TED, N_{TE}, for the hydrogen saturated MFI zeolite cutouts. The calculations were performed with the PBE/DZVP/GEN-A2* methodology employing 96 cores.

System	N_{aux}	Total density fitting timings [s]				RAM per core [GB]				N_{TE}
		TED	DIRECT	ASYMPTOTIC	MIXED	TED	DIRECT	ASYMPTOTIC	MIXED	
MFI-1	33608	3613.8	3999.3	1790.8	988.7	0.13	0.04	0.04	0.07	736
MFI-2	66512	26610.6	18194.4	5714.7	4818.8	0.52	0.08	0.08	0.15	1575
MFI-3	99416	89800.1	47012.6	12551.5	13455.6	1.15	0.12	0.12	0.24	2415
MFI-4	132320	207549.0	90143.7	20769.0	26196.7	2.03	0.16	0.16	0.32	3256

The total CPU density fitting timings as a function of the number of auxiliary functions, N_{aux}, are graphically illustrated in Figure 5.4.

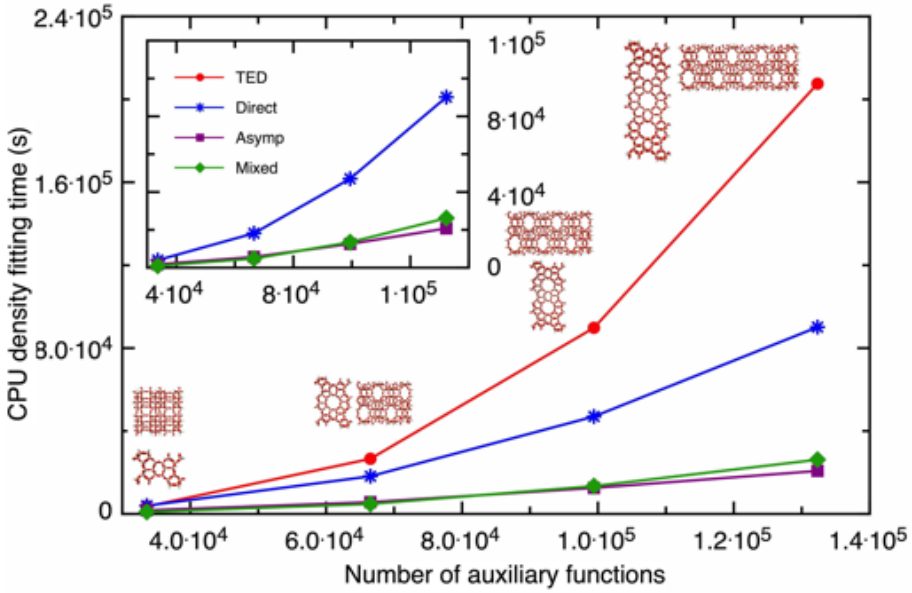

Figure 5.4: CPU time for density fitting as a function of the number of auxiliary functions for the zeolite cutouts. inset depicts different MINRES approaches. To guide the eye the individual data points are connected by lines. See text for further details.

As Table 5.3 and Figure 5.4 show the density fitting timings for the MFI zeolite cutouts are also considerably reduced by the MINRES approaches compared to TED. Qualitatively, a similar picture as for the fullerenes and water clusters was obtained. A remarkable difference, however, is observed for the density fitting with the ASYMPTOTIC and MIXED MINRES approaches. Whereas in all other systems the MIXED MINRES is always faster than the ASYMPTOTIC one, we find for the MFI cutouts a crossover at the trimer (MFI-3 in Table 5.3) with around 100,000 auxiliary functions. A more detailed analysis of this behaviour revealed that on the here used hardware architecture the direct calculation of two-center ERIs in the ASYMPTOTIC MINRES accesses mainly CPU cache which is faster than RAM access. Thus, the direct calculation of two-center ERIs becomes faster than their loading from RAM. The CPU timings for the single point SCF energy calculations are also reduced by MINRES compared to TED for the MFI zeolite cutouts. For the largest system, MFI-4, speed up factors of 2 and 4 for the single point energy calculations are obtained with MINRES DIRECT and MINRES ASYMPTOTIC or MIXED, respectively.

Table 5.3 also compares the RAM demand for TED and MINRES density fitting. Because our MINRES implementation requires only the storage of vectors rather than matrices, the RAM demand is considerably reduced compared to TED. This holds even for MINRES MIXED that stores near-field two-center ERIs. By construction, the RAM demand for MINRES DIRECT and ASYMPTOTIC is identical and scales linear with the number of auxiliary functions as Table 5.3 shows. This allows MINRES density fitting with hundreds of thousands to millions of auxiliary functions on current computer architectures.

5.1.4 Alkanes

To investigate the MINRES performance also for larger basis sets we performed 24 core parallel benchmark calculations on n-alkane chains ranging from $C_{100}H_{202}$ to $C_{300}H_{602}$ employing aug-cc-pVTZ basis sets in combination with GEN-A2* auxiliary function sets. As Figure 5.5 shows the crossover between MINRES and TED is at the $C_{150}H_{302}$ n-alkane with around 16,000 basis and 27,000 auxiliary functions. For these systems ASYMPTOTIC MINRES scales clearly sub-quadratic, $N^{1.4}$, with respect to the number of auxiliary functions.

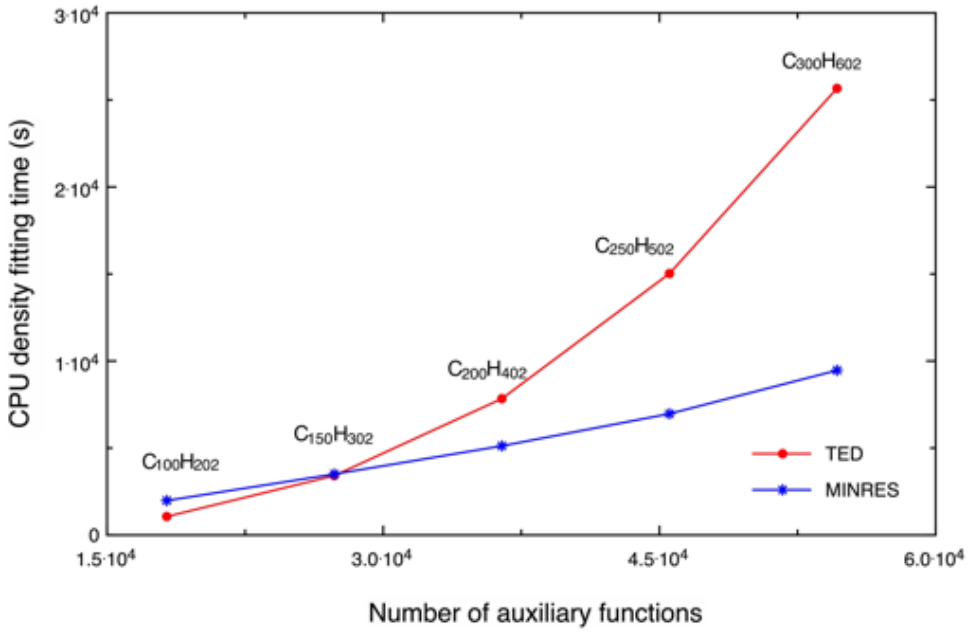

Figure 5.5: CPU density fitting timings versus the number of auxiliary functions for n-alkane chains employing the PBE/aug-cc-pVTZ/GEN-A2* level of theory using 24 cores. MINRES timings refer to the ASYMPTOTIC approach for the calculation of the two-center ERIs.

In conclusion, the implementation of the Krylov subspace MINRES algorithm for the solution

of the inhomogeneous fitting equation system in deMon2k has removed the computational bottleneck for the variational fitting of the Coulomb potential. Because MINRES also resolves the numerical problems of the density fitting arising from finite precision arithmetic a robust and reliable fitting implementation is achieved. The comparison of the different MINRES variants, DIRECT, ASYMPTOTIC and MIXED, showed that MINRES ASYMPTOTIC is the method of choice. It combines superbly computational performance with a low memory demand. As a result, subquadratic scaling with respect to the number of auxiliary functions in combination with excellent parallel scalability is achieved. The key to success is the use of the blocked atomic preconditioners which improve significantly the MINRES convergence. All our larger test calculations showed improved computational performance with respect to the TED implementation. Even for systems with over 100,000 auxiliary functions the CPU time for the density fitting with MINRES was negligible compared to the total time for the single point energy calculation.

5.2 Structure optimizations

After we have analyzed the performance of MINRES in single point SCF energy calculations, we now turn to structure optimizations. The fundamental difference between TED and MINRES in structure optimization is the $\mathbf{G}^{-1}$ matrix usage for the calculation. For TED, this matrix is calculated at the beginning of SCF procedure and it is updated by a quasi-Newton method throughout all optimization steps. On the other hand, MINRES solves iteratively the associated equation system every time this is needed in each optimization step. To further discuss this, we optimized a *heme* (porphyrin with an iron ion acting as a tetradentate ligand) organic system with 48 atoms using 24 cores. Using the PBE/DZVP/GEN-A2* level of theory this translates to 561 orbitals and 3,400 auxiliary functions. The optimization thresholds were the default ones in deMon2k and the optimization was performed using internal redundant coordinates. Although this is a small molecule the use of TED fitting fails in the optimization due to the failed convergence in the solution of the fitting equation, Eq. (4.1), at optimization step 27. Analyzing the TED in more detail, we found that 191 eigenvalues were truncated from the $\mathbf{G}^{-1}$ matrix at the beginning of the optimization. In Figure 5.6 the structure is depicted alongside with the energies (in a.u., red dots) and maximum absolute gradient components (MAG, in a.u./Bohr, blue stars) at each optimization step.

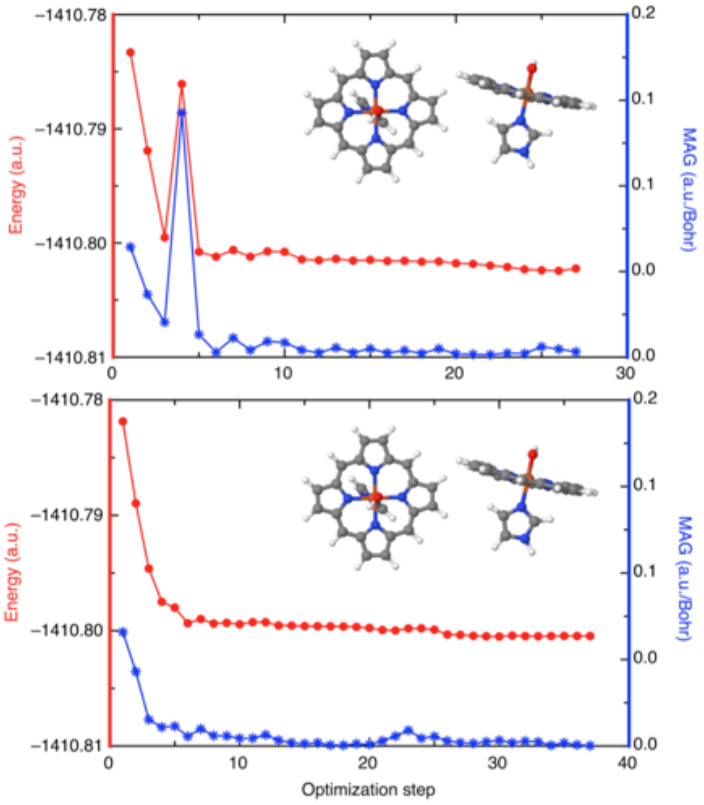

Figure 5.6: Heme structure optimization step energies (red dots) and maximum absolute gradient components (blue stars) vs the number of optimization steps from a failed structure optimization using TED (top) and a successful structure optimization using MINRES (bottom), respectively. The optimizations were performed using the PBE/DZVP/GEN-A2* methodology. The points are connected to guide the eye.

As can be seen from Figure 5.6 (top) at step 4 of the optimization with TED fitting, a peak in the energy as well as in the maximum gradient component occurs. In the corresponding MINRES optimization this peak is not occurring. This already demonstrates that the two optimization paths differ in an early stage. The breaking of the TED optimization is mainly due to the use of the same $\mathbf{G}^{-1}$ obtained at the beginning of the SCF for the initial structure. In more detail, the $\mathbf{G}$ matrix is geometry dependent, therefore, $\mathbf{G}$ changes during structure optimization, giving rise to a different eigenvalue spectrum when using TED fitting. This eigenvalue spectrum inside the TED can give a different number of truncated eigenvalues, N_{TE}, which can have less or more truncated eigenvalues. Analyzing in detail optimization step 27 using TED fitting, we found 186 truncated eigenvalues for this structure, whereas the initial structure had 191 truncated eigenvalues. On the other hand, if the optimization is carried out using the MINRES algorithm, as

shown in Figure 5.6 (bottom), it yields a smoother optimization convergence. Thus, the use of the correct **G** matrix for each geometry point on the PES yields a smoother optimization profile.

Taking advantage of the good parallel scaling of MINRES, we tested its performance for the structure optimization for larger molecules. Our first example is the optimization of the C_{720} fullerene from the test set in Section 5.1.1 using the same PBE/DZVP/GEN-A2* level of theory, and the default optimization thresholds of deMon2k. The optimization was again performed in internal redundant coordinates. For the optimization of this fullerene, 160 Intel(R) Xeon(R) CPU E5-2667 v4 @ 3.20GHz cores with 3 GB in RAM per core were used. This system has 10,800 basis functions and 66,960 auxiliary functions which makes the structural optimization quite a challenge. The molecular structure alongside with the relative energies in a.u. (red squares) and MAG in a.u./Bohr (blue stars) are depicted in Figure 5.7. From this Figure, a smooth convergence for the optimization of this system can be seen. At the step 68 the structure was optimized.

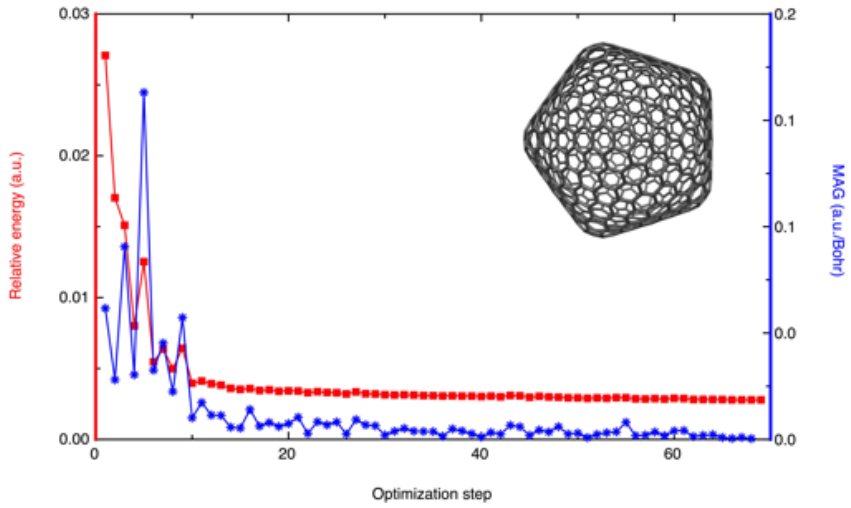

Figure 5.7: C_{720} fullerene structure optimization relative energies (red squares) and maximum absolute gradient component (blue stars) vs the number of optimization steps from a successful structure optimization using the MINRES approach. The PBE/DZVP/GEN-A2* methodology was used. The points are connected to guide the eye.

Because of the excellent performance in the optimization for large systems with large auxiliary function sets we took a third system to test the MINRES algorithm in optimizations. An oligothiophene with 100 thiophene units (Figure 5.8) was optimized using the VWN/DZVP/GEN-A2* level of theory and a tolerance of 10^{-4} a.u. in the root mean squared forces optimization thresh-

old. This kind of system is of great technological interest due to its importance in semiconductor electronics and data processing[173–176]. In this respect, oligothiophenes represent a particular attractive group of compounds because of their stability in air and easy handling in several organic solvents[177,178]. Moreover, they can be rather straightforward chemically altered by many standard reactions for heteroaromatics[179]. Therefore, oligothiophenes have attracted considerable attention in materials science as part of thin-film organic transistors and organic photovoltaic cells[180,181]. Some other properties related to these systems and their unusual electrical and optical behaviour have been intensively investigated in the past few years, too. Of special interest are electroluminescence[182–184], photoconductivity and electrical conductivity[185–187].

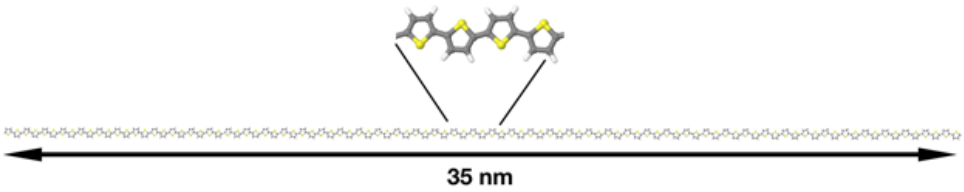

Figure 5.8: Optimized planar oligothiophene containing 100 thiophene units. The optimization was performed in Cartesian coordinates. White, grey and yellow colors are hydrogen, carbon and sulfur atoms, respectively.

All this shows that large oligothiophenes are of considerable interest to nanoscience. In order to test our MINRES implementation for structure optimization we optimized a 100 thiophene units containing oligothiophene. This molecule contains 702 atoms and its length is roughly 35 nm. The optimizations were carried out without any symmetry constraints but using two kinds of coordinated systems for the optimizer. In particular, we used Cartesian and internal redundant coordinates for these optimizations. In both cases the initial start structure was assumed planar. This planarity was conserved in the Cartesian optimization as the corresponding optimized structure in Figure 5.8 shows. On the other hand, the optimization of the oligothiophene in internal redundant coordinates converged to a curved structure as shown in Figure 5.9. The energy difference between these two structures is only 0.4 kcal/mol with the planar structure being lower in energy. At this stage of this study is still unclear if both optimized structures are indeed minima. This demonstrates the challenge of nanosystem structure optimization in terms of accuracy and diversity of the potential energy surface. It also underlines the need for corresponding frequency analyses if true minimum structures are sought.

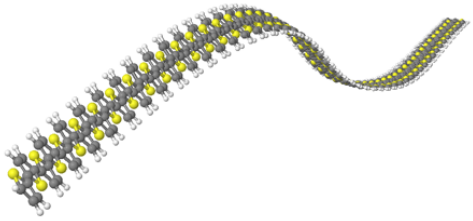

Figure 5.9: Optimized curved oligothiophene containing 100 thiophene units. The optimization was performed using internal redundant coordinates. White, grey and yellow colors are hydrogen, carbon and sulfur atoms, respectively.

5.3 Polarizabilities

In order to test the MINRES implementation for response property calculations with the ADPT approach, we performed static and dynamic polarizability calculations for a series of molecules. Experimental structures were used in this validation. The average polarizability and anisotropic polarizability are obtained by the following formulas, calculated in the principal axes system of the polarizability tensor.

$$\bar{\alpha}(\omega) = \frac{1}{3}(\alpha_{xx}(\omega) + \alpha_{yy}(\omega) + \alpha_{xx}(\omega)) \tag{5.1}$$

and

$$|\Delta\alpha(\omega)|^2 = \frac{1}{2}[(\alpha_{xx}(\omega) - \alpha_{yy}(\omega))^2 + (\alpha_{xx}(\omega) - \alpha_{zz}(\omega))^2 + (\alpha_{yy}(\omega) - \alpha_{zz}(\omega))^2]. \tag{5.2}$$

The static and dynamical polarizabilities, α, and anisotropic polarizabilities, $\Delta\alpha$, of these systems were already obtained with TED density fitting by Carmona[188]. Therefore, we used these data as a theoretical reference for our MINRES calculations in Table 5.4 and 5.5. The level of theory for these calculations was PBE/TZVP-FIP1[189,190] /GEN-A2*. It is worth of noticing that the TED calculations that Carmona presented in his thesis were not performed using the EN2 algorithm. Instead, they refer to direct analytic solutions of the ADPT response equation system, Eq. (4.90), with an explicit TED of the response matrix.

In Table 5.4 the static average polarizabilities, $\bar{\alpha}$, and the anisotropic polarizabilities, $|\Delta\bar{\alpha}|$, in a.u. are compared between the TED approach and the EN2 + MINRES solver of this work. As reference, we also list corresponding experimental results in the last two columns of Table 5.4.

Table 5.4: Comparison of static GGA ADPT polarizabilities and polarizability anisotropies in a.u. of small molecules. The theoretical results were obtained with MINRES and TED density fitting using PBE/TZVP-FIP1/GEN-A2* level of theory. For all calculations the experimental structures were used.

	This work		Ref. [188]		Exp.	
Molecule	$\bar{\alpha}$	$\lvert\Delta\bar{\alpha}\rvert$	$\bar{\alpha}$	$\lvert\Delta\bar{\alpha}\rvert$	$\bar{\alpha}$	$\lvert\Delta\bar{\alpha}\rvert$
HF	5.88	1.20	5.88	1.20	5.40[b]	1.35[c]
CH_4	17.10		17.12		17.27[d]	
C_2H_2	23.15	12.80	23.28	12.74	22.68[d]	11.83[e]
CH_3F	17.31	1.30[a]	17.51	1.50[a]	17.32[d]	1.41[f]
HCl	17.64[a]	2.09	17.59[a]	2.11	17.54[g]	1.47[h]
H_2S	24.27	0.32[a]	24.28	0.24[a]	24.66[i]	0.67[j]
CH_2F_2	18.40	1.99	18.40	2.00	18.20[d]	1.70[f]
OCS	33.86	25.41	33.91	25.46	34.33[d]	26.26[e]
SO_2	25.66	13.41	25.63	13.32	25.49[k]	12.98[k]
CHF_3	19.63	1.45[a]	19.59	1.29[a]	18.69[d]	1.46[f]
CF_4	19.92		19.95		19.53[d]	
CS_2	53.98	55.45	53.93	53.34	55.38[d]	57.38[e]
MAE	2.2%	14.9%	2.3%	17.4%		

[a] Calculated dynamic values at $\lambda = 632.8$ nm. [b] Static value from refractive index dispersion [191]. [c] Static value from molecular beam electric resonance [192]. [d] Static value from refractive index dispersion [193]. [e] Deduced from static estimates [194]. [f] Dynamic values ($\lambda = 632.8$ nm) from Ref [195]. [g] Depolarized light scattering at experimental $\lambda = 632.8$ nm [196]. [h] Static value from molecular beam electric resonance [193]. [i] Extrapolated static value from dispersion of dynamic mean polarizability [197]. [j] Dynamic value from Kerr effect at ($\lambda = 632.8$ nm) [198]. [k] Static value from refractive index and Rayleigh scattering dispersion [197].

From Table 5.4 it is important to note that the mean absolute error (MAE) in percentage, of the two different theoretical approaches compared to experimental data is in the same range of 2% for the static polarizability. Therefore, our implementation gives a good agreement with the experiment and slightly improves the anisotropic polarizabilities compared to TED density fitting, lowering the corresponding MAE from 17.4% to 14.9%.

Furthermore, in Table 5.5 the dynamical polarizabilities for another test set are compared. The

calculations were performed with the same methodology, i.e. PBE/TZVP-FIP1/GEN-A2*. The table compares the EN2 + MINRES dynamic polarizabilities and anisotropic polarizabilites of this work with those from TED density fitting and experimental references.

Table 5.5: Comparison of dynamic GGA ADPT polarizabilities and polarizability anisotropies in a.u. for small molecules. The theoretical results were obtained using the PBE/TZVP-FIP1/GEN-A2* level of theory with MINRES and TED density fitting employing the experimental wavelengths. For all calculations the experimental structures were used.

	This Work		Ref. [188]		Exp.	
Molecule	$\bar{\alpha}$	$\lvert\Delta\bar{\alpha}\rvert$	$\bar{\alpha}$	$\lvert\Delta\bar{\alpha}\rvert$	$\bar{\alpha}$	$\lvert\Delta\bar{\alpha}\rvert$
NH_3	15.07	2.55	15.05	2.48	14.98[a]	1.94[a]
H_2O	10.40	0.19	10.38	0.25	9.92[b]	0.66[b]
N_2	11.80	5.24	11.90	5.31	11.92[a]	4.70[a]
CO	13.59	3.56	13.62	3.56	13.34[a]	3.59[a]
NO	11.83	5.81	12.00	5.79	11.74[a]	5.70[a]
O_2	10.58	7.30	10.64	7.27	10.78[a]	7.42[a]
N_2O	19.77	19.50	19.78	19.53	20.24[a]	19.97[a]
CO_2	17.72	13.77	17.74	13.83	17.75[a]	14.17[a]
Cl_2	30.54	15.68	30.53	15.70	31.11[a]	17.54[a]
C_2H_4	28.73	13.01	28.72	13.10	28.48[a]	12.21[c]
C_2H_6	29.99	4.67	30.01	4.67	30.16[a]	5.20[a]
C_6H_6	72.17	40.79	72.27	40.56	70.18[a]	37.93[a]
MAE	1.6%	13.2%	1.6%	12.3%		

[a]Depolarized light scattering at $\lambda = 632.8$ nm [196]. [b]Depolarization ratio from Rayleigh scattering at $\lambda = 514.5$ nm [199]. [c] Depolarized light scattering at $\lambda = 632.8$ nm [197].

In the last row of Table 5.5 the MAE in percentage is given. In this case, the MAE for the dynamic polarizability is 1.6% compared to experiment whereas the anisotropic dynamical polarizabilities are in the range of 13% and 12% using EN2 + MINRES and TED, respectively. Thus, the MINRES anisotropic dynamical polarizabilities show a slightly larger deviation from experiment than those obtained with the TED density fitting. We attribute this to the TED of the

response matrix **R** used by Carmona[188]. This can hide pathological large Coulomb response matrix elements due to too small energy gaps between occupied and unoccupied GGA MO manifolds[164], see Eq. (4.86). These large Coulomb response matrix contributions can make the **R** matrix ill-conditioned. Whereas in TED the corresponding small absolute eigenvalues of **R** are eliminated, they persists inside the MINRES solution.

To conclude our validation of the new EN2 + MINRES implementation for polarizability calculations, we present in Table 5.6 hyperpolarizability results for some small molecules. Again, we compare with results from Carmona[188] and corresponding experimental data. For these calculations the PBE/TZVP-FIP1/GEN-A2* level of theory was used. The listed dynamical polarizabilities in Table 5.6 refer to the second harmonic generation (SHG) with $\omega_1 = \omega_2 = \omega$. To facilitate direct comparison with experimental data we performed all calculations in the so-called EFISH (electric-field-induced second harmonic) orientation. In this orientation the permanent dipole moment of the molecule is aligned along the z component of the external field. The average hyperpolarizability is calculated by the following expression

$$\bar{\beta}(\omega_3;\omega_1,\omega_2) \equiv \bar{\beta}(\omega) = \frac{1}{5}\sum_{i}^{x,y,z}\left[\beta_{zii}(\omega) + \beta_{izi}(\omega) + \beta_{iiz}(\omega)\right]. \tag{5.3}$$

Table 5.6: Comparison of static and dynamic GGA ADPT average hyperpolarizabilities, $\bar{\beta}$ (in a.u.), of small molecules using MINRES and TED density fitting. The theoretical results were obtained using PBE/TZVP-FIP1/GEN-A2* level of theory.

	This Work		Ref. [188]		Exp.
Molecule	$\bar{\beta}$	$\bar{\beta}(\omega)$	$\bar{\beta}$	$\bar{\beta}(\omega)$	$\bar{\beta}(\omega)$
CH_3OH	-46.44	-60.34	-46.75	-61.05	-35.0± 2.1[a]
CH_3F	-55.57	-67.06	-57.54	-69.19	-57.0± 4.2[b]
CH_3Cl	-4.21	-6.11	-3.06	-4.44	13.3± 1.4[c]
CHF_3	-35.84	-35.58	-31.34	-36.09	-25.2± 0.9[b]
CF_3Cl	-81.11	-86.43	-71.60	-82.55	-69.2 ± 2.8[b]
$CHCl_3$	-12.06	-16.65	-13.76	-18.20	1.2 ± 2.6[c]
$CFCl_3$	-39.04	-36.32	-36.35	-37.44	-30.9 ± 9.6[c]
MAE	207%	275%	194%	258%	

[a]Dynamic value at $\lambda = 694.3$ nm [200]. [b]Dynamic value at $\lambda = 694.3$ nm [201]. [c]Dynamic value at $\lambda = 694.3$ nm [202].

The dynamical polarizabilities in Table 5.6 were calculated at $\omega = 694.3$ nm. All calcula-

tions used analytically evaluated exchange-correlation kernels[203] and finite difference kernel derivatives[204]. As can be seen from Table 5.6, the calculated hyperpolarizabilities show much larger deviations from the experimental data than the previously discussed polarizabilities. The main reason are the previously described pathologies in the GGA functional that at the current level of knowledge can only be overcome by the use of hybrid functionals[170]. Therefore, we discourage GGA hyperpolarizability calculations.

5.3.1 Fullerene and Amylose Polarizabilities

As has been seen, MINRES density fitting is well suited for response property calculations of small systems. Because we have optimized the full set of fullerenes from section 5 using the PBE/DZVP/GEN-A2* level of theory, we decided to calculate their polarizabilities. We can directly compare with the published results from Mejía *et al.*[163] where the VWN/DZVP/GEN-A2 level of theory was used. Between these two methodologies there is a huge difference in terms of the number of auxiliary functions used. For the biggest fullerene C_{960}, Mejía *et al.* reported 14,400 basis functions and more than 32,000 auxiliary functions. However, with our methodology the same fullerene has the same number of basis functions whereas the auxiliary functions dimension grows to 89,280. Undoubtedly, this is a challenging task! Figure 5.10 compares the static DZVP polarizabilities per atom from Mejía *et al.*[163] using VWN/DZVP/GEN-A2 (red dots) with our PBE/DZVP/GEN-A2* results (blue stars). As this graphic reveals both calculations show similar trends The main difference is a slightly higher polarizability per atom with PBE/GEN-A2*. We attribute this mainly to the extended auxiliary function set used in our calculations. All polarizability calculations were performed with a residual tolerance in the EN2 algorithm of 5.0×10^{-6} and the default TOL for MINRES of 10^{-7} in the residual norm approximation.

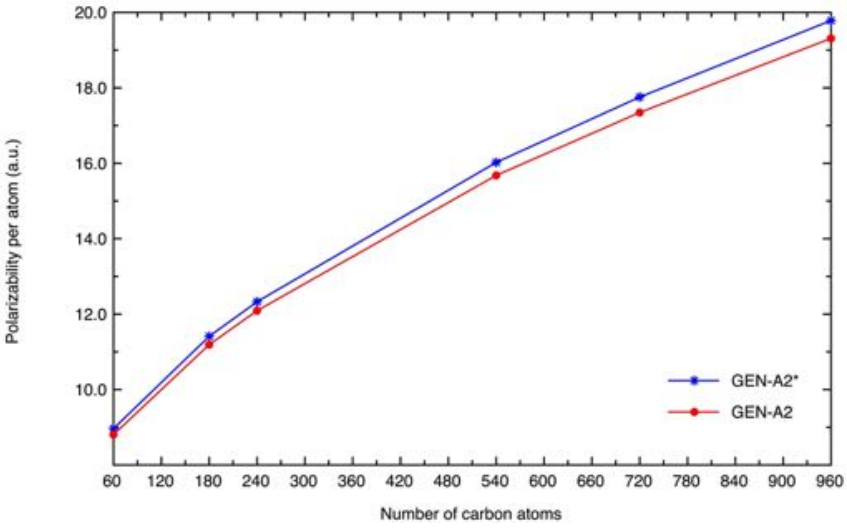

Figure 5.10: Polarizability per carbon atom for icosahedral fullerenes from C_{60} to C_{960} using PBE/DZVP/GEN-A2* and VWN/DZVP/GEN-A2 levels of theory. To guide the eye the individual data points are connected by lines.

In order to test large system polarizability calculations with MINRES density fitting employing augmented basis sets we calculated the average static polarizabilities of amylose systems with 8 to 64 units of α-D-glucose with the PW86[72]/EPR-III[205]/GEN-A2* methodology. The results were compared with the ones obtained with the current TED density fitting implementation. The structures of these compounds were built using the SWEET tool[206]. In fact, it is important to mention that the largest system shown here has over 1,300 atoms, 28,275 basis functions and more than 112,000 auxiliary functions. The geometry for this system is depicted in Figure 5.11.

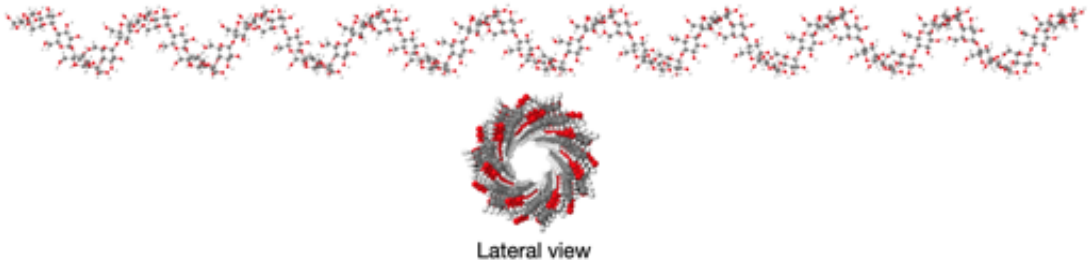

Figure 5.11: Amylose chain containing 64 α-D-glucose units.

In this case the PW86 functional is used because it is another GGA functional for which analytic kernel evaluation[203] is available in deMon2k. The obtained average static polarizability values for these systems are given in Table 5.7 and they are plotted as polarizability per atom in

Figure 5.12 for a direct comparison between MINRES and TED density fitting as function of the number of atoms, N_{atom}. Also in Table 5.7 the number of α-D-glucose units, the number of atoms N_{atom}, the number of auxiliary functions, N_{aux} and the number of basis functions, N_{bas}, for each amylose, are given. The red dots in Figure 5.12 correspond to the calculations performed with the current ADPT methodology using TED density fitting and the inverse of the Coulomb matrix as preconditioner for the EN2 algorithm. The blue stars correspond to the new EN2 + MINRES algorithm (see Figure 4.8) performed after a SCF calculation with MINRES density fitting. Thus, these calculations are free of any $\mathbf{G}^{-1}$ operation.

Table 5.7: Number of α-D-glucose units and atoms, N_{atom}, for each amylose system alongside with the number of auxiliary functions, N_{aux} as well as the number of basis functions, N_{bas}. The calculated average static polarizabilities (in a.u.) are obtained with the PW86/EPR-III/GEN-A2* methodology employing MINRES and TED density fitting.

Glucose units	N_{atom}	N_{aux}	N_{bas}	$\bar{\alpha}_{MINRES}$	$\bar{\alpha}_{TED}$
8	171	14168	3579	786.2	786.2
16	339	28152	7107	1570.4	1570.5
32	675	56120	14163	3140.6	-
48	1011	84088	21219	4730.0	-
64	1347	112056	28275	6307.6	-

From Table 5.7 it can be observed that the average static polarizabilities for the amyloses containing 8 and 16 α-D-glucose units are in excellent agreement between the two response property methodologies, MINRES (blue stars in Figure 5.12) and TED (red dots in Figure 5.12), respectively. However, the polarizabilities for the 32, 48 and 64 unit amyloses were impossible to calculate with the TED methodology, because of the memory demand needed by this density fitting approach. The polarizability calculations were performed with 96 Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20GHz cores with 4 GB of RAM per core. Our new MINRES implementation overcomes this problem as Table 5.7 shows.

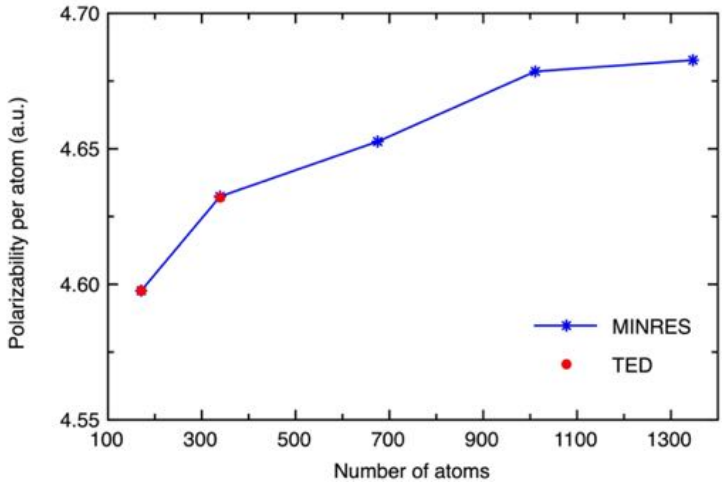

Figure 5.12: Comparison between MINRES (blue stars) and TED (red dots) calculated average static polarizabilities per atom vs the number of atoms, N_{atom}, for the amylose chains containing up to 64 α-D-glucose units. To guide the eye, the individual data points are connected by lines.

As Figure 5.12 shows, the polarizability per atom grows monotonically, although with a reduction of the slope with increasing system size. Compared to the fullerenes in Figure 5.10 the polarizability per atom of the amylose structures is by a factor of 3 to 4 smaller and increases much less with system size. We attribute this to a missing π-system delocalized over the full molecule. Moreover, the differences in the polarizability per atom are small compared to the ones for the fullerenes, thus, a strong cooperative effect is not observed for the amyloses. The largest amylose system is about 20 nm long and its diameter is around 1.2 nm. The here presented results demonstrate that first-principles response property calculations for nanostructures can be performed employing moderate computational architectures with the newly developed mathematical framework for density fitting.

5.3.2 DNA Polarizabilities

In order to test the possibility to apply the newly developed EN2 + MINRES algorithm for the solution of ADPT equation systems of microbiologically relevant systems, we calculated the polarizabilities of deoxyribonucleic acid (DNA) double strands. To this end, we studied 3 DNA fragments containing 4, 8 and 16 adenine-thymine base pairs, respectively. These structures are shown in Figures 5.13 as compound a, b and c, respectively. The structures were generated with the Accelrys Discovery Studio Client in the DNA B-strand conformation with both ends capped with hydroxyl groups. The phosphate groups were protonated in order to obtain neutral

systems.

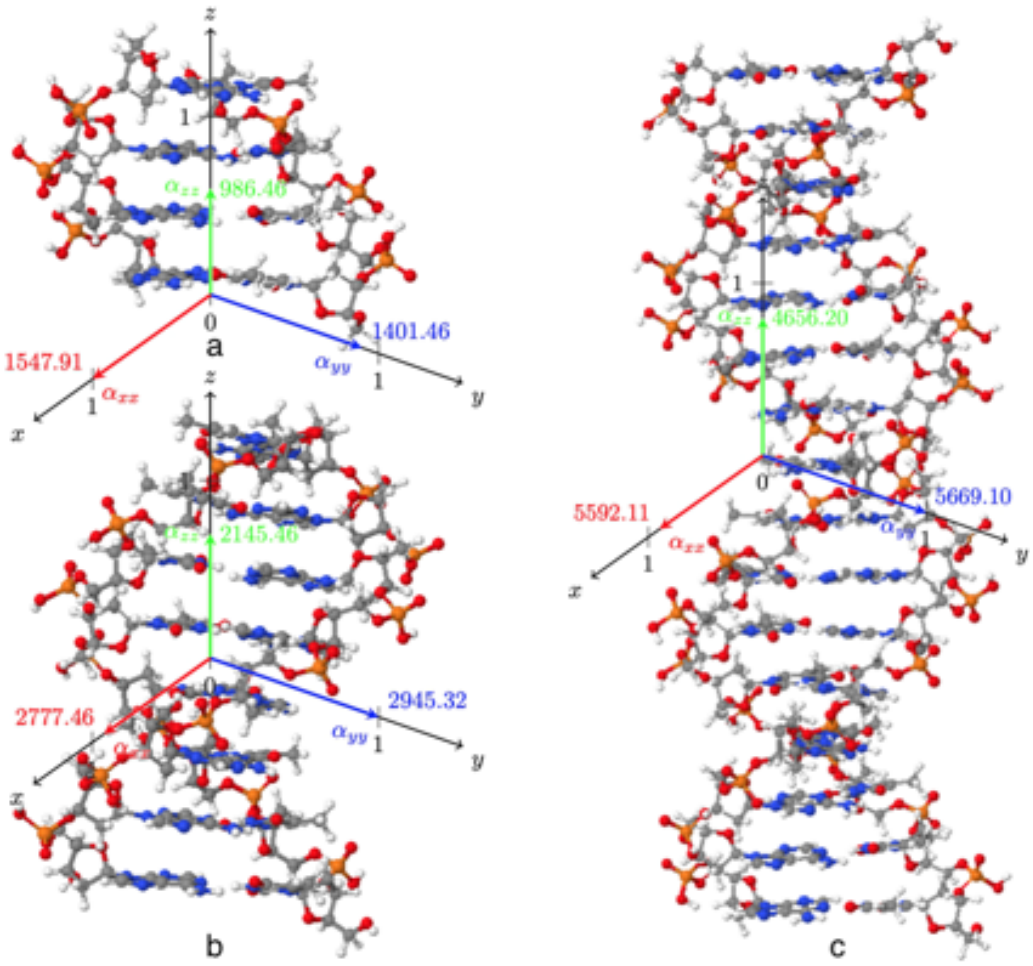

Figure 5.13: Ball and stick representation of the DNA double strand in the B-strand conformation with 4 (a), 8 (b) and 16 (c) adenine-thymine base pairs. The colors of the atoms are gray for carbon, blue for nitrogen, orange for phosphorus and white for hydrogens, respectively. For each structure the coordinated axes with their corresponding normalized polarizability tensor components are plotted.

Aiming to test MINRES density fitting also for LDA, the polarizabilities of these systems were calculated with the VWN/DZVP/GEN-A2* methodology using 96 Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20GHz cores with 4 GB of RAM per core. Table 5.8 lists the obtained average static polarizabilities with MINRES and TED density fitting alongside with the number of base pairs, the number of atoms, N_{atom}, the number of auxiliary functions, N_{aux}, and the number of basis functions, N_{bas}. The averaged polarizability per atom for these systems from the TED (red dots) and MINRES (blue stars) approaches are compared in Figure 5.14.

Table 5.8: Number of adenine-thymine base pairs and atoms, N_{atom}, for each DNA fragment alongside with the number of auxiliary functions, N_{aux}, as well as the number of basis functions, N_{bas}. The calculated average static polarizabilities in a.u. obtained with the VWN/DZVP/GEN-A2* methodology are compared using the MINRES and TED approach for density fitting.

Base pairs	N_{atom}	N_{aux}	N_{bas}	$\bar{\alpha}_{MINRES}$	$\bar{\alpha}_{TED}$
4	260	17208	2904	1326.7	1326.7
8	524	35020	5896	2659.6	-
16	1052	70644	11880	5305.8	-

From Table 5.8 it can be seen that the average static polarizabilities for the DNA fragment containing 4 base pairs are the same for response property calculations with MINRES and TED density fitting. However, once again the polarizabilities for the 8 and 16 base pair DNA fragments were impossible to calculate with the TED approach due to the memory demand needed for the Coulomb matrix inversion. On the other hand, with the MINRES approach these calculations were possible. In Figure 5.14 the red dot is the calculated polarizability per atom using the TED density fitting methodology. It is in perfect agreement to the one obtained with MINRES. As Table 5.8 shows, the polarizability increases linearly with the length of the DNA fragments. Interestingly, the polarizability per atom is larger for smaller fragments than for larger ones. As a result, a negative slope for the average polarizability per atom with respect to the DNA fragment size is observed in Figure 5.14. This result is in qualitative agreement with recent experimental reports from the literature[207,208]. In fact, it explains why DNA molecules do not behave as rod-shaped objects in the dielectrophoresis process and why their polarization is more related to the counter-ion cloud fluctuation than to their structure. In order to gain more insight, in Figure 5.13 the polarizability tensor components are depicted, too. These tensor components are normalized to their largest value. Looking at the tensor components from Figure 5.13, it is possible to see that the largest contribution corresponds to the α_{xx} component for the 4 units fragment, whereas for the 8 and 16 units the largest value switches to the α_{yy} component. Thus, the polarizability tensor does not reflect the geometrical anisotropy of the systems. Nevertheless, the average polarizability as an extensive quantity grows with the molecular size as Table 5.8 shows. In fact, polarizabilities of DNA systems have been recently studied extensively by several authors[209–217] using theoretical and experimental methodologies. In particular, a frequency-dependent calculation of the dielectric polarizability in solution, based on molecular dynamics simulation has been recently reported[218]. In here, the authors propose that the co- and counter-ion diffusion, in response to the ion concentration gradients associated with the deformation of the electric double layer (EDL), effectively reduces the polarization. Additionally,

Jung and coworkers explain that there is also a cross-over frequency where the polarizability is roughly independent of the ion diffusion coefficient in terms of the solvent quality. These studies have been reported in the literature a few years ago for this kind of system in aqueous salt solutions[219–221]. Although our polarizability calculations are size extensive, the changes in the polarizability per atom for the amyloses and DNA fragments are too small to draw definite conclusions about cooperative or anti-cooperative effects in these systems. Therefore, further studies that go beyond the scope of this thesis, namely to show the effect of MINRES density fitting in response property calculations, are needed.

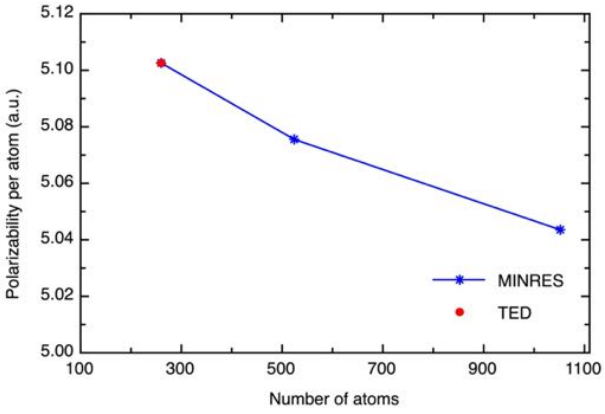

Figure 5.14: Comparison between MINRES (blue stars) and TED (red dots) calculated average static polarizabilities per atom in a.u. for the DNA fragments versus the number of atoms.

In any case, it is important to underline that the largest DNA strand, depicted in Figure 5.13c, has the molecular formula $C_{320}H_{402}O_{188}N_{112}P_{30}$ with a total of 1,052 atoms. This biological system contains 11,880 basis functions and more than 70,000 auxiliary functions. The system size is roughly 5.2 nm long, thus, lying at the nanoscale. Therefore, these results clearly show that the here developed scheme for density fitting can be applied efficiently to microbiologically relevant systems.

5.4 Nuclear spin-spin coupling constants

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool for the determination of the electronic structure and nuclear configuration of molecules. It is of paramount importance in areas such as biochemistry and organometallic chemistry as well as inorganic and organic chemistry as a whole. The chemical shifts (magnetic shieldings) and the indirect nuclear spin-spin coupling constants (NSSCCs) are the spectroscopic parameters measured in high resolution

NMR experiments. In many cases they provide sufficient information for an accurate structure assignment. However, with increasing system complexity the interpretation of these data can become complicated even for an expert spectroscopist. Undoubtedly, the reliable theoretical prediction of NMR parameters could significantly ease this interpretation work.

Because NSSCCs depend strongly on the electron density between the pair of coupling nuclear spins they provide information on the chemical bonds as well as of the coupling nuclei, complementing the information obtained from the chemical shifts. Unfortunately, the working equations for NSSCC calculations are more complicated as for the chemical shielding, which also increases the complexity in these calculations. Many years ago, Ramsey proposed[222–226] four individual mechanisms that are together responsible for the measured NSSCCs. These are the Fermi-contact (FC), spin-dipole (SD), paramagnetic spin-orbit (PSO) and diamagnetic spin-orbit (DSO) contributions. The FC and SD contributions arise from a spin polarization of the closed shell density, also named spin magnetization, due to a nuclear spin magnetic moment perturbation. The resulting magnetic field interacts with another nuclear spin magnetic moment which yields the coupling between these two nuclear spins. The PSO and DSO contributions arise from a current density in the electron cloud due to the presence of a nuclear spin moment. As a result, a change in the magnetic field associated to the electron distribution occurs which interacts with another nuclear magnetic spin moment. This constitutes the nuclear spin-spin coupling mechanism for the PSO and DSO terms.

For ADFT NSSCC calculations we augment the ADFT energy expression, Eq. (3.35), by the PSO, DSO, FC and SD terms. This yields

$$\begin{aligned} E = &\sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu}^{PSO+DSO} + \sum_{\mu,\nu} P_{\mu\nu}^{\alpha-\beta} H_{\mu\nu}^{FC+SD} \\ &+ \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu || \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} || \bar{l} \rangle + E_{xc}[\tilde{\rho}]. \end{aligned} \tag{5.4}$$

For a detailed derivation of this energy expression we refer the interested reader to [227].

In Eq. (5.4) $P_{\mu\nu}^{\alpha-\beta}$ denotes elements of the spin density matrix

$$P_{\mu\nu}^{\alpha-\beta} = P_{\mu\nu}^{\alpha} - P_{\mu\nu}^{\beta},$$

that arise from the spin polarization of the closed-shell density due to the FC and SD operators.

Differentiating the ADFT energy expression in Eq. (5.4) with respect to the Cartesian components of the nuclear magnetic dipole moments of nuclei P and Q yields the (reduced) nuclear spin-spin coupling tensor K_{PQ} with elements

$$K_{P_\lambda Q_\eta} = \frac{\partial^2 E}{\partial \mu_{P_\lambda} \partial \mu_{Q_\eta}}. \tag{5.5}$$

Here λ and η denote Cartesian components x, y and z. Evaluating this second energy derivative yields

$$\begin{aligned} K_{P_\lambda Q_\eta} = &\sum_{\mu,\nu} P_{\mu\nu}^{\alpha-\beta(\lambda),FC} H_{\mu\nu}^{(\eta),FC} + \sum_{\omega=x}^{z} \sum_{\mu,\nu} P_{\mu\nu,\omega}^{\alpha-\beta(\lambda),SD} H_{\mu\nu,\omega}^{(\eta),SD} + \\ &\sum_{\mu,\nu} P_{\mu\nu}^{(\lambda),PSO} H_{\mu\nu}^{(\eta),PSO} + \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu}^{(\lambda\eta),DSO}. \end{aligned} \tag{5.6}$$

The λ and η superscripts enclosed in parentheses in Eq. (5.6) indicate differentiation with respect to μ_{P_λ} or μ_{Q_η}, respectively. To avoid cluttering of notation we will avoid from now on the explicit notation of the nuclear centers and assume that λ always refers to nucleus P and η to nucleus Q. The sum over Cartesian coordinates in the SD term of Eq. (5.6) arises from the orientation of the spin quantization axis along the x, y and z axes of the coordinate system. Because our focus is here on the calculation of the perturbed density matrix elements we will not further discuss the calculation of the perturbed operator matrix elements in Eq. (5.6). We refer the interested reader to [168].

The perturbed density matrix elements in Eq. (5.6) can be calculated according to McWeeny's SCP. Assuming perturbation independent basis and auxiliary functions, we find as generic expression for the perturbed spin density matrix elements

$$P_{\mu\nu}^{\sigma(\lambda),X} = 2 \sum_{i}^{occ} \sum_{a}^{uno} \frac{\mathcal{K}_{ia}^{\sigma(\lambda),X}}{\epsilon_i^\sigma - \epsilon_a^\sigma} (c_{\mu a}^\sigma c_{\nu i}^\sigma + s\, c_{\mu i}^\sigma c_{\nu a}^\sigma), \tag{5.7}$$

where the X superscript represents FC, SD or PSO contributions, while σ denotes either the α or β spin component. The scaling factor s is -1 for the imaginary PSO perturbation operator and 1 otherwise. The $\mathcal{K}_{ia}^{\sigma(\lambda),X}$ denotes a perturbed spin Kohn-Sham matrix element in MO

representation. It is given by

$$K_{ia}^{\sigma(\lambda),X} = \sum_{\mu,\nu} c_{\mu i}^{\sigma} K_{\mu\nu}^{(\lambda),X} c_{\nu a}^{\sigma}, \tag{5.8}$$

where the $K_{\mu\nu}^{(\lambda),X}$ matrix elements are given by

$$K_{\mu\nu}^{(\lambda),X} = H_{\mu\nu}^{(\lambda),X} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle \left[x_{\bar{k}}^{(\lambda),X} + z_{\bar{k}}^{(\lambda),X} \right]. \tag{5.9}$$

The first term in Eq. (5.9) is the perturbed core Hamiltonian for the X contribution. The exchange-correlation perturbed fitting coefficients are

$$z_{\bar{k}}^{(\lambda),X} = \sum_{\bar{l},\bar{m}} G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} | f_{xc}[\tilde{\rho}] | \bar{m} \rangle x_{\bar{m}}^{(\lambda),X}. \tag{5.10}$$

The similarity between these equations and the ones from the already discussed response property calculations is obvious. Thus, the EN2 + MINRES algorithm can be used for direct ADPT calculation of NSSCCs.

To test this new implementation, we took three amylose systems (8, 16 and 32 units) from section 5.3.1 and calculated selected nuclear spin-spin couplings in these systems. The calculations were performed with 24 cores using the PBE/DZVP/GEN-A2 level of theory. The structures of these systems are depicted in Figure 5.15. This Figure also shows the atoms for which nuclear spin-spin couplings are calculated. The thresholds for the ADPT solutions were the same as in the polarizability calculations, i.e. 5.0×10^{-6} for the EN2 residual tolerance and 10^{-7} in the approximated MINRES residual norm. The calculated NSSCC comparison is shown in Table 5.9. All values are given in Hz. The Δ columns in Table 5.9 list the differences in the NSSCC values from the TED and EN2 + MINRES calculations. As these columns show the calculated NSSCCs from TED and MINRES density fitting are in general very similar. Noticeable differences are only found for the Fermi-Contact (FC) term. However, they are still well below the intrinsic accuracy of GGA calculated NSSCCs which is in the range of 3 to 5 Hz.

In the first column of Table 5.9 the NSSCC contribution is listed for the specified atom pairs. The following columns give the corresponding values for each contribution in Hz. Also, the absolute difference, Δ, between the coupling constants from the two different density fittings is given in Hz. This data structure is used for each of the three amylose systems in Table 5.9.

In general, the obtained results with TED and MINRES are the same. It is worth noticing, that the FC contribution in Table 5.9 is the one that changes most from one methodology to another. Nevertheless, the differences are in the range of tenths of Hz, at most. For all other contributions the use of any of these methodologies leads to the same result. This further shows the stability and robust performance of MINRES inside deMon2k, making it suitable for NMR property calculations such as NSSCCs despite the system complexity or the number of auxiliary functions used.

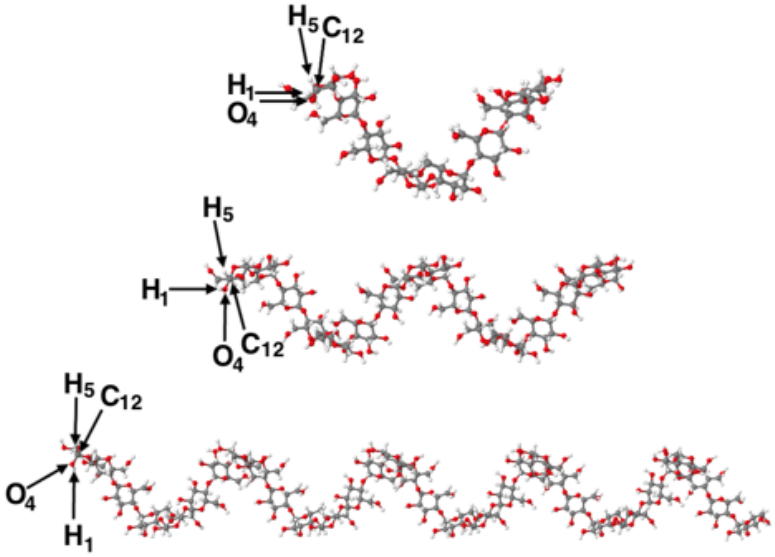

Figure 5.15: Amylose systems with 8 (top), 16 (middle) and 32 (bottom) units, respectively. The atomic symbols (H_1, H_5, C_{12} and O_4) indicate the four centers for which nuclear spin-spin couplings are calculated. The atomic color code is: hydrogen (white), carbon (gray) and oxygen (red).

Table 5.9: NSSCC contributions calculated using the PBE/DZVP/GEN-A2 level of theory for three amylose structures containing 8, 16 and 32 units, respectively, using the TED and EN2 + MINRES approaches. The calculated values are shown in Hz alongside with their absolute differences Δ.

Atom pair	**8 Units amylose**			**16 Units amylose**			**32 Units amylose**		
H1-O4	TED	MINRES	Δ	TED	MINRES	Δ	TED	MINRES	Δ
FC	-36.69	-36.82	0.13	-36.70	-37.01	0.31	-30.84	-31.08	0.24
SD	0.83	0.93	0.10	0.83	0.86	0.03	0.55	0.54	0.01
DSO	-0.10	-0.10	0.00	-0.10	-0.10	0.00	-0.15	-0.15	0.00
PSO	-9.35	-9.35	0.00	-9.35	-9.35	0.00	-9.42	-9.42	0.00
Total	-45.31	-45.35	0.04	-45.32	-45.60	0.28	-39.86	-40.11	0.25
H1-H5									
FC	8.26	8.23	0.03	8.26	8.19	0.07	20.08	20.04	0.04
SD	0.25	0.26	0.01	0.25	0.26	0.01	0.01	0.01	0.00
DSO	0.69	0.69	0.00	0.69	0.69	0.00	-1.67	-1.67	0.00
PSO	-0.78	-0.78	0.00	-0.78	-0.78	0.00	2.10	2.10	0.00
Total	8.43	8.41	0.02	8.42	8.36	0.06	20.52	20.48	0.04
H1-C12									
FC	-0.45	-0.31	0.14	-0.45	-0.27	0.18	-2.64	-2.50	0.14
SD	-0.02	-0.02	0.00	-0.02	-0.02	0.00	-0.03	-0.03	0.00
DSO	-0.06	-0.06	0.00	-0.06	-0.06	0.00	-0.01	-0.01	0.00
PSO	0.12	0.12	0.00	0.12	0.12	0.00	-0.14	-0.14	0.00
Total	-0.41	-0.26	0.15	-0.41	-0.23	0.18	-2.83	-2.68	0.15
O4-H5									
FC	-1.87	-2.06	0.19	-1.86	-1.96	0.10	-16.93	-16.89	0.04
SD	0.14	0.14	0.00	0.14	0.14	0.00	0.12	0.12	0.00
DSO	0.01	0.01	0.00	0.01	0.01	0.00	-0.01	-0.01	0.00
PSO	-0.26	-0.26	0.00	-0.26	-0.26	0.00	0.08	0.08	0.00
Total	-1.98	-2.17	0.19	-1.97	-2.08	0.11	-16.74	-16.70	0.04
O4-C12									
FC	16.85	16.92	0.07	16.84	16.89	0.05	21.63	21.65	0.02
SD	-1.83	-1.83	0.00	-1.83	-1.83	0.00	-1.12	-1.12	0.00
DSO	-0.07	-0.07	0.00	-0.07	-0.07	0.00	-0.08	-0.08	0.00
PSO	-0.65	-0.65	0.00	-0.65	-0.65	0.00	2.57	2.57	0.00
Total	14.29	14.36	0.07	14.29	14.34	0.05	23.00	23.02	0.02
H5-C12									
FC	175.07	174.98	0.09	175.05	174.95	0.10	204.81	204.75	0.06
SD	0.22	0.22	0.00	0.22	0.22	0.00	0.30	0.30	0.00
DSO	0.61	0.61	0.00	0.61	0.61	0.00	0.56	0.56	0.00
PSO	-0.80	-0.80	0.00	-0.80	-0.80	0.00	-1.07	-1.07	0.00
Total	175.11	175.01	0.10	175.09	174.99	0.10	204.61	204.53	0.08

6

Conclusions and Perspectives

Contents

6.1 Conclusions

From the results obtained in this thesis the following conclusions can be drawn:

- The implementation of the Krylov subspace MINRES algorithm for the solution of the inhomogeneous fitting equation system in deMon2k has removed the computational bottleneck for the variational fitting of the Coulomb potential. Because MINRES also resolves the numerical problems of the density fitting arising from finite precision arithmetic a robust and reliable fitting implementation is achieved. The comparison of the different MINRES variants, DIRECT, ASYMPTOTIC and MIXED, showed that MINRES ASYMPTOTIC is the method of choice. It combines superb computational performance with a low memory demand. As a result, subquadratic scaling with respect to the number of auxiliary functions in combination with excellent parallel scalability is achieved. The key to success is the use of blocked atomic preconditioners in the form of modified Cholesky or truncated eigenvalue decomposed (TED) atomic blocks. Both atomic block preconditioners improve significantly the MINRES convergence. In this work, all our larger test calculations showed improved computational performance with respect to the TED density fitting implementation. Moreover, speed up factors of 4 in single point energy calculations for systems with over 100,000 auxiliary functions were obtained using the density fitting with MINRES compared to the standard TED approach.

- The parallel scaling of MINRES is close to ideal due to the matrix-vector multiplications. This yields, in combination with the double asymptotic expansion of the two-center ERIs, a straightforward parallel implementation of the MINRES algorithm.

- The modified Cholesky preconditioner is well suited for atomic blocks of the Coulomb matrix that become indefinite. A particular convenience of this approach is that we can work with the full rank of the matrix. However, TED preconditioner has proven to be as good as modified Cholesky for medium size auxiliary function sets like GEN-A2 or GEN-A2* and even better when large auxiliary function sets such as GEN-A3* and GEN-A4* are used where modified Cholesky decomposition-based preconditioners fail.

- Structural optimizations were possible for large systems like fullerenes with up to 960 atoms using the PBE/DZVP/GEN-A2* methodology. This is a breakthrough because with the current implementation of deMon2k this type of structure optimizations were not possible due to the large auxiliary function set dimensions. Moreover, the geometry optimization of an oligothiophene system consisting of 100 thiophene units with a lateral

extension of 35 nm was possible to be performed with the new MINRES implementation at the VWN/DZVP/GEN-A2* level of theory. Unexpectedly, this optimization yielded two distinct results when performed in Cartesian or internal redundant coordinates. Although the geometries of these minima are obviously different, their energy separation at the VWN/DZVP/GEN-A2* level of theory is below 0.5 kcal/mol. This demonstrates the challenge of nanosystem structure optimization in terms of accuracy and diversity of the potential energy surface.

- The static and dynamic polarizabilities obtained with the EN2 + MINRES algorithm showed no significant deviation from the ADPT implemented years ago using the TED of the response matrix for the solution of the ADPT equation system. On the other hand, the hyperpolarizability calculations showed larger deviations compared to the polarizabilities. We attribute this to the TED of the response matrix in the older ADPT implementation that can hide certain pathologies of the underlying GGA functional.

- The optimized fullerene structures were used to obtain their polarizabilities using the GEN-A2* auxiliary function set. These results show a slightly higher polarizability per atom than the ones already reported using the GEN-A2 auxiliary function set. Although different functionals were used in these calculations, we attribute this difference mainly to the extended auxiliary function set in our calculations. Moreover, the EN2 + MINRES algorithm also permitted us to calculate the polarizabilities of amyloses and DNA fragments with GEN-A2* auxiliary function sets. For these systems, where we could compare TED and MINRES density fitting, we found in all cases excellent agreements between these approaches. However, the polarizabilities of the larger amyloses and DNA fragments were only accessible by the newly developed EN2 + MINRES algorithm because of its greatly reduced RAM demand. In this respect, it is important to note that the largest amylose system contains more than 1,300 atoms, 28,275 basis functions and 112,056 auxiliary functions. Although the main aim of these polarizability calculations was to validate EN2 + MINRES algorithm and to explore its potentiality for large scale polarizability calculations, the obtained results are also of scientific interest. Whereas the studied fullerenes show a strong continuous increase of the polarizability per atom with system size, the amyloses show only a moderate increase. Moreover, it seems that amylose polarizability per atom reaches already a saturation for the largest studied system with 64 glucose units. On the other hand, the polarizability per atom of the DNA fragments decreases with system size and, therefore, shows just the opposite trend as for the fullerenes and amyloses.

- NSSCCs between selected atoms of three amyloses were calculated with PBE/DZVP/GEN-A2 level of theory using TED and MINRES density fitting approaches. The differences between both approaches is less than 1 Hz for all compared NSSCC contributions. These differences are well below the intrinsic accuracy of GGA functionals for spin-spin coupling constants which is around 3 to 5 Hz.

- Finally, it is important to remark, that the sizes of the largest systems studied in this thesis by ab initio methodologies, namely the C_{960} fullerene, the oligothiophene containing 100 thiophene units, the amylose with 64 glucose units and the DNA fragment with 16 base pairs, are all at the nanometric length scale with dimensions in the range of 3, 35, 20 and 5 nm, respectively.

6.2 Perspectives

Based on the achievements presented in this thesis the following new research topics can be explored:

- In this work we showed how the TED of the Coulomb matrix can be substituted throughout the deMon2k program, i.e. in the SCF energy calculation, the structure optimization and finally in perturbative response calculations for polarizabilities, hyperpolarizabilities and nuclear spin-spin coupling constants. With this development the linear algebra bottleneck associated to the density fitting is eliminated. It is important to note that this bottleneck was twofold, namely in terms of computational as well as random access memory (RAM) demand. Both problems were overcome with the here presented adaptation and implementation of the Krylov subspace method MINRES. As a result, only the Kohn-Sham or Fock matrix transformations and diagonalizations in Roothaan-Hall type SCF approaches remain as limiting steps for massively parallel large scale first-principles ADFT calculations. In large scale (more than 1000 atoms) parallel deMon2k calculations these linear algebra steps account for more than 50% of the CPU time and introduce a dominant cubic scaling. Again, they are also critical in terms of RAM allocation. However, according to Stewart[172], the Kohn-Sham matrix diagonalization is not a prerequisite for obtaining a self-consistent density matrix through a SCF procedure. Instead, a self-consistent density matrix, i.e. SCF convergence, can also be obtained by the annihilation of all MO Kohn-Sham matrix elements that connect occupied and virtual molecular orbitals. This method is called pseudodiagonalization (PD) and is applicable, at least in

principle, in the framework of localized MOs. Thus, the PD algorithm seems promising to tackle the twofold bottleneck of Roothaan-Hall type SCF approaches. Preliminary work in our laboratories has shown that ADFT SCF convergence can be reached with a modified PD algorithm and that the orbital localization is kept during the PD ADFT SCF procedure. Further work in this direction is currently underway in our laboratories.

- The here developed MINRES density fitting approach can be further enhanced if a geometry inspection is performed. Thus, the preconditioner could be built on top of this information will be optimal for each case of study. Likewise, it could be also possible to build a hybrid preconditioner as a mixture between an atomic and a molecular one. Hence, the preconditioner will contain more information about the molecular structure and the MINRES algorithm will converge in fewer iterations. This will be of particular importance for response calculations with MINRES density fitting that still suffer from slower MINRES convergence.

- Last but not least, the newly developed MINRES density fitting approach is currently implemented by several research groups into special property branches of deMon2k including analytic second ADFT energy derivatives and time-dependent ADFT, to name a few.

MINRES algorithm

Schemes A1, A2 and A3 show the MINRES iterative algorithm alongside its dependencies as implemented in deMon2k.

Scheme A.1: Algorithm **MINRES** for solving $\mathbf{Gx} = \mathbf{J}$. It estimates $\phi = ||\mathbf{r}^{(k)}||$, $\psi = ||\mathbf{Gr}^{(k)}||$, $\chi = ||\mathbf{x}^{(k)}||$, $\mathcal{A} = ||\mathbf{G}||$ and $\kappa = \text{cond}(\mathbf{G})$.

MINRES ($\mathbf{G}$, $\mathbf{J}$, maxit) $\longrightarrow \mathbf{x}$, ϕ, ψ, χ, $||\mathbf{G}||$, κ

$\beta_1 = ||\mathbf{J}||, \quad \mathbf{v}^{(0)} = \mathbf{0}, \quad \beta_1\mathbf{v}^{(1)} = \mathbf{J}, \quad \phi_0 = \tau_0 = \beta_1, \quad \chi_0 = 0, \quad \kappa = 1$

$\delta_1^{(1)} = \gamma_{\min} = 0, \quad c_0 = -1, \quad s_0 = 0, \quad \mathbf{d}^{(0)} = \mathbf{d}^{(-1)} = \mathbf{x}^{(0)} = \mathbf{0}, \quad k = 1$

while no stopping condition is true

LanczosStep ($\mathbf{G}$, $\mathbf{v}^{(k)}$, $\mathbf{v}^{(k-1)}$, β_k) $\rightarrow \alpha_k$, β_{k+1}, $\mathbf{v}^{(k+1)}$

$\delta_k^{(2)} = c_{k-1}\delta_k^{(1)} + s_{k-1}\alpha_k, \quad \gamma_k^{(1)} = s_{k-1}\delta_k^{(1)} - c_{k-1}\alpha_k$

$\epsilon_{k+1}^{(1)} = s_{k-1}\beta_{k+1}, \quad \delta_{k+1}^{(1)} = -c_{k-1}\beta_{k+1}$

SymOrtho ($\gamma_k^{(1)}$, β_{k+1}) $\rightarrow c_k$, s_k, $\gamma_k^{(2)}$

$\tau_k = c_k\phi_{k-1}, \quad \phi_k = s_k\phi_{k-1}, \quad \psi_{k-1} = \phi_{k-1}\sqrt{[\gamma_k^{(1)}]^2 + [\delta_{k+1}^{(1)}]^2}$

if $k = 1 \quad \mathcal{A}_k = \sqrt{\alpha_1^2 + \beta_2^2}$ **else** $\mathcal{A}_k = \max\left\{\mathcal{A}_{k-1}, \sqrt{\beta_k^2 + \alpha_k^2 + \beta_{k+1}^2}\right\}$ **end**

if $\gamma_k^{(2)} \neq 0$,

$\mathbf{d}^{(k)} = \left(\mathbf{v}^{(k)} - \delta_k^{(2)}\mathbf{d}^{(k-1)} - \epsilon_k^{(1)}\mathbf{d}^{(k-2)}\right)/\gamma_k^{(2)}, \quad \mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + \tau_k\mathbf{d}^{(k)}, \quad \chi_k = ||\mathbf{x}^{(k)}||$

$\gamma_{\min} = \min\left\{\gamma_{\min}, \quad \gamma_k^{(2)}\right\}, \quad \kappa = \mathcal{A}_k/\gamma_{\min}$

end

$k \leftarrow k + 1$

end

$\mathbf{x} = \mathbf{x}^{(k)}, \quad \phi = \phi_k, \quad \psi = \phi_k\sqrt{[\gamma_{k+1}^{(1)}]^2 + [\delta_{k+2}^{(1)}]^2}, \quad \chi = \chi_k, \quad \mathcal{A} = \mathcal{A}_k$

end

Scheme A.2: Algorithm **LanczosStep** to generate a new Lanczos vector.

LanczosStep ($\mathbf{G}$, $\mathbf{v}^{(k)}$, $\mathbf{v}^{(k-1)}$, β_k) $\rightarrow \alpha_k$, β_{k+1}, $\mathbf{v}^{(k+1)}$

$\mathbf{p}^{(k)} = \mathbf{Gv}^{(k)}, \qquad \alpha_k = [\mathbf{v}^{(k)}]^T\mathbf{p}^{(k)}, \qquad \mathbf{p}^{(k)} \leftarrow \mathbf{p}^{(k)} - \alpha_k\mathbf{v}^{(k)}$

$\mathbf{v}^{(k+1)} = \mathbf{p}^{(k)} - \beta_k\mathbf{v}^{(k-1)}, \qquad \beta_{k+1} = ||\mathbf{v}^{(k+1)}||$

if $\beta_{k+1} \neq 0, \qquad \mathbf{v}^{(k+1)} \leftarrow \mathbf{v}^{(k+1)}/\beta_{k+1}$ **end**

Scheme A.3: Algorithm **SymOrtho** performing a Givens rotation.

SymOrtho $(a, b) \rightarrow c, s, r$

if $b = 0$

$s = 0,\ r = |a|,$ **if** $a = 0,\ c = 1$ **else** $c = \mathrm{sign}(a)$ **end**

elseif $a = 0$

$c = 0,\ s = \mathrm{sign}(b),\ r = |b|$

elseif $|b| \geq |a|$

$\tau = a/b,\ s = \mathrm{sign}(b)/\sqrt{1+\tau^2},\ c = s\tau,\ r = b/s$

elseif $|a| > |b|$

$\tau = b/a,\ c = \mathrm{sign}(a)/\sqrt{1+\tau^2},\ s = c\tau,\ r = a/c$

end

B

Lanczos algorithm

Lanczos algorithm

As an example for the Lanczos iteration algorithm take the 3×3 symmetric matrix $\mathbf{A}$ as follows

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \tag{B.1}$$

We now use the Lanczos algorithm to construct the orthogonal basis that transforms $\mathbf{A}$ into tridiagonal form. To this end, we initialize the Lanczos iterations for $k = 0$ with

$$\mathbf{v}^{(0)} = \mathbf{0} \quad \text{and} \quad \beta_0 = 1. \tag{B.2}$$

For $k = 1$ we set

$$\mathbf{v}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \beta_1 = 1, \tag{B.3}$$

which yields for α_1 in the first Lanczos step ($k = 1$)

$$\alpha_1 = \mathbf{v}^{(1)^T} \mathbf{A} \mathbf{v}^{(1)} = (1, 0, 0) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = (1, 0, 0) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 1. \tag{B.4}$$

For the second orthogonal vector follows

$$\begin{aligned} \mathbf{v}^{(2)} &= \mathbf{A}\mathbf{v}^{(1)} - \alpha_1 \mathbf{v}^{(1)} - \beta_1 \mathbf{v}^{(0)} \\ &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{B.5}$$

The corresponding β_2 is calculated as the norm of $\mathbf{v}^{(2)}$ and, therefore, its value is 1. With this

calculation the first Lanczos step is finished. In the second Lanczos step ($k = 2$), we first calculate α_2 as

$$\alpha_2 = \mathbf{v}^{(2)^T}\mathbf{A}\mathbf{v}^{(2)} = (0,0,1)\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}0\\ 0\\ 1\end{pmatrix} = (0,0,1)\begin{pmatrix}1\\ 1\\ 0\end{pmatrix} = 0. \tag{B.6}$$

With this α_2 value, the third orthogonal vector is calculated as

$$\begin{aligned}\mathbf{v}^{(3)} &= \mathbf{A}\mathbf{v}^{(2)} - \alpha_2\mathbf{v}^{(2)} - \beta_1\mathbf{v}^{(1)}\\ &= \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}0\\ 0\\ 1\end{pmatrix} - 0 - 1\begin{pmatrix}1\\ 0\\ 0\end{pmatrix} = \begin{pmatrix}1\\ 1\\ 0\end{pmatrix} - \begin{pmatrix}1\\ 0\\ 0\end{pmatrix} = \begin{pmatrix}0\\ 1\\ 0\end{pmatrix}.\end{aligned} \tag{B.7}$$

Because the norm of $\mathbf{v}^{(3)}$ is 1, β_3 is 1, too. This concludes the second Lanczos step. In the third Lanczos step ($k = 3$), we first calculate α_3 as

$$\alpha_3 = \mathbf{v}^{(3)^T}\mathbf{A}\mathbf{v}^{(3)} = (0,1,0)\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}0\\ 1\\ 0\end{pmatrix} = (0,1,0)\begin{pmatrix}0\\ 1\\ 1\end{pmatrix} = 1. \tag{B.8}$$

With this α_3 the new orthogonal vector is the null vector with a corresponding vanishing β. Thus, the Lanczos iterations are converged. Collecting the orthogonal $\mathbf{v}$ vectors in the $\mathbf{V}$ matrix and the obtained α and β values in the $\mathbf{T}$ matrix yields

$$\mathbf{V} = \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix} \quad \text{and} \quad \mathbf{T} = \begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}. \tag{B.9}$$

Note that the tridiagonal matrix $\mathbf{T}$ has the correct form. The obtained matrices can be used to calculate the $\mathbf{A}$ matrix by

$$\mathbf{A} = \mathbf{V}\mathbf{T}\mathbf{V}^T, \tag{B.10}$$

which confirms the results from the Lanczos iterations

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \tag{B.11}$$

C

QR decomposition

QR decomposition

The QR decomposition (also called QR factorization) of a real squared matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{QR} \tag{C.1}$$

where $\mathbf{Q}$ is an orthogonal matrix ($\mathbf{Q^T Q} = \mathbf{E}$) and $\mathbf{R}$ is an upper triangular matrix.

To calculate the QR factorization of $\mathbf{A}$ we first construct an orthonormal basis by Gram-Schmidt[228] orthogonalization. To this end, we express $\mathbf{A}$ in terms of its column vectors as

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \cdots, \mathbf{a}_n) \tag{C.2}$$

With these column vectors the Gram-Schmidt orthogonalization is given by

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\
\mathbf{u}_2 &= \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|}\mathbf{u}_1, \qquad \mathbf{q}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\
\mathbf{u}_3 &= \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|}\mathbf{u}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|}\mathbf{u}_2, \qquad \mathbf{q}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \\
&\vdots \\
\mathbf{u}_n &= \mathbf{a}_n - \sum_{i=1}^{n-1} \frac{\mathbf{a}_n \cdot \mathbf{u}_i}{\|\mathbf{u}_i\|}\mathbf{u}_i, \qquad \mathbf{q}_n = \frac{\mathbf{u}_n}{\|\mathbf{u}_n\|},
\end{aligned} \tag{C.3}$$

The resulting orthonormal $\mathbf{q}$ vectors build the $\mathbf{Q}$ matrix as

$$\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \cdots, \mathbf{q}_n). \tag{C.4}$$

To calculate the $\mathbf{R}$ matrix from Eq. (C.1) we explore the orthonormality of $\mathbf{Q}$, i.e. $\mathbf{Q}^T\mathbf{Q} = \mathbf{E}$. Thus, it follows

$$\mathbf{R} = \mathbf{Q^T A} = \begin{pmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix} = \begin{pmatrix} \mathbf{q}_1^T \cdot \mathbf{a}_1 & \mathbf{q}_1^T \cdot \mathbf{a}_2 & \dots & \mathbf{q}_1^T \cdot \mathbf{a}_n \\ 0 & \mathbf{q}_2^T \cdot \mathbf{a}_2 & \dots & \mathbf{q}_2^T \cdot \mathbf{a}_n \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \mathbf{q}_n^T \cdot \mathbf{a}_n \end{pmatrix}. \tag{C.5}$$

Here we have used the fact that $\mathbf{q}_i^T \cdot \mathbf{a}_j = 0 \; \forall \; i > j$.

As example, consider the following matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}. \tag{C.6}$$

Performing the Gram-Schmidt procedure, we obtain

$$\begin{aligned}
\mathbf{u}_1 = \mathbf{a}_1 &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\
\mathbf{q}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||} &= \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \\
\mathbf{u}_2 = \mathbf{a}_2 - (\mathbf{a}_2 \cdot \mathbf{q}_1)\mathbf{q}_1 &= \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}, \\
\mathbf{q}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||} &= \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}, \\
\mathbf{u}_3 = \mathbf{a}_3 - (\mathbf{a}_3 \cdot \mathbf{q}_1)\mathbf{q}_1 - (\mathbf{a}_3 \cdot \mathbf{q}_2)\mathbf{q}_2, & \\
\mathbf{q}_3 = \frac{\mathbf{u}_3}{||\mathbf{u}_3||} &= \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \qquad .
\end{aligned} \tag{C.7}$$

Thus the $\mathbf{Q}$ and $\mathbf{R}$ matrices are obtained as

$$\mathbf{Q} = [\mathbf{q}_1|\mathbf{q}_2|\mathbf{q}_3] = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \tag{C.8}$$

$$\mathbf{R} = \begin{pmatrix} \mathbf{q}_1^T \cdot \mathbf{a}_1 & \mathbf{q}_1^T \cdot \mathbf{a}_2 & \mathbf{q}_1^T \cdot \mathbf{a}_3 \\ 0 & \mathbf{q}_2^T \cdot \mathbf{a}_2 & \mathbf{q}_2^T \cdot \mathbf{a}_3 \\ 0 & 0 & \mathbf{q}_3^T \cdot \mathbf{a}_3 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{pmatrix}. \tag{C.9}$$

D

Successive QR decomposition by Givens rotations

Instead of performing QR decomposition of a matrix **A** in one step as described in appendix C, it can be also performed successively, e.g. by a sequence of Givens rotations. Such an approach is particularly advantageous if the **A** matrix itself is stepwise constructed as in MINRES. A Givens rotation can be used to annihilate a specific matrix (or vector) element in the form

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}, \tag{D.1}$$

From the conditions,

$$ca_1 + s \cdot a_2 \equiv r \qquad \wedge \qquad -sa_1 + ca_2 \equiv 0, \tag{D.2}$$

and the normalization condition for c (cosine) and s (sine),

$$c^2 + s^2 = 1, \tag{D.3}$$

follows

$$r = \sqrt{a_1^2 + a_2^2}, \quad c = \frac{a_1}{r} \quad \text{and} \quad s = \frac{a_2}{r}. \tag{D.4}$$

As an example, we now apply successive Givens rotation to the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{pmatrix} \tag{D.5}$$

In order to transform it into a triangular form.

We start by eliminating a_{41} through a_{31}. Thus, we find

$$r = \sqrt{a_{31}^2 + a_{41}^2} = \sqrt{1^2 + 1^2} = \sqrt{2}, \quad c = \frac{a_{31}}{r} = \frac{1}{\sqrt{2}} \quad \text{and} \quad s = \frac{a_{41}}{r} = \frac{1}{\sqrt{2}}. \tag{D.6}$$

Thus, the first Givens rotation matrix has the following form

$$\mathbf{\Omega}^{(1)} = \begin{pmatrix} & & 0 & 0 \\ & \mathbf{E} & 0 & 0 \\ 0 & 0 & c & s \\ 0 & 0 & -s & c \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}. \tag{D.7}$$

Applying this rotation matrix to the coefficient matrix $\mathbf{A}$ we obtain the first rotated matrix, $\mathbf{A}^{(1)}$, given by

$$\mathbf{A}^{(1)} = \mathbf{\Omega}^{(1)}\mathbf{A} = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ \sqrt{2} & 3/\sqrt{2} & \sqrt{2} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix}. \tag{D.8}$$

The next rotation eliminates $a_{31}^{(1)}$ through $a_{21}^{(1)}$. Thus, we find

$$r = \sqrt{a_{21}^{(1)^2} + a_{31}^{(1)^2}} = \sqrt{1+2} = \sqrt{3}, \quad c = \frac{a_{21}^{(1)}}{r} = \frac{1}{\sqrt{3}}, \quad \text{and} \quad s = \frac{a_{31}^{(1)}}{r} = \frac{\sqrt{2}}{\sqrt{3}}. \tag{D.9}$$

Hence, the second Givens rotation has the following form

$$\mathbf{\Omega}^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{3} & \sqrt{2}/\sqrt{3} & 0 \\ 0 & -\sqrt{2}/\sqrt{3} & 1/\sqrt{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{D.10}$$

Again, applying this rotation matrix to the coefficient matrix $\mathbf{A}^{(1)}$ we obtained a second rotated matrix, $\mathbf{A}^{(2)}$ which is given by

$$\mathbf{A}^{(2)} = \mathbf{\Omega}^{(2)}\mathbf{A}^{(1)} = \begin{pmatrix} 1 & -1 & 4 \\ \sqrt{3} & 7/\sqrt{3} & 0 \\ 0 & -5/\sqrt{6} & \sqrt{6} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix}. \tag{D.11}$$

The third rotation will eliminate a_{21} through a_{11}, Thus, we find

$$r = \sqrt{a_{11}^{(2)^2} + a_{21}^{(2)^2}} = \sqrt{1+3} = 2, \quad c = \frac{a_{11}^{(2)}}{r} = \frac{1}{2}, \quad \text{and} \quad s = \frac{a_{21}^{(2)}}{r} = \frac{\sqrt{3}}{2}. \tag{D.12}$$

This gives us the third rotation matrix in the following form

$$\mathbf{\Omega}^{(3)} = \begin{pmatrix} 1/2 & \sqrt{3}/2 & 0 & 0 \\ -\sqrt{3}/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{D.13}$$

Applying $\mathbf{\Omega}^{(3)}$ to $\mathbf{A}^{(2)}$ we obtain $\mathbf{A}^{(3)}$ as follows

$$\mathbf{A}^{(3)} = \mathbf{\Omega}^{(3)}\mathbf{A}^{(2)} = \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5/\sqrt{3} & -2\sqrt{3} \\ 0 & -5/\sqrt{6} & \sqrt{6} \\ 0 & -5/\sqrt{2} & -\sqrt{2} \end{pmatrix}. \tag{D.14}$$

Now we need to eliminate the subdiagonal rows from the second column. Eliminating $a_{42}^{(3)}$ through $a_{32}^{(3)}$ gives us

$$\begin{aligned} r &= \sqrt{a_{32}^{(3)^2} + a_{42}^{(3)^2}} = \sqrt{(-5/\sqrt{6})^2 + (-5/\sqrt{2})^2} = 10/\sqrt{6}, \\ c &= \frac{a_{32}^{(3)}}{r} = -\frac{1}{2}, \quad \text{and} \quad s = \frac{a_{42}^{(3)}}{r} = -\frac{\sqrt{3}}{2}. \end{aligned} \tag{D.15}$$

Therefore, the fourth rotation matrix is given by

$$\mathbf{\Omega}^{(4)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & -\sqrt{3}/2 \\ 0 & 0 & \sqrt{3}/2 & -1/2 \end{pmatrix}. \tag{D.16}$$

The fourth times rotated matrix, $\mathbf{A}^{(4)}$, is obtained as

$$\mathbf{A}^{(4)} = \mathbf{\Omega}^{(4)}\mathbf{A}^{(3)} = \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5/\sqrt{3} & -2\sqrt{3} \\ 0 & 10/\sqrt{6} & 0 \\ 0 & 0 & 2\sqrt{2} \end{pmatrix}. \tag{D.17}$$

To eliminate $a_{32}^{(4)}$ through $a_{22}^{(4)}$ we find

$$\begin{aligned} r &= \sqrt{a_{22}^{(4)^2} + a_{32}^{(4)^2}} = \sqrt{(5/\sqrt{3})^2 + (10/\sqrt{6})^2} = 5, \\ c &= \frac{a_{22}^{(4)}}{r} = \frac{1}{\sqrt{3}}, \quad \text{and} \quad s = \frac{a_{32}^{(4)}}{r} = \frac{2}{\sqrt{6}}. \end{aligned} \tag{D.18}$$

Therefore, the fifth rotation matrix is given by

$$\mathbf{\Omega}^{(5)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} & 0 \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{D.19}$$

The $\mathbf{A}^{(5)}$ matrix is obtained as follows

$$\mathbf{A}^{(5)} = \mathbf{\Omega}^{(5)}\mathbf{A}^{(4)} = \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 2\sqrt{2} \\ 0 & 0 & 2\sqrt{2} \end{pmatrix}. \tag{D.20}$$

Ultimately, in order to eliminate the last element, $a_{43}^{(5)}$ through $a_{33}^{(5)}$ we find

$$\begin{aligned} r &= \sqrt{a_{33}^{(5)^2} + a_{43}^{(5)^2}} = \sqrt{(2\sqrt{2})^2 + (2\sqrt{2})^2} = 4, \\ c &= \frac{a_{33}^{(5)}}{r} = \frac{1}{\sqrt{2}}, \quad \text{and} \quad s = \frac{a_{43}^{(5)}}{r} = \frac{1}{\sqrt{2}}. \end{aligned} \tag{D.21}$$

Thus, the sixth rotation matrix is given by

$$\mathbf{\Omega}^{(6)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}. \tag{D.22}$$

Finally, the $\mathbf{A}^{(6)}$ in triangular form is obtained as

$$\mathbf{A}^{(6)} = \mathbf{\Omega}^{(6)}\mathbf{A}^{(5)} = \begin{pmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} \equiv \mathbf{R}. \tag{D.23}$$

The $\mathbf{Q}$ matrix of the full decomposition is formed by the transpose of the successively applied rotations through the following product

$$\mathbf{Q} = \left(\mathbf{\Omega}^{(6)}\mathbf{\Omega}^{(5)}\mathbf{\Omega}^{(4)}\mathbf{\Omega}^{(3)}\mathbf{\Omega}^{(2)}\mathbf{\Omega}^{(1)}\right)^{T}, \tag{D.24}$$

which yields the following explicit form of the orthogonal matrix

$$\mathbf{Q} = \begin{pmatrix} 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix}. \tag{D.25}$$

Energy differences and SCF cycles for benchmark calculations

Table E.1: Absolute energy differences in a.u. with respect to TED standard approach and number of TED and MINRES SCF cycles for single point energy calculations of the fullerene benchmark set.

System	Absolute energy difference			Number of SCF cycles			
	DIRECT	ASYMPTOTIC	MIXED	TED	DIRECT	ASYMPTOTIC	MIXED
C_{20}	8.78E-05	8.77E-05	8.77E-05	13	13	13	13
C_{60}	1.35E-05	1.34E-05	1.36E-05	11	11	13	11
C_{180}	1.81E-03	1.82E-03	1.81E-03	14	14	14	14
C_{240}	2.67E-03	2.68E-03	2.67E-03	14	14	14	14
C_{540}	6.34E-03	6.34E-03	6.34E-03	13	15	15	15
C_{720}	8.63E-03	8.63E-03	8.64E-03	15	16	16	15
C_{960}	9.67E-03	9.68E-03	9.68E-03	16	15	15	15

Table E.2: Absolute energy differences in a.u. with respect to TED standard approach and number of TED and MINRES SCF cycles for single point energy calculations of the water cluster benchmark set.

System	Absolute energy difference			Number of SCF cycles			
	DIRECT	ASYMPTOTIC	MIXED	TED	DIRECT	ASYMPTOTIC	MIXED
$(H_2O)_{50}$	2.87E-03	2.87E-03	2.87E-03	13	13	13	13
$(H_2O)_{100}$	5.15E-03	5.15E-03	5.15E-03	14	15	20	15
$(H_2O)_{200}$	8.94E-03	8.94E-03	8.94E-03	15	35	36	35
$(H_2O)_{300}$	1.27E-02	1.27E-02	1.17E-02	15	38	38	36
$(H_2O)_{400}$	1.58E-02	1.58E-02	1.58E-02	15	39	41	41
$(H_2O)_{500}$	1.91E-02	1.91E-02	1.91E-02	16	43	40	39

Table E.3: Absolute energy differences in a.u. with respect to TED standard approach and number of TED and MINRES SCF cycles for single point energy calculations of the zeolite benchmark set.

System	Absolute energy difference			Number of SCF cycles			
	DIRECT	ASYMPTOTIC	MIXED	TED	DIRECT	ASYMPTOTIC	MIXED
MFI-1	4.53E-03	4.50E-03	4.52E-03	13	14	15	14
MFI-2	9.90E-03	9.69E-03	9.61E-03	14	15	15	15
MFI-3	1.77E-02	1.77E-02	1.76E-02	14	17	17	17
MFI-4	2.19E-02	2.19E-02	2.19E-02	16	17	17	17

Bibliography

[1] J. L. Whitten, *J. Chem. Phys.* **58**, 4496 (1973).

[2] B. I. Dunlap, J. Connolly, J. R. Sabin, *J. Chem. Phys.* **71**, 3396 (1979).

[3] B. I. Dunlap, J. Connolly, J. R. Sabin, *J. Chem. Phys.* **71**, 4993 (1979).

[4] J. Mintmire, B. Dunlap, *Phys. Rev. A* **25**, 88 (1982).

[5] J. Mintmire, J. Sabin, S. Trickey, *Phys. Rev. B* **26**, 1743 (1982).

[6] O. Vahtras, J. Almlöf, M. Feyereisen, *Chem. Phys. Lett.* **213**, 514 (1993).

[7] R. R. Zope, B. I. Dunlap, *J. Chem. Phys.* **124**, 044107 (2006).

[8] B. I. Dunlap, N. Rösch, S. Trickey, *Mol. Phys.* **108**, 3167 (2010).

[9] L. N. Wirz, S. S. Reine, T. B. Pedersen, *J. Chem. Theory Comput.* **13**, 4897 (2017).

[10] Q. Sun, T. C. Berkelbach, J. D. McClain, G. K.-L. Chan, *J. Chem. Phys.* **147**, 164119 (2017).

[11] H. Früchtl, R. A. Kendall, R. J. Harrison, K. G. Dyall, *Int. J. Quantum Chem.* **64**, 63 (1997).

[12] S. Hamel, M. E. Casida, D. R. Salahub, *J. Chem. Phys.* **114**, 7342 (2001).

[13] R. Polly, H.-J. Werner, F. R. Manby, P. J. Knowles, *Mol. Phys.* **102**, 2311 (2004).

[14] D. Mejía-Rodríguez, A. M. Köster, *J. Chem. Phys.* **141**, 124114 (2014).

[15] S. F. Manzer, E. Epifanovsky, M. Head-Gordon, *J. Chem. Theory Comput.* **11**, 518 (2015).

[16] F. A. Delesma, G. Geudtner, D. Mejía-Rodríguez, P. Calaminici, A. M. Köster, *J. Chem. Theory Comput.* **14**, 5608 (2018).

[17] S. Saebo, P. Pulay, *Annu. Rev. Phys. Chem.* **44**, 213 (1993).

[18] H.-J. Werner, F. R. Manby, P. J. Knowles, *J. Chem. Phys.* **118**, 8149 (2003).

[19] M. Schütz, H.-J. Werner, R. Lindh, F. R. Manby, *J. Chem. Phys.* **121**, 737 (2004).

[20] R. A. DiStasio, Y. Jung, M. Head-Gordon, *J. Chem. Theory Comput.* **1**, 862 (2005).

[21] U. Bozkaya, D. Sherril, *J. Chem. Phys.* **147**, 044104 (2017).

[22] C. Peng, J. Calvin, E. Valeev, *Int. J. Quantum Chem.* **119**, e25894 (2019).

[23] L. Michal, *J. Chem. Phys.* **152**, 044104 (2019).

[24] R. T. Gallant, A. St-Amant, *Chem. Phys. Lett.* **256**, 569 (1996).

[25] C. Fonseca Guerra, J. G. Snijders, G. te Velde, E. J. Baerends, *Theor. Chem. Acc.* **99**, 391 (1998).

[26] Y. Jung, A. Sodt, P. M. W. Gill, M. Head-Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6692 (2005).

[27] A. Sodt, J. E. Subotnik, M. Head-Gordon, *J. Chem. Phys.* **125**, 194109 (2006).

[28] A. Sodt, M. Head-Gordon, *J. Chem. Phys.* **128**, 104106 (2008).

[29] S. Reine, E. Tellgren, A. Krapp, T. Kargaard, T. Helgaker, B. Jansik, S. Hazst, P. Salek, *J. Chem. Phys.* **129**, 104101 (2008).

[30] D. P. Tew, *J. Chem. Phys.* **148**, 011102 (2018).

[31] D. N. Laikov, *Chem. Phys. Lett.* **281**, 151 (1997).

[32] A. M. Köster, J. U. Reveles, J. M. del Campo, *J. Chem. Phys.* **121**, 3417 (2004).

[33] U. Birkenheuer, A. B. Gordienko, V. A. Nasluzov, M. K. Fuchs-Rohr, N. Rösch, *Int. J. Quantum Chem.* **102**, 743 (2005).

[34] A. V. Bienvenu, G. Knizia, *J. Chem. Theory Comput.* **14**, 1297 (2018).

[35] J. L. Morales, J. Nocedal, *SIAM J. Optim.* **10**, 1079 (2000).

[36] V. D. Domínguez-Soria, G. Geudtner, J. L. Morales, P. Calaminici, A. M. Köster, *J. Chem. Phys.* **131**, 124102 (2009).

[37] C. Paige, M. Saunders, *SIAM J. Numer. Anal.* **12**, 617 (1975).

[38] S. Choi, C. Paige, M. Saunders, *SIAM J. Sci. Comput.* **33**, 1810 (2011).

[39] J. Mehra, H. Rechenberg, *The Historical Development of Quantum Theory, Vol. 5: Erwin Schrödinger and the Rise of Wave Mechanics. Part 2: The creation of Wave Mechanics, Early Response and Applications 1925-1926*, Spinger-Verlag (1987).

[40] E. Schrödinger, *Phys. Rev.* **28**, 1049 (1926).

[41] W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, John Wiley & Sons (2001).

[42] M. Born, R. Oppenheimer, *Ann. Phys.* **389**, 457 (1927).

[43] P. W. Atkins, R. S. Friedman, *Molecular Quantum Mechanics*, Oxford University Press (2011).

[44] D. J. Griffiths, *Introduction to Quantum Mechanics*, Pearson (2005).

[45] M. Born, V. Fock, *Z. Phys. A* **51**, 165 (1928).

[46] N. C. Handy, A. M. Lee, *Chem. Phys. Lett.* **252**, 425 (1996).

[47] W. Kolos, L. Wolniewicz, *J. Chem. Phys.* **41**, 3663 (1964).

[48] G. Geudtner, P. Calaminici, J. Carmona-Espíndola, J. M. del Campo, V. D. Domínguez-Soria, R. F. Moreno, G. U. Gamboa, A. Goursot, A. M. Köster, J. U. Reveles, T. Mineva, J. Vázquez-Pérez, A. Vela, B. Zúñiga-Gutiérrez, D. R. Salahub, *WIREs: Comput. Mol. Sci.* **2**, 548 (2012).

[49] B. O. Roos, P. R. Taylor, P. E. Sigbahn, *Chem. Phys.* **48**, 157 (1980).

[50] G. D. Purvis III, R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).

[51] J. Rayleigh, *Phil. Trans.* **161**, 77 (1870).

[52] W. Ritz, *J. Reine Angew. Math.* **135**, 1 (1909).

[53] C. Eckart, *Phys. Rev.* **36**, 878 (1930).

[54] R. K. Nesbet, *Variational Principles and Methods in Theoretical Physics and Chemistry*, Cambridge University Press (2002).

[55] L. H. Thomas, *Math. Proc. Camb. Philos. Soc.* **23**, 542 (1927).

[56] E. Fermi, *Rend. Accad. Naz. Lincei* **6**, 32 (1927).

[57] P. A. Dirac, *Math. Proc. Camb. Philos. Soc.* **26**, 376 (1930).

[58] E. Wigner, F. Seitz, *Phys. Rev.* **43**, 804 (1933).

[59] E. Wigner, F. Seitz, *Phys. Rev.* **46**, 509 (1934).

[60] E. Teller, *Rev. Mod. Phys.* **34**, 627 (1962).

[61] J. C. Slater, *Phys. Rev.* **81**, 385 (1951).

[62] J. C. Slater, K. H. Johnson, *Phys. Rev. B* **5**, 844 (1972).

[63] P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864 (1964).

[64] W. Kohn, *Rev. Mod. Phys.* **71**, 1253 (1999).

[65] P. A. M. Dirac, *Math. Proc. Camb. Philos. Soc.* **35**, 416 (1939).

[66] R. G. Parr, W. Yang, *Annu. Rev. Phys. Chem.* **46**, 701 (1995).

[67] W. Kohn, L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).

[68] K. Capelle, *Braz. J. Phys.* **36**, 1318 (2006).

[69] J. C. Slater, *Phys. Rev.* **34**, 1293 (1929).

[70] S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).

[71] D. C. Langreth, M. Mehl, *Phys. Rev. B* **28**, 1809 (1983).

[72] J. P. Perdew, W. Yue, *Phys. Rev. B* **33**, 8800 (1986).

[73] A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).

[74] R. Colle, O. Salvetti, *Theor. Chem. Acc.* **37**, 329 (1975).

[75] R. Colle, O. Salvetti, *J. Chem. Phys.* **79**, 1404 (1983).

[76] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).

[77] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).

[78] A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).

[79] A. D. Becke, *Int. J. Quantum Chem.* **36**, 599 (1989).

[80] A. M. Köster, G. Geudtner, P. Calaminici, M. E. Casida, V. Domínguez, R. Flores-Moreno, G. Gamboa, A. Goursot, T. Heine, A. Ipatov, F. Janetzko, J. del Campo, J. Reveles, A. Vela, B. Zúñiga-Gutiérrez, D. R. Salahub, *The deMon Developers, Cinvestav, México* (2011).

[81] D. R. Salahub, S. Y. Noskov, B. Lev, R. Zhang, V. Ngo, A. Goursot, P. Calaminici, A. M. Köster, A. Álvarez Ibarra, D. Mejía-Rodríguez, J. Rezac, C. F, A. de la Lande, *Molecules* **20**, 4780 (2015).

[82] A. de la Lande, A. Álvarez-Ibarra, K. Hasnaoui, F. Cailliez, X. Wu, T. Mineva, J. Cuny, P. Calaminici, L. López-Sosa, G. Geudtner, I. Navizet, C. García Iriepa, D. Salahub, A. M. Köster, *Molecules* **24**, 1653 (2019).

[83] S. F. Boys, *Proc. R. Soc. Lond. A* **200**, 542 (1950).

[84] J. Pople, R. Nesbet, *J. Chem. Phys.* **22**, 571 (1954).

[85] C. Roothaan, *Rev. Mod. Phys.* **32**, 179 (1960).

[86] J. Binkley, J. Pople, P. Dobosh, *Mol. Phys.* **28**, 1423 (1974).

[87] B. Ford, G. Hall, *Comput. Phys. Commun.* **8**, 337 (1974).

[88] C. C. J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951).

[89] G. G. Hall, *Proc. R. Soc. A* **205**, 541 (1951).

[90] H. Sambe, R. H. Felton, *J. Chem. Phys.* **62**, 1122 (1975).

[91] M. Casida, C. Daul, A. Goursot, A. Köster, L. Petterson, E. Proynov, A. S. Amant, D. R. Salahub, H. Duarte, N. Godbout, J. Guan, C. Jamoroski, M. Leboeuf, V. Malkin, O. Malkina, F. Sima, A. Vela, deMon-KS version 3.4, deMon software (1996).

[92] J. Andzelm, in *Density Functional Methods in Chemistry*, pp. 155–174, Springer (1991).

[93] M. Feyereisen, G. Fitzgerald, A. Komornicki, *Chem. Phys. Lett.* **208**, 359 (1993).

[94] B. Dunlap, *J. Mol Struct.* **529**, 37 (2000).

[95] A. M. Köster, *J. Chem. Phys.* **118**, 9943 (2003).

[96] S. Trickey, J. Alford, J. Boetgger, in *Computational Materials Science*, Theoretical and Computational Chemistry, pp. 171 – 228, Elsevier (2004).

[97] J. Andzelm, E. Radzio, D. R. Salahub, *J. Comput. Chem.* **6**, 520 (1985).

[98] J. Andzelm, N. Russo, D. R. Salahub, *J. Chem. Phys.* **87**, 6562 (1987).

[99] P. Calaminici, F. Janetzko, A. M. Köster, R. Mejía-Olvera, B. Zúñiga-Gutiérrez, *J. Chem. Phys.* **126**, 044108 (2007).

[100] A. M. Köster, *J. Chem. Phys.* **104**, 4114 (1996).

[101] A. Álvarez-Ibarra, *Asymptotic Expansion of Molecular Integrals in Self-Consistent Auxiliary Density Functional Methods*, Ph.D. Thesis, Cinvestav (2013).

[102] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **240**, 283 (1995).

[103] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, *Theor. Chem. Acc.* **97**, 119 (1997).

[104] F. Weigend, *Phys. Chem. Chem. Phys.* **4**, 4285 (2002).

[105] A. M. Köster, G. Geudtner, G. U. Gamboa, A. Álvarez-Ibarra, P. Calaminici, R. Flores-Moreno, A. Goursot, A. de la Lande, D. Mejía-Rodríguez, T. Mineva, L. G. M. Petterson, J. M. Vázquez-Pérez, B. Zúñiga-Gutiérrez, The deMon2k Users Guide, http://www.demon-software.com/public_html/download/manual/manual.pdf, accessed: 16-04-2020.

[106] E. Baerends, D. Ellis, P. Ros, *Chem. Phys.* **2**, 41 (1973).

[107] L. Versluis, T. Ziegler, *J. Chem. Phys.* **88**, 322 (1988).

[108] B. I. Dunlap, *J. Phys. Chem. A* **107**, 10082 (2003).

[109] A. M. Köster, *Entwicklung einer LCGTO-Dichtefunktionalmethode mit Hilfsfunktionen*, Habilitation thesis, Universität Hannover (1998).

[110] L. Belpassi, F. Tarantelli, A. Sgamellotti, H. M. Quiney, *J. Chem. Phys.* **124**, 124104 (2006).

[111] F. Janetzko, A. M. Köster, D. R. Salahub, *J. Chem. Phys.* **128**, 024102 (2008).

[112] A. M. Köster, J. M. del Campo, F. Janetzko, B. Zúñiga-Gutiérrez, *J. Chem. Phys.* **130**, 114106 (2009).

[113] B. G. Johnson, D. A. Holder, in *The J-Matrix Method*, pp. 311–343, Springer (2008).

[114] P. Pulay, *Chem. Phys. Lett.* **73**, 393 (1980).

[115] P. Pulay, *J. Comput. Chem.* **3**, 556 (1982).

[116] P. Salek, S. Hazst, L. Thogersen, P. Jorgensen, P. Manninen, J. Olsen, B. Jansk, S. Reine, F. Pawlowski, E. Tellgren, T. Helgaker, S. Coriani, *J. Chem. Phys.* **126**, 114110 (2007).

[117] C. G. Broyden, *IMA J. Appl. Math.* **6**, 76 (1970).

[118] R. Fletcher, *Comput. J.* **13**, 317 (1970).

[119] D. Goldfarb, *Math. Comput.* **24**, 23 (1970).

[120] D. Shanno, *Math. Comput.* **24**, 647 (1970).

[121] R. Fletcher, *Practical Methods of Optimization*, Wiley (1980).

[122] N. Godbout, D. R. Salahub, J. Andzelm, E. Wimmer, *Can. J. Phys.* **70**, 560 (1992).

[123] A. Krylov, *Izvestija AN SSSR* **7**, 491 (1931).

[124] G. H. Golub, V. L. C. F., *Matrix Computations*, The Johns Hopkins University Press (2013).

[125] W. Arnoldi, *Quart. Appl. Math.* **9**, 17 (1951).

[126] M. Hestenes, E. Stiefel, *J. Res. N.B.S.* **49**, 409 (1952).

[127] Y. Saad, M. Schultz, *SIAM J. Sci. Comput.* **7**, 856 (1986).

[128] H. van der Vorst, *SIAM J. Sci. Comput.* **13**, 631 (1992).

[129] T. Eirola, O. Nevanlinna, *Linear Algebra Appl.* **121**, 511 (1989).

[130] J. N. Pedroza-Montero, J. L. Morales, A. Álvarez Ibarra, G. Geudtner, P. Calaminici, A. M. Köster, *J. Chem. Theory Comput.* **16**, 2965 (2020).

[131] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, USA, 2 edition (2003).

[132] M. Benzi, *J. Comput. Phys.* **182**, 418 (2002).

[133] J. Liu, A. L. Marsden, *J. Comput. Phys.* **383**, 72 (2018).

[134] G. Dillon, V. Kalantzis, Y. Xi, Y. Saad, *SIAM J. Sci. Comput.* **40**, A2234 (2018).

[135] J. Scott, M. Tůma, *Numer. Algorithms* **79**, 1147 (2018).

[136] S. Cheng, N. Higham, *SIAM J. Matrix Anal. and Appl.* **19**, 1097 (1998).

[137] C. Mészáros, *Comput. Math. Appl.* **50**, 1157 (2005).

[138] A. J. Rothman, E. Levina, J. Zhu, *Biometrika* **97**, 539 (2010).

[139] E. D. Nino-Ruiz, A. Sandu, X. Deng, in *Proceedings of the 6th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems*, ScalA '15, Association for Computing Machinery, New York, NY, USA (2015).

[140] T. Jindrová, J. Mikyska, *Fluid Phase Equilib.* **393**, 7 (2015).

[141] E. D. Nino-Ruiz, A. Sandu, X. Deng, *J. Comput. Sci.* **36**, 100654 (2019).

[142] X. Kang, X. Deng, *J. Stat. Comput. Simul.* **90**, 443 (2020).

[143] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, *LAPACK Users' Guide*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 3rd edition (1999).

[144] J. F. H. Lew-Yee, R. Flores-Moreno, J. L. Morales, J. M. del Campo, *J. Chem. Theory Comput.* **16**, 1597 (2020).

[145] J. N. Pedroza-Montero, F. A. Delesma, J. L. Morales, P. Calaminici, A. M. Köster, *J. Chem. Phys. Just accepted* (2020).

[146] G. Seifert, *J. Phys. Chem. A* **111**, 5609 (2007).

[147] A. Álvarez-Ibarra, A. M. Köster, *J. Chem. Phys.* **139**, 024102 (2013).

[148] R. Fournier, *J. Chem. Phys.* **92**, 5422 (1990).

[149] A. Komornicki, G. Fitzgerald, *J. Chem. Phys.* **98**, 1398 (1993).

[150] S. M. Colwell, C. W. Murray, N. C. Handy, R. D. Amos, *Chem. Phys. Lett.* **210**, 261 (1993).

[151] A. M. Lee, S. M. Colwell, *J. Chem. Phys.* **101**, 9704 (1994).

[152] C. Ochsenfeld, M. Head-Gordon, *Chem. Phys. Lett.* **270**, 399 (1997).

[153] V. Weber, A. M. N. Niklasson, M. Challacombe, *Phys. Rev. Lett.* **92**, 193002 (2004).

[154] V. Weber, A. M. N. Niklasson, M. Challacombe, *J. Chem. Phys.* **123**, 044106 (2005).

[155] A. M. N. Niklasson, V. Weber, *J. Chem. Phys.* **127**, 064105 (2007).

[156] J. Kussmann, C. Ochsenfeld, *J. Chem. Phys.* **127**, 204103 (2007).

[157] S. Coriani, S. Host, B. Jansik, L. Thogersen, J. Olsen, P. Jorgensen, S. Reine, F. Pawlowski, T. Helgaker, P. Salek, *J. Chem. Phys.* **126**, 154108 (2007).

[158] M. Kobayashi, T. Touma, H. Nakai, *J. Chem. Phys.* **136**, 084108 (2012).

[159] R. Dovesi, B. Kirtman, L. Maschio, J. Maul, F. Pascale, M. Rérat, *J. Phys. Chem. C* **123**, 8336 (2019).

[160] L. Maschio, B. Kirtman, *J. Chem. Theory Comput.* **16**, 340 (2020).

[161] C. J. Schattenberg, K. Reiter, F. Weigend, M. Kaupp, *J. Chem. Theory Comput.* **16**, 931 (2020).

[162] R. McWeeny, *Methods of Molecular Quantum Mechanics*, London: Academic Press (1992).

[163] D. Mejía-Rodríguez, R. I. Delgado Venegas, P. Calaminici, A. M. Köster, *J. Chem. Theory Comput.* **11**, 1493 (2015).

[164] J. N. Pedroza-Montero, F. A. Delesma-Díaz, R. I. Delgado-Venegas, P. Calaminici, A. M. Köster, *Theor. Chem. Acc.* **135**, 230 (2016).

[165] R. Flores-Moreno, J. Melin, J. V. Ortiz, G. Merino, *J. Chem. Phys.* **129**, 224105 (2008).

[166] R. Flores-Moreno, *J. Chem. Theory Comput.* **6**, 48 (2010).

[167] R. I. Delgado-Venegas, D. Mejía-Rodríguez, R. Flores-Moreno, P. Calaminici, A. M. Köster, *J. Chem. Phys.* **145**, 224103 (2016).

[168] B. Zúniga-Gutiérrez, L. G. Cota, J. N. Pedroza-Montero, A. M. Köster, *to be submitted*.

[169] R. Flores-Moreno, A. M. Köster, *J. Chem. Phys.* **128**, 134105 (2008).

[170] F. Delesma-Díaz, *Range Separated Hybrid Functionals in Auxiliary Density Functional Theory*, Ph.D. Thesis, Cinvestav (2020).

[171] D. Mejía-Rodríguez, *Low-Order Scaling Methods for Auxiliary Density Functional Theory*, Ph.D. Thesis, Cinvestav (2015).

[172] J. J. P. Stewart, P. Császár, P. Pulay, *J. Comput. Chem.* **3**, 227 (1982).

[173] S. Dallakyan, M. Chandross, S. Mazumdar, *Phys. Rev. B* **68**, 075204 (2003).

[174] D. Gentili, F. Di Maria, F. Liscio, L. Ferlauto, F. Leonardi, L. Maini, M. Gazzano, S. Milita, G. Barbarella, M. Cavallini, *J. Mater. Chem.* **22**, 20852 (2012).

[175] L. Zhang, N. S. Colella, B. P. Cherniawski, S. C. B. Mannsfeld, A. L. Briseno, *ACS Appl. Mater. Interfaces* **6**, 5327 (2014).

[176] C. J. Lee, F. M. Jradi, V. D. Mitchell, J. White, C. R. McNeill, J. Subbiah, S. Marder, D. J. Jones, *J. Mater. Chem. C* **8**, 567 (2020).

[177] M. Zhao, B. P. Singh, P. N. Prasad, *J. Chem. Phys.* **89**, 5535 (1988).

[178] H. Thienpont, G. L. J. A. Rikken, E. W. Meijer, W. ten Hoeve, H. Wynberg, *Phys. Rev. Lett.* **65**, 2141 (1990).

[179] M. B. Smith, J. March, *J. Med. Chem.* **50**, 2279 (2007).

[180] A. Mishra, M. Fischer, P. Bäuerle, *Angew. Chem. Int. Ed.* **48**, 2474 (2009).

[181] J. Letizia, J. Rivnay, A. Facchetti, M. Ratner, T. Marks, *Adv. Funct. Mater.* **20**, 50 (2010).

[182] A. Holmes, D. Bradley, A. Brown, P. Burn, J. Burroughes, R. Friend, N. Greenham, R. Gymer, D. Halliday, R. Jackson, A. Kraft, J. Martens, K. Pichler, I. Samuel, *Synt. Met.* **57**, 4031 (1993).

[183] S. Karg, W. Riess, M. Meier, M. Schwoerer, *Synt. Met.* **57**, 4186 (1993).

[184] M. Schwoerer, *Phys. Bl.* **50**, 52 (1994).

[185] S. Roth, W. Graupner, *Synt. Met.* **57**, 3623 (1993).

[186] M. Willander, A. Assadi, C. Svensson, *Synt. Met.* **57**, 4099 (1993).

[187] Y. Ohmori, K. Muro, K. Yoshino, *Synt. Met.* **57**, 4111 (1993).

[188] J. Carmona-Espíndola, *Time-Dependent Auxiliary Density Perturbation Theory: Development, Implementation and Applications*, Ph.D. Thesis, Cinvestav (2012).

[189] J. Guan, P. Duffy, J. T. Carter, D. P. Chong, K. C. Casida, M. E. Casida, M. Wrinn, *J. Chem. Phys.* **98**, 4753 (1993).

[190] D. P. Chong, *J. Chin. Chem. Soc.* **39**, 375 (1992).

[191] A. J. Perkins, *J. Phys. Chem.* **68**, 654 (1964).

[192] D. M. Bishop, L. M. Cheung, *J. Phys. Chem.* **11**, 119 (1982).

[193] M. A. Spackman, *J. Phys. Chem.* **93**, 7594 (1989).

[194] G. R. Alms, A. Burnham, W. H. Flygare, *J. Chem. Phys.* **63**, 3321 (1975).

[195] C. K. Miller, B. J. Orr, J. F. Ward, *J. Chem. Phys.* **74**, 4858 (1981).

[196] N. J. Bridge, A. D. Buckingham, J. W. Linnett, *Proc. R. Soc. London. Series A.* **295**, 334 (1966).

[197] M. P. Bogaard, A. D. Buckingham, R. K. Pierens, A. H. White, *J. Chem. Soc., Faraday Trans. 1* **74**, 3008 (1978).

[198] M. Bogaard, A. Buckingham, G. Ritchie, *Chem. Phys. Lett.* **90**, 183 (1982).

[199] W. F. Murphy, *J. Chem. Phys.* **67**, 5877 (1977).

[200] G. Maroulis, *Chem. Phys. Lett.* **98**, 3022 (1993).

[201] J. Ward, I. Bigio, *Phys. Rev. A* **11**, 60 (1975).

[202] C. Miller, J. Ward, *Chem. Phys. Lett.* **16**, 1179 (1977).

[203] B. Zúñiga-Gutiérrez, A. M. Köster, *Mol. Phys.* **114**, 1026 (2016).

[204] S. V. Shedge, J. Carmona-Espíndola, S. Pal, A. M. Köster, *J. Phys. Chem A* **114**, 2357 (2010).

[205] N. Rega, M. Cossi, V. Barone, *J. Chem. Phys.* **105**, 11060 (1996).

[206] A. Bohne, E. Lang, C.-W. von der Lieth, *Mol. Mod. Ann.* **4**, 33 (1998).

[207] S. Tuukkanen, J. J. Toppari, A. Kuzyk, L. Hirviniemi, V. P. Hytönen, T. Ihalainen, P. Törmä, *Nano Lett.* **6**, 1339 (2006).

[208] S. Tuukkanen, A. Kuzyk, J. J. Toppari, H. Häkkinen, V. P. Hytönen, E. Niskanen, M. Rinkiö, P. Törmä, *Nanotechnology* **18**, 295204 (2007).

[209] P. Mignon, S. Loverix, J. Steyaert, P. Geerlings, *Nucleic Acids Res.* **33**, 1779 (2005).

[210] J. Regtmeier, R. Eichhorn, L. Bogunovic, A. Ros, D. Anselmetti, *Anal. Chem.* **82**, 7141 (2010).

[211] H. Zhao, *Phys. Rev. E* **84**, 021910 (2011).

[212] A. Cuervo, P. D. Dans, J. L. Carrascosa, M. Orozco, G. Gomila, L. Fumagalli, *Proc. Natl. Acad. Sci.* **111**, E3624 (2014).

[213] Y. Yang, K. Lao, D. Wilkins, A. Grisafi, M. Ceriotti, R. DiStassio Jr., *Sci. Data* **6**, 152 (2019).

[214] W. Cao, M. Chern, A. M. Dennis, K. A. Brown, *Nano Lett.* **19**, 5762 (2019).

[215] S. Mishra, A. K. Mondal, S. Pal, T. K. Das, E. Z. B. Smolinsky, G. Siligardi, R. Naaman, *J. Phys. Chem. C* **124**, 10776 (2020).

[216] K. Indumathi, A. Abiram, G. Praveena, *Mol. Phys.* **118**, e1584682 (2020).

[217] R. Cheng, J. Martens, T. D. Fridgen, *Phys. Chem. Chem. Phys.* **22**, 11546 (2020).

[218] G. Jung, S. Kasper, F. Schmid, *J. Electrochem. Soc.* **166**, B3194 (2019).

[219] F. van der Touw, M. Mandel, *Biophys. Chem.* **2**, 231 (1974).

[220] H. Washizu, K. Kikuchi, *J. Phys. Chem. B* **110**, 2855 (2006).

[221] S. Tomic, S. D. Babi, T. Vuletic, S. Krca, D. Ivankovic, L. Griparic, R. Podgornik, *Phys. Rev. E* **75**, 021905 (2007).

[222] N. F. Ramsey, *Phys. Rev.* **78**, 699 (1950).

[223] N. F. Ramsey, *Phys. Rev.* **77**, 567 (1950).

[224] N. F. Ramsey, *Phys. Rev.* **86**, 243 (1952).

[225] N. F. Ramsey, E. M. Purcell, *Phys. Rev.* **85**, 143 (1952).

[226] N. F. Ramsey, *Phys. Rev.* **91**, 303 (1953).

[227] B. Zúñiga-Gutiérrez, *Magnetic Shielding Spin-Spin Coupling Constant and Magnetiz-*

ability Tensors Calculated with Auxiliary Density Perturbation Theory, Ph.D. Thesis, Cinvestav (2011).

[228] E. Cheney, D. Kincaid, *Linear Algebra: Theory and Applications*, Jones and Bartlett Publishers (2009).