CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO
DEPARTAMENTO DE QUÍMICA

Métodos con bajo escalamiento computacional para la teoría del funcional de la densidad auxiliar

T E S I S

Que presenta

DANIEL MEJÍA RODRÍGUEZ

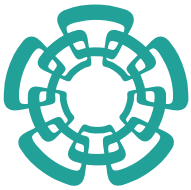
Para obtener el grado de

DOCTOR EN CIENCIAS

EN LA ESPECIALIDAD DE
CIENCIAS QUÍMICAS

Director de la Tesis:
Prof. Dr. Andreas M. Köster

México, D.F. NOVIEMBRE 2015



Center for Research and Advanced Studies
Of The National Polytechnic Institute

Zacatenco Campus
Chemistry Department

Low-order scaling methods for auxiliary density functional theory

A dissertation presented by
Daniel Mejía Rodríguez

In partial fulfillment of the requirements
for the degree of
Doctor of Science

In the specialty of
Chemical Science

Thesis advisor:
Prof. Dr. Andreas M. Köster

Mexico, D.F. November 2015

To my parents, whose education and words of encouragement have taken me to this place.

To my brother and sister, for supporting, helping and standing me.

To my lovely wife, for convincing me to pursue the Ph.D., for her love and support during all these years. We are going to be three now and many new exciting experiences will follow!

Acknowledgments

I WOULD LIKE TO EXPRESS MY GRATITUDE to my supervisor, Dr. Andreas M. Köster, whose expertise, understanding and patience added considerably to my graduate experience. His guidance helped me in all the time of research and writing of this thesis.

I THANK the other members of the committee: Dr. J. Vince Ortiz, Dr. Carlos Amador Bedolla, Dra. María de los Ángeles Paz Sandoval, Dr. Aarón Rojas Aguilar and Dr. Alberto Vela Amieva for their careful reading and insightful comments which helped to improve this thesis.

I THANK my fellow labmates and members of the Department of Chemistry for the stimulating discussions and all the fun we have had during the past five years.

I THANK the Center for Research and Advanced Studies of the National Polytechnic Institute (CINVESTAV) and the economical support of the National Council for Science and Techonology (CONACyT) through the Ph. D. fellowship No. 47922.

Contents

List of Tables

List of Figures

List of Abbreviations

ADFT	Auxiliary density functional theory
ADMM	Auxiliary density-matrix method
ADPT	Auxiliary density perturbation theory
AJD	Approximate joint diagonalization
ALDA	Adiabatic local density approximation
AO	Atomic orbital
ARI	Atomic resolution-of-the-identity
BLAS	Basis linear algebra subprograms
BOMD	Born-Oppenheimer molecular dynamics
Bz	η^6-benzyl
CADF	Concentric atomic density fitting
Cb	η^4-cyclobutadienyl
CCSD	Coupled-cluster with single and double excitations
CCSD(T)	Coupled-cluster with single, double and perturbative triple excitations
CLMO	Cholesky localized molecular orbitals
CMO	Canonical molecular orbital
CO	Cluster orbital
Cp	η^5-cyclopentadienyl
CPKS	Coupled-perturbed Kohn-Sham
CPSCF	Coupled-perturbed self-consistent field
CPU	Central processing unit
deMon2k	Computational chemistry code: density of Montreal
DF	Density fitting
DFA	Density functional approximation
DFT	Density functional theory
DIIS	Direct inversion of iterative subspace
DZ	Dunning's double-zeta augmented correlation consistent basis set
DZVP	Double-zeta valence polarized basis set
EN	Eirola-Nevanlinna
ER	Edminston-Ruedenberg
ERI	Electron repulsion integral
EXX	Exact exchange
FB	Foster-Boys
GAMESS	Computational chemistry code: General atomic and molecular electronic structure system
GEN-An	Automatically generated auxiliary function set of An quality
GGA	Generalized gradient approximation
GPT	Gaussian product theorem

GPU	Graphical processing unit
GTO	Gaussian-type orbital
HF	Hartree-Fock
HGTF	Hermite Gaussian-type function
KS	Kohn-Sham
LAPACK	Linear algebra package
LCGTO	Linear combination of Gaussian-type orbitals
LDA	Local density approximation
LDF	Local density fitting
LMO	Localized molecular orbital
LT	Laplace transform
MAD	Mean absolute deviation
MaxAD	Maximum absolute deviation
MD	Mean deviation
MKL	Intel® math kernel library
MO	Molecular orbital
MP2	Møller-Plesset second-order perturbation theory
PARI	Pair atomic resolution-of-the-identity
PES	Potential-energy surface
PM	Pipek-Mezey
PPP	Pariser-Parr-Pople model
QFMM	Quantum fast multipole method
QZ	Dunning's quadruple-zeta augmented correlation consistent basis set
RH	Roothaan-Hall
RI	Resolution-of-the-identity
RPA	Random phase approximation
SCF	Self-consistent field
SCP	Self-consistent perturbation
SOS	Scaled opposite-spin
SVD	Singular value decomposition
TF	Thomas-Fermi
TFD	Thomas-Fermi-Dirac
TFDW	Thomas-Fermi-Dirac-von Weizäcker
TZ	Dunning's triple-zeta augmented correlation consistent basis set
UEG	Uniform electron gas

Nomenclature

$\mathbf{A}$	Coulomb response matrix
c	Molecular orbital coefficient matrix
$\breve{c}$	Orthogonalized molecular orbital coefficient matrix
$\mathcal{E}$	Electronic energy
$\mathcal{E}_2$	Error of the variational fitting of the Coulomb potential
$\mathcal{E}_2^{EXX}$	Error of the variational fitting of the exact exchange potential
E_{DFJ}	Density-fitted Coulomb energy
E_{DFJK}	Density-fitted Coulomb and exact exchange energy
E_{DFK}	Density-fitted exact exchange energy
E_{EXX}	Exact exchange energy
E_{HF}	Hartree-Fock energy
E_{xc}	Exchange-correlation energy
$\mathbf{E}$	Identity matrix
$\mathbf{F}$	Exchange-correlation kernel matrix
G	Number of grid points
$\mathbf{G}$	Coulomb matrix
$\mathbf{G}_i$	Local Coulomb matrix for orbital i
$\mathbf{H}_0$	Approximated inverse matrix for the Eirola-Nevanlinna algorithm
$\mathbf{J}$	Coulomb vector
$\bar{k}$	Hermite Gaussian auxiliary function
$\mathbf{K}$	Kohn-Sham matrix
M	Number of auxiliary functions
M_{local}	Number of local auxiliary functions
N	Number of basis functions
N_{batch}	Number of integral batches
N_{CPU}	Number of CPUs in a parallel task
N_{leaf}	Maximum number of LMO centroids inside a leaf node
N_{local}	Number of local basis functions
N_{occ}	Number of occupied orbitals
N_{uno}	Number of unoccupied orbitals
$\mathbf{P}$	Density matrix
$\mathbf{R}$	Response matrix
$\mathbf{S}$	Overlap matrix
v_{xc}	Exchange-correlation potential
$\mathbf{W}$	Energy-weighted density matrix
$\mathbf{X}$	Exact exchange potential matrix
$\mathbf{x}$	Density fitting coefficient vector
$\mathbf{z}$	Exchange-correlation fitting coefficient vector

$\mathbf{Z}_i$	Inverse Cholesky factor of the local Coulomb matrix of orbital i
δ	Dirac delta
δ_{ij}	Kronecker delta
$\boldsymbol{\varepsilon}$	Diagonal Lagrange multiplier matrix
λ	Perturbation parameter
μ	Atomic orbital
Ψ	Total wavefunction
$\Psi_{\mathbf{R}}$	Electronic wavefunction
ψ	Molecular orbital
ρ	Electron density
$\tilde{\rho}$	Auxiliary density
Θ	Nuclear wavefunction
ζ	Gaussian function exponent
$\boldsymbol{\nabla}$	Gradient operator
∇^2	Laplacian operator

Related Work

Publications related to the thesis:

- Robust and Efficient Variational Fitting of Fock Exchange
  D. Mejía-Rodríguez, A. M. Köster
  J. Chem. Phys. **141**, 124114 (2014)

- QM/MM Calculations with deMon2k
  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. Rezác, F. Cailliez, A. de la Lande
  Molecules **20**, 4780 (2015)

- Robust and Efficient Auxiliary Density Perturbation Theory Calculations
  D. Mejía-Rodríguez, R. Delgado Venegas, P. Calaminici, A. M. Köster
  J. Chem. Theory Comput. **11**, 1493 (2015)

- Hybrid Functionals with Variationally Fitted Exact Exchange
  D. Mejía-Rodríguez, X. Huang, J. M. del Campo, A. M. Köster
  Adv. Quantum Chem. **71**, 41 (2015)

- Influence of Spin Multiplicity on the Melting of Na_{55}^{+}
  J. M. Vásquez-Pérez, G. U. Gamboa, D. Mejía-Rodríguez, A. Álvarez-Ibarra, G. Geudtner, P. Calaminici, A. M. Köster
  J. Phys. Chem. Lett. **6**, 4646 (2015)

- Second Analytic Derivatives in Auxiliary Density Functional Theory
  R. I. Delgado-Venegas, D. Mejía-Rodríguez, Roberto Flores-Moreno, P. Calaminici, Andreas M. Köster
  In preparation

Low-order scaling methods for auxiliary density functional theory

Resumen

Esta tesis presenta la derivación e implementación de dos metodologías con bajo escalamiento computacional usadas en el marco de la teoría del funcional de la densidad auxiliar (ADFT).

Primero, se presenta el desarrollo e implementación de la aproximación variacional local del intercambio exacto (LDF-EXX) para la evaluación rápida y eficiente de dicha contribución. Esta metodología permite cálculos Hartree-Fock y cálculos con funcionales híbridos en ADFT con rapidez extraordinaria. Aún más, la metodología LDF-EXX logra reducir el cuello de botella computacional de aproximaciones variacionales previamente reportadas. La metodología LDF-EXX resultante tiene una eficiencia computacional comparable a la teoría del funcional de la densidad de Kohn y Sham con un funcional basado en la aproximación del gradiente generalizado.

En segundo lugar, se presenta el desarrollo e implementación de un nuevo solucionador iterativo para el sistema de ecuaciones de respuesta que se origina en la teoría de perturbaciones de la densidad auxiliar (ADPT). Este nuevo método iterativo resuelve el cuello de botella computacional de los cálculos ADPT y permite el cómputo de propiedades de respuesta de nanosistemas, en tan solo unas horas, utilizando arquitecturas computacionales modestas.

Las contribuciones alcanzadas en esta tesis abren la posibilidad de cálculos de gran escala en tiempos razonables cuando se utilizan funcionales híbridos. Además, permiten el cómputo eficiente de propiedades de respuesta de nanosistemas. El potencial de éstas dos metodologías nuevas es demostrado a través de aplicaciones seleccionadas que incluyen problemas químicos actuales.

Low-order scaling methods for auxiliary density functional theory

Abstract

This thesis presents the derivation and implementation of two new low-order scaling methodologies to be used in the framework of auxiliary density functional theory (ADFT).

First, we present the development and implementation of the local density fitting exact exchange (LDF-EXX) approach for the rapid and efficient evaluation of exact exchange. This methodology enables Hartree-Fock and hybrid Kohn-Sham density functional theory calculations with remarkable speed-ups compared to standard exact exchange implementations. Furthermore, the here presented LDF-EXX methodology circumvents the bottlenecks of former density fitting and resolution-of-the-identity approximation to exact exchange, resulting in a much faster density fitting approach. The resulting LDF-EXX methodology has a computational efficiency comparable to density-fitting Kohn-Sham density functional theory calculations performed with a generalized-gradient-approximation functional.

Second, we present the development and implementation of a new iterative solver for the response equation system of auxiliary density perturbation theory (ADPT). This new iterative approach resolves the current bottleneck of ADPT calculations and enables the computation of molecular response properties of nanosystems in just a few hours on moderate parallel computational architectures.

The contributions achieved in this thesis open the avenue for large-scale all-electron calculations with hybrid functionals in reasonable timings. In addition, they enable the very fast computation of molecular properties of nanosystems. Selected applications of these to new methodologies to current chemical problems exhibit the potential of the here presented low-order scaling methods.

Quantum mechanics always seem to require infinitely many dimensions; I don't think I can cope with so many—I'm going to have about a 100 or so—that ought to be enough, don't you think?

Alan Turing

0 Introduction and Objectives

0.1 Introduction

The key role of modeling, theory and computation to enable the growth of Chemistry as a science has been recognized in the international community by the several Nobel Prizes in Chemistry awarded to theoretical developments, the last two being awarded in 1998 and 2013. Most modern chemical research cannot be performed without the inclusion of some computational component, whether it is a simple visualization of a molecule or a high-level quantum mechanical calculation. However, the inclusion of quantum mechanical calculations into mainstream Chemistry has not been easy or straightforward. Perhaps because, as noted by Mulliken and Roothaan,[1] there has been thousands of chemists that without any invocation to electronic structure calculations gathered an enormous amount

of knowledge and built the foundations of the chemical sciences. The difficulty to merge theory and computation into mainstream Chemistry has led to an apparent division in the science; in words of Longuet-Higgins[2]: "*It has always seemed to me that there are three kinds of Chemistry: experimental, theoretical and computational.*"

The division that Longuet-Higgins saw in Chemistry during the 1970's was not new. In fact, it might have it's origins some centuries ago and can be explained by the discussion about whether a mathematical formulation of Chemistry is possible, or even desirable. For instance, in 1786 Kant[3,4] wrote that Chemistry will never be a genuine science because it cannot be formulated in mathematical terms;[a] in the early 1800's Gay-Lussac[5] believed that "*…we are perhaps not far removed from time when we shall submit the bulk of chemical phenomena to computation.;*"[b] and Comte[6] argued that if mathematical doctrines would ever become predominant in Chemistry, this would mean the rapid decline of the science.[c] The division between those that wished for a mathematical theory of Chemistry and those who didn't remained several years more. With the advent of quantum mechanics around 1920's this debate heated up. In particular, Dirac's 1929 statement[7]

> *The general theory of quantum mechanics is now almost complete, the imperfections that still remain being in connection with the exact fitting in of the theory with relativity ideas. These gives rise to difficulties only when high-speed particles are involved,*

[a] Solange als noch für die chymischen Wirkungen der Materien aufeinander kein Begriff ausgefunden wird, der sich construieren läßt, d. i. kein Gesetz der Annäherung oder Entfernung der Theile angeben läßt, nach welchem etwa in Propertionen ihrer Dichtigkeiten u. d. g. ihre Bewegungen samt ihren Folgen sich im Raume a priori anschaulich machen und darstellen lassen (eine Forderung, die schwerlich jemals erfüllt werden wird), so kann Chymie nichts mehr als systematische Kunst, oder Experimentallehre, niemals aber eigentliche Wissenschaft werden, weil die Principien derselben blos empirisch sind und keine Darstellung a priori in der Anschauung erlauben, folglich die Grundsätze chymischer Erscheinungen ihrer Möglichkeit nach nicht im mindesten begreiflich machen, weil sie der Anwendung der Mathematik unfähig wird.

[b] J'espère donner par là une preuve de ce qu'ont avancé des chimistes trés-distingues, qu'on n'est peut-être pas éloigné de l'époque à laquelle on pourra soumettre au calcul la plupart des phénomènes chimiques.

[c] Toute tentative de faire rentrer les questions chimiques dans le domaine des doctrines mathématiques, doit être réputée jusqu'ici, et sans doute à jamais, profondément irrationnelle, comme étant antipathique à la nature des phénomènes…J'ai fait ressortit, dans le volume précédent, le tort général fait jusquu'ici à la physique par l'abus de l'analyse mathématique. Mais là, il ne s'agissait que de l'usage irréfléchi d'un instrument, qui, judicieusement dirigé, est susceptible, pour un tel ordre de recherches, d'une admirable efficacité. Ici, au contraire, on ne doit pas craindre de garantir que si, par une aberration heurusement presque impossible, l'emploi de l'analyse mathématique acquérait jaimais, en chimie, une semblable prépondérance, il déterminerait inévitablement, et sans aucune compensation, dans l'economie entière de cette science, une immense et rapide rétrogradation, en substitutant l'empire des conceptions vagues à celui des notions positives, et una facile verbiage algébrique à une laborieuse explorations des faits.

and are therefore of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions, in which it is, indeed, usually sufficiently accurate if one neglects relativity variation of mass with velocity and assumes only Coulomb forces between the various electrons and the nuclei. The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these equations leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

has generated a lot of controversy. For many, Dirac's claim is reductionist and completely wrong.[4,8,9] For many others, the statement is, as Parr and Crawford wrote: "*both the hope and despair of valence theoreticians […].*"[10] Certainly, Dirac was wrong to state that relativity has no importance in atomic and molecular structure, but what about the mathematical theory to describe the whole of chemistry? As Kutzelnigg noted, this is a nontrivial philosophical question.[4] However, it is a fact that quantum mechanics has enabled the computation and prediction of many properties of chemical interest. According to Mulliken and Roothaan, the slow acceptance of quantum mechanics followed from the fact that quantum mechanics was not successful to make quantitative predictions because the answers to the easy problems were already obtained in the laboratory and the more difficult problems were too complex, in terms of the mathematical analysis and the computational effort required, for quantum mechanics to provide an answer.[1] In Frederick Soddy's words:[11] "*…for the most part, chemistry is still too complicated a science for the theories to be a substitute for the facts as any real theory should be.*"

Nowadays, almost 60 years after Mulliken's and Roothaan's discussion and 90 years after Dirac's claim, the application of quantum mechanics to complex chemical phenomena is still a challenge. Even today, the simulation of a real chemical laboratory experiment by quantum mechanical methods still requires vast methodological improvements. To

put this in perspective, think about the Hartree-Fock (HF) method. HF is the simplest approach for solving the Schrödinger equation[12–16] and describes electron-electron interactions within a mean-field approach.[17–20] Even this approach scales conventionally as $\mathcal{O}(N^4)$, where $\mathcal{O}()$ denotes the order of the asymptotic scaling behavior. This means that to calculate a system 10 times larger than a reference one, the computational time required increases by a factor of 10^4 with respect to the time needed for the smaller reference system. The increase becomes even more dramatic if electron correlation effects, neglected in the HF approach, are accounted for because the scaling behavior is, at least, $\mathcal{O}(N^5)$. This was, and still is, the reason for the excitement density functional theory (DFT)[21] in the formulation of Kohn and Sham (KS-DFT)[22] generated in the quantum chemical community. It incorporates electron correlation without increasing the scaling. Nevertheless, the formal scaling of KS-DFT is still $\mathcal{O}(N^4)$.

To illustrate how prohibitive even an $\mathcal{O}(N^4)$ scaling is for the calculation of large systems, we can think about it in terms of "Moore's Law".[23,24] It is an empirical observation made by Gordon E. Moore in 1965, co-founder of Intel Corporation and Fairchild Semiconductor, that describes the doubling of components per integrated circuit every two years. Moore's Law has been astonishingly valid over the last decades. The factor of 10^4 for a 10-times larger system can be described as roughly 2^{13}, which would—with Moore's assumption—correspond to 26 years of computer development. In other words, one would need to wait 26 years for the computers to evolve to perform an HF calculation for a 10-times larger system within the same time frame as today for the reference one.[25]

Furthermore, the increase of power consumption has led to the end of frequency (processor speed) scaling as the dominant computer architecture paradigm since the early 2000's. Moore's law has been applicable since then by adding extra processors for parallel computing. However, the potential speed-up of a parallel computing platform is not linear. Instead, the small portion of a program that cannot be parallelized limit the overall speed-up as given by Amdahl's Law.[26] If α is the fraction of running time a program spends on

non-parallelizable parts, then the maximum speed-up is

$$S(P) = \lim_{P\to\infty} \frac{1}{\frac{1-\alpha}{P} + \alpha} = \frac{1}{\alpha} \,, \tag{1}$$

where P is the number of processors available. If only 1% of the program cannot be, or is not, parallelized, only a speed-up of 100 can be achieved regardless of the number of processors given to the task. This is clearly not an option for any enthusiastic researcher attempting to grasp deeper insights into molecular processes in chemistry, biochemistry, or material science.

The present work continues the effort of developing "*desirable approximate methods [...]*" that lower the computational demand of existing algorithms to make them applicable to larger systems in reasonable time frames. In particular, a new formulation for calculating the exact exchange contribution, which lowers the computational demand of traditional algorithms, is presented. Exact exchange is a key ingredient for HF, post-HF, as well as for some of the most popular Density Functional Approximations (DFA) used in DFT calculations. In addition, a new iterative solver for Auxiliary Density Perturbation Theory (ADPT), which enables property calculations on systems with several hundred of atoms in just hours, is also presented.

To this end, the thesis is organized as follows. In chapter 1 a general revision of quantum mechanics, with a particular focus on Kohn and Sham's formulation of Density Functional Theory (KS-DFT), is given. Chapter 2 presents the Auxiliary Density Functional Theory (ADFT) formalism, a very efficient alternative to KS-DFT. In chapter 3, the development of efficient algorithms for calculating exact exchange contributions in the framework of ADFT are presented. A new iterative solver for ADPT is derived and presented in chapter 4. Selected applications of the algorithms in chapters 3 and 4 are presented on chapter 5. Finally, the last chapter summarizes the progress achieved in this PhD and suggests some future developments that arise naturally as extension of the here presented algorithms.

0.2 Objectives

The main objective of this work was to continue the efforts of developing new all-electron first principle "*desirable approximate methods [...]*" by the development and implementation of two new low-order scaling methodologies. These new methodologies should allow the computation of larger systems in reasonable time frames. In order to achieve this goal, we proposed the following specific objectives:

1. Development and implementation of a new low-order scaling exact exchange approximation.
   (a) Development of the new low-order scaling exact exchange approximation.
   (b) Serial and parallel implementation of the exact exchange potential.
   (c) Serial and parallel implementation of the exact exchange energy gradients.
   (d) Validation, benchmarking and optimization of the new exact exchange implementation.
2. Development and implementation of a new low-order scaling ADPT solver.
   (a) Theoretical development of the new low-order scaling ADPT solver.
   (b) Serial and parallel implementation of the new ADPT solver.
   (c) Validation, benchmarking and optimization of the new ADPT solver implementation.

Gar manches rechnet Erwin schon
Mit seiner Wellenfunktion.
Nur wissen möcht man gerne wohl,
Was man sich dabei vorstell'n soll.

Erwin with his psi can do
Calculations quite a few
But one thing has not been seen:
Just what does psi really mean?

Erwin Hückel

1

Quantum Chemistry Fundamentals

1.1 The Schrödinger Equation

According to the postulates of quantum mechanics, the state of a system is fully described by a function $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, t)$. Ψ is called the wavefunction, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are the spatial and spin coordinates of particles $1, 2, \ldots, n$ that constitute the system and t is the time variable. The wavefunction $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, t)$ evolves in time according to the equation

$$\hat{H}\Psi = i\hbar\frac{\partial\Psi}{\partial t} \tag{1.1}$$

Equation (1.1) is the celebrated time-dependent Schrödinger equation introduced by Erwin Schrödinger in 1926.[12–16] The operator $\hat{H}$ is the Hamilton operator. The time-dependent Schrödinger equation can often be separated into equations for the time and space varia-

tion of the wavefunction. The time-independent wavefunction $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, satisfies the equation

$$\hat{H}\Psi = E\Psi \tag{1.2}$$

For an isolated system with N electrons and M nuclei the explicit form of the nonrelativistic Hamilton operator is given, in atomic units which will be used throughout this discussion, by:

$$\hat{H} = -\frac{1}{2}\sum_{i}^{N}\nabla_i^2 - \sum_{A}^{M}\frac{1}{2M_A}\nabla_A^2 + \sum_{i}^{N}\sum_{j>i}^{N}\frac{1}{|\,\mathbf{r}_i - \mathbf{r}_j\,|} - \sum_{i}^{N}\sum_{A}^{M}\frac{Z_A}{|\,\mathbf{r}_i - \mathbf{R}_A\,|} + \sum_{A}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\,\mathbf{R}_A - \mathbf{R}_B\,|} \tag{1.3}$$

where $\mathbf{r}$ and $\mathbf{R}$ denote the spatial coordinates of the electrons and nuclei, respectively. The atomic mass of nucleus A is M_A and it's charge Z_A. The physical meaning of the terms on the right-hand-side (rhs) of Equation (1.3) are, respectively: kinetic energy of the N electrons, kinetic energy of the M nuclei, electrostatic repulsion energy between the $N(N-1)/2$ electron pairs, electrostatic attraction between the N electrons and the M nuclei, and electrostatic repulsion between the $M(M-1)/2$ nuclei pairs. It is therefore customary to express Equation (1.3) as:

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{ne} + \hat{V}_{nn} \tag{1.4}$$

The term $\hat{V}_{ne}$ couples electronic and nuclear motions, complicating the solution of Equation (1.2). 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.[27] The separation is based on the fact that nuclei are several thousand times heavier than electrons. In a classical dynamical sense, the electrons can be regarded as particles that follow the nuclear motion adiabatically, meaning that they are dragged along with the nuclei without requiring relaxation time.[28] In terms of Equation (1.2) this means that

we can assume a quasi-separable ansatz of the form

$$\Psi(\mathbf{x}, \mathbf{R}) = \Psi_{\mathbf{R}}(\mathbf{x}) \cdot \Theta(\mathbf{R}) \tag{1.5}$$

The quantum mechanical consequence of the mass difference is that the nuclear components of the wavefunction are much more localized in space. Thus, in a small space around the nuclei where the nuclear wavefunction is different from zero it rises much more steeply than the electronic one, which means that $\boldsymbol{\nabla}_A \Theta(\mathbf{R}) \gg \boldsymbol{\nabla}_A \Psi_{\mathbf{R}}(\mathbf{x})$, from which we may approximate

$$\hat{T}_n \Psi(\mathbf{x}, \mathbf{R}) = -\sum_A^M \frac{1}{2M_A} \left[\Psi_{\mathbf{R}}(\mathbf{x}) \nabla_A^2 \Theta(\mathbf{R}) + 2\boldsymbol{\nabla}_A \Psi_{\mathbf{R}}(\mathbf{x}) \cdot \boldsymbol{\nabla}_A \Theta(\mathbf{R}) + \Theta(\mathbf{R}) \nabla_A^2 \Psi_{\mathbf{R}}(\mathbf{x}) \right] \tag{1.6}$$

$$\approx -\Psi_{\mathbf{R}}(\mathbf{x}) \sum_A^M \frac{1}{2M_A} \nabla_A^2 \Theta(\mathbf{R}) \tag{1.7}$$

Substitution of Equation (1.5) into Equation (1.2) and using approximation (1.7) yields:

$$\frac{[\hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}]\Psi_{\mathbf{R}}(\mathbf{x})}{\Psi_{\mathbf{R}}(\mathbf{x})} = E - \frac{[\hat{T}_N + \hat{V}_{nn}]\Theta(\mathbf{R})}{\Theta(\mathbf{R})} \tag{1.8}$$

Thus, it is clear that the left-hand-side (lhs) can only be a function of $\mathbf{R}$ alone

$$\begin{aligned} \frac{[\hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}]\Psi_{\mathbf{R}}(\mathbf{x})}{\Psi_{\mathbf{R}}(\mathbf{x})} &= \mathcal{E}(\mathbf{R}) \\ [\hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}]\Psi_{\mathbf{R}}(\mathbf{x}) &= \mathcal{E}(\mathbf{R})\Psi_{\mathbf{R}}(\mathbf{x}) \\ \hat{H}_e \Psi_{\mathbf{R}}(\mathbf{x}) &= \mathcal{E}(\mathbf{R})\Psi_{\mathbf{R}}(\mathbf{x}) \end{aligned} \tag{1.9}$$

Equation (1.9) is known as the electronic Schrödinger equation, which yield a set of orthonormalized eigenfunctions $\Psi_{\mathbf{R}}(\mathbf{x})$ with corresponding eigenvalues $\mathcal{E}(\mathbf{R})$. For each solution there will be a corresponding nuclear Schrödinger equation. For the electronic

ground state, that is of importance here, the nuclear Schrödinger equation is given by:

$$[\hat{T}_n + \hat{V}_{nn} + \mathcal{E}_0(\mathbf{R})]\Theta(\mathbf{R}) = E\Theta(\mathbf{R})$$
$$\hat{H}_n\Theta(\mathbf{R}) = E\Theta(\mathbf{R}) \tag{1.10}$$

The usual procedure to solve Equation (1.10) is to solve first Equation (1.9), then substitute the electronic energy into the nuclear Schrödinger equation and solve it.

Varying the nuclear positions maps out a multi-dimensional potential energy surface (PES). In Born-Oppenheimer molecular dynamics (BOMD), the motions of the nuclei in this potential are usually solved using Newtonian mechanics. Note that each electronic state will give rise to a different PES, and that crossing between two surfaces is not allowed by the adiabatic theorem.[29] To simplify notation we drop the parametric dependency of the electronic wavefunction on $\mathbf{R}$ from now on. Furthermore, the term $\hat{V}_{nn}$ will be included in the electronic Hamiltonian since it only adds a constant that shifts the electronic energy as:

$$E_e = \mathcal{E} + V_{nn} \tag{1.11}$$

Even with the BO approximation, the Schrödinger equation cannot be solved analytically for a system with two or more electrons. Thus, only approximate solutions are available for most chemical systems of interest. Several methods have been developed for this purpose,[28] the most relevant being the ones based on the Rayleigh-Ritz variational method.[30,31] The variational principle in quantum mechanics states that any approximate wavefunction will always have an energy expectation value that is above that of the ground state energy[32,33]

$$E_e[\Psi_t] \geqslant E_e[\Psi_0] \tag{1.12}$$

where Ψ_t and Ψ_0 denote a trial and ground state wavefunctions, respectively. In other words, the energy expectation value of any appropriate trial wavefunction will provide an upper bound to the exact ground state energy. 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.

1.2 Density Functional Theory

Thomas and Fermi suggested the use of the electronic density $\rho(\mathbf{r})$ instead of the wavefunction $\Psi(\mathbf{x})$ to calculate atomic properties.[34,35] Further works by Dirac[36] as well as Wigner and Seitz[37,38] improved the model by introducing a local expression for the exchange potential. Several years later, Slater introduced the idea of approximating the Fock exchange operator of the Hartree-Fock method by an average local potential based on the free-electron gas model.[39] The result was an exchange potential expressed solely in terms of $\rho(\mathbf{r})$. Further development led to the so-called $X\alpha$ methodology of Slater and Johnson.[40] All these methods suggested that a theory based on $\rho(\mathbf{r})$ instead of $\Psi(\mathbf{x})$ could actually be accurate enough to calculate molecules. The main advantage of such a methodology comes from the fact that the electronic density $\rho(\mathbf{r})$ depends only on 3 instead of $3N$ spatial variables. In terms of the wavefunction, $\rho(\mathbf{r})$ is defined as:

$$\rho(\mathbf{r}) = N \int \dots \int |\Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 \, d\mathbf{r}_2 \, d\mathbf{r}_3 \dots d\mathbf{r}_N \tag{1.13}$$

Thus, the electronic density is obtained by integrating $N-1$ electrons out of the wavefunction. It represents the probability of finding an electron in position $\mathbf{r}$ when the other $N-1$ electrons are in arbitrary positions. In addition of being a much more simpler mathematical entity, $\rho(\mathbf{r})$ is an observable, making it more attractive than the rather mysterious $\Psi(\mathbf{x})$. Nonetheless, all the methods lacked a mathematically rigorous foundation and some of them failed to describe very "simple" phenomena such as chemical bonding.[41] A rigorous mathematical foundation for an *ab initio* theory based solely on $\rho(\mathbf{r})$ was first given by Hohenberg and Kohn in 1964, starting what we know today as Density Functional Theory.[21] 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 electronic density $\rho(\mathbf{r})$, apart from a trivial additive constant.*

The proof for the first Hohenberg-Kohn theorem is very simple. Assume that there exist two potentials $v(\mathbf{r})$ and $v'(\mathbf{r})$ differing by more than a constant and giving rise to the same ground-state density $\rho(\mathbf{r})$. Obviously, $v(\mathbf{r})$ and $v'(\mathbf{r})$ correspond to two different Hamilton operators $\hat{H}$ and $\hat{H}'$, which in turn give rise to two different wavefunctions $\Psi(\mathbf{x})$ and $\Psi'(\mathbf{x})$. Because of the variational principle the following inequality, written in Dirac's notation,[42] must hold

$$E_0 = \langle \Psi | \hat{H} | \Psi \rangle \leqslant \langle \Psi' | \hat{H} | \Psi' \rangle \tag{1.14}$$

Assuming that the ground state is non-degenerate, the inequality strictly holds. Because we have identical ground-state densities for the two Hamilton operators by construction, we can rewrite Equation (1.14) as

$$E_0 < \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle = E_0' + \int \left[v(\mathbf{r}) - v'(\mathbf{r}) \right] \, \rho(\mathbf{r}) \, \mathrm{d}\mathbf{r} \tag{1.15}$$

Similarly, taking $\Psi(\mathbf{x})$ as the trial wavefunction for $\hat{H}'$ yields

$$E_0' < \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle = E_0 + \int \left[v'(\mathbf{r}) - v(\mathbf{r}) \right] \, \rho(\mathbf{r}) \, \mathrm{d}\mathbf{r} \tag{1.16}$$

Finally, adding Equations (1.15) and (1.16) we obtain

$$E_0 + E_0' < E_0' + E_0 \tag{1.17}$$

which is clearly a contradiction. Thus, the theorem has been proven by *reductio ad absurdum*. This contradiction confirms that, for v-representable densities, the knowledge of the ground-state density uniquely determines the external potential of a non-degenerate quantum mechanical system. The following mapping can be defined based on the first

Hohenberg-Kohn theorem:

$$\rho(\mathbf{r}) \mapsto N,\ v(\mathbf{r}) \mapsto \hat{H} \mapsto \Psi(\mathbf{x})[\rho] \mapsto E[\rho] \tag{1.18}$$

Therefore, the ground-state wavefunction and the corresponding energy can be expressed as a functional of $\rho(\mathbf{r})$:

$$E_0[\rho] = F[\rho] + \int v(\mathbf{r})\,\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} = \langle \Psi_0[\rho]|\hat{H}|\Psi_0[\rho]\rangle \tag{1.19}$$

where $F[\rho]$ is the universal Hohenberg-Kohn functional given by

$$F[\rho] = \langle \Psi_0|\hat{T}_e|\Psi_0\rangle + \langle \Psi_0|\hat{V}_{ee}|\Psi_0\rangle \tag{1.20}$$

The name universal arises because $F[\rho]$ does not depend on the external potential $v(\mathbf{r})$ and, therefore, it is a universal functional of $\rho(\mathbf{r})$.[43]

Second Hohenberg-Kohn theorem: *For a positive semi-definite trial density $\rho_t(\mathbf{r})$ that integrates to the number of electrons of the system, $E[\rho_t] \geqslant E_0$, where E_0 is the ground state energy of the system.*

For a trial density $\rho_t(\mathbf{r})$ that is not the ground-state density, the first Hohenberg-Kohn theorem states that $\rho_t(\mathbf{r})$ corresponds to a different external potential $v_t(\mathbf{r})$, and, therefore, to a different wavefunction $\Psi_t(\mathbf{x})$. If we use $\rho_t(\mathbf{r})$ as trial for a problem having external potential $v(\mathbf{r})$, it follows from the variational principle:

$$E[\rho_t] = F[\rho_t] + \int v(\mathbf{r})\,\rho_t(\mathbf{r})\,\mathrm{d}\mathbf{r} = \langle \Psi_t|\hat{H}|\Psi_t\rangle > E_0 \tag{1.21}$$

This theorem establishes the variational principle for DFT. An electronic density different from the non-degenerate ground-state one will provide an upper bound for the ground-state energy of a given external potential.

More general derivations of the Hohenberg-Kohn theorems, that circumvent the v-rep-

resentable assumption and extend the applicability of DFT, have been given by Levy[44] with the so-called constrained-search formulation, and by Lieb[45] with the convex-conjugate (Legendre transform) formulation.[46] DFT has also been extended to excited-states in both time-dependent[47–49] and time-independent[50–55] formulations.

If we would knew the exact universal functional $F[\rho]$, 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 closed form.[43] In fact, none of the formulations of Hohenberg-Kohn, Levy or Lieb give even a hint for $F[\rho]$. To proceed, the energy functional defined in Equation (1.19) can be rewritten as

$$\begin{aligned} E[\rho] &= F[\rho] + \int v(\mathbf{r})\,\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ &= T[\rho] + V_{ee}[\rho] + \int v(\mathbf{r})\,\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ &= T[\rho] + J[\rho] + V_{ee}^{nc}[\rho] + \int v(\mathbf{r})\,\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{aligned} \tag{1.22}$$

where $T[\rho]$ is the kinetic energy and $V_{ee}[\rho]$ is the electron-electron interaction energy. $V_{ee}[\rho]$ can be further split into a classical part, $J[\rho]$, and a non-classical part, $V_{ee}^{nc}[\rho]$. The problem with the density functional in Equation (1.22) is how to calculate $T[\rho]$ and $V_{ee}^{nc}[\rho]$. The traditional Thomas-Fermi (TF) model ignores $V_{ee}^{nc}[\rho]$ and takes $T[\rho]$ by applying locally the uniform electron gas (UEG) expression for the kinetic energy:[34,35]

$$E_{TF}[\rho] = T_{TF}[\rho] + J[\rho] + \int v(\mathbf{r})\,\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{1.23}$$

The TF model fails to predict the correct behavior for atomic densities, gives total atomic energies 10 to 50 % too low[43] and fails to predict binding of molecules.[41,56,57]

One step further was taken by Dirac who introduced $V_{ee}^{nc}[\rho]$ as the exchange energy of an UEG[36] given rise to the TF-Dirac (TFD) model. However, the inclusion of exchange did not improve the description of the density and made the calculated energies even worse.[43] For this reason von Weizsäcker[58] suggested a gradient correction of $T_{TF}[\rho]$ which remedies three defects of the TF and TFD models: (i) $\rho(\mathbf{r})$ is finite at the nuclei, (ii) bind-

ing of atoms occurs and negative ions are stable and, (iii) $\rho(\mathbf{r})$ has exponential falloff for neutral atoms and molecules.[57] Although the TFD-von Weizsäcker (TFDW) model shows a qualitative improvement, it is still far from being a quantitative model to describe chemical phenomena. Higher order corrections, up to fourth order, provide higher accuracy but still describe only an average behavior of the electrons.[43]

1.3 The Kohn-Sham Methodology

Kohn and Sham gave a route to overcome, at least partially, the extremely difficult task of finding an accurate enough kinetic energy functional $T[\rho]$.[22] 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, $T_c[\rho]$,

$$T[\rho] = T_s[\rho] + T_c[\rho] \tag{1.24}$$

The subscripts s and c stand for *single-particle* and *correlation*, respectively.[59] In the non-interacting system, the total kinetic energy $T_s[\rho]$ is just the sum of the single-particle kinetic energies:

$$\begin{aligned} T_s[\rho] &= \langle \Psi[\rho] | \hat{T} | \Psi[\rho] \rangle \\ &= \sum_{i}^{\text{occ}} \langle \psi_i | \hat{T} | \psi_i \rangle \end{aligned} \tag{1.25}$$

where $\Psi[\rho]$ is the Slater determinant[60] forming the density $\rho(\mathbf{r})$, and $\psi_i(\mathbf{r})$ are the single-particle orbitals of the non-interacting system. Using Equations (1.22) and (1.24) we can rewrite the energy functional as

$$\begin{aligned} E[\rho] &= T_s[\rho] + T_c[\rho] + J[\rho] + V_{ee}^{nc}[\rho] + \int v(\mathbf{r})\, \rho(\mathbf{r})\, \mathrm{d}\mathbf{r} \\ &= T_s[\rho] + J[\rho] + E_{\text{xc}}[\rho] + V[\rho] \end{aligned} \tag{1.26}$$

where

$$V[\rho] = \int v(\mathbf{r})\, \rho(\mathbf{r})\, \mathrm{d}\mathbf{r} \tag{1.27}$$

and $E_{xc}[\rho]$ is the *exchange-correlation functional* that contains $T_c[\rho]$ and $V_{ee}^{nc}[\rho]$. Equation (1.26) is formally exact but, unfortunately, $E_{xc}[\rho]$ remains unknown. The practical advantage of writing the energy functional as Equation (1.26) is that $E_{xc}[\rho]$ is typically much smaller than the known terms $T_s[\rho]$, $J[\rho]$ and $V[\rho]$, thus, reasonably simple approximations for $E_{xc}[\rho]$ may provide accurate enough results for $E[\rho]$.[59]

Since $T_s[\rho]$ is not an explicit functional of $\rho(\mathbf{r})$, Equation (1.26) cannot be directly minimized. 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

$$\frac{\delta E[\rho]}{\delta \rho(\mathbf{r})} = \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta V[\rho]}{\delta \rho(\mathbf{r})} \equiv 0 \tag{1.28}$$

The functional derivative $\frac{\delta V[\rho]}{\delta \rho(\mathbf{r})}$ yields the external potential, v; the term $\frac{\delta J[\rho]}{\delta \rho(\mathbf{r})}$ yields the Hartree (Coulomb) potential, v_H; 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}. Consider now a system of non-interacting particles moving in a potential $v_s(\mathbf{r})$. For this system, the minimization condition is just

$$\frac{\delta E_s[\rho_s]}{\delta \rho_s(\mathbf{r})} = \frac{\delta T_s[\rho_s]}{\delta \rho_s(\mathbf{r})} + \frac{\delta V_s[\rho_s]}{\delta \rho_s(\mathbf{r})} \equiv 0 \tag{1.29}$$

Comparing Equations (1.28) and (1.29) we find that both minimizations have the same solution, $\rho_s(\mathbf{r}) \equiv \rho(\mathbf{r})$, if:

$$\frac{\delta V_s[\rho]}{\delta \rho(\mathbf{r})} = \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta V[\rho]}{\delta \rho(\mathbf{r})} \tag{1.30}$$

i.e.

$$v_s(\mathbf{r}) = v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) + v(\mathbf{r}) \tag{1.31}$$

Consequently, one can calculate the density of the interacting system with external potential $v(\mathbf{r})$ by solving the equations of a non-interacting system with external potential $v_s(\mathbf{r})$.

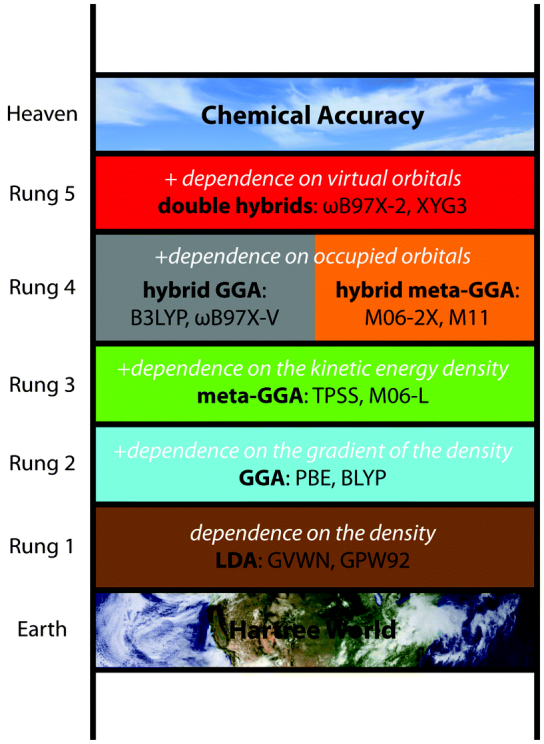


Figure 1.1: Perdew's Jacob's Ladder of density functional approximations. Reproduced from reference [61] with permission from the PCCP Owner Societies.

To this end, Kohn and Sham represented the non-interacting wavefunction as a single Slater determinant and obtained $T_s[\rho]$ through the Kohn-Sham (KS) orbitals $\psi_i(\mathbf{r})$ in the same way as in Equation (1.25). Solving the set of single-particle Schrödinger equations

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{1.32}$$

yields orbitals that reproduce the density of the original system

$$\rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \tag{1.33}$$

Since $v_s(\mathbf{r})$ depends on $\rho(\mathbf{r})$, the problem of solving Equations (1.32), known as KS equations, is a nonlinear one. The usual way of solving this problem will be discussed in the next chapter.

KS-DFT is astounding in its simplicity, yet it delivers, in principle, the exact density and exact total energy of any interacting, correlated electronic system. Everything hinges

on the functional $E_{xc}[\rho]$ and its functional derivative $v_{xc}[\rho]$.[62] Due to the central role of KS-DFT in electronic structure theory and material sciences[33], and the dependence of KS-DFT accuracy on $E_{xc}[\rho]$, authors of successful exchange-correlation energy functionals have some of the most cited papers in both physics and chemistry.[62,63] In order to give a quick overview of existing $E_{xc}[\rho]$, we turn to Perdew's Jacob's Ladder.[64,65] Perdew proposed a hierarchy of density functional approximations (DFA) for $E_{xc}[\rho]$ ordered by the use of increasingly complex ingredients to construct $E_{xc}[\rho]$. Figure 1.1 shows a pictorial representation of Perdew's Jacob's ladder, where the lower end is the "Hartree World" and the upper one is the "Heaven of Chemical Accuracy". Higher levels usually—but not always—yield more accurate results and are computationally more demanding than the lower ones. The first rung is the Local Density Approximation (LDA) and uses only $\rho(\mathbf{r})$ as its ingredient; the second rung is the generalized gradient approximation (GGA), which adds the density gradient, $\boldsymbol{\nabla}\rho(\mathbf{r})$, to its ingredients; the third rung is known as meta-GGA and adds the Laplacian of the density, $\nabla^2\rho(\mathbf{r})$, and the kinetic energy density, $\tau(\mathbf{r})$, or at least one of them, to its ingredients; the fourth rung has been named hyper-GGA and includes the exact exchange energy density; and the fifth rung adds exact partial correlation. Nowadays, several of the most popular DFAs belong to the second, third and fourth rungs. Many highly accurate functionals are of the meta-GGA type and, thus, their calculation is computationally expensive compared to a standard LDA or GGA functional. However, the simplicity and accuracy achieved by some second rung functionals attracted the attention of quantum chemists to KS-DFT.[62] Therefore, some groups are still devoted to the development of accurate, and simple, GGA functionals. For example, the Vela-Medel-Trickey (VMT)[66] functional improves substantially the energetics of the widely used Perdew-Burke-Ernzerhof (PBE)[67] form while being as simple in form. Further improvements may be achieved by imposing additional physically meaningful constraints like in the VT{8,4} functional[68] which imposes the large s constraint:

$$\lim_{s\to\infty} s^{1/2} F_{xc}(\rho, s) < \infty \tag{1.34}$$

where $F_{xc}(\rho, s)$ is the enhancement factor and s is the dimensionless gradient. Besides energetics, this efforts have also been directed to improve the calculation of properties. In particular, the semi-empirical functionals of Keal and Tozer KT1, KT2 and KT3,[69,70] were developed to improve nuclear magnetic resonance (NMR) calculations without affecting the geometrical and energetical performance. More recently, Carmona *et. al*[71] developed a GGA functional which has a correct asymptotic potential (CAP) and improves the calculation of molecular polarizabilities.

The developments of these, and many other, highly accurate DFAs have enabled an exponentially growing attention for DFT, to the point where KS-DFT is becoming the standard tool in electronic structure theory calculations.[33]

Looking toward the future, it seems certain that colossal rewards lie ahead from large-scale quantum-mechanical calculations of the structure of matter.

R. S. Mulliken and C. C. J. Roothaan

2

Auxiliary Density Functional Theory

2.1 The LCGTO Formalism

The KS equations, Equation (1.32), can be obtained by varying the Kohn-Sham energy expression, Equation (1.26), subject to the orthonormality constraint

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} \tag{2.1}$$

where δ_{ij} is the Kronecker delta. In order to perform such variation it is convenient to expand the molecular orbitals, $\psi_i(\mathbf{r})$, in a linear combination of atomic orbitals (LCAO):

$$\psi_i(\mathbf{r}) = \sum_{\mu} c_{\mu i}\, \mu(\mathbf{r}) \tag{2.2}$$

Here, $\mu(\mathbf{r})$ represents an atomic orbital or, more general, a basis function, and $c_{\mu i}$ a molecular orbital (MO) coefficient. In deMon2k[72–74] 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). An unnormalized Cartesian GTO is given by:[75]

$$\mu(\mathbf{r}) = (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} \sum_{k}^{K_\mu} d_k e^{\zeta_k (\mathbf{r}-\mathbf{A})^2} \tag{2.3}$$

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_μ, 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 under which minimization of Equation (1.26) is performed. Using the explicit form for all the known terms in Equation (1.26) and assuming a closed-shell system yields:

$$E[\rho] = -\sum_{i}^{\text{occ}} \langle \psi_i \,|\, \nabla^2 \psi_i \rangle - \sum_{A} \int \frac{Z_A}{|\,\mathbf{r} - \mathbf{A}\,|}\, \rho(\mathbf{r}) \mathrm{d}\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\, \rho(\mathbf{r}_2)}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|} \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2 + E_{\text{xc}}[\rho] \tag{2.4}$$

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

$$\rho(\mathbf{r}) = 2 \sum_{i}^{\text{occ}} |\,\psi_i\,|^2 = 2 \sum_{i}^{\text{occ}} \sum_{\mu,\nu} c_{\mu i}\, c_{\nu i}\, \mu(\mathbf{r})\, \nu(\mathbf{r}) = \sum_{\mu,\nu} P_{\mu\nu}\, \mu(\mathbf{r})\, \nu(\mathbf{r}) \tag{2.5}$$

where we have introduced the density matrix, $\mathbf{P}$, with elements:

$$P_{\mu\nu} = 2 \sum_{i}^{\text{occ}} c_{\mu i}\, c_{\nu i} \tag{2.6}$$

Substituting Equations (2.2), (2.5) and (2.6) into Equation (2.4) yields

$$E[\rho] = -\frac{1}{2} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu \,|\, \nabla^2 \nu \rangle - \sum_{\mu,\nu} \sum_{A} P_{\mu\nu} \left\langle \mu \left| \frac{Z_A}{|\,\mathbf{r} - \mathbf{A}\,|} \right| \nu \right\rangle + \frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \iint \frac{\mu(\mathbf{r}_1)\, \nu(\mathbf{r}_1)\, \sigma(\mathbf{r}_2) \tau(\mathbf{r}_2)}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|} \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2 + E_{\text{xc}}[\rho] \tag{2.7}$$

The first two-terms of Equation (2.7) 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 the notation, we introduce the core Hamiltonian matrix, $\mathbf{H}$, with elements

$$H_{\mu\nu} = -\frac{1}{2}\langle \mu \,|\, \nabla^2 \nu \rangle - \sum_A \left\langle \mu \left| \frac{Z_A}{|\,\mathbf{r} - \mathbf{A}\,|} \right| \nu \right\rangle \tag{2.8}$$

and a shorthand notation for the ERIs

$$\langle \mu\nu \| \sigma\tau \rangle = \iint \frac{\mu(\mathbf{r}_1)\,\nu(\mathbf{r}_1)\,\sigma(\mathbf{r}_2)\tau(\mathbf{r}_2)}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|} \mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{2.9}$$

to rewrite Equation (2.7) as

$$E[\rho] = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu \| \sigma\tau \rangle + E_{\mathrm{xc}}[\rho] \tag{2.10}$$

Note that the notation for the ERIs differs from the conventional Dirac notation used for the rest of the integrals. In the ERI notation the double vertical bar represents the two-electron Coulomb 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 the functions 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. Imposing the orthonormality constraint, Equation (2.1), in the LCGTO formalism leads to the Lagrange functional

$$\mathcal{L}[\rho; \mathbf{c}] = E[\rho] - 2 \sum_{i,j}^{\text{all}} \varepsilon_{ij} \left(\sum_{\mu,\nu} c_{\mu i} S_{\mu\nu} c_{\nu j} - \delta_{ij} \right) \tag{2.11}$$

The variation of the Lagrange functional,

$$\frac{\partial \mathcal{L}[\rho;\mathbf{c}]}{\partial c_{\mu i}} = \frac{\partial E[\rho]}{\partial c_{\mu i}} - 4\sum_{j}^{\text{occ}}\sum_{\nu} S_{\mu\nu}c_{\nu j}\varepsilon_{ji} \tag{2.12}$$

$$= 4\sum_{\nu}\left(H_{\mu\nu} + \sum_{\sigma,\tau} P_{\sigma\tau}\langle\mu\nu\|\sigma\tau\rangle + \langle\mu|v_{\text{xc}}\,|\nu\rangle\right)c_{\nu i} - 4\sum_{j}^{\text{all}}\sum_{\nu} S_{\mu\nu}c_{\nu j}\varepsilon_{ji} \quad , \tag{2.13}$$

must vanish at a stationary point. To obtain these equations the variation of $E_{\text{xc}}[\rho]$ is performed using the chain rule

$$\frac{\partial E_{\text{xc}}[\rho]}{\partial c_{\mu i}} = \int \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})}\frac{\partial\rho(\mathbf{r})}{\partial c_{\mu i}}\mathbf{dr} = 4\sum_{\nu} c_{\nu i}\int v_{\text{xc}}[\rho;\mathbf{r}]\,\mu(\mathbf{r})\,\nu(\mathbf{r})\,\mathbf{dr} \tag{2.14}$$

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

$$K_{\mu\nu} \equiv \frac{\partial E[\rho]}{\partial P_{\mu\nu}} = H_{\mu\nu} + \sum_{\sigma,\tau} P_{\sigma\tau}\langle\mu\nu\|\sigma\tau\rangle + \langle\mu|v_{\text{xc}}[\rho]\,|\nu\rangle \tag{2.15}$$

Substitution of Equation (2.15) into Equation (2.13) under the minimization condition yields

$$\mathbf{K}\mathbf{c}_i = \sum_{j}^{\text{all}} \mathbf{S}\mathbf{c}_j\varepsilon_{ji} \tag{2.16}$$

Which is a generalized eigenvalue equation.[76] There is one Equation of the form (2.16) for each molecular orbital. Collecting all equations into a single matrix equation yields

$$\mathbf{K}\mathbf{c} = \mathbf{S}\mathbf{c}\boldsymbol{\varepsilon} \tag{2.17}$$

This set of equations have the same form as the famous Roothaan-Hall (RH) equations[77,78] appearing in HF. Here, $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{\text{occ}}, \ldots, \mathbf{c}_{\mathbf{all}})$, is a square matrix composed from all the occupied molecular coefficient vectors as well as the virtual molecular orbital coefficients vectors. It is important to note that Equation (2.17) is a nonlinear generalized eigenvalue equation because the KS matrix depends on, at least, the occupied subspace of

$\mathbf{c}$. Therefore, Equation (2.17) has to be solved iteratively starting from an educated guess for $\mathbf{c}$. It is also important to note that $\mathbf{K}$ is invariant under separate unitary transformations of the occupied and unoccupied subspaces of $\mathbf{c}$. This property follows from the fact that the electronic density and, hence, the density matrix are invariant under such transformations

$$\mathbf{P} = 2\mathbf{c}_{\text{occ}}\mathbf{c}_{\text{occ}}^T \tag{2.18}$$

$$= 2\mathbf{c}_{\text{occ}}\mathbf{U}_{\text{occ}}\mathbf{U}_{\text{occ}}^T\mathbf{c}_{\text{occ}}^T \tag{2.19}$$

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 ε must have, in general, the following block diagonal form:

$$\varepsilon = \begin{pmatrix} \varepsilon_{\mathbf{occ}} & 0 \\ 0 & \varepsilon_{\mathbf{uno}} \end{pmatrix} \tag{2.20}$$

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

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{\text{occ}} & 0 \\ 0 & \mathbf{U}_{\text{uno}} \end{pmatrix} \tag{2.21}$$

such that $\mathbf{U}^T\varepsilon\mathbf{U}$ 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, ε is a diagonal matrix.

2.1.1 Computational scaling of the KS-DFT LCGTO method

Let us now analyze the computational complexity of the RH method for solving the KS equations in the LCGTO ansatz. The computation of the full core Hamiltonian matrix $\mathbf{H}$

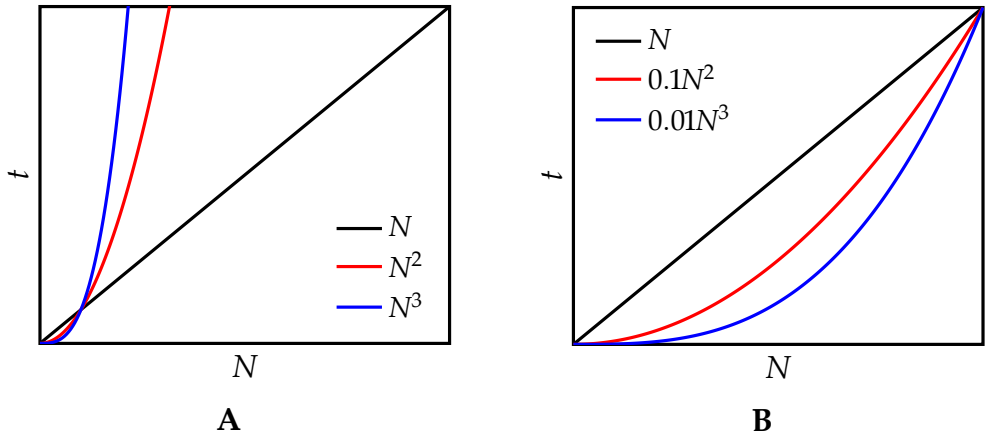


Figure 2.1: Influence of the prefactor on the relative speed of algorithms with different scalings.

scales as N^2, where N is the number of basis functions, because there are a total of N^2 elements that need to be computed. 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 contribution to the KS matrix has a formal N^4 scaling, because there are a total of N^4 ERIs. This contribution is not constant because it depends on $\mathbf{P}$. The exchange-correlation contribution has a formal $N^2 \times G$ scaling, where G is the number of grid points necessary for the numerical integration needed to compute the exchange-correlation contribution.

Besides the construction of $\mathbf{K}$, several matrix operations are needed in order to solve the RH generalized eigenvalue equation. These operations have a formal scaling of N^3 but can be performed, up to a certain basis set size, in a very efficient manner using optimized computational libraries for such operations. Many of these libraries are based on the Basic Linear Algebra Subroutines (BLAS)[79–86] and the Linear Algebra Package (LAPACK),[87] and are machine-specific optimized, like Intel's Math Kernel Library (MKL) and AMD's Core Math Library (ACML).

The scaling prefactor is a very important aspect to take into account when the asymptotic region of an algorithm has not yet been reached. Figure (2.1) shows two fictitious examples where the timings of three different modules with linear (—), quadratic (—) and cubic (—) scalings are compared. In Figure (2.1A) all prefactors are 1 and the cubic scaling module quickly becomes the most demanding task. However, if the prefactor of

this cubic module can be made very small, Figure (2.1B), the cubic scaling module can be faster than the quadratic, and even linear, modules within a certain size range. This is the case for matrix operations performed with the Intel MKL libraries. Moreover, the prefactor for the N^4 Coulomb contribution, the highest scaling part of a LCGTO KS-DFT calculation, can also be reduced to $\frac{1}{8}$ by taking into account the permutational symmetry of the integrals. Nevertheless, even by employing the permutational symmetry of the ERIs, the calculation of the Coulomb contribution rapidly becomes the computationally most demanding part and other techniques are needed in order to reduce its computational load.

2.2 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 approximation. This technique was introduced by Dunlap and co-workers,[88–91] inspired by a former work of Sambe and Felton.[92] It became widespread available for use since its introduction into the deMon-KS[93] and DGauss[94] programs more than 20 years ago. It is equivalent to the application of the truncated resolution of the identity (RI)[95,96] for Coulomb integrals used in other programs, specially from the "wavefunction community". An extensive review of the influence of the variational fitting technique on electronic structure calculations can be found in references [97] and [98].

The variational approximation of the Coulomb potential, as implemented in deMon2k, is based on the minimization of the error:

$$\mathcal{E}_2 = \frac{1}{2} \iint \frac{[\rho(\mathbf{r}_1) - \tilde{\rho}(\mathbf{r}_1)]\,[\rho(\mathbf{r}_2) - \tilde{\rho}(\mathbf{r}_2)]}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|} \mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \geqslant 0 \tag{2.22}$$

The here appearing approximated density, $\tilde{\rho}(\mathbf{r})$, is expanded as a linear combination of

primitive Hermite-Gaussian type functions[99] (HGTFs), $\bar{k}(\mathbf{r})$:

$$\tilde{\rho}(\mathbf{r}) = \sum_{\bar{k}} x_{\bar{k}}\, \bar{k}(\mathbf{r}) \tag{2.23}$$

From now on these HGTFs will be called *auxiliary functions* and will be denoted by latin letters with a bar above them. An unnormalized auxiliary function $\bar{k}(\mathbf{r})$, centered on atom A with exponent $\zeta_{\bar{k}}$ has the following form:

$$\bar{k}(\mathbf{r}) = \left(\frac{\partial}{\partial A_x}\right)^{\bar{k}_x} \left(\frac{\partial}{\partial A_y}\right)^{\bar{k}_y} \left(\frac{\partial}{\partial A_z}\right)^{\bar{k}_z} e^{-\zeta_{\bar{k}}(\mathbf{r}-\mathbf{A})^2} \tag{2.24}$$

As for the GTOs, all parameters appearing in Equation (2.24) remain constant during an electronic structure calculation unless the geometry of the molecule is changed. In deMon2k the auxiliary functions are grouped in *s*, *spd* and *spdfg* sets sharing the same exponent within each set.[100,101] Specially developed integral recurrence relations[99,102] ensure maximum performance in the analytic molecular integral calculations with these auxiliary function sets. Expanding $\rho(\mathbf{r})$ and $\tilde{\rho}(\mathbf{r})$ in Equation (2.22) yields:

$$\mathcal{E}_2 = \frac{1}{2}\sum_{\mu,\nu}\sum_{\sigma,\tau} P_{\mu\nu}P_{\sigma\tau}\langle\mu\nu\|\sigma\tau\rangle - \sum_{\mu,\nu}\sum_{\bar{k}} P_{\mu\nu}\langle\mu\nu\|\bar{k}\rangle x_{\bar{k}} + \frac{1}{2}\sum_{\bar{k},\bar{l}} x_{\bar{k}}\langle\bar{k}\|\bar{l}\rangle x_{\bar{l}} \tag{2.25}$$

Since $\mathcal{E}_2$ is positive semi-definite,[103] the following inequality holds:

$$\frac{1}{2}\sum_{\mu,\nu}\sum_{\sigma,\tau} P_{\mu\nu}P_{\sigma\tau}\langle\mu\nu\|\sigma\tau\rangle \geqslant \sum_{\mu,\nu}\sum_{\bar{k}} P_{\mu\nu}\langle\mu\nu\|\bar{k}\rangle x_{\bar{k}} - \frac{1}{2}\sum_{\bar{k},\bar{l}} x_{\bar{k}}\langle\bar{k}\|\bar{l}\rangle x_{\bar{l}} \tag{2.26}$$

Note that the equality holds only when $\rho(\mathbf{r})$ equals $\tilde{\rho}(\mathbf{r})$. Thus, any approximated density will provide a lower bound to the true Coulomb repulsion energy. In this context it is worth to point out a common misconception in the literature regarding variationally density fitting. There are several works in which auxiliary function sets are "optimized" in order to reproduce the total energy corresponding to calculations without fitting.[104–107] However, such an approach is as useful as optimizing basis sets to total energies for the

basis set limit. 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 this in mind we now use inequality (2.26) in order to write a new variational energy expression based on Equation (2.10)

$$E[\rho] = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} \langle \bar{k} \| \bar{l} \rangle x_{\bar{l}} + E_{\mathrm{xc}}[\rho] \tag{2.27}$$

The fitting coefficients $\{x_{\bar{k}}\}$ are obtained from the minimization of $\mathcal{E}_2$:

$$\frac{\partial \mathcal{E}_2}{\partial x_{\bar{m}}} = -\sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{m} \rangle + \sum_{\bar{l}} x_{\bar{l}} \langle \bar{l} \| \bar{m} \rangle \equiv 0 \; \forall \; \bar{m} \tag{2.28}$$

The set of Equations (2.28) can be written in a more concise form as the linear equation system

$$\mathbf{G}\,\mathbf{x} = \mathbf{J} \tag{2.29}$$

where

$$\mathbf{G} = \begin{pmatrix} \langle \bar{1} \| \bar{1} \rangle & \langle \bar{1} \| \bar{2} \rangle & \cdots & \langle \bar{1} \| \bar{m} \rangle \\ \langle \bar{2} \| \bar{1} \rangle & \langle \bar{2} \| \bar{2} \rangle & \cdots & \langle \bar{2} \| \bar{m} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \bar{m} \| \bar{1} \rangle & \langle \bar{m} \| \bar{2} \rangle & \cdots & \langle \bar{m} \| \bar{m} \rangle \end{pmatrix} \tag{2.30}$$

and

$$\mathbf{J} = \begin{pmatrix} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{1} \rangle \\ \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{2} \rangle \\ \vdots \\ \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{m} \rangle \end{pmatrix} \tag{2.31}$$

are the Coulomb matrix and Coulomb vector, respectively. The fitting coefficients are collected in the vector $\mathbf{x}$. A straightforward solution of Equation (2.29) is obtained by the

inversion of $\mathbf{G}$:

$$\mathbf{x} = \mathbf{G}^{-1}\,\mathbf{J} \tag{2.32}$$

However, the inversion of $\mathbf{G}$ can be numerically unstable if large auxiliary function sets are used. Normalization of the auxiliary functions with respect to the Coulomb norm,

$$\langle \bar{k} \| \bar{k} \rangle = 1 \; \forall \; \bar{k} \tag{2.33}$$

ensures certain numerical control of $\mathbf{G}$, nevertheless, $\mathbf{G}$ can still be ill-conditioned. Therefore, deMon2k solves Equation (2.29) by means of a robust numerical solver based on a quasi Newton method.[108]

Once the fitting Equation (2.29) has been solved—for a particular density—the KS matrix, $\mathbf{K}$, can be obtained by varying Equation (2.27) with respect to the density matrix

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} + \langle \mu | v_{\mathrm{xc}} | \nu \rangle \tag{2.34}$$

Thus, the four-center ERI calculation has been substituted by two steps involving three- and two-center ERIs. The first step involves the calculation of $\mathbf{J}(\mathbf{P})$ and solving Equation (2.29). The second step is the calculation of $\mathbf{K}$. Note that these two steps must be performed in every SCF iteration due to the dependence on $\mathbf{P}$. The formal scaling for this approach is $N^2 \times M$, where M is the number of auxiliary functions. Usually $M \leqslant 3N$, thus, the scaling can be seen as N^3 with a prefactor larger than 1. As we have seen, the prefactor can be made smaller by taking into account the permutational symmetry of the ERIs, introducing a prefactor of $\frac{1}{2}$. Furthermore, the N^3 can be reduced by integral screening, asymptotic expansion techniques and taking into account the set structure with common exponents of the auxiliary functions.[99,102,109] Integral screening can be further improved by incremental building of target quantities during the SCF, e.g. for the Coulomb vector $\mathbf{J}^{(I+1)}$, in the Ith $+\,1$ SCF iteration as

$$\mathbf{J}^{(I+1)} = \mathbf{J}^{(I)} + \Delta\mathbf{J}^{(I+1)} \tag{2.35}$$

where

$$\Delta J_{\bar{k}}^{(I+1)} = \sum_{\mu,\nu} \Delta P_{\mu\nu}^{(I+1)} \langle \mu\nu \| \bar{k} \rangle \tag{2.36}$$

and

$$\Delta \mathbf{P}^{(I+1)} = \mathbf{P}^{(I+1)} - \mathbf{P}^{(I)} \tag{2.37}$$

A similar strategy can be used for the Coulomb contribution to the KS matrix by employing $\Delta \mathbf{x}^{I+1}$. Since $\Delta \mathbf{P}^{I+1}$ and $\Delta \mathbf{x}^{I+1}$ tend to the null matrix and null vector, respectively, as convergence is approaching, this technique allows to screen out a larger and larger number of integrals when the SCF reaches convergence. This leads to an algorithm where the most expensive part, in terms of computational demand, is the calculation of the exchange-correlation potential.

2.3 Exchange-Correlation Potential from Fitted Densities

The use of auxiliary functions for the calculation of the exchange-correlation potential has a long history in DFT methods.[92,110] Traditionally, these methods either fit the potential[92] or fit the non-integral powers of $\rho(\mathbf{r})$[111] by using a set of Cartesian GTO auxiliary functions. The former approach is not variational and, therefore, reliable forces (and higher order derivatives) cannot be obtained.[112,113] The latter usually needs different auxiliary function sets for different powers of $\rho(\mathbf{r})$ and its gradient, making the treatment rather cumbersome and limited. A different approach is 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.[114–117] The resulting approximation has been named Auxiliary Density Functional Theory (ADFT).[118] In ADFT it is essential that $\tilde{\rho}(\mathbf{r})$ inherits some properties of $\rho(\mathbf{r})$, specifically, that $\tilde{\rho}(\mathbf{r}) \geqslant 0$ and $\int \tilde{\rho}(\mathbf{r}) \mathrm{d}\mathbf{r} = N_e$. Normalization to the number of electron, N_e, 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 property for $\tilde{\rho}(\mathbf{r})$ is less straightforward. Fortunately, regions where $\tilde{\rho}(\mathbf{r}) < 0$ are rather small and usually occur

when $\rho(\mathbf{r}) \approx 0$,[119] therefore, screening of this points does not impact the accuracy of the methodology.[116]

The energy expression in ADFT takes the form:

$$E[\rho] = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle \, x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} \, \langle \bar{k} \| \bar{l} \rangle \, x_{\bar{l}} + E_{\mathrm{xc}}[\tilde{\rho}] \tag{2.38}$$

In deMon2k, this approach is called *AUXIS* and is the default method for calculating the exchange-correlation contributions. The corresponding Kohn-Sham matrix elements are given by:

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle \, x_{\bar{k}} + \frac{\partial E_{\mathrm{xc}}[\tilde{\rho}]}{\partial P_{\mu\nu}} \tag{2.39}$$

The last term of Equation (2.39) can be evaluated in a similar manner as in Equation (2.14) to yield

$$\frac{\partial E_{\mathrm{xc}}[\tilde{\rho}]}{\partial P_{\mu\nu}} = \int \frac{\delta E_{\mathrm{xc}}[\tilde{\rho}(\mathbf{r})]}{\delta \tilde{\rho}(\mathbf{r})} \frac{\partial \tilde{\rho}(\mathbf{r})}{\partial P_{\mu\nu}} \mathrm{d}\mathbf{r} \; = \sum_{\bar{k}} \frac{\partial x_{\bar{k}}}{\partial P_{\mu\nu}} \int v_{\mathrm{xc}}[\tilde{\rho};\, \mathbf{r}] \, \bar{k}(\mathbf{r}) \, \mathrm{d}\mathbf{r} \tag{2.40}$$

with:

$$v_{\mathrm{xc}}[\tilde{\rho};\, \mathbf{r}] \equiv \frac{\delta E_{\mathrm{xc}}[\tilde{\rho}]}{\delta \tilde{\rho}(\mathbf{r})} \tag{2.41}$$

The derivatives of the Coulomb fitting coefficients are obtained using Equations (2.31) and (2.32) to yield:

$$\frac{\partial x_{\bar{k}}}{\partial P_{\mu\nu}} = \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}} \langle \bar{l} \| \mu\nu \rangle \tag{2.42}$$

Note that $G^{-1}_{\bar{k}\bar{l}}$ refers to the $\bar{k},\bar{l}$ element of $\mathbf{G}^{-1}$. To simplify notation, we now introduce the exchange-correlation fitting coefficient vector, $\mathbf{z}$, with elements

$$z_{\bar{k}} = \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}} \langle \bar{l} | v_{\mathrm{xc}}[\tilde{\rho}] \rangle \; , \tag{2.43}$$

in order to rewrite Equation (2.39) as

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle \, (x_{\bar{k}} + z_{\bar{k}}) \tag{2.44}$$

It is important to note that $\mathbf{z}$ is spin-dependent and accounts for the difference between the α and β KS matrices in open-shell calculations. Also note that the set of Equations (2.43) can be formulated as an inhomogeneous equation system of the form:

$$\mathbf{G}\mathbf{z} = \mathbf{L} \tag{2.45}$$

where

$$\mathbf{L} = \begin{pmatrix} \langle v_{\mathrm{xc}}|\bar{1}\rangle \\ \langle v_{\mathrm{xc}}|\bar{2}\rangle \\ \vdots \\ \langle v_{\mathrm{xc}}|\bar{m}\rangle \end{pmatrix} \tag{2.46}$$

Domínguez-Soria *et al.*[108] also proposed a pre-conditioned conjugate gradient iterative solver for Equation (2.45) coupled to the Coulomb fitting solver. In order to keep the approach variational, $\tilde{\rho}(\mathbf{r})$ must be taken unaltered from the solution of Equation (2.32) to calculate $v_{\mathrm{xc}}[\tilde{\rho}]$. However this is not mandatory for the calculation of the Coulomb contribution. Therefore two sets of Coulomb fitting coefficients $\mathbf{x}$ are generally available in a deMon2k calculations, one set is obtained from the solution of Equation (2.32) and is used to calculate $v_{\mathrm{xc}}[\tilde{\rho}]$, the other set results from SCF convergence and acceleration techniques—such as auxiliary density mixing[118] and DIIS[118,120,121]—and is used to build $\mathbf{K}$.

2.3.1 Computational scaling of the ADFT LCGTO method

Because the approximated density is a linear combination of auxiliary functions, the density calculation at each grid point becomes linear, i. e., the numerical integration scaling becomes $M \times G$. Furthermore, since the auxiliary functions decay exponentially, the number of auxiliary functions with non-negligible values at each grid point are nearly constant for sufficiently large systems. Therefore, only M_{local} auxiliary functions need to be calculated for the density at each grid point. Adding one atom to the system does not change M_{local}. As a consequence, the numerical integration scales linear in G with a constant prefactor equal to M_{local}. The Coulomb contribution can be computed in an almost

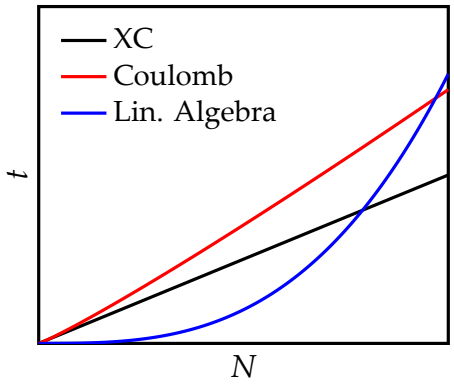


Figure 2.2: Schematic representation of the computational scaling for the three most demanding modules of ADFT in deMon2k. XC refers to the exchange-correlation contribution.

linear scaling effort by using the double-asymptotic ERI expansion technique.[109] Moreover, the iterative solver for both the Coulomb fitting, Equation (2.32), and the exchange-correlation coefficients, Equation (2.45), have a subquadratic scaling behavior.[108] All other calculations needed during each SCF iteration are matrix-matrix operations. The standard recommendation is to use vendor-optimized mathematical libraries or the Automatically Tuned Linear Algebra Software (ATLAS) for these N^3 operations to guarantee very small prefactors. Figure (2.2) shows a fictitious example for the computational scaling of ADFT. In this example, the linear algebra part has a prefactor in the order of 10^{-7} while the (almost) linear scaling modules have prefactors greater than 1. Nowadays, the linear algebra becomes the most dominant part in the deMon2k code when $N \sim 20,000$; this is reached for systems with around $1,000$ atoms or more.

When teaching chemistry students, I explain that DFT is some algorithm meaning unreliable, while ab initio *is Latin for too expensive.*

Kieron Burke

3

Low-Order Scaling Exact Exchange Evaluation

The initial idea of DFT was to represent the total energy as a functional of the density, $E[\rho]$. Unfortunately, the form of $E[\rho]$ is unknown. The so far developed approximations to $E[\rho]$ failed, in many cases, to reproduce even the most elementary properties of quantum systems. This was the reason to resort to an implicit representation of T_s in terms of the KS orbitals. Then, only an explicit representation of E_{xc} in terms of $\rho(\mathbf{r})$ is required in this variant of DFT. Unfortunately, the explicitly density-dependent approximations for E_{xc} currently available exhibit also some important deficiencies, such as the inabilities to properly bind atomic negative ions, reproduce the London dispersion force, and describe strongly correlated systems.[122,123] The obvious next step for an improvement is a represen-

tation of E_{xc}, or at least E_x, in terms of the KS orbitals. The formally correct way of implementing this is the so-called *Optimized (Effective) Potential Method* (OPM or OEP), in which the many-body problem is approached by the simultaneous solution of the KS equations and an integral equation which determines v_{xc}.[122] Unfortunately, the OPM turns out to be computationally demanding even for rather simple orbital-dependent expressions. For this reason, applications of orbital-dependent xc-functionals often rely on the so-called *generalized Kohn-Sham* (GKS) or *Hartree-Fock-Kohn-Sham* (HFKS) approach.[43,122] This idea was originally suggested by Kohn and Sham.[22] The explicit E_x is known as exact exchange (EXX) functional and is defined as

$$E_{\mathrm{EXX}}[\psi] = -\sum_{i,j}^{\mathrm{occ}} \iint \frac{\psi_i(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\psi_i(\mathbf{r}_2)\psi_j(\mathbf{r}_2)}{|\,\mathbf{r}_1-\mathbf{r}_2\,|}\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{3.1}$$

$$= -\sum_{i,j}^{\mathrm{occ}} \langle ij\|ij\rangle \tag{3.2}$$

Inserting the LCGTO expansion, Equation (3.2) transforms into

$$E_{\mathrm{EXX}} = -\frac{1}{4}\sum_{\mu,\nu}\sum_{\sigma,\tau} P_{\mu\nu}P_{\sigma\tau}\langle\mu\sigma\|\nu\tau\rangle \tag{3.3}$$

where we have used the definition of the closed-shell density matrix, Equation (2.6). The total energy in an EXX only HFKS calculation, which coincides with the HF energy, is given by

$$E[\rho] = \sum_{\mu,\nu} P_{\mu\nu}H_{\mu\nu} + \sum_{\mu,\nu}\sum_{\bar{k}} P_{\mu\nu}\langle\mu\nu\|\bar{k}\rangle x_{\bar{k}} - \frac{1}{2}\sum_{\bar{k},\bar{l}} x_{\bar{k}}\,\langle\bar{k}\|\bar{l}\rangle x_{\bar{l}} - \frac{1}{4}\sum_{\mu,\nu}\sum_{\sigma,\tau} P_{\mu\nu}P_{\sigma\tau}\langle\mu\sigma\|\nu\tau\rangle \tag{3.4}$$

where the variationally fitted Coulomb potential was used. There are two main problems with an approach based on Equation (3.4). First, computing the EXX term introduces a formal N^4 scaling that renders the approach unsuitable for large systems. Second, the energy is not self-interaction free. The self-interaction arises because the variationally fitted Coulomb energy represents a lower bound to the real Coulomb energy. This difference,

albeit a small one, will make the potential for each particle a little bit too attractive.

The first problem may be tackled by noting two very important properties of the exchange contribution. The four-center ERIs appearing in Equation (3.3) have non-negligible contributions only when $\mu(\mathbf{r}_1)$ is close in space to $\sigma(\mathbf{r}_1)$ and $\nu(\mathbf{r}_2)$ is close to $\tau(\mathbf{r}_2)$. In LCGTO, this property is linked to the Gaussian Product Theorem (GPT), which states that the product of two Gaussians is another Gaussian centered somewhere on the line between the two original Gaussians. In the case of two spherically symmetric GTOs:

$$\mathcal{N}_\mu \, e^{-\zeta_\mu (\mathbf{r}-\mathbf{A})^2} \, \mathcal{N}_\sigma \, e^{-\zeta_\sigma (\mathbf{r}-\mathbf{B})^2} = \mathcal{N}_\mu \, \mathcal{N}_\sigma \, e^{-\xi_p (\mathbf{A}-\mathbf{B})^2} \, e^{-\zeta_p (\mathbf{r}-\mathbf{P})^2} \tag{3.5}$$

where

$$\zeta_p = \zeta_\mu + \zeta_\sigma \tag{3.6}$$

$$\xi_p = \zeta_\mu \, \zeta_\sigma / \zeta_p \tag{3.7}$$

$$\mathbf{P} = (\zeta_\mu \, \mathbf{A} + \zeta_\sigma \, \mathbf{B}) / \zeta_p \tag{3.8}$$

and $\mathcal{N}_\mu$, $\mathcal{N}_\sigma$ are normalization constants. For our discussion, the following two facts arising from the GPT are most important. First, the exponent of the product, ζ_p, is always larger than the exponents of the two original functions. Thus the product function decays more rapidly than the original functions. Second, the product pre-factor,

$$\mathcal{N}_\mu \, \mathcal{N}_\sigma \, e^{-\xi_p (\mathbf{A}-\mathbf{B})^2} \ , \tag{3.9}$$

decays exponentially with the squared distance $(\mathbf{A} - \mathbf{B})^2$. The GPT also holds for non-spherically symmetric GTOs. To illustrate this further, take for example the product between two *d*-type functions centered on carbon atoms of the linear $C_{20}H_{42}$ described with a polarized valence basis set. The first column of Figure (3.1) shows the 10^{-6} a.u. isosurface of four d_{xy} functions centered on the first four carbon atoms of the chain. The second column shows the same isosurfaces for the products between the d_{xy} function of

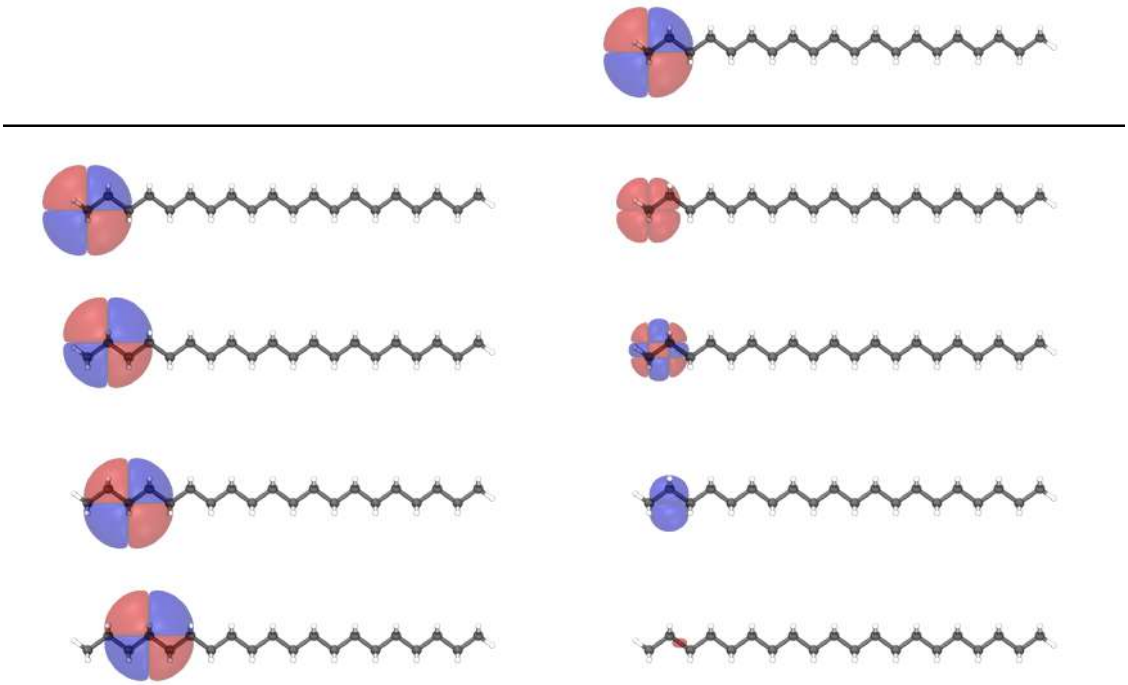


Figure 3.1: Product of two d_{xy} functions centered on carbon atoms of the $C_{20}H_{42}$. The first column shows contour plots of the d_{xy} functions of the first four carbon atoms in the chain. The second column shows contour plots of the product between d_{xy} of the first carbon atom (header) and the d_{xy} function depicted on the left of the product.

C1—shown at the header of the second column—and the d_{xy} functions depicted in the corresponding row. Note that all products are more compact than the original functions as a consequence of the larger exponent. The product between the d_{xy} functions centered on C1 and C4 is already below 10^{-5} a.u. for every point in space, and the product for the functions centered on C1 and C5 (not shown) is already below 10^{-10} a.u. for all points in space! This example clearly shows that if the centers of the two original functions, $\mu(\mathbf{r})$ and $\sigma(\mathbf{r})$, are further away from each other, then the distribution $\rho_{\mu\sigma}(\mathbf{r}) = \mu(\mathbf{r})\,\sigma(\mathbf{r})$ will vanish for any $\mathbf{r}$. As the system size increases, the number of significant distributions $\rho_{\mu\sigma}(\mathbf{r}) = \mu(\mathbf{r})\sigma(\mathbf{r})$ approaches N from above as N becomes larger.[124] Therefore, the number of non-negligible four center ERIs approaches N^2 asymptotically from above.[125,126] The EXX term has another important feature: the elements of both density matrices couple coordinates of electron one ($\mathbf{r}_1$) with coordinates of electron two ($\mathbf{r}_2$). In insulating systems $P_{\mu\nu}$ has a non-negligible value only if $\mu(\mathbf{r})$ is close to $\nu(\mathbf{r})$. A direct consequence of the GPT and the locality of $\mathbf{P}$ in insulating systems is that all four AOs appearing in the four-

center ERIs must be close in space. These, along with appropriate thresholding of $K_{\mu\nu}$ elements, are the fundamentals behind linear-scaling EXX algorithms like ONX[124,127–129] and LinK.[130,131]

Despite the existence of these linear-scaling EXX algorithms, much effort has been posed into the development of new algorithms to compute EXX. This is partially because the linear-scaling is only achieved with respect to system size. However, computing exact exchange on a given system while enlarging the basis set still results in the unfavorable N^4 scaling. Several attempts have been made to overcome this situation. Examples are Friesner's pseudospectral method,[132–135] which involves the use of both numerical grids and analytical two-electron integrals, and the somewhat related chain-of-spheres exchange (COSX) developed by Neese and coworkers,[136–138] that exploits the short-range nature of EXX by calculating this contribution via a semi-numerical integration.

Another group of algorithms aiming in the same direction are based on the approximation of the orbital distribution functions $\rho_{ij}(\mathbf{r}) = \psi_i(\mathbf{r})\psi_j(\mathbf{r})$, $\rho_{\mu i}(\mathbf{r}) = \mu(\mathbf{r})\psi_i(\mathbf{r})$ or $\rho_{\mu\nu}(\mathbf{r}) = \mu(\mathbf{r})\nu(\mathbf{r})$. The aim is to reduce the formal scaling in the same manner as in the variational fitting of the Coulomb potential. Examples are the RI-K,[104,139–142] density-fitting (DF),[143] Cholesky decomposition (CD) of the four-center ERI supermatrix,[144–146] atomic RI (ARI),[147] pair atomic RI (PARI),[148–152] concentric atomic density fitting (CADF)[153] and local density-fitting (LDF)[154] algorithms.

In the following, we present our LDF-EXX approach,[155] that has proven particularly efficient and robust.[74,155,156] Therefore, it can be used in DFT without compromising its efficiency.

3.1 Variational Fitting of the Exact Exchange Potential

Similar as for the variational fitting of the Coulomb potential, the orbital distribution $\rho_{ij}(\mathbf{r})$ can be approximated as a linear combination of atom-centered auxiliary functions:

$$\tilde{\rho}_{ij}(\mathbf{r}) = \sum_{\bar{k}} x_{\bar{k}ij}\, \bar{k}(\mathbf{r}) \tag{3.10}$$

The fitting functional is defined as

$$\mathcal{E}_2^{\mathrm{EXX}} = -\sum_{i,j}^{\mathrm{occ}} \iint \frac{\left[\rho_{ij}(\mathbf{r}_1) - \tilde{\rho}_{ij}(\mathbf{r}_1)\right]\left[\rho_{ij}(\mathbf{r}_2) - \tilde{\rho}_{ij}(\mathbf{r}_2)\right]}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|}\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{3.11}$$

As shown in Appendix A, the exact exchange fitting functional $\mathcal{E}_2^{\mathrm{EXX}}$ is negative semidefinite. Therefore, the fitting procedure seeks to maximize $\mathcal{E}_2^{\mathrm{EXX}}$. Due to the negative semidefinite nature of $\mathcal{E}_2^{\mathrm{EXX}}$, the following inequality must hold

$$-\sum_{i,j}^{\mathrm{occ}}\langle ij\|ij\rangle \leqslant \sum_{i,j}^{\mathrm{occ}}\sum_{\bar{k},\bar{l}} x_{\bar{k}ij}\langle\bar{k}\|\bar{l}\rangle x_{\bar{l}ij} - 2\sum_{i,j}^{\mathrm{occ}}\sum_{\bar{k}}\langle ij\|\bar{k}\rangle x_{\bar{k}ij} \tag{3.12}$$

The maximization of $\mathcal{E}_2^{\mathrm{EXX}}$ corresponds to a minimization of the fitted exact exchange energy given by the rhs of inequality (3.12).

Introducing the LCGTO expansion of the orbital distributions and substituting Equation (3.10) into Equation (3.11) yields

$$\mathcal{E}_2^{\mathrm{EXX}} = -\sum_{i,j}^{\mathrm{occ}}\langle ij\|ij\rangle + 2\sum_{i,j}^{\mathrm{occ}}\langle ij\|\bar{k}\rangle x_{\bar{k}ij} - \sum_{\bar{k},\bar{l}} x_{\bar{k}ij}\langle\bar{k}\|\bar{l}\rangle x_{\bar{l}ij} \tag{3.13}$$

The expansion coefficients $x_{\bar{k}ij}$ are determined by the maximization condition

$$\frac{\partial\mathcal{E}_2^{\mathrm{EXX}}}{\partial x_{\bar{k}ij}} = \langle ij\|\bar{k}\rangle - \sum_{\bar{l}} x_{\bar{l}ij}\langle\bar{l}\|\bar{k}\rangle \equiv 0 \;\;\forall\;\; \bar{k}, i, j \tag{3.14}$$

From Equation (3.14) one can define a set of linear equation systems, one for each ρ_{ij} distribution, which in matrix notation take the form:

$$\mathbf{G}\mathbf{x}_{ij} = \mathbf{J}_{ij} \tag{3.15}$$

with

$$\mathbf{J}_{ij} = \begin{pmatrix} \langle ij\|\bar{1}\rangle \\ \langle ij\|\bar{2}\rangle \\ \vdots \\ \langle ij\|\bar{m}\rangle \end{pmatrix} \tag{3.16}$$

After the fitting equations have been solved, the fitted exact exchange can be written in a more compact form as

$$\begin{aligned} E_{\mathrm{EXX}} &= -\sum_{i,j}^{\mathrm{occ}} \sum_{\bar{k},\bar{l}} x_{\bar{k}ij} G_{\bar{k}\bar{l}} x_{\bar{l}ij} && (3.17) \\ &= -\sum_{i,j}^{\mathrm{occ}} \sum_{\bar{k},\bar{l}} \langle ij\|\bar{k}\rangle G^{-1}_{\bar{k}\bar{l}} \langle \bar{l}\|ij\rangle && (3.18) \\ &= -\frac{1}{4} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\sigma\|\bar{k}\rangle G^{-1}_{\bar{k}\bar{l}} \langle \bar{l}\|\tau\nu\rangle && (3.19) \end{aligned}$$

Using either Equation (3.18) or Equation (3.19) avoids the explicit calculation of the three-index exact exchange fitting coefficients appearing in Equation (3.17). Note, however, that this is the result of solving the set of Equations (3.15). If this fitted exact exchange energy is added to the Hartree energy with Coulomb fitting, the self-interaction-free density fitting Hartree-Fock energy expression is obtained:[155]

$$\begin{aligned} E_{\mathrm{HF}} &= \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\bar{k}} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu\|\bar{k}\rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} G_{\bar{k}\bar{l}} x_{\bar{l}} \\ &\quad -\frac{1}{4} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\sigma\|\bar{k}\rangle G^{-1}_{\bar{k}\bar{l}} \langle \bar{l}\|\tau\nu\rangle && (3.20) \end{aligned}$$

It is important to note that E_{HF} is self-interaction free only when the same auxiliary function set is used for both Coulomb and exact exchange fittings. In order to show that E_{HF} is self-interaction free, consider the self-interaction exact exchange fitting coefficient

$$x_{\bar{k}ii} = \sum_{\mu,\nu} \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}} \langle \bar{l}\|ii\rangle \tag{3.21}$$

Summing over all these coefficients yields

$$\sum_{i}^{\mathrm{occ}} x_{\bar{k}ii} = \frac{1}{2} \sum_{\mu,\nu} \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}} \langle \bar{l} \| \mu\nu \rangle P_{\mu\nu} = \frac{1}{2} x_{\bar{k}} \tag{3.22}$$

Thus, it follows that the self-interaction energy contained in the Coulomb fitting coefficients is canceled by the diagonal exact exchange fitting coefficients.

Moreover, in the simultaneous Coulomb and exact exchange fittings with a common auxiliary function set, an advantageous error compensation occur. The error compensation arises from the fact that the total energy calculated only with the variational Coulomb fitting is a lower bound to the true energy, $E \geqslant E^{\mathrm{DFJ}}$, while the total energy calculated only with the variational exact exchange fitting correspond to an upper bound to the true energy, $E \leqslant E^{\mathrm{DFK}}$. From these inequalities follows

$$E - E_{\mathrm{DFK}} \leqslant 0 \leqslant E - E_{\mathrm{DFJ}} \tag{3.23}$$

and further

$$E - E_{\mathrm{DFK}} \leqslant E - E_{\mathrm{DFJK}} \leqslant E - E_{\mathrm{DFJ}} \tag{3.24}$$

Thus, the absolute deviation from the four-center energy while performing both fittings simultaneously is smaller than the maximum absolute deviation of either fit alone.

The downside of the straightforward implementation of an algorithm based on Equations (3.17)–(3.19) is that the formal scaling still is N^4 with respect to system size. However, a formal N^3 scaling is achieved with respect to the basis set size. Take for example Equation (3.18). The three-center ERIs must be computed and transformed into MO representation. If one decides to perform the transformation for both MOs in one step, that is

$$\langle ij \| \bar{k} \rangle = \sum_{\mu,\nu} c_{\mu i} c_{\nu j} \langle \mu\nu \| \bar{k} \rangle \quad \forall \quad i, j, \bar{k} \quad , \tag{3.25}$$

the resulting algorithm scales as $N_{\mathrm{occ}}^2 \times N^2 \times M$, which can be related to an N^5 scaling. A

more efficient algorithm arises by splitting the ERI transformation into two steps:

$$\begin{aligned} \langle i\nu\|\bar{k}\rangle &= \sum_{\mu} c_{\mu i}\langle\mu\nu\|\bar{k}\rangle \quad \forall \quad i,\nu,\bar{k} \\ \langle ij\|\bar{k}\rangle &= \sum_{\nu} c_{\nu j}\langle i\nu\|\bar{k}\rangle \quad \forall \quad i,j,\bar{k} \end{aligned} \tag{3.26}$$

The first transformation scales as $N_{\text{occ}} \times N^2 \times M$ and the second as $N_{\text{occ}}^2 \times N \times M$, which gives two N^4 steps. Another quartic step, common to both approaches, arises from the multiplication

$$\sum_{\bar{k}} \langle ij\|\bar{k}\rangle G^{-1}_{\bar{k}\bar{l}} \quad \forall \quad i,j,\bar{l} \tag{3.27}$$

Changing basis set keeps N_{occ} constant, thus, the N^3 scaling with respect to the basis set size. The scaling of the ERI calculation and transformation to the MO basis can be reduced by one order of magnitude in sufficiently large systems, nevertheless, the quartic scaling multiplication remains present. Hence, the application of DF or RI exact exchange has historically been limited to compact systems with large basis sets.[104,140,147,152,155]

3.2 Local-Density-Fitting Exact Exchange

To overcome this situation, local versions of the RI and DF algorithms have been proposed recently. These algorithms are now called ARI,[147] PARI[151,152] and LDF.[154,155] In this section we will present the development of an LDF-EXX algorithm which has proven to be robust and efficient.[74,155,156]

The basis of our LDF approach is the already mentioned invariance of the Kohn-Sham energy and the Kohn-Sham matrix to orthogonal transformations of the MOs. In this way, the delocalized canonical MOs (CMOs) can be transformed into spatially localized MOs (LMOs) by minimizing, or maximizing, an appropriate functional. The transformation of CMOs into LMOs using an orthogonal matrix $\mathbf{U}$ is given by:

$$\psi_I(\mathbf{r}) = \sum_{i}^{\text{occ}} U_{Ii}\psi_i(\mathbf{r}) \tag{3.28}$$

The inverse transformation is accomplished by using $\mathbf{U}^T$ due to the properties of orthogonal matrices. Transforming the CMOs appearing in Equation (3.2) into LMOs yields:

$$\begin{aligned} E_{\text{EXX}} &= -\sum_{I,J,K,L}^{\text{occ}} \sum_{i,j}^{\text{occ}} U_{Ii} U_{Jj} \langle IJ \| KL \rangle U_{Ki} U_{Lj} \\ &= -\sum_{I,J,K,L}^{\text{occ}} \delta_{IK} \langle IJ \| KL \rangle \delta_{JL} \\ &= -\sum_{I,J}^{\text{occ}} \langle IJ \| IJ \rangle \end{aligned} \tag{3.29}$$

where the capital letters label LMOs. Equation (3.29) differs from Equation (3.2) only in the representation of MOs used. The total E_{EXX} remains constant, however, each individual term of the sum may be different. The objective of using LMOs is to minimize the number of LMO pairs that have non-negligible contributions to E_{EXX}. This can be accomplished if the centers of LMOs $\psi_I(\mathbf{r})$ and $\psi_J(\mathbf{r})$ are far away from each other and their spatial extents are minimized. In addition, the computation of the EXX contribution to the Kohn-Sham matrix can also benefit from the use of LMOs. The contribution of the EXX to the $K_{\mu\nu}$ element is given by:

$$\begin{aligned} \frac{\partial E_{\text{EXX}}}{\partial P_{\mu\nu}} = \; X_{\mu\nu} = & -\frac{1}{2} \sum_{\sigma,\tau} \sum_{\bar{k},\bar{l}} P_{\sigma\tau} \langle \mu\sigma \| \bar{k} \rangle G^{-1}_{\bar{k}\bar{l}} \langle \bar{l} \| \nu\tau \rangle \\ = & -\sum_{i}^{\text{occ}} \sum_{\bar{k},\bar{l}} \langle i\mu \| \bar{k} \rangle G^{-1}_{\bar{k}\bar{l}} \langle \bar{l} \| \nu i \rangle \end{aligned} \tag{3.30}$$

where the sum over occupied orbitals can be performed with any representation of the MOs. Irrespective from the MO representation used, non-vanishing contributions to $X_{\mu\nu}$ arise only from MOs close in space to $\mu(\mathbf{r})$ and $\nu(\mathbf{r})$. For CMOs this will be, generally, the whole molecule, whereas for LMOs this can be much more restricted.

In order to illustrate the advantage of computing $X_{\mu\nu}$ with LMOs instead of CMOs, Figure 3.2 shows the 10^{-6} a.u. isosurface for the products between a CMO or an LMO and four different d_{xy} AOs of $C_{20}H_{42}$. The chosen AOs are centered on C1, C10, C15 and C20 in order to span the whole molecule. It can be seen that all the products of the CMO have

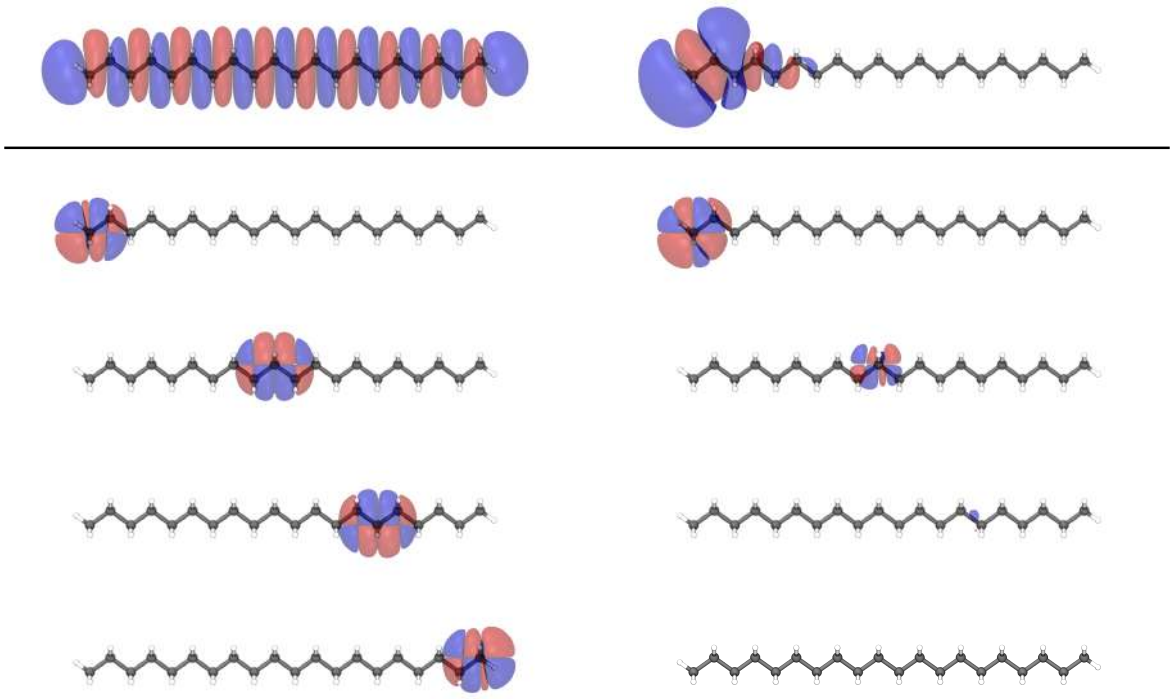


Figure 3.2: Product of a canonical molecular orbital (CMO) and a localized molecular orbital (LMO) with different atomic orbitals (AOs) of $C_{20}H_{42}$. The header shows the CMO and the LMO in the first and second column, respectively. The first column shows the product of d_{xy} AOs centered at C1, C10, C15 and C20 with the CMO and the second column shows the products of the same AOs with the LMO. The molecular orbitals are shown at 10^{-3} a.u. amplitudes and the products at 10^{-6} a.u. amplitudes.

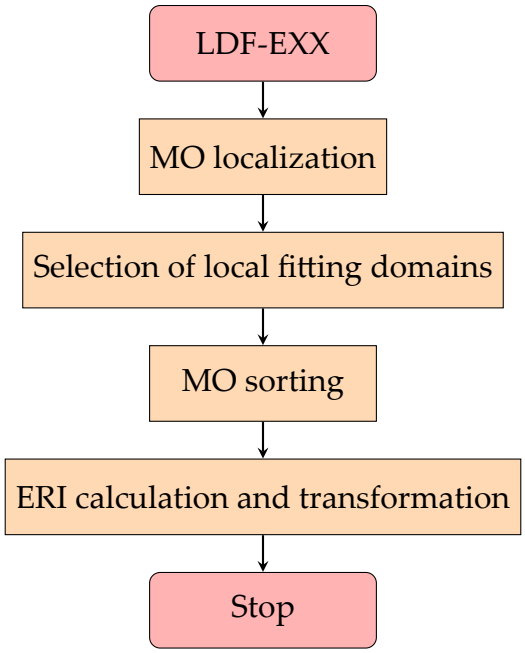


Figure 3.3: Algorithmic flowchart of the LDF-EXX computation.

non-vanishing contributions, whereas the products of the LMO decay with the distance between the LMO center and the AO. In this way, many $\psi_i(\mathbf{r})\mu(\mathbf{r})$ combinations can be discarded and, consequently, the computational effort diminished. Furthermore, due to the variational nature of the EXX fitting, one can restrict the auxiliary functions set for each LMO to only those around its center. Very good approximations are obtained provided that the local auxiliary function set chosen describes appropriately the Hilbert space spanned by the non-negligible products of $\psi_i(\mathbf{r})$. Thus, in our LDF-EXX approach each $\psi_i(\mathbf{r})$ has a local basis set $\mathcal{B}_i$ and a local auxiliary function set $\mathcal{A}_i$. These local sets will be called the local fitting domains.

The resulting LDF-EXX algorithm is based on the same equations as DF-EXX. However, the equations are solved using LMOs and local fitting domains. As a consequence, a reduction of the computational complexity is achieved, leading to an almost linear scaling algorithm. A simplified flowchart of the processes involved in our LDF-EXX algorithm is shown in Figure 3.3. In the following, each process will be discussed in detail.

3.2.1 Molecular orbital localization

The MO localization is a very important step and, ultimately, determines the speed and accuracy of the LDF-EXX algorithm. This is due to the fact that tightly localized MOs have smaller fitting domains. Therefore, the number of operations needed to compute $\mathbf{X}$ is minimized by using tightly localized MOs. Unfortunately, the scaling of several localization schemes is at least cubic. In order to overcome this disadvantage, we developed an approach that provides tightly localized MOs without too much computation.

The first step for the MO localization is obtaining a set of MOs from the pivoted Cholesky decomposition of the density matrix as proposed by Aquilante *et al.*[157] The pivoted Cholesky decomposition can be performed in a linear scaling fashion for sparse matrices, nevertheless, the current implementation of deMon2k uses the LAPACK subroutines which have an $N^2 \times N_{\text{occ}}$ scaling. One of the advantages of using these Cholesky MOs is that they are already somewhat local and provide an adequate starting point for some other localization schemes such as Foster-Boys (FB),[158–160] Edminston-Ruedenberg (ER),[160] or Pipek-Mezey (PM).[161,162]

The FB localization minimizes the spatial extension of the MOs. This is equivalent to maximize the sum of squares of distances of orbital centroids from the origin of the coordinate system:

$$FB[\psi] = \sum_{i}^{\text{occ}} \langle i|\mathbf{r}|i\rangle^2 \tag{3.31}$$

The ER method maximizes the self-repulsion energy

$$ER[\psi] = \sum_{i}^{\text{occ}} \langle ii\|ii\rangle \ , \tag{3.32}$$

and the PM approach maximizes the Mulliken charge[163] of each orbital

$$PM[\psi] = \sum_{i}^{\text{occ}} \sum_{A} \langle i|P_A|i\rangle^2 \tag{3.33}$$

where P_A is an operator that projects onto the basis functions centered on atom A.

For our purposes, the most natural localization approach would be ER, because maximization of Equation (3.32) minimizes the exchange integrals

$$\langle ij||ij\rangle \quad \forall \quad i \neq j \tag{3.34}$$

Unfortunately, the ER localization has a formal scaling of at least N^5.[162,164] Therefore, we decided to implement the more economical FB algorithm as it is the second obvious choice for our purpose. Several implementations of the FB localization follow the original prescription of Edmiston and Ruedenberg that determines the optimum orthogonal transformation of the MOs by consecutive two by two rotations until convergence is reached.[160,162] For the occupied space, the FB functional usually have strong and isolated maxima and has been successfully optimized using this method.[165] The objective is to find the optimal two by two rotation

$$\begin{aligned} \psi_i'(\mathbf{r}) &= \cos\gamma\ \psi_i(\mathbf{r}) + \sin\gamma\ \psi_j(\mathbf{r}) \\ \psi_j'(\mathbf{r}) &= -\sin\gamma\ \psi_i(\mathbf{r}) + \cos\gamma\ \psi_j(\mathbf{r}) \end{aligned} \tag{3.35}$$

that maximizes the chosen functional. Orbital-stability conditions lead to:[162]

$$\cos 4\gamma = -\frac{A_{ij}}{\sqrt{A_{ij}^2 + B_{ij}^2}} \quad , \quad 0 \leqslant \gamma < \frac{\pi}{2} \tag{3.36}$$

where A_{ij} and B_{ij} are defined as:

$$A_{ij} = \langle i|\mathbf{r}|j\rangle^2 - \frac{1}{4}\left(\langle i|\mathbf{r}|i\rangle - \langle j|\mathbf{r}|j\rangle\right)^2 \tag{3.37}$$

$$B_{ij} = \langle i|\mathbf{r}|j\rangle \cdot \left(\langle i|\mathbf{r}|i\rangle - \langle j|\mathbf{r}|j\rangle\right) \tag{3.38}$$

The orbital-stability criteria also implies that

$$B_{ij} = 0 \tag{3.39}$$

$$A_{ij} < 0 \tag{3.40}$$

for all pairs of MOs. As already noted by Pipek and Mezey, this set of relations could be satisfied if all off-diagonal matrix elements $\langle i|\mathbf{r}|j\rangle$ were zero, or in other words, if the Boys localized orbitals were all eigenfunctions of the vector operator $\mathbf{r}$. Unfortunately, the components x, y and z of $\mathbf{r}$ in the finite N-dimensional matrix representation of the occupied subspace do not commute[166,167] and, consequently, the eigenvalue problem of $\mathbf{r}$ cannot be solved. However, the FB localized MOs do minimize the squares of the off-diagonal matrix elements of $\mathbf{r}$.[162] Thus, FB localization can be seen as the approximate joint diagonalization (AJD)[168–170] of the three dipole matrices $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{Z}}$.

In AJD, one seeks to minimize the sum of squares of the off-diagonal elements of all the matrices that need to be joint diagonalized. This is accomplished also by consecutive two by two rotations in an extended Jacobi fashion. In order to obtain the Jacobi angles for the set of dipole matrices, we define a 2×2 real symmetric matrix $\boldsymbol{\mathcal{G}}$ as

$$\boldsymbol{\mathcal{G}} = \begin{pmatrix} \mathcal{X}_{ii} - \mathcal{X}_{jj} & \mathcal{Y}_{ii} - \mathcal{Y}_{jj} & \mathcal{Z}_{ii} - \mathcal{Z}_{jj} \\ 2\mathcal{X}_{ij} & 2\mathcal{Y}_{ij} & 2\mathcal{Z}_{ij} \end{pmatrix} \begin{pmatrix} \mathcal{X}_{ii} - \mathcal{X}_{jj} & 2\mathcal{X}_{ij} \\ \mathcal{Y}_{ii} - \mathcal{Y}_{jj} & 2\mathcal{Y}_{ij} \\ \mathcal{Z}_{ii} - \mathcal{Z}_{jj} & 2\mathcal{Z}_{ij} \end{pmatrix} \tag{3.41}$$

In terms of the previously defined quantities A_{ij} and B_{ij} the matrix $\boldsymbol{\mathcal{G}}$ can be written as

$$\boldsymbol{\mathcal{G}} = 4 \begin{pmatrix} \langle i|\mathbf{r}|j\rangle^2 - A_{ij} & \frac{1}{2}B_{ij} \\ \frac{1}{2}B_{ij} & \langle i|\mathbf{r}|j\rangle^2 \end{pmatrix} \tag{3.42}$$

The Jacobi angles can now be computed in closed form as

$$\cos\gamma = \sqrt{\frac{x+1}{2}} \quad , \quad \sin\gamma = \frac{y}{\sqrt{2x+2}} \tag{3.43}$$

where $[x, y]^T$ is any eigenvector with $x \geqslant 0$ associated to the largest eigenvalue of $\boldsymbol{\mathcal{G}}$. This particular choice constraints the rotation angle to $\gamma \in [-\frac{\pi}{4}, \frac{\pi}{4}]$ and leads to a locally quadratic convergent joint diagonalization algorithm.[168,169] It is important to note that PM localization can also be casted in terms of an AJD, however, in this case the problem con-

Table 3.1: Maximum orbital spreads [a.u.] for Cholesky localized molecular orbitals (CLMOs) and Foster-Boys localized molecular orbitals (FBLMOs) of coronene with increasing basis set size. Also shown are maximum orbital spreads for canonical molecular orbitals (CMOs) and FBLMOs from the trust-region minimization in reference [165].

	cc-pVDZ		cc-pVTZ		cc-pVQZ	
	Occ	Virt	Occ	Virt	Occ	Virt
CMO	7.498	10.806	7.496	11.613	7.497	12.310
CLMO	3.204	7.149	6.150	9.257	6.627	9.709
FBLMO	2.290	2.951	2.296	3.459	2.298	4.208
Ref. [165]	2.288	3.004	2.292	3.261	3.012	7.614

sists of N_{atom} matrices.

The FB implementation in deMon2k performs the Jacobi rotations with the angles defined by Equation (3.43), since this choice yields a much more stable algorithm. In order to demonstrate the effectiveness of this MO localization scheme, Table 3.1 shows the maximum orbital spreads for different MO localization approaches for the HF wavefunction of coronene. The orbital spread of the ith MO is defined as:

$$\sigma_i = \sqrt{\langle i|\mathbf{r}^2|i\rangle - \langle i|\mathbf{r}|i\rangle^2} \tag{3.44}$$

It is a measure of the spatial extent of that orbital and thus of its locality. Coronene, also known as superbenzene, has been chosen because it represents a non-metallic highly delocalized system. Table 3.1 shows that Cholesky Localized MOs (CLMOs) are weakly localized Canonical MOs (CMOs). Increasing the cardinality of the Dunning basis sets results in less local CLMOs. Nevertheless, CLMOs do represent a good starting point for the FB localization. FB Localized MOs (FBLMOs) are much more local than both, CMOs and CLMOs. The occupied FBLMOs maximum spread remains almost constant independent of the cardinality of the basis set. However, the maximum spread for the unoccupied FBLMOs deteriorates with increasing basis set size, but still remains much smaller than for the CMOs. It is important to note that the FBLMOs obtained with the AJD algorithm are, at worst, equally localized than the ones obtained with the more sophisticated trust-region Newton method (TRNM) described in reference [165]. Even more, the FBLMOs obtained through AJD for the cc-pVQZ basis set are, markedly, more compact than those

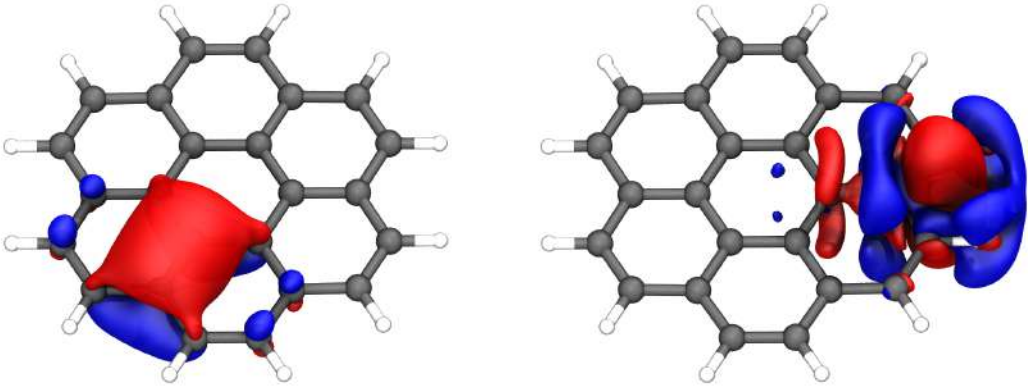


Figure 3.4: Most delocalized occupied and virtual HF coronene molecular orbital isosurfaces (0.03 a.u.).

obtained with the TRNM. This emphasize the use of CLMOs, instead of CMOs, as starting orbitals for the FB localization procedure.

In order to relate the orbital spreads to an easily understandable picture, Figure 3.4 shows the 0.03 a.u. isosurface for the least localized occupied and virtual FBLMOs of coronene. These orbitals are identical to the ones presented in reference [165] and are typical examples of FB localized orbitals produced by mixing several σ and π MOs.

3.2.2 Selection of the local fitting domains

Once the MOs have been localized, the next step in the LDF-EXX algorithm is the selection of the local fitting domains. The flowchart for the selection of the local fitting domains is depicted in Figure 3.5. The first step is to calculate the atomic Löwdin populations[171] for each orbital. Löwdin populations are calculated from the MO coefficients corresponding to the symmetric-orthogonalized AO basis.[172] The Löwdin MO coefficients can be obtained from the original ones by the transformation:

$$\breve{\mathbf{c}} = \mathbf{S}^{1/2}\mathbf{c} \tag{3.45}$$

The atomic Löwdin populations for a given orbital $\psi_i(\mathbf{r})$ are then defined as:

$$q_{iA} = \sum_{\mu \in A} \breve{c}^2_{\mu i} \tag{3.46}$$

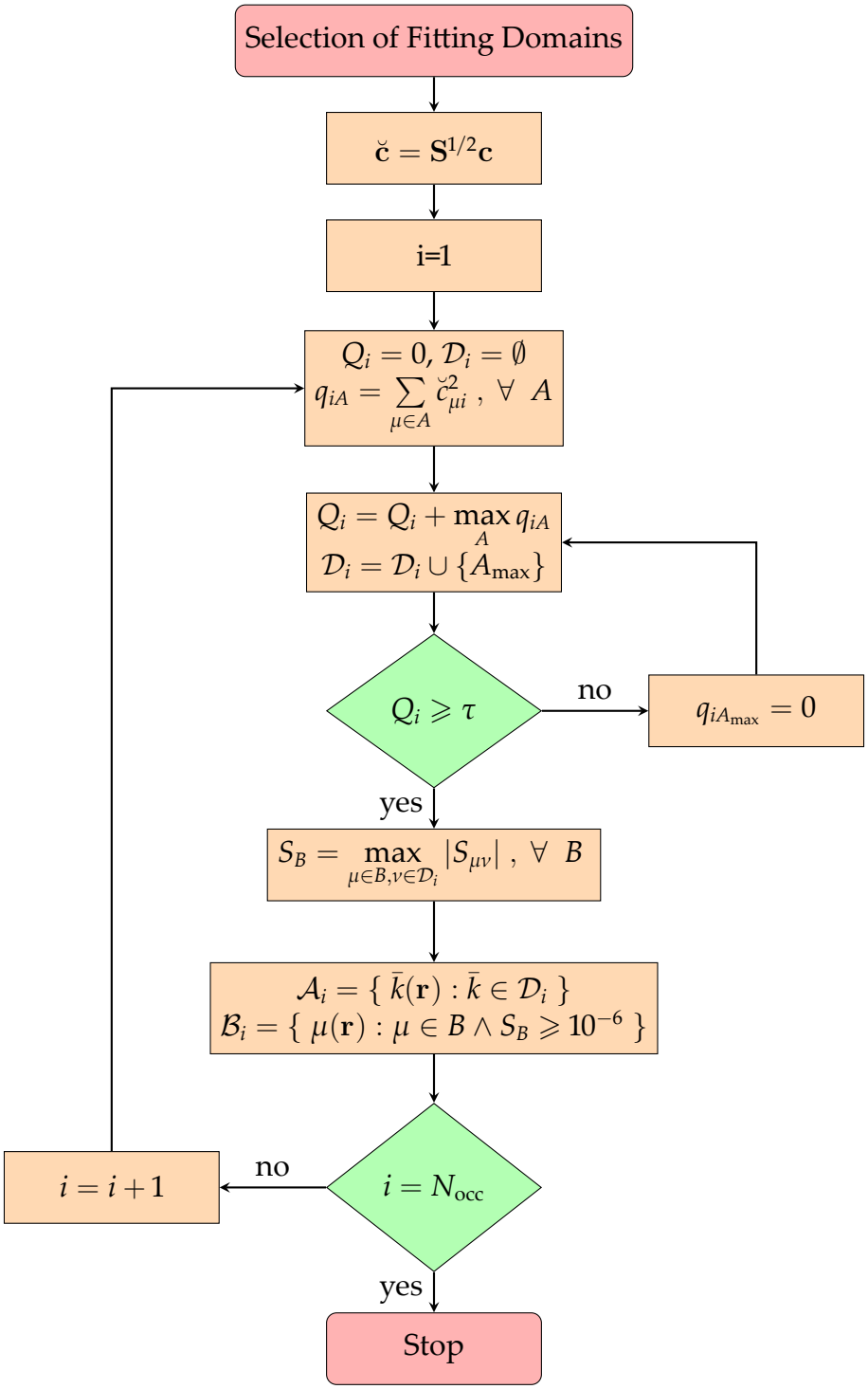


Figure 3.5: Flowchart for the selection of the local fitting domains. The vector $\mathbf{q}_i$ collects the Löwdin atomic populations q_{iA} for orbital $\psi_i(\mathbf{r})$ and $A_{\max}$ denotes the position of the largest element of $\mathbf{q}_i$.

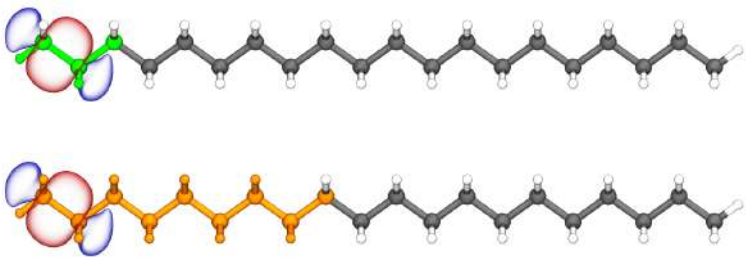


Figure 3.6: Auxiliary functions (top, green) and basis functions (bottom, orange) fitting domains for the least localized FBLMO of $C_{20}H_{42}$.

Once all q_{iA}'s for a given MO have been obtained, a local set of atoms, $\mathcal{D}_i$, is built with those atoms with the n largest atomic populations. The number of elements n of the local set is defined as the minimum number of atoms that achieve

$$Q_i = \sum_{A}^{n} q_{iA} \geqslant \tau \tag{3.47}$$

if the populations were ordered from largest to smallest. Then, all auxiliary functions centered on atoms in $\mathcal{D}_i$ define the local auxiliary function set $\mathcal{A}_i$. Additionally, all AOs centered on atoms with significant overlap to any atom in $\mathcal{D}_i$ define the local basis set $\mathcal{B}_i$. Figure 3.6 shows an actual example of atoms contributing to both $\mathcal{A}_i$ (green) and $\mathcal{B}_i$ (orange) for the least localized FBLMO of $C_{20}H_{42}$ and a $\tau = 0.99$. It can be seen that the number of auxiliary functions in each $\mathcal{A}_i$ is very small as compared to the total number of auxiliary functions of the system.

It is important to note that the Löwdin population analysis is not rotationally invariant for Cartesian representations of basis sets.[173] As a consequence, the selection of the fitting domains will not be, in general, rotationally invariant. We have not seen any problems related to the selection of the fitting domains probably because of the rather strict threshold τ used in the selection. This (almost) rotational invariance property may not be shared by other selection schemes like the one described in reference [154], where coarser criteria are used.

3.2.3 Molecular orbital sorting

The parallelization paradigm used for the LDF-EXX algorithm is based on the distribution of the LMOs. For example, if a system contains 160 LMOs and the calculation is performed with 16 CPUs, then, each CPU will calculate and transform the ERIs corresponding only to 10 LMOs. In order to maximize the speed-up obtained from such a parallel calculation, all the LMOs assigned to a given CPU must be close to each other. In this way, ERIs can be reused for several, if not all, the assigned LMOs. If the LMOs are not close to each other, the ERIs can't be reused and the number of floating-point operations will increase. Furthermore, it is possible that many CPUs end up calculating the same ERIs diminishing the parallelization efficiency.

In order to avoid such situations, the LMOs are ordered according to the spatial location of their centroids. The ordering is performed through the octal tree algorithm, also known as octree.[174] The term octree is used to describe a class of hierarchical data structures whose common property is that they are based on the principle of recursive decomposition of space.[175] The first node of the tree, the *root*, is a cube. Each node has either eight children or no children. The eight children form a $2 \times 2 \times 2$ regular subdivision of the parent node. A node with children is called an *internal node*. A node without children is called a *leaf*.[176] In Figure 3.7, an schematic representation of a simple octree is given. The root node contains all data and its considered to be a *Level 0* node. The root node is subdivided into eight *Level 1* children. Six of these Level 1 nodes are empty and are represented as non-filled circles. Since there is no point in subdividing the empty nodes, they are classified as *leaves* because they are the ending points of the tree. The remaining two Level 1 nodes contain some amount of data and are represented as filled circles. These nodes can either be subdivided into eight children or be classified as *leaves* depending on the amount of data contained within them. If the amount of data is greater than a predefined threshold then the subdivision process continues and they are classified as *internal nodes*. This is exactly the case depicted in Figure 3.7, given rise to sixteen *Level 2* nodes—eight for each one of the Level 1 internal nodes. In the example, all of the Level 2 nodes are leaves because

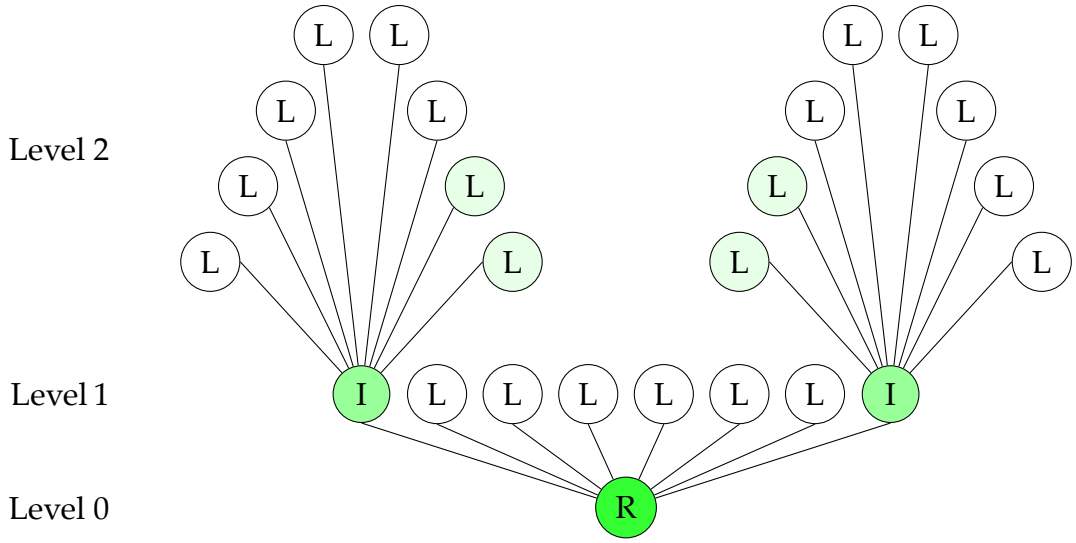


Figure 3.7: Schematic representation of an octree where each node is depicted as a circle. The node labeled with an "R" is the *root node*, the nodes labeled with an "I" are *internal nodes*, and the nodes labeled with an "L" are *leaf nodes*. A filled circle means that the given node contains some data, and the amount of these data is represented by the intensity of the filling.

they are empty or contain little amount of data. The octree generation can also be stopped when a specified resolution is achieved independently of the amount of data remaining in the higher level nodes. For example, in order to reach a resolution of 256 eight levels are required ($2^8 = 256$). This means that the Level 8 nodes have a side length that is $\frac{1}{256}$ of that of the root node. If the desired resolution is 256, then the octree generation is stopped under all circumstances at Level 8, no matter the amount of data present in the Level 8 leaf nodes.

Our octree implementation in deMon2k uses both stopping criteria. The maximum amount of LMOs centroids for each leaf node in a parallel calculation is given by:

$$N_{\text{leaf}} = \left\lfloor \frac{N_{\text{occ}}}{N_{\text{CPU}}} \right\rfloor + 1 \tag{3.48}$$

where N_{occ} is the number of occupied LMOs in the system and the floor function $\lfloor x \rfloor$ returns the largest integer not greater than x. In the case of serial calculations N_{leaf} is defined as:

$$N_{\text{leaf}} = \left\lfloor \frac{N_{\text{occ}}}{N_{\text{batch}}} \right\rfloor + 1 \tag{3.49}$$

where N_{batch} is the number of batches needed to calculate the exact exchange matrix $\mathbf{X}$

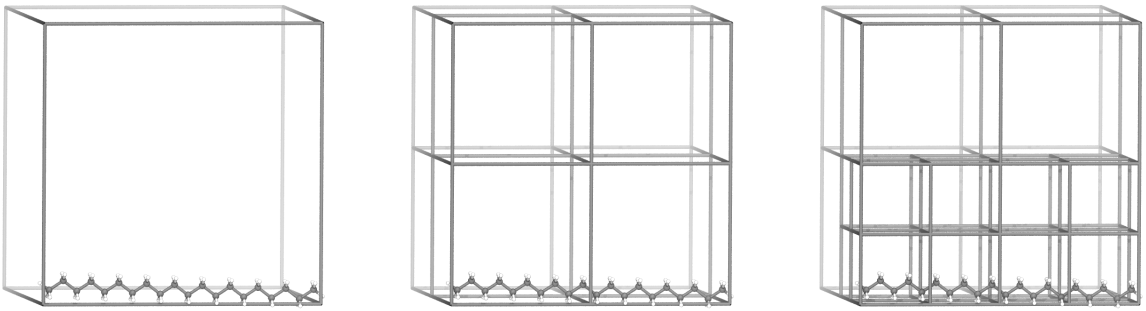


Figure 3.8: Octree generated for $C_{20}H_{42}$ and 4 CPUs. The left image shows the root node, the middle image also shows all the generated level 1 nodes and the right image shows, in addition, all level 2 nodes.

due to memory restrictions. In order to avoid memory overflows, the second stopping criteria in deMon2k is a maximum number of nodes including internal and leaf nodes. The current setting is that no more than $40,000$ nodes may be created. This number of nodes is achieved if *every* node is subdivided until Level 5. However, in most cases several leaf nodes are empty, allowing non-empty leaf nodes of levels greater than 5.

To further illustrate the algorithm, an actual octree generated for $C_{20}H_{42}$ and 4 CPUs is shown in Figure 3.8. The first step is to determine the stopping criteria for the octree generation according to the total amount of data present and the number of CPUs in the task. The data that needs to be distributed are the 81 occupied LMOs of $C_{20}H_{42}$. According to Equation (3.48), the maximum number of LMOs centered on a given leaf node is

$$N_{\text{leaf}} = \left\lfloor \frac{81}{4} \right\rfloor + 1 = 21 \tag{3.50}$$

The next step is to generate the root node, i. e. a cube that encloses the whole system. The root node is shown on the leftmost image in Figure 3.8. Since the root node contains more LMO centers, 81, than the previously defined maximum, 21, it is subdivided into 8 smaller cubes. These are the Level 1 nodes and are shown in the middle image of Figure 3.8. Six Level 1 nodes are empty—the four nodes in the back and the two upper nodes in the front—and they are classified as leaves. The remaining two Level 1 nodes contain 40 and 41 LMO centers, respectively. Since our stopping criteria has not been met, these two Level 1 nodes are classified as internal nodes and further subdivided (right image of Figure 3.8). Only four Level 2 nodes have LMO centers within them. These four nodes contain 20, 20, 20

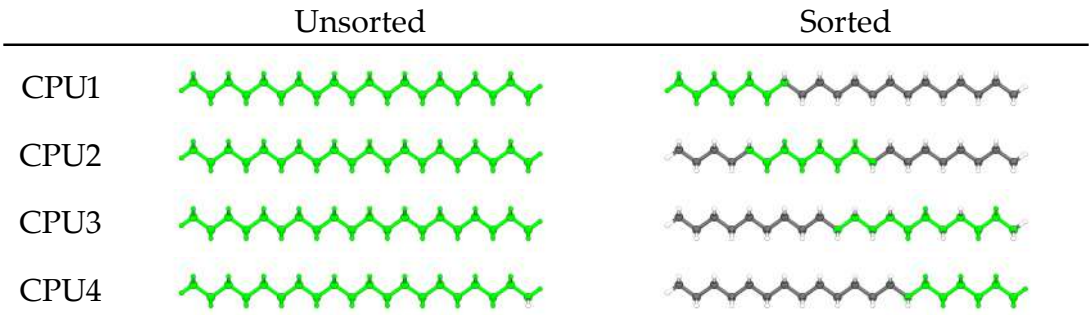


Figure 3.9: Auxiliary function center distribution (green) before and after the octree algorithm.

and 21 LMO centers. Thus, the stopping criteria is reached for all nodes and the octree generation is stopped. In fact, Figure 3.7 can be seen as an schematic representation of Figure 3.8. Each one of the four non-empty leaf nodes have LMOs centered close in space to each other. The contributions to the exact exchange matrix $\mathbf{X}$ of the LMOs centered on one non-empty leaf node are computed by one of the CPUs, maximizing the reuse of computed ERIs.

This maximization is depicted in Figure 3.9 where the centers of the auxiliary functions whose ERIs must be computed are shown in green for each one of the 4 CPUs, in the $C_{20}H_{42}$ example, before and after the sorting algorithm. Each CPU has the same number of LMOs assigned for both the unsorted and sorted cases. It can be seen that in the unsorted case every CPU must address almost every auxiliary function of the molecule. After the octree algorithm, each CPU addresses only a fraction of the auxiliary functions of the molecule. Therefore, more three-center integrals, $\langle \mu\nu \| \bar{k} \rangle$, can be reused to compute half-transformed integrals of the type $\langle \mu i \| \bar{k} \rangle$ appearing in Equation (3.30).

3.2.4 ERI computation and transformation

The last step in LDF-EXX is the computation and transformation of the ERIs in order to calculate the exact exchange matrix elements, $X_{\mu\nu}$, according to Equation (3.30). Due to the local fitting domains, each LMO needs the computation of only a limited amount of AO ERIs, $\langle \mu\nu \| \bar{k} \rangle$, and its corresponding transformation to $\langle \mu i \| \bar{k} \rangle$. In order to simplify the

notation we define a new set of matrices $\mathbf{E}_i$ with elements:

$$E_{\mu\bar{k},i} = \langle \mu i \| \bar{k} \rangle \tag{3.51}$$

In general, all $\mathbf{E}_i$ matrices have different dimensions. In addition, each LMO has an associated Coulomb matrix $\mathbf{G}_i$. As in the previous case, the dimension of the $\mathbf{G}_i$ matrices can be different. Then, the contribution of each LMO to $\mathbf{X}$ is given by:

$$\mathbf{X}_i = -\mathbf{E}_i \mathbf{G}_i^{-1} \mathbf{E}_i^T \tag{3.52}$$

where $\mathbf{X}_i$ and $\mathbf{X}$ can have, also, different dimensionalities. This property follows from the locality of all $\mathcal{B}_i$. Thus, the last step is to sum the elements of $\mathbf{X}_i$ into the appropriate blocks of $\mathbf{X}$. The existence of one $\mathbf{G}_i$ matrix for each LMO means that N_{occ} matrices need to be inverted, as shown in Equation (3.52). At first glance, this seems to have a huge impact on the computational efficiency of LDF-EXX. However, this is not the case. Since all $\mathcal{A}_i$ are local, the addition of one atom to a sufficiently large system does not alter the size of the $\mathcal{A}_i$ sets. Thus, even when matrix inversion has a formal M_{local}^3 scaling, growing the system size translates into an N_{occ} scaling because M_{local} remains constant. The same is true for the computational cost of the $\mathbf{E}_i$ matrices.

An important property of our LDF-EXX algorithm is that, different to the PARI[151,152] and CADF[153] approaches, the fitting remains negative definite. As a consequence, all $\mathbf{G}_i$ can be decomposed as:

$$\mathbf{G}_i = \mathbf{L}_i \mathbf{L}_i^T \tag{3.53}$$

With the aid of the Cholesky vectors $\mathbf{L}_i$, we can transform Equation (3.52) into:

$$\mathbf{X}_i = -\mathbf{E}_i \mathbf{Z}_i \mathbf{Z}_i^T \mathbf{E}_i^T \tag{3.54}$$

$$= -\mathbf{H}_i \mathbf{H}_i^T \tag{3.55}$$

where

$$\mathbf{Z}_i = (\mathbf{L}_i^{-1})^T \tag{3.56}$$

and

$$\mathbf{H}_i = \mathbf{E}_i \mathbf{Z}_i \tag{3.57}$$

The advantage of employing Equation (3.55) instead of Equation (3.52) is that computing the inverse of a triangular matrix is much more efficient than computing the inverse of a general dense matrix. Moreover, the symmetric operation appearing in Equation (3.55) can be performed using less memory than the operations appearing in Equation (3.52). It is important to note that the inverse Cholesky factors $\mathbf{Z}_i$ can be obtained in a direct form with the AINV algorithm,[177–180] however, this option has yet to be explored.

In this way, the scaling of the ERI part of the LDF-EXX algorithm has been reduced to N_{occ} with respect to the system size and to M_{local}^3 with respect to the local auxiliary function set size. The memory requirements of the algorithm are also reduced by the LDF-EXX approach. In order to calculate all ERIs and obtain the $\mathbf{X}$ matrix, enough memory to store all $\mathbf{E}_i$ matrices, the LMO coefficients $\mathbf{c}$ and the resulting $\mathbf{X}$ matrix is needed. Once all $\mathbf{E}_i$ matrices are computed, a loop over the occupied LMOs is performed in order to obtain the $\mathbf{X}_i$ contributions. Therefore, the memory to store one $\mathbf{G}_i$ matrix and one $\mathbf{X}_i$ matrix is also required. Note that the memory of $\mathbf{G}_i$ can be reused to store $\mathbf{Z}_i$ in the same way as the memory of $\mathbf{E}_i$ can be reused to store $\mathbf{H}_i$. Thus, the total memory needed is proportional to:

$$T = N_{\text{local}}^2 + M_{\text{local}}^2 + \sum_i^{\text{occ}} N_{\text{local},i} \times M_{\text{local},i} + N^2 + N \times N_{\text{occ}} \tag{3.58}$$

Here, the terms on the rhs corresponds to $\mathbf{X}_i$, $\mathbf{G}_i$, all $\mathbf{E}_i$, $\mathbf{X}$ and $\mathbf{c}$, respectively. If the available memory is not enough to hold all T real numbers, then, batching over the LMOs occurs. The sum over all occupied orbitals in Equation (3.58) and the number of occupied orbitals N_{occ}, are restricted only to those LMOs active in a given batch. As a consequence, the memory required to calculate all quantities in a batch is decreased.

Analytical energy derivatives with respect to nuclear positions are important quantities for geometry optimizations, vibrational frequency analysis and BOMD simulations. The gradient of E with respect to all nuclear displacements is a vector which collects all these derivatives. In Cartesian coordinates, the gradient vector is given by:

$$\mathbf{g} = \left(\frac{\partial E}{\partial A_x} \quad \frac{\partial E}{\partial A_y} \quad \frac{\partial E}{\partial A_z} \quad \frac{\partial E}{\partial B_x} \quad \frac{\partial E}{\partial B_y} \quad \frac{\partial E}{\partial B_z} \quad \cdots \right)^T \tag{3.59}$$

In order to simplify the notation, we will denote the derivative of a function, f, with respect to one of its parameters, λ, as:

$$\frac{\partial f}{\partial \lambda} \equiv f^{(\lambda)} \tag{3.60}$$

To proceed, let us write an ADFT energy expression including both E_{xc} and E_{EXX} as:

$$E[\rho] = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} x_{\bar{k}} \langle \bar{k} \| \mu\nu \rangle P_{\mu\nu} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} G_{\bar{k}\bar{l}} x_{\bar{l}} + E_{\text{xc}}[\tilde{\rho}] + \alpha E_{\text{EXX}} \tag{3.61}$$

where α denotes the fraction of E_{EXX} mixed into a particular global hybrid functional. Thus, an arbitrary component of $\mathbf{g}$ is given by:[181,182]

$$\begin{aligned} E^{(\lambda)} \;=\; & \sum_{\mu,\nu} P^{(\lambda)}_{\mu\nu} H_{\mu\nu} + \sum_{\mu,\nu} P_{\mu\nu} H^{(\lambda)}_{\mu\nu} + \sum_{\mu,\nu} \sum_{\bar{k}} x_{\bar{k}} \langle \bar{k} \| \mu\nu \rangle^{(\lambda)} P_{\mu\nu} + \\ & \sum_{\mu,\nu} \sum_{\bar{k}} x_{\bar{k}} \langle \bar{k} \| \mu\nu \rangle P^{(\lambda)}_{\mu\nu} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} G^{(\lambda)}_{\bar{k}\bar{l}} x_{\bar{l}} + E^{(\lambda)}_{\text{xc}} + \alpha E^{(\lambda)}_{\text{EXX}} \end{aligned} \tag{3.62}$$

The derivatives of the elements of $\mathbf{H}$, $\mathbf{G}$ and the three-center ERIs are obtained through integral recurrence relations. The derivative of E_{xc} is obtained via the chain rule and is given by:

$$\begin{aligned} E_{\text{xc}}[\tilde{\rho}]^{(\lambda)} \;&=\; \int \frac{\delta E_{\text{xc}}[\tilde{\rho}]}{\delta \tilde{\rho}(\mathbf{r})} \frac{\partial \tilde{\rho}(\mathbf{r})}{\partial \lambda} \mathbf{dr} \\ &=\; \sum_{\bar{k}} x^{(\lambda)}_{\bar{k}} \langle v_{\text{xc}}[\tilde{\rho}] \,|\, \bar{k} \rangle + \sum_{\bar{k}} x_{\bar{k}} \langle v_{\text{xc}}[\tilde{\rho}] \,|\, \bar{k}^{(\lambda)} \rangle \end{aligned} \tag{3.63}$$

Furthermore, the here appearing elements of $\mathbf{x}^{(\lambda)}$ can be obtained by deriving Equation (2.29),

$$\mathbf{G}^{(\lambda)}\mathbf{x} + \mathbf{G}\,\mathbf{x}^{(\lambda)} = \mathbf{J}^{(\lambda)}\,, \tag{3.64}$$

from where it follows

$$\mathbf{x}^{(\lambda)} = \mathbf{G}^{-1}\left(\mathbf{J}^{(\lambda)} - \mathbf{G}^{(\lambda)}\mathbf{x}\right) \tag{3.65}$$

Substituting Equation (3.65) into Equation (3.63) yields:

$$\begin{aligned} E_{\mathrm{xc}}[\tilde{\rho}]^{(\lambda)} \;=\; & \sum_{\mu,\nu}\sum_{\bar{k}} P^{(\lambda)}_{\mu\nu}\langle\mu\nu\|\bar{k}\rangle z_{\bar{k}} + \sum_{\mu,\nu}\sum_{\bar{k}} P_{\mu\nu}\langle\mu\nu\|\bar{k}\rangle^{(\lambda)} z_{\bar{k}} - \\ & \sum_{\bar{k},\bar{l}} x_{\bar{k}} G^{(\lambda)}_{\bar{k}\bar{l}} z_{\bar{l}} + \sum_{\bar{k}} x_{\bar{k}}\langle v_{\mathrm{xc}}[\tilde{\rho}]\,|\,\bar{k}^{(\lambda)}\rangle \end{aligned} \tag{3.66}$$

Finally, in order to obtain $E^{(\lambda)}_{\mathrm{EXX}}$, let us rewrite E_{EXX} as

$$E_{\mathrm{EXX}} = -\frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu}\sum_{i}^{\mathrm{occ}}\sum_{\bar{k}} x_{\bar{k}\mu i}\, J_{\bar{k}\nu i}\;, \tag{3.67}$$

where

$$J_{\bar{k}\nu i} = \langle\bar{k}\|\nu i\rangle\;, \tag{3.68}$$

and

$$x_{\bar{k}\mu i} = \sum_{\bar{l}} G^{-1}_{\bar{k}\bar{l}}\, J_{\bar{l}\mu i} \tag{3.69}$$

Remember that in the LDF-EXX approach, the indices appearing in Equations (3.67)-(3.69) are restricted according to the fitting domains of each $\psi_i(\mathbf{r})$. The derivative of Equation (3.67) is given by:

$$E^{(\lambda)}_{\mathrm{EXX}} = -\frac{1}{2}\sum_{\mu,\nu} P^{(\lambda)}_{\mu\nu}\sum_{i}^{\mathrm{occ}}\sum_{\bar{k}} x_{\bar{k}\mu i}J_{\bar{k}\nu i} - \frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu}\sum_{i}^{\mathrm{occ}}\sum_{\bar{k}} x^{(\lambda)}_{\bar{k}\mu i}J_{\bar{k}\nu i} - \frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu}\sum_{i}^{\mathrm{occ}}\sum_{\bar{k}} x_{\bar{k}\mu i}J^{(\lambda)}_{\bar{k}\nu i} \tag{3.70}$$

The derivatives of the $x_{\bar{k}\mu i}$ fitting coefficients can be obtained by using a modified form of Equation (3.65) to yield

$$E_{\mathrm{EXX}}^{(\lambda)} = - \frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu}^{(\lambda)} \sum_{i}^{\mathrm{occ}} \sum_{\bar{k}} x_{\bar{k}\mu i} J_{\bar{k}\nu i} + \frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu} \sum_{i}^{\mathrm{occ}} \sum_{\bar{k},\bar{l}} x_{\bar{k}\mu i} G_{\bar{k}\bar{l}}^{(\lambda)} x_{\bar{l}\nu i} - \frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu} \sum_{i}^{\mathrm{occ}} \sum_{\bar{k}} J_{\bar{k}\mu i}^{(\lambda)} x_{\bar{k}\nu i} - \frac{1}{2}\sum_{\mu,\nu} P_{\mu\nu} \sum_{i}^{\mathrm{occ}} \sum_{\bar{k}} x_{\bar{k}\mu i} J_{\bar{k}\nu i}^{(\lambda)} \tag{3.71}$$

The last two terms of Equation (3.71) are formally equivalent, however, we kept them explicitly separated because

$$P_{\mu\nu}^{(\lambda)} = 2\sum_{i}^{\mathrm{occ}} c_{\mu i}^{(\lambda)} c_{\nu i} + 2\sum_{i}^{\mathrm{occ}} c_{\mu i} c_{\nu i}^{(\lambda)} \ , \tag{3.72}$$

arises naturally from the substitution of

$$J_{\bar{k}\nu i}^{(\lambda)} = \sum_{\mu} \langle \bar{k} \| \mu\nu \rangle^{(\lambda)} c_{\mu i} + \sum_{\mu} \langle \bar{k} \| \mu\nu \rangle c_{\mu i}^{(\lambda)} \tag{3.73}$$

into Equation (3.71). Thus, Equation (3.71) can be rewritten as

$$E_{\mathrm{EXX}}^{(\lambda)} = \sum_{\mu,\nu} P_{\mu\nu}^{(\lambda)} X_{\mu\nu} + \sum_{\bar{k},\bar{l}} \Gamma_{\bar{k}\bar{l}} G_{\bar{k}\bar{l}}^{(\lambda)} - \sum_{i}^{\mathrm{occ}} \sum_{\bar{k}} \sum_{\sigma} \sum_{\mu,\nu} P_{\mu\nu} x_{\bar{k}\mu i} \langle \bar{k} \| \nu\sigma \rangle^{(\lambda)} c_{\sigma i} \ , \tag{3.74}$$

where we have defined an auxiliary matrix $\mathbf{\Gamma}$, with elements

$$\Gamma_{\bar{k}\bar{l}} = \frac{1}{2}\sum_{i}^{\mathrm{occ}} \sum_{\mu,\nu} x_{\bar{k}\mu i} P_{\mu\nu} x_{\bar{l}\nu i} \tag{3.75}$$

Substituting Equations (3.66) and (3.74) into Equation (3.62) yields

$$\begin{aligned} E^{(\lambda)} &= \sum_{\mu,\nu} K_{\mu\nu} P^{(\lambda)}_{\mu\nu} + \sum_{\mu,\nu} P_{\mu\nu} H^{(\lambda)}_{\mu\nu} + \\ &\sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle^{(\lambda)} \left(x_{\bar{k}} + z_{\bar{k}} \right) - \sum_{\bar{k},\bar{l}} x_{\bar{k}} G^{(\lambda)}_{\bar{k}\bar{l}} \left(\frac{1}{2} x_{\bar{l}} + z_{\bar{l}} \right) + \\ &\sum_{\bar{k}} \langle v_{\mathrm{xc}} | \bar{k}^{(\lambda)} \rangle x_{\bar{k}} + \alpha \sum_{\bar{k},\bar{l}} \Gamma_{\bar{k}\bar{l}} G^{(\lambda)}_{\bar{k}\bar{l}} - \alpha \sum_{i}^{\mathrm{occ}} \sum_{\bar{k}} \sum_{\sigma} \sum_{\mu,\nu} P_{\mu\nu} x_{\bar{k}\mu i} \langle \bar{k} \| \nu\sigma \rangle^{(\lambda)} c_{\sigma i} \end{aligned} \tag{3.76}$$

The last two terms of the rhs of Equation (3.76) are the only additions to the energy derivative when performing an hybrid ADFT calculation in comparison to a "pure" ADFT one. In the same manner as in "pure" ADFT, the derivative of $\mathbf{P}$ can be eliminated by means of the Roothaan-Hall equations (2.17) and the orthonormality constraint, Equation (2.1). First, note that the first term of the rhs of Equation (3.76) can be written as

$$\begin{aligned} \sum_{\mu,\nu} K_{\mu\nu} P^{(\lambda)}_{\mu\nu} &= 2 \sum_{i}^{\mathrm{occ}} \sum_{\mu,\nu} c^{(\lambda)}_{\mu i} K_{\mu\nu} c_{\nu i} + 2 \sum_{i}^{\mathrm{occ}} \sum_{\mu,\nu} c_{\mu i} K_{\mu\nu} c^{(\lambda)}_{\nu i} \\ &= 4 \sum_{i}^{\mathrm{occ}} \sum_{\mu,\nu} c^{(\lambda)}_{\mu i} K_{\mu\nu} c_{\nu i} \end{aligned} \tag{3.77}$$

Moreover, substitution of Equation (2.17) into Equation (3.77) yields

$$\sum_{\mu,\nu} K_{\mu\nu} P^{(\lambda)}_{\mu\nu} = 4 \sum_{i}^{\mathrm{occ}} \sum_{\mu,\nu} c^{(\lambda)}_{\mu i} S_{\mu\nu} c_{\nu i} \varepsilon_i \tag{3.78}$$

The derivative of the orthonormality constraint,

$$\sum_{\mu,\nu} c^{(\lambda)}_{\mu i} S_{\mu\nu} c_{\nu j} + \sum_{\mu,\nu} c_{\mu i} S^{(\lambda)}_{\mu\nu} c_{\nu j} + \sum_{\mu,\nu} c_{\mu i} S_{\mu\nu} c^{(\lambda)}_{\nu j} = 0 \quad \forall \quad i,j \ , \tag{3.79}$$

allows the elimination of the derivatives of the MO coefficients by substitution of the identity

$$\begin{aligned} -\sum_{\mu,\nu} c_{\mu i} S^{(\lambda)}_{\mu\nu} c_{\nu i} &= \sum_{\mu,\nu} c^{(\lambda)}_{\mu i} S_{\mu\nu} c_{\nu i} + \sum_{\mu,\nu} c_{\mu i} S_{\mu\nu} c^{(\lambda)}_{\nu i} \\ &= 2\sum_{\mu,\nu} c^{(\lambda)}_{\mu i} S_{\mu\nu} c_{\nu i} \ , \end{aligned} \tag{3.80}$$

into Equation (3.78):

$$\begin{aligned} \sum_{\mu,\nu} K_{\mu\nu} P^{(\lambda)}_{\mu\nu} &= -2\sum_{i}^{\text{occ}} \sum_{\mu,\nu} c_{\mu i} S^{(\lambda)}_{\mu\nu} c_{\nu i} \varepsilon_i \\ &= -\sum_{\mu,\nu} W_{\mu\nu} S^{(\lambda)}_{\mu\nu} \ , \end{aligned} \tag{3.81}$$

where $W_{\mu\nu}$, an element of the closed-shell energy-weighted density matrix, is defined as:

$$W_{\mu\nu} = 2\sum_{i}^{\text{occ}} \varepsilon_i c_{\mu i} c_{\nu i} \tag{3.82}$$

By substituting Equation (3.81) into Equation (3.76), an expression for calculating an element of g without the need of derivatives of the density matrix or the MO coefficients is obtained:

$$\begin{aligned} E^{(\lambda)} = & -\sum_{\mu,\nu} W_{\mu\nu} S^{(\lambda)}_{\mu\nu} + \sum_{\mu,\nu} P_{\mu\nu} H^{(\lambda)}_{\mu\nu} + \\ & \sum_{\mu,\nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle^{(\lambda)} \left(x_{\bar{k}} + z_{\bar{k}} \right) - \sum_{\bar{k},\bar{l}} x_{\bar{k}} G^{(\lambda)}_{\bar{k}\bar{l}} \left(\frac{1}{2} x_{\bar{l}} + z_{\bar{l}} \right) + \\ & \sum_{\bar{k}} \langle v_{\text{xc}}[\tilde{\rho}] \,|\, \bar{k}^{(\lambda)} \rangle x_{\bar{k}} + \alpha \sum_{\bar{k},\bar{l}} \Gamma_{\bar{k}\bar{l}} G^{(\lambda)}_{\bar{k}\bar{l}} - \alpha \sum_{i}^{\text{occ}} \sum_{\bar{k}} \sum_{\sigma} \sum_{\mu,\nu} P_{\mu\nu} x_{\bar{k}\mu i} \langle \bar{k} \| \nu\sigma \rangle^{(\lambda)} c_{\sigma i} \end{aligned} \tag{3.83}$$

As already noted, the additional terms appearing in an hybrid ADFT calculation are the last two terms of Equation (3.83). All other terms are already coded into deMon2k and their implementation will not be further discussed. For a competent discussion of the calculations of these terms we refer the interested reader to references [181] and [182]. The

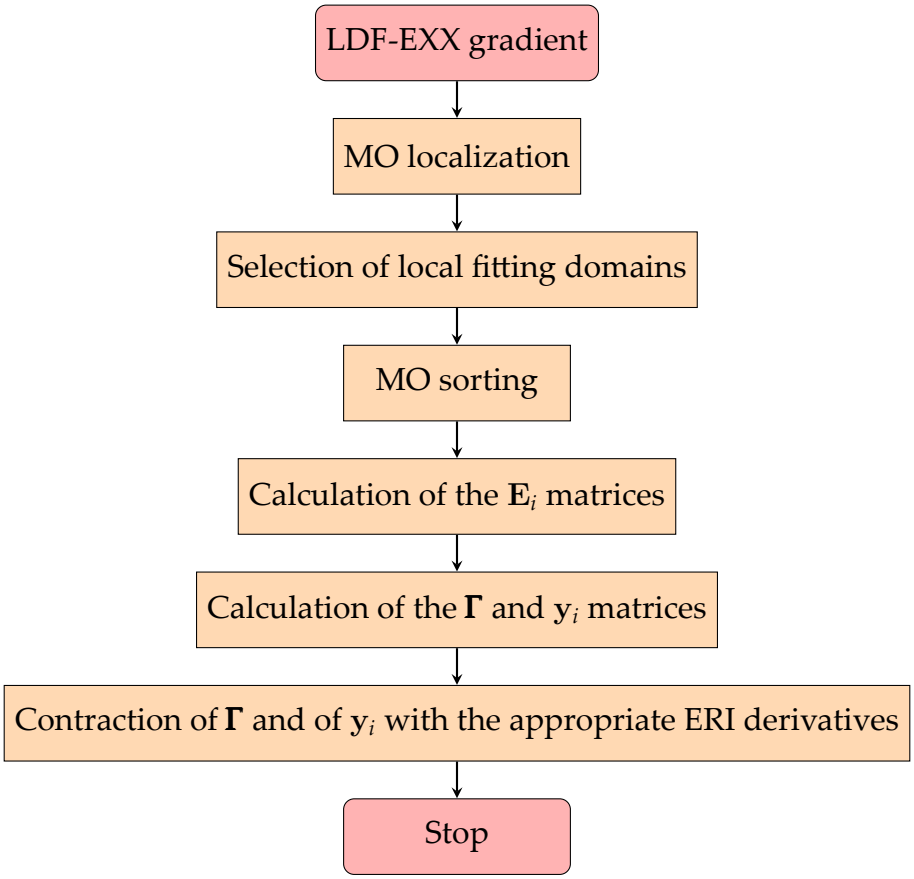


Figure 3.10: Algorithmic flowchart for the calculation of the LDF-EXX gradient.

terms characteristic to EXX gradients are calculated according to the algorithmic flowchart depicted in Figure (3.10). Note that all steps until the calculation of the $\mathbf{E}_i$ matrices are also performed during each SCF iteration, therefore, the same subroutines are used for these steps. After all $\mathbf{E}_i$ matrices have been computed, the auxiliary matrix $\mathbf{\Gamma}$ and the transformed LDF-EXX fitting coefficients $\mathbf{y}_i$ are obtained according to the algorithm represented in Figure 3.11. As in the SCF, the $\mathbf{Z}_i$ matrices are obtained by inversion of a Cholesky factor of each $\mathbf{G}_i$. Then, the half-transformed LDF-EXX coefficients $\mathbf{x}_i$ are obtained as:

$$\mathbf{x}_i = \mathbf{G}_i^{-1}\mathbf{E}_i^T = \mathbf{Z}_i\mathbf{Z}_i^T\mathbf{E}_i^T \tag{3.84}$$

For this transformation, the memory of each $\mathbf{E}_i$ can be reused for each $\mathbf{x}_i$. The next step is to transform the $\mathbf{x}_i$ coefficients to MO representation according to

$$\gamma_{\bar{k}ij} = \sum_{\mu\in\mathcal{B}_i} x_{\bar{k}\mu i} c_{\mu j} \;\; \forall \;\; \bar{k} \in \mathcal{A}_i \;\; \wedge \;\; j \leqslant N_{\text{occ}} \tag{3.85}$$

Thus, a new field with dimensions $M_{\text{local}} \times N_{\text{occ}}$ is needed. This field is immediately used to obtain the corresponding contribution to $\mathbf{\Gamma}$ as

$$\mathbf{\Gamma}_i = \sum_{j}^{\text{occ}} \gamma_{\bar{k}ij}\gamma_{\bar{l}ij} \;\; \forall \;\; \bar{k},\bar{l} \in \mathcal{A}_i \tag{3.86}$$

Note that in Equation (3.86), only a submatrix $\mathbf{\Gamma}_i$ is obtained, and the elements of this submatrix have to be summed into the appropriate blocks of the full $\mathbf{\Gamma}$ matrix. In order to store $\mathbf{\Gamma}_i$, the memory of $\mathbf{Z}_i$ is reused. Furthermore, only a given column of $\mathbf{\Gamma}$ is read and written at a time, thus avoiding the allocation of a full M^2 field. For the efficient calculation of the last term in Equation (3.83) we introduce new auxiliary coefficients $y_{\bar{k}\sigma i}$ and $y_{\bar{k}\sigma\nu}$ defined as:

$$y_{\bar{k}\sigma i} = \sum_{j}^{\text{occ}} \gamma_{\bar{k}ij}\, c_{\sigma j} \;\; \forall \;\; \bar{k} \in \mathcal{A}_i \;\; \wedge \;\; \sigma \in \mathcal{B}_i \tag{3.87}$$

and

$$y_{\bar{k}\sigma\nu} = \sum_{i}^{\text{occ}} y_{\bar{k}\nu i} c_{\sigma i} \tag{3.88}$$

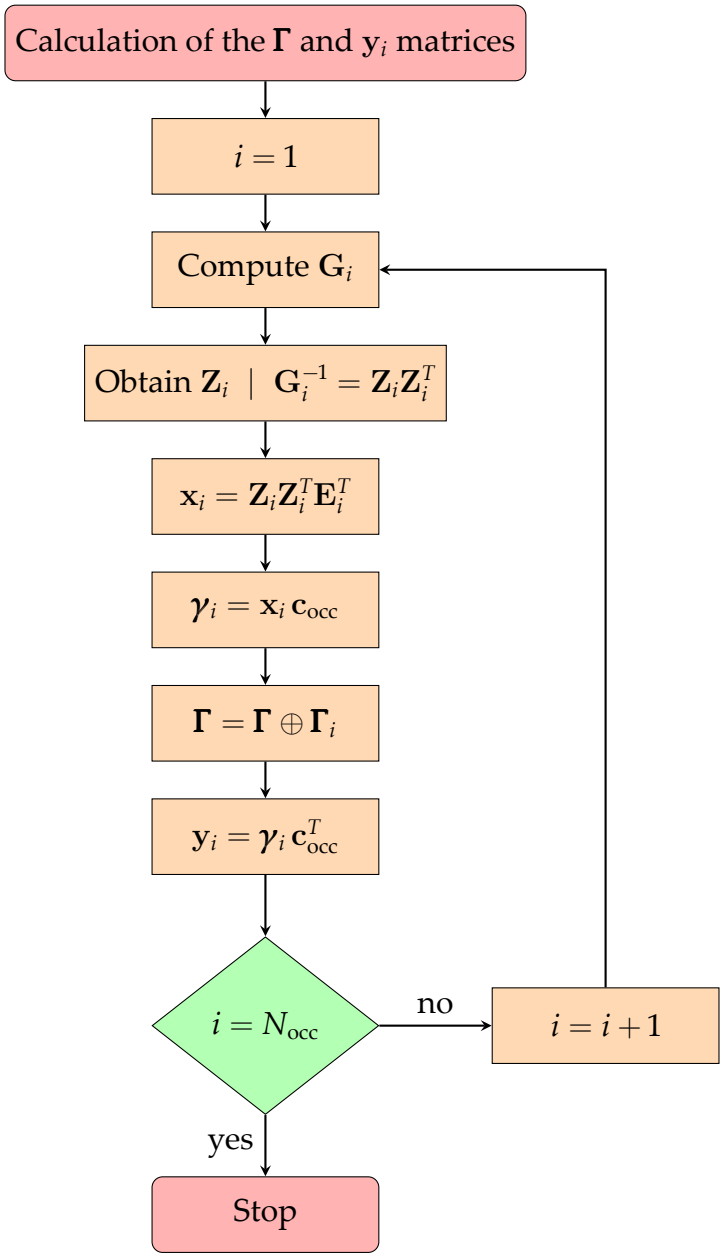


Figure 3.11: Flowchart for the calculation of the LDF-EXX gradient auxiliary matrices.

The auxiliary coefficients $y_{\bar{k}\sigma\nu}$ can be directly contracted with the ERI derivatives appearing in the last term of Equation (3.83). Thus, the calculation of this term becomes:

$$\sum_{i}^{\text{occ}} \sum_{\bar{k}} \sum_{\sigma} \sum_{\mu,\nu} P_{\mu\nu} x_{\bar{k}\mu i} \langle \bar{k} \| \sigma\nu \rangle^{(\lambda)} c_{\sigma i} = \sum_{\bar{k}} \sum_{\sigma,\nu} y_{\bar{k}\sigma\nu} \langle \bar{k} \| \sigma\nu \rangle^{(\lambda)} \tag{3.89}$$

The memory needed to store all $\mathbf{y}_i$ coefficients is reused from the one used to store all $\mathbf{x}_i$ coefficients. The flowchart depicted in Figure 3.11 shows the steps performed to compute $\mathbf{\Gamma}$ and the auxiliary coefficients $\mathbf{y}_i$. The direct sum

$$\mathbf{\Gamma} = \mathbf{\Gamma} \oplus \mathbf{\Gamma}_i \tag{3.90}$$

appearing in Figure 3.11 means that the calculated $\mathbf{\Gamma}_i$ is summed into the appropriate blocks of $\mathbf{\Gamma}$. As already mentioned, $\mathbf{\Gamma}$ is updated vector by vector and stored on disk. On the other hand, all $\mathbf{y}_i$ coefficients are kept in memory, transformed into shell blocks to yield $y_{\bar{k}\sigma\nu}$ and contracted with the corresponding shell blocks of ERI derivatives $\langle \bar{k} \| \sigma\nu \rangle^{(\lambda)}$.

In terms of memory, only one additional temporary field, with dimension $M_{\text{local}} \times N_{\text{occ}}$, needs to be allocated. The advantage is that no N_{local}^2 and N^2 fields, for the local contribution of the EXX matrix $\mathbf{X}_i$ and the EXX matrix $\mathbf{X}$, respectively, are needed. Therefore, the memory needed for the LDF-EXX gradients is given by:

$$T = M_{\text{local},i}^2 + N_{\text{occ}} \times M_{\text{local},i} + \sum_{i}^{\text{occ}} N_{\text{local},i} \times M_{\text{local},i} + N \times N_{\text{occ}} \tag{3.91}$$

If the memory is not sufficient to store all quantities, then the calculation is split into batches over the occupied LMOs, in an analog manner as for the LDF-EXX potential computation. In this case the loop over all occupied orbitals appearing in Figure (3.11) is restricted to only those LMOs active in a given batch. Thus, only a subset of the $y_{\bar{k}\sigma i}$ coefficients and Γ_i matrices are computed. The contraction of the contribution to the $y_{\bar{k}\sigma\nu}$ coefficients with the ERI derivatives is done in each batch. Differently, the contraction of $\mathbf{\Gamma}$ with $\mathbf{G}^{(\lambda)}$ is performed after all $\mathbf{\Gamma}_i$ contributions are summed, i.e. at the end of all

batches.

The parallelization of the LDF-EXX gradients is performed exactly as for the LDF-EXX potential. The LMOs are distributed among all cores. Moreover, the octree algorithm is also employed to maximize the reuse of three-center ERIs. When memory is not sufficient batching also occurs. Thus, it can be seen that the LDF-EXX gradient computation is consistent with the LDF-EXX potential one, up to the point that the same batching and parallelization paradigms are used.

3.4 Validation and Benchmarks

3.4.1 Validation of thermochemical parameters

In order to validate our LDF-EXX approach, we compare standard heats of formation, ΔH_f^{298K}, obtained for the B3LYP,[183] PBE0,[184,185] and M06-2X[186] DFAs and the HF method, employing the LDF-EXX algorithm and the standard four-center ERI exact exchange implemented in NWChem,[187] with available experimental data.[188] Note that the hybrid GGAs, B3LYP and PBE0, are implemented in both hybrid ADFT and DF-DFT versions. However, the hybrid meta-GGA can only be implemented for the DF-DFT approach due to the explicit dependence on the KS-MOs via the kinetic energy density, $\tau(\mathbf{r})$. The corresponding energy expressions are given by:

$$E_{\mathrm{xc}}^{\mathrm{B3LYP}} = 0.81\,E_{\mathrm{c}}^{\mathrm{LYP}} + 0.19\,E_{\mathrm{c}}^{\mathrm{VWN}} + 0.80\,E_{\mathrm{x}}^{\mathrm{Dirac}} + 0.72\,\Delta E_{\mathrm{x}}^{\mathrm{B88}} + 0.20\,E_{\mathrm{EXX}} \tag{3.92}$$

$$E_{\mathrm{xc}}^{\mathrm{PBE0}} = E_{\mathrm{c}}^{\mathrm{PBE}} + 0.75\,E_{\mathrm{x}}^{\mathrm{PBE}} + 0.25\,E_{\mathrm{EXX}} \tag{3.93}$$

$$E_{\mathrm{xc}}^{\mathrm{M06-2X}} = E_{\mathrm{c}}^{\mathrm{M06-2X}} + E_{\mathrm{x}}^{\mathrm{M06-2X}} + 0.54\,E_{\mathrm{EXX}} \tag{3.94}$$

Here Δ means that only the "non-local" part of the Becke88 exchange functional[189] is used. It is important to note that the implementation of the M06 correlation functionals differ slightly from the ones proposed in reference [186] by modifying the self-correlation correction term inside the VS98 (VSXC)[190,191] contribution and the M06 part, as suggested by Gräfenstein *et al.*[192] This modification avoids a singularity occurring in the equal-spin part

of the correlation energy, leaving the total energy values essentially unchanged. As mentioned in reference [192], a similar modification can also be used in the PKZB[193,194] and TPSS[195] meta-GGA functionals.

Standard heats of formation were obtained by the method proposed by Curtiss *et al.*[196] for the 223 molecules of the G3/99 test set,[196,197] using the Def2-TZVPP[198] basis set in spherical representation and the B3LYP/6-31G(2df,p) optimized geometries.[199] Geometries, zero-point energy corrections and enthalpy corrections at 298 K were obtained at the B3LYP/6-31G(2df,p) level of theory using Gaussian 09.[200] In short, in order to calculate ΔH_f^{298K} for a general molecule with formula $A_xB_yC_z$, the following steps must be performed:

1. Calculate the zero-point corrected atomization energy D_0 as:

$$D_0(A_xB_yC_z) = \Big(xE_e(A) + yE_e(B) + zE_e(C)\Big) - E_e(A_xB_yC_z) + E_{zpe}(A_xB_yC_z) \tag{3.95}$$

2. Calculate the enthalpy of formation at 0 K as:

$$\Delta H_f^{0K}(A_xB_yC_z) = \Big(x\Delta H_f^{0K}(A) + y\Delta H_f^{0K}(B) + z\Delta H_f^{0K}(C)\Big) - D_0(A_xB_yC_z) \tag{3.96}$$

3. Calculate ΔH_f^{298K} as:

$$\begin{aligned}\Delta H_f^{298K}(A_xB_yC_z) = {} & \Delta H_f^{0K}(A_xB_yC_z) + H_{298K}(A_xB_yC_z) - H_{0K}(A_xB_yC_z) - \\ & x\Big(H_{298K}(A) - H_{0K}(A)\Big) - y\Big(H_{298K}(B) - H_{0K}(B)\Big) - \\ & z\Big(H_{298K}(C) - H_{0K}(C)\Big)\end{aligned} \tag{3.97}$$

In the above scheme, $\Delta H_f^{0K}(X)$ and $H_{298K}(X) - H_{0K}(X)$ for a given atom X is taken from reference [196]. Table 3.2 shows mean deviations (MD), mean absolute deviations (MAD) and maximum absolute deviations (MaxAD), in kcal/mol, of the calculated HF and experimental ΔH_f^{298K}.[188] Three different HF approaches were used for this purpose. The first approach, labeled as NWChem, is the standard four-center HF implementation of NWChem and is included as reference. The second and third approaches use the LDF-

Table 3.2: Errors for Hartree-Fock standard heats of formation [kcal/mol] with respect to experiment. All quantities were calculated with the Def2-TZVPP basis set in spherical representation. The NWChem column refers to a standard four-center HF implementation. The A2*/A2* uses the GEN-A2* auxiliary function set for both the SCF and the final energy calculation, whereas the A2/A2* calculations were performed with the GEN-A2 for the SCF and the GEN-A2* auxiliary function set for the final energy calculation. See text for more details.

	NWChem	A2*/A2*	A2/A2*
MD	212.8	212.7	210.2
MAD	212.8	212.7	210.2
MaxAD	582.0	582.0	577.2

EXX algorithm and differ only in the auxiliary function set used. The results shown in the column "A2*/A2*" were obtained by using the GEN-A2*[201] auxiliary function set. In the case of the "A2/A2*" approach, the wavefunction was obtained with the GEN-A2 auxiliary function set and an additional non-self-consistent energy calculation was performed using the GEN-A2* auxiliary function set. The results shown in Table 3.2 demonstrates that the LDF-EXX HF approach is almost indistinguishable to a standard four-center HF implementation in terms of accuracy achieved with an appropriate auxiliary function set. Furthermore, even the GEN-A2 auxiliary function set provides accurate enough HF MOs. This follows from the fact that the A2/A2* approach gives results within 1.5 kcal/mol of the A2*/A2* approach, even though the A2/A2* energy is non-self-consistent. The individual ΔH_f^{298K} for the 223 molecules of the G3/99 test set are listed in Table B.1 of Appendix B. It is important to note that, in the case of the A2*/A2* approach, ΔH_f^{298K} deviations for the Si-containing molecules are specific for the basis set used (Def2-TZVPP) and are due to the automatically generated auxiliary function set. These deviations vanish when using, for example, the more systematically augmented Dunning basis sets.

Table 3.3: Errors [kcal/mol] for hybrid DFT standard heats of formation employing the Def2-TZVPP basis set in spherical representation. MD is the mean deviation, MAD is the mean absolute deviation and MaxAD is the maximum absolute deviation. See Table 3.2 for the used abbreviations.

	B3LYP			PBE0			M06-2X		
	NWChem	A2*/A2*	A2/A2*	NWChem	A2*/A2*	A2/A2*	NWChem	A2*/A2*	A2/A2*
MD	3.8	4.9	3.4	−4.2	−3.6	−5.6	−1.3	−0.9	−2.6
MAD	5.0	5.8	4.8	6.4	5.9	7.6	2.4	2.3	3.5
MaxAD	20.9	21.5	24.2	36.3	35.9	42.3	16.2	15.9	20.6

The ΔH_f^{298K} for the full G3/99 test set were also computed for three hybrid DFAs: B3LYP, PBE0 and M06-2X. Table 3.3 shows MD, MAD and MaxAD for these hybrid DFAs as implemented in deMon2k. Note that B3LYP and PBE0 results were obtained with the ADFT methodology, whereas the M06-2X were obtained with the DF-DFT one. Since the fraction of EXX mixed in these three hybrid DFAs is rather small—see Equations (3.92)-(3.94)—the major differences between NWChem and deMon2k results arise from the different DFT methodologies employed. In particular, Table 3.3 shows that B3LYP is not as well suited for ADFT calculations as for DFT, and that PBE0 is better suited for ADFT calculations than for DFT ones.

In addition to ΔH_f^{298K}, we also compared chemical reaction barrier heights calculated with our LDF-EXX algorithm. For this purpose, the HTBH38/08 and NHTBH38/08 databases were employed.[202–204] The HTBH database contains 19 hydrogen transfer reactions with values for the forward and reverse classical reactions barrier heights. The NHTBH database contains 19 reactions, including heavy-atom transfer, bimolecular nucleophilic substitution, association, and unimolecular reactions. The "best estimates" provided in the Minnesota databases[205] are used as experimental references for the validation. All calculations were performed with the spherical representation of the Def2-TZVPP basis set and with the structures optimized at the QCISD/MG3 level.[204] Table 3.4 shows MADs of the three hybrid DFAs for the chemical reaction barrier heights. Note that ADFT results are consistently better than their DFT counterparts. Furthermore, the A2/A2* approach gives more accurate results than the A2*/A2*. This result is counter-intuitive, given that the GEN-A2* set is larger than the GEN-A2 set, however, it is a direct consequence of ill-conditioned $\mathbf{G}$ matrices. These ill-conditioned matrices are more likely to occur when a calculation is performed with very large auxiliary functions sets, such as the GEN-A2* set. Clearly, these validation calculations indicate that the solution method for the fitting equation system, Equation (2.29), must be revisited if accuracies below 1 kcal/mol should be addressed. However, this is outside the scope of this work.

Table 3.4: Mean absolute errors [kcal/mol] for hybrid DFT chemical reaction barrier heights employing the Def2-TZVPP basis set in spherical representation. See Table 3.2 for the used abbreviations.

	B3LYP			PBE0			M06-2X		
	NWChem	A2*/A2*	A2/A2*	NWChem	A2*/A2*	A2/A2*	NWChem	A2*/A2*	A2/A2*
HTBH forward	4.29	4.16	3.91	4.05	3.66	3.61	1.04	1.10	1.08
HTBH backward	4.65	4.62	4.47	4.90	4.48	4.77	1.38	1.53	1.51
NHTBH forward	6.09	6.01	5.11	4.60	4.54	3.67	1.76	1.76	1.63
NHTBH backward	5.24	5.21	4.61	4.45	4.38	3.89	1.92	1.88	1.84

The results of the M06-2X DFA are also interesting. Remember that the M06-2X implementation in deMon2k is a slightly modified form of the original M06-2X functional. Nevertheless, the results of a standard implementation of this functional using four-center ERIs are almost identical to the DF-DFT results obtained with deMon2k. We have also tested the adaptive grid accuracy for the M06-2X chemical reaction barrier heights since it has been reported that reaction energies obtained with meta-GGAs may be very sensitive to the integration grid employed, specially for the M06 functional family.[206] Therefore, we performed additional calculations with the M06-2X functional and the FINE adaptive grid of deMon2k, yielding essentially the same results as the ones reported in Table 3.4.

In summary, the LDF-EXX approach is almost indistinguishable to standard four-center EXX implementations for computing thermochemical data with either HF or hybrid DFAs. Furthermore, chemical reaction barrier heights are consistently better for all hybrid DFAs and the LDF-EXX ADFT method, especially with the non-self-consistent A2/A2* approach. Finally, the modified version of the M06 suite implemented in deMon2k is very stable and yields accurate results with the default grid settings of the program.

A remarkable feature of these results is that neither the GEN-A2 nor the GEN-A2* auxiliary functions sets were developed for the variational approximation of exact exchange. Yet, their performance can be compared to the rather large auxiliary function sets developed by Weigend and specifically designed to fit the exact exchange.[104,207] Furthermore, the GEN-A2 and GEN-A2* auxiliary functions sets are automatically generated and can adapt to whichever basis set is employed.

3.4.2 Validation of geometrical parameters

The LDF-EXX energy gradients were validated by optimizing the 148 molecules of the G2/97 molecular test set.[196] The optimizations were performed with the Def2-TZVPP basis set using either the standard HF implementation of NWChem or the LDF-EXX HF implementation of deMon2k. Both codes used their default settings for SCF and optimization convergence criteria. Table 3.5 shows MD, MAD and MaxAD for all bond lengths and bond angles of the full G2/97 set. For this purpose, we defined a bond between two atoms

Table 3.5: Deviation of bond lengths and bond angles of LDF-EXX with respect to standard HF. See Table 3.2 for the used abbreviations.

	Bonds [pm]		Angles [°]	
	A2*/A2*	A2/A2*	A2*/A2*	A2/A2*
MD	0.01	0.16	0.00	0.00
MAD	0.02	0.18	0.05	0.07
MaxAD	0.29	3.00	0.63	0.62

when their distance was smaller than $\frac{3}{2}$ times the sum of their covalent radii. Moreover, bond angles were defined only for bonded triads. With these definitions, the results presented in Table 3.5 correspond to 751 bond lengths and 1202 bond angles. As Table 3.5 shows the GEN-A2*/GEN-A2* approach yields optimized geometrical parameters that are basically indistinguishable from a standard four-center HF implementation. Excellent results can also be obtained with the GEN-A2/GEN-A2* approach, with a MAD of only 0.18 pm. Only 8 bonds, of the 751 compared, deviated more than 1 pm and involved either S or Cl atoms. The largest deviation was in the Cl−N bond present in nitrosyl chloride (ClNO), which was 3.0 pm shorter than the NWChem reference. This underlines the excellent cost-performance ratio that can be obtained with the GEN-A2/GEN-A2* LDF-EXX approach even when, as already mentioned, neither the GEN-A2 nor GEN-A2* auxiliary functions sets were developed to fit exact exchange.

3.4.3 Timings

In order to demonstrate the computational performance of the LDF-EXX approach, we compared the average time needed to compute the Fock, or KS, matrix in one SCF cycle for three hydrocarbon systems with different dimensionalities. All timings reported in this section were obtained on an Intel® Xeon® X5675 processor with a clock rate of 3.07 GHz and 4 GB of memory, unless otherwise stated. As one-dimensional systems we used linear alkane chains, $C_{10n}H_{20n+2}$, with $1 \leqslant n \leqslant 10$. Thus, 10 carbon atoms are added to the chain by each increase of n. The geometrical parameters that define the chains are given in Table 3.6. The average time, in minutes, for one Fock matrix build (EXX and Coulomb) using the 6-31G** basis set[208,209] is shown in Figure 3.12. The results are compared with two

Table 3.6: Geometrical parameters of the linear alkane chains.

Parameter	Value	Units
C-C	1.500	Å
C-H	1.000	Å
∠ C-C-C	109.47	°
∠ C-C-H	109.47	°
∠ C-C-C-C	180.0	°

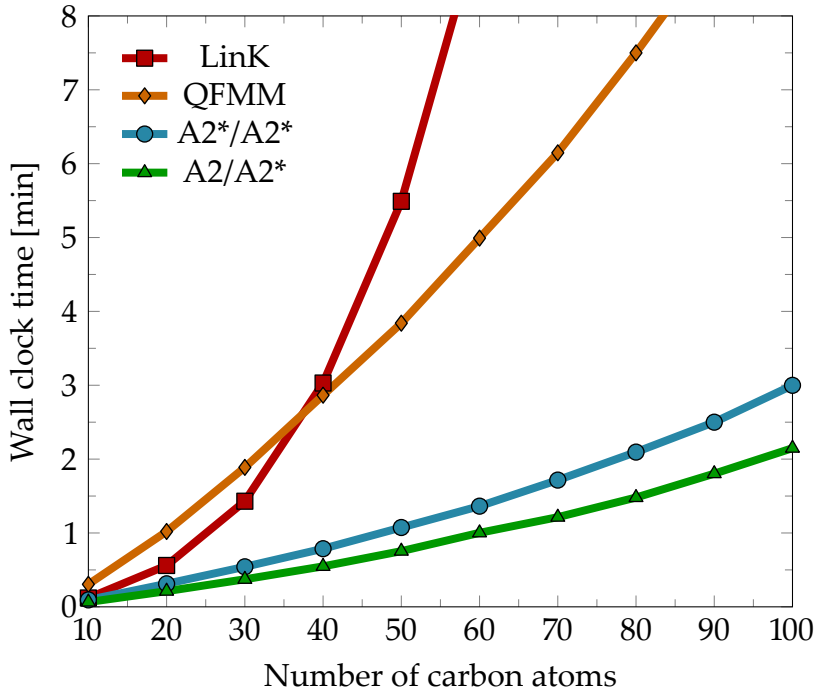


Figure 3.12: Fock matrix build average timings [min] for linear alkane chains employing the 6-31G** basis set. LinK — refers to the default HF implementation in GAMESS, QFMM — refers to the GAMESS HF implementation using the quantum fast multipole method to calculate Coulomb interactions, A2*/A2* — refers to the LDF-EXX implementation in deMon2k using the GEN-A2*/GEN-A2* approach, and A2/A2* — refers to the LDF-EXX implementation in deMon2k using the GEN-A2/GEN-A2* approach.

different four-center HF implementations of GAMESS.[210] The data labeled "LinK" refer to the default HF GAMESS implementation which uses the LinK algorithm[130,131] to compute the exchange contribution; the timings labeled "QFMM" were obtained using the quantum fast multipole method also implemented in GAMESS[211–215] to calculate the Coulomb contribution and the LinK algorithm for the exchange part; the two remaining data sets, "A2/A2*" and "A2*/A2*", refer to timings obtained with the LDF-EXX algorithm for the exact exchange contribution and the double asymptotic expansion for the Coulomb part.[109] It is evident that the LDF-EXX HF method implemented in deMon2k is faster than any of the

HF methodologies implemented in GAMESS. For the $C_{80}H_{162}$ chain, LDF-EXX achieves already a speed-up of one order of magnitude with respect to the standard HF implementation of GAMESS. Even when compared to the combined QFMM+LinK methodology, LDF-EXX is 3.5× faster with the GEN-A2*/GEN-A2* approach and 5.0× faster with the GEN-A2/GEN-A2* approach. It is important to note that the timings of the GEN-A2/GEN-A2* approach include the contribution of the final non-self-consistent energy calculation with the GEN-A2* auxiliary function set. Furthermore, the LDF-EXX algorithm implemented in deMon2k builds the full exchange contribution each cycle, thus, only the Coulomb part takes advantage of the incremental building discussed in Sections 2.2 and 2.3. In contrast, both HF implementations in GAMESS take advantage of the incremental building of the Fock matrix for both exchange and Coulomb contributions. It is also worth noting that the *total energy* difference between the HF solutions of the two codes did not exceed 1.0 kcal/mol for $C_{100}H_{202}$ when comparing the GEN-A2*/GEN-A2* approach.

Changing the basis set from 6-31G** to cc-pVTZ[216] results in a marked increase of the speed-up factor. Taking as example the $C_{40}H_{82}$ chain, the speed-up factor changed from 3.8× to 17.5× when comparing the timings of the GEN-A2*/GEN-A2* with the LinK approach. In fact, convergence of the HF wavefunction is achieved within 50 min with the LDF-EXX algorithm, which is half of the time needed to complete one SCF iteration with GAMESS! This is inline with the previous discussion about the scaling with respect to the basis set size, namely, the M_{local}^3 scaling for LDF-EXX compared to the N^4 scaling of a standard HF implementation. Note that the comparison is not with respect to the faster QFMM+LinK method because the requested memory for this approach (17.5 GB) exceeded the computational resources assigned. Moreover, the size of the molecule is not big enough to notice a significant difference between LinK and QFMM+LinK (see Figure 3.12).

Timings of the LDF-EXX energy gradients were also compared to the standard HF implementation and the linear scaling QFMM+LinK HF implementations of GAMESS. Figure 3.13 shows timings to compute the complete HF energy gradient (one electron, Coulomb and exchange) employing the 6-31G** basis set. Again the improved perfor-

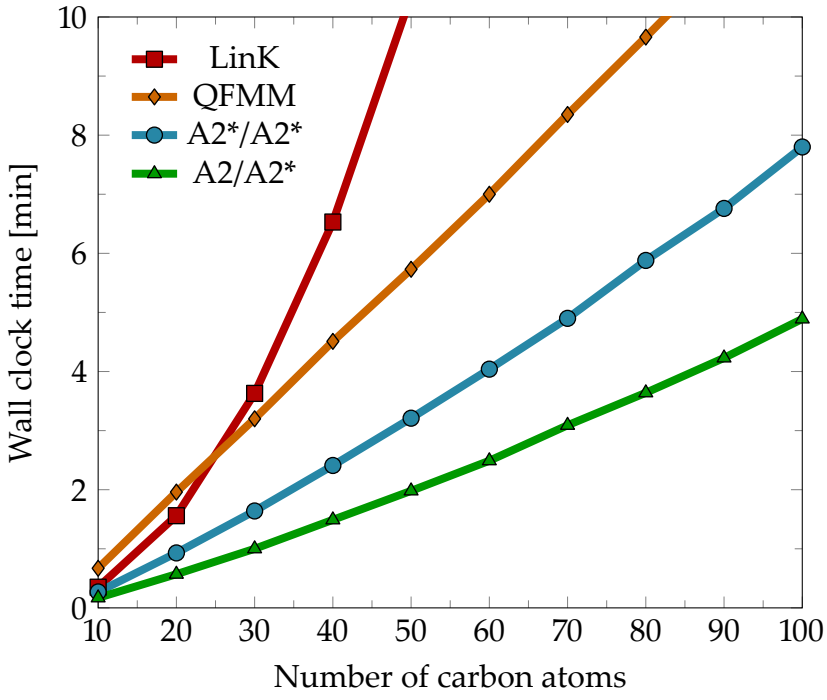


Figure 3.13: Timings [min] for the full HF energy gradient of linear alkane chains employing the 6-31G** basis set. LinK — refers to the default HF implementation in GAMESS, QFMM — refers to the GAMESS HF implementation using the quantum fast multipole method to calculate Coulomb interactions, A2*/A2* — refers to the LDF-EXX implementation in deMon2k using the GEN-A2*/GEN-A2* approach, and A2/A2* — refers to the LDF-EXX implementation in deMon2k using the GEN-A2/GEN-A2* approach.

mance of our new LDF-EXX algorithm is clearly visible.

Noticing such huge speed-up factors of LDF-EXX with respect to traditional HF implementations, we decided to perform the comparisons of two-dimensional and three-dimensional systems with respect to the computationally more efficient DF-DFT and ADFT approaches. As two-dimensional systems, we used saturated graphite sheets, $C_{2n^2}H_{4n}$, with $2 \leqslant n \leqslant 8$. Figure 3.14 shows the average time to compute one SCF iteration of LDF-EXX HF in comparison with the corresponding times for DF-DFT PBE and ADFT PBE calculations on these systems. As Figure 3.14 shows, LDF-EXX HF and DF-DFT GGA calculations need almost the same time for one SCF iteration. In contrast, the ADFT approach achieves an almost linear scaling behavior very soon, and keeps far apart from the other two methodologies with increasing system sizes. It is important to note that the times shown in Figure 3.14 do not include the final LDF-EXX energy calculation. Besides benchmarking the computational efficiency of the LDF-EXX algorithm, saturated graphite sheets also test the MO localization algorithm because they are highly delocalized systems.

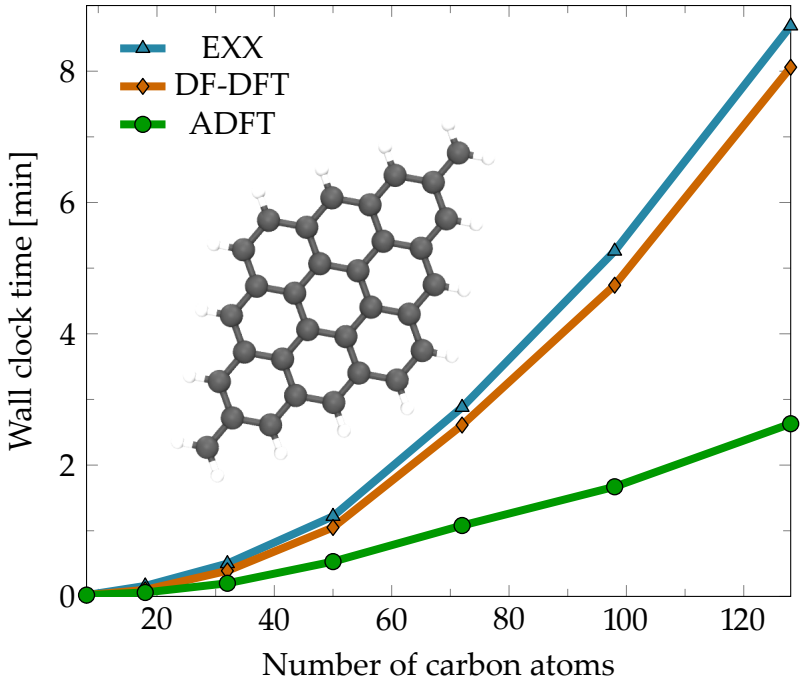


Figure 3.14: Fock and Kohn-Sham matrix build average timings [min] for saturated graphite sheets employing the 6-31G** basis set. EXX — refers to the LDF-EXX HF calculations using the GEN-A2*/GEN-A2* approach, DF-DFT — refers to the DF-DFT PBE calculations, and ADFT — refers to the ADFT PBE calculations.

The MO localization algorithm took no more than 10 seconds per iteration in the largest system, $C_{128}H_{32}$. This represent less than 2% of the total computational time.

Finally, saturated diamond unit cells $C_{2n^3}H_{6n^2}$, with $2 \leqslant n \leqslant 6$, were used to benchmark three-dimensional systems. These systems have n diamond unit cells in each Cartesian direction. Figure 3.15 shows the average time needed for one SCF iteration for LDF-EXX HF, DF-DFT PBE and ADFT PBE calculations on these systems. The results are similar to the ones for the saturated graphite sheets in Figure 3.14. As expected, ADFT calculations are much faster than LDF-EXX HF or DF-DFT ones. However, one important difference with respect to Figure 3.14 arises: the LDF-EXX HF calculations are always faster than DF-DFT ones!

In conclusion, LDF-EXX calculations show, by and large, the same computational efficiency as DF-DFT ones with a GGA-type functional. Furthermore, LDF-EXX can achieve speed-ups of 1 to 2 orders of magnitude with respect to traditional four-center implementations even with basis sets as small as the double-ζ 6-31G**. With larger basis sets this difference becomes even larger. Thus, our new LDF-EXX algorithm can be straightforward used

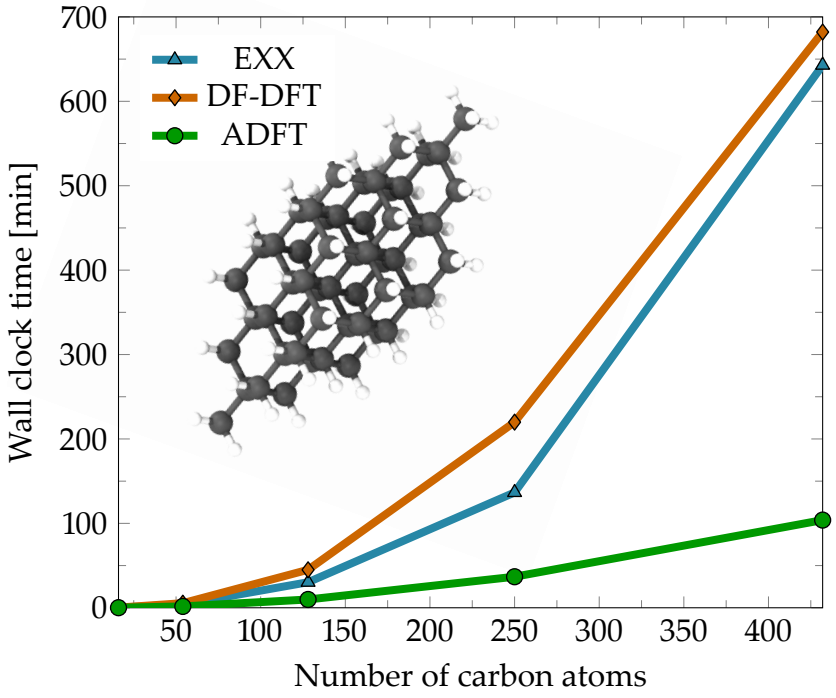


Figure 3.15: Fock and Kohn-Sham matrix build average timings [min] for saturated diamond unit cells employing the 6-31G** basis set. EXX — refers to the LDF-EXX HF calculations using the GEN-A2*/GEN-A2* approach, DF-DFT — refers to the DF-DFT PBE calculations, and ADFT — refers to the ADFT PBE calculations.

for hybrid DF-DFT calculations with a minimum impact on the scaling of this methodology. On the other hand, LDF-EXX will still be the computational bottleneck in ADFT hybrid calculations. A possibility to overcome this computational bottleneck is the combination of LDF-EXX with the so-called Auxiliary Density Matrix Method (ADMM).[217,218] The working formulas for such a combination are presented in section 6.2.2 of this thesis.

3.4.4 Parallelization

The parallelization of the code was benchmarked with the same one-, two- and three-dimensional systems with and without the octree sorting algorithm. The wall clock time needed to build one EXX contribution with 1, 2, 4, 8 and 16 CPUs is shown in Figure 3.16. The benchmarked systems were the linear alkane chain $C_{50}H_{102}$, the saturated graphite sheet $C_{128}H_{32}$ and the saturated diamond $C_{128}H_{96}$ employing the 6-31G** and the GEN-A2/GEN-A2* approach. As Figure 3.16 shows, the use of the octree algorithm to sort the LMOs among the CPUs makes a huge impact on the efficiency of the LDF-EXX algorithm

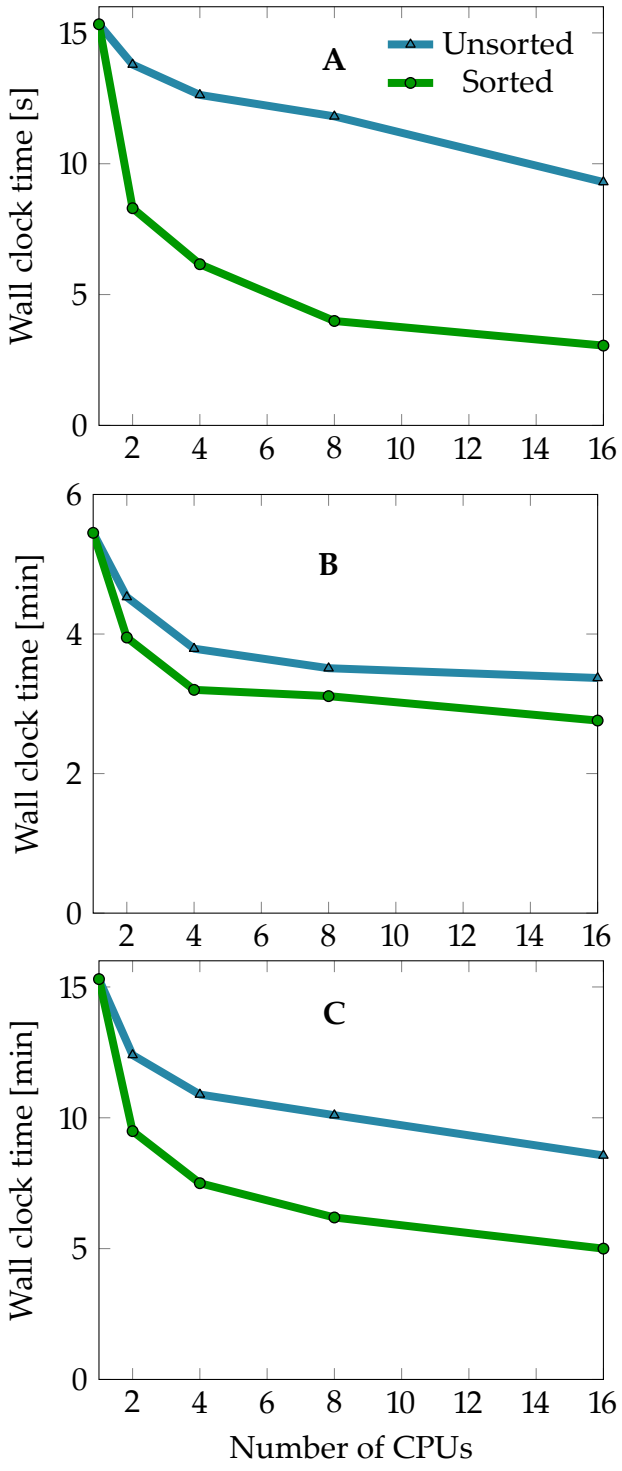


Figure 3.16: Influence of the octree algorithm on the average time to build one EXX contribution in parallel for one-, two- and three-dimensional systems. Plot **A** shows the time [s] for the linear alkane chain $C_{50}H_{102}$. Plot **B** shows the time [min] for the saturated graphite sheet $C_{128}H_{32}$. Plot **C** shows the time [min] for the saturated diamond $C_{128}H_{96}$. All calculations employed the 6-31G** basis set and the GEN-A2/GEN-A2* approach.

parallelization, specially for the one-dimensional systems. For this case, the sorting algorithm can speed-up the parallelization of the code almost by a factor of 3 for small number of processor count. The octree algorithm has also a positive impact on the efficiency of the LDF-EXX parallelization for two-dimensional and three-dimensional systems. However, as Figure 3.16B shows, the time to build one EXX potential for the two-dimensional saturated graphite sheets is less accelerated by the sorting algorithm. This behavior is a consequence of the extension of LMOs, which are less local in conjugated systems compared to non-conjugated ones. The three-dimensional saturated diamond unit cells corroborate this observation. Figure 3.16C shows that the boost achieved by the octree sorting algorithm is intermediate between the one-dimensional and three-dimensional cases.

It is important to note that the time needed to sort the LMOs is negligible (around 5 s for these systems) and, therefore, LMOs are always sorted. Also note that when $N_{\text{CPU}} \geqslant N_{\text{occ}}$ the sorting algorithm is not needed.

I would like to emphasize strongly me belief that the era of computing chemists, when hundreds if not thousands of chemists will go to the computing machine instead of the laboratory for increasingly many facets of chemical information, is already at hand. There is only one obstacle, namely that someone must pay for the computing time.

R. S. Mulliken

4

Low-Order Scaling Response Property Evaluation

Many molecular properties can be calculated as the derivative of the total energy with respect to internal or external perturbations. These properties are often observables and, therefore, permit the direct measurement of the molecular electronic structure response to a given perturbation. For this reason they are often named response properties. The connection between energy derivatives and molecular properties can be best seen by expanding the energetic response in a Taylor series around the perturbations. Thus, the perturbed energy can be written as:

$$E(\lambda) = E_0 + \left.\frac{dE}{d\lambda}\right|_{\lambda=0} \lambda + \frac{1}{2}\left.\frac{d^2E}{d\lambda^2}\right|_{\lambda=0} \lambda^2 + \cdots \tag{4.1}$$

where E_0 is the unperturbed reference energy and λ is used to denote a perturbation. Such perturbations can be electromagnetic fields, external charge distributions or nuclear displacements, to name a few. There is a vast number of response properties that are of interest to various research fields. Many of them correspond to second- and higher-order derivatives of the energy and, therefore, these properties depend on the response of the density matrix $\mathbf{P}$. In quantum chemistry, this response can be obtained either from the response of the molecular orbitals or directly from the response of the density matrix. The former yields the so-called Coupled-Perturbed SCF (CPSCF) methods.[219–226] that reduce to the Coupled-Perturbed Kohn-Sham (CPKS) approach in the framework of Kohn-Sham DFT. The latter is the starting point for McWeeny's self-consistent perturbation (SCP) theory.[227–232] Unfortunately, the CPKS equation system possesses a very large dimension and the SCP problem can become rather difficult to converge. Therefore, many recent developments in molecular response theory have been aimed in developing new algorithms that can circumvent these problems.[233–241] All these methods have to solve a nonlinear equation system and, thus, rely on iterative solvers. Noniterative molecular response algorithms have also been developed,[242–250] among which Auxiliary Density Perturbation Theory (ADPT) is a prominent example.[242,243,251]

4.1 Auxiliary Density Perturbation Theory

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 elements is substituted by the response of the Coulomb and exchange-correlation fitting coefficients. For the detailed derivation of the static and frequency-dependent ADPT equations we refer the interested reader to references [242], [243], [251] and [252]. Here, we will take only the results for the linear response and perturbation-independent basis set as an example.

According to McWeeny's SCP method, an element of the dynamic first-order perturbed

density matrix is given by

$$P^{(\lambda)}_{\mu\nu}(\omega) = 2\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \frac{\mathcal{K}^{(\lambda)}_{ia}(\omega)}{\varepsilon_i - \varepsilon_a - \omega} c_{\mu i} c_{\nu a} + 2\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \frac{\mathcal{K}^{(\lambda)}_{ia}(\omega)}{\varepsilon_i - \varepsilon_a + \omega} c_{\mu a} c_{\nu i} \tag{4.2}$$

where $\boldsymbol{\mathcal{K}}^{(\lambda)}(\omega)$ is the perturbed Kohn-Sham matrix in MO representation. An element of $\boldsymbol{\mathcal{K}}^{(\lambda)}(\omega)$ is obtained according to

$$\mathcal{K}^{(\lambda)}_{ia}(\omega) = \sum_{\mu,\nu} c_{\mu i} K^{(\lambda)}_{\mu\nu}(\omega) c_{\nu a} \tag{4.3}$$

The perturbed ADFT Kohn-Sham matrix elements are given by:

$$K^{(\lambda)}_{\mu\nu}(\omega) = H^{(\lambda)}_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle \left[x^{(\lambda)}_{\bar{k}}(\omega) + z^{(\lambda)}_{\bar{k}}(\omega) \right] \tag{4.4}$$

The here appearing perturbed core-Hamiltonian, $\mathbf{H}^{(\lambda)}$, depends on the particular perturbation being studied. Independent of the form of $\mathbf{H}^{(\lambda)}$, the perturbed fitting coefficients $\mathbf{x}^{(\lambda)}$ and $\mathbf{z}^{(\lambda)}$ must be computed in order to obtain $\mathbf{P}^{(\lambda)}$. To this end, ADPT takes advantage of the perturbed Coulomb fitting equations which, for perturbation-independent auxiliary function sets, take the form

$$\sum_{\bar{l}} G_{\bar{k}\bar{l}} x^{(\lambda)}_{\bar{l}}(\omega) = \sum_{\mu,\nu} \langle \bar{k} \| \mu\nu \rangle P^{(\lambda)}_{\mu\nu}(\omega) \tag{4.5}$$

Equations (4.2) and (4.5) can be combined into a single one. Multiplying Equation (4.2) by $\langle \mu\nu \| \bar{k} \rangle$ and summing over all AO pairs yields

$$\begin{aligned} \sum_{\mu,\nu} P^{(\lambda)}_{\mu\nu}(\omega) \langle \mu\nu \| \bar{k} \rangle \;\; = \;\; & 2\sum_{\mu,\nu}\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \frac{\mathcal{K}^{(\lambda)}_{ia}(\omega)}{\omega_{ia} - \omega} c_{\mu i} c_{\nu a} \langle \mu\nu \| \bar{k} \rangle + \\ & 2\sum_{\mu,\nu}\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \frac{\mathcal{K}^{(\lambda)}_{ia}(\omega)}{\omega_{ia} + \omega} c_{\nu i} c_{\mu a} \langle \mu\nu \| \bar{k} \rangle \end{aligned} \tag{4.6}$$

where we have introduced $\omega_{ia} \equiv \varepsilon_i - \varepsilon_a$. The rhs of Equation (4.6) can be further simplified

by performing the sum over the AOs and using the permutational symmetry of the ERIs:

$$\sum_{\mu,\nu} P^{(\lambda)}_{\mu\nu}(\omega)\langle\mu\nu\|\bar{k}\rangle = 4\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \mathcal{K}^{(\lambda)}_{ia}(\omega)\frac{\omega_{ia}}{\omega_{ia}^2-\omega^2}\langle ia\|\bar{k}\rangle \tag{4.7}$$

Substituting the rhs of Equation (4.7) into the rhs of Equation (4.5) yields

$$\sum_{\bar{l}} G_{\bar{k}\bar{l}} x^{(\lambda)}_{\bar{l}}(\omega) = 4\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \mathcal{K}^{(\lambda)}_{ia}(\omega)\frac{\omega_{ia}}{\omega_{ia}^2-\omega^2}\langle ia\|\bar{k}\rangle \tag{4.8}$$

To proceed further, the perturbed Kohn-Sham matrix $\boldsymbol{\mathcal{K}}(\omega)$ appearing in Equation (4.8) is expanded:

$$\begin{aligned}\sum_{\bar{l}} G_{\bar{k}\bar{l}} x^{(\lambda)}_{\bar{l}}(\omega) &= 4\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \mathcal{H}^{(\lambda)}_{ia}\frac{\omega_{ia}}{\omega_{ia}^2-\omega^2}\langle ia\|\bar{k}\rangle + \\ &\quad 4\sum_{\bar{l}}\sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \left[x^{(\lambda)}_{\bar{l}}(\omega) + z^{(\lambda)}_{\bar{l}}(\omega)\right]\langle\bar{l}\|ia\rangle\frac{\omega_{ia}}{\omega_{ia}^2-\omega^2}\langle ia\|\bar{k}\rangle\end{aligned} \tag{4.9}$$

Collecting all terms that depend on the perturbed fitting coefficients transforms Equation (4.9) into:

$$\sum_{\bar{l}} G_{\bar{k}\bar{l}} x^{(\lambda)}_{\bar{l}}(\omega) - 4\sum_{\bar{l}} A_{\bar{k}\bar{l}}(\omega)\left[x^{(\lambda)}_{\bar{l}}(\omega) + z^{(\lambda)}_{\bar{l}}(\omega)\right] = 4b^{(\lambda)}_{\bar{k}}(\omega) \tag{4.10}$$

where

$$A_{\bar{k}\bar{l}}(\omega) = \sum_{i}^{\text{occ}}\sum_{a}^{\text{occ}}\langle\bar{k}\|ia\rangle\frac{\omega_{ia}}{\omega_{ia}^2-\omega^2}\langle ia\|\bar{l}\rangle \tag{4.11}$$

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

$$b^{(\lambda)}_{\bar{k}}(\omega) = \sum_{i}^{\text{occ}}\sum_{a}^{\text{uno}} \mathcal{H}^{(\lambda)}_{ia}\frac{\omega_{ia}}{\omega_{ia}^2-\omega^2}\langle ia\|\bar{k}\rangle \tag{4.12}$$

is an element of the perturbation vector $\mathbf{b}^{(\lambda)}(\omega)$. Finally, the perturbed exchange-correlation coefficients are obtained from Equation (2.45) as:

$$\mathbf{z}^{(\lambda)}(\omega) = \mathbf{G}^{-1}\mathbf{L}^{(\lambda)}(\omega) \tag{4.13}$$

with

$$
\begin{aligned}
L_{\bar{k}}^{(\lambda)}(\omega) &= \iint \frac{\delta v_{\mathrm{xc}}[\tilde{\rho}](\mathbf{r}_1)}{\delta\tilde{\rho}(\mathbf{r}_2)}\,\bar{k}(\mathbf{r}_1)\,\tilde{\rho}^{(\lambda)}(\mathbf{r}_2,\omega)\,\mathrm{d}\mathbf{r}_1\;\mathrm{d}\mathbf{r}_2 \\
&= \sum_{\bar{l}} x_{\bar{l}}^{(\lambda)}(\omega) \iint f_{\mathrm{xc}}[\tilde{\rho}](\mathbf{r}_1,\mathbf{r}_2)\,\bar{k}(\mathbf{r}_1)\,\bar{l}(\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_1\;\mathrm{d}\mathbf{r}_2 \\
&= \sum_{\bar{l}} x_{\bar{l}}^{(\lambda)}(\omega)\langle\,\bar{l}\,|\,f_{\mathrm{xc}}[\tilde{\rho}]\,|\,\bar{k}\,\rangle
\end{aligned}
\tag{4.14}
$$

In Equation (4.14), we used the following definition for the adiabatic exchange-correlation kernel:

$$
f_{\mathrm{xc}}[\tilde{\rho}](\mathbf{r}_1,\mathbf{r}_2) \equiv \frac{\delta^2 E_{\mathrm{xc}}[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r}_1)\,\delta\tilde{\rho}(\mathbf{r}_2)} \tag{4.15}
$$

The adiabatic kernel is, formally, nonlocal in space but local in time, i.e. frequency independent. For local and semilocal functionals the kernel further simplifies to:[43,122]

$$
f_{\mathrm{xc}}[\tilde{\rho}](\mathbf{r}_1,\mathbf{r}_2) = f_{\mathrm{xc}}[\tilde{\rho}](\mathbf{r}_1,\mathbf{r}_2)\delta(\mathbf{r}_1-\mathbf{r}_2) \tag{4.16}
$$

For the case of LDA DFAs, this approximation is known as the Adiabatic Local Density Approximation (ALDA).[43,122] Substitution of Equations (4.13) and (4.14) into Equation (4.10) yields the following linear equation system for $\mathbf{x}^{(\lambda)}(\omega)$:

$$
\begin{aligned}
\left[\frac{1}{4}\mathbf{G} - \mathbf{A}(\omega)\left(\mathbf{E}+\mathbf{G}^{-1}\mathbf{F}\right)\right]\mathbf{x}^{(\lambda)}(\omega) &= \mathbf{b}^{(\lambda)}(\omega) \\
\mathbf{R}(\omega)\mathbf{x}^{(\lambda)}(\omega) &= \mathbf{b}^{(\lambda)}(\omega)
\end{aligned}
\tag{4.17}
$$

where an element of the kernel matrix, $\mathbf{F}$ is given by:

$$
F_{\bar{k}\bar{l}} = \langle\,\bar{k}\,|\,f_{\mathrm{xc}}[\tilde{\rho}]\,|\,\bar{l}\,\rangle \tag{4.18}
$$

Different to CPKS, Equation (4.17) can be solved by the explicit inversion of the ADPT response matrix, $\mathbf{R}(\omega)$, because the dimension of the problem is reduced from $(N_{\mathrm{occ}} \times N_{\mathrm{uno}}) \times (N_{\mathrm{occ}} \times N_{\mathrm{uno}})$ in CPKS to $M \times M$ in ADPT. This procedure, however, is prone to numerical instabilities that can be encountered during the matrix inversion. Therefore,

this step is performed based on the singular value decomposition (SVD) of the response matrix. It is important to note that even when the matrices $\mathbf{G}$, $\mathbf{A}(\omega)$ and $\mathbf{F}$ are symmetric matrices, the full response matrix is generally not symmetric. Thus, the SVD of $\mathbf{R}(\omega)$ is carried out by the eigenvalue decomposition (EVD) of $\mathbf{R}(\omega)\left[\mathbf{R}(\omega)\right]^T$ or its transpose. The disadvantage of this step is that if $\mathbf{R}(\omega)$ is ill-conditioned, then $\mathbf{R}(\omega)\left[\mathbf{R}(\omega)\right]^T$ will be even worse conditioned. In order to avoid such problem, one can try to obtain an explicitly symmetric form of Equation (4.17) and perform the SVD over this symmetric matrix. It is straightforward to show that introducing the identity

$$\mathbf{E} = \mathbf{G}^{-1}\mathbf{A}(\omega)\mathbf{A}(\omega)^{-1}\mathbf{G} \tag{4.19}$$

into Equation (4.17) yields an equivalent symmetric problem with the form

$$\begin{aligned} \mathbf{R}(\omega)\mathbf{G}^{-1}\mathbf{A}(\omega)\mathbf{y}^{(\lambda)}(\omega) &= \mathbf{b}^{(\lambda)}(\omega) \\ \left[\frac{1}{4}\mathbf{A}(\omega) - \mathbf{A}(\omega)\mathbf{G}^{-1}\mathbf{A}(\omega) - \mathbf{A}(\omega)\mathbf{G}^{-1}\mathbf{F}\mathbf{G}^{-1}\mathbf{A}(\omega)\right]\mathbf{y}^{(\lambda)}(\omega) &= \mathbf{b}^{(\lambda)}(\omega) \end{aligned} \tag{4.20}$$

where

$$\mathbf{y}^{(\lambda)}(\omega) = \mathbf{A}(\omega)^{-1}\mathbf{G}\mathbf{x}^{(\lambda)}(\omega) \tag{4.21}$$

For ease of notation, we will denote this symmetric matrix as $\mathbf{R}'(\omega)$. Equation (4.20) still can have some numerical issues despite being a real and symmetric linear equation system. Let us illustrate most of the problems encountered while solving the ADPT equation system. Figure 4.1 shows a color map of the difference, in a.u., between $\mathbf{R}(\omega)$ and $\left[\mathbf{R}(\omega)\right]^T$ for one water molecule described at the VWN/DZVP/GEN-A2 level of theory. If $\mathbf{R}(\omega)$ would be symmetric, then the color map shown in Figure 4.1 would be a blank square. However, it can be seen that $\mathbf{R}(\omega)$ is not symmetric even for this very simple system. In order to obtain a direct solution of the ADPT response equations, either Equation (4.17) or Equation (4.20) must be solved. The numerical stability of any of these methods will be governed by the condition number, κ, of the corresponding matrix. If the spectral norm

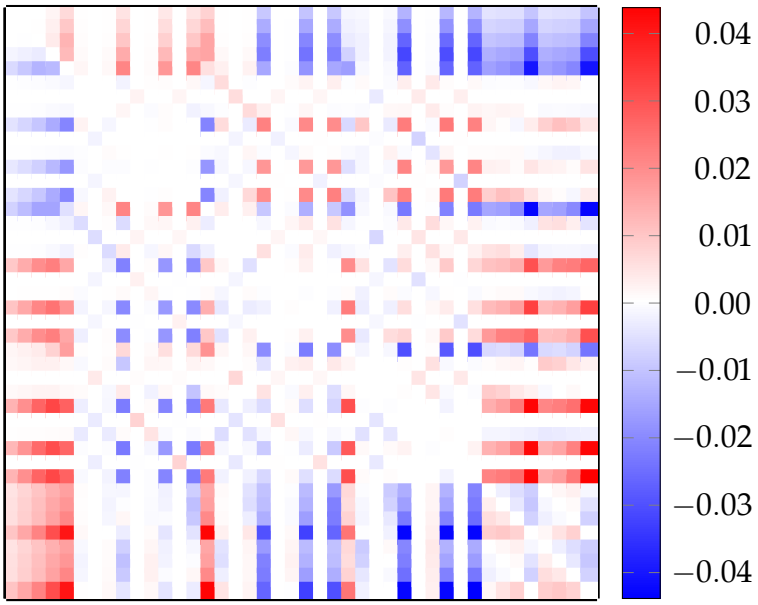


Figure 4.1: Color map showing the difference [a.u.] between the ADPT response matrix and its transpose for one water molecule at the VWN/DZVP/GEN-A2 level of theory.

is chosen, then κ can be computed as

$$\kappa(\mathbf{A}) = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})} \tag{4.22}$$

for any arbitrary matrix $\mathbf{A}$. Here, $\sigma_{\max}(\mathbf{A})$ and $\sigma_{\min}(\mathbf{A})$ denote the largest and smallest singular values of matrix $\mathbf{A}$, respectively. By definition, singular values are positive semidefinite. All nonzero singular values correspond to the positive square roots of the nonzero eigenvalues of $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$.[253] The utility of defining the condition number κ in terms of singular values is that *every* real matrix has a real decomposition of the form $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\mathbf{\Sigma}$ is a rectangular diagonal matrix that collects all singular values.[254]

In order to gain insight into the numerical stability of solving the ADPT equation system, it is instructive to plot all singular values of the matrices $\mathbf{R}(\omega)$, $\mathbf{R}'(\omega)$ and $\mathbf{R}(\omega)\left[\mathbf{R}(\omega)\right]^T$. Figure 4.2 shows a plot of the singular values of all these matrices on a logarithmic scale. From this figure follows that the symmetric approach has the worst condition number, 10^9. Then, $\mathbf{R}(\omega)\left[\mathbf{R}(\omega)\right]^T$ follows relatively close with a condition number of around 10^8. Finally, the original matrix has a condition number of around 10^4. Thus, it is preferable to

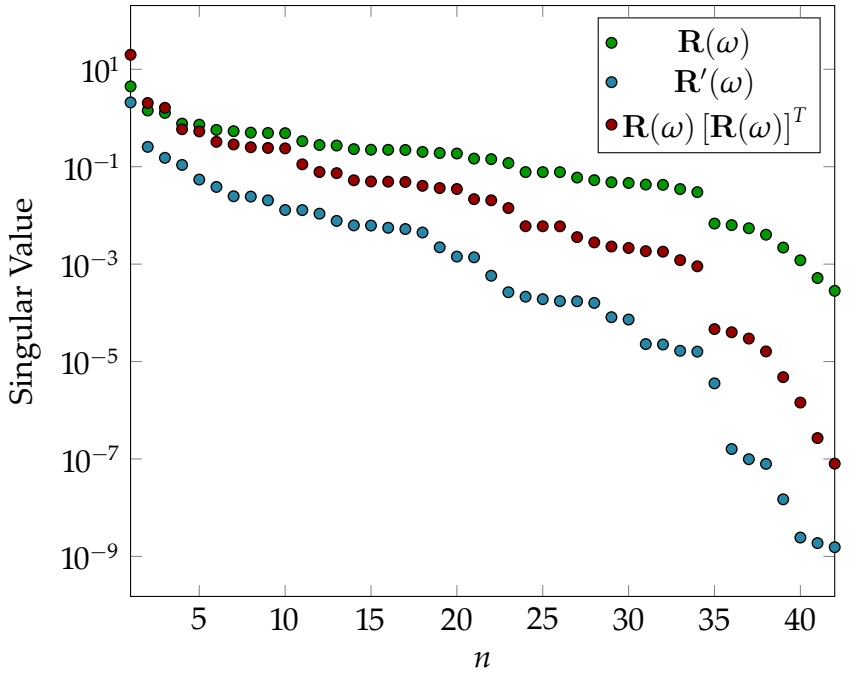


Figure 4.2: Singular values of the ADPT response matrix and the corresponding symmetrized response matrices for one water molecule employing the VWN/DZVP/GEN-A2 level of theory.

avoid any of the modifications to the original matrix in order to ensure the best numerical stability attainable for the solution of the ADPT equation system. Furthermore, the condition number worsens with increasing system size. When using the VWN/DZVP/GEN-A2 level of theory, the condition number of the ADPT response matrix $\mathbf{R}(\omega)$ for the C_4H_{10} linear alkane chain is around 3×10^6, the corresponding one for the $C_{10}H_{22}$ chain is around 1×10^7 and the one for $C_{20}H_{42}$ is 3×10^7. We note that the default method to solve the ADPT response equation system in deMon2k prior to this work was to obtain $\mathbf{R}(\omega)^{-1}$ as:

$$\mathbf{R}(\omega)^{-1} = [\mathbf{R}(\omega)]^T \left(\mathbf{R}(\omega)\,[\mathbf{R}(\omega)]^T\right)^{-1} \tag{4.23}$$

The numerical stability was controlled by the SVD of $\mathbf{R}(\omega)\,[\mathbf{R}(\omega)]^T$.

Besides these numerical instabilities, solving either Equation (4.17) or (4.20) can become computationally demanding. Most evident is the M^3 scaling for the matrix diagonalization needed for the SVD. However, the real bottleneck of ADPT arises in the actual computation of $\mathbf{R}(\omega)$. In order to obtain the ADPT response matrix, four different matrices must be computed: $\mathbf{G}$, $\mathbf{G}^{-1}$, $\mathbf{A}(\omega)$ and $\mathbf{F}$. When the ADPT module is called, the matrix $\mathbf{G}$

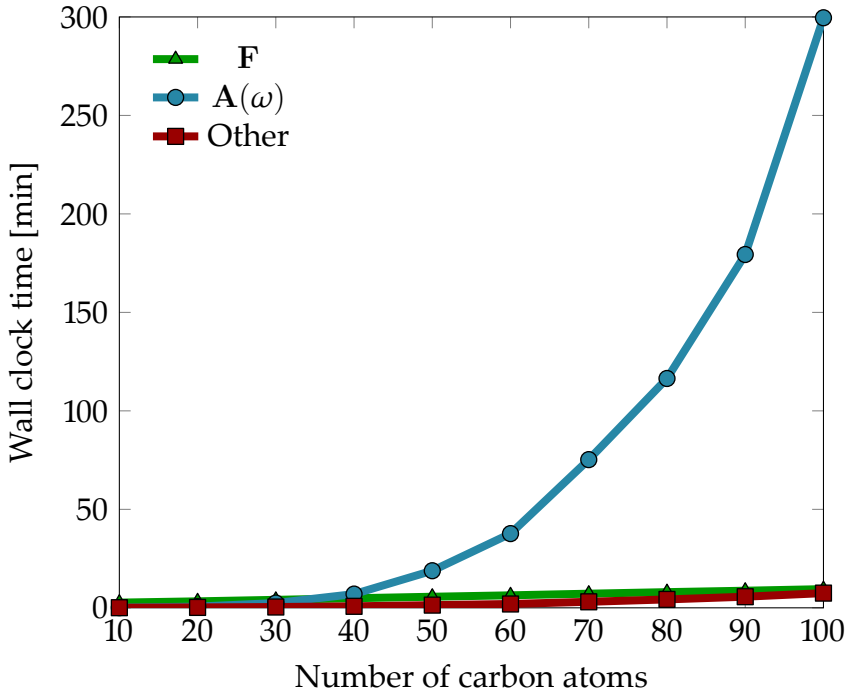


Figure 4.3: Wall clock times of the principal ADPT modules for the polarizability calculations of linear alkane chains. The VWN/DZVP/GEN-A2 level of theory was used. See text for more details. A small offset was introduced to **F** for visibility.

and its inverse are already computed and stored. The computation of the kernel matrix, **F**, involves a numerical integration which scales as $M_{\text{local}}^2 \times G$. Finally, computation of $\mathbf{A}(\omega)$ involves three quartic-scaling steps: the first one is an $M \times N^2 \times N_{\text{occ}}$ scaling step that transforms the three-center ERIs from $\langle \bar{k} \| \mu\nu \rangle$ to $\langle \bar{k} \| i\nu \rangle$, the second one is an $M \times N \times N_{\text{occ}} \times N_{\text{uno}}$ scaling step that transforms $\langle \bar{k} \| i\nu \rangle$ into $\langle \bar{k} \| ia \rangle$, and the third involves the contraction of the ERIs over the MO indices. Thus, the ADPT scaling is dominated by the computation of the Coulomb response matrix which represents its computational bottleneck. To illustrate this, static polarizability calculations for linear alkane chains with up to 100 carbon atoms were performed. In this case, the VWN/DZVP/GEN-A2 level of theory was used. The calculations were performed on a single Intel® Xeon® X5675 @ 3.07 GHz processor with 4 GB of allocatable memory. Figure 4.3 shows the wall clock time needed to compute **F**, $\mathbf{A}(\omega)$ and all other steps involved in the calculation of the polarizability tensor $\boldsymbol{\alpha}$. It can be seen that the computation of the kernel matrix possesses almost linear scaling behavior. This linear scaling is due to the fact that M_{local} remains constant and, thus, this part of the code scales only with the number of grid points G. In contrast, the $\mathbf{A}(\omega)$ calculation

possesses a quartic scaling and, thus, quickly becomes the bottleneck for the polarizability tensor calculation. It is evident that if ADPT is to be used to describe response properties of nanosystems a new way for solving the response equation system is needed.

The explicit computation of all matrices involved in the response equation system can be avoided if an iterative procedure is used to obtain $\mathbf{x}^{(\lambda)}(\omega)$. In this way, only the action of $\mathbf{R}(\omega)$ on a trial vector is required. Along this line, an iterative solver for the ADPT response equation system is presented in the following section. The newly developed iterative solver allows the reduction of the computational scaling of the algorithm and works directly on $\mathbf{R}(\omega)$, avoiding the ill-conditioned matrices $\mathbf{R}'(\omega)$ and $\mathbf{R}(\omega)\left[\mathbf{R}(\omega)\right]^T$.[255]

4.2 Iterative Solution to the ADPT Equations

It is widely known that the solution of very large linear equation systems can be achieved in the most efficient and stable form by iterative methods. This is, in part, because the coefficient matrix is used only to compute its action on a given vector. In this form, the elements of the matrix do not need to be available at all times, saving a huge amount of memory. It is therefore convenient to derive an iterative algorithm that avoids the explicit calculation of the Coulomb matrix $\mathbf{A}(\omega)$, the current bottleneck of ADPT. To this end, we now analyze how to compute the action of the response matrix on a trial vector $\mathbf{p}$ in the most efficient manner. It is evident that such an action can be divided into three: the action of $\mathbf{G}$, the action of $\mathbf{G}^{-1}\mathbf{F}$ and the action of $\mathbf{A}(\omega)$.

The action of $\mathbf{G}$ on a trial vector is extensively used in the Coulomb solver proposed by Domínguez-Soria *et al.*[108] The complexity of this step has a formal M^2 scaling, but it can be further lowered by means of the double-asymptotic ERI expansion.[109]

The action of $\mathbf{G}^{-1}\mathbf{F}$ on the trial vector $\mathbf{p}$ can be subdivided into two steps. The result of the first one is the kernel vector $\mathbf{f}$ with elements:

$$\begin{aligned} f_{\bar{k}} &= \sum_{\bar{l}} \langle \bar{k} \,|\, f_{\mathrm{xc}} \,|\, \bar{l} \rangle p_{\bar{l}} \\ &= \langle \bar{k} \,|\, f_{\mathrm{xc}} \,|\, p \rangle \end{aligned} \tag{4.24}$$

where $p(\mathbf{r})$ is obtained as

$$p(\mathbf{r}) = \sum_{\bar{k}} p_{\bar{k}} \bar{k}(\mathbf{r}) \tag{4.25}$$

The value of $p(\mathbf{r})$ at every grid point is obtained in an analog way as for the auxiliary density, thus, it possesses an $M_{\text{local}} \times G$ scaling. The evaluation of the vector $\mathbf{f}$ can also be performed with an $M_{\text{local}} \times G$ scaling, since $\mathbf{f}$ is obtained as:

$$f_{\bar{k}} = \sum_{g}^{G} w_g \cdot \bar{k}(\mathbf{r}_g) \cdot f_{\text{xc}}[\tilde{\rho}(\mathbf{r}_g)] \cdot p(\mathbf{r}_g) \tag{4.26}$$

Thus, only M_{local} contributions to $\mathbf{f}$ are computed at every grid point. In the second step, $\mathbf{f}$ is transformed with the inverse Coulomb matrix. In order to perform this transformation, $\mathbf{G}^{-1}$ is read from disk. An alternative option is to solve, by means of the preconditioned conjugate gradient solver used to obtain $\mathbf{z}$,[108] the linear equation system:

$$\mathbf{G}\mathbf{f}' = \mathbf{f} \tag{4.27}$$

This avoids the need for an explicit $\mathbf{G}^{-1}$ matrix for ADPT (see also section 6.2.1). Finally, the action of $\mathbf{A}(\omega)$ on $\mathbf{p}$ yields a vector $\mathbf{a}(\omega)$ with elements

$$a_{\bar{k}}(\omega) = \sum_{\bar{l}} A_{\bar{k}\bar{l}}(\omega)\, p_{\bar{l}} \tag{4.28}$$

Substituting Equation (4.11) into Equation (4.28) yields

$$a_{\bar{k}}(\omega) = \sum_{\bar{l}} \sum_{i}^{\text{occ}} \sum_{a}^{\text{uno}} \langle \bar{k} \| ia \rangle \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \langle ia \| \bar{l} \rangle p_{\bar{l}} \tag{4.29}$$

Introducing the LCGTO expansion, Equation (4.29) transforms into

$$a_{\bar{k}}(\omega) = \sum_{\sigma,\tau} \langle \bar{k} \| \sigma\tau \rangle \sum_{i}^{\text{occ}} \sum_{a}^{\text{uno}} c_{\sigma i} c_{\tau a} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \sum_{\mu,\nu} c_{\mu i} c_{\nu a} \sum_{\bar{l}} \langle \mu\nu \| \bar{l} \rangle p_{\bar{l}} \tag{4.30}$$

where we have ordered the summations according to the discussion of the following indi-

vidual steps.

The last sum on the rhs of Equation (4.30) is arithmetically equivalent to the one used to obtain the Kohn-Sham matrix in the ADFT direct SCF approach.[99,256] Thus, the same subroutines, with the appropriate input fields, can be used. The formal scaling of this step is $N^2 \times M$, but it can be further reduced by integral screening and the double asymptotic ERI expansion. Performing the sum for all AO pairs defines a new matrix $\mathbf{Q}$ with elements

$$Q_{\mu\nu} = \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle p_{\bar{k}} \tag{4.31}$$

Substituting Equation (4.31) into Equation (4.30) yields

$$a_{\bar{k}}(\omega) = \sum_{\sigma,\tau} \langle \bar{k} \| \sigma\tau \rangle \sum_{i}^{\text{occ}} \sum_{a}^{\text{uno}} c_{\sigma i} c_{\tau a} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \sum_{\mu,\nu} c_{\mu i} c_{\nu a} Q_{\mu\nu} \tag{4.32}$$

The next step is the transformation of the matrix $\mathbf{Q}$ into its MO representation, namely,

$$\boldsymbol{\mathcal{Q}} = \mathbf{c}_{\text{occ}}^T \, \mathbf{Q} \, \mathbf{c}_{\text{uno}} \tag{4.33}$$

which can be decomposed as two matrix-matrix multiplications, one with $N^2 \times N_{\text{occ}}$ scaling and the other with $N \times N_{\text{occ}} \times N_{\text{uno}}$ scaling. These two cubic steps are performed with optimized BLAS subroutines. This ensures near peak performance and, therefore, the cubic scaling can be hidden for most system sizes. Afterwards, $\boldsymbol{\mathcal{Q}}$ is scaled with the orbital energies to yield a new matrix whose elements are given by

$$\mathcal{Q}'_{ia}(\omega) = \mathcal{Q}_{ia} \frac{\omega_{ia}}{\omega_{ia}^2 - \omega^2} \tag{4.34}$$

This transformation has an $N_{\text{occ}} \times N_{\text{uno}}$ scaling. With these steps Equation (4.32) is transformed into:

$$a_{\bar{k}}(\omega) = \sum_{\sigma,\tau} \langle \bar{k} \| \sigma\tau \rangle \sum_{i}^{\text{occ}} \sum_{a}^{\text{uno}} c_{\sigma i} c_{\tau a} \mathcal{Q}'_{ia}(\omega) \tag{4.35}$$

The next step is the back-transformation of $\boldsymbol{\mathcal{Q}}'(\omega)$ into AO representation according to

$$\mathbf{T}(\omega) = \mathbf{c}_{\text{occ}}\boldsymbol{\mathcal{Q}}'(\omega)\mathbf{c}_{\text{uno}}^{T} \tag{4.36}$$

which can be performed in an equivalent way as discussed for Equation (4.33). Finally, an element of the resulting $\mathbf{a}(\omega)$ vector is computed as:

$$a_{\bar{k}}(\omega) = \sum_{\sigma,\tau}\langle\bar{k}\|\sigma\tau\rangle T_{\sigma\tau}(\omega) \tag{4.37}$$

This last step is almost arithmetic equivalent to the computation of the Coulomb vector $\mathbf{J}$, Equation (2.31). However, note that $\mathbf{T}(\omega)$ is not a symmetric matrix. Thus, care must be taken when re-using the subroutines for computing $\mathbf{J}$. Besides this small technical detail, this last step preserves the $M \times N^2$ formal scaling and can also take advantage of integral screening and the double asymptotic ERI expansion.

In summary, we have shown that the action of $\mathbf{R}(\omega)$ can be decomposed into several steps which do not exceed a formal cubic scaling. The calculation of the kernel vector $\mathbf{f}$ exhibits a linear scaling and three other steps show, at worst, quadratic scalings, namely the computation of $\mathbf{Gp}$, $\mathbf{G}^{-1}\mathbf{f}$, and $\boldsymbol{\mathcal{Q}}'(\omega)$. Two other steps, the formation of $\mathbf{Q}$ and $\mathbf{a}(\omega)$ can be computed with subquadratic scaling by using integral screening and the double asymptotic ERI expansion. Finally, the two remaining steps involve four cubic scaling matrix-matrix multiplications that benefit from a very low prefactor by using optimized BLAS subroutines and, consequently, remain "hidden" for a wide range of system sizes before becoming the computational bottleneck.

In addition to the reduction of the computational complexity, the iterative solver also achieves a reduction in the memory demand for the perturbation calculation. Explicit construction of the Coulomb response matrix $\mathbf{A}(\omega)$ requires sufficient memory to store $M \times N_{\text{occ}} \times N_{\text{uno}}$ real numbers, many times forcing its computation to be performed in batches. With the new solver, two $N \times N$ matrices and few M-dimensional vectors are needed. In fact, the small amount of memory required by the iterative ADPT solver allows

the storage of near-field ERIs in memory. As a consequence, the mixed SCF paradigm, where near-field ERIs are stored in memory and far-field ERIs are recalculated as needed, becomes also applicable for the ADPT solver.[103]

In the previous paragraphs, we have shown that the iterative solution of the ADPT response equation not only reduces the formal scaling of the algorithm, but also allows the usage of highly efficient algorithms to compute the computationally most-demanding tasks, i.e., optimized BLAS subroutines for matrix-matrix multiplications and an analog of the mixed SCF scheme for the ERI calculation. However, given that the ADPT response matrix $\mathbf{R}(\omega)$ is not symmetric, special solvers must be used. Traditional algorithms developed for these type of equation systems include BiCG, BiCGSTAB,[257] CGS,[258] QMR[259] and GMRES.[260] Some of them lack the desirable global residual minimization properties or require the storage of a large number of basis vectors. Moreover, BiCG, BiCGSTAB and CGS work with the "squared" linear system in order to deal with a symmetric matrix. As we have seen, working with the squared system might introduce severe numerical instabilities to the algorithm. As an alternative, a new class of efficient nonsymmetric iterative solvers have been introduced on the basis of the Eirola-Nevanlinna (EN) algorithm.[261–264] It has been shown that the EN algorithm performs equally well as the more traditional GMRES approach.[265] One advantage of the EN algorithm is that it improves an approximate inverse coefficient matrix each iteration. This property is particularly important when dealing with several perturbations because the improved $\mathbf{R}(\omega)^{-1}$ can be used as a preconditioner for subsequent response equation systems. In the following subsection we will give a brief description of the EN algorithm and its implementation in the ADPT response branch of deMon2k.

4.2.1 Eirola-Nevanlinna algorithm

In 1989, Eirola and Nevanlinna proposed an iterative algorithm to solve nonsymmetric linear equation systems which is accelerated via rank-one updates of an approximate inverse coefficient matrix.[261] Figure 4.4 shows the original EN algorithm, EN1, which has been rewritten in terms of the ADPT response equation system. To simplify the notation,

procedure EN1
  Initialization: $\mathbf{x}_0^{(\lambda)}$, $\mathbf{H}_0$ arbitrary, $\mathbf{r}_0 = \mathbf{b}^{(\lambda)} - \mathbf{R}\mathbf{x}_0^{(\lambda)}$
  for $j = 0,\ n$ **do**
    if $\|\mathbf{r}_j\|_\infty \leq \tau$ **then**
      Exit
    end if
    $\mathbf{u}_j = \mathbf{H}_j(\mathbf{E} - \mathbf{R}\mathbf{H}_j)\mathbf{r}_j$
    $\mathbf{v}_j = \frac{1}{|\mathbf{R}\mathbf{u}_j|^2}(\mathbf{E} - \mathbf{H}_j^T\mathbf{R}^T)\mathbf{R}\mathbf{u}_j$
    $\mathbf{H}_{j+1} = \mathbf{H}_j + \mathbf{u}_j\mathbf{v}_j^T$
    $\mathbf{x}_{j+1}^{(\lambda)} = \mathbf{x}_j^{(\lambda)} + \mathbf{H}_{j+1}\mathbf{r}_j$
    $\mathbf{r}_{j+1} = \mathbf{b}^{(\lambda)} - \mathbf{R}\mathbf{x}_{j+1}^{(\lambda)}$
  end for
end procedure

Figure 4.4: Original Eirola-Nevanlinna algorithm EN1. The matrix $\mathbf{R}$ is the coefficient matrix, $\mathbf{H}_0$ is an approximation to the inverse coefficient matrix and $\mathbf{x}_0^{(\lambda)}$ is a guess for the solution vector.

the ω dependency of vectors and matrices will be dropped in this subsection. $\mathbf{H}_0$ is a guess to the inverse response matrix and $\mathbf{x}_0^{(\lambda)}$ is a guess for the perturbed fitting coefficients. The algorithm starts by calculating the residual $\mathbf{r}_0 = \mathbf{b}^{(\lambda)} - \mathbf{R}\mathbf{x}_0^{(\lambda)}$ and improves $\mathbf{x}_j^{(\lambda)}$ on each iteration by minimization of the residual. In exact arithmetic the EN1 algorithm yields the solution in at most M steps, for an $M \times M$ linear equation system, under the assumption that all $\mathbf{H}_j$ are nonsingular. The disadvantage of the EN1 algorithm is that the action of $\mathbf{R}$ on different vectors has to be computed four times. An alternative algorithm, developed for situations where it is more convenient to calculate a linear combination of $j+1$ vectors instead of multiplying a vector by $\mathbf{R}$, was also given in the original EN paper[261] and we will call it EN2. For stability reasons, Vuik and van der Vorst proposed a slight modification of the EN2 algorithm based on the modified Gram-Schmidt orthogonalization and a scaling invariant approach.[262,264] Figure 4.5 shows the EN2 algorithm including the modifications of Vuik and van der Vorst. Only two actions of the ADPT response matrix are needed per iteration of the EN2 algorithm, making it computationally more efficient than EN1. In EN2, however, the updates to the approximate inverse are not performed explicitly. Instead, the action of the updates are carried through the vectors $\mathbf{c}$ and $\mathbf{u}$. Other modifications of the EN2 algorithm have also been suggested,[262–264] but the here presented EN2 algorithm requires one action of $\mathbf{R}$ less than these modified algorithms. In

procedure EN2
  Initialization: $\mathbf{x}_0^{(\lambda)}$, $\mathbf{H}_0$ arbitrary, $\mathbf{r}_0 = \mathbf{b}^{(\lambda)} - \mathbf{R}\mathbf{x}_0^{(\lambda)}$
  for $j = 0,\ n$ **do**
    if $\|\mathbf{r}_j\|_\infty \leq \tau$ **then**
      Exit
    end if
    $\gamma = \frac{\mathbf{r}_j^T \mathbf{H}_0^T \mathbf{R}^T \mathbf{r}_j}{\mathbf{r}_j^T \mathbf{H}_0^T \mathbf{R}^T \mathbf{R}\mathbf{H}_0 \mathbf{r}_j}$
    $\boldsymbol{\xi} = \mathbf{r}_j - \gamma \mathbf{R}\mathbf{H}_0\mathbf{r}_j$
    $\boldsymbol{\eta} = \gamma \mathbf{H}_0 \mathbf{r}_j$
    for $m = 0,\ j-1$ **do**
      $\alpha = \mathbf{c}_m^T \boldsymbol{\xi}$
      $\boldsymbol{\xi} = \boldsymbol{\xi} - \alpha \mathbf{c}_m$
      $\boldsymbol{\eta} = \boldsymbol{\eta} + \alpha \mathbf{u}_m$
    end for
    $\mathbf{c}_j = \mathbf{R}\mathbf{H}_0\boldsymbol{\xi}$
    $\mathbf{u}_j = \mathbf{H}_0\boldsymbol{\xi}$
    for $m = 0,\ j-1$ **do**
      $\beta = -\mathbf{c}_m^T \mathbf{c}_j$
      $\mathbf{c}_j = \mathbf{c}_j + \beta \mathbf{c}_m$
      $\mathbf{u}_j = \mathbf{u}_j + \beta \mathbf{u}_m$
    end for
    $\mathbf{c}_j = \frac{\mathbf{c}_j}{\|\mathbf{c}_j\|_2}$
    $\mathbf{u}_j = \frac{\mathbf{u}_j}{\|\mathbf{c}_j\|_2}$
    $\mathbf{x}_{j+1}^{(\lambda)} = \mathbf{x}_j^{(\lambda)} + \boldsymbol{\eta} + \mathbf{u}_j \mathbf{c}_j^T \boldsymbol{\xi}_j$
    $\mathbf{r}_{j+1} = \boldsymbol{\xi}_j - \mathbf{c}_j \mathbf{c}_j^T \boldsymbol{\xi}_j$
  end for
end procedure

Figure 4.5: Modified Eirola-Nevanlinna algorithm, EN2, based on a modified Gram-Schmidt orthogonalization and designed to be scaling invariant. The matrix $\mathbf{R}$ is the coefficient matrix, $\mathbf{H}_0$ is an approximation to the inverse coefficient matrix and $\mathbf{x}_0^{(\lambda)}$ is a guess for the solution vector.

order to avoid memory overflows a restarted version of the EN2 algorithm, that drops all $\mathbf{c}$ and $\mathbf{u}$ vectors after r iterations, is the one actually implemented in deMon2k. Since all iterative ADPT calculations performed so far have converged in 9 iteration or less, r was set to 15. This choice of r allows more flexibility for hard-cases and keeps the memory demand fixed.

A very important choice to be made is that of $\mathbf{H}_0$, i.e., the guess for $\mathbf{R}^{-1}$. The only condition imposed by the algorithm is that $\mathbf{H}_0$ has to be nonsingular. The identity matrix, $\mathbf{E}$, is a valid guess, however, choosing $\mathbf{H}_0 = \mathbf{E}$ has led to non-converging cases. As a consequence, we decided to use $\mathbf{H}_0 = \mathbf{G}^{-1}$ as guess. This choice is motivated by several observations. First, $\mathbf{G}^{-1}$ is formally a nonsingular matrix. Second, the Coulomb matrix appears in the definition of $\mathbf{R}$, thus, $\mathbf{R}^{-1}$ can be obtained as a perturbation to $\mathbf{G}^{-1}$ according to the Woodbury formula.[266,267] Third, setting $\mathbf{H}_0 = \mathbf{G}^{-1}$ ensures the solution of the zeroth-order approximation to Equation (4.9), i.e. the solution to:

$$\mathbf{G}\mathbf{x}^{(\lambda)} = 4\mathbf{b}^{(\lambda)} \tag{4.38}$$

Note that Equation (4.9) can also be used to obtain an SCP solution but previous tests have shown that convergence is very problematic. Therefore, we decided to use the EN2 algorithm with $\mathbf{H}_0 = \mathbf{G}^{-1}$ because it combines the reduced dimensionality obtained in Equation (4.9) with a very robust numerical solver.

4.3 Validation and Benchmarks

4.3.1 Validation

In order to validate our new iterative procedure, we compared static and dynamic polarizabilities of small molecules obtained with the direct and the new iterative solver. The experimental geometries were used for this comparison.[268] In addition, the DFT-optimized valence triple-ζ plus polarization (TZVP)[269] basis sets augmented by field-induced polarization (FIP) functions were employed. These TZVP-FIP1 basis sets are described in detail

in references [270], [271] and [272]. The reported mean polarizabilities are calculated from the diagonal elements of the polarizability tensor as

$$\bar{\alpha}(\omega) = \frac{1}{3}\left[\alpha(\omega)_{xx} + \alpha(\omega)_{yy} + \alpha(\omega)_{zz}\right] \tag{4.39}$$

The corresponding polarizability anisotropies are calculated in the principal axes system of the polarizability tensor according to

$$|\Delta\alpha(\omega)|^2 = \frac{1}{2}\left[\left(\alpha(\omega)_{xx} - \alpha(\omega)_{yy}\right)^2 + \left(\alpha(\omega)_{xx} - \alpha(\omega)_{zz}\right)^2 + \left(\alpha(\omega)_{yy} - \alpha(\omega)_{zz}\right)^2\right] \tag{4.40}$$

Table 4.1 shows static polarizabilities and polarizability anisotropies [a.u.] obtained from the direct and iterative solution of the ADPT equations. We included experimental results as reference but will not further discuss them here because ADPT accuracy has been extensively discussed in the literature.[242,243] Instead, we focus on the comparison of the results obtained with the two different solvers for the ADPT equation system. The agreement between direct and iterative solutions is almost perfect with a maximum absolute deviation of 0.01 a.u. for $\bar{\alpha}(\omega)$ and 0.02 a.u. for $|\Delta\alpha(\omega)|$. Further tightening of the convergence criteria does not change these results. Dynamic polarizabilities, listed in Table 4.2, show exactly the same behavior. The small differences arise from the SVD performed in the direct solution of the ADPT response equation system. These results demonstrate that the iterative solver does not alter the previously reported accuracy of ADPT. Furthermore, it is expected that $\mathbf{x}^{(\lambda)}$ contains less numerical noise since the SVD step has been completely eliminated. It is also important to mention that each perturbation converged in less than five iterations, demonstrating that the chosen EN2 algorithm in combination with the $\mathbf{G}^{-1}$ start guess is well suited for solving the ADPT response equation system. To further demonstrate the suitability of the EN2 algorithm, we computed the polarizability dispersion for the Li atom. These calculations were performed with the objective to test the performance of the iterative algorithm near excitation poles, where the polarizability diverges towards $\pm\infty$. Figure 4.6 shows the polarizability dispersion of the lithium atom

Table 4.1: Comparison of static LDA polarizabilities [a.u.] and polarizability anisotropies [a.u.] of small molecules obtained with the direct and iterative solver for the ADPT equation system. For comparison, the available experimental data are listed, too. For all molecules, the experimental geometries are used.

Molecule	Direct		Iterative		Expt.	
	$\bar{\alpha}$	$\lvert\Delta\alpha\rvert$	$\bar{\alpha}$	$\lvert\Delta\alpha\rvert$	$\bar{\alpha}$	$\lvert\Delta\alpha\rvert$
HF	5.89	1.13	5.89	1.13	5.40[b]	1.35[c]
CH_4	17.38		17.38		17.27[d]	
C_2H_2	23.40	12.52	23.39	12.52	22.68[d]	11.83[e]
CH_3F	17.69	1.49[a]	17.69	1.49[a]	17.32[d]	1.41[f]
HCl	18.02[a]	1.96	18.02[a]	1.96	17.54[g]	1.47[h]
H_2S	24.72	0.57[a]	24.72	0.57[a]	24.66[i]	0.67[j]
CH_2F_2	18.46	1.92[a]	18.46	1.92[a]	18.20[d]	1.70[h]
OCS	34.10	25.25	34.10	25.25	34.33[d]	26.26[e]
SO_2	25.65	13.26	25.64	13.26	25.49[k]	12.98[k]
CHF_3	19.65	1.41[a]	19.65	1.41[a]	18.69[d]	1.46[f]
CF_4	19.93		19.93		19.53[d]	
CS_2	54.27	54.95	54.26	54.93	55.38[d]	57.38[e]

[a] Calculated dynamic values at experimental λ.
[b] Static value from refractive index dispersion (Ref. 273).
[c] Static value from molecular beam electric resonance (Ref. 274).
[d] Static value from refractive index dispersion (Ref. 275).
[e] Deduced from static estimates of Ref. 276.
[f] Dynamic values at $\lambda = 632.8$ nm (Ref. 277).
[g] Depolarized light scattering at $\lambda = 632.8$ nm (Ref. 278).
[h] Static value from molecular beam electric resonance (Ref. 275).
[i] Extrapolated static value from dispersion dynamic mean polarizability (Ref. 279).
[j] Dynamic value at $\lambda = 632.8$ nm from Kerr effect (Ref. 280).
[k] Static value from refractive index and Rayleigh scattering dispersion (Ref. 279).

Table 4.2: Dynamic LDA polarizabilities [a.u.] and polarizability anisotropies [a.u.] of small molecules obtained with the direct and iterative solver for the ADPT equation system. For comparison, the available experimental data are listed, too. For all molecules, the experimental geometries are used.

Molecule	Direct		Iterative		Expt.	
	$\bar{\alpha}$	$\lvert\Delta\alpha\rvert$	$\bar{\alpha}$	$\lvert\Delta\alpha\rvert$	$\bar{\alpha}$	$\lvert\Delta\alpha\rvert$
NH_3	15.29	2.64	15.29	2.64	14.98[a]	1.94[a]
H_2O	10.46	0.14	10.46	0.13	9.92[b]	0.66[b]
N_2	11.99	5.06	11.99	5.05	11.95[a]	4.70[a]
CO	13.70	3.34	13.70	3.35	13.34[a]	3.59[a]
NO	11.97	5.54	11.97	5.53	11.74[a]	5.70[a]
O_2	10.56	6.18	10.56	6.18	10.78[a]	7.42[a]
N_2O	19.82	19.46	19.82	19.45	20.24[a]	19.97[a]
CO_2	17.72	13.88	17.72	13.89	17.75[a]	14.17[a]
Cl_2	31.03	15.45	31.03	15.43	31.11[a]	17.54[a]
C_2H_4	29.12	12.86	29.10	12.84	28.48[a]	12.21[a]
C_2H_6	30.43	4.66	30.43	4.67	30.10[a]	5.20[a]
C_6H_6	72.68	40.57	72.67	40.56	70.18[a]	37.93[c]

[a] Depolarized light scattering at $\lambda = 632.8$ nm (Ref. 278).
[b] Depolarized ratio from Rayleigh scattering at $\lambda = 514.5$ nm (Ref. 281).
[c] Depolarized light scattering at $\lambda = 632.8$ nm (Ref. 279).

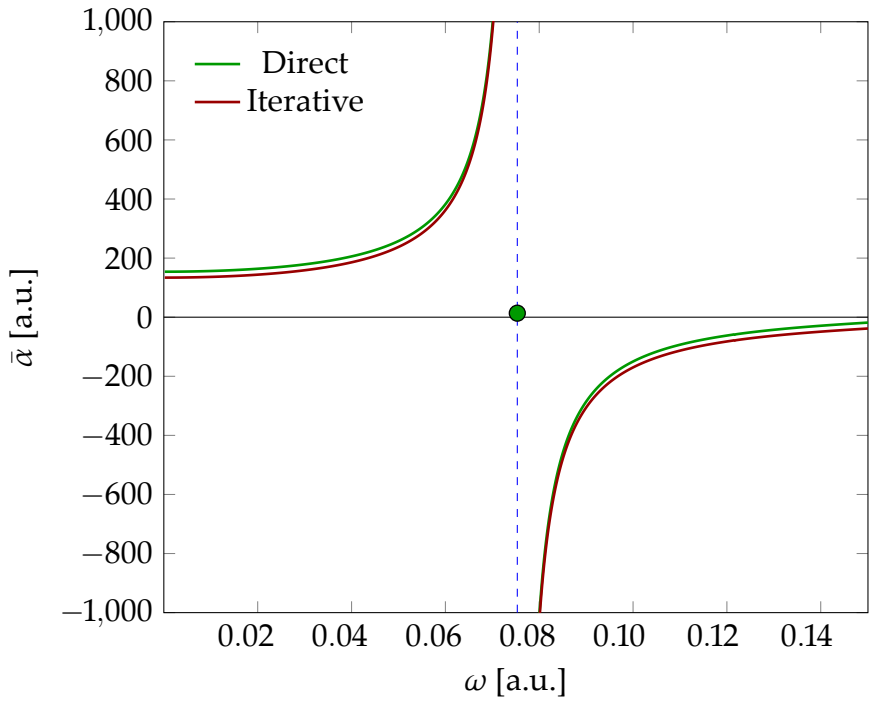


Figure 4.6: Polarizability dispersion for the Li atom at the PBE/aug-cc-pVTZ/GEN-A2* level of theory. The green point near the center of the image demonstrates the odd behavior of the direct solver when $|\omega - \omega_{ia}| < 10^{-4}$ a.u. for this system. A small offset of 10 a.u. was included to aid the visibility.

obtained with the PBE/aug-cc-pVTZ/GEN-A2* level of theory. It can be seen that both solvers, direct and iterative, overlap for almost all points. The only difference is found for the closest point to the pole (green point of Figure 4.6). The average polarizability obtained with the direct approach is around 13 a.u., while the corresponding iterative average polarizability is above $200,000$ a.u.! Thus, the direct solution shows an odd behavior very close to the singularity. This kind of odd behavior of the direct solver is also seen in the polarizability anisotropy which can be as large as 3 a.u. with the direct approach. For the Li atom this anisotropy should be of course zero by symmetry and, therefore, this large values arise from numerical instabilities in the direct approach. In contrast, the iterative polarizability anisotropy does not exceed 0.03 a.u. for any point calculated with the iterative approach.

In summary, we have shown that the here presented ADPT iterative solver provides the same results as the traditional ADPT solver for both static and dynamic polarizabilities. Furthermore, we also demonstrated that the iterative solver is numerically more stable near excitation poles than the traditional ADPT solver.

4.3.2 Benchmarks

In order to test the real computational complexity of the code, static polarizability calculations for linear alkane chains with up to 100 carbon atoms, at the VWN/DZVP/GEN-A2 level of theory, were performed. A single Intel® Xeon® X5675 @ 3.07 GHz processor with a maximum of 4 GB of allocatable memory was used for all calculations. Figure 4.7 shows the wall clock time needed to obtain the full polarizability tensor using the direct and the iterative ADPT solvers. The polarizability module in deMon2k performs the following steps:

1. If requested, rotate the molecule and the converged MOs to EFISH orientation.
2. Compute $\mathbf{G}$ and its inverse $\mathbf{G}^{-1}$.
3. Solve Coulomb fitting equation system to obtain $\mathbf{x}$.
4. Build the perturbation vectors $\mathbf{b}^{(\lambda)}(\omega)$.
5. Solve the response equation system to obtain $\mathbf{x}^{(\lambda)}(\omega)$ and calculates $\mathbf{z}^{(\lambda)}(\omega)$.
6. Build the perturbed density matrices $\mathbf{P}^{(\lambda)}(\omega)$.
7. Calculate the polarizability tensor elements as $\alpha_{\lambda\eta}(\omega) = \sum_{\mu,\nu} P^{(\lambda)}_{\mu\nu}(\omega) \langle\, \mu \,|r_\eta\, \nu \,\rangle$.

The difference between the direct and iterative solvers lies only in step 5. For the direct solver, step 5 can be further divided as:

5. Solves the response equation system to obtain $\mathbf{x}^{(\lambda)}(\omega)$ and calculates $\mathbf{z}^{(\lambda)}(\omega)$.
   (a) Build the Coulomb response matrix $\mathbf{A}(\omega)$.
   (b) Build the exchange-correlation kernel matrix $\mathbf{F}$.
   (c) Build the ADPT response matrix $\mathbf{R}(\omega) = \frac{1}{4}\mathbf{G} - \mathbf{A}(\omega)\left(\mathbf{E} - \mathbf{G}^{-1}\mathbf{F}\right)$.
   (d) Obtain $\mathbf{R}(\omega)^{-1} = \mathbf{R}(\omega)^T \left(\mathbf{R}(\omega)\mathbf{R}(\omega)^T\right)^{-1}$.
   (e) Obtain $\mathbf{x}^{(\lambda)}(\omega) = \mathbf{R}(\omega)^{-1}\mathbf{b}^{(\lambda)}(\omega)$.

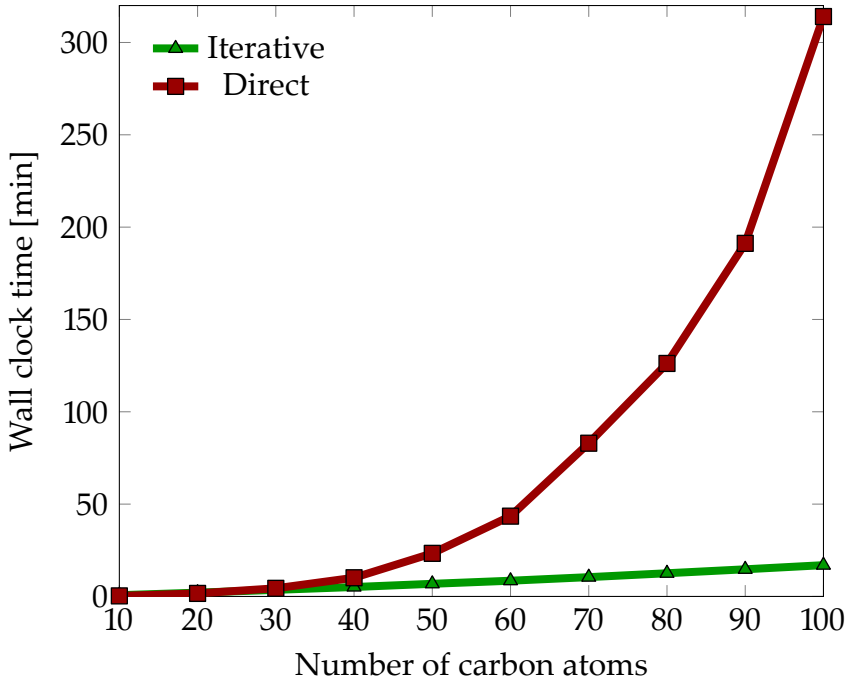


Figure 4.7: Timings [min] for serial static polarizability tensor calculations of linear alkane chains with the direct and iterative ADPT solvers. The VWN/DZVP/GEN-A2 level of theory was employed.

(f) Obtain $\mathbf{z}^{(\lambda)}(\omega) = \mathbf{G}^{-1}\mathbf{F}\mathbf{x}^{(\lambda)}(\omega)$.

In the case of the iterative solver, the EN2 algorithm is used in step 5 to directly obtain $\mathbf{x}^{(\lambda)}(\omega)$. The perturbed exchange-correlation fitting coefficients are then obtained by computing the action of $\mathbf{F}$ on $\mathbf{x}^{(\lambda)}(\omega)$ and, subsequently, the action of $\mathbf{G}^{-1}$ on the resulting vector. As can be seen from Figure 4.7, the direct and iterative solvers show similar performance until the 30 carbon alkane chain. However, for larger alkane chains the quartic scaling of the explicit building of the Coulomb response matrix $\mathbf{A}(\omega)$ (see also Figure 4.3) in the direct solver becomes dominant, rendering this approach no longer competitive with the iterative solver. The iterative solver is already 2× faster than the direct one for the $C_{40}H_{82}$ chain, and achieves a 20× speed-up for the largest alkane chain tested. This means that the polarizability of the $C_{100}H_{202}$ alkane can be obtained within 17 min with the iterative solver instead of 5.5 h with the direct one! The average polarizability obtained with both solvers differ by less than 0.1%, 1,076.69 a.u. vs. 1,077.57 a.u., for the direct and iterative solvers, respectively.

To further analyze the scaling behavior of the most time-consuming steps involved in

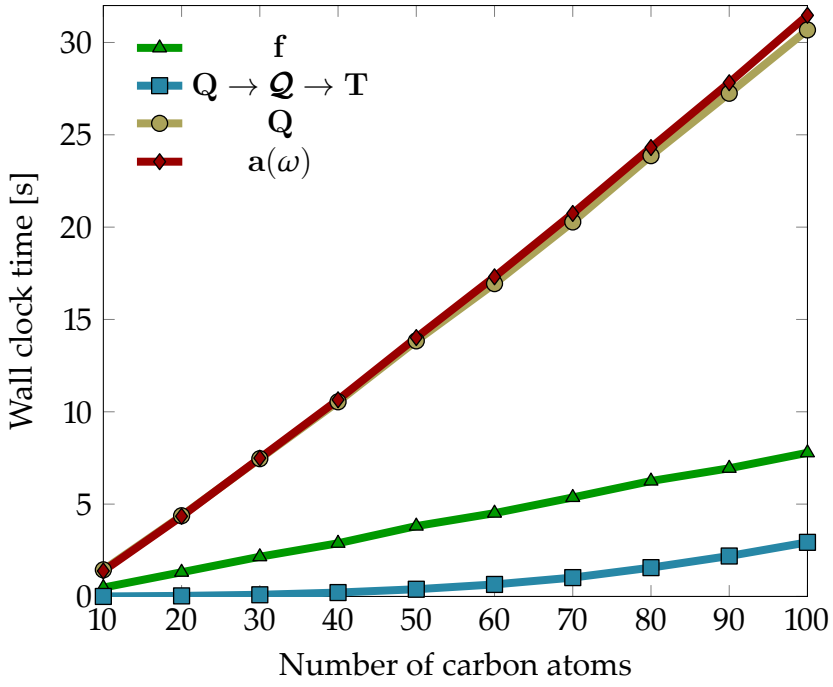


Figure 4.8: Average time [s] needed for each module in one iteration of the EN ADPT solver. The data labeled as $\mathbf{f}$ refer to Equation (4.24); $\mathbf{Q} \rightarrow \boldsymbol{\mathcal{Q}} \rightarrow \mathbf{T}$ refers to the matrix-matrix multiplication steps described in Equations (4.33)-(4.36); $\mathbf{Q}$ refers to Equation (4.31); and $\mathbf{a}(\omega)$ refers to Equation (4.37). As benchmark systems the same linear alkane chains as in Figure 4.7 are used.

the iterative solution to the ADPT response equations system, Figure 4.8 depicts timings [s] for the individual steps of one EN2 iteration. Note that all these steps are related to the calculation of the action of $\mathbf{R}(\omega)$ on some trial vectors, and that two of such actions are needed in each EN2 iteration. The most time-demanding steps involve the calculation and contraction of the three-center ERIs, namely, the computation of the matrix $\mathbf{Q}$ (Equation 4.31) and the vector $\mathbf{a}(\omega)$ (Equation 4.37). These steps represent 84% of the total time of the EN2 solver for the $C_{100}H_{202}$ chain. Note that the double asymptotic ERI expansion[109] is already used here. The other two remaining steps have very low computational demand for these systems. The action of the kernel matrix to obtain $\mathbf{f}$ (Equation 4.24) represents only 10 % of the total iterative algorithm time. Finally, the matrix-matrix multiplications appearing in Equations (4.33)-(4.36) represent only 4 % of the total time. The remaining 2 % is distributed among all other matrix-vector and vector-vector operations appearing in Figure 4.5. It is important to note that matrix-matrix multiplications are obtained with the standard cubic scaling algorithm implemented in the BLAS subroutines. As a consequence, these steps will eventually become the most computationally demanding ones.

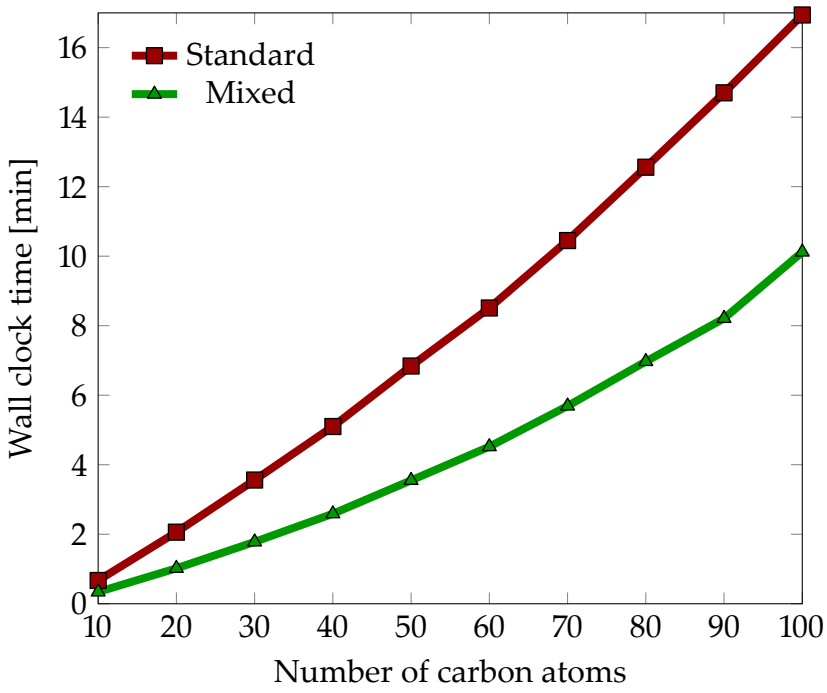


Figure 4.9: Performance comparison [min] between the standard and mixed ADPT iterative solvers. As benchmark systems the same linear alkane chains as in Figure 4.7 are used.

An alternative to overcome the cubic scaling is the introduction of the Laplace transform of the energy denominator of Equation (4.34) and the additional implementation of a sparse matrix-matrix multiplication for the resulting scaled density matrices. This alternative has yet to be explored but is kept as a perspective to develop a fully linear-scaling ADPT approach.

Another advantage of the iterative ADPT solver is its low-memory demand compared to that of the direct solver. As already mentioned, this opens up the possibility to store all near-field ERIs and recompute only the far-field ones by the double asymptotic expansion in each EN iteration. This new mixed ERIs[256] ADPT approach has a direct influence on the most time-demanding parts of the ADPT iterative solver. Usage of the mixed ADPT approach introduces an additional speed-up factor of 2 for the ADPT iterative solver. This is shown in Figure 4.9 where the performance of the standard ADPT iterative solver is compared to the one of the mixed ADPT iterative solver. The speed-up achieved by the mixed ADPT iterative solver should be more important for situations where many iterations are needed to achieve convergence or when more equations systems have to be solved, for example, in the computation of nuclear displacements or nuclear spin-spin coupling con-

stants.

The new ADPT iterative solver has also been parallelized. The parallelization also takes advantage of the low memory-demand of the iterative solver, consequently, architectures with a moderate amount of memory per processor, i.e. 2 GB, can be efficiently used. This is particularly important for the step that substitutes the calculation of the full exchange-correlation kernel matrix, $\mathbf{F}$. When the ADPT direct solver is invoked in parallel mode, all processors allocate the full M^2 $\mathbf{F}$ matrix and calculate the contribution from a distributed set of grid points. This matrix can be allocated by all processors simultaneously only when $M \leqslant 14,000$ in a 2 GB/processor architecture. In contrast, the ADPT iterative solver makes no use of M^2 matrices. The scalability of the parallel version of the ADPT iterative solver was tested by performing static polarizability calculations on the C_{720} fullerene at the VWN/DZVP/GEN-A2 level of theory. All calculations were performed on the WestGrid of Compute Canada using 5, 6, 7, 8 and 9 nodes composed of 12 Intel® Xeon® E5649 @ 2.53 GHz processors. Each one of these nodes has 2 GB of memory per core, for a total of 24 GB per node. This calculation employed more than 10,000 basis functions and almost 25,000 auxiliary functions. Therefore, a direct ADPT calculations of this system cannot be performed on this computational architecture. Figure 4.10 depicts the wall clock time [h] for solving the ADPT response equation system and to calculate the full polarizability tensor. Note that the ADPT iterative solver and all other steps needed to obtain the polarizability tensor are very well parallelized. Also note that the time needed to solve the ADPT response equation system was less than 7 hours with 5 nodes. Even more, with 9 nodes the total calculation time, including SCF and ADPT, is less than 10 hours! Another important result obtained for the C_{720}, and also for C_{960}, is that the computation is still dominated by the calculation of the ERIs. Therefore, the cubic scaling steps that remain in the iterative ADPT solver have still a minor impact for these system sizes.

In summary, the ADPT iterative solver is well suited for serial as well as parallel calculations, reduces the formal scaling to $N^2 \times N_{\text{occ}}$ and achieves almost linear scaling for systems with hundreds of atoms. As a consequence, the ADPT iterative solver allows the computation of polarizabilities of nanosystems in just hours. Extension to other molec-

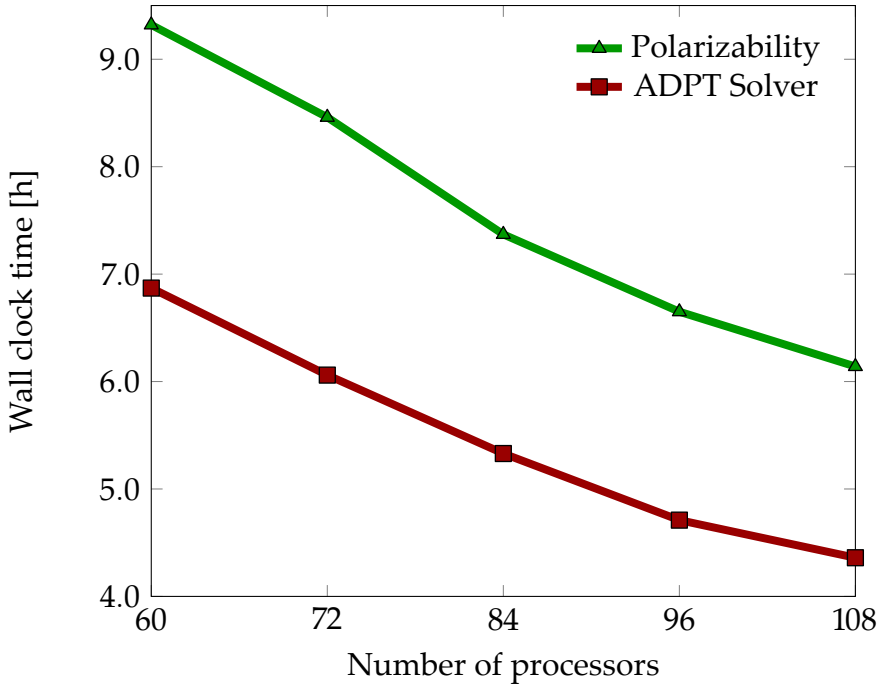


Figure 4.10: Scalability of the ADPT iterative solver tested by calculating the full polarizability tensor of the C_{720} fullerene at the VWN/DZVP/GEN-A2 level of theory. The wall clock time [h] vs. the number of cores in the parallel calculation is depicted.

ular response properties, like higher-order polarizabilities and nuclear displacements, is straightforward and its currently under investigation in our laboratory.

Solutions to problems

are easy to find:

the problem's a great

contribution.

What is truly an art

is to wring from your mind

a problem to fit

a solution.

Piet Hein

5 Applications

5.1 The Effect of Exact-Exchange in Simple Metal Clusters

Atomic clusters and small nanoparticles are recognized as distinct physical objects with their own properties. This became most clear by the experimental discovery of the electronic shell structure in alkali metal clusters.[282–284] With this discovery an emphasis was put on the quantized delocalization of valence electrons in the mean field created by the metal ions.[282–284] This behavior suggests the jellium model, which is defined by a Hamiltonian that treats the valence electrons as usual but the ionic cores as a uniform positively charged background. The model leads to a description of the electron density in terms of "Cluster Orbitals" (CO) that extend over the entire cluster. Most interesting, these "Cluster Orbitals" can also be found in canonical Kohn-Sham calculations.[285–289]

In the spherical jellium model, the ionic background density is that of a uniformly charged sphere. If the electron density is also assumed to be spherical, the COs will have good angular momentum quantum numbers l,m and their angular parts can be written in terms of spherical harmonics.[290] This gives rise to electronic shells very much as in the atomic case. The most prominent "magic numbers" observed in mass abundances, ionization potentials and electron affinities correspond to the filling of major spherical shells and are, in general, correctly reproduced for alkali metals and some noble metals.[282–284,290–295] In the monoelectronic case the quantum energies are given by[296]

$$E_{nl} = \left(2n + l + \frac{3}{2}\right)\hbar\omega \ . \tag{5.1}$$

Therefore, each level is separated from the next one by the same energy difference. As a consequence of Equation (5.1) each level has a constant $2n + l$ value and is $n_{\max}$-fold degenerate, where $n_{\max}$ is the maximum n appearing in the level. The energetic sequence is

$$(1S)\ (1P)\ (1D\ 2S)\ (1F\ 2P)\ (1G\ 2D\ 3S)$$

For many-electron systems one has to consider additionally the potential from the presence of other electrons, which leads to the energetic splitting of the degenerate levels. For higher levels, the energetic differences of shells with different angular quantum numbers become so large that the energetic sequence is affected, e.g. the $1H$ shell ($2n + l = 7$) is lower in energy than the $3S$ shell ($2n + l = 6$).[297] The modified energetic sequence is

$$1S\ 1P\ 1D\ 2S\ 1F\ 2P\ 1G\ 2D\ 1H\ 3S\ 2F\ \ldots$$

If the assumption of a uniformly positively charged background is lifted, additional splitting must occur for F and higher CO shells, due to the finite point group symmetry of the actual cluster. Most recently, this effect has been observed in the Na_{55}^{+} cluster[298] and is depicted in Figure 5.1 for the icosahedral point group symmetry. In many cases this subshell splitting is overlaid by Jahn-Teller[299] distortions. As a consequence, low-spin

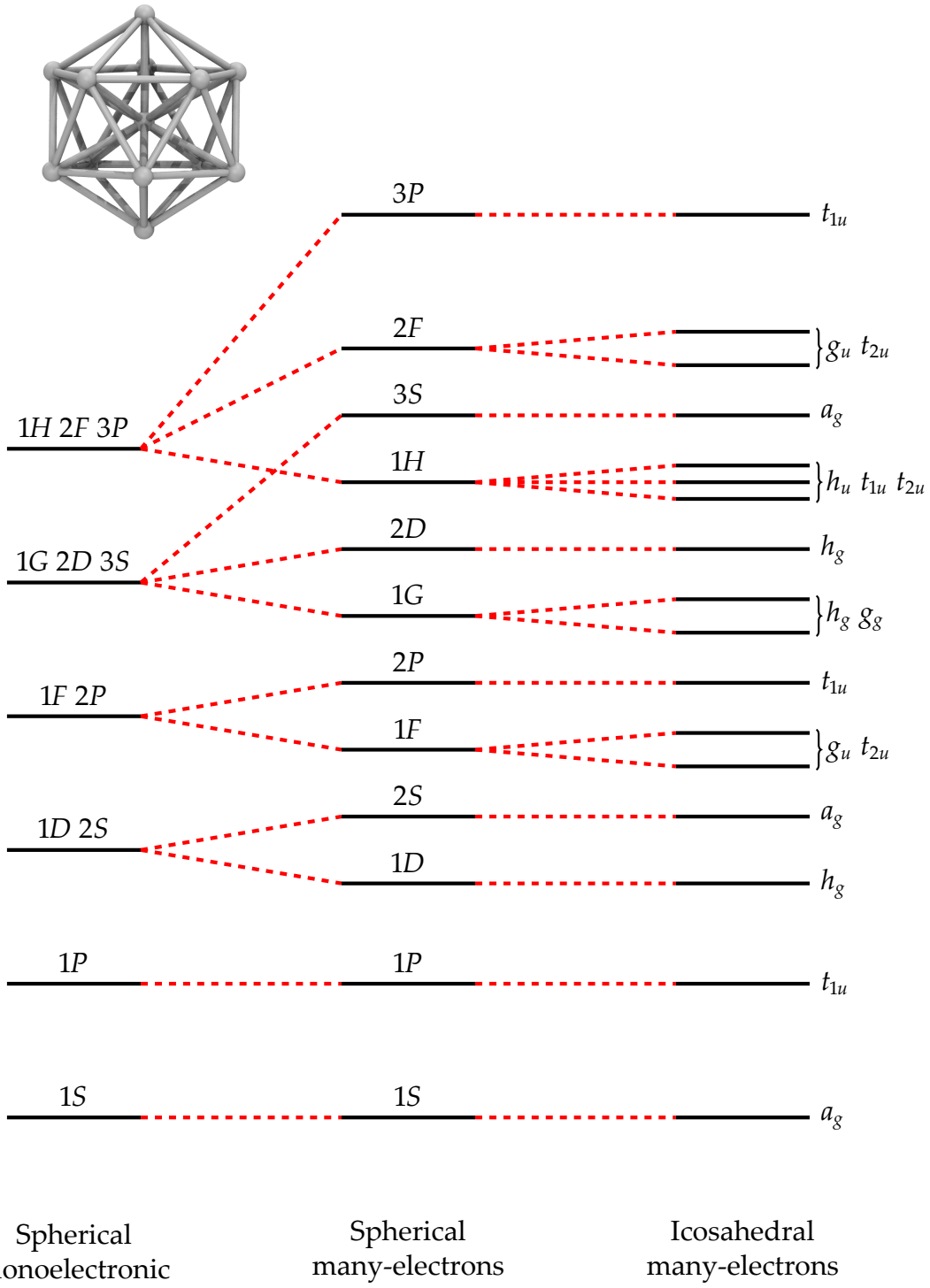


Figure 5.1: Electronic shell structure of the spherical jellium model with an harmonic potential. In the spherical monoelectronic case the energy levels have a constant $2n + l$ value. When electron-electron interaction appears the spherical energy levels split. These many-electron levels are, in turn, further split due to the finite point group symmetry of a given cluster.

configurations (singlets or doublets) are the ground-states of small simple metal clusters. In contrast, some mixed metal clusters have strong spin magnetic moments and have valence electron configurations similar to Mn and Cr.[289,300] These very stable high-spin metal clusters can have potential applications in spintronic devices, where information is stored via electron spins rather than charges.

Different to mixed metal clusters, it is unclear if high-spin configurations of simple metal clusters can become relevant. In particular, we seek if these high-spin simple metal clusters provide explanations to more recent experimental features that cannot be fully explained by electronic shell closing effects. A prominent example are the trends in the melting temperatures of small sodium clusters.[298,301,302]

Therefore, a systematic study of small simple metal clusters in high-spin configurations is presented, where all the calculated high-spin states correspond to electronic configurations where subshell closing of a spin manifold occurs.

5.1.1 Sodium clusters

Surprisingly, the trends in the melting temperature of small sodium clusters cannot be straightforwardly explained by their electronic shell structure. In particular, the relative high melting temperature of Na_{55}^{+} around 290 K is not obviously related to its electronic shell structure.[301,302] The 54 valence electron of Na_{55}^{+} give rise to the following electronic configuration

$$1S^2\ 1P^6\ 1D^{10}\ 2S^2\ 1F^{14}\ 2P^6\ 1G^{14}$$

Thus, 4 electrons are missing for closing the 1G shell, which occurs in Na_{59}^{+}. This is in complete agreement with the abundance of cationic sodium clusters in mass spectra,[303] where a pronounced peak for Na_{59}^{+} is observed. This enhanced stability, however, does not correspond to a higher melting temperature. Instead, the less stable Na_{55}^{+} cluster shows a considerably higher melting temperature than the more stable Na_{59}^{+} cluster. In order to explain this discrepancy the concept of geometrical shell closing has been introduced.[301] In the particular case of Na_{55}^{+} a closed-shell icosahedron is assumed based on the comparison

of the measured photoelectron spectra and the calculated density of states.[304] However, a closed-shell Na_{55}^{+} cluster will undergo Jahn-Teller distortion. Figure 5.1 shows that for an icosahedral cluster the 1G shell is split into a five-fold degenerated h_g subshell and a four-fold degenerated g_g subshell. In fact, the fourteen electrons on the 1G shell can be arranged as $(h_g)^{10}\ (g_g)^4$, where the α and β spin manifolds of the h_g subshell are closed and only the α manifold of the g_g subshell is closed. This, of course, will favor a perfectly icosahedral quintet state. A fundamental question then arises. Can the subshell closing of only one spin manifold explain the discrepancy between the magic numbers in cluster melting and cluster abundance in mass spectrometry? In order to gain insight into this question we searched high-spin configurations of small sodium clusters, up to Na_{55}^{+}, were shell or subshell closing of a spin manifold occurs. We will devote particular attention to Na_{55}^{+} to try to explain the discrepancy in the magic numbers.

The equilibrium geometries of small sodium clusters have been extensively studied with different theoretical methods. For example, Martins, Buttet and Car studied sodium clusters Na_n and Na_n^+ with $n \leqslant 8$ and $n = 13$ by means of the local spin density approximation and a pseudopotential approximation to treat core electrons.[305] Röthlisberger and Andreoni also studied small sodium clusters within the local spin density approximation and employing pseudopotentials.[306,307] More recently, Solov'yov, Solov'yov and Greiner used the B3LYP DFA to find the equilibrium geometries of neutral and singly charged sodium clusters consisting of up to 20 atoms.[308] Furthermore, Aguado and Kostko studied the equilibrium geometries of neutral and anionic sodium clusters with up to 80 atoms based on global minima found with the Gupta potential.[309] The ground-state geometries of these studies were used to compare how relevant a given high-spin configuration might be. The study begins by comparing the energies obtained at the PBE/DZVP/GEN-A2 level theory. Further computations with the hybrid DFA PBE0 are then presented for the most relevant cases.

According to the jellium energy levels for many-electron systems, shown in Figure 5.1, the first high-spin configuration with a closed subshell corresponds to the five-electron configuration $1S^2 1P^3$. This configuration is possible only for tetrahedral or cubic symme-

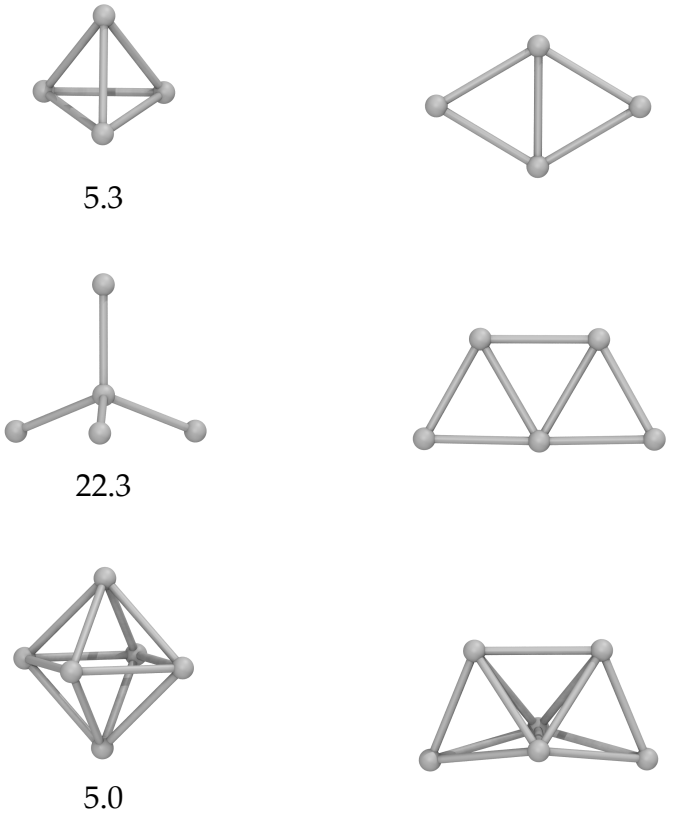


Figure 5.2: Optimized structures of Na_4^- (top), Na_5 (middle) and Na_6^+ (bottom). The left column shows the high-spin structure while the right column shows the corresponding ground states. The relative stabilities [kcal/mol] of the high-spin clusters at the PBE/DZVP/GEN-A2 level of theory are shown below the structure. Bonds are drawn between all nearest neighbors.

try with three-fold degeneracies. Thus, Na_4^-, Na_5 and Na_6^+ clusters in T_d, T_d and O_h symmetry, respectively, were investigated. Figure 5.2 shows the optimized structures of these clusters along with the corresponding ground states. Also shown in Figure 5.2 are the relative stabilities [kcal/mol] of the high-spin quartets. None of the quartet clusters were the ground-states, however, both Na_4^- and Na_6^+ clusters are within 5 kcal/mol from their corresponding ground-states.

The next high-spin configuration occurs for a T_d symmetry with $(a_1)^2\,(t_2)^6\,(e)^2$ triplet configuration, where the $1D$ shell is split into two subshells with e and t_2 irreducible representations. This T_d symmetry and electron count can be realized with the Na_{10} cluster, shown in Figure 5.3. However, the calculation predicts the $2S$ shell to be below of the $1D$ shell leading to an electronic $(a_1)^2\,(t_2)^6\,(a_1)^1\,(e)^1$ configuration. As a consequence, the Na_{10} triplet state lies around 13 kcal/mol above the almost degenerated D_{4d} and C_2 singlet

ground-states.

An eleven electron quartet in O_h symmetry follows. The Na_{12}^+ cluster can be arranged into a cuboctahedron, shown in Figure 5.3, with an empty center to fulfill the symmetry and electron count requirements. The expected electronic configuration, $(a_{1g})^2\,(t_{1u})^6\,(t_{2g})^3$, is in fact obtained. However, this hollow geometry lies around 13 kcal/mol above the C_{2v} doublet ground state.

The thirteen electron Na_{12}^- cluster can be arranged into an icosahedron where each one of the sodium atoms occupies a vertex position. The 1D shell does not split in icosahedral symmetry, therefore, the thirteen electrons can be configured into a sextet as $1S^2\,1P^6\,1D^5$. However, the Na_{12}^- icosahedron is a hollow structure. Similar to the Na_{12}^+ cuboctahedron, this vacancy diminishes significantly the stability of the cluster. The ground-state Na_{12}^- doublet was obtained by re-optimizing the Na_{12} neutral cluster described in reference [308]. The structures of the Na_{12}^- icosahedral sextet and the re-optimized doublet are shown in Figure 5.3. Due to the vacancy, the Na_{12}^- sextet turned out to be 10 kcal/mol less stable than the low-spin Na_{12}^- cluster.

A very similar I_h structure, having the same electron configuration as Na_{12}^-, can be obtained by placing one extra sodium atom in the center of the icosahedron. This neutral Na_{13} sextet is, so far, the most relevant high-spin configuration found in the series. Although it is not the ground-state, it is just 1.3 kcal/mol less stable than the C_1 global minimum reported in reference [309] and 3.3 kcal/mol less stable than another C_1 structure obtained by re-optimization of the minimum reported in reference [308]. Given the small energetical difference, and the high degeneracy of a sextet state, the icosahedral Na_{13} cluster could explain why the measured dipole moment of the thirteen atom sodium cluster essentially vanishes.[310] Furthermore, this very small energetical difference should enable the detection of the Na_{13} in a molecular beam through a Stern-Gerlach experiment. Unfortunately, no reference to Stern-Gerlach experiments performed on Na_{13} clusters was found in the literature. Figure 5.4 shows the structures of the icosahedral Na_{13} cluster and the minimum structures reported in references [309] and [308]. Solov'yov's structure resemble two interpenetrating icosahedra, while Aguado's structure is closer to a capped

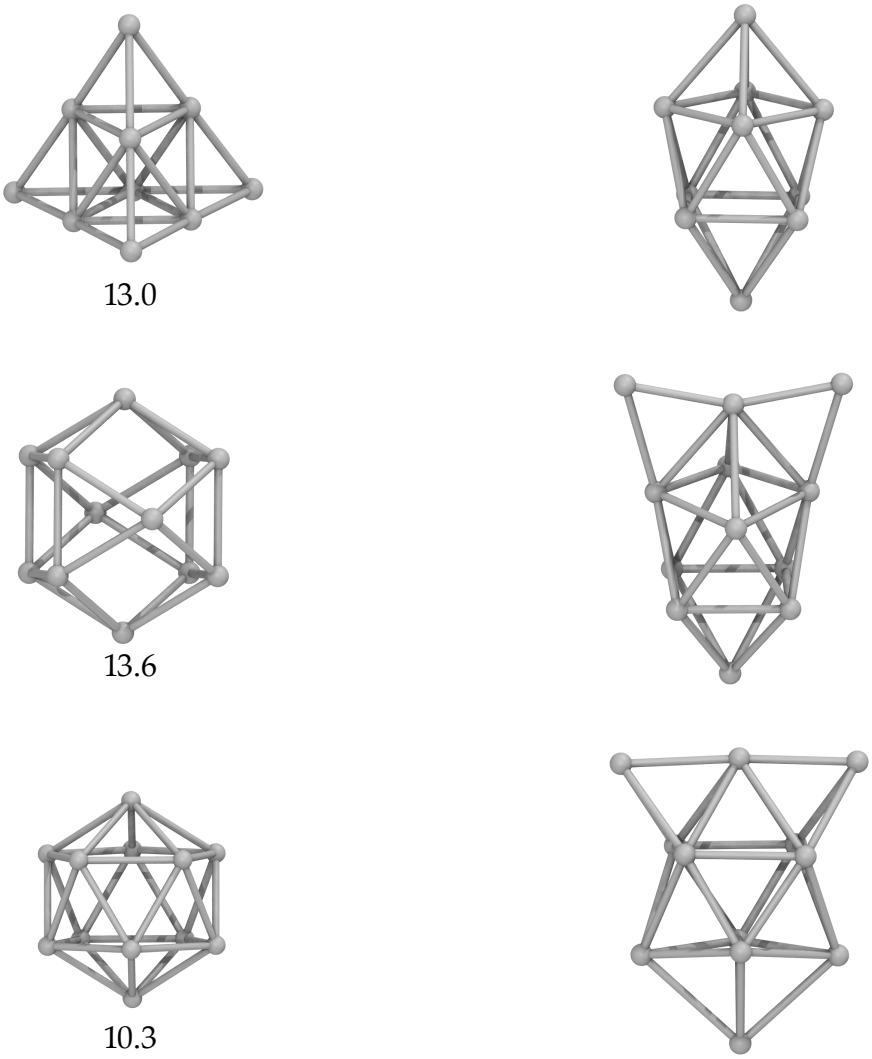


Figure 5.3: Optimized structures of Na_{10} (top), Na_{12}^{+} (middle) and Na_{12}^{-} (bottom). The left column shows the high-spin structure while the right column shows the corresponding ground states. The relative stabilities [kcal/mol] of the high-spin clusters at the PBE/DZVP/GEN-A2 level of theory are shown below the structure. Bonds are drawn between all nearest neighbors.

tetrahedron. The COs of the icosahdral Na_{13} are shown in Figure 5.5. Note that both α and β spin manifolds completely fill the 1S and 1P shells, however, only the α manifold closes the 1D shell.

From now on, we search only for structures without vacancies. This restriction follows from the fact that smaller hollow structures in high-spin configurations are at least 10 kcal/mol less stable than other low-spin minima found. Unfortunately, other T_d and O_h high-spin structures did not follow the energetic ordering of the the spherical jellium model and resulted to be rather unstable. For example, the Na_{25} O_h cluster in a quartet state is less stable than the C_1 doublet ground-state by more than 60 kcal/mol.

Two more clusters, with I_h symmetry and possible high-spin configurations remain in our investigated series. The first one is the Na_{45} neutral cluster, which can adopt an

$$(a_g)^2\ (t_{1u})^6\ (h_g)^{10}\ (a_g)^2\ (t_{2u})^6\ (g_u)^8\ (t_{1u})^6\ (h_g)^5$$

configuration. This configuration results from the splitting of the 1G shell into h_g and g_g subshells. The α manifold of the h_g subshell is then completely filled. This icosahedral cluster is constructed on top of the Na_{13} I_h one, where the second icosahedral shell takes the anti-Mackay[311] positions and has a rhombic triacontahedron shape. In fact, this high-spin icosahedral cluster is the most stable Na_{45} cluster around the rhombic triacontahedron structure. The Na_{45} sextet is around 1 kcal/mol more stable than both slightly distorted doublet and quartet clusters. The structures for the Na_{45} sextet and doublet are shown in Figure 5.6. The quartet geometry is essentially the same as the doublet one. Also shown in Figure 5.6 are two other doublet cluster geometries. One based on an incomplete second icosahedral Mackay shell and the other one obtained by re-optimizing the global minimum reported in reference [309]. It is interesting to note that the Na_{45} sextet is also more stable than the incomplete Mackay shell structure. However, the re-optimized Na_{45} cluster turned out to be more stable than the high-spin Na_{45} by almost 5 kcal/mol. This neutral Na_{45} sextet may also be detected in a molecular beam by Stern-Gerlach-type deflection but, again, no reference has been found for this experiment.

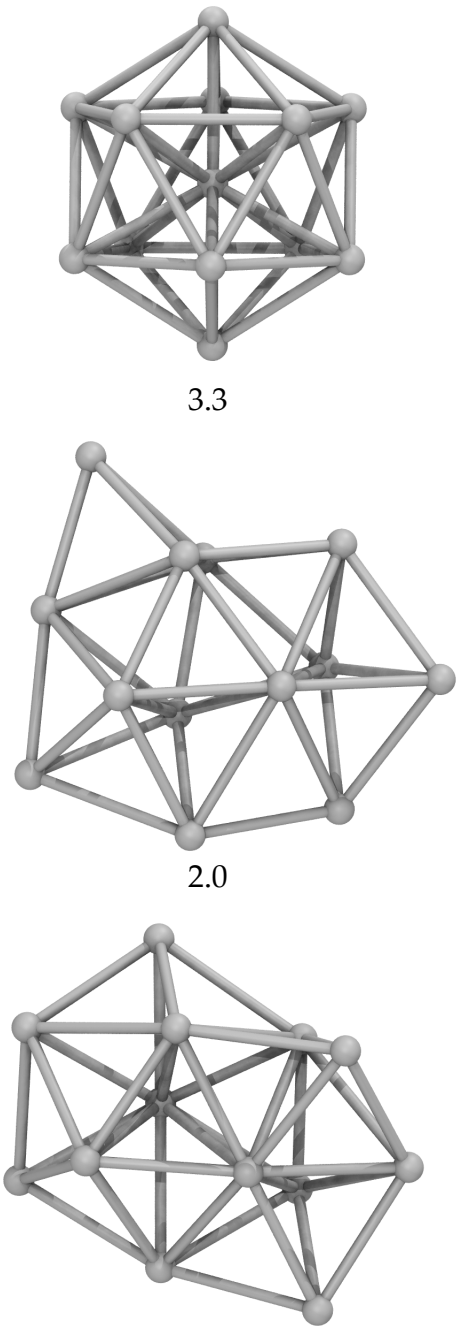


Figure 5.4: Optimized Na_{13} clusters. The icosahedral Na_{13} sextet cluster is shown at the top, the re-optimized doublet from reference [309] is shown at the middle and the one from reference [308] at the bottom. Relative stabilities [kcal/mol] at the PBE/DZVP/GEN-A2 level of theory are shown below each structure. Bonds are drawn between all nearest neighbors.

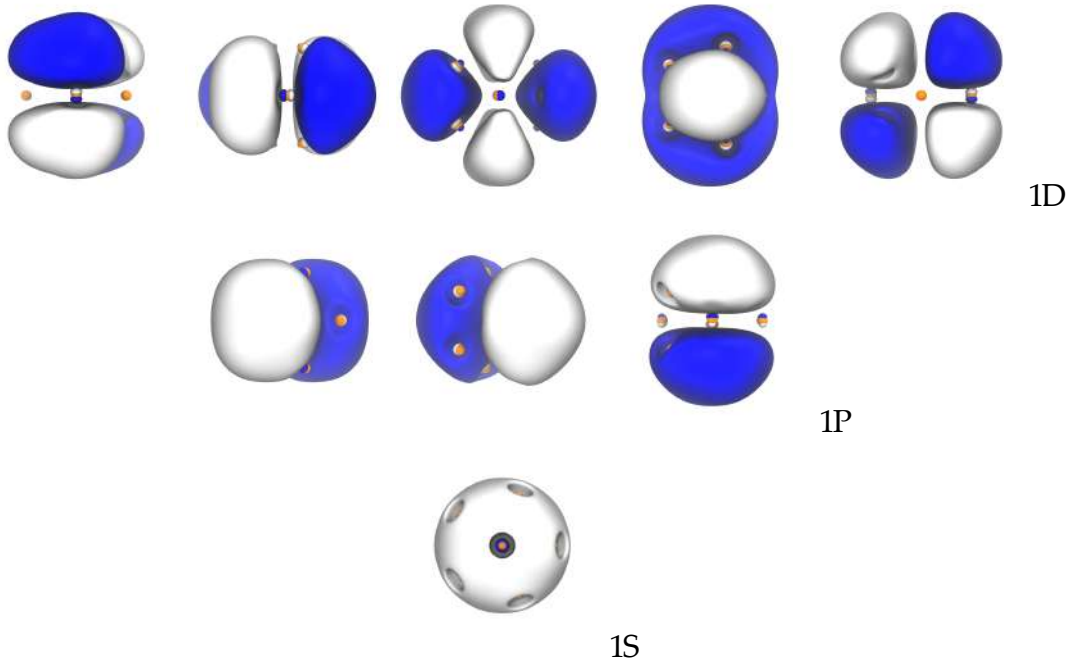


Figure 5.5: Cluster orbitals of the icosahedral Na_{13} cluster. The sodium atoms are shown as orange spheres.

The last cluster studied is the already mentioned Na_{55}^{+} cluster in I_h symmetry. Its valence electrons can be configured as

$$(a_g)^2\ (t_{1u})^6\ (h_g)^{10}\ (a_g)^2\ (t_{2u})^6\ (g_u)^8\ (t_{1u})^6\ (h_g)^{10}\ (g_g)^4$$

which results from the splitting of the 1G shell. The electron configuration is very similar to the previously discussed Na_{45} example. In this case, the β manifold also fills the h_g subshell. The four remaining electrons of Na_{55}^{+} close the α g_g subshell. This icosahedral cluster is the most prominent example in the series because all calculations establish an I_h structure as the most stable one. Different to the singlet case, the high-spin quintet configuration given above provides a very stable electronic environment that avoids any Jahn-Teller distortion. In fact, we have found that the quintet configuration is the ground state of the Na_{55}^{+} cluster. It is 1.6 kcal/mol and 2.5 kcal/mol more stable than the slightly distorted triplet and singlet states. Although the relative stabilities of the low-spin configurations do not automatically discard them, the high-spin state can further be related to other experimental data. One of the most important ones is the comparison of the measured photoelectron spectra[304] with the calculated density of states (DOS). The com-

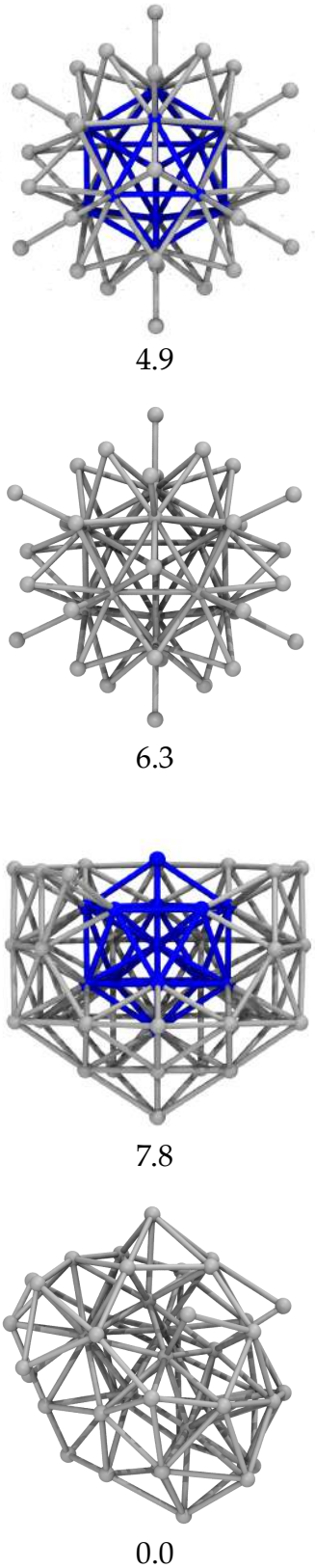


Figure 5.6: Optimized Na_{45} clusters. The icosahedral Na_{45} sextet is shown at the top and its first icosahedral shell is highlighted in blue. The upper middle image shows the C_i Na_{45} doublet structure. The lower middle image shows the Na_{45} doublet cluster which was built by removing 10 atoms of the 55 atom icosahedral structure. Finally, the global minimum structure is shown at the bottom. Relative stabilities [kcal/mol] at the PBE/DZVP/GEN-A2 level of theory are shown below each structure. Bonds are drawn to highlight the structure elements of these clusters.

parison between the calculated density of states of the singlet and quintet Na_{55}^+ and the measured photoelectron spectra is shown in Figure 5.7. The subshell closing of the quintet yields a characteristic shoulder in the frontier orbital DOS at around -4.5 eV. This shoulder is due to the half occupation of the 1G g_g subshell by the α-spin manifold. Therefore, this shoulder is absent in the singlet DOS. This qualitative difference may bridge directly to experiment, because DOS structures are often reproduced by photoelectron spectra. In fact, the experimental photoelectron spectrum[304] of Na_{55}^+, shown as the top image of Figure 5.7, possesses a narrow peak for the valence electrons with a characteristic shoulder, just like our quintet DOS. The structure of the photoelectron spectrum has been used as an argument for the spherical geometry of the cluster. Here, we suggest that the high-spin quintet structure is responsible for the shoulder in the photoelectron spectrum. The g_g subshell closing stabilizes the high-spin quintet state and avoids Jahn-Teller distortions of the cluster. What remains to be seen is how this stabilization of the highly symmetric I_h structure by subshell closing influences the cluster dynamics. As previously mentioned, the trends on the melting temperatures of sodium clusters cannot be straightforwardly explain by shell closing arguments. Although it has been argued that the geometrical shell closing is responsible for the maxima of melting temperatures, there has not been an explanation as to why the melting temperature of Na_{55}^+ is considerably larger than that of the next icosahedral cluster, Na_{147}^+, or the Na_{59}^+ cluster with a completely filled 1G shell. Further studies of the stability of the quintet Na_{55}^+ cluster through Born-Oppenheimer molecular dynamics indeed confirm that the subshell closing affects the dynamics of Na_{55}^+ and predicts a higher melting temperature.[298]

Including exact exchange, through the PBE0 hybrid DFA, introduces important changes in the relative stability of the high-spin clusters. The relative energy between the Na_{13} doublet of reference [309] and that of reference [308] changes from 2.0 kcal/mol with PBE to 1.9 kcal/mol with PBE0. However, the relative energy between the doublet ground-state and the high-spin sextet reduces from 3.3 kcal/mol with PBE to only 1.2 kcal/mol with PBE0. Different to PBE, the hybrid PBE0 predicts that the high-spin sextet is the second most stable structure. Furthermore, PBE0 also increases the distance of the surface Na

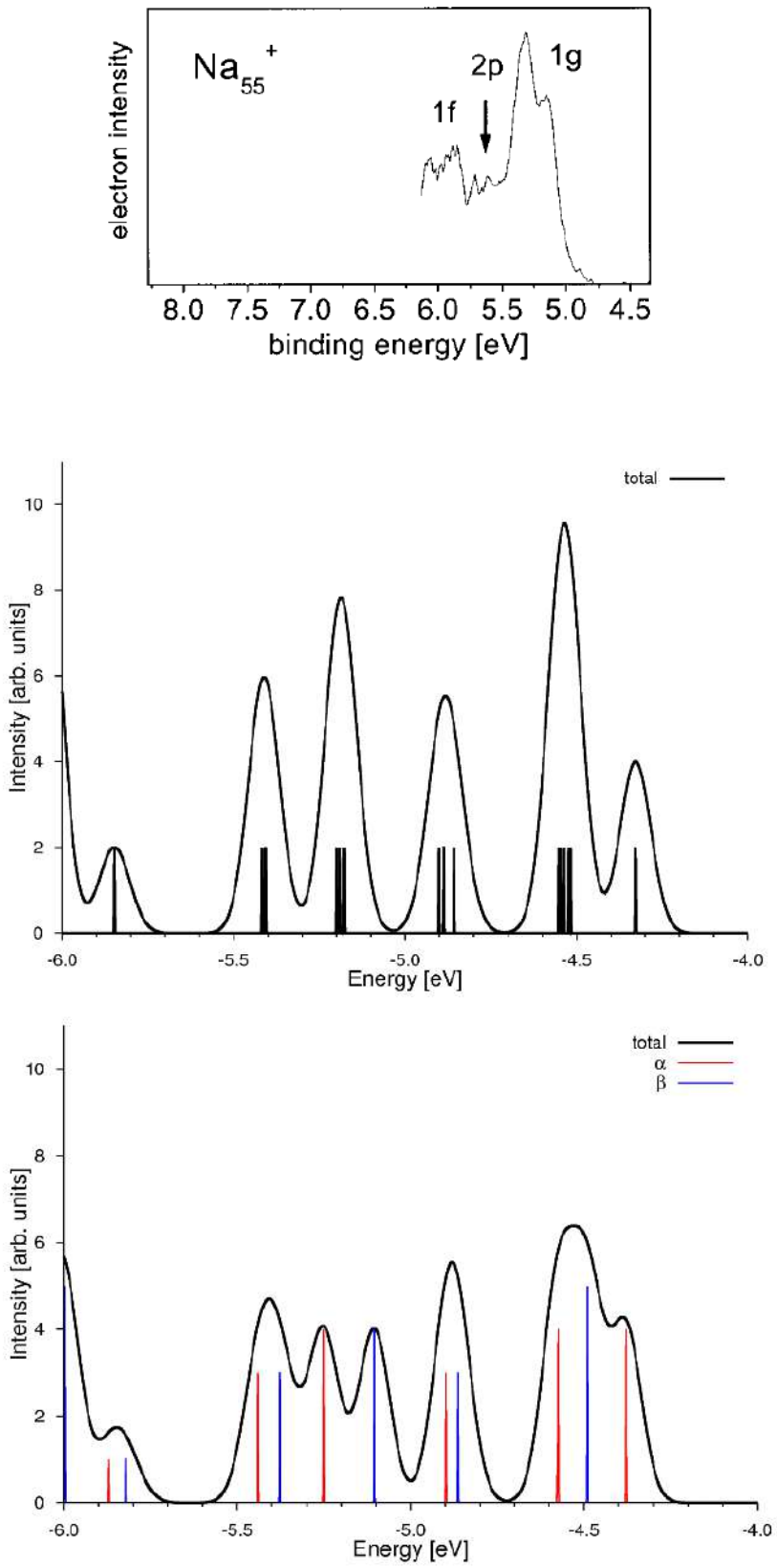


Figure 5.7: Comparison of the measured photoelectron spectra (top) and the calculated density of states of the singlet (middle) and quintet (bottom) Na_{55}^{+} cluster. The photoelectron spectra is a reprinted figure with permission from [G. Wrigge, M. Astruc Haffmann and B. v. Issendorff, Phys. Rev. A, 65, 063201, 2002] Copyright 2002 by the American Physical Society.

Table 5.1: Geometrical parameters and relative energies of the icosahedral sodium clusters employing the PBE and PBE0 functionals. Calculations were performed with the DZVP basis and GEN-A2 auxiliary function set. Hybrid calculations used the GEN-A2/GEN-A2* approach. The distance to the center of mass, r_{COM}, for each type of nonredundant sodium atom is given in Å and the relative energy, ΔE, of the high-spin icosahedral cluster with respect to the ground state is given in kcal/mol.

	PBE			PBE0		
	Na_{13}	Na_{45}	Na_{55}^+	Na_{13}	Na_{45}	Na_{55}^+
		3.372	3.512		3.383	3.531
r_{COM}	3.535	5.639	6.065	3.548	5.660	6.081
		6.628	7.047		6.672	7.107
ΔE	3.3	4.9	0.0	1.2	0.5	0.0

atoms to the center of the icosahedron by 1 pm.

A similar effect is also seen for the icosahedral Na_{45}cluster. In this case, the hybrid PBE0 predicts that the high-spin sextet is only 0.5 kcal/mol less stable than the doublet ground-state. Thus, this high-spin structure may also contribute to the measured dipole moment and photoelectron spectra. The distance to the center of the cluster for the three nonredundant sodium atoms (which are not located at the center themselves) is also elongated with PBE0. The first icosahedral shell is the least affected being shifted only by 1 pm. The two kinds of surface atoms in the second icosahedral shell are 2 pm and 4 pm farther away to the center compared to their PBE counterparts.

The PBE0 functional affects the energetics of the Na_{55}^+ icosahedral cluster much less than that of Na_{13} and Na_{45}. In fact, PBE0 predicts that the quintet is 1.9 kcal/mol more stable than the triplet and 2.9 kcal/mol more stable than the singlet. This means that PBE0 stabilizes the high-spin quintet by 0.4 kcal/mol compared to PBE. The optimized geometry of the icosahedral Na_{55}^+ quintet with PBE0 follows the same trend as for the smaller ichosahedral clusters. The distance to the center of the first icosahedral shell is 2 pm larger with PBE0 compared to PBE. Also, the distance to the center for the second icosahedral shell atoms are 2 pm and 6 pm larger with PBE0. Table 5.1 summarizes the above discussed results by comparing the optimized geometrical parameters [Å] and relative energies [kcal/mol] of the high-spin icosahedral sodium clusters with both PBE and PBE0. Note that Na_{13} and

Na_{45} are $(h_g)^5$ sextets and Na_{55}^+ is a $(g_g)^4$ quintet. Unexpectedly, the inclusion of exact exchange stabilized the high-spin configurations of the three icosahedral clusters compared to PBE. This effect has already been observed in transition metal complexes, mostly involving iron, where the high-spin stabilization depended linearly on the percentage of exact exchange used in hybrid functionals.[312,313] Reiher *et al.*[312] reparametrized the B3LYP hybrid functional to reproduce the experimental spin-state energetics. They found that a 15 % mixing of exact exchange was optimal. It is thus instructive to think that the energetical results obtained with PBE and PBE0 provide bounds for the real spin-state energetics of these clusters. Therefore, we can clearly state that the Na_{55}^+ quintet is the ground state. Furthermore, the Na_{13} cluster should also be considered for the correct description of the experimental data available.

In summary, we have found that spin quenching does occur for small sodium clusters. However, all three high-spin icosahedral sodium clusters, Na_{13}, Na_{45} and Na_{55}^+, should be considered for the correct description of experimental features. We suggest that subshell closing is an important mechanism to be considered for the stability of high-spin metal clusters. In particular, the subshell closing occurring in the Na_{55}^+ cluster stabilizes the icosahedral structure predicted by photoelectron spectroscopy.[304] Furthermore, the subshell closing mechanism might be responsible for the unusually high melting temperature of the Na_{55}^+ cluster. The subshell closing might also be important for the future search of magnetic superatoms[300] and magnetic metal clusters with potential applications in spintronics.

5.1.2 Copper clusters

The study of high-spin copper clusters is interesting in many respects. Different to the sodium atom, the presence of *d* orbitals near the valence *s*-electron introduces some directionality in the bonds formed in the cluster.[314] Therefore, it is surprising that canonical Kohn-Sham calculations also predict the COs obtained with the jellium model.[285] Furthermore, the copper clusters are much more stable than their sodium counterparts.[315] As a consequence, copper clusters are more likely to be found in real life applications. In par-

Table 5.2: Geometrical parameters and relative energies of the icosahedral copper clusters described with PBE and PBE0. Calculations were performed with the DZVP basis and GEN-A2 auxiliary function set. Hybrid calculations used the GEN-A2/GEN-A2* approach. The distance to the center of mass, r_{COM}, for each type of nonredundant copper atom is given in Å and the relative energy, ΔE, of the high-spin icosahedral cluster with respect to the ground state is given in kcal/mol.

	PBE			PBE0		
	Cu_{13}	Cu_{45}	Cu_{55}^+	Cu_{13}	Cu_{45}	Cu_{55}^+
		2.403	2.472		2.389	2.458
r_{COM}	2.460	3.937	4.258	2.448	3.905	4.216
		4.686	4.904		4.661	4.884
ΔE	8.5	29.4	0.0	4.5	9.6	0.0

ticular, high-spin clusters can be ideal for molecular electronic devices, as the coupling could be altered by charging or weak fields.[316]

Therefore, it is important to establish if the subshell closing mechanism plays the same crucial role in determining the stability of high-spin copper clusters as in the sodium case. To this end, we focused on the same icosahedral arrangements than the ones obtained for Na_{13}, Na_{45} and Na_{55}^+. The same levels of theory, PBE/DZVP/GEN-A2 and PBE0/DZVP with the GEN-A2/GEN-A2* approach, were used. These results are summarized in Table 5.2. Inclusion of exact exchange also stabilizes the high-spin states of the copper clusters. However, Cu_{13} and Cu_{45} are more than two times less stable than their sodium counterparts. The same degree of stabilization as in the sodium cluster case is obtained with the PBE0 hybrid functional for the Cu_{55}^+ high-spin cation, namely, 0.4 kcal/mol compared to the PBE results. Based on these results we can clearly state that the high-spin configurations of Cu_{13} and Cu_{45} do not play an important role for the description of measured quantities. However, it is important to note that the high-spin Cu_{55}^+ quintet is also predicted to be the ground-state. In this case, PBE predicts that the quintet Cu_{55}^+ is 1.3 kcal/mol more stable than the triplet and 2.0 kcal/mol more stable than the singlet. On the other hand, PBE0 predicts the quintet to be 1.8 kcal/mol more stable than the triplet and 2.4 kcal/mol more stable than the singlet. This result encourages further research on magnetic clusters made from simple metals, as they could be incorporated more easily into real life applica-

tions.

Different to the sodium clusters case, exact exchange made the copper clusters a little bit more compact. The hybrid PBE0 predicts that the icosahedral shell of Cu_{13} is 1 pm closer to the center compared to the PBE result. The same is true for the first icosahedral shells of the Cu_{45} and Cu_{55}^{+} clusters. Furthermore, the second icosahedral shell gets also compacted when exact exchange is included.

In summary, the subshell closing stabilization mechanism is also present in copper clusters. This shows that high-spin cluster configurations should not be neglected in further research, in particular those involving the closure of shells and subshells. Further work to gain insights into the dynamical stabilization of the subshell closing is also being carried out in our laboratory.

5.2 The Accuracy of DFT for Transition Metal Compounds

Predicting chemical bonding within organic compounds is relatively straightforward. This picture changes dramatically when turning to transition metal compounds, specially when dealing with metal atoms having partially filled *d*-shells.[317] DFT has become the preferred method for electronic structure theory because its cost scales more favorably with system size than does the cost of correlated wavefunction methods, and yet it competes well in accuracy.[318] The advantages of DFT are specially important for transition metals. The reason for this is electron correlation. Due to the partially filled d shells and nearly degenerate $(n+1)s$ and nd shells, systems containing transition metals often have many low-lying nearly degenerate states.[319] The correlation effects on geometries, densities and energies, due to the near-degeneracy, can be very large.[318] Furche and Perdew showed that GGAs and meta-GGAs functionals yield geometries that are in very good agreement to experimental gas phase structures for 3*d* metals.[319] In the same study, they argued against the use of hybrid functionals because the more strongly correlated the system, the less exact exchange is needed for its accurate description.[319] One of the best DFT descriptions obtained so far for 3*d* transition metal clusters was presented by Calaminici *et al.* by employing DFT

optimized basis sets.[201] The use of optimized basis sets was also emphasized for transition metal carbonyls by showing that double-ζ optimized basis sets can outperform the larger correlation consistent triple- and quadruple-ζ basis sets.[320]

However, some problems remain that affect the accuracy of DFT despite the active research and development of new exchange and correlation DFAs.[321] In short, two main problems have been identified. First, pure dispersion interactions are not well reproduced with many functionals and, second, Coulomb self-interaction is not correctly canceled out by the exchange DFAs leading to overstabilization where delocalized electrons and low-spin configurations occur. On the other hand, inclusion of exact exchange through hybrid functionals overstabilizes high-spin states.[312,313,317,322]

Despite the issues remaining with DFT, it still yields more accurate results than other high-level correlated *ab initio* methods like CCSD(T) for transition metal compounds.[323] Nevertheless, most theoretical studies on transition metal compounds have relied on small basis sets or introduced frozen-core approximations to alleviate the computational complexity of the computation.[317] In this study, we take advantage of the low computational complexity of ADFT and LDF-EXX hybrid ADFT methods to employ the augmented correlation-consistent basis sets proposed by Dunning,[216,324–327] with double-, triple- and quadruple-ζ quality. Standard ADFT calculations are performed with the widely used PBE functional and the GEN-A2 auxiliary function set, whereas hybrid ADFT calculations used the PBE0 functional and the GEN-A2/GEN-A2* approach.

Figure 5.8 shows the geometrical parameters used to compare the influence of both the basis set and exact exchange. In particular, we will focus on metal-carbon bond distances and the deviation of the hydrogen atoms from the plane of the aromatic rings, where applicable. Experimental structures used for comparison were, preferably, from gas-phase data and can be found in references [328–340]. Table 5.3 lists bond lengths [Å] and angles [°] for the sixteen transition metal compounds studied. To ease the notation we will call the aug-cc-pVXZ basis set XZ only. Note that going from DZ to TZ and QZ can show sometimes an oscillatory behavior. For instance, the Ti−C bond length in $Ti(CH_3)_4$ changes from 2.084 Å with DZ to 2.095 Å with TZ and goes back to 2.090 Å with QZ.

Table 5.3: Bond distances [Å] and bond angles [°] of selected transition metal organometallic compounds. Experimental structures were taken from references [328–340]. Cb stands for η^4-cyclobutadiene, Cp for η^5-cyclopentadienyl and Bz for η^6-benzene substituents.

Molecule	Point Group	Structural Parameter	PBE DZ	PBE TZ	PBE QZ	PBE0 DZ	PBE0 TZ	PBE0 QZ	Expt.
$Ti(CH_3)_4$	T_d	r(TiC)	2.084	2.095	2.090	2.076	2.087	2.084	2.14
		r(CH)	1.117	1.106	1.105	1.109	1.098	1.098	
		∠(HCH)	108.0	108.4	108.5	107.8	108.1	108.3	
$TiCl_3CH_3$	C_{3v}	r(TiC)	2.037	2.049	2.057	2.019	2.033	2.041	2.047
		r(TiCl)	2.198	2.200	2.201	2.189	2.192	2.195	2.185
		r(CH)	1.116	1.104	1.103	1.107	1.096	1.096	1.098
		∠(ClTiCl)	113.3	113.5	113.3	113.8	114.0	113.8	113.0
		∠(HCH)	110.3	110.6	110.1	110.5	110.6	110.3	109.9
$TiCp_2Cl_2$	C_2	r(TiCp)	2.161	2.168	2.170	2.154	2.160	2.164	
		r(TiC)	2.419	2.432	2.436	2.389	2.401	2.411	2.372
		r(CC)	1.429	1.413	1.413	1.418	1.404	1.404	1.397
		r(CH)	1.099	1.088	1.088	1.091	1.080	1.081	1.113
		r(TiCl)	2.347	2.340	2.330	2.341	2.335	2.326	2.318
		∠(CpTiCp)	130.8	130.5	129.6	131.5	131.1	130.2	128.1
		∠(ClTiCl)	97.54	97.91	98.10	96.78	97.90	97.31	97.1
		∠($\overline{\alpha}$)	5.9	6.0	6.3	3.0	3.5	4.0	8.8
$CrBz(CO_3)_3$	C_{3v}	r(CrBz)	1.705	1.704	1.700	1.690	1.687	1.684	
		r($\overline{CrC}_{Bz}$)	2.224	2.214	2.212	2.204	2.195	2.194	2.208
		r(CC)	1.427	1.414	1.416	1.415	1.404	1.406	1.417
		r(CH)	1.104	1.089	1.091	1.093	1.082	1.083	
		r(CrC_{CO})	1.838	1.839	1.838	1.840	1.841	1.842	1.863
		r(CO)	1.174	1.166	1.165	1.157	1.149	1.148	
		∠($CCrC_{CO}$)	86.4	87.0	86.9	85.9	86.4	86.4	88.6
		∠($\overline{\alpha}$)	3.3	3.9	3.5	3.2	3.5	3.2	
$CrCp_2$	D_{5h}	r(CrCp)	1.757	1.769	1.792	1.768	1.852	1.807	1.798
		r(CrC)	2.145	2.149	2.166	2.147	2.208	2.172	2.169
		r(CC)	1.447	1.434	1.431	1.432	1.414	1.417	1.431
		r(CH)	1.101	1.088	1.088	1.092	1.081	1.080	1.108
		∠(α)	0.54	1.3	2.3	0.7	0.5	1.0	2.9
$CrBz_2$	D_{6h}	r(CrBz)	1.623	1.627	1.624	1.622	1.623	1.619	1.613
		r(CrC)	2.163	2.158	2.157	2.155	2.149	2.147	2.152
		r(CC)	1.430	1.417	1.419	1.419	1.408	1.409	1.423
		r(CH)	1.105	1.093	1.094	1.095	1.084	1.084	1.090
		∠(α)	4.5	4.4	3.8	4.0	3.7	3.4	4.7

Table 5.3: continued

Molecule	Point Group	Structural Parameter	PBE DZ	PBE TZ	PBE QZ	PBE0 DZ	PBE0 TZ	PBE0 QZ	Expt.
$MnCp(CO)_3$	C_s	r(MnCp)	1.776	1.780	1.790	1.769	1.772	1.781	
		r($\overline{MnC}_{Cp}$)	2.157	2.154	2.159	2.145	2.143	2.149	2.165
		r($\overline{CC}$)	1.440	1.426	1.424	1.427	1.415	1.414	1.42
		r($\overline{CH}$)	1.099	1.087	1.088	1.091	1.080	1.081	
		r($\overline{MnC}_{CO}$)	1.782	1.784	1.783	1.783	1.785	1.785	1.80
		r($\overline{CO}$)	1.171	1.163	1.162	1.153	1.145	1.144	1.15
		$\angle(\overline{CMnC}_{CO})$	90.9	91.6	91.8	90.6	91.1	89.7	92
		$\angle(\alpha)$	0.7	1.2	1.0	0.8	0.6	0.4	
$MnCp_2$	D_{5h}	r(MnCp)	2.049	2.059	2.082	2.050	2.057	2.071	2.046
		r(MnC)	2.385	2.389	2.407	2.382	2.383	2.394	2.383
		r(CC)	1.437	1.424	1.422	1.425	1.414	1.412	1.429
		r(CH)	1.100	1.087	1.089	1.092	1.081	1.081	1.133
		$\angle(\alpha)$	-1.4	-1.1	-0.5	-1.5	-1.5	-1.2	0
$FeEt(CO)_4$	C_{2v}	r(FeEt)	2.006	2.012	2.014	1.970	1.977	1.980	1.995
		r(FeC)	2.129	2.132	2.134	2.094	2.099	2.101	2.117
		r(CC)	1.428	1.411	1.412	1.420	1.405	1.405	1.419
		r(CH)	1.103	1.090	1.091	1.095	1.083	1.084	1.072
		r(FeC_{ax})	1.802	1.799	1.798	1.806	1.802	1.802	1.815
		r(FeC_{eq})	1.791	1.791	1.789	1.780	1.779	1.778	1.806
		r($\overline{CO}$)	1.164	1.155	1.154	1.146	1.139	1.137	1.143
		$\angle(C_{ax}FeC_{ax})$	176.0	175.2	175.8	174.2	172.9	174.0	
		$\angle(C_{eq}FeC_{eq})$	112.8	111.4	111.7	112.5	111.0	111.0	111.7
		$\angle$(HCH)	114.9	114.8	114.6	114.1	113.9	113.9	
		$\angle(\alpha)$	-21.9	-21.5	-21.3	-23.8	-23.3	-23.1	
$FeCb(CO)_3$	C_s	r(FeCb)	1.772	1.771	1.779	1.747	1.747	1.753	
		r($\overline{FeC}$)	2.054	2.048	2.054	2.028	2.024	2.028	2.063
		r(C−C)	1.475	1.461	1.460	1.465	1.452	1.451	1.456
		r(C=C)	1.462	1.448	1.446	1.448	1.436	1.435	1.456
		r($\overline{CH}$)	1.097	1.084	1.086	1.089	1.077	1.078	
		r($\overline{FeC}_{CO}$)	1.777	1.777	1.775	1.770	1.769	1.768	1.819
		r($\overline{CO}$)	1.167	1.159	1.157	1.149	1.141	1.140	1.131
		$\angle(\overline{CFeC}_{CO})$	91.8	92.0	92.1	91.3	91.5	89.6	95.5
		$\angle(\alpha)$	-7.8	-7.3	-6.2	-8.0	-8.0	-8.0	
$FeCp_2$	D_{5h}	r(FeCp)	1.640	1.640	1.658	1.652	1.651	1.668	1.660
		r(FeC)	2.050	2.044	2.057	2.052	2.045	2.058	2.064
		r(CC)	1.445	1.433	1.432	1.430	1.420	1.418	1.440
		r(CH)	1.100	1.088	1.089	1.091	1.080	1.080	1.104
		$\angle(\alpha)$	1.8	2.1	2.1	2.0	2.0	2.8	3.7

Table 5.3: continued

Molecule	Point Group	Structural Parameter	PBE DZ	PBE TZ	PBE QZ	PBE0 DZ	PBE0 TZ	PBE0 QZ	Expt.
$CoCp(CO)_2$	C_s	r(CoCp)	1.712	1.723	1.712	1.698	1.707	1.699	
		r($\overline{CoC}_{Cp}$)	2.106	2.108	2.099	2.088	2.089	2.083	2.129
		r($\overline{CC}$)	1.443	1.429	1.429	1.429	1.418	1.417	1.450
		r($\overline{CH}$)	1.099	1.087	1.088	1.091	1.079	1.080	1.083
		r($\overline{CoC}_{CO}$)	1.733	1.734	1.735	1.733	1.732	1.734	1.679
		r($\overline{CO}$)	1.170	1.161	1.161	1.151	1.144	1.143	1.191
		$\angle(\overline{CCoC}_{CO})$	98.1	98.7	97.8	97.9	98.6	97.9	98.4
		$\angle(\alpha)$	1.7	1.7	1.7	1.4	1.4	1.0	
NiCp(NO)	C_{5v}	r(NiCp)	1.724	1.735	1.727	1.717	1.726	1.721	
		r(NiC)	2.116	2.118	2.112	2.104	2.105	2.101	2.11
		r(CC)	1.442	1.428	1.428	1.429	1.417	1.417	1.43
		r(CH)	1.099	1.087	1.087	1.091	1.080	1.080	
		r(NiN)	1.626	1.624	1.625	1.592	1.591	1.591	1.626
		r(NO)	1.181	1.172	1.173	1.162	1.154	1.154	1.165
		$\angle(\alpha)$	0.6	1.6	1.6	0.8	1.4	1.1	
$NiCp_2$	D_{5h}	r(NiCp)	1.813	1.822	1.815	1.822	1.828	1.824	1.828
		r(NiC)	2.188	2.189	2.183	2.189	2.189	2.185	2.196
		r(CC)	1.440	1.426	1.426	1.426	1.415	1.415	1.430
		r(CH)	1.100	1.087	1.087	1.091	1.080	1.080	1.083
		$\angle(\alpha)$	0.5	0.4	0.4	0.1	0.4	0.3	0.3
$Zn(CH_3)_2$	D_{3d}	r(ZnC)	1.952	1.955	1.956	1.940	1.943	1.943	1.929
		r(CH)	1.113	1.102	1.102	1.058	1.095	1.095	1.09
		$\angle$(HCH)	107.9	108.0	108.2	107.7	107.9	108.0	107.7
$ZnCp(CH_3)$	C_s	r(ZnCp)	1.920	1.941	1.940	1.908	1.921	1.921	1.932
		r($\overline{ZnC}$)	2.278	2.286	2.288	2.262	2.268	2.266	2.280
		r($\overline{CC}$)	1.438	1.424	1.422	1.426	1.414	1.413	1.422
		r($\overline{CH}$)	1.100	1.087	1.090	1.092	1.080	1.083	1.08
		r(ZnC_{Me})	1.963	1.967	1.955	1.951	1.955	1.946	1.903
		r($\overline{CH}_{Me}$)	1.111	1.099	1.100	1.103	1.093	1.093	1.12
		$\angle$(HCH)	108.9	109.1	108.5	108.7	108.9	108.2	
		$\angle(\alpha)$	0.9	0.5	0.3	0.9	0.9	0.1	

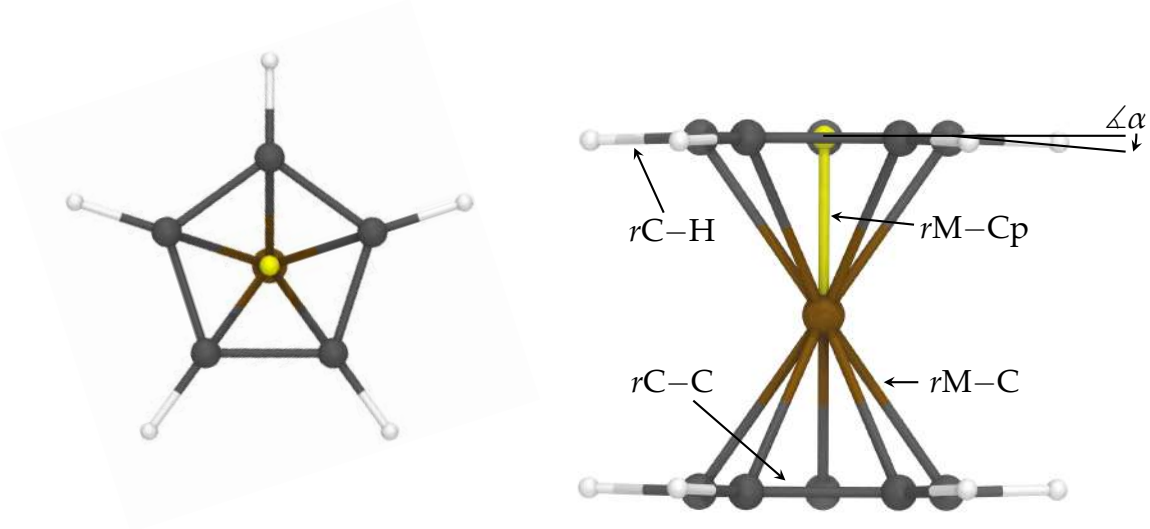


Figure 5.8: Top and side views of a generic metal sandwich complex. The side view shows the geometrical parameters being compared in this study. In particular, the angle α measures the deviation of the hydrogens from the plane of the aromatic ring. As indicated by the arrow, positive values refer to bending towards the metal atom.

Table 5.4: Mean absolute deviations [pm] of relevant bond lengths of transition metal organometallic compounds.

Structural	PBE			PBE0		
Parameter	DZ	TZ	QZ	DZ	TZ	QZ
r(M−X)	2.1	2.2	1.9	2.3	2.3	2.0
r(M−Cp)	1.7	1.5	1.3	1.3	3.2	1.2

These oscillations are usually rather small. By and large, PBE performs rather good to describe the geometries of these complex systems. The structures obtained with PBE0 are slightly worse than those of PBE. However, some particular cases should be noted. These cases will be discussed taking into account the mean absolute deviations (MAD) for the 26 M−C, M−N and M−Cl bonds and the 5 M−Cp distances shown in Table 5.4. Note that increasing the basis set size improves the PBE M−X and M−Cp distances, however, an important deterioration is seen for the M−Cp distance when increasing the basis set size, from DZ to TZ, with PBE0. This deterioration is caused by the very poor description of the Cr−Cp distance with the PBE0 functional. This distance is either 3 pm too short with the DZ basis or 5 pm too long with the TZ basis. A 1 pm accuracy is recovered with the very large QZ basis set. All other M−Cp PBE0 distances are predicted within 1 pm accuracy. Some important deviations, common to both PBE and PBE0, should also be noted. The

most important one is the Co−C_{CO} bond distance in $CoCp(CO)_2$, which is predicted to be 6 pm too large with both functionals and all basis sets! The experimental reference for this structure comes from gas-phase electron diffraction, and the authors already noted that this Co−C bond was unusually short.[335] Consequently, the experimental C−O bond length is elongated and the predicted bond lengths turned out to be 3 pm too short for PBE and 5 pm too short for PBE0. Another important deviation occurs for the Zn−C_{Me} bond length in the methyl(cyclopentadienyl)zinc complex. Both PBE and PBE0 predict elongated bonds deviating by up to 6 pm from the gas-phase experimental measurement. The consistency between PBE and PBE0 bond lengths, along with some large deviations in the experimental measurements (see for example [328] and [329]) suggest that some of these experimental measures should be revisited.

The bending of the hydrogen atoms out of the plane of the aromatic rings are almost equally well described with PBE and PBE0. Again, the largest difference is obtained for chromocene with more than 2° deviation from the experimental measurement. The C−C and C−H bonds are consistently shorter, by 1 pm, with PBE0 than with PBE. The shortening of these bonds makes the PBE0 predictions worse. The same is true for carbonyl C−O bonds and the nitrosyl N−O bond, where the bonds are shortened by almost 2 pm.

Altogether, PBE and PBE0 describe the structural parameters of the here studied organometallic compounds reasonably well if large basis sets are used, as the here employed aug-cc-pVQZ basis. Structural details are qualitatively correct, as shown by the α bending angles in Table 5.3. Nevertheless, unexpected deviations may occur as discussed above.

Another important aspect we studied on these compounds are spin-state splittings. Especially large differences between PBE and its hybrid PBE0 can be found for chromocene, manganocene and nickelocene for these energy splittings. Table 5.5 lists relative energies [kcal/mol] for a low-spin, an intermediate-spin and a high-spin configuration for each of these metallocenes. For chromocene, PBE and PBE0 predict triplet ground states with all basis sets employed. However, with PBE0 the splitting between the ground-state triplet and the quintet vanishes with DZ, is only 3.2 kcal/mol with TZ and 1.4 kcal/mol with QZ. In contrast, the PBE splitting is more than 10 kcal/mol for all basis sets. A similar behavior

Table 5.5: Relative energies [kcal/mol] of different spin states of chromocene, manganocene and nickelocene. Calculations were performed with PBE and PBE0 and the aug-cc-pVXZ basis set (X = D,T,Q). For PBE0 calculations, the GEN-A2/GEN-A2* approach was used.

Molecule	PBE			PBE0		
	DZ	TZ	QZ	DZ	TZ	QZ
$^{1}CrCp_2$	20.3	21.4	22.6	37.9	39.3	45.9
$^{3}CrCp_2$	0.0	0.0	0.0	0.0	0.0	0.0
$^{5}CrCp_2$	18.3	15.1	10.9	0.0	3.2	1.4
$^{2}MnCp_2$	0.0	0.0	0.0	8.3	7.9	6.7
$^{4}MnCp_2$	27.3	24.7	24.0	21.7	21.6	21.5
$^{6}MnCp_2$	19.0	16.6	14.1	0.0	0.0	0.0
$^{1}NiCp_2$	16.2	15.9	15.8	22.5	22.1	21.8
$^{3}NiCp_2$	0.0	0.0	0.0	0.0	0.0	0.0
$^{5}NiCp_2$	116.8	114.6	113.8	126.4	122.8	122.9

is seen also for nickelocene, where both PBE and PBE0 predict triplet ground-states. The difference arises in the singlet-triplet splitting. Similar to chromocene, PBE0 stabilizes the high-spin state by around 7 kcal/mol.

For manganocene, PBE and PBE0 predict different ground-states with rather large separations from other spin multiplicities. In this case, PBE favors the low-spin doublet state by more than 14 kcal/mol while PBE0 favors the sextet state by more than 7 kcal/mol. This difference has already been documented between the OPTX-PBE "pure"- and B3LYP hybrid-functionals and has been used to tune hybrid functionals for metallic systems.[312,317] It has been pointed out that Mn complexes have low crossover energy barriers between several electronic states.[341] Photoelectronic, electron paramagnetic resonance and gas-phase structure determination have shown the presence of spin equilibrium in manganocene and dimethylmanganocene with $^{2}E_{2g}$ and $^{6}A_{1g}$ as the only well-defined identified states.[342–345] However the doublet $^{2}E_{2g}$ state is favored at very low-temperatures (below 90 K) by about 0.5 kcal/mol.[341,343,346] The here presented results are gas-phase 0 K calculations therefore, the correct spin-state description is obtained by the "pure" functional PBE. It is important to note, however, that the structural parameters presented in Table 5.3 correspond to the high-spin sextet because the electron diffraction experiment was measured at around 100 °C, where a 60:40 mixture of sextet and doublet manganocene is present.[340,346]

In summary, the "pure" DFT description outperforms the hybrid one for the here studied organometallic compounds, even when the very large aug-cc-pVQZ basis sets was employed. PBE geometries are almost at basis set convergence, however, the relative energies can still vary up to 5 kcal/mol when going from aug-cc-pVTZ to aug-cc-PVQZ. Our results discourage the use of hybrid functionals for the calculation of transition metal systems even if only one metal atom is present! PBE0 bond lengths showed unexpected deterioration with increasing basis set size and the spin-state splittings are even qualitatively not reliable. Furthermore, we would like to emphasize that tuning the amount of exact exchange present in hybrid functionals to obtain the correct spin-state splitting will certainly alter, usually for worse, its performance for other properties. An alternative approach could be the optimization of the Dunning basis sets for "pure" DFT functionals.

5.3 Static Polarizabilities of Giant Fullerenes: Where is the limit?

The study of size-dependent properties of nanoparticles has become a frontier line of research in nanoscience. In particular, the study of the electric-dipole polarizability of giant fullerenes has attracted considerable attention.[243,347–352] Unlike most nanoparticles, fullerenes are distinctly different in that they have all the atoms on the surface whereas solid spherical nanoparticles have only roughly $N^{1/3}$ atoms on the surface.[349,351] This feature results in a rapid increase in their volume as the number of atoms increases. As the electric-dipole polarizability is proportional to volume, fullerenes are most suitable for investigating quantum size effects manifested in polarizability.

Several studies on fullerene polarizability have been based on classical models or have employed drastic approximations. The results of these models, however, show rather large variance. For example, the predictions for the C_{540} polarizability vary from 5844 a.u. (Pariser-Parr-Pople model[347]) over 6924 a.u. (Penn model-linear response theory[348]) and up to 12612 a.u. (tight binding[348]).

In this section, we present first-principle static and dynamic polarizabilities of icosahedral fullerenes, ranging from C_{60} to C_{960}, obtained with Auxiliary Density Perturbation

Theory (ADPT). These calculations provide an accurate description of the electronic structure since they are based on all-electron first-principles density functional theory. It is important to note that the only assumption made is that the vibronic contributions are negligible, thus, only the electronic part is shown. This is different to previous studies where, in addition, some other approximate techniques have been invoked to lower the computational complexity of the linear response.[349,350]

The selected methodology was VWN/DZVP/GEN-A2 since it has demonstrated good performance to accuracy ratio.[243,351] This basis and auxiliary function set combination results in more than 14,000 basis functions and 32,000 auxiliary functions for the 960 fullerene. It is important to note that no simplification due to symmetry has been used. Therefore, the iterative ADPT solver can be used for other nonsymmetric systems with the same computational efficiency. Optimized geometries, at the same level of theory, were used for all fullerenes.[243,351] The optimized structures are shown in Figure 5.9 with the view aligned along one of the C_3 axes. All twelve five-membered rings are colored in red in order to aid the view to distinguish the vertices of the icosahedron. Note that C_{60} and C_{180} have rather spherical shapes but the larger fullerenes have clearly polyhedrical shapes. At some point, the icosahedral faces will be so large that they will basically behave as graphene sheets. As a consequence, the specific polarizability (per carbon atom) of these giant fullerenes should be that of the graphene. Here, we attempt to find, by an all-electron first principles calculation, at which size the polyhedral faces can be rendered as graphene sheets.

Table 5.6 lists the obtained ADPT static polarizabilities [a.u.] along with previous results obtained in the literature.[243,347–352] The semiempirical PPP approach clearly underestimates the polarizability of the larger fullerenes. It is interesting to note that for C_{720} the PPP polarizability is only 76 % that of VWN/DZVP/GEN-A2. Furthermore, Ruiz *et al.*[347] obtained that the PPP polarizabilities are 75 % that of a spherical conducting shell. Therefore, the all-electron ADPT results predict that the giant fullerenes behave almost as spherical conducting shells starting from C_{720}. The sum-over-states polarizabilities, corrected by random phase approximation (SOS-RPA), are in very good agreement to the VWN/DZVP/GEN-A2 for all the series. The same is true for the PBE0/SVPD polarizabil-

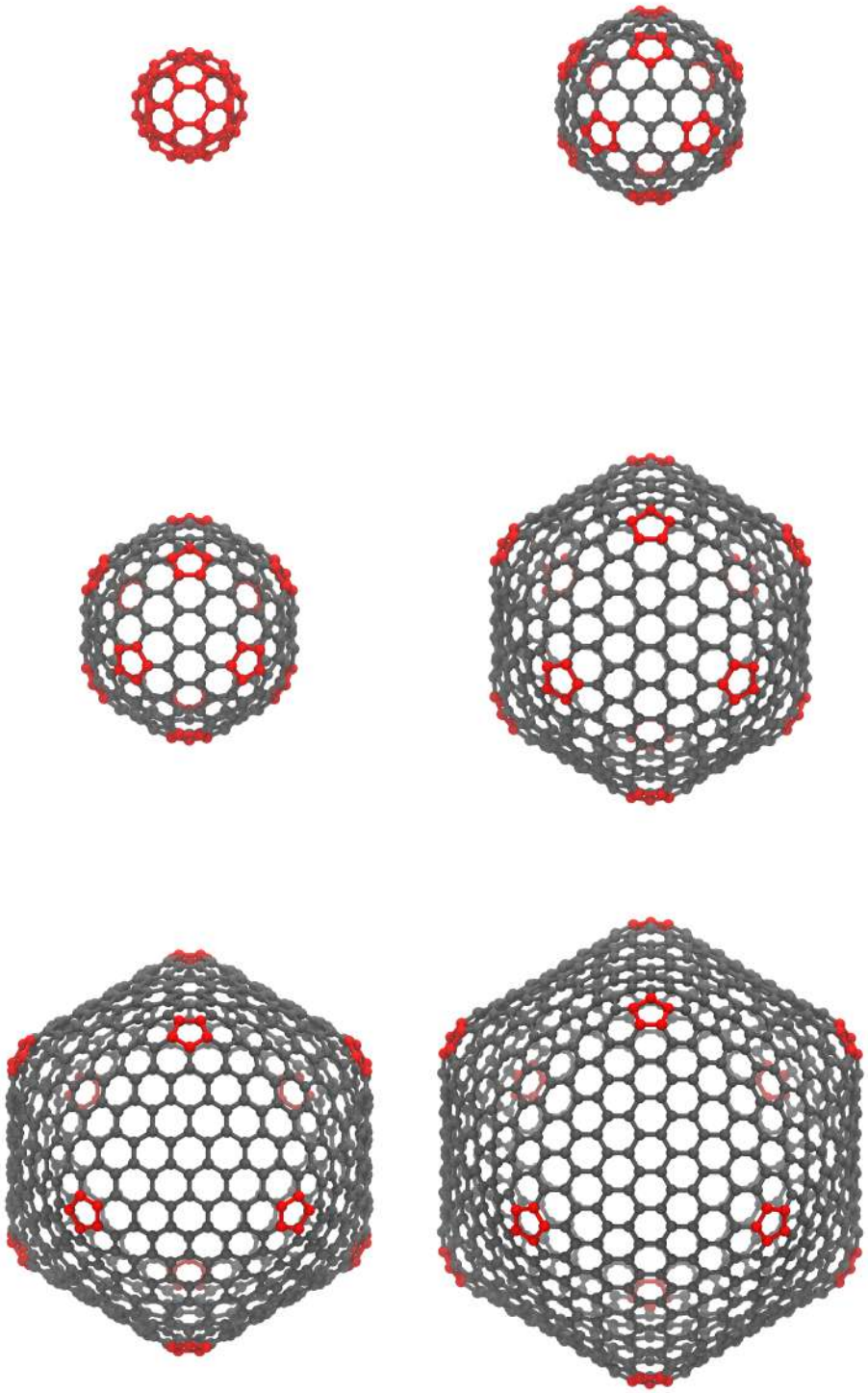


Figure 5.9: Optimized structures of the calculated icosahedral fullerenes. Geometries were optimized at the VWN/DZVP/GEN-A2 level of theory without any simplification due to symmetry. The view is aligned along one of the C_3 axes and all twelve five-membered rings are colored in red to facilitate the identification of the vertices of the icosahedron.

Table 5.6: Static average polarizability [a.u.] of icosahedral fullerenes with different methodologies. The Parriser-Parr-Pole (PPP) results were taken from reference [347]. Classical Penn model results were taken from reference [348]. The sum-over-states results corrected by random phase approximation (SOS-RPA) were taken from reference [349]. PBE0 refers to the hybrid PBE0/SVPD results from reference [350]. The last two columns are results obtained with the ADPT VWN/DZVP/GEN-A2 methodology solving the response equations with the direct [243,351] and iterative solver, respectively.

Molecule	PPP	Penn	SOS-RPA	PBE0	Direct	Iterative
C_{60}	526	434	533	551	529	528
C_{180}	1404		2025		2014	2014
C_{240}	2018	2317	2915	2018	2902	2902
C_{540}	6222	6929	7794	8395	8464	8466
C_{720}	9494		12471	12342		12489
C_{960}		14747	18524			18538

ities, except for C_{240}, where PBE0/SVPD severely underestimates the average polarizability. This underestimation led Rappoport and Furche to state that there is a minimum in the average polarizability per carbon atom in the icosahedral fullerene series.[350] Figure 5.10 shows that no such minimum is found by ADPT, neither with the direct nor the iterative solvers. This agrees with the SOS-RPA results reported in reference [349]. Both PBE0/SVPD and VWN/DZVP/GEN-A2 specific polarizabilities (per atom) exhibit a negative curvature possibly indicating a saturation for larger icosahedral fullerenes. In order to obtain the asymptotic polarizability per atom, we fitted the ADPT data to an asymptotic model, yielding

$$\alpha/N = 7.77562 + 16.5394\left(1 - e^{-0.0012237\times N}\right) \tag{5.2}$$

as the best fit. Using Equation (5.2) we get an asymptotic limit of 24.3 a.u. per carbon atom. If the same trend is followed by larger fullerenes, the C_{3840} would attain an specific polarizability only 0.1 a.u. smaller than the asymptotic limit. It is important to mention that Gueorguiev *et al.* already suggested that quantum size effects should be present, at least, until C_{3840}.[348] This statement was based on the ratio of the static polarizability calculated with a quantum mechanical approach and that of a classical model.

Dynamical polarizabilities are much less studied than their static counterparts. Here we compare calculated and experimental dynamical polarizabilities of C_{60}. Furthermore, we present as a showcase application dynamical polarizabilities of the two largest fullerees,

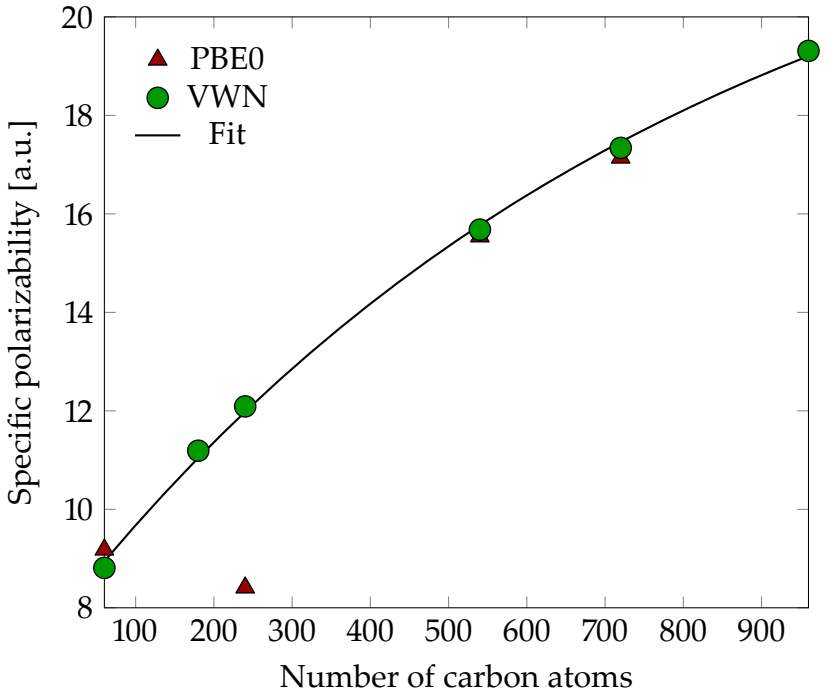


Figure 5.10: Specific polarizability (per carbon atom) of icosahedral fullerenes from C_{60} to C_{960}. PBE0 results are taken from reference [350]. VWN results were obtained with the new ADPT iterative solver. The solid line was obtained by fitting the VWN data to an asymptotic model. See text for more details.

C_{720} and C_{960}. Figure 5.11 shows the excellent agreement of ADPT dynamical polarizability for C_{60} compared to available experimental data.[353–356] This is inline with previous reports of our group concerning the accuracy of ADPT dynamical polarizabilities.[243,351]

The specific (per carbon atom) dynamical polarizabilities of C_{720} and C_{960} are shown in Figure 5.12. To the best of our knowledge, this is the first report of all-electron dynamical polarizabilities of such giant fullerenes. Note that the range for the y-axis in Figure 5.12 is so large because the most energetic wavelength used, 1047 nm, is close to the first excitation pole of C_{720} and is behind the first excitation pole of C_{960}. This explains the jump of the specific polarizability of C_{720} to very large values and the sharp reduction of the specific polarizability of C_{960}.

In summary, we provide ADPT static polarizabilities for icosahedral fullerenes from C_{60} to C_{960}. Our results follow closely those reported for the SOS-RPA method[349] and do not show the minimum in the specific polarizability for C_{240} previously reported in reference [350]. In fact, the specific polarizability of giant fullerenes grows monotonically, although with a negative curvature, suggesting a saturation for larger systems. Fitting the

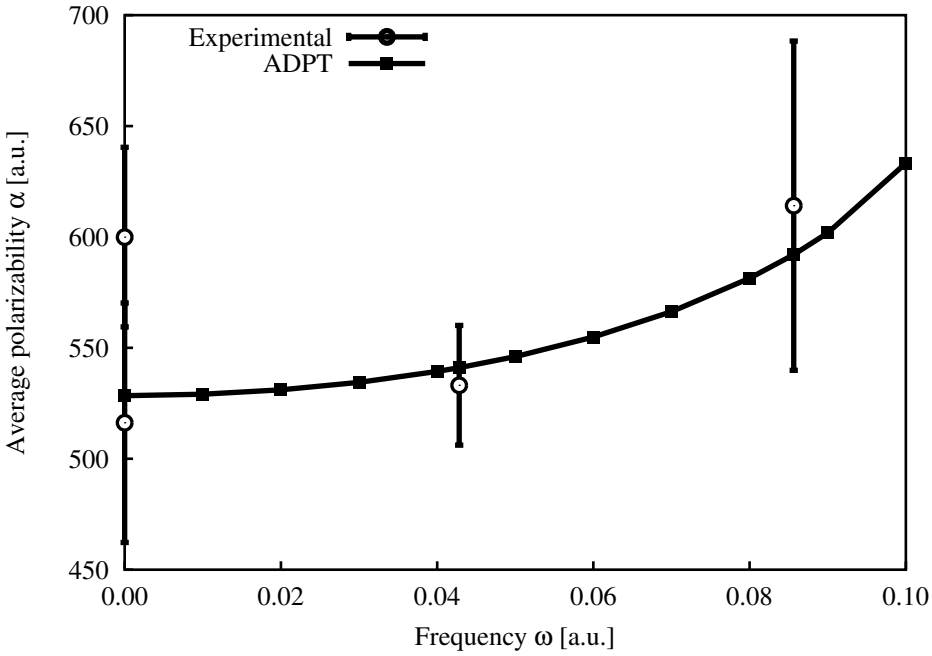


Figure 5.11: Calculated and experimental C_{60} dynamical polarizabilities. The average dynamical polarizabilities were calculated at the VWN/GEN-A2/DZVP level of theory.

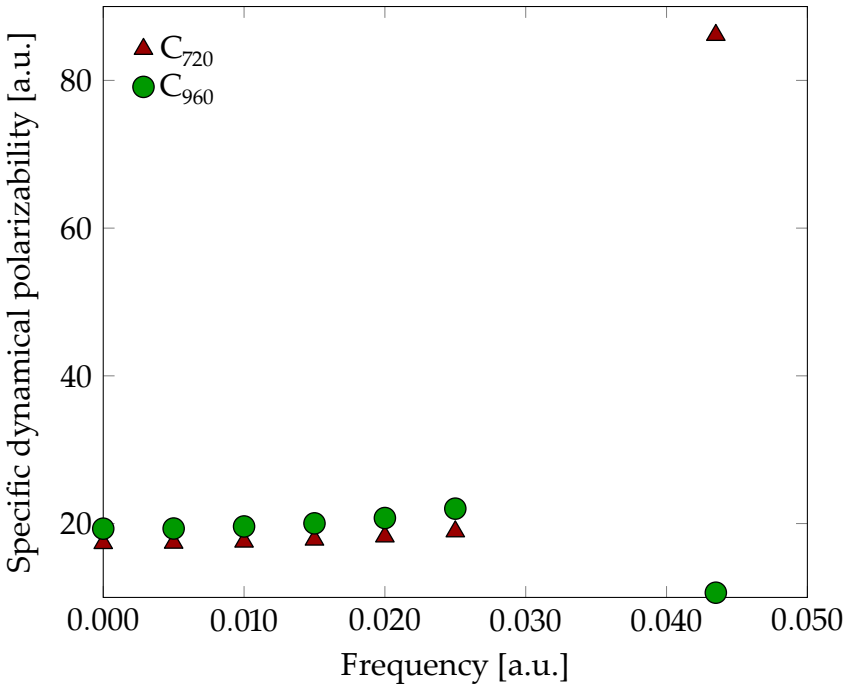


Figure 5.12: Specific dynamical polarizabilities of C_{720} and C_{960} calculated at the VWN/DZVP/GEN-A2 level of theory.

calculated ADPT static polarizabilities to an asymptotic model predicts that saturation is achieved already for C_{3840} with a polarizability close to 24.2 a.u. per carbon atom. Furthermore, C_{60} ADPT dynamical polarizabilities are in excellent agreement with experimental ones. Thus, the EN iterative solver does not alter the ADPT accuracy for the calculated dynamical polarizabilities.[243,351] Finally, the first all-electron dynamical polarizability calculations for C_{720} and C_{960} were presented. The dynamical polarizabilities were obtained for six different laser wavelengths, one of which is very close to the first excitation pole of C_{720} and is beyond the first pole of C_{960}. It is important to note that these results were obtained without the inclusion of simplifications due to the icosahedral symmetry of the systems. As a consequence, ADPT can now be used to predict molecular response properties for nanosystems with many thousands of basis and auxiliary functions.

6

Conclusions and Perspectives

6.1 Conclusions

Quantum mechanics has enabled the computation and prediction of many chemical properties of interest. However, the simulation of "real life" chemical problems is still out of our reach. Therefore, the main objective of this work was to develop new approximate methods for applying quantum mechanics to more complex atomic and molecular systems at a first-principle level of theory.

In this respect, we have shown the development and application of such approximate methods for two typical quantum chemical applications. On one hand, we developed a low-scaling methodology for the computation of exact exchange, which we called local density fitting exact exchange (LDF-EXX)[74,155,156] that can be used in molecular energy

calculations and structure optimizations. On the other hand, we developed a low-scaling algorithm for the solution of the response equation system arising in auxiliary density perturbation theory (ADPT)[242,243] for the calculation of molecular response properties. The resulting methodology is based on the Eirola-Nevanlinna[261] algorithm and has proven to be robust and efficient.[255]

The LDF-EXX methodology extends and improves previous efforts to reduce the computational complexity for calculating exact exchange.[104,139–154] Ultimately, this development resulted in the implementation of hybrid functionals in deMon2k[156] and opened the possibility to extend deMon2k's capabilities beyond density functional theory (DFT). The performance achieved by LDF-EXX is based on an efficient molecular orbital localization algorithm and on the use of very simple auxiliary function sets. To this end, we have substituted the standard Jacobi-sweep-based localization proposed many years ago by Edminston and Ruedenberg[160] by the more efficient approximate joint diagonalization (AJD) approach.[168–170] The usefulness of the new AJD-based localization approach is not limited to LDF-EXX but can be used in other quantum chemical methodologies too, like the Individual Gauge for Localized Orbital approach for the computation of nuclear magnetic resonance shielding tensors.[357] Different to other approximate exact exchange algorithms, LDF-EXX was validated with the relatively simple GEN-A*n* and GEN-A*n** auxiliary function sets.[201] These auxiliary function sets yielded 1 kcal/mol accuracies for heats of formation and chemical reaction barrier heights. Furthermore, the GEN-A*n* set showed remarkable speed-ups compared to standard electron repulsion integral calculations and also to conventional density fitting and resolution-of-the-identity implementations.

Although LDF-EXX achieves the above mentioned remarkable performance, this is only a minor step forward for the application of quantum mechanical methods to the mathematical description of "*…the whole of chemistry […].*" To this end, implementation of LDF-EXX on new architectures, like hybrid CPU-GPU computational clusters, can improve the size-range where LDF-EXX is applicable. However, this new implementation should solve some of the remaining issues with the here presented LDF-EXX implementation. Although the LDF-EXX can be used with success for non-insulators, like sodium and cop-

per clusters, the AJD localization algorithm struggles to optimize the objective functional. Moreover, the presented parallelization is far from being adequate for massive parallel high-performance computing environments.

The new ADPT iterative solver also shows remarkable performance. This advance is of great importance for the direct comparison between theory and experiment and the prediction of new material properties. With this approach, molecular property calculations of nanosystems can be performed in just a few hours.[255] Thus, systematic screening of molecular databases with a first-principle methodology becomes feasible. The main advantage of the ADPT iterative solver is that it can be straightforward extended for the computation of other molecular response properties, like static and dynamic hyperpolarizabilities[252] and Fukui functions,[358] to name a few. The ADPT iterative solver has also great potential to be used in the calculation of analytic second derivatives in ADFT, which are currently under development in our laboratory.

One important aspect of the ADPT iterative solver is that, besides lowering the computational complexity, it also diminishes the memory demand of the code. Therefore, the computations can be performed on modest and economical computational architectures. This low memory demand has also enabled the use of a mixed ADPT approach, which is an analog of the mixed self-consistent-field procedure developed in our group.[256] Furthermore, this diminished memory requirements can be useful for a GPU version of deMon2k, too.

Both of these low-order scaling methods were applied to chemical problems of current interest. First, we studied the effect of EXX on the electronic structure of sodium and copper clusters. We established the subshell closing of a spin-manifold as an important stabilizing mechanism for high-spin clusters. In particular, high-spin configurations of icosahedral Na_{13}, Na_{45} and Na_{55}^{+} clusters were found to be remarkably stable. For Na_{13} it was found that the 1D cluster shell was closed by the α-manifold. The PBE results predicts the sextet Na_{13} to be only 3 kcal/mol less stable than the ground state. Furthermore, inclusion of EXX, via the PBE0 functional, reduces this energy difference further, making the sextet only 1 kcal/mol less stable than the ground state. A similar situation occurs for

Na_{45}. In this case, the predicted electronic configuration,

$$1S^2\ 1P^6\ 1D^{10}\ 2S^2\ 1F^{14}\ 2P^6\ 1G^5\ ,$$

allowed an stable sextet configuration due to the splitting of the 1G shell into h_g and g_g subshells with five- and four-fold degeneracies. Thus, the h_g subshell was also closed by the α-manifold. The PBE functional predicts this sextet to be 5 kcal/mol less stable than the ground state. Again, EXX stabilized the sextet and predicts it to be only 0.5 kcal/mol less stable than the doublet. Both of the sextet neutral cluster, Na_{13} and Na_{45}, are predicted to be sufficiently stable to be detected by Stern-Gerlach-type experiments. Unfortunately, no reference to an experimental result for this cluster sizes was found in the literature.

The Na_{55}^+ cluster is a special case. Both PBE and PBE0 predict that the high-spin quintet is the ground state. In this case, inclusion of EXX added an extra 0.4 kcal/mol stabilization. The stability of the quintet is explained by its electronic configuration,

$$1S^2\ 1P^6\ 1D^{10}\ 2S^2\ 1F^{14}\ 2P^6\ 1G^{14}\ ,$$

that indicates the complete closure of the h_g subshell by both, α- and β-spin, manifolds and the closure of the g_g subshell by the α-manifold. The calculated density of states of the quintet Na_{55}^+ cluster shows a characteristic shoulder, due to the g_g subshell closing by only one spin manifold. This shoulder is also seen in the experimentally measured photoelectron spectrum of the Na_{55}^+ cluster,[304] supporting our finding of a quintet ground state. The proposed subshell closing mechanism could provide an explanation as to why the melting temperature of the Na_{55}^+ cation is abnormally high[301,302]. Further studies in our group confirmed that the subshell closing mechanism indeed affects the dynamical behavior of Na_{55}^+ and predicts the melting temperature to be above 300 K.[298]

The subshell closing mechanism was also studied in icosahedral copper clusters. It was shown that also for these clusters it can be an important stabilizing mechanism. Different to the sodium case, the Cu_{13} and Cu_{45} sextets are far less stable as compared to their

corresponding low-spin ground states. However, the Cu_{55}^{+} quintet was also found to be the ground state for copper. This suggests that subshell closing may be an important mechanism for simple metal cluster stabilization, independent of the nature of the cluster. Indeed, this could lead to "real life" applications. Unlike sodium clusters, copper clusters could provide the first stable magnetic cluster to be applied in the emerging field of spintronics.

We also studied the effect of exact exchange and basis set size on the quality of the description of transition metal complexes. To this end, we used the PBE and PBE0 functionals with the augmented correlation consistent basis sets of Dunning. We found that PBE results improve systematically with increasing basis set size. However, this was not the case for PBE0, were we found a poorer description of the bond lengths using the aug-cc-pVTZ than with the aug-cc-pVDZ basis. We also encountered some other issues due to the inclusion of EXX for the description of these complexes. In particular, C−H, C−O and N−O bond lengths became shorter for all basis sets. In addition, the spin-state splittings can be qualitatively changed by hybrid functionals. Specifically, the PBE0 functional overstabilizes high-spin configurations and incorrectly predicts a sextet ground state for manganocene at 0 K. Therefore, we advocate the use of "pure" DFT functionals for calculation of the transition metal compounds, even if only one transition metal atom is in the system.

Finally, the static and dynamic polarizabilities of giant icosahedral fullerenes were also studied. We found that the specific polarizability (per carbon atom) increases monotonically with the system size and that a previously reported minimum for C_{240} does not exist.[350] We also predicted that, starting with the C_{3840} icosahedral fullerene, the specific polarizability will achieve its asymptotic maximum at around 24.3 a.u. per carbon atom. Interestingly, it was previously predicted that quantum-size effects should be observed, at least, until C_{3840}.[348] We also presented, to the best of our knowledge, the first all-electron dynamical polarizability calculation for C_{720} and C_{960}. We predict that the usual laser used for this measurements, 1047 nm, will provide an energy very close to the first excitation pole of both C_{720} and C_{960}. For the C_{720} this energy will be below the first pole, provoking

a huge increase in the specific polarizability. On the other hand, the laser energy will be above the first excitation pole of C_{960} and a steep decrease in the specific polarizability should be observed.

Vast improvements are still needed for real chemical simulations. Both LDF-EXX and the new ADPT iterative solver contribute to this ultimate goal. The advancement made is just a step of an interstellar journey. Nevertheless, I think that quantum chemical calculations will, step by step, achieve Dirac's vision and provide answers to the real and complex problems as described by Mulliken and Roothaan.[1] In other words, we will build, step by step, *the snare that will some day catch the sun:*

> *Through all the ages of history there were men to whom this whisper had come of hidden things about them. They could no longer lead ordinary lives nor content themselves with the common things of this world once they had heard this voice. And mostly they believed not only that all this world was as it were a painted curtain before things unguessed at, but that these secrets were Power. Hitherto Power had come to men by chance, but now there were these seekers seeking, seeking among rare and curious and perplexing objects, sometimes finding some odd utilisable thing, sometimes deceiving themselves with fancied discovery, sometimes pretending to find. The world of every day laughed at these eccentric beings, or found them annoying and ill-treated them, or was seized with fear and made saints and sorcerers and warlocks of them, or with covetousness and entertained them hopefully; but for the greater part heeded them not at all. Yet they were of the blood of him who had first dreamt of attacking the mammoth; every one of them was of his blood and descent; and the thing they sought, all unwittingly, was the snare that will some day catch the sun.*
>
> (*The World Set Free, H. G. Wells*)

6.2.1 Eirola-Nevanlinna Coulomb fitting

The partition of the electron repulsion integrals (ERIs) into near-field and far-field has led to the proposal of the mixed SCF scheme.[256] In this approach, all near-field ERIs are stored in memory and only the far-field ERIs are recomputed in each SCF step employing fast asymptotic expansions.[109] The mixed SCF scheme solves the ERI bottleneck of ADFT. New computational bottlenecks then appear. On one hand, the numerical integration of the exchange-correlation potential, even though performed with the linear expanded auxiliary density, appears again as a time consuming step. On the other, linear algebra tasks are more and more dominating the CPU time consumption of ADFT calculations. Whereas the numerical integration of the exchange-correlation potential is already addressed by a more aggressive grid point screening, little has been done so far for the linear algebra tasks beyond the use of highly optimized external linear algebra libraries. In particular, matrix diagonalization consumes already for medium sized systems a large portion of the computational time in ADFT calculations.[156]

The largest matrix that has currently to be diagonalized is the Coulomb matrix $\mathbf{G}$. This diagonalization, used for the singular value decomposition (SVD) of $\mathbf{G}$, can be completely avoided if a sufficiently robust numerical solver is employed. The Eirola-Nevanlinna (EN) solver, applied to the ADPT response equation system, has shown to yield results in perfect agreement to those obtained by the SVD of the response matrix. Moreover, excellent convergence has been observed with the $\mathbf{G}^{-1}$ start guess. Therefore, we propose the use of the EN algorithm for the solution of the fitting equation system

$$\mathbf{G}\mathbf{x} = \mathbf{J} \tag{6.1}$$

The guess for the inverse of $\mathbf{G}$ can either be the identity matrix or can be obtained through a low-cost approach, like the approximate inverse (AINV) from Benzi and Tuma.[177–180] AINV is based on the pivoted inverse Cholesky decomposition of the matrix. It is very

efficient and can be easily parallelized. If the EN algorithm proves to be robust enough, an SVD-free deMon2k version can be achieved. This will reduce the linear algebra bottleneck of current ADFT calculations in deMon2k.

6.2.2 Auxiliary Density Matrix Method

As concluded in section 3.4.3 the LDF-EXX will represent the computational bottleneck in ADFT hybrid calculations. To circumvent this bottleneck we propose a new algorithm that combines LDF-EXX with the Auxiliary Density Matrix Method (ADMM)[217,218] from the literature. Note that ADMM is only intended for hybrid DFT calculations. In ADMM the exact exchange energy, $K(\mathbf{P})$, is split into two parts. The first part, $k(\mathbf{p})$, represents the exact exchange energy calculated with a smaller secondary basis set. The second part, $X(\mathbf{P}) - x(\mathbf{p})$, is a correction term evaluated by DFT exchange energies in the primary and secondary basis set, respectively. Thus, the ADMM exact exchange energy contribution to a hybrid functional is given by:

$$K(\mathbf{P}) = k(\mathbf{p}) + X(\mathbf{P}) - x(\mathbf{p}) \tag{6.2}$$

The here appearing density matrices, $\mathbf{P}$ and $\mathbf{p}$, are calculated from the primary and secondary basis set, respectively. In fact, $\mathbf{p}$ is obtained by projection of $\mathbf{P}$ on the smaller secondary basis set. In the original ADMM implementations[217,218] only projection of the basis set is possible. However, when combining ADMM with LDF-EXX two possible projection choices arise. Either the primary basis set or the primary auxiliary function set can be projected. The former choice provides the greatest reduction of the computational scaling for the computation of the three-center ERIs. The latter will provide the greatest reduction for the calculation and inversion of all charge density ERI matrices $\mathbf{G}_i$ needed in LDF-EXX. Because the three-center ERI computation and transformation is the bottleneck of LDF-EXX calculations, projection of the primary basis set will result in the computationally most efficient algorithm. Therefore, we suggest this approach for the combination of LDF-EXX with ADMM.

In order to project the primary basis set, we follow the ADMM2 approximation of Guidon *et al.*[217] In ADMM2 the projection is based on a least-squares fitting of the projected MOs, $\check{\psi}_i(\mathbf{r})$, obtained by minimizing

$$W_2 = \sum_i^{\text{occ}} \langle\, \psi_i(\mathbf{r}) - \check{\psi}_i(\mathbf{r}) \,|\, \psi_i(\mathbf{r}) - \check{\psi}_i(\mathbf{r}) \,\rangle \tag{6.3}$$

with respect to the projected MO coefficients. Expansion of the expectation value in Equation (6.3) yields

$$W_2 = \sum_i^{\text{occ}} \langle\, \psi_i \,|\, \psi_i \,\rangle - 2\langle\, \psi_i \,|\, \check{\psi}_i \,\rangle + \langle\, \check{\psi}_i \,|\, \check{\psi}_i \,\rangle \tag{6.4}$$

Equation (6.4) can be rewritten in matrix notation as:

$$W_2 = \mathbf{C}^T\mathbf{S}\mathbf{C} - 2\mathbf{c}^T\mathbf{Q}\mathbf{C} + \mathbf{c}^T\mathbf{s}\mathbf{c} \tag{6.5}$$

where the columns of $\mathbf{C}$ and $\mathbf{c}$ contain the MO coefficients of the primary and secondary basis, respectively. The corresponding overlap matrices for these basis sets are denoted by $\mathbf{S}$ and $\mathbf{s}$. The matrix $\mathbf{Q}$ is the mixed overlap matrix between the secondary and primary basis sets. Variation of Equation (6.5) with respect to the MO coefficients of the secondary basis yields

$$\frac{\partial W_2}{\partial c_{\alpha i}} = -2\sum_\mu Q_{\alpha\mu}C_{\mu i} + 2\sum_\beta s_{\alpha\beta}c_{\beta i} \equiv 0 \quad \forall \quad \alpha \;\wedge\; i \leqslant N_{\text{occ}} \tag{6.6}$$

where α and β label AOs of the secondary basis. From Equation (6.6) follows:

$$c_{\beta i} = \sum_\alpha \sum_\mu s^{-1}_{\beta\alpha} Q_{\alpha\mu} C_{\mu i} \quad \forall \quad \beta \;\wedge\; i \leqslant N_{\text{occ}} \tag{6.7}$$

We can recast this set of equations into matrix notation as

$$\mathbf{c} = \mathbf{s}^{-1}\mathbf{Q}\mathbf{C} = \mathbf{T}\mathbf{C} \tag{6.8}$$

where $\mathbf{T} = \mathbf{s}^{-1}\mathbf{Q}$ is the transformation matrix from the primary MOs to the secondary MOs. The projected density matrix is obtained using the projected MO coefficients of Equation (6.8):

$$\mathbf{p} = \mathbf{c}\mathbf{c}^T = \mathbf{T}\mathbf{P}\mathbf{T}^T \tag{6.9}$$

The ADMM-EXX energy, K, is computed from $\mathbf{P}$ and $\mathbf{p}$ according to Equation (6.2). Different projection strategies that preserve either the normalization or the orthonormalization of the MOs have also been proposed,[217,218] but we will focus here on the above described unconstrained projection.

The corresponding contribution of the ADMM-EXX to an element of the Kohn-Sham matrix is given by:

$$K_{\mu\nu} = \frac{\partial X}{\partial P_{\mu\nu}} + \sum_{\alpha,\beta} \left(\frac{\partial k}{\partial p_{\alpha\beta}} - \frac{\partial x}{\partial p_{\alpha\beta}} \right) \frac{\partial p_{\alpha\beta}}{\partial P_{\mu\nu}} \tag{6.10}$$

The first term of the rhs of Equation (6.10) is just the contribution to the Kohn-Sham matrix from the chosen correction exchange functional evaluated with the primary density. The first term inside the parenthesis is an element of the LDF-EXX contribution to the Kohn-Sham matrix, Equation (3.52), evaluated with the secondary basis, i.e.

$$\mathbf{k} = \sum_{i}^{\text{occ}} \mathbf{e}_i \mathbf{Z}_i \mathbf{Z}_i^T \mathbf{e}_i^T \tag{6.11}$$

where an element of the ERI matrix $\mathbf{e}_i$ is given by:

$$e_{\alpha\bar{k},i} = \langle \alpha i \| \bar{k} \rangle \tag{6.12}$$

The second term inside the parenthesis is the contribution to the Kohn-Sham matrix from the exchange DFA evaluated with the secondary density calculated from the projected density matrix $\mathbf{p}$. Finally, the partial derivative of an element of the projected density matrix with respect to an element of the primary density matrix can be easily obtained from Equation (6.9) as

$$\frac{\partial p_{\alpha\beta}}{\partial P_{\mu\nu}} = T_{\alpha\mu} T_{\beta\nu} \tag{6.13}$$

Thus, the ADMM-LDF-EXX contribution to the Kohn-Sham matrix is given by:

$$\mathbf{K} = \mathbf{X} + \mathbf{T}^T (\mathbf{k} - \mathbf{x}) \mathbf{T} \tag{6.14}$$

Here $\mathbf{X}$ and $\mathbf{x}$ are the contribution of the mapping exchange potential evaluated with the primary and secondary densities, respectively, and $\mathbf{k}$ is the exact exchange potential evaluated with the secondary basis. Note that the auxiliary density, $\tilde{\rho}(\mathbf{r})$, has yet not been introduced. Therefore, Equations (6.2)-(6.14) constitute the working equations for the ADMM-LDF-EXX approach within the DF-DFT method. We have previously seen that LDF-EXX HF calculations are as fast as DF-DFT GGA ones. As a consequence, ADMM-LDF-EXX might be even counterproductive, i.e. slowing down DF-DFT hybrid computations because of the two additional numerical integrations needed! This situation can be alleviated with the ADFT methodology, where the numerical calculation of the exchange functionals are performed with the auxiliary density and, therefore, are far less demanding.

The ADMM-LDF-EXX Approach within ADFT

The above presented ADMM-LDF-EXX approach can be straightforward incorporated into the ADFT methodology. The main difference to conventional ADFT arises from the fact that a secondary auxiliary density, $\check{\rho}(\mathbf{r})$, must be obtained by the variational fitting of the secondary density. Thus, a new fitting equation system,

$$\mathbf{G}\check{\mathbf{x}} = \check{\mathbf{J}} \ , \tag{6.15}$$

must be solved. Note that the same auxiliary function set is used for the primary and secondary variational density fitting. In Equation (6.15), the Coulomb vector $\check{\mathbf{J}}$ is defined

as

$$\check{\mathbf{J}} = \begin{pmatrix} \sum_{\alpha,\beta} \langle \bar{1} \| \alpha\beta \rangle p_{\alpha\beta} \\ \sum_{\alpha,\beta} \langle \bar{2} \| \alpha\beta \rangle p_{\alpha\beta} \\ \vdots \\ \sum_{\alpha,\beta} \langle \bar{m} \| \alpha\beta \rangle p_{\alpha\beta} \end{pmatrix} \tag{6.16}$$

Once Equation (6.15) has been solved, the computation of $x(\mathbf{p})$ is carried out according to:

$$x(\mathbf{p}) = E_{\mathrm{x}}[\check{\rho}] \tag{6.17}$$

Finally, the contribution of the ADMM-EXX to an element of the Kohn-Sham matrix is obtained as:

$$K_{\mu\nu} = \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle z_{\bar{k}} + \sum_{\alpha,\beta} T_{\alpha\mu} \left(k_{\alpha\beta} - \sum_{\bar{k}} \langle \alpha\beta \| \bar{k} \rangle \check{z}_{\bar{k}} \right) T_{\beta\nu} \tag{6.18}$$

where

$$\check{z}_{\bar{k}} = \sum_{\bar{k},\bar{l}} G^{-1}_{\bar{k}\bar{l}} \langle\, \bar{l} \,|\, v_{\mathrm{xc}}[\check{\rho}] \,\rangle \tag{6.19}$$

When using ADFT with the ADMM-LDF-EXX approach, two additional ERI computations, building $\check{\mathbf{J}}$ and computing $\mathbf{K}$, and two additional numerical integrations must be performed. The ERI computations can be performed very efficiently by employing the double-asymptotic ERI expansion and an incremental building of the contributions. As already noted, the numerical integrations of DFAs are performed with the use of auxiliary densities and, therefore, are not too demanding.

Analytical ADFT Gradients for ADMM-LDF-EXX

The ADMM-LDF-EXX approach introduces four new contributions to ADFT gradients. First, the contribution from $X[\tilde{\rho}]$; second, the contribution from $x[\check{\rho}]$; third, the contribution from $k(\mathbf{p})$; and fourth, the contribution from the projection. The first contribution is obtained using Equation (3.66) without modifications. The second contribution, corre-

sponding to $x[\check{\rho}]$, is given by

$$\begin{aligned} E_{\mathrm{x}}[\check{\rho}]^{(\lambda)} &= \int \frac{\delta E_{\mathrm{x}}[\check{\rho}]}{\delta \check{\rho}(\mathbf{r})} \frac{\partial \check{\rho}(\mathbf{r})}{\partial \lambda} \mathrm{d}\mathbf{r} \\ &= \sum_{\bar{k}} \check{x}_{\bar{k}}^{(\lambda)} \langle v_{\mathrm{xc}}[\check{\rho}] \,|\, \bar{k} \rangle + \sum_{\bar{k}} \check{x}_{\bar{k}} \langle v_{\mathrm{xc}}[\check{\rho}] \,|\, \bar{k}^{(\lambda)} \rangle \end{aligned} \tag{6.20}$$

The derivatives of the secondary fitting coefficients, $\check{\mathbf{x}}$, can be obtained by deriving Equation (6.15), from where it follows

$$\begin{aligned} E_{\mathrm{x}}[\check{\rho}]^{(\lambda)} &= \sum_{\alpha,\beta} \sum_{\bar{k}} p_{\alpha\beta}^{(\lambda)} \langle \alpha\beta \| \bar{k} \rangle \check{z}_{\bar{k}} + \sum_{\alpha,\beta} \sum_{\bar{k}} p_{\alpha\beta} \langle \alpha\beta \| \bar{k} \rangle^{(\lambda)} \check{z}_{\bar{k}} - \\ &\quad \sum_{\bar{k},\bar{l}} \check{x}_{\bar{k}} G_{\bar{k}\bar{l}} \check{z}_{\bar{l}} + \sum_{\bar{k}} \check{x}_{\bar{k}} \langle v_{\mathrm{xc}}[\check{\rho}] \,|\, \bar{k}^{(\lambda)} \rangle \end{aligned} \tag{6.21}$$

which is equivalent to Equation (3.66) evaluated with the secondary basis and the secondary quantities $\mathbf{p}$, $\check{\mathbf{x}}$, and $\check{\mathbf{z}}$. For the third contribution, corresponding to $k(\mathbf{p})$, an equivalent to Equation (3.74) is obtained, namely

$$k^{(\lambda)} = \sum_{\alpha,\beta} p_{\alpha\beta}^{(\lambda)} k_{\alpha\beta} + \sum_{\bar{k},\bar{l}} \check{\Gamma}_{\bar{k}\bar{l}} G_{\bar{k}\bar{l}}^{(\lambda)} - \sum_{i}^{\mathrm{occ}} \sum_{\bar{k}} \sum_{\alpha,\beta} \sum_{\gamma} p_{\alpha\beta} \check{x}_{\bar{k}\alpha i} \langle \bar{k} \| \beta\gamma \rangle^{(\lambda)} c_{\gamma i} \tag{6.22}$$

Only $\mathbf{p}^{(\lambda)}$ remains to be evaluated. Deriving Equation (6.9) yields

$$\mathbf{p}^{(\lambda)} = \mathbf{T}^{(\lambda)} \mathbf{P} \mathbf{T}^{T} + \mathbf{T} \mathbf{P}^{(\lambda)} \mathbf{T}^{T} + \mathbf{T} \mathbf{P} \left(\mathbf{T}^{T} \right)^{(\lambda)} \tag{6.23}$$

The here appearing transformation matrix derivatives are given as:

$$\begin{aligned} \mathbf{T}^{(\lambda)} &= \left(\mathbf{s}^{-1} \right)^{(\lambda)} \mathbf{Q} + \mathbf{s}^{-1} \mathbf{Q}^{(\lambda)} \\ &= -\mathbf{s}^{-1} \mathbf{s}^{(\lambda)} \mathbf{s}^{-1} \mathbf{Q} + \mathbf{s}^{-1} \mathbf{Q}^{(\lambda)} \end{aligned} \tag{6.24}$$

To proceed we collect the terms involving $p_{\alpha\beta}^{(\lambda)}$ of Equations (6.21) and (6.22) and introduce

$$x_{\alpha\beta} \equiv \sum_{\bar{k}} \langle \alpha\beta \| \bar{k} \rangle \check{z}_{\bar{k}} \tag{6.25}$$

to ease the notation. The calculation of the other terms in Equations (6.21) and (6.22) is straightforward according to the discussion in section 3.3 and, therefore, will not be further discussed here. Thus, we find by substituting Equation (6.23) into the difference of these two terms:

$$\begin{aligned}\sum_{\alpha,\beta}\left(k_{\alpha\beta}-x_{\alpha\beta}\right)p_{\alpha\beta}^{(\lambda)} = \;& 2\;\sum_{\mu,\nu}\sum_{\alpha,\beta}\left(k_{\alpha\beta}-x_{\alpha\beta}\right)T_{\alpha\mu}^{(\lambda)}P_{\mu\nu}T_{\beta\nu} + \\ & \sum_{\mu,\nu}\sum_{\alpha,\beta}\left(k_{\alpha\beta}-x_{\alpha\beta}\right)T_{\alpha\mu}P_{\mu\nu}^{(\lambda)}T_{\beta\nu}\end{aligned} \tag{6.26}$$

The second term of the rhs of Equation (6.26) along with the first term of the rhs of Equation (3.66) can be immediately recognized as

$$\sum_{\mu,\nu}K_{\mu\nu}P_{\mu\nu}^{(\lambda)} \tag{6.27}$$

which can just be absorbed in the Pulay forces. The remaining first term of the rhs of Equation (6.26) can be evaluated by substituting Equation (6.24) into it, yielding

$$\begin{aligned}2\sum_{\mu,\nu}\sum_{\alpha,\beta}\left(k_{\alpha\beta}-x_{\alpha\beta}\right)T_{\alpha\mu}^{(\lambda)}P_{\mu\nu}T_{\beta\nu} = \;& -2\sum_{\mu,\nu}\sum_{\alpha,\beta}\sum_{\gamma,\delta}\left(k_{\alpha\beta}-x_{\alpha\beta}\right)s_{\alpha\gamma}^{-1}s_{\gamma\delta}^{(\lambda)}T_{\delta\mu}P_{\mu\nu}T_{\beta\nu} + \\ & 2\sum_{\mu,\nu}\sum_{\alpha,\beta}\sum_{\gamma}\left(k_{\alpha\beta}-x_{\alpha\beta}\right)s_{\alpha\gamma}^{-1}Q_{\gamma\mu}^{(\lambda)}P_{\mu\nu}T_{\beta\nu}\end{aligned} \tag{6.28}$$

which can be rewritten as

$$2\sum_{\mu,\nu}\sum_{\alpha,\beta}\left(k_{\alpha\beta}-x_{\alpha\beta}\right)T_{\alpha\mu}^{(\lambda)}P_{\mu\nu}T_{\beta\nu} = -\sum_{\alpha,\beta}w_{\alpha\beta}s_{\alpha\beta}^{(\lambda)} + \sum_{\alpha}\sum_{\mu}\check{W}_{\alpha\mu}Q_{\alpha\mu}^{(\lambda)} \tag{6.29}$$

where we introduced the generalized energy-weighted density matrices

$$\mathbf{w} = 2\mathbf{s}^{-1}\left(\mathbf{k}-\mathbf{x}\right)\mathbf{p} \tag{6.30}$$

$$\check{\mathbf{W}} = 2\mathbf{s}^{-1}\left(\mathbf{k}-\mathbf{x}\right)\mathbf{TP} \tag{6.31}$$

In summary, only two subroutines must be added to the code. One that builds the generalized energy-weighted density matrices, $\mathbf{w}$ and $\breve{\mathbf{W}}$, and another one that contracts $\breve{\mathbf{W}}$ with the mixed-overlap derivatives $\mathbf{Q}^{(\lambda)}$. The remaining contributions can be obtained by calling the same subroutines programmed for the LDF-EXX gradient within ADFT, provided that they are fed with the appropriate quantities—primary or secondary.

ADMM mapping

After the general theory for ADFT ADMM-LDF-EXX calculations has been derived, the basis set mapping and the influence of the exchange correction functional was studied. To this end HF ΔH_f^{298K} for the full G3/99 molecular test set were computed and compared to the corresponding LDF-EXX results. Merlot *et al.* have already shown that there is no single exchange functional and secondary basis set combination that performs equally well for every primary basis set and molecule.[218] Thus, we decided to test the KT3 exchange functional[70] along with the 3-21G[359,360] secondary basis set because this combination was the best to approximate the B3LYP cc-pVTZ[216] results in reference [218]. In addition, we also probed the recently developed CAP exchange functional[71] because its potential decays asymptotically like the EXX potential. Finally the EV93 exchange functional[361] was also tested because it reproduces the optimized effective potential of noble gases. To investigate the influence of the secondary basis set we also used the DFT optimized DZVP basis.[100,269] The Def2-TZVPP[198] basis was always used as primary basis set along with the GEN-A2* auxiliary function set.

Individual ΔH_f^{298K} for the 223 molecules of the G3/99 test set are shown in Table B.2 of Appendix B. Table 6.1 lists a summary of these results in terms of MADs, MDs and MaxADs for the ADMM-LDF-EXX calculated ΔH_f^{298K} with respect to the corresponding LDF-EXX HF reference ones. The KT3 functional shows the best performance among the three exchange functional when the 3-21G secondary basis is used. However, it shows the largest deviations when DZVP is used as secondary basis. Of the other two exchange functionals, CAP shows the best performance for both secondary basis sets. This result was expected due to the asymptotic properties of the CAP potential which EV93 and KT3

Table 6.1: Deviations [kcal/mol] of ADMM-LDF-EXX HF standard heats of formation with respect to LDF-EXX HF ones. See text for details.

	CAP	EV93	KT3
3-21G			
MD	21.0	26.2	19.6
MAD	21.0	26.2	19.6
MaxAD	43.6	53.5	40.5
DZVP			
MD	−0.2	0.4	1.8
MAD	−0.2	0.4	1.8
MaxAD	−6.7	−6.9	−0.8

do not possess. It is also worth noting that the mapping compared HF and not hybrid calculations, as has been the usual case.[217,218] In this way, any effect of the ADMM approach will not be hidden by the small fraction of EXX generally mixed into hybrid DFAs. When using the ADMM approach for hybrid ADFT calculations, the deviations shown in Table 6.1 are expected to decrease, at least, by the same fraction as EXX is mixed into a given hybrid.

According to these results, the ADMM-LDF-EXX approach is a very promising methodology that can overcome the computational bottleneck of hybrid ADFT calculations. Furthermore, the CAP exchange functional is the best choice to perform the ADMM mapping. However, a thorough validation for several hybrid functionals has yet to be performed. This validation must include relative energy comparisons, such as ΔH_f^{298K} and chemical reaction barrier heights, as well as optimized geometry comparisons. This validation is very important to ensure that ADMM-LDF-EXX does not change the overall shape of the potential energy surface and, therefore, to establish if ADMM-LDF-EXX can be used for BOMD calculations. If ADMM-LDF-EXX is shown to be accurate enough, then it is also useful to benchmark its performance for one-, two- and three-dimensional systems. These benchmarks are very useful to tune the programming and achieve the best performance of the code. In this way, ADMM-LDF-EXX can be used for hybrid ADFT calculations with small impact on the very well-known ADFT performance.

The excellent performance achieved by our LDF-EXX algorithm encourages the development of post-Hartree-Fock methods based on the same philosophy. These so-called local correlation methods have been available since many years.[362–367] Many of them are based on the orbital invariant formulation of Møller-Plesset perturbation theory[368] proposed by Saebø and Pulay.[365,367] In particular, Werner, Schütz and coworkers have provided very significant contributions to the local correlation method. They have achieved linear or near linear scaling for LMP2,[369,370] LCCSD[371,372] and LCCSD(T),[373–375] where L means that the method is formulated in terms of localized orbitals. Furthermore, local density-fitting MP2 (LDF-MP2) have also been successfully applied by them.[376,377]

A different approach towards local electron correlation methods is based on the Laplace transform of the energy denominators.[378–387] It has also been extended for the density-fitting approach with localized molecular orbitals.[386,387] The Pulay and Saebø's methodology needs the solution of a linear equation system because the first-order amplitudes no longer decouple. In comparison, Almlöf's Laplace transform technique needs a numerical integration of the Laplace integral, although it has been shown that it requires only a few quadrature points.[385]

The LDF-MP2 envisioned for deMon2k is based on the same principles than the here presented LDF-EXX. Particularly important is the use of the GEN-A*n* auxiliary function sets to ensure the maximum performance attainable. In order to use localized molecular orbitals, it will rely on the elimination of the energy denominators by a Laplace transform, as suggested by Häser and Almlöf.[378,379] We also propose the implementation of the scaled opposite-spin (SOS) variant[386,387] because it is a more efficient approach to calculate the second-order correlation energies. In SOS-MP2, the second-order correction to the energy is given by:

$$E_{\text{SOS-MP2}} = 1.3 \sum_{i,a}^{\alpha} \sum_{j,b}^{\beta} \frac{\langle ia \| jb \rangle^2}{\Delta_{ij}^{ab}} = 1.3 E_{\text{OS}} \tag{6.32}$$

where i, j label occupied molecular orbitals (MOs), a, b label unoccupied MOs and

$$\Delta_{ij}^{ab} = \varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j \tag{6.33}$$

The spin-dependence of the opposite-spin contribution has been transferred to the sum limits. Despite of its simpler form, the SOS-MP2 version improves both the relative energies as well as the gradients.[387] Introducing the Laplace transform of the denominator yields:

$$E_{\mathrm{OS}} = \sum_{i,a}^{\alpha} \sum_{j,b}^{\beta} \langle ia \| jb \rangle^2 \int_0^{\infty} e^{-\Delta_{ij}^{ab} t} \, dt \tag{6.34}$$

where t is the Laplace variable. By solving the Laplace integral with a numerical quadrature, Equation (6.34) can be approximated as:

$$E_{\mathrm{OS}} \approx \sum_q w_q \sum_{i,a}^{\alpha} \sum_{j,b}^{\beta} \langle ia \| jb \rangle^2 e^{-\Delta_{ij}^{ab} t_q} \tag{6.35}$$

The orbital energies appearing in the exponential can be factored out and included directly into the definition of the MOs. In this way, each quadrature point will have its own set of MOs defined as:

$$\begin{aligned} \psi_I(\mathbf{r}) &= \psi_i(\mathbf{r}) e^{\varepsilon_i t_q / 2} \\ \psi_A(\mathbf{r}) &= \psi_a(\mathbf{r}) e^{-\varepsilon_a t_q / 2} \end{aligned} \tag{6.36}$$

The SOS-MP2 energy can now be written in terms of the newly transformed MOs as:

$$E_{\mathrm{OS}} \approx \sum_q w_q \sum_{I,A}^{\alpha} \sum_{J,B}^{\beta} \langle IA \| JB \rangle^2 \tag{6.37}$$

The advantage of using Equation (6.37) is that it is invariant under unitary transformation of the new set of MOs. As a consequence, localized MOs (LMOs) can be used and the orbital products (transition densities) can be restricted to only the spatially close ones. Furthermore, these orbital products can be variationally fitted to a linear combination of

auxiliary functions, exactly in the same way as it was done for the LDF-EXX. In this case, the transition densities are approximated as:

$$\rho_{IA}(\mathbf{r}) \approx \sum_{\bar{k}} x_{\bar{k}IA}\bar{k}(\mathbf{r}) \tag{6.38}$$

The fitting coefficients are obtained by the minimization of the error in the Coulomb metric, in an analog manner to Equation (3.14). Thus, we find for the fitting coefficients:

$$x_{\bar{k}IA} = \sum_{\bar{l}} \langle \bar{k}||\bar{l}\rangle^{-1}\langle \bar{l}||IA\rangle \tag{6.39}$$

As in LDF-EXX, the auxiliary functions can also be restricted to only those close in space to $\psi_I(\mathbf{r})$ and $\psi_A(\mathbf{r})$. Using Equations (6.38) and (6.39) we can rewrite the SOS-MP2 energy as:

$$E_{\mathrm{OS}} \approx \sum_{q} w_q \sum_{I,A}^{\alpha}\sum_{J,B}^{\beta}\sum_{\bar{k},\bar{l}} B_{\bar{k}}^{IA}B_{\bar{k}}^{JB}B_{\bar{l}}^{IA}B_{\bar{l}}^{JB} \tag{6.40}$$

with

$$B_{\bar{k}}^{IA} = \sum_{\bar{l}} \langle IA||\bar{l}\rangle \langle \bar{l}||\bar{k}\rangle^{-1/2} \tag{6.41}$$

In order to further simplify Equation (6.40) we introduce a new pair of matrices $\mathbf{X}^{\alpha}$ and $\mathbf{X}^{\beta}$ defined as:

$$X_{\bar{k}\bar{l}}^{\alpha} = \sum_{I,A}^{\alpha} B_{\bar{k}}^{IA}B_{\bar{l}}^{IA} \tag{6.42}$$

$$X_{\bar{k}\bar{l}}^{\beta} = \sum_{J,B}^{\beta} B_{\bar{k}}^{JB}B_{\bar{l}}^{JB} \tag{6.43}$$

Substituting the definition of the $\mathbf{X}$ matrices into Equation (6.40) yields

$$E_{\mathrm{OS}} \approx \sum_{q} w_q \sum_{\bar{k},\bar{l}} X_{\bar{k}\bar{l}}^{\alpha}X_{\bar{k}\bar{l}}^{\beta} \tag{6.44}$$

We thus see that the equations arising in our proposed SOS-MP2 methodology share some common features to both LDF-EXX and ADPT. The same algorithms for orbital localiza-

tion and space subdivision (octree) can boost the performance of this SOS-MP2 approach. Furthermore, the $\mathbf{X}$ matrices can be built with the same subroutines as the ones used to compute the Coulomb response matrix $\mathbf{A}$. As a result, we expect that a very efficient SOS-MP2 algorithm can be implemented in deMon2k. This algorithm can be accelerated by a GPU implementation similar to reference [388].

Other highly-correlated approaches may follow this MP2 proposal, leading to local coupled-clusters and configuration interaction algorithms. In addition, electron propagator methods can also be programmed with this philosophy in mind.[389,390]

6.2.4 Additional future work

In the following, we will enumerate some of the additional future works that could follow the here presented developments.

1. Introduction of Benzi and Tuma approximate inverse (AINV)[177–180] algorithm for the direct computation of the inverse Cholsesky factors in LDF-EXX.

2. Improvement of the LDF-EXX parallelization by employing the octree algorithm with more resolution. The parallelization approach may also be revisited and be completely changed from the current molecular orbital partition to an atomic orbital partition.

3. Development of a second-order orbital localization approach to find a more appropriate set of localized orbitals for both LDF-EXX and the above presented local SOS-MP2 approach.

4. Introduction of the range-separated Coulomb operator for the implementation of range-separated hybrid functionals.

5. Introduction of the Laplace transform for the $\frac{\omega^2}{\omega^2-\omega_{ia}^2}$ factor appearing in the ADPT iterative solver. In this way, the energy denominator disappears and the intermediate transformation to molecular orbital representation can be avoided. Furthermore, the

introduction of sparse matrix-matrix multiplication subroutines can render a truly linear-scaling APDT approach.

A

Negative definiteness of $\mathcal{E}_2^{\mathrm{EXX}}$

We wish to show that the exact exchange fitting functional

$$\mathcal{E}_2 = -\sum_{i,j}^{\mathrm{occ}} \iint \frac{\left[\rho_{ij}(\mathbf{r}_1) - \tilde{\rho}_{ij}(\mathbf{r}_1)\right]\left[\rho_{ij}(\mathbf{r}_2) - \tilde{\rho}_{ij}(\mathbf{r}_2)\right]}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|}\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{A.1}$$

is negative semidefinite. If we write $\rho_{ij}(\mathbf{r}_1) - \tilde{\rho}_{ij}(\mathbf{r}_1) = F_{ij}(\mathbf{r}_1)$, the fitting functional may be rewritten in the form

$$\mathcal{E}_2 = -\sum_{i,j}^{\mathrm{occ}} \iint \frac{F_{ij}(\mathbf{r}_1)F_{ij}(\mathbf{r}_2)}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|}\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{A.2}$$

We may now define a potential as

$$\phi_{ij}(\mathbf{r}_1) = \int \frac{F_{ij}(\mathbf{r}_2)}{|\,\mathbf{r}_1 - \mathbf{r}_2\,|}\mathrm{d}\mathbf{r}_2 \tag{A.3}$$

Taking the Laplacian, with respect to the coordinates $\mathbf{r}_1$, on both sides of Equation (A.3) yields:

$$\begin{aligned}\nabla^2_{\mathbf{r}_1}\phi_{ij}(\mathbf{r}_1) &= \int \frac{1}{|\,\mathbf{r}_1-\mathbf{r}_2\,|}\nabla^2_{\mathbf{r}_1}F_{ij}(\mathbf{r}_2)\mathrm{d}\mathbf{r}_2 + 2\int \boldsymbol{\nabla}_{\mathbf{r}_1}F_{ij}(\mathbf{r}_2)\cdot\boldsymbol{\nabla}_{\mathbf{r}_1}\frac{1}{|\,\mathbf{r}_1-\mathbf{r}_2\,|}\mathrm{d}\mathbf{r}_2 + \\ &\int F_{ij}(\mathbf{r}_2)\nabla^2_{\mathbf{r}_1}\frac{1}{|\,\mathbf{r}_1-\mathbf{r}_2\,|}\mathrm{d}\mathbf{r}_2 \end{aligned} \tag{A.4}$$

Since $F_{ij}(\mathbf{r}_2)$ does not depend on $\mathbf{r}_1$ the first two terms on the rhs of Equation (A.4) vanish. Thus, the only term remaining involves the Laplacian of the Coulomb operator. It is straightforward to show that

$$\nabla^2_{\mathbf{r}_1}\frac{1}{|\,\mathbf{r}_1-\mathbf{r}_2\,|} = 0 \;\;\forall\;\; \mathbf{r}_1 \neq \mathbf{r}_2\;. \tag{A.5}$$

However, this result does not hold for $\mathbf{r}_1 = \mathbf{r}_2$, as indicated in Equation (A.5), because the derivatives are undefined at these points. Instead, the Laplacian of the Coulomb operator at these singularities is given by[391]

$$\nabla^2_{\mathbf{r}_1}\frac{1}{|\,\mathbf{r}_1-\mathbf{r}_2\,|} = -4\pi \;\;\forall\;\; \mathbf{r}_1 = \mathbf{r}_2\;, \tag{A.6}$$

from where it follows

$$\nabla^2_{\mathbf{r}_1}\frac{1}{|\,\mathbf{r}_1-\mathbf{r}_2\,|} = -4\pi\delta(\mathbf{r}_1-\mathbf{r}_2) \tag{A.7}$$

Using Equation (A.7) we can transform Equation (A.4) into

$$\begin{aligned}\nabla^2_{\mathbf{r}_1}\phi_{ij}(\mathbf{r}_1) &= -4\pi\int F_{ij}(\mathbf{r}_2)\delta(\mathbf{r}_1-\mathbf{r}_2)\mathrm{d}\mathbf{r}_2 \\ &= -4\pi F_{ij}(\mathbf{r}_1)\end{aligned} \tag{A.8}$$

We may now use Equations (A.3) and (A.8) to rewrite the exact exchange fitting functional in the form

$$\mathcal{E}_2 = \frac{1}{4\pi}\sum_{i,j}^{\mathrm{occ}}\int \phi_{ij}(\mathbf{r})\nabla^2\phi_{ij}(\mathbf{r})\mathrm{d}\mathbf{r} \tag{A.9}$$

where we have dropped the subindeces for the coordinate of integration and derivation. We now use the vector identity[391]

$$\boldsymbol{\nabla} \cdot (\phi \boldsymbol{\nabla} \psi) = \phi \nabla^2 \psi + \boldsymbol{\nabla} \phi \cdot \boldsymbol{\nabla} \psi \tag{A.10}$$

which holds for any pair of functions. Hence, we are left with

$$\mathcal{E}_2 = \frac{1}{4\pi} \sum_{i,j}^{\text{occ}} \int \boldsymbol{\nabla} \cdot \left[\phi_{ij}(\mathbf{r}) \boldsymbol{\nabla} \phi_{ij}(\mathbf{r}) \right] \mathrm{d}\mathbf{r} \; - \frac{1}{4\pi} \sum_{i,j}^{\text{occ}} \int \boldsymbol{\nabla} \phi_{ij}(\mathbf{r}) \cdot \boldsymbol{\nabla} \phi_{ij}(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{A.11}$$

Applying the divergence theorem to the first term of the rhs of Equation (A.11) yields a surface integral over an enclosing infinite sphere:

$$\mathcal{E}_2 = \frac{1}{4\pi} \sum_{i,j}^{\text{occ}} \oint_S \left[\phi_{ij}(\mathbf{r}) \boldsymbol{\nabla} \phi_{ij}(\mathbf{r}) \right] \cdot \mathbf{n} \, \mathrm{d}S - \frac{1}{4\pi} \sum_{i,j}^{\text{occ}} \int \boldsymbol{\nabla} \phi_{ij}(\mathbf{r}) \cdot \boldsymbol{\nabla} \phi_{ij}(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{A.12}$$

Assuming that the potentials and their gradients will vanish at infinity, the surface integral will vanish, which results in

$$\mathcal{E}_2 = -\frac{1}{4\pi} \sum_{i,j}^{\text{occ}} \int |\boldsymbol{\nabla} \phi_{ij}(\mathbf{r})|^2 \mathrm{d}\mathbf{r} \tag{A.13}$$

Since the integral of a square is necessarily positive then the sum appearing in Equation (A.13) is positive definite. Thus, it follows that $\mathcal{E}_2$ is negative semidefinite, vanishing only when $\rho_{ij}(\mathbf{r}) = \tilde{\rho}_{ij}(\mathbf{r})$ for every point in space.

Note that the only assumption made is that $\phi_{ij}(\mathbf{r})$ must vanish at infinity, which holds for any Gaussian type orbital. This result was previously derived by Slater in Appendix 19 of Reference 392.

B

Validation Tables

Table B.1: Hartree-Fock standard heats of formation [kcal/mol] for the 223 neutral molecules of the G3/99 test set. B3LYP/6-31G(2df,p) optimized geometries and the Def2-TZVPP basis set in spherical representation were used in all calculations. A2*/A2* calculations were performed with the GEN-A2* auxiliary function set for both the SCF and the final energy calculation, whereas the A2/A2* were perfomed with the GEN-A2 auxiliary function set for the SCF and the GEN-A2* auxiliary function set for the final energy calculation.

Molecule	NWChem	A2*/A2*	A2/A2*	Expt.
LiH	57.7	57.7	57.6	33.3
^{2}BeH	82.0	82.0	82.0	81.7
^{2}CH	169.8	169.8	169.7	142.5
CH_2 ($^3B_1$)	129.0	129.0	129.0	93.7
CH_2 ($^1A_1$)	157.3	157.3	157.1	102.8
$^3CH_3$	99.4	99.4	99.4	35.0
CH_4	73.8	73.8	73.7	−17.9
^{3}NH	117.9	117.9	117.8	85.2
NH_2 ($^2B_1$)	109.9	109.9	109.8	45.1
NH_3	86.8	86.8	86.6	−11.0
^{2}OH	47.9	47.9	47.8	9.4
H_2O	19.6	19.6	19.5	−57.8
HF	−20.6	−20.6	−20.6	−65.1
SiH_2 ($^1A_1$)	106.3	104.6	105.8	65.2
SiH_2 ($^3B_1$)	110.5	108.8	110.0	86.2
$^2SiH_3$	92.9	91.2	92.2	47.9
SiH_4	73.5	71.8	73.1	8.2
$^2PH_2$	77.4	77.4	77.0	33.1
PH_3	70.7	70.7	70.1	1.3
H_2S	48.2	48.2	47.7	−4.9
HCl	7.5	7.5	7.3	−22.1
Li_2	72.3	72.3	72.3	51.6
LiF	−31.1	−31.1	−31.1	−80.1
HC≡CH	165.9	166.0	166.0	54.2
$H_2C{=}CH_2$	147.3	147.3	146.9	12.5
$H_3C{-}CH_3$	140.3	140.3	139.6	−20.1
CN	196.3	196.4	196.3	104.9
HCN	145.3	145.3	145.3	31.5
CO	58.9	59.0	58.8	−26.4
^{2}HCO	106.7	106.7	106.2	10.0
H_2CO	90.5	90.5	89.9	−26.0
CH_3OH	95.4	95.4	94.5	−48.0
N_2	114.0	114.1	113.9	0.0
H_2NNH_2	194.2	194.1	193.0	22.8
NO	121.4	121.6	120.6	21.6
$^3O_2$	87.5	87.5	87.1	0.0
H_2O_2	101.9	101.9	100.6	−32.5
F_2	73.8	73.8	72.5	0.0
CO_2	49.1	49.1	49.2	−94.1
Na_2	51.8	51.8	51.7	34.0
Si_2	185.4	182.0	185.4	139.9
P_2	113.8	113.8	114.0	34.3
$^3S_2$	82.0	82.1	80.9	30.7
Cl_2	39.9	40.0	36.7	0.0
NaCl	−15.2	−15.2	−15.1	−43.6
SiO	57.9	56.3	57.8	−24.6
CS	141.3	141.3	141.2	66.9
SO	72.6	72.7	71.7	1.2
ClO	83.4	83.4	81.4	24.2

Table B.1: (continued)

Molecule	NWChem	A2*/A2*	A2/A2*	Expt.
ClF	41.1	41.2	39.4	−13.2
Si_2H_6	131.5	128.1	128.7	19.1
CH_3Cl	78.2	78.2	76.7	−19.6
H_3C-SH	117.2	117.2	115.6	−5.5
HOCl	69.0	69.0	67.0	−17.8
SO_2	82.5	82.6	79.8	−71.0
BF_3	−163.1	−163.1	−164.0	−271.4
BCl_3	−18.2	−18.2	−22.7	−96.3
AlF_3	−181.4	−181.4	−181.3	−289.0
$AlCl_3$	−77.3	−77.2	−78.8	−139.7
CF_4	−60.5	−60.6	−63.4	−223.0
CCl_4	111.1	111.1	103.7	−22.9
OCS	93.7	93.7	93.8	−33.1
CS_2	141.5	141.6	141.9	28.0
F_2CO	9.1	9.0	6.7	−149.1
SiF_4	−241.1	−242.3	−241.3	−386.0
$SiCl_4$	−58.6	−59.8	−63.0	−158.4
NNO	204.6	204.6	204.7	19.6
ClNO	160.3	160.3	158.1	12.4
NF_3	136.8	136.7	134.0	−31.6
PF_3	−99.0	−98.9	−101.8	−229.1
O_3	236.7	236.7	234.4	34.1
F_2O	138.5	138.4	136.3	5.9
ClF_3	112.6	112.7	110.3	−38.0
$F_2C=CF_2$	51.4	52.1	46.3	−157.4
$Cl_2C=CCl_2$	174.4	174.8	164.9	−3.0
F_3C-CN	120.3	120.5	118.2	−118.4
$HC\equiv C-CH_3$	224.3	224.3	224.1	44.2
$H_2C=C=CH_2$	225.6	225.6	224.4	45.5
C_3H_4 (cyclopropene)	251.4	251.4	249.6	66.2
$H_2C=CH-CH_3$	209.6	209.6	208.3	4.8
C_3H_6 (cyclopropane)	219.9	219.9	218.2	12.7
$CH_3-CH_2-CH_3$	205.5	205.4	204.0	−25.0
C_4H_6 (Z-1,3-butadiene)	276.4	276.4	274.6	26.3
C_4H_6 (2-butyne)	284.4	284.4	283.7	34.8
C_4H_6 (methylenecyclopropane)	298.3	298.3	295.2	47.9
C_4H_6 (bicyclo[1.1.0]butane)	309.8	309.8	307.4	51.9
C_4H_6 (cyclobutene)	292.2	292.1	288.7	37.4
C_4H_8 (cyclobutane)	284.7	284.6	281.8	6.8
C_4H_8 (isobutene)	271.9	271.9	269.5	−4.0
C_4H_{10} (trans butane)	270.8	270.7	268.6	−30.0
C_4H_{10} (isobutane)	270.1	270.0	267.8	−32.1
C_5H_8 (spiropentane)	368.6	368.5	364.6	44.3
C_6H_6 (benzene)	361.8	361.8	357.5	19.7
CH_2F_2	19.0	18.9	17.6	−107.7
CHF_3	−21.6	−21.6	−23.6	−166.6
CH_2Cl_2	84.8	84.8	81.8	−22.8
$CHCl_3$	95.2	95.3	90.4	−24.7
CH_3NH_2	160.2	160.2	159.3	−5.5
CH_3CN	199.8	199.8	199.5	18.0
CH_3NO_2 (nitromethane)	242.6	242.5	239.2	−17.8

Table B.1: (continued)

Molecule	NWChem	A2*/A2*	A2/A2*	Expt.
CH_3ONO (methyl nitrite)	240.9	240.8	237.7	−15.9
CH_3SiH_3	130.5	128.8	129.5	−7.0
HCO_2H	79.4	79.4	78.0	−90.5
HCO_2CH_3	152.5	152.5	150.3	−85.0
CH_3CONH_2	204.1	204.0	201.3	−57.0
C_2H_5N (aziridine)	242.6	242.5	240.7	30.2
NCCN (cyanogen)	284.0	284.1	284.0	73.3
$NH(CH_3)_2$	230.6	230.5	228.9	−4.4
$CH_3CH_2NH_2$	223.6	223.5	221.9	−11.3
$H_2C{=}C{=}O$ (ketene)	150.2	150.2	149.5	−11.4
C_2H_4O (oxirane)	177.9	177.8	176.0	−12.6
CH_3CHO	146.4	146.4	144.8	−39.7
$O{=}CH{-}CH{=}O$	161.5	161.5	159.0	−50.7
CH_3CH_2OH	157.1	157.1	155.5	−56.2
$(CH_3)_2O$	167.5	167.4	165.8	−44.0
C_2H_4S (thiooxirane)	190.3	190.3	187.6	19.6
$(CH_3)_2S{=}O$	218.4	218.4	213.7	−36.2
CH_3CH_2SH	182.2	182.2	179.8	−11.1
$(CH_3)_2S$	184.4	184.4	181.7	−8.9
$H_2C{=}CHF$	120.1	120.1	118.7	−33.2
CH_3CH_2Cl	141.3	141.3	139.1	−26.8
$H_2C{=}CHCl$	149.3	149.3	147.2	8.9
$H_2C{=}CHCN$	272.1	272.2	271.4	43.2
$(CH_3)_2C{=}O$	204.4	204.3	201.4	−51.9
CH_3CO_2H	136.0	135.9	133.2	−103.4
CH_3CFO	100.3	100.3	97.7	−105.7
CH_3COCl	139.6	139.6	136.2	−58.0
$CH_3CH_2CH_2Cl$	206.5	206.5	203.6	−31.5
$(CH_3)_2CHOH$	219.2	219.1	216.7	−65.2
$CH_3{-}O{-}CH_2CH_3$	229.3	229.3	227.0	−51.7
$(CH_3)_3N$	299.8	299.7	297.0	−5.7
C_4H_4O (furan)	275.3	275.2	271.5	−8.3
C_4H_4S (thiophene)	294.0	294.1	289.0	27.5
C_4H_5N (pyrrole)	331.5	331.5	327.7	25.9
C_5H_5N (pyridine)	379.1	379.1	374.4	33.6
H_2	25.8	25.8	25.8	0.0
^2SH	60.2	60.2	60.0	34.2
$^2C{\equiv}CH$	215.6	215.7	215.7	135.1
$HC{=}CH_2$ ($^2A'$)	178.5	178.6	178.1	71.6
$CH_3C{=}O$ ($^2A'$)	163.0	163.0	161.5	−2.4
$CH_2{-}OH$ (^2A)	115.7	115.7	114.9	−4.1
CH_3O ($^2A'$)	113.5	113.4	112.7	4.1
CH_3CH_2O ($^2A''$)	174.7	174.7	173.2	−3.7
CH_3S ($^2A'$)	125.9	125.9	124.6	29.8
CH_2CH_3 ($^2A'$)	163.4	163.4	162.7	28.9
$(CH_3)_2CH$ ($^2A'$)	226.5	226.4	225.1	21.5
$^2C(CH_3)_3$	289.5	289.4	287.1	12.3
NO_2 ($^2A_1$)	181.5	181.6	180.2	7.9
$CH_3CH{=}C{=}CH_2$	289.1	289.1	287.1	38.8
C_5H_8 (isoprene)	339.8	339.8	336.7	18.0
C_5H_{10} (cyclopentane twist)	331.2	331.1	327.6	−18.3

Table B.1: (continued)

Molecule	NWChem	A2*/A2*	A2/A2*	Expt.
C_5H_{12} (n-pentane)	336.1	336.1	333.3	−35.1
$C(CH_3)_4$ (neopentane)	335.3	335.2	331.9	−40.2
C_6H_8 (1,3-cyclohexadiene)	394.8	394.8	389.9	25.4
C_6H_8 (1,4-cyclohexadiene)	392.9	392.9	387.9	25.0
C_6H_{12} (cyclohexane chair)	390.8	390.7	386.5	−29.5
C_6H_{14} (n-hexane)	401.5	401.4	398.0	−39.9
C_6H_{14} (3-methylpentane)	403.5	403.4	399.9	−41.1
$C_6H_5{-}CH_3$ (toluene)	425.6	425.6	420.1	12.0
C_7H_{16} (n-heptane)	466.8	466.8	462.7	−44.9
C_8H_8 (cyclooctatetraene)	531.7	531.7	524.5	70.7
C_8H_{18} (n-octane)	532.1	532.1	527.3	−49.9
$C_{10}H_8$ (naphtalene)	587.7	587.7	578.2	35.9
$C_{10}H_8$ (azulene)	631.0	631.0	621.4	69.1
$CH_3CO_2CH_3$ (Z-methylacetate)	209.6	209.5	206.2	−98.4
$(CH_3)_3COH$ (t-butanol)	282.2	282.0	278.6	−74.7
$C_6H_5NH_2$ (aniline)	439.0	439.0	433.3	20.8
C_6H_5OH (phenol)	374.4	374.4	368.6	−23.0
C_4H_6O (divinyl ether)	299.4	299.4	296.7	−3.3
C_4H_8O (tetrahydrofuran)	285.7	285.6	282.2	−44.0
C_5H_8O (cyclopentanone)	327.6	327.5	322.5	−45.9
$C_6H_4O_2$ (benzoquinone)	393.0	393.0	384.6	−29.4
$C_6H_4N_2$ (pyrimidine)	392.8	392.7	387.9	46.8
$(CH_3)_2SO_2$	221.0	221.1	215.8	−89.2
C_6H_5Cl (chlorobenzene)	364.8	364.8	358.0	12.4
$NC(CH_2)_2CN$ (succinonitrile)	393.9	394.0	392.9	50.1
$C_4H_4N_2$ (pyrazine)	399.5	399.4	394.4	46.9
C_4H_4O (3-butyn-2-one)	296.0	296.1	294.0	15.6
C_4H_6O (E-crotonaldehyde)	276.0	276.0	273.1	−24.0
$C_4H_6O_3$ (acetic anhydride)	266.7	266.7	261.6	−136.8
C_4H_6S (2,5-dihydrothiophene)	306.7	306.7	301.4	20.8
$(CH_3)_2CHCN$	329.7	329.7	328.0	5.6
C_4H_8O (methyl ethyl ketone)	269.5	269.5	265.9	−57.1
$(CH_3)_2CHCHO$	277.0	276.9	273.8	−51.6
$C_4H_8O_2$ (1,4-dioxane)	305.4	305.3	300.8	−75.5
C_4H_8S (tetrahydrothiophene)	305.0	305.0	300.2	−8.2
$(CH_3)_3CCl$	268.4	268.3	263.9	−43.5
C_4H_9Cl (n-butyl chloride)	271.9	271.9	268.3	−37.0
C_4H_9N (tetrahydropyrrole)	351.9	351.8	348.2	−0.8
$C_4H_9NO_2$ (2-nitrobutane)	433.9	433.9	428.5	−39.1
$(CH_3CH_2)_2O$	291.2	291.2	288.2	−60.3
$CH_3CH(OCH_3)_2$ (dimethyl acetal)	311.7	311.6	307.8	−93.1
$(CH_3)_3CSH$	311.7	311.6	307.1	−26.2
$C_4H_{10}S_2$ (diethyl disulfide)	352.4	352.5	345.7	−17.9
$(CH_3)_3CNH_2$	350.5	350.4	346.9	−28.9
$(CH_3)_4Si$	301.3	299.5	298.2	−55.7
C_5H_6S (2-methyl thiophene)	357.1	357.1	350.7	20.0
C_5H_7N (N-methyl pyrrole)	400.6	400.6	395.8	24.6
$C_5H_{10}O$ (tetrahydropyran)	347.4	347.3	343.0	−53.4
$(CH_3CH_2)_2C{=}O$	334.7	334.7	330.4	−61.6
$C_5H_{10}O_2$ (isopropyl acetate)	334.6	334.6	329.6	−115.1
$C_5H_{10}S$ (tetrahydrothiopyran)	368.8	368.7	363.1	−15.2

Table B.1: (continued)

Molecule	NWChem	A2*/A2*	A2/A2*	Expt.
$C_5H_{11}N$ (piperidine)	413.2	413.1	408.7	−11.3
$C_5H_{12}O$ (t-butyl methyl ether)	358.8	358.7	354.5	−67.8
$C_6H_4F_2$ (1,3-difluorobenzene)	307.2	307.3	300.3	−73.9
$C_6H_4F_2$ (1,4-difluorobenzene)	308.3	308.4	301.2	−73.3
C_6H_5F (fluorobenzene)	334.3	334.4	328.6	−27.7
$C_6H_{14}O$ (diisopropyl ether)	419.0	418.9	414.3	−76.3
PF_5	−186.3	−185.7	−188.4	−381.1
SF_6	−48.0	−47.7	−50.3	−291.7
P_4	185.3	185.4	175.4	14.1
SO_3	111.9	112.1	109.0	−94.6
SCl_2	70.4	70.5	65.1	−4.2
$POCl_3$	22.1	22.5	15.3	−133.8
PCl_5	66.6	67.0	57.7	−86.1
Cl_2O_2S	114.6	114.8	107.6	−84.8
PCl_3	31.0	31.1	25.6	−69.0
Cl_2S_2	107.2	107.4	100.2	−4.0
$SiCl_2$ ($^1A_1$)	21.0	19.3	19.2	−40.3
CF_3Cl	−14.3	−14.3	−18.7	−169.5
C_2F_6	−50.7	−50.0	−55.5	−321.3
$^2CF_3$	9.0	8.9	6.5	−111.3
$^2C_6H_5$ (phenyl radical)	392.2	392.3	387.8	81.2
M.D.	212.8	212.7	210.2	
M.A.D.	212.8	212.7	210.2	
Max. Dev.	582.0	582.0	577.2	

Table B.2: ADMM Hartree-Fock standard heats of formation [kcal/mol] for the 223 neutral molecules of the G3/99 test set. B3LYP/6-31G(2df,p) optimized geometries, the Def2-TZVPP basis set in spherical representation and the GEN-A2*/GEN-A2* approach were used in all calculations.

Molecule	LDF-EXX	CAP 3-21G	KT3 3-21G	EV93 3-21G	CAP DZVP	KT3 DZVP	EV93 DZVP	Expt.
LiH	57.7	58.3	58.1	57.5	57.8	57.8	57.8	33.3
^{2}BeH	82.0	82.8	82.5	82.8	84.8	85.2	85.0	81.7
^{2}CH	169.8	173.8	173.4	174.2	169.4	169.5	169.1	142.5
CH_2 ($^3B_1$)	129.0	131.7	130.9	131.4	128.6	128.8	128.4	93.7
CH_2 ($^1A_1$)	157.3	162.4	161.8	163.3	156.3	156.6	156.0	102.8
$^3CH_3$	99.4	103.6	102.6	103.2	98.4	98.6	98.1	35.0
CH_4	73.8	78.9	77.7	78.8	72.4	72.8	72.0	−17.9
^{3}NH	117.9	122.4	121.6	122.3	117.6	117.8	117.7	85.2
NH_2 ($^2B_1$)	109.9	116.3	114.9	115.9	109.1	109.4	109.2	45.1
NH_3	86.8	92.3	91.0	91.5	85.4	85.9	85.5	−11.0
^{2}OH	47.9	51.1	50.5	50.1	47.8	47.9	47.8	9.4
H_2O	19.6	22.7	22.8	21.8	18.9	19.3	19.1	−57.8
HF	−20.6	−19.9	−19.1	−20.5	−20.9	−20.9	−21.1	−65.1
SiH_2 ($^1A_1$)	104.6	115.0	114.4	117.0	105.5	105.0	105.4	65.2
SiH_2 ($^3B_1$)	108.8	120.1	118.8	120.0	110.5	110.6	111.0	86.2
$^2SiH_3$	91.2	107.7	106.8	109.3	92.8	92.6	93.2	47.9
SiH_4	71.8	92.5	92.4	95.9	73.0	72.7	73.5	8.2
$^2PH_2$	77.4	91.5	90.4	93.9	77.4	77.5	78.0	33.1
PH_3	70.7	89.3	88.2	93.4	70.3	70.5	71.2	1.3
H_2S	48.2	60.4	58.7	62.6	48.3	48.7	49.5	−4.9
HCl	7.5	13.0	12.0	13.8	8.0	8.3	8.6	−22.1
Li_2	72.3	72.6	73.6	72.9	72.2	72.6	73.1	51.6
LiF	−31.1	−35.1	−30.7	−33.4	−31.5	−31.7	−32.1	−80.1
HC≡CH	166.0	171.4	170.4	171.9	162.4	165.1	164.2	54.2
$H_2C{=}CH_2$	147.3	156.6	155.0	157.2	144.9	146.3	145.3	12.5
$H_3C{-}CH_3$	140.3	150.8	149.4	152.2	138.0	139.2	137.6	−20.1
CN	196.4	205.5	204.0	207.1	194.1	195.6	195.0	104.9
HCN	145.3	157.1	155.9	159.0	142.2	144.2	143.7	31.5
CO	59.0	70.9	71.1	74.0	56.9	58.1	57.1	−26.4
^{2}HCO	106.7	119.5	119.4	122.2	105.2	106.4	105.6	10.0
H_2CO	90.5	104.9	104.8	107.9	89.0	90.0	89.2	−26.0
CH_3OH	95.4	105.4	105.4	108.1	94.6	95.3	94.4	−48.0
N_2	114.1	135.7	132.8	137.5	111.7	113.5	113.4	0.0

Table B.2: (continued)

Molecule	LDF-EXX	CAP 3-21G	KT3 3-21G	EV93 3-21G	CAP DZVP	KT3 DZVP	EV93 DZVP	Expt.
H_2NNH_2	194.1	207.0	205.1	207.9	193.1	194.1	193.8	22.8
NO	121.6	140.1	138.2	141.8	121.3	122.2	121.9	21.6
$^3O_2$	87.5	105.4	104.6	107.4	89.4	89.9	89.6	0.0
H_2O_2	101.9	107.3	108.1	108.6	104.0	104.4	104.2	−32.5
F_2	73.8	72.3	75.7	76.9	78.0	78.3	78.1	0.0
CO_2	49.1	68.8	68.5	72.4	47.2	49.2	48.0	−94.1
Na_2	51.8	51.1	52.3	50.5	50.1	51.9	50.8	34.0
Si_2	182.0	196.1	194.1	197.1	187.2	187.0	187.5	139.9
P_2	113.8	137.0	133.9	140.5	116.5	116.3	117.5	34.3
$^3S_2$	82.1	102.5	98.6	104.1	86.4	85.5	86.5	30.7
Cl_2	40.0	48.6	46.1	50.4	43.6	42.9	44.1	0.0
NaCl	−15.2	−14.5	−18.8	−18.4	−15.4	−15.1	−15.2	−43.6
SiO	56.3	66.4	68.8	72.5	57.8	57.6	57.5	−24.6
CS	141.3	156.9	154.3	158.6	141.2	141.6	141.5	66.9
SO	72.7	87.7	87.9	92.2	74.5	74.6	75.0	1.2
ClO	83.4	90.2	89.8	93.1	85.9	85.9	86.3	24.2
ClF	41.2	41.1	43.5	46.3	44.0	44.0	44.4	−13.2
Si_2H_6	128.1	166.8	165.9	172.2	130.8	130.0	130.4	19.1
CH_3Cl	78.2	88.3	86.3	89.3	78.7	78.9	78.8	−19.6
H_3C-SH	117.2	132.8	130.3	135.1	117.2	117.7	117.9	−5.5
HOCl	69.0	72.9	73.2	75.7	71.4	71.3	71.9	−17.8
SO_2	82.6	118.2	117.4	126.9	87.1	88.0	89.2	−71.0
BF_3	−163.1	−157.8	−150.9	−148.5	−158.4	−158.3	−161.2	−271.4
BCl_3	−18.2	0.5	−4.8	−1.1	−18.5	−18.8	−20.6	−96.3
AlF_3	−181.4	−182.5	−174.2	−171.8	−174.1	−174.0	−174.8	−289.0
$AlCl_3$	−77.2	−61.2	−67.9	−62.4	−75.0	−75.7	−77.5	−139.7
CF_4	−60.6	−44.9	−39.7	−31.6	−55.2	−53.9	−56.5	−223.0
CCl_4	111.1	137.1	132.1	139.6	117.2	116.0	116.9	−22.9
OCS	93.7	114.2	111.8	118.3	93.0	94.6	94.3	−33.1
CS_2	141.6	161.4	157.1	164.6	141.8	142.6	142.9	28.0
F_2CO	9.0	26.6	29.8	35.6	10.6	12.3	10.4	−149.1
SiF_4	−242.3	−232.6	−223.7	−214.3	−228.3	−228.2	−231.0	−386.0
$SiCl_4$	−59.8	−24.0	−32.5	−22.9	−51.3	−52.3	−53.4	−158.4
NNO	204.6	238.2	233.7	241.0	203.9	206.6	206.5	19.6

Table B.2: (continued)

Molecule	LDF-EXX	CAP 3-21G	KT3 3-21G	EV93 3-21G	CAP DZVP	KT3 DZVP	EV93 DZVP	Expt.
ClNO	160.3	181.9	179.0	185.4	162.2	163.0	163.5	12.4
NF_3	136.7	153.1	156.1	163.9	143.5	144.5	143.5	−31.6
PF_3	−98.9	−90.9	−84.7	−74.0	−88.1	−88.3	−88.9	−229.1
O_3	236.7	264.8	262.9	268.1	241.6	242.6	242.1	34.1
F_2O	138.4	143.2	146.5	150.4	145.1	145.5	145.1	5.9
ClF_3	112.7	117.8	123.2	131.3	123.4	122.7	123.8	−38.0
$F_2C{=}CF_2$	52.1	74.7	79.7	90.7	57.2	59.5	57.0	−157.4
$Cl_2C{=}CCl_2$	174.8	203.4	197.7	206.2	179.2	179.4	180.4	−3.0
$F_3C{-}CN$	120.5	147.2	149.4	160.4	122.3	125.6	122.8	−118.4
$HC{\equiv}C{-}CH_3$	224.3	234.9	233.1	236.8	220.4	224.0	222.2	44.2
$H_2C{=}C{=}CH_2$	225.6	237.0	234.7	238.4	222.5	225.4	223.7	45.5
C_3H_4 (cyclopropene)	251.4	269.1	266.4	271.4	248.2	251.3	249.7	66.2
$H_2C{=}CH{-}CH_3$	209.6	224.1	222.2	226.3	206.4	208.8	207.0	4.8
C_3H_6 (cyclopropane)	219.9	235.2	233.1	237.9	217.2	219.4	217.6	12.7
$CH_3{-}CH_2{-}CH_3$	205.4	221.3	219.7	224.1	202.5	204.4	202.0	−25.0
C_4H_6 (Z-1,3-butadiene)	276.4	294.9	292.2	297.7	272.7	276.1	274.1	26.3
C_4H_6 (2-butyne)	284.4	299.7	297.3	303.0	279.9	284.2	281.3	34.8
C_4H_6 (methylenecyclopropane)	298.3	317.1	314.6	320.9	294.8	298.3	296.0	47.9
C_4H_6 (bicyclo[1.1.0]butane)	309.8	334.0	331.0	338.6	306.6	310.3	307.8	51.9
C_4H_6 (cyclobutene)	292.1	313.9	311.7	318.2	289.1	292.5	290.3	37.4
C_4H_8 (cyclobutane)	284.6	305.7	304.5	311.1	282.1	284.9	282.3	6.8
C_4H_8 (isobutene)	271.9	291.8	289.5	295.5	268.3	271.3	269.2	−4.0
C_4H_{10} (trans butane)	270.7	292.1	290.2	296.3	267.0	269.7	266.5	−30.0
C_4H_{10} (isobutane)	270.0	290.8	289.1	294.8	266.4	269.1	266.1	−32.1
C_5H_8 (spiropentane)	368.5	393.9	391.0	399.9	364.6	368.7	365.7	44.3
C_6H_6 (benzene)	361.8	385.6	382.7	391.8	356.8	363.0	360.0	19.7
CH_2F_2	18.9	28.8	32.5	37.5	20.4	20.9	19.4	−107.7
CHF_3	−21.6	−9.0	−4.1	2.6	−18.2	−17.4	−19.5	−166.6
CH_2Cl_2	84.8	100.6	97.5	102.2	87.3	87.2	87.5	−22.8
$CHCl_3$	95.3	116.6	112.4	118.5	99.7	99.1	99.8	−24.7
CH_3NH_2	160.2	172.1	170.7	173.7	158.2	159.3	158.3	−5.5
CH_3CN	199.8	216.5	214.6	219.3	196.3	199.2	198.0	18.0
CH_3NO_2 (nitromethane)	242.5	281.0	277.9	285.4	243.7	246.0	245.0	−17.8
CH_3ONO (methyl nitrite)	240.8	271.8	270.2	278.9	242.6	244.2	243.0	−15.9

Table B.2: (continued)

Molecule	LDF-EXX	CAP 3-21G	KT3 3-21G	EV93 3-21G	CAP DZVP	KT3 DZVP	EV93 DZVP	Expt.
CH_3SiH_3	128.8	150.8	149.6	153.9	129.5	129.6	129.5	−7.0
HCO_2H	79.4	97.5	97.8	101.3	78.5	79.9	78.8	−90.5
HCO_2CH_3	152.5	175.7	176.6	183.6	151.3	153.1	151.2	−85.0
CH_3CONH_2	204.0	227.7	226.3	231.4	201.4	204.1	202.2	−57.0
C_2H_5N (aziridine)	242.5	262.3	259.9	266.1	240.3	242.3	241.0	30.2
NCCN (cyanogen)	284.1	305.9	303.1	310.4	280.2	284.3	282.9	73.3
$NH(CH_3)_2$	230.5	248.7	247.1	252.5	228.1	229.5	227.9	−4.4
$CH_3CH_2NH_2$	223.5	241.6	239.7	244.4	221.1	222.9	221.2	−11.3
$H_2C{=}C{=}O$ (ketene)	150.2	165.3	164.2	168.1	148.5	150.9	149.7	−11.4
C_2H_4O (oxirane)	177.8	196.3	195.4	201.7	176.7	178.4	176.8	−12.6
CH_3CHO	146.4	165.9	165.1	169.5	144.3	146.3	144.6	−39.7
$O{=}CH{-}CH{=}O$	161.5	190.1	189.6	195.9	159.4	162.1	160.4	−50.7
CH_3CH_2OH	157.1	173.2	172.9	177.3	155.4	157.0	155.2	−56.2
$(CH_3)_2O$	167.4	183.6	183.7	189.5	166.3	167.3	165.4	−44.0
C_2H_4S (thiooxirane)	190.3	208.8	205.6	211.7	190.5	191.3	190.8	19.6
$(CH_3)_2S{=}O$	218.4	254.4	251.2	260.3	221.4	222.3	222.8	−36.2
CH_3CH_2SH	182.2	202.6	199.8	206.0	181.5	182.7	182.0	−11.1
$(CH_3)_2S$	184.4	203.8	200.5	206.2	184.4	184.9	184.6	−8.9
$H_2C{=}CHF$	120.1	131.5	132.2	136.4	119.4	121.0	119.7	−33.2
CH_3CH_2Cl	141.3	156.3	154.0	158.3	141.1	142.0	141.0	−26.8
$H_2C{=}CHCl$	149.3	163.1	160.5	164.4	148.6	149.9	149.4	8.9
$H_2C{=}CHCN$	272.2	292.1	289.5	295.4	268.0	271.9	270.3	43.2
$(CH_3)_2C{=}O$	204.3	228.6	227.3	233.1	202.0	204.8	202.6	−51.9
CH_3CO_2H	135.9	158.8	158.6	163.8	134.5	137.0	135.1	−103.4
CH_3CFO	100.3	120.6	121.7	127.6	99.8	102.1	100.2	−105.7
CH_3COCl	139.6	163.5	161.2	167.0	139.4	141.3	140.2	−58.0
$CH_3CH_2CH_2Cl$	206.5	226.7	224.1	230.0	205.6	207.3	205.6	−31.5
$(CH_3)_2CHOH$	219.1	240.9	240.1	246.0	217.0	219.3	216.8	−65.2
$CH_3{-}O{-}CH_2CH_3$	229.3	251.6	251.2	258.8	227.3	229.1	226.4	−51.7
$(CH_3)_3N$	299.7	323.3	321.6	329.1	296.5	298.5	296.1	−5.7
C_4H_4O (furan)	275.2	298.6	297.3	306.3	273.3	277.2	275.1	−8.3
C_4H_4S (thiophene)	294.1	319.8	316.0	325.1	292.4	295.7	294.8	27.5
C_4H_5N (pyrrole)	331.5	353.7	351.5	359.7	327.5	332.0	329.9	25.9
C_5H_5N (pyridine)	379.1	408.8	405.3	415.7	375.2	380.5	378.0	33.6

Table B.2: (continued)

Molecule	LDF-EXX	CAP 3-21G	KT3 3-21G	EV93 3-21G	CAP DZVP	KT3 DZVP	EV93 DZVP	Expt.
H_2	25.8	27.3	26.7	26.6	25.6	25.2	25.1	0.0
^2SH	60.2	67.5	66.5	68.5	60.6	60.8	61.2	34.2
$^2C{\equiv}CH$	215.7	219.4	218.2	219.6	213.2	215.6	214.4	135.1
$HC{=}CH_2$ ($^2A'$)	178.6	185.5	184.0	185.8	176.5	178.1	176.8	71.6
$CH_3C{=}O$ ($^2A'$)	163.0	181.0	180.0	184.3	161.1	163.1	161.6	−2.4
$CH_2{-}OH$ (^2A)	115.7	125.2	125.0	127.1	115.1	115.7	115.0	−4.1
CH_3O ($^2A'$)	113.4	124.2	123.5	126.6	112.7	113.2	112.3	4.1
CH_3CH_2O ($^2A''$)	174.7	191.3	190.1	194.6	173.3	174.7	172.9	−3.7
CH_3S ($^2A'$)	125.9	136.6	134.7	137.6	126.2	126.4	126.3	29.8
CH_2CH_3 ($^2A'$)	163.4	172.8	171.2	173.7	161.4	162.5	161.0	28.9
$(CH_3)_2CH$ ($^2A'$)	226.4	241.0	238.9	243.0	223.6	225.5	223.2	21.5
$^2C(CH_3)_3$	289.4	309.0	306.6	312.3	285.8	288.5	285.5	12.3
NO_2 ($^2A_1$)	181.6	213.5	209.7	215.8	182.6	184.4	184.1	7.9
$CH_3CH{=}C{=}CH_2$	289.1	306.0	303.3	308.8	285.1	289.0	286.5	38.8
C_5H_8 (isoprene)	339.8	363.8	360.8	368.2	335.6	339.9	337.5	18.0
C_5H_{10} (cyclopentane twist)	331.1	356.0	355.0	363.0	327.6	331.2	327.7	−18.3
C_5H_{12} (n-pentane)	336.1	363.0	360.8	368.7	331.6	335.1	331.2	−35.1
$C(CH_3)_4$ (neopentane)	335.2	360.3	358.3	364.9	330.6	333.9	330.3	−40.2
C_6H_8 (1,3-cyclohexadiene)	394.8	422.9	420.6	429.7	390.0	395.5	391.9	25.4
C_6H_8 (1,4-cyclohexadiene)	392.9	421.9	419.6	429.0	388.0	393.6	390.1	25.0
C_6H_{12} (cyclohexane chair)	390.7	422.0	421.0	430.6	386.2	390.8	386.6	−29.5
C_6H_{14} (n-hexane)	401.4	433.9	431.4	441.0	396.2	400.5	395.9	−39.9
C_6H_{14} (3-methylpentane)	403.4	435.0	432.7	441.9	398.2	402.5	398.2	−41.1
$C_6H_5{-}CH_3$ (toluene)	425.6	454.7	451.3	462.3	420.3	427.2	423.5	12.0
C_7H_{16} (n-heptane)	466.8	504.8	502.0	513.3	460.7	465.9	460.5	−44.9
C_8H_8 (cyclooctatetraene)	531.7	566.7	563.0	574.4	524.9	533.2	528.6	70.7
C_8H_{18} (n-octane)	532.1	575.7	572.6	585.6	525.4	531.3	525.2	−49.9
$C_{10}H_8$ (naphtalene)	587.7	628.3	623.4	639.6	581.3	592.0	586.6	35.9
$C_{10}H_8$ (azulene)	631.0	672.8	667.5	683.5	624.3	635.1	629.8	69.1
$CH_3CO_2CH_3$ (Z-methylacetate)	209.5	237.6	238.0	246.7	207.8	210.6	207.9	−98.4
$(CH_3)_3COH$ (t-butanol)	282.0	308.4	307.7	314.8	279.1	282.3	279.3	−74.7
$C_6H_5NH_2$ (aniline)	439.0	469.8	466.0	477.2	434.0	441.0	437.7	20.8
C_6H_5OH (phenol)	374.4	404.0	401.3	412.3	370.7	377.3	374.1	−23.0
C_4H_6O (divinyl ether)	299.4	323.6	322.0	330.2	297.4	300.9	298.7	−3.3

Table B.2: (continued)

Molecule	LDF-EXX	CAP 3-21G	KT3 3-21G	EV93 3-21G	CAP DZVP	KT3 DZVP	EV93 DZVP	Expt.
C_4H_8O (tetrahydrofuran)	285.6	312.2	312.5	322.0	284.1	287.0	284.0	−44.0
C_5H_8O (cyclopentanone)	327.5	360.9	360.1	369.2	324.9	329.5	325.9	−45.9
$C_6H_4O_2$ (benzoquinone)	393.0	439.1	436.2	448.5	389.6	396.5	393.0	−29.4
$C_6H_4N_2$ (pyrimidine)	392.7	428.8	424.6	436.3	389.5	394.0	391.8	46.8
$(CH_3)_2SO_2$	221.1	279.2	274.2	288.2	227.3	229.4	230.5	−89.2
C_6H_5Cl (chlorobenzene)	364.8	392.3	387.9	398.7	361.7	367.4	364.7	12.4
$NC(CH_2)_2CN$ (succinonitrile)	394.0	427.6	424.7	434.4	388.0	394.2	391.5	50.1
$C_4H_4N_2$ (pyrazine)	399.4	435.5	431.1	443.0	396.6	400.7	398.9	46.9
C_4H_4O (3-butyn-2-one)	296.1	319.7	317.7	324.5	292.0	297.0	294.2	15.6
C_4H_6O (E-crotonaldehyde)	276.0	304.5	302.4	309.7	272.6	276.4	273.8	−24.0
$C_4H_6O_3$ (acetic anhydride)	266.7	307.2	307.1	318.7	264.9	269.2	265.7	−136.8
C_4H_6S (2,5-dihydrothiophene)	306.7	333.4	330.2	339.2	305.6	308.3	306.9	20.8
$(CH_3)_2CHCN$	329.7	356.7	354.6	362.2	325.0	329.4	326.7	5.6
C_4H_8O (methyl ethyl ketone)	269.5	298.6	297.2	304.9	266.2	269.8	266.7	−57.1
$(CH_3)_2CHCHO$	276.9	306.3	305.1	312.4	273.9	277.4	274.4	−51.6
$C_4H_8O_2$ (1,4-dioxane)	305.3	340.1	341.3	353.7	305.1	307.9	304.4	−75.5
C_4H_8S (tetrahydrothiophene)	305.0	332.9	330.4	339.3	304.3	306.5	304.6	−8.2
$(CH_3)_3CCl$	268.3	292.0	289.5	296.1	266.4	268.6	266.1	−43.5
C_4H_9Cl (n-butyl chloride)	271.9	297.6	294.7	302.3	270.3	272.7	270.2	−37.0
C_4H_9N (tetrahydropyrrole)	351.8	379.9	378.7	388.1	348.5	351.9	349.0	−0.8
$C_4H_9NO_2$ (2-nitrobutane)	433.9	487.9	484.3	497.0	433.3	438.0	434.8	−39.1
$(CH_3CH_2)_2O$	291.2	319.8	318.8	328.3	288.5	291.1	287.5	−60.3
$CH_3CH(OCH_3)_2$ (dimethyl acetal)	311.6	343.4	344.4	356.5	310.3	312.8	309.1	−93.1
$(CH_3)_3CSH$	311.6	340.8	337.8	346.2	308.9	311.5	309.3	−26.2
$C_4H_{10}S_2$ (diethyl disulfide)	352.5	391.7	385.4	397.4	353.5	354.8	353.5	−17.9
$(CH_3)_3CNH_2$	350.4	378.0	376.2	383.5	346.1	349.6	346.6	−28.9
$(CH_3)_4Si$	299.5	325.7	322.5	329.8	298.0	300.5	298.2	−55.7
C_5H_6S (2-methyl thiophene)	357.1	387.3	383.2	394.1	354.8	358.9	357.3	20.0
C_5H_7N (N-methyl pyrrole)	400.6	428.5	426.3	437.2	396.3	401.3	398.8	24.6
$C_5H_{10}O$ (tetrahydropyran)	347.3	380.2	380.2	391.2	345.0	348.8	344.9	−53.4
$(CH_3CH_2)_2C{=}O$	334.7	368.6	367.4	376.9	330.5	334.9	331.0	−61.6
$C_5H_{10}O_2$ (isopropyl acetate)	334.6	373.7	373.4	385.6	331.5	335.9	331.7	−115.1
$C_5H_{10}S$ (tetrahydrothiopyran)	368.7	402.6	400.0	410.2	367.0	370.2	367.5	−15.2
$C_5H_{11}N$ (piperidine)	413.1	447.6	446.2	456.7	409.3	413.7	410.0	−11.3

Table B.2: (continued)

Molecule	LDF-EXX	CAP 3-21G	KT3 3-21G	EV93 3-21G	CAP DZVP	KT3 DZVP	EV93 DZVP	Expt.
$C_5H_{12}O$ (t-butyl methyl ether)	358.7	390.4	389.6	399.7	355.3	358.6	354.8	−67.8
$C_6H_4F_2$ (1,3-difluorobenzene)	307.3	337.3	337.4	351.0	307.0	313.2	309.6	−73.9
$C_6H_4F_2$ (1,4-difluorobenzene)	308.4	338.5	338.5	352.2	308.1	314.2	310.6	−73.3
C_6H_5F (fluorobenzene)	334.4	361.0	359.7	371.0	331.7	337.8	334.6	−27.7
$C_6H_{14}O$ (diisopropyl ether)	418.9	457.6	455.9	468.4	414.7	418.6	414.0	−76.3
PF_5	−185.7	−168.3	−159.1	−141.1	−167.4	−166.7	−168.9	−381.1
SF_6	−47.7	−21.4	−14.3	10.2	−22.4	−21.4	−22.4	−291.7
P_4	185.4	246.4	235.1	254.5	195.3	192.2	193.8	14.1
SO_3	112.1	167.4	164.0	177.8	119.5	121.5	122.9	−94.6
SCl_2	70.5	89.7	84.3	92.6	76.8	75.4	77.3	−4.2
$POCl_3$	22.5	72.9	65.5	78.8	32.3	31.6	32.8	−133.8
PCl_5	67.0	117.0	104.2	118.8	81.1	78.2	81.0	−86.1
Cl_2O_2S	114.8	175.5	167.6	184.2	126.9	127.3	129.7	−84.8
PCl_3	31.1	60.0	52.8	63.2	39.0	37.3	38.9	−69.0
Cl_2S_2	107.4	137.9	129.6	141.6	116.6	114.6	117.3	−4.0
$SiCl_2$ ($^1A_1$)	19.3	33.5	29.6	35.5	24.1	23.1	23.2	−40.3
CF_3Cl	−14.3	5.0	8.3	16.6	−8.9	−8.0	−9.9	−169.5
C_2F_6	−50.0	−19.3	−11.2	3.0	−42.5	−39.5	−44.0	−321.3
$^2CF_3$	8.9	19.5	24.1	30.0	12.4	13.1	11.2	−111.3
$^2C_6H_5$ (phenyl radical)	392.3	414.0	410.9	419.4	387.7	393.9	390.0	81.2
M.D.	212.7	233.7	232.3	238.9	212.5	214.5	213.1	
M.A.D.	212.7	233.7	232.3	238.9	212.5	214.5	213.1	
Max. Dev.	582.0	625.6	622.5	635.5	575.3	581.2	575.1	

Bibliography

(1) R. S. Mulliken, C. C. J. Roothaan, "Broken Bottlenecks and the Future of Molecular Quantum Mechanics", *Proc. Natl. Acad. Sci.* **45**, 394 (1959).

(2) C. Longuet-Higgins, "Closing Remarks", *Faraday Discuss. Chem. Soc.* **62**, 347 (1977).

(3) I. Kant, *Methaphysische Anfangsgründe der Naturwissenschaft*, Hartknoch, Leipzig, Germany (1786).

(4) W. Kutzelnigg, "Perspective on 'Quantum Mechanics of Many-Electron Systems'", *Theor. Chem. Acc.* **103**, 182 (2000).

(5) J. L. Gay-Lussac, "Mémoire sur la Combinaison des Substances Gazeuses, les Unes Avec les Autres", *Mem. Soc. d'Arcueil* **2**, 207 (1809).

(6) A. Comte, *Cours de Philosophie Positive. Tome IV*, Schleicher, Paris, France (1830).

(7) P. A. M. Dirac, "Quantum Mechanics of Many-Electron Systems", *Proc. R. Soc. Lond. Ser. A* **123**, 714 (1929).

(8) A. Simões, "Dirac's Claim and the Chemists", *Phys. Perspect.* **4**, 253 (2002).

(9) G. F. Thomas, "The Emancipation of Chemistry", *Found. Chem.* **14**, 109 (2012).

(10) R. G. Parr, B. L. Crawford Jr., "National Academy of Sciences Conference on Quantum-Mechanical Methods in Valence Theory", *Proc. Natl. Acad. Sci.* **38**, 547 (1952).

(11) F. Soddy, "The Wider Aspects of The Discovery of Atomic Disintegration Contrasting the Experimental Facts with the Mathematical Theories", 4th Conference of Nobel Prizewinners at Lindau (1954).

(12) E. Schrödinger, "Quantisierung als Eigenwertproblem (Erste Mitteilung)", *Ann. Phys.* **79**, 361 (1926).

(13) E. Schrödinger, "Quantisierung als Eigenwertproblem (Zweite Mitteilung)", *Ann. Phys.* **79**, 489 (1926).

(14) E. Schrödinger, "Quantisierung als Eigenwertproblem (Dritte Mitteilung)", *Ann. Phys.* **80**, 437 (1926).

(15) E. Schrödinger, "Quantisierung als Eigenwertproblem (Vierte Mitteilung)", *Ann. Phys.* **81**, 109 (1926).

(16) E. Schrödinger, "An Undulatory Theory of the Mechanics of Atoms and Molecules", *Phys. Rev.* **28**, 1049 (1926).

(17) D. R. Hartree, "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. I. Theory and Methods", *Proc. Cambridge Philos. Soc.* **24**, 89 (1928).

(18) V. A. Fock, "Näherungsmethode zur Lösung des Quantenmechanischen Mehrkörperproblems", *Z. Phys.* **61**, 126 (1930).

(19) V. A. Fock, "'Self-Consistent Field' mit Austausch für Natrium", *Z. Phys.* **62**, 795 (1930).

(20) V. A. Fock, "Über Austauschenergie", *Z. Phys.* **81**, 195 (1933).

(21) P. Hohenberg, W. Kohn, "Inhomogeneous Electron Gas", *Phys. Rev.* **136**, B864 (1964).

(22) W. Kohn, L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", *Phys. Rev.* **140**, A1133 (1965).

(23) G. E. Moore, "Cramming More Components onto Integrated Circuits", *Electronics Magazine* **38**, 114 (1965).

(24) G. E. Moore, "Cramming More Components onto Integrated Circuits", *Proc. IEEE* **86**, 82 (1998).

(25) C. Ochsenfeld, J. Kussmann, D. S. Lambrecht, "Linear-Scaling Methods in Quantum Chemistry", in: *Reviews in Computational Chemistry*, (Eds. K. B. Lipkowitz, T. R. Cundari) Volume 23, pp. 1–82, John Wiley & Sons, Inc., Hoboken, NJ, USA (2007).

(26) G. M. Amdahl, "Validity of the Single Processor Approach to Achieving Large Scale Computing Capabilities", in: *Proceedings of the April 18-20, 1967, Spring Joint Computer Conference* AFIPS '67 (Spring), pp. 483–485, ACM, New York, NY, USA (1967), doi:10.1145/1465482.1465560.

(27) M. Born, R. A. Oppenheimer, "Zur Quantentheorie der Molekeln", *Ann. Phys.* **84**, 457 (1927).

(28) P. W. Atkins, R. Friedman, *Molecular Quantum Mechanics*, 4th edition, Oxford University Press, Oxford, Great Britain (2005).

(29) M. Born, V. A. Fock, "Beweis des Adiabatensatzes", *Z. Phys.* **51**, 165 (1928).

(30) J. W. Rayleigh, "In Finding the Correction for the Open End of an Organ-Pipe", *Phil. Trans.* **161**, 77 (1870).

(31) W. Ritz, "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik", *J. Reine Angew. Math.* **135**, 1 (1908).

(32) C. Eckart, "The Theory and Calculation of Screening Constants", *Phys. Rev.* **36**, 878 (1930).

(33) R. K. Nesbet, *Variational Principles and Methods in Theoretical Physics and Chemistry*, Cambridge University Press, Cambridge, United Kingdom (2004).

(34) L. H. Thomas, "The Calculation of Atomic Fields", *Proc. Cambridge Phil. Soc.* **23**, 542 (1927).

(35) E. Fermi, "Un Metodo Statistico per la Determinazione di alcune Proprietà dell'Atomo", *Rend. Lincei* **6**, 602 (1927).

(36) P. A. M. Dirac, "Note on Exchange Phenomena in the Thomas-Fermi Atom", *Math. Proc. Cambridge* **26**, 376 (1930).

(37) E. Wigner, F. Seitz, "On the Constitution of Metallic Sodium", *Phys. Rev.* **43**, 804 (1933).

(38) E. Wigner, F. Seitz, "On the Consitution of Metallic Sodium. II", *Phys. Rev.* **46**, 509 (1934).

(39) J. C. Slater, "A Simplification of the Hartree-Fock Method", *Phys. Rev.* **81**, 385 (1951).

(40) J. C. Slater, K. H. Johnson, "Self-Consistent-Field $X\alpha$ Method for Polyatomic Molecules and Solids", *Phys. Rev. B* **5**, 844 (1972).

(41) E. Teller, "On the Stability of Molecules in the Thomas-Fermi Theory", *Rev. Mod. Phys.* **34**, 627 (1962).

(42) P. A. M. Dirac, "A New Notation For Quantum Mechanics", *Math. Proc. Cambridge* **35**, 416 (1939).

(43) R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, The International Series of Monographs in Chemistry, Oxford University Press, New York, USA (1989).

(44) M. Levy, "Universal Variational Functionals of Electron Densities, First-Order Density Matrices, and Natural-Spin Orbitals and Solution of the v-Representability Problem", *Proc. Natl. Acad. Sci.* **76**, 6062 (1979).

(45) E. H. Lieb, "Density Functionals for Coulomb Systems", *Int. J. Quantum Chem.* **24**, 243 (1983).

(46) H. Eschrig, *The Fundamentals of Density Functional Theory*, Edition am Gutenbergplatz, Leipzig, Germany (2003).

(47) E. Runge, E. K. U. Gross, "Density-Functional Theory for Time-Dependent Systems", *Phys. Rev. Lett.* **52**, 997 (1984).

(48) H. Kohl, R. M. Dreizler, "Time-Dependent Density-Functional Theory: Conceptual and Practical Aspects", *Phys. Rev. Lett.* **56**, 1993 (1986).

(49) M. E. Casida, "Time-Dependent Density-Functional Response Theory for Molecules", in: *Recent Advances in Density Functional Methods* World Scientific, Singapore (1995).

(50) A. Görling, "Density-Functional Theory for Excited States", *Phys. Rev. A* **54**, 3912 (1996).

(51) A. Görling, "Density-Functional Theory Beyond the Hohenberg-Kohn Theorem", *Phys. Rev. A* **59**, 3359 (1999).

(52) M. Levy, A. Nagy, "Variational Density-Functional Theory for an Individual Excited State", *Phys. Rev. Lett.* **83**, 4361 (1999).

(53) A. Nagy, M. Levy, "Variational Density-Functional Theory for Degenerate Excited States", *Phys. Rev. A* **63**, 052502 (2001).

(54) P. W. Ayers, M. Levy, "Time-Independent (Static) Density-Functional Theories for Pure Excited States: Extensions and Unification", *Phys. Rev. A* **80**, 012508 (2009).

(55) P. W. Ayers, M. Levy, A. Nagy, "Time-Independent Density-Functional Theory for Excited States of Coulomb Systems", *Phys. Rev. A* **85**, 042518 (2012).

(56) N. Balàzs, "Formation of Stable Molecules Within the Statistical Theory of Atoms", *Phys. Rev.* **156**, 42 (1967).

(57) E. Lieb, "Thomas-Fermi and Related Theories of Atoms and Molecules", *Rev. Mod. Phys.* **53**, 603 (1981).

(58) C. F. von Weizsäcker, "Zur Theorie der Kernmassen", *Z. Phys.* **96**, 431 (1935).

(59) K. Capelle, "A Bird's-Eye View of Density-Functional Theory", *Braz. J. Phys.* **36**, 1318 (2006).

(60) J. C. Slater, H. C. Verna, "The Theory of Complex Spectra", *Phys. Rev.* **34**, 1293 (1929).

(61) N. Mardirossian, M. Head-Gordon, "ωB97X-V: A 10-Parameter, Range-Separated Hybrid, Generalized-Gradient Approximation Density Functional with Nonlocal Correlation, Designed by a Survival-of-the-Fittest Strategy", *Phys. Chem. Chem. Phys.* **16**, 9904 (2014).

(62) A. D. Becke, "Perspective: Fifty Years of Density-Functional Theory in Chemical Physics", *J. Phys. Chem.* **140**, 18A301 (2014).

(63) S. Redner, "Citation Statistics From 110 Years of Physical Review", *Physics Today* **58**, 49 (2005).

(64) J. P. Perdew, K. Schmidt, "Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy", *AIP Conf. Proc.* **577**, 1 (2001).

(65) J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, "Prescription for the Design and Selection of Density Functional Approximations: More Constraint Satisfaction with Fewer Fits", *J. Chem. Phys.* **123**, 062201 (2005).

(66) A. Vela, V. Medel, S. B. Trickey, "Variable Lieb-Oxford Bound Satisfaction in a Generalized Gradient Exchange-Correlation Functional", *J. Chem. Phys.* **130**, 244103 (2009).

(67) J. Perdew, K. Burke, M. Ernzerhof, "Generalized Gradient Approximation Made Simple", *Phys. Rev. Lett.* **77**, 3865 (1997).

(68) A. Vela, J. C. Pacheco-Kato, J. L. Gázquez, J. M. del Campo, S. B. Trickey, "Improved Constraint Satisfaction in a Simple Generalized Gradient Approximation Exchange Functional", *J. Chem. Phys.* **136**, 144115 (2012).

(69) T. W. Keal, D. J. Tozer, "The Exchange-Correlation Potential in Kohn-Sham Nuclear Magnetic Resonance Shielding Calculations", *J. Chem. Phys.* **119**, 3015 (2003).

(70) T. W. Keal, D. J. Tozer, "A Semiempirical Generalized Gradient Approximation Exchange-Correlation Functional", *J. Chem. Phys.* **121**, 5654 (2004).

(71) J. Carmona-Espíndola, J. L. Gázquez, A. Vela, S. B. Trickey, "Generalized Gradient Approximation Exchange Energy Functional with Correct Asymptotic Behavior of the Corresponding Potential", *J. Chem. Phys.* **142**, 054105 (2015).

(72) A. M. Köster, G. Geudtner, P. Calaminici, M. E. Casida, V. D. Domínguez, R. Flores-Moreno, G. U. Gamboa, T. Heine, A. Ipatov, F. Janetzko, J. M. del Campo, J. U. Reveles, A. Vela, B. Zuñiga-Gutiérrez, D. R. Salahub, "deMon2k, Version 3; The deMon Developers: Cinvestav", México, D. F. (2011). See also: www.demon-software.com.

(73) G. Geudtner, P. Calaminici, J. Carmona-Espíndola, J. M. del Campo, V. D. Domínguez-Soria, R. Flores-Moreno, G. U. Gamboa, A. Goursot, A. M. Köster, J. U. Reveles, T. Mineva, J. M. Vásquez-Pérez, A. Vela, D. R. Salahub, "deMon2k", *WIREs Comput. Mol. Sci.* **2**, 548 (2012).

(74) D. R. Salahub, S. Y. Noskov, B. Lev, R. Zhang, V. Ngo, A. Goursot, P. Calaminici, A. M. Köster, A. Alvarez-Ibarra, D. Mejia-Rodriguez, J. Rezac, F. Cailliez, A. de la Lande, "QM/MM Calculations with deMon2k", *Molecules* **20**, 4780 (2015).

(75) S. F. Boys, "Electronic Wave Functions. I. A General Method of Calculation for the Stationary States of Any Molecular System", *Proc. R. Soc. Lond. A* **200**, 542 (1950).

(76) B. Ford, G. Hall, "The Generalized Eigenvalue Problem in Quantum Chemistry", *Comp. Phys. Comm.* **8**, 337 (1974).

(77) C. C. J. Roothaan, "New Developments in Molecular Orbital Theory", *Rev. Mod. Phys.* **23**, 69 (1951).

(78) G. G. Hall, "The Molecular-Orbital Theory of Chemical Valency. VIII. A Method of Calculating Ionization Potentials", *Proc. Roy. Soc.* **A205**, 541 (1951).

(79) C. L. Lawson, R. J. Hanson, D. Kincaid, F. T. Krogh, "Basic Linear Algebra Subprograms for FORTRAN usage", *ACM Trans. Math. Soft.* **5**, 308 (1979).

(80) J. J. Dongarra, J. Du Croz, S. Hammarling, R. J. Hanson, "An Extended Set of FORTRAN Basic Linear Algebra Subprograms", *ACM Trans. Math. Soft.* **14**, 1 (1988).

(81) J. J. Dongarra, J. Du Croz, S. Hammarling, R. J. Hanson, "Algorithm 656: An Extended Set of FORTRAN Basic Linear Algebra Subprograms", *ACM Trans. Math. Soft.* **14**, 18 (1988).

(82) J. J. Dongarra, J. Du Croz, I. S. Duff, S. Hammarling, "A Set of Level 3 Basic Linear Algebra Subprograms", *ACM Trans. Math. Soft.* **16**, 1 (1990).

(83) J. J. Dongarra, J. Du Croz, I. S. Duff, S. Hammarling, "Algorithm 679: A Set of Level 3 Basic Linear Algebra Subprograms", *ACM Trans. Math. Soft.* **16**, 18 (1990).

(84) L. S. Blackford, J. Demmel, J. Dongarra, I. Duff, S. Hammarling, G. Henry, M. Heroux, L. Kaufman, A. Lumsdaine, A. Petitet, R. Pozo, K. Remington, R. C. Whaley, "An Updated Set of Basic Linear Algebra Subprograms (BLAS)", *ACM Trans. Math. Soft.* **28**, 135 (2002).

(85) J. Dongarra, "Basic Linear Algebra Subprograms Technical Forum Standard", *Int. J. High Perform. C.* **16**, 1 (2002).

(86) J. Dongarra, "Basic Linear Algebra Subprograms Technical Forum Standard", *Int. J. High Perform. C.* **16**, 115 (2002).

(87) 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*, 3rd edition, Society for Industrial and Applied Mathematics, Philadelphia, PA (1999).

(88) B. I. Dunlap, J. W. D. Conolly, J. R. Sabin, "On Some Approximations in Applications of $X\alpha$ Theory", *J. Chem. Phys.* **71**, 3396 (1979).

(89) B. I. Dunlap, J. W. D. Conolly, J. R. Sabin, "On First-Row Diatomic Molecules and Local Density Models", *J. Chem. Phys.* **71**, 4993 (1979).

(90) J. W. Mintmire, B. I. Dunlap, "Fitting the Coulomb Potential Variationally in Linear-Combination-of-Atomic-Orbitals Density-Functional Calculations", *Phys. Rev. A* **25**, 88 (1982).

(91) J. W. Mintmire, J. R. Sabin, S. B. Trickey, "Local-Density-Functional Methods in Two-Dimensionally Periodic Systems. Hydrogen and Beryllium Monolayers", *Phys. Rev. B* **26**, 1743 (1982).

(92) H. Sambe, R. H. Felton, "A New Computational Approach to Slater's SCF-$X\alpha$ Equation", *J. Chem. Phys.* **62**, 1122 (1975).

(93) M. E. Casida, C. Daul, A. Goursot, A. M. Köster, L. G. M. Pettersson, E. Proynov, A. St-Amant, D. R. Salahub, H. Duarte, N. Godbout, J. Guan, C. Jamorski, M. Leboeuf, V. Malkin, O. Malkina, F. Sima, A. Vela, "deMon-KS Version 3.4", deMon Software, Montréal (1996).

(94) J. Andzelm, "DGauss: Density Functional–Gaussian Approach. Implementation and Applications", in: *Density Functional Methods in Chemistry* , (Eds. J. K. Labanowski, J. W. Andzelm) Springer, New York, USA (1991).

(95) M. Feyereisen, G. Fitzgerald, A. Komornicki, "Use of Approximate Integrals in Ab Initio Theory. An Application in MP2 Energy Calculations", *Chem. Phys. Lett.* **208**, 359 (1993).

(96) O. Vahtras, J. Almlöf, M. W. Feyereisen, "Integral Approximations for LCAO-SCF Calculations", *Chem. Phys. Lett.* **213**, 514 (1993).

(97) B. I. Dunlap, "Robust and Variational Fitting: Removing the Four-Center Integrals From Center Stage in Quantum Chemistry", *J. Mol. Struct.: THEOCHEM* **529**, 37 (2000).

(98) B. I. Dunlap, N. Rösch, S. B. Trickey, "Variational Fitting Methods for Electronic Structure Calculations", *Mol. Phys.* **108**, 3167 (2010).

(99) A. M. Köster, "Hermite Gaussian Auxiliary Functions for the Variational Fitting of the Coulomb Potential in Density Functional Methods", *J. Chem. Phys.* **118**, 9943 (2003).

(100) J. Andzelm, E. Radzio, D. R. Salahub, "Compact Basis Sets for LCAO-LSD Calculations. Part I: Method and Bases for Sc to Zn", *J. Comput. Chem.* **6**, 520 (1985).

(101) J. Andzelm, N. Russo, D. R. Salahub, "Ground and Excited States of Group IVA Diatomics From Local-Spin-Density Calculations: Model Potentials for Si, Ge, and Sn", *J. Chem. Phys.* **87**, 6562 (1987).

(102) A. M. Köster, "Efficient Recursive Computation of Molecular Integrals for Density Functional Methods", *J. Chem. Phys.* **104**, 4114 (1996).

(103) A. Álvarez Ibarra, "Asymptotic Expansion of Molecular Integrals in Self-Consistent Auxiliary Density Functional Methods", Ph. D. Thesis; Cinvestav (2013).

(104) F. Weigend, "A Fully Direct RI-HF Algorithm: Implementation, Optimised Auxiliary Basis Sets, Demonstration of Accuracy and Efficiency", *Phys. Chem. Chem. Phys.* **4**, 4285 (2002).

(105) K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, "Auxiliary Basis Sets to Approximate Coulomb Potentials", *Chem. Phys. Lett.* **240**, 283 (1995).

(106) K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, "Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials", *Theor. Chem. Acc.* **97**, 119 (1997).

(107) F. Weigend, "Accurate Coulomb-Fitting Basis Sets for H to Rn", *Phys. Chem. Chem. Phys.* **8**, 1057 (2006).

(108) V. D. Domínguez-Soria, G. Geudtner, J. L. Morales, P. Calaminici, A. M. Köster, "Robust and Efficient Density Fitting", *J. Chem. Phys.* **131**, 124102 (2009).

(109) A. Alvarez-Ibarra, A. M. Köster, "Double Asymptotic Expansion of Three-Center Electronic Repulsion Integrals", *J. Chem. Phys.* **139**, 024102 (2013).

(110) E. J. Baerends, D. E. Ellis, P. Ros, "Self-Consistent Molecular Hatree-Fock-Slater Calculations I. The Computational Procedure", *Chem. Phys.* **2**, 41 (1973).

(111) B. I. Dunlap, "Geometry Optimization Using Local Density Functional Methods", *J. Phys. Chem.* **90**, 5524 (1986).

(112) L. Versluis, T. Ziegler, "The Determination of Molecular Structures by Density Functional Theory. The Evaluation of Analytical Energy Gradients by Numerical Integration", *J. Chem. Phys.* **88**, 322 (1988).

(113) B. I. Dunlap, "Analytic and Variational $X\alpha$ in the Slater-Roothaan Method", *J. Phys. Chem. A* **107**, 10082 (2003).

(114) D. Laikov, "Fast Evaluation of Density Functional Exchange-Correlation Terms Using the Expansion of the Electron Density in Auxiliary Basis Sets", *Chem. Phys. Lett.* **281**, 151 (1997).

(115) A. M. Köster, "Entwicklung einer LCGTO-Dichtefunktionalmethode mit Hilfsfunktionen", Habilitation Thesis; Universität Hannover (1998).

(116) A. M. Köster, J. U. Reveles, J. M. del Campo, "Calculation of Exchange-Correlation Potentials with Auxiliary Function Densities", *J. Chem. Phys.* **121**, 3417 (2004).

(117) U. Birkenheuer, A. Gordienko, V. Nasluzov, M. Fuchs-Rohr, N. Rösch, "Model Density Approach to the Kohn-Sham Problem: Efficient Extension of the Density Fitting Technique", *Int. J. Quantum Chem.* **102**, 743 (2005).

(118) A. M. Köster, J. M. del Campo, F. Janetzko, B. Zuniga-Gutierrez, "A MinMax Self-Consistent-Field Approximation for Auxiliary Density Functional Theory", *J. Chem. Phys.* **130**, 114106 (2009).

(119) B. G. Johnson, D. A. Holder, "A Generalized Formulation of Density Functional Theory with Auxiliary Basis Sets", in: *The J-Matrix Method. Developments and Applications* Springer, New York, USA (2008).

(120) P. Pulay, "Convergence Acceleration of Iterative Sequences, the Case of SCF Iteration", *Chem. Phys. Lett.* **73**, 393 (1980).

(121) P. Pulay, "Improved SCF Convergence Acceleration", *J. Comput. Chem.* **3**, 556 (1982).

(122) E. Engel, R. M. Dreizler, *Density Functional Theory. An Advanced Course*, Springer-Verlag Berlin Heidelberg, Berlin, Germany (2011).

(123) K. Burke, "Perspective on Density Functional Theory", *J. Chem. Phys.* **136**, 150901 (2012).

(124) E. Schwegler, M. Challacombe, "Linear Scaling Computation of the Fock Matrix", *J. Chem. Phys.* **106**, 5526 (1997).

(125) V. Dyczmons, "No N^4-Dependence in the Calculation of Large Molecules", *Theor. Chim. Acta* **28**, 307 (1973).

(126) R. Ahlrichs, "Methods for Efficient Evaluation of Integrals for Gaussian Type Basis Sets", *Theor. Chim. Acta* **33**, 157 (1974).

(127) E. Schwegler, M. Challacombe, "Linear Scaling Computational of the Hartree-Fock Exchange Matrix", *J. Chem. Phys.* **105**, 2726 (1996).

(128) E. Schwegler, M. Challacombe, M. Head-Gordon, "Linear Scaling Computation of the Fock Matrix. II. Rigorous Bounds on Exchange Integrals and Incremental Fock Build", *J. Chem. Phys.* **106**, 9708 (1997).

(129) E. Schwegler, M. Challacombe, "Linear Scaling Computation of the Fock Matrix. III. Formation of the Exchnage Matrix With Permutational Symmetry", *Theor. Chem. Acc.* **104**, 344 (2000).

(130) C. Ochsenfeld, C. A. White, M. Head-Gordon, "Linear and Sublinear Scaling Formation of Hartree-Fock Type Exchange Matrices", *J. Chem. Phys.* **109**, 1663 (1998).

(131) W. Liang, Y. Shao, C. Ochsenfeld, A. T. Bell, M. Head-Gordon, "Fast Evaluation of a Linear Number of Local Exchange Matrices", *Chem. Phys. Lett.* **358**, 43 (2002).

(132) R. A. Friesner, "Solution of Self-Consistent Field Electronic Structure Equations by a Pseudospectral Method", *Chem. Phys. Lett.* **116**, 39 (1985).

(133) R. A. Friesner, "A Solution of the Hartree-Fock Equations by a Pseudospectral Method: Application to Diatomic Molecules", *J. Chem. Phys.* **85**, 1462 (1986).

(134) R. A. Friesner, "Solution of the Hartree-Fock Equations for Polyatomic Molecules by a Pseudospectral Method", *J. Chem. Phys.* **86**, 3522 (1987).

(135) B. H. Greeley, T. V. Russo, D. T. Mainz, R. A. Friesner, J.-M. Langlois, W. A. Goddard III, R. E. Donelly, M. N. Ringnalda, "New Pseudospectral Algorithms for Electronic Structure Calculations: Length Scale Separation and Analytical Two-Electron Integral Corrections", *J. Chem. Phys.* **101**, 4028 (1994).

(136) S. Kossmann, F. Neese, "Comparison of Two Efficient Approximate Hartree-Fock Approaches", *Chem. Phys. Lett.* **481**, 240 (2009).

(137) F. Neese, F. Wennmohs, A. Hansen, U. Becker, "Efficient, Approximate and Parallel Hartree-Fock and Hybrid DFT Calculations. A "Chain-of-Spheres" Algorithm for the Hartree-Fock Exchange", *Chem. Phys.* **356**, 98 (2009).

(138) R. Iszák, F. Neese, "An Overlap Fitted Chain of Spheres Exchange Method", *J. Chem. Phys.* **135**, 144105 (2011).

(139) R. A. Kendall, H. A. Früchtl, "The Impact of the Resolution of the Identity Approximate Integral Method on Modern Ab Initio Algorithm Development", *Theor. Chem. Acc.* **97**, 158 (1997).

(140) H. A. Früchtl, R. A. Kendall, R. J. Harrison, K. G. Dyall, "An Implementation of RI-SCF on Parallel Computers", *Int. J. Quantum Chem.* **64**, 63 (1997).

(141) S. Hamel, M. E. Casida, D. R. Salahub, "Assessment of the Quality of Orbital Energies in Resolution-of-the-Identity Hartree-Fock Calculations Using deMon Auxiliary Basis Sets", *J. Chem. Phys.* **114**, 7342 (2001).

(142) S. Hamel, M. E. Casida, D. R. Salahub, "Exchange-Only Optimized Effective Potential for Molecules From Resolution-of-the-Identity Techniques: Comparison with Local Density Approximation", *J. Chem. Phys.* **116**, 8276 (2002).

(143) S. Reine, E. Tellgren, A. Krapp, T. Kjærgaard, T. Helgaker, B. Jansik, S. Høst, P. Salek, "Variational and Robust Density Fitting of Four-Center Two-Electron Integrals in Local Metrics", *J. Chem. Phys.* **129**, 104101 (2008).

(144) F. Aquilante, T. B. Pedersen, R. Lindh, "Low-Cost Evaluation of the Exchange Fock-Matrix From Cholesky and Density Fitting Representations of the Electron Repulsion Integrals", *J. Chem. Phys.* **126**, 194106 (2007).

(145) L. Boman, H. Koch, A. S. de Merás, "Method Specific Cholesky Decompositions: Coulomb and Exchange Energies", *J. Chem. Phys.* **129**, 154107 (2008).

(146) F. Aquilante, T. B. Pedersen, R. Lindh, "Density Fitting With Auxiliary Basis Sets From Cholesky Decompositions", *Theor. Chem. Acc.* **124**, 1 (2009).

(147) A. Sodt, M. Head-Gordon, "Hartree-Fock Exchange Computed Using the Atomic Resolution of the Identity Approximation", *J. Chem. Phys.* **128**, 104106 (2008).

(148) C. Fonseca Guerra, J. G. Snijders, G. te Velde, E. J. Baerends, "Towards and Order-N DFT Method", *Theor. Chem. Acc.* **99**, 391 (1998).

(149) M. A. Watson, N. C. Handy, A. J. Cohen, "Density Functional Calculations, Using Slater Basis Sets, With Exact Exchange", *J. Chem. Phys.* **119**, 6475 (2003).

(150) M. Krykunov, T. Ziegler, E. Van Lenthe, "Hybrid Density Functional Calculations of Nuclear Magnetic Shieldings Using Slater-Type Orbitals and the Zeroth-Order Regular Approximation", *Int. J. Quantum Chem.* **109**, 1676 (2009).

(151) P. Merlot, T. Kjærgaard, T. Helgaker, R. Lindh, F. Aquilante, S. Reine, T. B. Pedersen, "Attractive Electron-Electron Interactions Within Robust Local Fitting Approximations", *J. Comp. Chem.* **34**, 1486 (2013).

(152) S. F. Manzer, E. Epifanovsky, M. Head-Gordon, "Efficient Implementation of the Pair Atomic Resolution of the Identity Approximation for Exact Exchange for Hybrid and Range-Separated Density Functionals", *J. Chem. Theory Comput.* **11**, 518 (2015).

(153) D. S. Hollman, H. F. Schaefer, E. F. Valeev, "Semi-Exact Concetric Atomic Density Fitting: Reduced Cost and Increased Accuracy Compared to Standard Density Fitting", *J. Chem. Phys.* **140**, 064109 (2014).

(154) R. Polly, H.-J. Werner, F. R. Manby, P. J. Knowles, "Fast Hartree-Fock Theory Using Local Density Fitting Approximations", *Mol. Phys.* **102**, 2311 (2004).

(155) D. Mejia-Rodriguez, A. M. Köster, "Robust and Efficient Variational Fitting of Fock Exchange", *J. Chem. Phys.* **141**, 124114 (2014).

(156) D. Mejia-Rodriguez, X. Huang, J. M. del Campo, A. M. Köster, "Hybrid Functionals with Variationally Fitted Exact Exchange", *Adv. Quantum Chem.* **71**, 41 (2015).

(157) F. Aquilante, T. B. Pedersen, A. S. de Merás, H. Koch, "Fast Noniterative Orbital Localization for Large Molecules", *J. Chem. Phys.* **125**, 174101 (2006).

(158) S. F. Boys, "Construction of Some Molecular Orbitals to be Approximately Invariant for Changes From One Molecule to Another", *Rev. Mod. Phys.* **32**, 296 (1960).

(159) J. M. Foster, S. F. Boys, "Canonical Configurational Interaction Procedure", *Rev. Mod. Phys.* **32**, 300 (1960).

(160) C. Edmiston, K. Ruedenberg, "Localized Atomic and Molecular Orbitals", *Rev. Mod. Phys.* **35**, 457 (1963).

(161) J. Pipek, "Localization Measure and Maximum Delocalization in Molecular Systems", *Int. J. Quantum Chem.* **36**, 487 (1989).

(162) J. Pipek, P. G. Mezey, "A Fast Intrinsic Localization Procedure Applicable for *ab initio* and Semiempirical Linear Combination of Atomic Orbital Wave Functions", *J. Chem. Phys.* **90**, 4916 (1989).

(163) R. S. Mulliken, "Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I", *J. Chem. Phys.* **23**, 1833 (1955).

(164) J. E. Subotnik, A. Sodt, M. Head-Gordon, "Localized Orbital Theory and Ammonia Triborane", *Phys. Chem. Chem. Phys.* **9**, 5522 (2007).

(165) B. Jansík, S. Høst, K. Kristensen, P. Jørgensen, "Local Orbitals by Minimizing Powers of the Orbital Variance", *J. Chem. Phys.* **134**, 194101 (2011).

(166) K. Jug, E. Fasold, A. M. Köster, "Charge and Valence in Lithium Compounds", *Chem. Phys. Lett.* **188**, 294 (1992).

(167) K. Jug, A. M. Köster, "Nonsymmetrical Orthogonalization as a Consequence of Charge and Moment Conservation", *Chem. Phys. Lett.* **192**, 437 (1992).

(168) A. Bunse-Gerstner, R. Byers, V. Mehrmann, "Numerical Methods for Simultaneous Diagonalization", *SIAM J. Matrix Anal. Appl.* **14**, 927 (1993).

(169) J.-F. Cardoso, A. Souloumiac, "Jacobi Angles for Simultaneous Diagonalization", *SIAM J. Matrix Anal. Appl.* **17**, 161 (1996).

(170) A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines, "A Blind Source Separation Technique Using Second-Order Statistics", *IEEE T. Signal Proces.* **45**, 434 (1997).

(171) P.-O. Löwdin, "On the Nonorthogonality Problem", *Adv. Quantum Chem.* **5**, 185 (1970).

(172) P.-O. Löwdin, "On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals", *J. Chem. Phys.* **18**, 365 (1950).

(173) G. Bruhn, E. R. Davidson, I. Mayer, A. E. Clark, "Löwdin Population Analysis With and Without Rotational Invariance", *Int. J. Quantum Chem.* **106**, 2065 (2006).

(174) C. L. Jackins, S. L. Tanimoto, "Oct-trees and Their Use in Representing Three-Dimensional Objects", *Computer Graphics and Image Processing* **14**, 249 (1980).

(175) H. Samet, *The Design and Analysis of Spatial Data Structures*, Addison-Wesley, New York, USA (1989).

(176) S. Lefebvre, S. Hornus, F. Neyret, "Octree Textures on the GPU", in: *GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation* Pearson Education, New Jersey, USA (2005).

(177) M. Benzi, C. D. Meyer, M. Tuma, "A Sparse Approximate Inverse Preconditioner for the Conjugate Gradient Method", *SIAM J. Sci. Comput.* **17**, 1135 (1996).

(178) M. Benzi, M. Tuma, "A Sparse Approximate Inverse Preconditioner for Nonsymmetric Linear Systems", *SIAM J. Sci. Comput.* **19**, 968 (1998).

(179) M. Benzi, J. K. Cullum, M. Tuma, "Robust Approximate Inverse Preconditioning for the Conjugate Gradient Method", *SIAM J. Sci. Comput.* **22**, 1318 (2000).

(180) A. Rafiei, F. Toutounian, "New Breakdown-Free Variant of AINV Method for Nonsymmetric Positive Definite Matrices", *J. Comput. Appl. Math.* **219**, 72 (2008).

(181) J. U. Reveles Ramírez, "Geometry Optimization in LCGTO-DFT Methods with Auxiliary Functions", Ph. D. Thesis; Cinvestav (2004).

(182) J. M. del Campo Ramírez, "Exploring Chemical Reactivity With Auxiliary Density Functional Theory", Ph. D. Thesis; Cinvestav (2008).

(183) P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. Frisch, "Ab Initio Calculations of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields", *J. Chem. Phys.* **98**, 11623 (1994).

(184) J. P. Perdew, M. Ernzerhof, K. Burke, "Rationale for Mixing Exact Exchange with Density Functional Approximations", *J. Chem. Phys.* **105**, 9982 (1996).

(185) C. Adamo, V. Barone, "Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model", *J. Chem. Phys.* **110**, 6158 (1999).

(186) Y. Zhao, D. G. Truhlar, "The M06 Suite of Density Functionals for the Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited Stated, and Transition Elements: Two New Functionals and Systematic Testing for Four M06-Class Functionals and 12 Other Functionals", *Theor. Chem. Acc.* **120**, 215 (2007).

(187) M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, W. A. de Jong, "NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations", *Comput. Phys. Commun.* **181**, 1477 (2010).

(188) M. W. Chase Jr., *NIST-JANAF Thermochemical Tables*, American Institute of Physics, New York, USA (1998).

(189) A. D. Becke, "Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior", *Phys. Rev. A.* **38**, 3098 (1988).

(190) T. van Voorhis, G. E. Scuseria, "A Novel Form for the Exchange-Correlation Energy Functional", *J. Chem. Phys.* **109**, 400 (1998).

(191) T. van Voorhis, G. E. Scuseria, "Erratum: "A Novel Form for the Exchange-Correlation Energy Functional [J. Chem. Phys. 109, 400 (1998)]", *J. Chem. Phys.* **129**, 219901 (2008).

(192) J. Gräfenstein, D. Izotov, D. Cremer, "Avoiding Singularity Problems Associated with meta-GGA (Generalized Gradient Approximation) Exchange and Correlation Functionals Containing the Kinetic Energy Density", *J. Chem. Phys.* **127**, 214103 (2007).

(193) J. P. Perdew, S. Kurth, A. Zupan, P. Blaha, "Accurate Density Functional with Correct Formal Properties. A Step Beyond the Generalized Gradient Approximation", *Phys. Rev. Lett.* **82**, 2544 (1999).

(194) J. P. Perdew, S. Kurth, A. Zupan, P. Blaha, "Erratum: Accurate Density Functional with Correct Formal Properties. A Step Beyond the Generalized Gradient Approximation [Phys. Rev. Lett. 82, 2544 (1999)]", *Phys. Rev. Lett.* **82**, 5179 (1999).

(195) J. M. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, "Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids", *Phys. Rev. Lett.* **91**, 146401 (2003).

(196) L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, "Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies for Formation", *J. Chem. Phys.* **106**, 1063 (1997).

(197) L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, J. A. Pople, "Gaussian-3 Theory for Molecules Containing First and Second-Row Atoms", *J. Chem. Phys.* **109**, 7764 (1998).

(198) F. Weigend, R. Ahlrichs, "Balanced Basis Sets for Split Valence, Triple Zeta Valence and Quadrupule Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy", *Phys. Chem. Chem. Phys.* **7**, 3297 (2007).

(199) L. A. Curtiss, P. C. Redfern, K. Raghavachari, J. A. Pople, "Gaussian-3X (G3X) Theory: Use of Improved Geometries, Zero-Point Energies and Hartree-Fock Basis Sets.", *J. Chem. Phys.* **114**, 108 (2001).

(200) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, "Gaussian 09 Revision D.01", Gaussian Inc. Wallingford CT (2009).

(201) P. Calaminici, F. Janetzko, A. M. Köster, R. Mejía-Olvera, B. Zuniga-Gutierrez, "Density Functional Theory Optimized Basis Sets for Gradient Corrected Functionals: 3d Transition Metal Systems", *J. Chem. Phys.* **126**, 044108 (2007).

(202) Y. Zhao, B. J. Lynch, D. G. Truhlar, "Multi-Coefficient Extrapolated Density Functional Theory for Thermochemistry and Thermochemical Kinetics", *Phys. Chem. Chem. Phys.* **7**, 43 (2005).

(203) Y. Zhao, N. E. Schultz, D. G. Truhlar, "Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics and Noncovalent Interactions", *J. Chem. Theory Comput.* **2**, 364 (2006).

(204) J. Zheng, Y. Zhao, D. G. Truhlar, "The DBH24/08 Database and Its Use to Assess Electronic Structure Model Chemistries for Chemical Reactions Barrier Heights", *J. Chem. Theory Comput.* **5**, 808 (2009).

(205) R. Peverati, D. G. Truhlar, "Quest for a Universal Density Functional: The Accuracy of Density Functionals Across a Broad Spectrum of Databases in Chemistry and Physics", *Philos. Trans. R. Soc. A* **372**, 20120476 (2014).

(206) S. E. Wheeler, K. N. Houk, "Integration Grid Errors for Meta-GGA-Predicted Reaction Energies: Origin of Grid Errors for the M06 Suite of Functionals", *J. Chem. Theory Comput.* **6**, 395 (2010).

(207) F. Weigend, "Hartree-Fock Exchange Fitting Basis Sets for H to Rn", *J. Comp. Chem.* **29**, 167 (2007).

(208) W. J. Hehre, R. Ditchfield, J. A. Pople, "Self–Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian–Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules", *J. Chem. Phys.* **56**, 2257 (1972).

(209) P. C. Hariharan, J. A. Pople, "The Effect of Polarization Functions on Molecular Orbital Hydrogenation Energies", *Theor. Chim. Acta* **28**, 213 (1973).

(210) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, "General Atomic and Molecular Electronic Structure System", *J. Comput. Chem.* **14**, 1347 (1993).

(211) E. O. Steinborn, K. Ruedenberg, "Rotation and Translation of Regular and Irregular Solid Spherical Harmonics", *Adv. Quantum Chem.* **7**, 1 (1973).

(212) L. Greengard, *The Rapid Evaluation of Potential Fields in Particle Systems*, MIT, Cambridge, USA (1987).

(213) C. H. Choi, J. Ivanic, M. S. Gordon, K. Ruedenberg, "Rapid and Stable Determination of Rotation Matrices Between Spherical Harmonics by Direct Recursion", *J. Chem. Phys.* **111**, 8825 (1999).

(214) C. H. Choi, K. Ruedenberg, M. S. Gordon, "New Parallel Optimal-Parameter Fast Multipole Method", *J. Comput. Chem.* **22**, 1484 (2001).

(215) C. H. Choi, "Direct Determination of Multipole Moments of Cartesian Gaussian Functions in Spherical Polar Coordinates", *J. Chem. Phys.* **120**, 3535 (2004).

(216) T. H. Dunning Jr., "Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen", *J. Chem. Phys.* **90**, 1007 (1989).

(217) M. Guidon, J. Hutter, J. VandeVondele, "Auxiliary Density Matrix Methods for Hartree-Fock Exchange Calculations", *J. Chem. Theory Comput.* **6**, 2348 (2010).

(218) P. Merlot, R. Izsák, A. Borgoo, T. Kjærgaard, T. Helgaker, S. Reine, "Charge-Constrained Auxiliary-Density-Matrix Methods for the Hartree-Fock Exchange Contribution", *J. Chem. Phys.* **141**, 094104 (2014).

(219) J. Gerrat, I. M. Mills, "Force Constants and Dipole-Moment Derivatives of Molecules from Perturbed Hartree-Fock Calculations. I", *J. Chem. Phys.* **49**, 1719 (1968).

(220) J. A. Pople, R. Krishnan, H. B. Schlegel, J. S. Binkley, "Derivative Studies in Hartree-Fock and Møller-Plesset Theories", *Int. J. Quantum Chem. Symp.* **16**, 225 (1979).

(221) R. Fournier, "Second and Third Derivatives of the Linear Combination of Gaussian Type Orbitals-Local Spin Density Energy", *J. Chem. Phys.* **92**, 5422 (1990).

(222) S. P. Karna, M. Dupuis, "Frequency Dependent Nonlinear Optical Properties of Molecules: Formulation and Implementation in the HONDO Program", *J. Comput. Chem.* **12**, 487 (1991).

(223) A. Komornicki, G. Fitzgerald, "Molecular Gradients and Hessians Implemented in Density Functional Theory", *J. Chem. Phys.* **98**, 1398 (1993).

(224) S. M. Colwell, C. W. Murray, N. C. Handy, R. D. Amos, "The Determination of Hyperpolarisabilities Using Density Functional Theory", *Chem. Phys. Lett.* **210**, 261 (1993).

(225) A. M. Lee, S. M. Colwell, "The Determination of Hyperpolarizabilities Using Density Functional Theoy with Nonlocal Functionals", *J. Chem. Phys.* **101**, 9704 (1994).

(226) Y. Yamaguchi, Y. Osamura, J. D. Goddard, H. F. Schaefer III, *A New Dimensional to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*, Oxford University Press, New York, USA (1994).

(227) R. McWeeny, "Perturbation Theory for the Fock-Dirac Density Matrix", *Phys. Rev.* **126**, 1028 (1962).

(228) G. Diercksen, R. McWeeny, "Self-Consistent Pertubation Theory. I. General Formulation and Some Applications", *J. Chem. Phys.* **44**, 3554 (1966).

(229) R. McWeeny, G. Diercksen, "Self-Consistent Perturbation Theory. II. Extension to Open Shells", *J. Chem. Phys.* **49**, 4852 (1968).

(230) J. L. Dodds, R. Mcweeny, A. J. Sadlej, "Self-Consistent Perturbation Theory. Generalization for Perturbation-Dependent Non-Orthogonal Basis Set", *Mol. Phys.* **34**, 1779 (1977).

(231) J. L. Dodds, R. McWeeny, W. T. Raynes, J. P. Riley, "SCF Theory for Multiple Perturbations", *Mol. Phys.* **33**, 611 (1977).

(232) R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press, London, UK (2001).

(233) C. Ochsenfeld, M. Head-Gordon, "A Reformulation of the Coupled Perturbed Self-Consistent Field Equations Entirely within a Local Atomic Orbital Density Matrix-Based Scheme", *Chem. Phys. Lett.* **270**, 399 (1997).

(234) V. Weber, A. M. N. Niklasson, M. Challacombe, "*Ab Initio* Linear Scaling Response Theory: Electric Polarizability by Perturbed Projection", *Phys. Rev. Lett.* **92**, 193002 (2004).

(235) V. Weber, A. M. Niklasson, M. Challacombe, "High-Order Response in $\mathcal{O}(N)$ by Perturbed Projection", *J. Chem. Phys.* **123**, 044106 (2005).

(236) A. M. Niklasson, V. Weber, "Linear Scaling Density Matrix Perturbation Theory for Basis-Set-Dependent Quantum Response Calculations: An Orthogonal Formulation", *J. Chem. Phys.* **127**, 064105 (2007).

(237) J. Kussmann, C. Ochsenfeld, "A Density Matrix-Based Method for the Linear-Scaling Calculation of Dynamic Second- and Third-Order Properties at the Hartree-Fock and Kohn-Sham Density Functional Theory Levels", *J. Chem. Phys.* **127**, 204103 (2007).

(238) S. Coriani, S. Høst, B. Jansik, L. Tøgersen, J. Olsen, P. Jørgensen, S. Reine, F. Pawłowski, T. Helgaker, P. Sałek, "Linear-Scaling Implementation of Molecular Response Theory in Self-Consistent Field Electronic-Structure Theory", *J. Chem. Phys.* **126**, 154108 (2007).

(239) M. Beer, C. Ochsenfeld, "Efficient Linear-Scaling Calculation of Response Properties: Density Matrix-Based Laplace-Transformed Coupled-Perturbed Self-Consistent Field Theory", *J. Chem. Phys.* **128**, 221102 (2008).

(240) T. Kjærgaard, P. Jørgensen, J. Olsen, S. Coriani, T. Helgaker, "Hartree-Fock and Kohn-Sham Time-Dependent Response Theory in a Second-Quantization Atomic-Orbital Formalism Suitable for Linear Scaling", *J. Chem. Phys.* **129**, 054106 (2008).

(241) M. Kobayashi, T. Touma, H. Nakai, "Dynamic Hyperpolarizability Calculations of Large Systems: The Linear-Scaling Divide-And-Conquer Approach", *J. Chem. Phys.* **136**, 084108 (2012).

(242) R. Flores-Moreno, A. M. Köster, "Auxiliary Density Perturbation Theory", *J. Chem. Phys.* **128**, 134105 (2008).

(243) J. Carmona-Espíndola, R. Flores-Moreno, A. M. Köster, "Time-Dependent Auxiliary Density Perturbation Theory", *J. Chem. Phys.* **133**, 084102 (2010).

(244) K. B. Sophy, S. Pal, "Density Functional Response Approach for the Linear and Nonlinear Electric Properties of Molecules", *J. Chem. Phys.* **118**, 10861 (2003).

(245) K. B. Sophy, S. Pal, "Electric Properties of BH, CO and H_2O Molecules by Density Functional Response Approach", *J. Mol. Struct.: THEOCHEM* **676**, 89 (2004).

(246) K. B. Sophy, P. Calaminici, S. Pal, "Density Functional Static Dipole Polarizability and First-Hyperpolarizability Calculations of Na_n ($n = 2, 4, 6, 8$) Clusters Using an Approximate CPKS Method and its Comparison with MP2 Calculations", *J. Chem. Theory Comput.* **3**, 716 (2007).

(247) K. B. Sophy, S. V. Shedge, S. Pal, "Noniterative Density Functional Response Approach: Application to Nonlinear Optical Properties of *p*-Nitroanaliline and its Methyl-Substituted Derivatives", *J. Phys. Chem. A* **112**, 11266 (2008).

(248) S. V. Shedge, J. Carmona-Espíndola, S. Pal, A. M. Köster, "Comparison of the Auxiliary Density Perturbation Theory and the Noniterative Approximation to the Coupled Perturbed Kohn-Sham Method: Case Study of the Polarizabilities of Disubstituted Azoarene Molecules", *J. Phys. Chem. A* **114**, 2357 (2010).

(249) S. V. Shedge, S. Pal, A. M. Köster, "Validation and Application of Auxiliary Density Perturbation Theory and Non-Iterative Approximation to Coupled-Perturbed Kohn-Sham Approach for Calculation of Dipole-Quadrupole Polarizability", *Chem. Phys. Lett.* **510**, 185 (2011).

(250) B. I. Dunlap, I. V. Schweigert, "Self-Consistent, Constrained Linear-Combination-of-Atomic-Potentials Approach to Quantum Mechanics", *J. Chem. Phys.* **134**, 044122 (2011).

(251) R. Flores Moreno, "Analytic Derivatives in LCGTO-DFT Pseudo-Potential Methods with Auxiliary Functions", Ph. D. Thesis; Cinvestav (2006).

(252) J. Carmona-Espíndola, R. Flores-Moreno, A. M. Köster, "Static and Dynamic First Hyperpolarizabilities From Time-Dependent Auxiliary Density Perturbation Theory", *Int. J. Quantum Chem.* **112**, 3461 (2012).

(253) G. Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, Wellesley, USA (2003).

(254) W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Fortran Numerical Recipes*, Cambridge University Press, Cambridge, UK (1992).

(255) D. Mejia-Rodriguez, R. I. Delgado Venegas, P. Calaminici, A. M. Köster, "Robust and Efficient Auxuliary Density Perturbation Theory Calculations", *J. Chem. Theory Comput.* **11**, 1493 (2015).

(256) A. Álvarez Ibarra, A. M. Köster, "A New Mized Self-Consistent Field Procedure", *Mol. Phys.* (2015), doi:10.1080/00268976.2015.1078009.

(257) H. A. van der Vorst, "Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems", *SIAM J. Sci. Stat. Comp.* **13**, 631 (1992).

(258) P. Sonneveld, "CGS, A Fast Lanczos-Type Solver for Nonsymmetric Linear Systems", *SIAM J. Sci. Stat. Comp.* **10**, 36 (1989).

(259) R. W. Freund, N. M. Nachtigal, "QMR: A Quasi-Minimal Residual Method for Non-hermitian Linear Systems", *Numer. Math.* **60**, 315 (1991).

(260) Y. Saad, M. Schultz, "GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems", *SIAM J. Sci. Stat. Comp.* **7**, 856 (1986).

(261) T. Eirola, O. Nevanlinna, "Accelerating with Rank-One Updates", *Linear Algebra Appl.* **121**, 511 (1989).

(262) C. Vuik, H. A. van der Vorst, "A Comparison of Some GMRES-like Methods", *Linear Algebra Appl.* **160**, 131 (1992).

(263) U. M. Yang, K. A. Gallivan, "A New Family of Preconditioned Iterative Solvers for Nonsymmetric Linear Systems", *Appl. Numer. Math.* **19**, 287 (1995).

(264) W. Gao, J. Xue, Y. Qu, "A New Implementation of EN method", *Appl. Math. Comput.* **98**, 199 (1999).

(265) Z. Zlatev, K. Georgiev, "Applying Approximate LU-Factorization as Preconditioners in Eight Iterative Methods for Solving Systems of Linear Equations", *Cent. Eur. J. Math.* **11**, 1510 (2013).

(266) G. H. Golub, C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, USA (2012).

(267) W. W. Hager, "Updating the Inverse of a Matrix", *SIAM Rev.* **31**, 221 (1989).

(268) D. R. Lide, *CRC Handbook of Chemistry and Physics*, CRC Press, Florida, USA (2005).

(269) N. Godbout, D. R. Salahub, J. Andzelm, E. Wimmer, "Optimization of Gaussian-Type Basis Sets for Local Spin Density Functional Calculations. Part I. Boron Through Neon, Optimization Technique and Validation", *Can. J. Chem.* **70**, 560 (1992).

(270) P. Calaminici, K. Jug, A. M. Köster, "Density Functional Calculations of Molecular Polarizabilities and Hyperpolarizabilities", *J. Chem. Phys.* **109**, 7756 (1998).

(271) P. Calaminici, K. Jug, A. M. Köster, V. E. Ingamells, M. G. Papadopoulos, "Polarizabilities of Azabenzenes", *J. Chem. Phys.* **112**, 6301 (2000).

(272) P. Calaminici, "Density Functional Calculations of Molecular Electric Properties in Iron Containing Systems", *Chem. Phys. Lett.* **374**, 650 (2003).

(273) A. J. Perkins, "The Refractive Index of Anhydrous Hydrogen Flouride", *J. Phys. Chem.* **68**, 654 (1964).

(274) D. M. Bishop, L. M. Cheung, "Vibrational Contributions to Molecular Dipole Polarizabilities", *J. Phys. Chem. Ref. Data* **11**, 119 (1982).

(275) M. A. Spackman, "Accurate Prediction of Static Dipole Polarizabilities with Moderately Sized Basis Sets", *J. Phys. Chem.* **93**, 7594 (1989).

(276) G. R. Alms, A. K. Burnham, W. H. Flygare, "Measurement of the Dispersion in Polarizability Anisotropies", *J. Chem. Phys.* **63**, 3321 (1975).

(277) C. K. Miller, B. J. Orr, J. F. Ward, "An Interacting Segment Model of Molecular Electric Tensor Properties: Application to Dipole Moments, Polarizabilities, and Hyperpolarizabilities for the Halogenated Methanes", *J. Chem. Phys.* **74**, 4858 (1981).

(278) N. J. Bridge, A. D. Buckingham, "The Polarization of Laser Light Scattered by Gases", *Proc. R. Soc. Lond. A* **295**, 334 (1966).

(279) M. P. Bogaard, A. D. Buckingham, R. K. Pierens, A. H. White, "Rayleigh Scattering Depolarization Ratio and Molecular Polarizability Anisotropy for Gases", *J. Chem. Soc., Faraday Trans. 1* **74**, 3008 (1978).

(280) M. P. Bogaard, A. D. Buckingham, G. L. D. Ritchie, "Temperature Dependence of the Kerr Effect of Hydrogen Sulphide", *Chem. Phys. Lett.* **90**, 183 (1982).

(281) W. F. Murphy, "The Rayleigh Depolarization Ratio and Rotational Raman Spectrum of Water Vapor and the Polarizability Components for the Water Molecule", *J. Chem. Phys.* **67**, 5877 (1977).

(282) W. D. Knight, K. Clemenger, W. A. de Heer, W. A. Saunders, M. Y. Chou, M. L. Cohen, "Electronic Shell Structure and Abundances of Sodium Clusters", *Phys. Rev. Lett.* **52**, 2141 (1984).

(283) W. Ekardt, "Dynamical Polarizability of Small Metal Particles: Self-Consistent Spherical Jellium Background Model", *Phys. Rev. Lett.* **52**, 1925 (1984).

(284) W. Ekardt, "Work Function of Small Metal Particles: Self-Consistent Spherical Jellium Background Model", *Phys. Rev. B* **29**, 1558 (1984).

(285) K. Jug, B. Zimmermann, A. M. Köster, "Growth Pattern and Bonding of Copper Clusters", *Int. J. Quantum Chem.* **90**, 594 (2002).

(286) A. W. Castleman, S. N. Khanna, "Clusters, Superatoms, and Building Blocks of New Materials", *J. Phys. Chem. C* **113**, 2664 (2009).

(287) A. M. Köster, P. Calaminici, E. Orgaz, D. R. Roy, J. U. Reveles, S. N. Khanna, "On the Ground State of Pd_{13}", *J. Am. Chem. Soc.* **133**, 12192 (2011).

(288) J. U. Reveles, A. M. Köster, P. Calaminici, S. N. Khanna, "Structural Changes in Pd_{13} Upon Charging and Oxidation/Reduction", *J. Chem. Phys.* **136**, 114505 (2012).

(289) V. M. Medel, A. C. Reder, V. Chankan, P. Sen, A. M. Köster, P. Calaminici, S. N. Khanna, "Nature of Valence Transition and Spin Moment in Ag_nV^+", *J. Am. Chem. Soc.* **136**, 8229 (2014).

(290) M. Brack, "The Physics of Simple Metal Clusters: Self-Consistent Jellium Model and Semiclassical Approaches", *Rev. Mod. Phys.* **65**, 677 (1993).

(291) I. Katakuse, T. Ichihara, Y. Fujita, T. Matsuo, T. Sakurai, H. Matsuda, "Mass Distributions of Copper, Silver and Gold Clusters and Electronic Shell Structure", *Int. J. Mass Spectrom.* **67**, 229 (1985).

(292) M. Y. Chou, M. L. Cohen, "Electronic Shell Structure in Simple Metal Clusters", *Phys. Lett.* **113A**, 420 (1986).

(293) I. Katakuse, T. Ichihara, Y. Fujita, T. Matsuo, T. Sakurai, H. Matsuda, "Mass Distributions of Negative Cluster Ions of Copper, Silver and Gold", *Int. J. Mass Spectrom.* **74**, 33 (1986).

(294) G. Alameddin, J. Hunter, D. Cameron, M. M. Kappes, "Electronic And Geometric Structure in Silver Clusters", *Chem. Phys. Lett.* **192**, 122 (1992).

(295) W. A. de Heer, "The Physic of Simple Metal Clusters: Experimental Aspects and Simple Models", *Rev. Mod. Phys.* **65**, 611 (1993).

(296) H. Nishioka, "Shells and Supershells in Metal Clusters", *Z. Phys. D* **19**, 19 (1991).

(297) F. Weigend, R. Ahlrichs, "Quantum Chemical Treatments of Metal Clusters", *Philos. Trans. R. Soc. A* **368**, 1245 (2010).

(298) J. M. Vásquez-Pérez, G. U. Gamboa, D. Mejía-Rodríguez, A. Álvarez Ibarra, G. Geudtner, P. Calaminici, A. M. Köster, "Influence of Spin Multiplicity on the Melting of Na_{55}^{+}", *J. Phys. Chem. Lett.* **6**, 4646 (2015).

(299) H. A. Jahn, E. Teller, "Stability of Polyatomic Molecules in Degenerate Electronic States. I. Orbital Degeneracy", *Proc. R. Soc. London Ser. A* **161**, 220 (1937).

(300) Z. Luo, A. W. Castleman, "Special and General Superatoms", *Acc. Chem. Res.* **47**, 2931 (2014).

(301) H. Haberland, T. Hippler, J. Donges, O. Kostko, M. Schmidt, B. von Issendorff, "Melting of Sodium Clusters: Where Do the Magic Numbers Come From?", *Phys. Rev. Lett.* **94**, 035701 (2005).

(302) S. Zamith, F. Chirot, J.-M. L'Hermite, "A Two-State Model Analysis of the Melting of Sodium Clusters: Insights in the Enthalpy-Entropy Compensation", *Eur. Phys. Lett.* **92**, 13004 (2010).

(303) T. P. Martin, T. Bergmann, H. Göhlich, T. Lange, "Shell Structure of Clusters", *J. Phys. Chem.* **95**, 6421 (1991).

(304) G. Wrigge, M. Astruc Hoffmann, B. von Issendorff, "Photoelectron Spectroscopy of Soidum Clusters: Direct Observation of the Electronic Shell Structure", *Phys. Rev. A* **65**, 063201 (2002).

(305) J. L. Martins, J. Buttet, R. Car, "Electronic and Structural Properties of Sodium Clusters", *Phys. Rev. B* **31**, 1804 (1985).

(306) U. Röthlisberger, W. Andreoni, "Structural and Electronic Properties of Sodium Microclusters ($n = 2 - 20$) at Low and High Temperatures: New Insights From *ab initio* Molecular Dynamics Studies", *J. Chem. Phys.* **94**, 8129 (1991).

(307) U. Röthlisberger, W. Andreoni, P. Giannozzi, "Thirteen-atom clusters: Equilibrium Geometries, Structural Transformations, and Trends in Na, Mg, Al, and Si", *J. Chem. Phys.* **96**, 1248 (1992).

(308) I. A. Solov'yov, A. V. Solov'yov, W. Greiner, "Structure and Properties of Small Sodium Clusters", *Phys. Rev. A* **65**, 053203 (2002).

(309) A. Aguado, O. Kostko, "First-Principles Determination of the Structure of Na_N and Na_N^- Clusters with up to 80 Atoms", *J. Chem. Phys.* **134**, 164304 (2011).

(310) J. Bowlan, A. Liang, W. A. de Heer, "How Metallic are Small Sodium Clusters?", *Phys. Rev. Lett.* **106**, 043401 (2011).

(311) A. L. Mackay, "A Dense Non-Crystallographic Packing of Equal Spheres", *Acta Crystallogr.* **15**, 916 (1962).

(312) M. Reiher, O. Salomon, B. A. Hess, "Reparametrization of Hybrid Functionals Based on Energy Differences of States of Different Multiplicity", *Theor. Chem. Acc.* **107**, 45 (2001).

(313) M. Radoń, "Revisiting the Role of Exact Exchange in DFT Spin-State Energetics of Transition Metal Complexes", *Phys. Chem. Chem. Phys.* **16**, 14479 (2014).

(314) G. Geudtner, V. D. Domínguez-Soria, P. Calaminici, A. M. Köster, "Molecular Graphs of Li_n, Na_n and Cu_n (n =6-9) Clusters from the Density and Molecular Electrostatic Potential", *Comp. Theor. Chem.* **1053**, 337 (2015).

(315) K. Jug, B. Zimmermann, P. Calaminici, A. M. Köster, "Structure and Stability of Small Copper Clusters", *J. Phys. Chem.* **116**, 4497 (2002).

(316) J. U. Reveles, P. A. Clayborne, A. C. Reber, S. N. Khanna, K. Pradhan, P. Sen, M. R. Pederson, "Designer Magnetic Superatoms", *Nature Chemistry* **1**, 310 (2009).

(317) M. Swart, "Metal-Ligand Bonding in Metallocenes: Differentation Between Spin State, Electrostatic and Covalent Bonding", *Inorg. Chim. Acta* **360**, 179 (2007).

(318) C. J. Cramer, D. G. Truhlar, "Density Functional Theory for Transition Metals and Transition Metal Chemistry", *Phys. Chem. Chem. Phys.* **11**, 10757 (2009).

(319) F. Furche, J. P. Perdew, "The Performance of Semilocal and Hybrid Density Functionals in 3*d* Transition-Metal Chemistry", *J. Chem. Phys.* **124**, 044103 (2006).

(320) B. S. Narendrapurapu, N. A. Richardson, A. V. Copan, M. L. Estep, Z. Yang, H. F. Schaefer III, "Investigating the Effects of Basis Set on Metal-Metal and Metal-Ligand Bond Distances in Stable Transition Metal Carbonyls: Performance of Correlation Consistent Basis Sets with 35 Denstity Functionals", *J. Chem. Theory Comput.* **9**, 2930 (2013).

(321) J. N. Harvey, "On the Accuracy of Density Functional Theory in Transition Metal Chemistry", *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **102**, 203 (2006).

(322) M. Swart, A. R. Groenhorf, A. W. Ehlers, K. Lammertsma, "Validation of Exchange-Correlation Functionals for Spin States of Iron Complexes", *J. Phys. Chem. A* **108**, 5479 (2004).

(323) X. Xu, W. Zhang, M. Tang, D. G. Truhlar, "Do Practical Standard Coupled Cluster Calculations Agree Better than Kohn-Sham Calculations with Currently Available Functionals When Compared to the Best Available Experimental Data for Dissociation Energies of Bonds to 3*d* Transition Metals?", *J. Chem. Theory Comput.* **11**, 2036 (2015).

(324) R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, "Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions", *J. Chem. Phys.* **96**, 6796 (1992).

(325) D. E. Woon, T. H. Dunning Jr., "Gaussian-Basis Sets for Use in Correlated Molecular Calculations. 3. The Atoms Aluminum Through Argon", *J. Chem. Phys.* **98**, 1358 (1993).

(326) K. A. Peterson, D. E. Woon, T. H. Dunning Jr., "Benchmark Calculations With Correlated Molecular Wave Functions. IV. The Classical Barrier Height of the H+H2 –> H2 + H Reaction", *J. Chem. Phys.* **100**, 7410 (1994).

(327) A. K. Wilson, T. van Mourik, T. H. Dunning Jr., "Gaussian Basis Sets for Use in Correlated Molecular Calculations. VI. Sextuple Zeta Correlation Consistent Basis Sets for Boron Through Neon", *J. Mol. Struct. THEOCHEM* **388**, 339 (1996).

(328) M. I. Davis, C. S. Speed, "Gas-Phase Electron Diffraction Studies of Some Iron Carbonyl Complexes", *J. Organomet. Chem.* **21**, 401 (1970).

(329) B. J. Drouin, S. G. Kukolich, "Microwave Spectra and the Molecular Structure of Tetracarbonylethyleneiron", *J. Am. Chem. Soc.* **121**, 4023 (1999).

(330) A. P. Cox, A. H. Brittain, "Microwave Spectrum and Structure of Cyclopentadienyl Nickel Nitrosyl", *Trans. Faraday Soc.* **66**, 557 (1970).

(331) K. Suryanarayana, B. P. Stoicheff, R. Turner, "High Resolution Raman Spectroscopy of Gases: XIII. Rotational Spectra and Structures of the Zinc-, Cadmium-, and Mercury-Dimethyl Molecules", *Can. J. Phys.* **38**, 1516 (1960).

(332) A. Haaland, S. Samdal, R. Seip, "The Molecular Structure of Monomeric Methyl-(Cyclopentadienyl)Zinc, $(CH_3)Zn(\eta\text{-}C_5H_5)$, Determined by Gas Phase Electron Diffraction", *J. Organomet. Chem.* **153**, 187 (1978).

(333) N.-S. Chiu, L. Schäfer, R. Seip, "On Internal Rotation in Gaseous Benzenechromium Tricarbonyl", *J. Organomet. Chem.* **101**, 331 (1975).

(334) V. M. Rayon, G. Frenking, "Bis(benzene)chromium is a δ-Bonded Molecule and Ferrocene is a π-Bonded Molecule", *Organometallics* **22**, 3304 (2003).

(335) B. Beagley, C. T. Parrott, V. Ulbrecht, G. G. Young, "A Gas-Phase Electron Diffraction Study of the Molecular Structure of Cyclopentadienyl Cobalt Dicarbonyl, $C_5H_5Co(CO)_2$", *J. Mol. Struct.* **52**, 47 (1979).

(336) I. W. Bassi, G. Allegra, R. Scordamaglia, G. Chioccola, "Crystal Structure of an Organometallic Complex with Titanium-Carbon σ Bonds. Tetrabenzyltitanium", *J. Am. Chem. Soc.* **93**, 3787 (1971).

(337) P. Briant, J. Green, A. Haaland, H. Moellendal, K. Rypdal, J. Tremmel, "The Methyl Group Geometry in Trichloromethyltitanium: A Reinvestigation by Gas Electron Diffraction", *J. Am. Chem. Soc.* **111**, 3434 (1989).

(338) J. H. Calloman, E. Hirota, T. Ijima, K. Kuchitsu, W. J. Lafferty, *Structure Data of Free Polyatomic Molecules*, Springer-Verlag, Berlin, Germany (1976).

(339) A. F. Berndt, R. F. Marsh, "The Crystal Structure of Cyclopentadienyl Manganese Tricarbonyl", *Acta Cryst.* **16**, 118 (1963).

(340) Z.-F. Zu, Y. Xie, W.-L. Feng, H. F. Schaefer III, "Systematic Investigation of Electronic and Molecular Structures for the First Transition Metal Series Metallocenes M(C_5H_5) (M = V, Cr, Mn, Fe, Co, and Ni)", *J. Phys. Chem. A* **107**, 2716 (2003).

(341) D. Cozak, F. Gauvin, "Electronic Ground States and Isotropic Proton NMR Shifts of Manganocene and Its Derivatives", *Organometallics* **6**, 1912 (1987).

(342) J. W. Rabalais, L. O. Werme, T. Bergmark, L. Karlsson, M. Hussain, K. Siegbahn, "Electron Spectroscopy of Open-Shell Systems: Spectra of Ni(C_5H_5)$_2$, Fe(C_5H_5)$_2$, Mn(C_5H_5)$_2$ and Cr(C_5H_5)$_2$", *J. Chem. Phys.* **57**, 1185 (1972).

(343) J. H. Ammeter, R. Bucher, N. Oswald, "High-Spin-Low-Spin Equilibrium of Manganocene and Dimethylmaganocene", *J. Am. Chem. Soc.* **96**, 7833 (1974).

(344) J. H. Ammeter, "EPR of Orbitally Degenerate Sandwich Compunds", *J. Magn. Reson.* **30**, 299 (1978).

(345) J. C. Smart, J. L. Robbins, "A Low Spin Manganocene and Its Novel Anionic Derivative. Synthesis and Characterization of Decamethylmanganocene Complexes", *J. Am. Chem. Soc.* **100**, 3936 (1978).

(346) Y. Garcia, P. Gütlich, "Thermal Spin Crossover in Mn(II), Mn(III), Cr(II) and Co(III) Coordination Compounds", in: *Spin Crossover in Transition Metal Compounds II* Springer, Berlin, Germany (2004).

(347) A. Ruiz, J. M. Gomez-Llorente, "Electronic Structure and Polarizabilities of Icosahedral Fullerenes: A Pariser-Parr-Pople Approach", *J. Chem. Phys.* **114**, 1272 (2001).

(348) G. K. Gueorguiev, J. M. Pacheco, D. Tománek, "Quantum Size Effects in the Polarizability of Carbon Fullerenes", *Phys. Rev. A* **92**, 215501 (2004).

(349) R. R. Zope, T. Baruah, M. R. Pederson, B. I. Dunlap, "Static Dieletric Response of Icosahedral Fullerenes From C_{60} to C_{2160} Characterized by an All-Electron Density Functional Theory", *Phys. Rev. B* **77**, 115452 (2008).

(350) D. Rappoport, F. Furche, "Property-Optimized Gaussian Basis Sets for Molecular Response Calculations", *J. Chem. Phys.* **133**, 134105 (2010).

(351) P. Calaminici, J. Carmona-Espíndola, G. Geudtner, A. M. Köster, "Static and Dynamic Polarizability of C_{540} Fullerene", *Int. J. Quantum Chem.* **112**, 3252 (2012).

(352) G. R. Lewis, W. E. Bunting, R. R. Zope, B. I. Dunlap, J. C. Ellenbogen, "Smooth Scaling of Valence Electronic Properties in Fullerenes: From One Carbon Atom, to C_{60} to Graphene", *Phys. Rev. A* **87**, 052515 (2013).

(353) R. Antoine, P. Dugourd, D. Rayane, E. Benichou, M. Broyer, F. Chanderzon, C. Guet, "Direct Measurement of the Electric Polarizability of Isolated C_{60} Molecules", *J. Chem. Phys.* **110**, 9771 (1999).

(354) A. Ballard, K. Bonin, J. Louderback, "Absolute Measurement of the Optical Polarizability of C_{60}", *J. Chem. Phys.* **113**, 5732 (2000).

(355) M. Berninger, A. Stefanov, S. Deachapunya, M. Arndt, "Polarizability Measurements of a Molecule via a Near-Field Matter-Wave Interferometer", *Phys. Rev. A* **76**, 013607 (2007).

(356) L. Hackermüller, K. Hornberger, S. Gerlich, M. Gring, H. Ulbricht, "Optical Polarizabilities of Large Molecules Measured in Near-Field Interferometry", *Appl. Phys. B* **89**, 469 (2007).

(357) W. Kutzelnigg, "Theories of Magnetic Susceptibilities and NMR Chemical Shifts in Terms of Localized Quantities", *Israel J. Chem.* **19**, 193 (1980).

(358) R. Flores-Moreno, J. Melin, J. V. Ortiz, G. Merino, "Efficient Evaluation of Analytic Fukui Functions", *J. Chem. Phys.* **129**, 224105 (2008).

(359) J. S. Binkley, J. A. Pople, W. J. Hehre, "Self-Consistent Molecular Orbital Methods. 21. Small Split-Valence Basis Sets for First-Row Elements", *J. Am. Chem. Soc.* **102**, 939 (1980).

(360) M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, W. J. Hehre, "Self-Consistent Molecular-Orbital Methods. 22. Small Split-Valence Basis Sets for Second-Row Elements", *J. Am. Chem. Soc.* **104**, 2797 (1983).

(361) E. Engel, S. H. Vosko, "Exact Exchange-Only Potentials and the Virial Relation as Microscopic Criteria for Generalized Gradient Approximations", *Phys. Rev. B* **47**, 13164 (1993).

(362) P. Otto, J. Ladik, "Intermolecular and Intramolecular Interactions Calculated with *Ab Initio* Perturbative Configuration Interaction Method Using Strongly Localized Orbitals", *Int. J. Quantum Chem.* **22**, 169 (1982).

(363) P. Pulay, "Localizability of Electron Correlation", *Chem. Phys. Lett.* **100**, 151 (1983).

(364) S. Saebø, P. Pulay, "Local Configuration Interaction: An Efficient Approach for Larger Molecules", *Chem. Phys. Lett.* **113**, 13 (1985).

(365) P. Pulay, S. Saebø, "Orbital Invariant Formulation and Second-Order Gradient Evaluation in Møller-Plesset Perturbation Theory", *Theor. Chim. Acta* **69**, 357 (1986).

(366) S. Saebø, "Strategies for Electron Correlation Calculations on Large Molecular Systems", *Int. J. Quantum Chem.* **42**, 217 (1992).

(367) S. Saebo, P. Pulay, "Local Treatment of Electron Correlation", *Annu. Rev. Phys. Chem.* **44**, 213 (1993).

(368) C. Møller, M. S. Plesset, "Note on an Approximation Treatment for Many-Electron Systems", *Phys. Rev.* **46**, 618 (1934).

(369) M. Schütz, G. Hetzer, H.-J. Werner, "Low-Order Scaling Local Electron Correlation Methods. I. Linear Scaling Local MP2", *J. Chem. Phys.* **111**, 5691 (1999).

(370) G. Hetzer, M. Schütz, H. Stoll, H.-J. Werner, "Low-Order Scaling Local Correlation Methods II: Splitting the Coulomb Operator in Linear Scaling Local Second-Order Møller-Plesset Perturbation Theory", *J. Chem. Phys.* **113**, 9443 (2000).

(371) C. Hampel, H.-J. Werner, "Local Treatment of Electron Correlation in Coupled Cluster Theory", *J. Chem. Phys.* **104**, 6286 (1996).

(372) M. Schütz, H.-J. Werner, "Low-Order Scaling Local Electron Correlation Methods. IV. Linear Scaling Local Coupled-Cluster (LCCSD)", *J. Chem. Phys.* **114**, 661 (2001).

(373) M. Schütz, "Low-Order Scaling Local Electron Correlation Methods. III. Linear Scaling Local Perturbative Triples Correction (T)", *J. Chem. Phys.* **113**, 9986 (2000).

(374) M. Schütz, H.-J. Werner, "Local Perturbative Triples Correction (T) With Linear Cost Scaling", *Chem. Phys. Lett.* **318**, 370 (2000).

(375) M. Schütz, "Low-Order Scaling Local Correlation Methods. V. Connected Triples Beyond (T): Linear Scaling Local CCSDT-1b", *J. Chem. Phys.* **116**, 8772 (2002).

(376) H.-J. Werner, F. R. Manby, P. J. Knowles, "Fast Linear Scaling Second-Order Møller-Plesset Perturbation Theory (MP2) Using Local and Density Fitting Approximations", *J. Chem. Phys.* **118**, 8149 (2003).

(377) M. Schütz, H.-J. Werner, R. Lindh, F. R. Manby, "Analytical Energy Gradient for Local Second-Order Møller-Plesset Perturbation Theory Using Density Fitting Approximations", *J. Chem. Phys.* **121**, 737 (2004).

(378) J. Almlöf, "Elimination of Energy Denominators in Møller-Plesset Perturbation Theory by a Laplace Transform Approach", *Chem. Phys. Lett.* **181**, 319 (1991).

(379) M. Häser, J. Almlöf, "Laplace Transform Techniques in Møller-Plesset Perturbation Theory", *J. Chem. Phys.* **96**, 489 (1992).

(380) G. Rauhut, P. Pulay, "Considerations Regarding the Local Treatment of Laplace Transform MPPT", *Chem. Phys. Lett.* **248**, 223 (1996).

(381) A. K. Wilson, J. Almlöf, "Møller-Plesset Correlation Energies in a Localized Orbital Basis Using a Laplace Transform Technique", *Theor. Chim. Acta* **95**, 49 (1997).

(382) P. Y. Ayala, G. E. Scuseria, "Linear Scaling Second-Order Moller-Plesset Theory in the Atomic Orbital Basis for Large Molecular Systems", *J. Chem. Phys.* **110**, 3660 (1999).

(383) P. Y. Ayala, K. N. Kudin, G. E. Scuseria, "Atomic Orbital Laplace-Transformed Second-Order Møller-Plesset Theory for Periodic Systems", *J. Chem. Phys.* **115**, 9698 (2001).

(384) D. S. Lambrecht, B. Doser, C. Ochsenfeld, "Rigorous Integral Screening for Electron Correlation Methods", *J. Chem. Phys.* **123**, 184102 (2005).

(385) D. Kats, D. Usvyat, M. Schütz, "On the Use of the Laplace Transform in Local Correlation Methods", *Phys. Chem. Chem. Phys.* **10**, 3430 (2008).

(386) Y. Jung, R. C. Lochan, A. D. Dutoi, M. Head-Gordon, "Scaled Opposite-Spin Second Order Møller-Plesset Correlation Energy: An Economical Electronic Structure Method", *J. Chem. Phys.* **121**, 9793 (2004).

(387) Y. Jung, Y. Shao, M. Head-Gordon, "Fast Evaluation of Scaled Opposite Spin Second-Order Møller-Plesset Correlation Energies Using Auxiliary Basis Expansions and Exploiting Sparsity", *J. Comp. Chem.* **28**, 1953 (2007).

(388) L. Vogt, R. Olivares-Anaya, S. Kermes, Y. Shao, C. Amador-Bedolla, A. Aspuru-Guzik, "Accelerating Resolution-of-the-Identity Second-Order Møller-Plesset Quantum Chemistry Calculations with Graphical Processing Units", *J. Phys. Chem. A* **112**, 2049 (2008).

(389) J. V. Ortiz, "Towards an Exact One-Electron Picture of Chemical Bonding", *Adv. Quantum Chem.* **35**, 33 (1999).

(390) R. Flores-Moreno, J. V. Ortiz, "Integral Approximations in *ab initio*, Electron Propagator Calculations", *J. Chem. Phys.* **131**, 124110 (2009).

(391) G. B. Arfken, H. J. Weber, F. E. Harris, *Mathematical Methods for Physicists. A Comprehensive Guide*, Academic Press, Oxford, United Kingdom (2013).

(392) J. C. Slater, *Quantum Theory of Atomic Structure*, McGraw-Hill, New York, USA (1960).

This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. A template that can be used to format a PhD thesis with this look and feel has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate or from its author, Jordan Suchow, at suchow@post.harvard.edu.