Improving Performance and Accuracy of Hybrid-Functional Based Molecular Dynamics in Plane Waves

THÈSE N^{O} 9041 (2018)

PRÉSENTÉE LE 16 NOVEMBRE 2018
À LA FACULTÉ DES SCIENCES DE BASE
LABORATOIRE DE CHIMIE ET BIOCHIMIE COMPUTATIONNELLES
PROGRAMME DOCTORAL EN CHIMIE ET GÉNIE CHIMIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Martin Peter BIRCHER

acceptée sur proposition du jury:

Dr A.-S. Chauvin, présidente du jury
Prof. U. Röthlisberger, directrice de thèse
Dr T. Laino, rapporteur
Prof. T. A. Wesolowski, rapporteur
Prof. O. Yazyev, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2018

Ja, mach nur einen Plan!
Sei nur ein großes Licht!
Und mach dann noch ‘nen zweiten Plan,
Geh‘n tun sie beide nicht.
Bertolt Brecht, ‘Die Dreigroschenoper’[1]

Acknowledgements

I am grateful to Prof. Ursula Röthlisberger for having offered me a position in her group and for all of our valuable discussions; Drs Simone Meloni and J. Magnus Haugaard Olsen for many a memorable scientific dispute; Dr. Swarnendu Bhattacharyya for our joyful collaboration; Prof. Michele Cascella for having me introduced to the fascinating field that is DFT-based *first principles* molecular dynamics; and my mother, Silvia, for careful reading of the manuscript. Dr. Teodoro Laino, Profs Tomasz Wesolowski and Oleg Yazyev as well as Dr. Anne-Sophie Chauvin are most gratefully acknowledged for their kindly accepting to be part of my thesis committee.

Most importantly - and above all else! - I wish to express my sincere gratitude and greatest esteem to my dear family, friends and mentors for their open ears and kind support.

Ecublens VD, in September 2018
Martin P. Bircher

Résumé

La simulation de la matière condensée avec la dynamique moléculaire *ab initio* (AIMD) s'appuie fortement sur la théorie de la fonctionnelle de densité de Kohn-Sham (KS-DFT). La précision de telles simulations est dictée par la fiabilité de la surface d'énergie potentielle (PES), de même que par l'échelle de temps disponible. Ces facteurs sont influencés de manière cruciale par le choix de la fonctionnelle d'échange-correlation (xc). Parmi les fonctionnelles xc les plus exactes disponibles de nos jours se trouvent les fonctionelles hybrides, qui incluent une partie d'échange exact. En combinaison avec une base d'ondes planes, populaire dans le domaine de l'AIMD, le surcoût computationnel engendré par l'évaluation des intégrales d'échange peut cependant significativement limiter l'échelle de temps accessible, ce qui peut entraver la convergence des observables. Cet ouvrage a pour but d'améliorer la disponibilité de fonctionnelles hybrides dans une base d'ondes planes tout en réduisant leur coût computationnel.

La première partie fournit une introduction à la AIMD. En partant de la fonction de partition quantique, les simplifications dues à l'approximation de Born-Oppenheimer sont démontrées. Suit une discussion de la base théorique de la KS-DFT, un aperçu des approximations faites en pratique ainsi que leur implémentation dans le formalisme d'ondes planes/pseudopotentiels.

La deuxième partie présente l'implémentation et les performances de deux familles populaires de fonctionnelles xc. Premièrement, la méthode d'attenuation de Coulomb (CAM) est déclinée pour les ondes planes et ses performances sont comparées aux bases centrées sur les noyaux atomiques. Une nouvelle fonctionnelle, CAM-O3LYP, est documentée. Celle-ci surpasse la très populaire CAM-B3LYP dans certains cas difficiles. Deuxièmement, l'implémentation en ondes planes des fonctionnelles de Minnesota M05 à M11 est présentée. Une analyse détaillée de la convergence suggère qu'une grille d'intégration suffisamment dense est exigée pour obtenir des différences énergétiques précises. Une comparaison d'enthalpies entre ondes planes et bases centrées sur les noyaux révèle une sensibilité inhabituelle de certaines fonctionelles par rapport au choix de base. Une explication est présentée en comparant les densités électroniques parmi des bases et fonctionnelles différentes. Ceci démontre que les ondes planes permettent d'obtenir des valeurs limites pour les fonctionnelles s'avérant difficiles à converger dans une base centrée sur les noyaux.

La troisième partie introduit une approche pour le calcul des intégrales d'échange exact d'un système isolé, basée sur la mise à l'échelle des coordonnées. En appliquant cette approche à l'expression Hartree-Fock, il est possible de réduire le coût des transformations de Fourier rapides (FFT) qui constituent le goulet d'étranglement calculatoire, tout en maintenant une bonne précision. L'implémentation et la parallélisation de cette méthode sont présentées. Des accélérations d'un ordre de grandeur sont atteintes, ce qui améliore significativement le bilan computationnel des fonctionnelles hybrides.

Dans la quatrième partie, des résultats additionnels sont résumés. Une extension de la collection de potentiels centrés sur les noyaux corrigeant pour la dispersion (DCACP) est proposée. Finalement, une méthode pour l'affinement de structures basée sur la différence de densité électronique est présentée (d3MD). L'ouvrage se termine par une synthèse des résultats et par une discussion de futures applications envisageables.

Mots-clés : dynamique moléculaire *ab initio* ; théorie de la fonctionnelle de la densité (DFT) ; fonctionnelles hybrides ; méthode d'attenuation de Coulomb (CAM) ; échange exact ; mise à l'échelle des coordonnées ; méthodes multigrilles ; base d'ondes planes ; d3MD

Abstract

The simulation of condensed matter using quantum mechanical methods in *first principles* Molecular Dynamics (FPMD) heavily relies on Kohn-Sham Density Functional Theory (KS-DFT) calculations. The accuracy of such simulations is governed by the reliability of the underlying potential energy surface (PES) and the accessible time scale; both factors are crucially influenced by the choice of exchange-correlation (xc) functional. Hybrid xc functionals, which include a fraction of exact exchange, are among the most accurate xc approximations available to date. However, in particular in combination with the plane wave basis commonly used in FPMD, the computational overhead due to the evaluation of exchange integrals significantly limits the accessible time scale of the simulation, thus hampering convergence of observables. This work aims at improving the availability of popular hybrid xc functionals in a plane wave basis while reducing their computational overhead.

Part I provides an introduction to the field of FPMD. Starting from the quantum mechanical partition function, we show the simplifications behind the single-state Born-Oppenheimer approximation, followed by a discussion of the theoretical foundations of KS-DFT, an overview of practical approximations, and their implementation in a plane wave/pseudopotential formalism.

In Part II, we outline the implementation and performance of two popular families of xc functionals. We first derive the plane wave expressions for the Coulomb-Attenuation Method (CAM) and compare the performance of the method in a plane wave basis to common atom-centred basis sets. A new xc functional, CAM-O3LYP, is reported, which outperforms the commonly employed CAM-B3LYP in some notorious cases. We then present the implementation of the M05 to M11 families of Minnesota functionals in a plane wave basis. A detailed convergence analysis suggests that a fine enough integration mesh has to be provided in order to obtain accurate energy differences, and a comparison of reaction enthalpies obtained in plane waves and atom-centred bases reveals an unusual sensitivity of certain Minnesota functionals to the choice of basis. We provide a rationale for the observed differences by comparing electron densities between different functionals and basis sets demonstrating that plane waves make it possible to conveniently obtain values for basis set limits for functionals that can be difficult to converge in atom-centred bases.

Part III introduces a coordinate-scaling scheme for the calculation of exact exchange integrals for isolated systems in a plane wave basis. By applying the scaling relations of the exact exchange functional to the Hartree-Fock term, we show that it is possible to effectively reduce the cost of the Fast Fourier Transforms (FFT) that constitute the bottleneck of the calculation while retaining good accuracy. We then present the practical implementation and parallelisation of the method in the CPMD code and show that speedups can reach values of up to one order of magnitude, thereby considerably improving the computational footprint of hybrid functionals.

In Part IV, we offer a perspective on further work by discussing recent additions to the library of dispersion-corrected atom-centred potentials (DCACP) and by introducing a method for structural refinement based on electron density difference maps, density-difference driven molecular dynamics (d3MD). We conclude by providing a brief overview of the results obtained in this work and by discussing future perspectives.

Keywords: *First-principles* Molecular Dynamics; Density Functional Theory (DFT); Hybrid Functionals; Coulomb-Attenuation Method (CAM); Exact Exchange; Coordinate-Scaling; Multigrid Methods; Plane Wave Basis; Density-Difference Driven Molecular Dynamics (d3MD)

Abstraktum

Quantenmechanische Simulationen kondensierter Materie mittels *ab initio* Molekulardynamik (AIMD) werden heute hauptsächlich mit der Kohn-Sham Dichtefunktionaltheorie durchgeführt. Die Genauigkeit einer AIMD-Simulation wird dabei sowohl von der Zuverlässigkeit der zugrundeliegenden Potentialhyperfläche (PES) als auch von der erreichbaren Zeitskala beeinflusst. Diese beiden Faktoren wiederum hängen von der Wahl des Austausch-Korrelations-Funktionals (xc-Funktional) ab. Hybridfunktionale, die auf den Austauschintegralen einer einzelnen Determinante basieren, gehören zu den präzisesten xc-Funktionalen, die zur Zeit verfügbar sind. Die Berechnung der Austauschintegrale ist jedoch sehr rechenintensiv, insbesondere wenn die Simulation in ebenen Wellen durchgeführt wird, die im Gebiet der AIMD sehr verbreitet sind. Dies limitiert die erreichbare Zeitskala, was sich negativ auf die Konvergenz der Observablen auswirkt. In dieser Dissertation soll aufgezeigt werden, wie die Verfügbarkeit beliebter Hybridfunktionale in ebenen Wellen verbessert und die Rechenkosten verringert werden können.

Im ersten Teil wird das Gebiet der AIMD vertieft eingeführt. Von der quantenmechanischen Zustandsfunktion her wird gezeigt, welche Vereinfachungen und Annahmen hinter der Born-Oppenheimer-Näherung liegen. Darauf folgt eine Diskussion der theoretischen Grundlagen der KS-DFT sowie ein Überblick häufig getroffener Näherungen und ihrer Implementierung in ebenen Wellen.

Im zweiten Teil wird die Implementierung zweier populärer Familien von xc-Funktionalen beschrieben. Auf der Basis der Ableitung der Coulomb-Attenuationsmethode (CAM) in ebenen Wellen werden die erhaltenen Resultate mit atomzentrierten Basissätzen verglichen. Ein neues xc-Funktional, genannt CAM-O3LYP, das sich in gewissen problematischen Systemen als zuverlässiger als das beliebte CAM-B3LYP herausstellt, wird ebenso vorgestellt. Darauffolgend wird die Implementierung der M05 bis M11-Familien der Minnesota-Funktionale in ebenen Wellen dargelegt. Anhand einer detaillierten Konvergenzanalyse wird gezeigt, dass zur Erreichung genauer Energieunterschiede ein dichtes Integrationsgitter vonnöten ist. Ein Vergleich von Reaktionsenthalpien aus Rechnungen mit atomzentrierten Basissätzen zeigt eine aussergewöhnliche Empfindlichkeit der Minnesota-Funktionale gegenüber dem Basissatz. Dies wird anhand eines Vergleichs der Elektronendichten zwischen verschiedenen Funktionalen und Basissätzen illustriert. Somit wird gezeigt, dass ebene Wellen für Konvergenzanalysen den Basissatz der Wahl darstellen.

Im dritten Teil wird eine Methode eingeführt, die die Berechnung von Austauschintegralen beträchtlich beschleunigt. Indem die Koordinatenskalierungsbeziehungen des exakten Austauschfunktionals auf die Austauschintegrale angewandt werden, gelingt es, die Rechenzeit derjenigen schnellen Fouriertransformationen (FFT), die den Engpass der Rechnung darstellen, zu reduzieren, ohne Abschläge in der Genaugikeit der Resultate einzugehen. Die Implementierung der Methode, die Effizienzsteigerungen von bis zu einer Zehnerpotenz erlaubt, wird aufgezeigt.

Im vierten Teil werden weitere Arbeiten, die im Rahmen dieser Dissertation durchgeführt wurden, zusammengefasst. Erstens wird die Erweiterung der Bibliothek dispersionskorrigierter atomzentrierter Potenziale (DCACP) dokumentiert. Zweitens wird, basierend auf Elektronendichteunterschieden zwischen zwei Strukturen, eine neue Methode zur Strukturaufklärung eingeführt, die dichtedifferenzgesteuerte Molekulardynamik (d3MD). Der Text wird mit einer kurzen Übersicht der Resultate dieser Dissertation sowie mit einem Ausblick auf zukünftige Entwicklungen und Anwendungen abgeschlossen.

Schlagwörter: *Ab initio* Molekulardynamik; Dichtefunktionaltheorie; Hybridfunktionale; Coulomb-Attenuationsmethode; Exakter Hartree-Fock Austausch; Koordinatenskalierung; Mehrfachgitteransatz; Basissatz ebener Wellen; dichtedifferenzgesteuerte Molekulardynamik (d3MD)

Contents

Abstract (French) **v**

Abstract (English) **vii**

Abstract (German) **ix**

0 Preface **3**

0.1 Theoretical Chemistry - No Dry Matter 4

1 Introduction **9**

1.1 *In Silico*: From Work Bench to Core Processing Unit 9

1.2 Time-Scale *vs.* Accuracy and the Aim of this Thesis 11

I A Quantum Picture of Chemistry **15**

2 The Foundations of Molecular Dynamics **19**

2.1 Classical Statistical Mechanics 20

2.2 Quantum Statistical Mechanics 23

2.3 Molecular Dynamics in the Born-Oppenheimer Formulation . 28

3 The Electronic Structure Problem **37**

3.1 The Universal Density Functional 37

3.2 Kohn-Sham Density Functional Theory 41

4 A Numerical Solution **65**

4.1 An Expansion in Plane Waves 65

II Climbing Jacob's Ladder in *first principles* Molecular Dynamics 75

5 À-la-carte Coulomb-Attenuation Method for Plane Waves 79
5.1 Introduction . . . 80
5.2 Theory . . . 85
5.3 Implementation . . . 88
5.4 Test Set and Computational Details . . . 91
5.5 Results and Discussion . . . 93
5.6 Conclusions . . . 104

6 Basis Set Dependence of Minnesota Functionals 105
6.1 Introduction . . . 106
6.2 Theory . . . 111
6.3 Computational Details . . . 117
6.4 Results and Discussion . . . 118
6.5 Computational Overhead . . . 138
6.6 Conclusions . . . 140

III Accelerating the Evaluation of Exact Exchange Integrals in Plane Waves 143

7 Coordinate Scaling Relations for Exact Exchange 147

8 Implementation of Coordinate-Scaled Exact Exchange 157
8.1 Introduction . . . 158
8.2 Fast Fourier Transforms in Plane-Wave Kohn-Sham DFT . . . 160
8.3 Coordinate-Scaled Exact Exchange . . . 165
8.4 Implementation of Coordinate-Scaled Exact Exchange . . . 166
8.5 Results and Discussion . . . 170
8.6 Summary . . . 180

IV Retrospectives and Perspectives 181

9 Further Work and Projects 183
9.1 Dispersion Correction for Plane Wave DFT . . . 183
9.2 Density-Difference Driven MD: d3MD . . . 193

10 Conclusions and Outlook 207

V Postscript 213

A Supporting Information 215
  A.1 Supporting Information for Chapter 5 215
  A.2 Supporting Information for Chapter 7 217

List of Figures 221

List of Tables 227

Abbreviations and Nomenclature 234

Bibliography 235

Curriculum Vitae 255

Index of Chapters 261

Das Problem ist zu schwierig für einen einzelnen Mann.
Ich werde mein Elektronengehirn befragen.
Dagobert Duck in 'Die Geldquelle', as translated by Dr. Erika Fuchs[2]

CHAPTER 0

Preface

WIGGLE, wiggle, bump. This is how the author pictured the atomistic contents of his round bottom flask in the wet lab, just seconds before he absent-mindedly inhaled some not so innocuous cauliflower-like fumes. He then fled the lab. No permanent damage was caused.

Admittedly, at the time, the organic molecules at the source of this olfactory nuisance were to him nothing but a collection of symbols, interconnected with sticks; and all of the time the sticks moved, pushed around other sticks, attracted even others, and this was how chemical bonds were formed. In his head and on paper, that is, not in reality. But the level of abstraction was intriguing and appealing, arrows and lines hastily sketched on a piece of paper. And while he was surely aware that different solvents would have different effects on the reactions he so much loved to draw, influencing their stereochemical outcome, inhibiting or accelerating certain processes, this was dubious knowledge that had to be learnt by heart. There was no space for solvent on the paper. It would obscure the mechanistic beauty. It was not really *interesting*. And above all, it seemed that there was little proof that was not phenomenological in nature. The concepts of theoretical and physical chemistry, back then, seemed to be of little help[i], and so they had no place within the author's world of pencil drawn arrows.

Especially thermodynamics. That one seemed particularly useless. Discussions with fellow organic chemistry enthusiasts would often follow along the same lines: One was able to calculate the rate of a reaction from experimental data, estimate its barrier, determine the reaction enthalpy, but to what avail? Even with the knowledge of all possible thermodynamic quantities, how would they tell anyone *why* the molecules behaved the way they did? Reaction rates could support a certain mechanism, measured enthalpies could make others unlikely, but the proof of a reaction following a particular path was often in the observed intermediates. Which, again, was the realm of a world impregnated with sweet solvent fumes, the wet lab.

---

[i] They seemed a powerful sedative, though.

Oh, woe was him who had to study for a thermodynamics exam. What dry matter. And how wrong the author was!

0.1 Theoretical Chemistry - No Dry Matter

Early on in their studies, the author and his colleagues came to appreciate thermodynamics for a particularly dubious reason: It came in handy as an excuse. For as soon as a chemical reaction would not take place in quite the way they had expected, they were quick to cite the same old unpredictable delinquent: Entropy. Even if they might not have fundamentally understood what it was (entropy was just disorder, or was it?), everything bizarre and surprising was always and exclusively due to entropy. The entropic factor was the perfect scapegoat.

But the author soon learned that there was more to theoretical chemistry than he thought. Soon, it was no dry matter anymore. Not at all.

0.1.1 Schrödinger

At the same time, some new perspectives surfaced in the author's quest for answers: Quantum Chemistry. The basic concepts of this discipline are so different to the classical world that calling them fascinating is for sure no overstatement. More than that, they suddenly made that whole bunch of theory lectures become much more appealing (not quite that appealing yet, but at least much more than before). It wasn't only about this infamous cat in a box! Elegant and tremendously powerful: The author had fallen for Quantum Mechanics, even if it was no love at first sight, but more of an infatuation with the seemingly incomprehensible.

The basic principle of quantum wave mechanics is magnificent: Given a wavefunction of a system, all its physical properties can be derived by applying the proper set of operators to it. The wavefunction itself is a mere mathematical tool; a cryptic key that encodes all one could want to know about the system. More commonly: Ψ. A molecular Ψ is, however, of abhorrent dimensionality: Every nucleus, every electron will bring its own set into the coordinates of the wavefunction, which means that the overall object is some kind of Pandora's box that cannot be fully visualised, except for the case of one single particle (and even then, this is only possible if

Ψ is real). In all other cases, any attempt at illustration will run out of dimensions that are available to the human eye. Admittedly, a chemist may hope for something rather more visual.

The basic principles of applying the Schrödinger equation to chemical matter are these[3–5]: Firstly, the problem cannot be solved analytically for virtually any case of interest. Secondly, it therefore needs to be represented numerically and thirdly, since memory and paper are both finite, it needs to be suitably approximated such that a numerical representation is feasible. In order to make the necessary approximations, there exist a hierarchy of methods, one more accurate than the other. However, the more elaborate the approximation to the wavefunction, the more computationally demanding are the methods to treat. This ultimately leads to a point where the most exact method cannot be used for a sizeably large system, and compromises between accuracy and computational overhead have to be made.

So could these methods provide the explanations the author had been seeking ever since the beginning of his studies? There was never any discussion of molecules in solution at finite temperature. Such a system would be too big to be treated by an accurate wavefunction approximation. To the author, the sheer idea of attempting it seemed outrageous. The path seemed clear: Calculations are done on the species of interest - and on this species only. This may include an occasional, small handful of solvent molecules around the centre of interest, but not much more. The system itself is in vacuum, and solvation effects will be mimicked by appropriate models,[6,7] they may for instance incorporate the dielectric constant of a solvent to make the molecule experience a more or less polar environment. Finally, there exist approaches to interpolate the results to room temperature. And that was that.

Insecurity soon struck. Sceptical of the author's newly found enthusiasm for the quantum world, his friends would ask how he imagined those models to capture what was going on in their round bottom flasks, among an uncountable number of molecules. How would five water molecules mimic Avogadro numbers of particles[8]? Can rate constants really be reliably predicted by only implicitly accounting for the Gibbs free energy of activation, $\Delta G^{\dagger}$, based on the enthalpy $\Delta H^{\dagger}$[9]? The author had no answer. And he soon learned that many a property rather accurately calculated at 0 K (and appropriately extrapolated) would not always reflect what was going on in a reaction vessel at room temperature.[10] And with some theoretical study[11] hinting at a behaviour conflicting with what the author had observed while performing some novel radical reaction during his bachelor thesis,[12] hope cast on the predictive power of these computational methods seemed rather lost again. What had happened? The entropic factor had seemingly stricken back.

And so it appeared that the universe of Chemistry had to remain divided into two worlds: The world of the wet lab and the world of the theoretician. There was a missing link in between. Even if one were to calculate the wavefunction of a representative system, the author wondered, how would one then describe a chemical reaction at finite temperature? Meanwhile, unanswered questions from courses in Inorganic Chemistry, Organic Chemistry and Biochemistry would keep piling up.

0.1.2 Boltzmann

Statistical Mechanics changed it all by providing the missing link. It is a beautiful discipline: By invoking a statistical picture of the constituents of a chemical system, the familiar thermodynamic quantities - entropy, enthalpy, free energy, internal energy - can be derived not from measurements, but from sleek, elegant equations and by making use of clear assumptions. On a more empirical level, any chemist is soon confronted with Ludwig von Boltzmann and his factor, e^{-E/k_BT}. But it is only in the study of statistical mechanics that this simple exponential reveals all of its elegance. We will see in the next part that it plays a fundamental role in one of the key concepts - the summit[13] - of statistical mechanics: The partition function. Much like in wave mechanics, it is the vault that contains all that is to be known. Not only can many properties be directly derived from it, but it also serves as the crucial normalisation factor in thermal averages of observables.

What before were pencil drawn lines and arrows of Organic Chemistry have become equations: Integrals, derivatives, sums. Statistical Mechanics provides much more than a model of what is going on in that round-bottom flask; the resulting equations are not rooted in (admittedly cleverly made) assumptions or intuition, but they are fundamental theorems of physics. Alas, in those equations it appears again, the untractability of any chemical system of interest. So would it in the end have been nothing but running in circles?

Surprisingly, it does not have to be. There exists a hypothesis in Statistical Mechanics that allows for averages to be obtained over time. Or: Averaging an observable over the partition function is equivalent to averaging it over time; and this can be carried out by simply following the time-evolution of a system of interest. This solves the tractability problem of the partition function, but it does not solve one other principal issue.

In statistical mechanics, the Hamiltonian is a far more than familiar sight, and that distinctly recalls the dimensionality problem of wavefunction theory: If the system is too large, compromises have to be made. Then again, Statistical Mechanics is about statistics, and statistics require many samples, many particles. A reagent surrounded by half a dozen solvent molecules might make as bad if not worse a model than the same system treated *in vacuo*, at 0 K, using wavefunction theory. So the system one explores should be representative of the chemical event one wants to observe, and that means including its surroundings beyond that of a small cluster. While it is possible to treat such larger systems with wavefunction methods, their accuracy will be limited, and their computational cost too high to allow for the time evolution of the system to be followed.

The solution to this problem is as compellingly elegant as it is deceptively simple. It lies in the electron density of the system. The basic theorems of Hohenberg and Kohn[14] in 1964 were the foundations of a revolution in electronic structure theory. They proved that instead of resorting to the wavefunction, all ground-state expectation values of the Hamiltonian can be obtained from the N-electron density, a 3D-Cartesian variable, in a *density functional* theory[14, 15] (DFT; and under appropriate conditions only, as we will see later on in the text). A variational principle exists, making the theory as straightforward (if not more) to use in practice as wavefunction theory. Since DFT casts the problem into that of finding a quantity that is a function of only 3 Cartesian coordinates, suitable algorithms make it possible to treat large systems, while scaling linearly with system size.[16]

Statistical mechanics and DFT are not only intimately linked by some fundamental concepts (there exists an ensemble formulation of DFT[17]), but both of them add up to draw a picture much greater than any of its single elements. It is this unique combination that makes the link between macroscopic and microscopic worlds computationally possible.

Cue numerics.

Chapter 1

Introduction

The numerical aspects of putting the concepts of statistical mechanics and electronic structure theory to good use can be (almost) as exciting as the fundamental theories themselves. For if Statistical Mechanics and Density Functional Theory provide us with the equations that we need to solve to describe a chemical process, we are unable to solve them analytically (let a part a few small systems where analytical solutions are at least partially available). For most of those equations, we are not even able to solve their exact forms numerically. Instead, we search for appropriate models that are computationally tractable. Whether in the world of wavefunctions or densities: As physically justified, as accurate and as computationally efficient as possible - these are the challenges in the field of Computational Chemistry, where phenomena from the work bench are elucidated *in silico* on core processing units (CPUs).

1.1 *In Silico*: From Work Bench to Core Processing Unit

Computational Chemistry is, in its present form, a rather young discipline at the interface of Biology, Chemistry, Physics, Mathematics and Computational Sciences. The synergistic combination of increasingly powerful computational infrastructure, advances in electronic structure methods and algorithmic developments have made it possible to elucidate physical, chemical or biochemical processes at the atomistic level by simulating their time-evolution in solution and at finite temperature *in silico*. While considerable progress has been achieved in the field over the last decades, making it possible to simulate increasingly large systems, some problems and processes remain difficult to describe; be it due to the lack of availability or tractability of a suitable method.

1.1.1 Computational Chemistry

The possibility of gaining some fundamental understanding of chemical processes is certainly desirable from a philosophical perspective, but its impact is by far not limited to a metaphysical point of view. The simulation of chemical compounds and their interactions at an atomistic level is of high practical relevance: Understanding the detailed mechanisms of a chemical reaction can allow for reactions to be tweaked and applied under optimal conditions, guaranteeing maximal yields or minimal environmental impact.[18] Comprehending the molecular action of drugs has helped develop new compounds and fight resistance.[19–22] It is thanks to modern Theoretical and Computational Chemistry that an atomistic understanding of these challenges has become much more easily accessible: Partition functions comprising a large number of particles are efficiently sampled using stochastic Monte Carlo approaches,[23] the dynamics of liquids such as water can be followed using (path-integral) Molecular Dynamics (MD),[24] reaction barriers *in vacuo* can be calculated from high-throughput methods,[18] the binding of thousands of compounds to a putative active site in a protein can be performed in docking calculations,[21] and the decomposition of compounds upon impact of a laser pulse can be simulated in the excited state[25–27] - to mention only a few. From the plethora of studies that have seen the light over the last few decades, it appears that every problem chemically conceivable cannot escape some theoretical treatment.

This would not have been possible without the existence of appropriate, affordable infrastructure: The rise of Computational Chemistry has been accompanied and encouraged by ever improving computational methods, efficiency and CPU power. Using classical potentials, proteins in a membrane can now routinely be simulated, coarse-graining methods allow for even larger systems and time scales to be investigated and using Density Functional Theory (DFT),[28] it has become possible to model hundreds of atoms quantum mechanically, based on *first principles.* Computers have become powerful enough that even for large systems, calculations are not restricted to a single geometry, but it has instead become possible to track the dynamical evolution of the system over the pico- over the nano- up to a millisecond[20] range, permitting the extraction of thermodynamic observables at finite temperature. With computational power still increasing, so is the size of the systems that can be tackled. The scalability of highly accurate, correlated wavefunction methods has come a long way as well,[5, 29] and while such methods may not lend themselves to dynamics over long time scales[30] (yet), efficient wavefunction algorithms are vital in facilitating the benchmarking of computationally more cost-effective methods such as DFT.

A part of the formalism at the base of *in silico* chemistry will be explored in detail in Part I of this text.

1.2 Time-Scale *vs.* Accuracy and the Aim of this Thesis

Numerical simulations, commonly referred to as 'runs', rely on models being computationally tractable. Computational tractability implies that in a given time, a run can be performed that is on one hand long enough (to ensure convergence of the statistical observables), and on the other hand of sufficient accuracy with respect to the underlying potential energy surface (to ensure that the computed observables make any sense). Both criteria are intimately linked, as we shall see in the following paragraphs.

1.2.1 Accuracy: A Three Fold Problem

The practicability of a simple static electronic structure calculation is limited by two factors: The accuracy of the model itself with respect to the system that it is applied to, as well as the computational cost of that particular model. Especially in wavefunction based theory, apart from execution time, sufficient memory to store the information on the system may be an issue; this is much less of a problem for a density-dependent theory such as DFT. Nonetheless, for any method that is not exact but an approximation of the exact form, some desirable approximation may be too computationally expensive for a system of a given size. In the field of *first principles* MD, one particularly prohibitive combination can be the unfortunate combination of hybrid exchange-correlation (xc) functionals and plane wave-based MD: Many of the popular exchange-correlation (xc) approximations used to date belong to an extension of the group of hybrid functionals[31] initially proposed by Becke, which are among the most accurate approximations available to date. However, their use in a delocalised and periodic basis such as plane waves will engender a large computational overhead that, depending on the system, can prove prohibitive. While sometimes, the system may be replaced by a smaller model system, finite size effects should not be underestimated especially at the *first principles* level (one example being BLA - bond length alteration in polyenes[32]). Hence, a balanced compromise between execution time and accuracy has to be found.

If the dynamics of the system should be explored to unravel the underlying thermodynamics and kinetics (implying finite temperature), yet another question poses itself. We will see in Part I that the accuracy of an observable obtained as a time-average over a MD trajectory depends on the time scale of the fluctuations of the observable. It is therefore directly linked to the length of the simulated dynamics. If the dynamics are too short, the property may not have been sufficiently sampled, and the time-average taken over the trajectory will exhibit a large statistical error. This may result in an unphysical prediction. Enhanced sampling methods exist to overcome the time scale problem,[33–37] but while they reduce the time that has to be simulated in order for a given process to take place, they cannot remedy for fallacies due to an inaccurate PES.

Errors introduced due to insufficient sampling can be much larger than the error of the underlying, static electronic structure calculation. In *first principles* based MD, the compromise that has to be found will be inclined towards low execution time, rather than towards highest possible accuracy of the underlying electronic structure method. Given the magnitude of thermal fluctuations at room temperature, compromises on the accuracy of the PES can be made, provided that the approximate PES be at least qualitatively correct and representative of the physics of the process. And so, the choice of methods in *first principles* Molecular Dynamics is a choice between the achievable time scales and the accuracy of the underlying potential energy surface (PES)[i]. In practice, the errors due to sampling and the underlying electronic structure method should be well balanced.

1.2.2 Structure and Aim of the Present Work

Higher accuracy can be attained in a three-fold way: Once by creating or implementing more accurate approximations in the electronic structure regime, once by making existing approaches more computationally efficient, and once by attempting to overcome the time scale issue by improving the sampling itself *via* enhanced sampling techniques.

Part I of this work provides a journey through the theoretical foundations of DFT-based *first principles* MD; it shall be dedicated to the pleasure of theory. By drawing our very own Quantum Picture of Chemistry, we will attempt to establish a sleek link between the commonly employed models and approximations in the field to the rigorous basis that is exact quantum statistical mechanics. We will encounter several issues that render some of the approximations computationally untractable, and step by step, we shall explore formulations to break the problem down to a level at which it can be, finally, solved numerically. A small part will cover the actual numerics of DFT-based MD, focusing on aspects that are most important for the subsequent topics of this thesis. Whenever a particular approach should pose problems in terms of computational tractability, it shall be explicitly mentioned.

---

[i] With the notable exception of those cases where a computationally more tractable method introduces unacceptable errors. For such systems, a highly accurate PES is unavoidable.

We will not be able to resolve all of the problems evoked in Part I - nor do we even attempt to -, but we can try to improve over some of them. This is the aim of the present work, guided by the wish to extend the accuracy of *first principles* Molecular Dynamics simulations for a broad variety of systems both in terms of the time scale of the dynamics as well as the quality of the underlying PES.

Along these lines, in Part II, we will describe the implementation and validation of more accurate approximations in DFT in a framework specific to *first principles* MD: In Chapter 5, we detail the first implementation of the successful Coulomb-attenuation method in a plane wave basis typical for *first principles* MD. Based on the promising results that have been documented for excited state calculations with the Coulomb-attenuated CAM-B3LYP[38] exchange-correlation (xc) functional, we will show that it is easily possible to construct a customised functional which outperforms CAM-B3LYP for some of the systems where the latter has its known weaknesses, while being slightly more accurate in systems where CAM-B3LYP excels. We will also provide a short comparison between plane waves and the abundantly used atom-centred Gaussian basis which is used for many static studies. Along a similar line, Chapter 6 will in particular focus on the difference between different families of basis functions: We will outline the implementation of the popular Minnesota family[39–46] of xc functionals in plane waves and show that, due to their highly empirical character, it may be necessary to take special measures in a plane wave basis in order to recover fully converged results. We will see how reaction enthalpies obtained in plane waves may differ from atom-centred bases, highlighting the strong basis-set dependency of the Minnesota xc family; and we shall see that given appropriate conditions, energetics may converge much quicker in plane waves than in atom-centred bases.

Then, in Part III, we will describe how the new methods from Part II (and many more) can be made more computationally tractable by exploiting some formal scaling properties of DFT. Chapter 7 will present the use of coordinate-scaling relations in the calculation of the exact exchange of the Kohn-Sham determinant. In Chapter 8, we will detail the implementation and parallelisation of this coordinate-scaling scheme and demonstrate that speedups scale favourably with the number of orbitals, reaching about one order of magnitude at above around 200 states. The speedups obtained with respect to the standard algorithm will make it possible to perform higher-accuracy simulations for larger systems, which is especially important in the excited state regime, providing yet another link to Chapters 5 and 6.

In Part IV, we will offer a recapitulation and discuss some perspectives related to the methodology and implementations carried out in this work. We shall also briefly review some additional projects that have been explored in the context of this thesis: Firstly, the extension of the library of dispersion-corrected atom-centred potentials (DCACP)[47] that have been shown to considerably increase the quality of the PES of weakly bound complexes in combination with generalised gradient approximation (GGA) xc functionals. Secondly, we will sketch a method to perform structural elucidation based on N-electron density difference maps and show preliminary results of performing such density-difference driven MD (d3MD) at the example of three text-book organic chemistry reactions, allowing to overcome the time scale problem due to the rare event that links reactants to products.

This dissertation hence covers diverse aspects of development in Computational Chemistry: From implementations of existing approaches in an alternative basis set to the efficiency-driven derivation of new computational algorithms. But for now, we will plunge into the theoretical realms of the quantum mechanical world of chemistry. Over the next few pages of Part I, it is mainly this what the author would love to share with the reader: His personal insight into why what we do is such great fun. He sincerely hopes that the reader will see it like this, too.

Part I

A Quantum Picture of Chemistry

Notes on Notation

We shall denote quantum mechanical operators by upright letters carrying a hat ($\hat{\mathrm{H}}$), whereas their classical analogues (functions) are denoted using calligraphic letters ($\mathcal{H}$). Vectors are represented by upright, bold letters ($\mathbf{p}, \mathbf{q}$) and abstract *state vectors* are identified using Dirac's Bra-Ket notation ($|\Psi\rangle$). We shall use $\mathbf{p}$ and $\mathbf{r}$ for general momentum and direct space coordinates, respectively, and p and r shall be associated scalars. Discrete functions will take an argument in capitals. Other definitions will be given *en passant* in their respective chapters. Where possible, definitions are given in the space spanned by the complete set of state vectors labelled by i, $\{\left|\Psi^i\right\rangle\}$; their projections and *only* their projections onto direct space are referred to as wavefunctions and denoted by $\Psi^i(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. Parametric dependencies shall be denoted by a semicolon.

CHAPTER 2

The Foundations of Molecular Dynamics

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \left|\Psi(t)\right\rangle = \hat{\mathrm{H}} \left|\Psi(t)\right\rangle \tag{2.1}$$

- THE time dependent Schrödinger equation[48] needs no introduction. Early on, many a chemistry student is confronted with this mysterious equation that seems to hold the answer to everything[i] a chemist could ever care about: Given the knowledge of the Hamiltonian $\hat{\mathrm{H}}$ of interest along with the wavefunction $\langle \mathbf{r}|\Psi(t)\rangle = \Psi\left(\mathbf{r}_1, \ldots, \mathbf{r}_N, t\right)$ of a molecule, all of the possible observables are accessible, and the systems' time-evolution can be predicted.

There are only two problems: Neither is there an analytical expression for the wavefunction of any system of chemical interest, nor would a pure state in the spirit of eq. 2.1 be sufficient to describe the statistics of chemical and biological phenomena that make chemistry so magic. The most basic reaction conditions such as pressure and temperature are not (or ill) defined if only a single molecule is considered; and most reactions - even intramolecular - will refuse to take place in absence of a suitable chemical environment. Chemistry is about more than *only* molecules; it is about the *ensemble* of molecules interacting with each other.

There are two principal problems; and this chapter shall be the starting point of our quest for a link between the microscopic particles and the macroscopic behaviour of a chemical system; be it in an Erlenmeyer flask or in a living being, but in both cases at finite temperature and pressure, constituted by several Avogadro numbers of molecules.

---

[i] Everything non-relativistic, that is.

2.1 Classical Statistical Mechanics

Long before the advent of quantum mechanics, Boltzmann[49] and Gibbs[50] brought the field of *Statistical Mechanics*[13,51,52] to life, linking microscopic properties at the atomistic level (single molecules wiggling and bumping into each other) to macroscopic properties (the reagent in an Erlenmeyer flask that is just foaming over and distributing a nasty odour throughout the laboratory due to the overzealous bumping and wiggling). So before we delve into the fruit of the quantum revolution of the 20th century, let us step back yet another 100 years.

2.1.1 The Fundamental Postulate in the Microcanonical Ensemble

The state of N particles is uniquely defined by the set of their positions $\mathbf{q} = \{\mathbf{q}_1, \ldots, \mathbf{q}_N\}$ and momenta $\mathbf{p} = \{\mathbf{p}_1, \ldots, \mathbf{p}_N\}$ that span the $6N$-dimensional *phase space* $\Gamma(\mathbf{p}, \mathbf{q})$, and their time-evolution in this space is governed by Hamilton's equation of motion:

$$\frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}(\mathbf{p}, \mathbf{q})}{\partial \mathbf{q}_i}, \tag{2.2}$$

$$\frac{\mathrm{d}\mathbf{q}_i}{\mathrm{d}t} = \frac{\partial \mathcal{H}(\mathbf{p}, \mathbf{q})}{\partial \mathbf{p}_i}, \tag{2.3}$$

where the classical Hamiltonian $\mathcal{H}(\mathbf{p}, \mathbf{q})$ is simply:

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + U\left(\mathbf{q}_1, \ldots, \mathbf{q}_N\right). \tag{2.4}$$

A system of constant particle number N, volume V and energy E will evolve on a hypersurface spanned by constant $E = \mathcal{H}(\mathbf{p}, \mathbf{q})$ according to eqs 2.2 and 2.3. Such a *microcanonical ensemble* of particles will explore all distinct points or *microstates* in phase space, $X = \{\mathbf{p}, \mathbf{q} \in \Gamma \mid \mathcal{H}(\mathbf{p}, \mathbf{q}) = E\}$, with an equal probability p_i proportional to the inverse of the total number of microstates W:

$$p_i = \frac{1}{W}. \tag{2.5}$$

This is the fundamental postulate of statistical mechanics.

It is the groundbreaking statement by Boltzmann[49] that links the number of accessible microstates W, a purely microscopic quantity, with a macroscopic thermodynamic property: The entropy S of the system.

$$S = -k_B \ln W. \tag{2.6}$$

Since $\ln W$ is dimensionless, the Boltzmann constant k_B ensures the right dimensionality. Gibbs went on to provide a generalised expression for eq. 2.6, considering that the probability of a microstate p_i will not be equidistributed once the constraint $E = \text{const.}$ is abandoned:

$$S = -k_B \sum_i p_i \ln p_i. \tag{2.7}$$

This definition opens the possibility of leaving the microcanonical ensemble.

2.1.2 The Canonical Partition Function

This fundamental law is the summit of statistical mechanics [...].
Richard P. Feynman on the Canonical Partition Function[13]

Consider a system closer to 'chemical reality': A collection of particles at constant N, T and volume V: The *canonical ensemble*. What is the probability $p_i(\Gamma)$ that this system is in a certain microstate? By considering two microcanonical ensembles coupled to each other, where one of the ensembles serves as a heat bath, the canonical equivalent of the phase space distribution function p_i can be found in the form of a simple exponential:

$$p_i(\Gamma) \propto e^{-\beta \mathcal{H}(\Gamma)}, \tag{2.8}$$

where the thermodynamic β has been introduced as $\beta = (k_B T)^{-1}$. This expression is commonly referred to as the *Boltzmann factor*. Unlike the microcanonical ensemble, which is restricted to a hypersurface in phase space, a system at constant temperature has no such limitation. It can in principle explore the whole of phase space, and the probability of a certain configuration is given by the relation between its energy and the average temperature. The equality in eq. 2.5 has given way to a proportionality since the infinite continuum of probabilities does not sum up to 1. It is the subsequent normalisation that will give rise to a crucial result, if not the most crucial result of statistical mechanics.

The normalisation over probabilities in eq. 2.8 reads:

$$\mathcal{Z}(\beta) = \int_\Gamma \mathrm{d}\Gamma e^{-\beta \mathcal{H}(\Gamma)}. \tag{2.9}$$

This is the *canonical partition function* $\mathcal{Z}(\beta)$: According to Feynman, this sleek expression is the summit of statistical mechanics.

The inconspicuous $\mathcal{Z}(\beta)$ is much more than just a normalisation factor. All important thermodynamic properties, such as the average total energy E, can be derived from the partition function:

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln \mathcal{Z}. \tag{2.10}$$

One can then introduce an associated thermodynamic potential, the Helmholtz free energy A^{53}:

$$\begin{aligned} \langle A \rangle &\equiv \langle E \rangle - TS \qquad (2.11) \\ &= -k_B T \ln \mathcal{Z}, \qquad (2.12) \end{aligned}$$

from which the entropy is easily derived:

$$S = -\frac{\partial A}{\partial T}. \tag{2.13}$$

All these thermodynamic quantities are defined in terms of the probabilistic behaviour of particles that constitute the system, linking the macroscopic and the microscopic worlds.

The same holds for any observable of the system. By combining eqs 2.8 and 2.9, after integration, one arrives at an expression for the expectation value for a macroscopic observable $\langle \mathcal{O} \rangle$:

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int_{\Gamma} \mathrm{d}\Gamma \, \mathcal{O}(\Gamma) e^{-\beta \mathcal{H}(\Gamma)}. \tag{2.14}$$

Every macroscopic property of interest can hence be derived from the partition function, which is based on the microscopic properties of the system. It is therefore no exaggeration if one is to attribute a paramount importance to the partition function. It is the summit that we ultimately want to reach in our quest for a theoretical description of chemistry.

In the following chapters, we will see how to formally describe an ensemble of particles on the quantum mechanical level. After all, chemistry with classical particles only would not be chemistry. We shall then discuss approaches and approximations that enable us to solve the problem in finite time and using finite computational resources.

2.2 Quantum Statistical Mechanics

If the time-dependent Schrödinger equation exactly describes the time evolution of a system (up to a phase), it is not possible to apply it to a statistical mixture of states by resorting to the basic *pure* state vector picture; a prerequisite if a thermodynamic ensemble should be described.[54,55] Indeed, if one wishes to describe a system in a superposition of states, $|\Psi^i(t)\rangle = \sum_j c^i_j(t)\,|\psi^i_j\rangle$, where $\{|\psi^i_j\rangle\}$ is the spectrum of $\hat{\mathrm{H}}$, this gives rise to cross terms of the form ${c^i_k}^*(t)c^i_j(t)$. These terms, mediating between states, do not describe a statistical mixture of states $\{j\}$ over time, but give instead the probability of finding the very system i in one of its eigenstates $|\psi^i_j\rangle$. Clearly, a thermodynamic ensemble cannot be constructed on this basis, since the complex amplitudes give rise to coherence between the states.

Should we, instead, want to describe a *statistical* mixture of N systems each in a state $|\Psi^i(t)\rangle$ - which may or may not be found in a superposition of states each, $|\Psi^i(t)\rangle = \sum_j c^i_j(t)\,|\psi^i_j\rangle$, - we shall have to resort to a different formalism. The solution is found in a *projector*.

2.2.1 About Pure and Mixed States: The Density Operator

Following von Neumann,[56] the state of a *mixed* system is described by a density operator constructed from the set of state vectors $\{\left|\Psi^i(t)\right\rangle\}$, each carrying a statistical probability p_i,

$$\hat{\gamma}(t) = \sum_i p_i \left|\Psi^i(t)\right\rangle \left\langle\Psi^i(t)\right| , \tag{2.15}$$

which is a Hermitian operator of trace 1. There exists a set of $|\psi_j\rangle$ diagonalising $\hat{\gamma}(t)$. For a single *pure state*, $\hat{\gamma}$ is idempotent with $\mathrm{Tr}\,\hat{\gamma}^2 = 1$ and only one $p_i \neq 0$. For a statistical mixture of states, we have in general that $p_i \neq 0$, and $\mathrm{Tr}\,\hat{\gamma}^2 \leq 1$; only in the particular case of only pure $\left|\Psi^i(t)\right\rangle$, equality holds. This case, with $\mathrm{Tr}\,\hat{\gamma}^2 \leq 1$, is referred to as a *mixed state*.

In the density operator formalism, a basis function $|\psi_j\rangle$ of any state $|\Psi^i(t)\rangle$ can be observed with a probability that corresponds to the square of its (time-dependent) expansion coefficient $|c^i_j(t)|^2$ weighted by the statistical term p_i associated to the state i. (We may prove this by projecting the above expression onto some basis.) Hence, if an observable is an eigenfunction of some $|\Psi^i(t)\rangle$, eq. 2.15 gives the probability of observing this particular state within a mixture of states, and *not* the probability of the measurement collapsing into an eigenstate (for if an observable is associated to an eigenvalue, there is no collapse). Instead, if the state is in a superposition of eigenstates associated to the observable, the probability is the product of the probability of collapse and the probability of observing a state that is collapsing. The key advantage of the density operator formalism lies in its versatility: Both pure and mixed states are described with equal ease. We should not fail to note that the Hamiltonian of the $|\Psi^i(t)\rangle$ does not need to be separable. It can just as well be a tensor product of Hamiltonians, $\hat{\mathrm{H}} = \hat{\mathrm{H}}_1 \otimes \hat{\mathrm{H}}_2 \otimes \cdots \otimes \hat{\mathrm{H}}_N$. The notion of *partial* traces Tr_i then allows for the information due to one specific Hamiltonian i to be *traced out*: The state of the complete system need not be completely known, and still, expectation values for a subsystem $\hat{\mathrm{H}}_i$ may be obtained, all the while taking the uncertainty due to the remaining subsystems into account.[54,56] For an ensemble of states, the p_i contain all the information on the composition of the system and its statistical uncertainties, and for a single pure state, $\hat{\gamma}$ is simply a convenient and phase-less alternative to the state vector picture.

It is the same thought exercise that we have evoked for the expectation values of the p_i that can now lead us to the expectation value of a general observable. One simply calculates the elements of $\hat{\gamma}(t)$ in a given basis that diagonalises $\hat{\mathrm{H}}$ and inspects the resulting equation: There are only diagonal terms. (One might be concerned with only one sub-system and the associated partial trace without loss of generality.) It is then straightforward to imagine that this operation must have been analogous to tracing over $\hat{\gamma}$, which leads us to:

$$\langle \hat{\mathrm{O}} \rangle = \mathrm{Tr}\left(\hat{\mathrm{O}}\hat{\gamma}\right). \tag{2.16}$$

The energy of the system is therefore obtained from

$$E = \mathrm{Tr}\left(\hat{\mathrm{H}}\hat{\gamma}\right). \tag{2.17}$$

For any $\hat{\gamma}$ with $\mathrm{Tr}(\hat{\gamma}) = 1$, von Neumann has demonstrated that the entropy can be directly derived from an expression very much reminiscent of the entropy definition due to Gibbs[50]: The quantum or *von Neumann* entropy is given by

$$S = -k_B \,\mathrm{Tr}\left(\hat{\gamma} \ln \hat{\gamma}\right). \tag{2.18}$$

So far, we have provided a comprehensive alternative to the use of the time-independent Schrödinger equation and we have found a fundamental link between quantum mechanics and statistical mechanics - but the question on the time evolution of $\hat{\gamma}$ has been left open. The time-dependent case is easily covered by inserting $\hat{\gamma}$ into the time-dependent Schrödinger equation:

In a striking analogy to classical mechanics, the time evolution of a mixed system is described by the *Liouville-von Neumann equation*:

$$\dot{\hat{\gamma}} = \frac{1}{i\hbar}\left[\hat{\mathrm{H}}, \hat{\gamma}\right], \tag{2.19}$$

where $[\ldots]$ denotes the commutator, the quantum analogue of the Poisson bracket. Contrary to $\left|\Psi^{(i)}\right\rangle$, $\hat{\gamma}$ is phase-independent. It is in this equation that the density operator formalism reveals all of its beauty (we note that nothing keeps us from applying it to a single state vector, in which case eq. 2.19 is nothing but a complementary expression for the time-dependent Schrödinger equation). By further exploring the analogy to classical mechanics, we may follow the classical Liouvillian, $\mathcal{L}$, and define the quantum Liouville operator as:

$$i\hat{\mathrm{L}} = \frac{1}{i\hbar}\left[\,\cdot\,, \hat{\mathrm{H}}\right]. \tag{2.20}$$

We find the time evolution of the density operator to be:

$$\dot{\hat{\gamma}}(t) = e^{-i\hat{\mathrm{L}}t}\hat{\gamma}(0), \tag{2.21}$$

just as we know it from classical Hamiltonian mechanics.

Consider now a system in some equilibrium ensemble. Just as its classical analogue, a quantum system at equilibrium is characterised by $[\hat{\gamma}, \hat{\mathrm{H}}] = 0$, and therefore, the $\{c_j(t)\}$ must obey $|c_j(t)|^2 = |c_j(0)|^2$, and the same goes for the $\{p_i\}$. We can therefore drop the time-dependency. In the canonical ensemble, the weights p_i will be given by a Boltzmann distribution (by definition): The probability associated with any state i is thus $p_i = \exp(-\beta E_i)$. We may hence rewrite $\hat{\gamma}$ as the *thermal density operator* $\hat{\rho}$ (note the omitted dependency on t)[51, 54]:

$$\hat{\rho} = \sum_i e^{-\beta\hat{\mathrm{H}}}\left|\Psi^i\right\rangle\left\langle\Psi^i\right|. \tag{2.22}$$

If one expands the above expression in an eigenbasis of $\hat{\mathrm{H}}$, one arrives at

$$\hat{\rho} = \sum_i e^{-\beta E_i}. \tag{2.23}$$

This is nothing but the familiar, classical partition function expressed in terms of discrete energies, rather than an integral. We now recognise that eq. 2.23 corresponds to the trace of $\hat{\rho}$ in its eigenbasis. Since the trace of an operator is independent of the choice of basis, the quantum partition function Z may be written more generally as:

$$Z = \mathrm{Tr}\left(\hat{\rho}\right). \tag{2.24}$$

Just as in eq. 2.14, any observable $\langle \hat{\mathrm{O}} \rangle$ can be obtained from Z, by tracing over $\hat{\rho}$ like it was done for the density operator $\hat{\gamma}$ in eq. 2.16:

$$\langle \hat{\mathrm{O}} \rangle = \frac{1}{Z} \mathrm{Tr}\left(\hat{\mathrm{O}}\hat{\rho}\right). \tag{2.25}$$

Based on the classical partition functions for various ensembles, their analogues for quantum systems are then easily derived. Given that the time evolution of the density operator is known exactly and that any observable can be obtained from tracing, the full quantum problem can be described by a set of simple, but strikingly beautiful equations.

Only, it cannot be *solved*. For if its formal properties are known, a practical solution of the above equations is impossible apart from a few small (toy) systems. The problem lies in the representation of $|\Psi\rangle$: Spanning a complex Hilbert space on their own, the state vectors remain evasive to direct interpretation. And even though our familiar Euclidean space is a special case of a Hilbert space (and the mathematical extension from 3D Cartesian to ∞-dimensional Hilbert space can be rather intuitive), projecting the state vectors onto a real space basis does not solve the problem: In real space, the dependency on $3N$ Cartesian coordinates implies that for all but the smallest systems, a single wavefunction $\langle \mathbf{r}|\Psi^i\rangle = \Psi^i(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ becomes in itself an untractable object; let alone the projection of $\hat{\gamma}$, which becomes nonlocal, thus exacerbating the dimensionality issue: $\gamma(\mathbf{r}'_1, \ldots, \mathbf{r}'_N; \mathbf{r}_1, \ldots, \mathbf{r}_N)$. As we shall see later on, we may safely integrate over *most* of the N coordinates, but that does not make things substantially easier.

But the situation is not all as grim as it seems. Let us first tackle the prominent issue of nonlocality of γ; any worries about the remaining (huge number of) dimensions of Ψ we shall have to postpone to a later point.

2.2.2 Time-evolution and Ensemble Averages: The Ergodic Hypothesis

If we know that the probability of our system of interest to be in a certain state is governed by an equation of the form 2.25, would it not be possible to observe the very same ensemble for a sufficiently large amount of time while keeping track of the states that are explored? The (partial) answer to this question is given by the ergodic hypothesis.[49, 51] According to the Liouville equation, the phase space measure is invariant under time for Hamiltonian systems. Therefore, if the microstates in phase space are distributed uniformly at some point in time, they shall remain so in the future and at all times. Therefore, for such a system, one has

$$\langle \hat{\mathrm{O}} \rangle = \frac{1}{Z} \operatorname{Tr}\left(\hat{\mathrm{O}}\hat{\rho}\right) = \lim_{\tau\to\infty} \frac{1}{\tau} \int_{t=0}^{t=\tau} \mathrm{d}t \, \operatorname{Tr}\left(\hat{\mathrm{O}} \left|\Psi(t)\right\rangle \left\langle\Psi(t)\right|\right), \tag{2.26}$$

where we write the trace over a pure-state density operator, which is equivalent to evaluating $\langle \hat{\mathrm{O}}(t) \rangle = \langle \Psi(t) | \, \hat{\mathrm{O}} \, | \Psi(t) \rangle$. This assumption is, however, not valid for *all* Hamiltonian systems; hence the name ergodic *hypothesis*, rather than theorem. (The main issue lies in the uniform distribution of the microstates at $t = 0$.)

The ergodic hypothesis offers an appealing alternative to the computation (or direct sampling[23]) of the partition function: If the time evolution of a system is followed for sufficiently large times τ, the time average of the observable will approach its ensemble-average. Even though ergodicity is not proved for many systems - and it is even disproved for others - evidence suggests that it holds in many cases; it is this assumption that opens the door to *Molecular Dynamics*[57] (MD) simulations. By computing the trajectory, *i.e.* the time evolution of a system under appropriate simulation conditions, all thermodynamic quantities become accessible without having to resort to the density matrix or any other form of the partition function.

We have now found an elegant way of sampling the distribution generated by $\hat{\rho}$ without computing $\hat{\rho}$ itself. However, the dimensionality problem associated to the wavefunction $\Psi(t)$ that is to be propagated over time has yet to be resolved. We shall start with a very basic problem founded in chemical intuition: How to separate a single, abstract state vector into nuclear and electronic components.

2.3 Molecular Dynamics in the Born-Oppenheimer Formulation

This is an annotated post-print version of a chapter published in:

Bircher, Martin P.; Liberatore, E.; Browning, N. J.; Brickel, S.; Hofmann, C.; Patoz, A.; Unke, O. T.; Zimmermann, T.; Chergui, M.; Hamm, P.; Keller, U.; Meuwly, M.; Woerner, H.-J.; Vaníček, J.; Rothlisberger, U. *Structural Dynamics* **2017**, *4*, 061510

We have seen that, given appropriate initial conditions, the ensemble average over the thermal density operator can be replaced by a time-average based on a pure state. We shall therefore base our discussion on the time-(in)dependent Schrödinger equation for a single state vector, circumventing the density operator.

2.3.1 Separating the Hamiltonian: The Born-Oppenheimer Approximation

The behaviour of a pure, closed, non-relativistic quantum system is completely characterised by the time-dependent Schrödinger equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \left|\Psi(t)\right\rangle = \hat{\mathrm{H}} \left|\Psi(t)\right\rangle , \tag{2.27}$$

where $\left|\Psi\right\rangle$ is a single state vector in Hilbert space. For a molecular system, the time-independent Hamiltonian $\hat{\mathrm{H}}_{\mathrm{mol}}$ is given by:

$$\hat{\mathrm{H}}_{\mathrm{mol}} = \sum_{\alpha} \frac{\hat{\mathbf{P}}_{\alpha}^{2}}{2M_{\alpha}} + \hat{\mathrm{H}}_{\mathrm{el}}, \tag{2.28}$$

$$\begin{aligned} \hat{\mathrm{H}}_{\mathrm{el}} &= \sum_{i} \frac{\hat{\mathbf{p}}_{i}^{2}}{2} + \sum_{i<j} \frac{1}{\hat{\mathrm{r}}_{\mathrm{ij}}} - \sum_{\alpha,i} \frac{Z_{\alpha}}{\left|\hat{\mathbf{R}}_{\gamma} - \hat{\mathbf{r}}_{i}\right|} + \sum_{\alpha<\beta} \frac{Z_{\alpha}Z_{\beta}}{\hat{\mathrm{R}}_{\alpha\beta}} && (2.29) \\ &= \hat{\mathrm{T}}_{\mathrm{el}} + \hat{\mathrm{V}}_{\mathrm{ee}} + \hat{\mathrm{V}}_{\mathrm{eN}} + \hat{\mathrm{V}}_{\mathrm{NN}}, && (2.30) \end{aligned}$$

where the *electronic Hamiltonian* $\hat{\mathrm{H}}_{\mathrm{el}}$ has been introduced. The eigenvalues of the position operators $\hat{\mathbf{r}}$ and $\hat{\mathbf{R}}$ are the set of electronic and nuclear coordinates described by the collective variables $\{\mathbf{r}\}$ and $\{\mathbf{R}\}$, respectively, while $\hat{\mathbf{p}}$ and $\hat{\mathbf{P}}$ denote the corresponding momentum operators. The molecular Hamiltonian contains kinetic terms due to the nuclei α and electrons i, the potential energy due to the interaction of electrons and nuclei ($\hat{\mathrm{V}}_{\mathrm{eN}}$), as well as nuclear-nuclear ($\hat{\mathrm{V}}_{\mathrm{NN}}$) and electron-electron ($\hat{\mathrm{V}}_{\mathrm{ee}}$) repulsion terms.

The projection of the time-dependent state vector in real space, $\Psi(\mathbf{r}, \mathbf{R}, t)$ depends both on nuclear and electronic coordinates and remains, in this form, an untractable object for most systems of interest. If the molecular Hamiltonian were separable, a simplification could easily be achieved by a factorisation of the wavefunction into a nuclear and electronic component.[54,55] However, due to the presence of $\hat{\mathrm{V}}_{eN}$, the molecular Hamiltonian is not separable. The question on how to separate nuclear and electronic degrees of freedom has therefore been of paramount importance to molecular quantum mechanics, be it for static or dynamic approaches.

Indeed, a first ansatz to this problem was proposed by Born and Oppenheimer as early as 1927,[58] and was later generalised by Born and Huang in 1954.[59] The different time scales of electronic and nuclear motion lie at the very heart of their approach. A non-separable Hamiltonian may be written as the tensor product of two subsystems; in particular, $\hat{\mathrm{H}}_{\mathrm{mol}} = \hat{\mathrm{H}}_{\mathrm{fast}} \otimes \hat{\mathrm{H}}_{\mathrm{slow}}$, where one Hamiltonian is due to the fast motion of the electrons, and the other is due to the slower, nuclear components.[55] The spectrum of $\hat{\mathrm{H}}_{\mathrm{fast}}$ can then be expanded in terms of *electronic* eigenstates $|\Phi\rangle$ by taking the limit of clamped (or 'frozen') nuclei, a limit in which the kinetic contribution of the nuclei vanishes. Only terms due to the electronic Hamiltonian remain, $\hat{\mathrm{H}}_{\mathrm{fast}} = \hat{\mathrm{H}}_{\mathrm{el}}$; since the potential energy still contains the terms $\hat{\mathrm{V}}_{\mathrm{NN}}$ and $\hat{\mathrm{V}}_{\mathrm{eN}}$, the dependency of the Hamiltonian on the nuclear coordinates $\mathbf{R}$ becomes parametric. For any nuclear configuration $\mathbf{R}$, one can therefore obtain a set of electronic eigenstates:

$$\hat{\mathrm{H}}_{\mathrm{el}}(\mathbf{R}) \, |\Phi_l; \mathbf{R}\rangle = E_l(\mathbf{R}) \, |\Phi_l; \mathbf{R}\rangle \, , \tag{2.31}$$

where the eigenfunctions $|\Phi_l; \mathbf{R}\rangle$ depend parametrically on $\mathbf{R}$ through $\hat{\mathrm{V}}_{eN}$. Since $\hat{\mathrm{H}}_{\mathrm{el}}$ and $\hat{\mathbf{R}}$ commute, the basis of the molecular Hamiltonian $\hat{\mathrm{H}}_{\mathrm{mol}}$ can be constructed from the direct product of eigenfunctions of $\hat{\mathbf{R}}$ and the eigenstates of the fast Hamiltonian,[55] $|\mathbf{R}, \Phi_l; \mathbf{R}\rangle = |\mathbf{R}\rangle \otimes |\Phi_l; \mathbf{R}\rangle$. After introducing a resolution of identity in this basis into the real-space projection of the state vector, $\Psi(\mathbf{r}, \mathbf{R}, t) = \langle \mathbf{r}, \mathbf{R} | \Psi(t) \rangle$, by orthogonality, one obtains a factorised expression for $\Psi(\mathbf{r}, \mathbf{R}, t)$:

$$\begin{aligned} \Psi(\mathbf{r}, \mathbf{R}, t) = \langle \mathbf{r}, \mathbf{R} | \Psi(t) \rangle &= \sum_l \langle \mathbf{r} | \Phi_l; \mathbf{R} \rangle \langle \mathbf{R}, \Phi_l; \mathbf{R} | \Psi(t) \rangle & (2.32) \\ &= \sum_l \Phi_l(\mathbf{r}; \mathbf{R}) \chi_l(\mathbf{R}, t). & (2.33) \end{aligned}$$

Here the nuclear components $\chi_l(\mathbf{R}, t)$ (or 'nuclear wavefunctions', although they are no wavefunctions *stricto sensu*) are projections of the configurational basis onto the Hilbert space vector $|\Psi(t)\rangle$, and the $\Phi_l(\mathbf{r}; \mathbf{R})$ are the electronic wavefunctions of a system with configuration $\{\mathbf{R}\}$. Note that the $\Phi_l(\mathbf{r}; \mathbf{R})$ are time independent. This is commonly referred to as the Born-Oppenheimer[58] or Born-Huang[59] ansatz for the total wavefunction. The problem has now been conveniently split in two parts: In line with chemical intuition, the electronic structure problem can be treated independently from the nuclear wavefunction; the latter is, however, less straightforward to interpret, since its definition involves a component-wise projection of the total wavefunction of the system.

2.3.2 Nuclear Dynamics with the Born-Huang Expansion

Depending on the expansion (or: truncation) adopted in eq. 2.32, the Born-Huang ansatz can give rise to a hierarchy of different nuclear dynamics regimes.

The description of the nuclear dynamics of the system is obtained by inserting the ansatz 2.32 into the time-dependent Schrödinger equation. The resulting coupled-channels equation describes the exact time-evolution of the nuclear dynamics. Written explicitly in terms of $\hbar$ rather than atomic units:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\chi_k(\mathbf{R}, t) = &\left[-\sum_\alpha \frac{\hbar^2}{2M_\alpha}\nabla_\alpha^2 + E_k(\mathbf{R})\right]\chi_k(\mathbf{R}, t) \\ &+ \sum_l \mathbf{C}_{kl}\chi_l(\mathbf{R}, t), \end{aligned} \tag{2.34}$$

$$\mathbf{C}_{kl}(\mathbf{R}, t) = -\sum_\alpha \frac{\hbar^2}{2M_\alpha} D_{kl}^\alpha(\mathbf{R}, t) + \sum_\alpha \frac{\hbar^2}{M_\alpha}\mathbf{d}_{kl}^\alpha(\mathbf{R}, t)\nabla_\alpha. \tag{2.35}$$

The terms collected under $\mathbf{C}_{kl}(\mathbf{R}, t)$ are mediating between different electronic states and are referred to as the *nonadiabatic coupling* terms, with the scalar quantity $D_{kl}^\alpha(\mathbf{R})$ being the *kinetic* coupling, and the vectorial quantity $\mathbf{d}_{kl}^\alpha(\mathbf{R})$ being the *derivative* coupling:

$$D_{kl}^\alpha(\mathbf{R}) = \langle\Phi_k; \mathbf{R}|\, \hat{\mathrm{D}}^\alpha(\mathbf{R})\, |\Phi_l; \mathbf{R}\rangle\,, \tag{2.36}$$

$$\hat{\mathrm{D}}^\alpha(\mathbf{R}) = \nabla_\alpha^2, \tag{2.37}$$

$$\mathbf{d}_{kl}^\alpha(\mathbf{R}) = \langle\Phi_k; \mathbf{R}|\, \hat{\mathrm{d}}^\alpha(\mathbf{R})\, |\Phi_l; \mathbf{R}\rangle\,, \tag{2.38}$$

$$\hat{\mathrm{d}}^\alpha(\mathbf{R}) = \nabla_\alpha. \tag{2.39}$$

These terms are responsible for part of the nuclear quantum effects.

In the *Born-Oppenheimer* (BO) approximation in its proper sense, these coupling terms are neglected by setting $D^{\alpha}_{kl}(\mathbf{R}) = 0$ and $\mathbf{d}^{\alpha}_{kl}(\mathbf{R}) = \mathbf{0}$ (which is equivalent to neglecting the effect of the nuclear kinetic energy operator on the electronic wavefunction). This results in a greatly simplified ansatz for the total wavefunction and a less complex nuclear dynamics:

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \Phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}, t), \tag{2.40}$$

$$i\hbar\frac{\partial}{\partial t}\chi(\mathbf{R}, t) = \left[-\sum_{\alpha}\frac{\hbar^2}{2M_\alpha}\nabla^2_\alpha + E(\mathbf{R})\right]\chi(\mathbf{R}, t). \tag{2.41}$$

This simplification is often justified by the fact that the coupling terms are small and weighted by the inverse of the (heavy) nuclear masses (which leads to the often stated simplification that the BO approximation separates nuclear from electronic degrees of freedom due to differences in mass).

With this particular split in mind, we still have not discussed the practical form of $\chi(\mathbf{R}, t)$, which depends on $3N$ nuclear coordinates. It is the most *brut* approximation that is the most common one: Chemical processes are governed by electrons rearranging - following an organic chemists' drawing scheme, one may chose to neglect the quantum nature of the nuclei and consider them to be simple, classical objects. The limit of classical nuclei (point particles or delta functions centred at $\mathbf{R}$) may be recovered by rewriting the nuclear components $\chi(\mathbf{R})$ in polar representation[16, 60, 61]:

$$\chi(\mathbf{R}, t) = A(\mathbf{R}, t)\exp\left[\frac{i}{\hbar}S(\mathbf{R}, t)\right] \tag{2.42}$$

Insertion of the above expression (2.42) into eq. (2.41), separating real and imaginary parts and taking the classical limit $\hbar \rightarrow 0$, results in an isomorphism with the classical Hamilton-Jacobi equation of motion:

$$\frac{\partial S(\mathbf{R}, t)}{\partial t} + \sum_{\alpha}\frac{1}{2M_\alpha}\left(\nabla_\alpha S(\mathbf{R}, t)\right)^2 + E(\mathbf{R}) = 0 \tag{2.43}$$

In BO dynamics, the nuclei evolve according to the forces due to a single electronic state. In the limit of classical nuclei, this results in delta functions being propagated on a single potential energy surface (PES), neglecting all nuclear quantum effects. Some of these nuclear quantum effects such as zero-point energy and tunnelling may be recovered by resorting to techniques such as the path integral formalism.[62] Still, in order to describe the complete array of nuclear quantum effects, the nonadiabatic coupling elements must imperatively be included, since any dynamics based on the BO formalism forbids the nuclei - described as classical point charges or nuclear wave packets - to switch between potential energy surfaces; BO dynamics is therefore a strict *single state* dynamics.

Should we wish to assess possible regions of breakdown, it proves convenient to rewrite the leading derivative coupling term in eq. 2.38 in a more intuitive form that depends on the eigenvalues of $\hat{\mathrm{H}}_{\mathrm{el}}$[63]:

$$\mathbf{d}^{\alpha}_{kl}(\mathbf{R}) = \frac{\langle \Phi_k; \mathbf{R} |\, \nabla_\alpha \hat{\mathrm{H}}_{\mathrm{el}} \,| \Phi_l; \mathbf{R} \rangle}{E_l(\mathbf{R}) - E_k(\mathbf{R})}. \tag{2.44}$$

This expression reveals that the nonadiabatic coupling terms become important whenever two eigenstates are close in energy. In the case of a conical intersection, where the eigenvalues of $\hat{\mathrm{H}}_{\mathrm{el}}$ are degenerate, they even diverge. In such systems, the Born-Oppenheimer approximation breaks down and with it the picture of classical nuclei evolving on one single potential energy surfaces is no longer suitable. Due to faster nuclear motion, the nuclear wavepacket cannot be thought of as being localised on one PES, but instead spreads over several electronic states. In order to properly describe the dynamics of those systems, it is therefore necessary to go beyond the BO and to include the nuclear quantum effects that lead to couplings between electronic states. We shall briefly review two semiclassical approaches that allow nonadiabatic effects to be incorporated into the dynamics.

Semiclassical Methods: Ehrenfest Dynamics An alternative possibility to include the effects of several electronic states within a semiclassical approach is based on an ansatz for the total wavefunction of:

$$\Psi(\mathbf{r}, \mathbf{R}; t) = \Phi(\mathbf{r}; t)\chi(\mathbf{R}; t) \exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} E_{el}(t')\mathrm{d}t'\right] \tag{2.45}$$

$$E_{\mathrm{el}}(t) = \iint \Phi^*(\mathbf{r}, t)\chi^*(\mathbf{R}, t)\hat{\mathrm{H}}_{\mathrm{el}}(\mathbf{r}, \mathbf{R})\chi(\mathbf{R}, t)\Phi(\mathbf{r}, t)\, \mathrm{d}\mathbf{R}\, \mathrm{d}\mathbf{r}. \tag{2.46}$$

This single-configuration ansatz gives rise to *Ehrenfest dynamics*.[61] In contrast to the BO formalism, the time-dependent electronic wavefunction $\Phi(\mathbf{r}; t)$ exhibits no dependence on the nuclear coordinates at all, not even parametrically. Instead, an additional exponential term, the *phase term* is introduced. By inserting this ansatz into the time-dependent Schrödinger equation, the evolution of the nuclear and electronic wavefunctions are given by the Time-Dependent Self-Consistent Field (TDSCF) equations[64]:

$$i\hbar\frac{\partial}{\partial t}\Phi(\mathbf{r},t) = \left[\hat{\mathrm{T}}_{\mathrm{el}} + \int \mathrm{d}\mathbf{R}\ \chi^*(\mathbf{R})\hat{\mathrm{V}}(\mathbf{r},\mathbf{R})\chi(\mathbf{R})\right]\Phi(\mathbf{r},t), \tag{2.47}$$

$$i\hbar\frac{\partial}{\partial t}\chi(\mathbf{R},t) = \left[\hat{\mathrm{T}}_{\mathrm{N}} + \int \mathrm{d}\mathbf{r}\ \Phi^*(\mathbf{r},t)\hat{\mathrm{H}}_{\mathrm{el}}(\mathbf{r},\mathbf{R})\Phi(\mathbf{r},t)\right]\chi(\mathbf{R},t), \tag{2.48}$$

where $\hat{\mathrm{V}} = \hat{\mathrm{V}}_{\mathrm{ee}} + \hat{\mathrm{V}}_{\mathrm{eN}} + \hat{\mathrm{V}}_{\mathrm{NN}}$. The nuclei evolve in a time-dependent *mean field* of the electronic states, and vice versa. As described above for the Born-Oppenheimer ansatz, the limit of classical nuclei can be recovered by inserting the polar representation of the nuclear wavefunction, (2.42), into the time derivative of eq. (2.48) and taking the limit $\hbar \to 0$.[61] The resulting equation is again isomorphic to the Hamilton-Jacobi equation and may be further recast to yield a Newton-like equation for the nuclei:

$$-\mathbf{F}_\alpha(t) = \nabla_{\mathbf{R}_\alpha}\int \mathrm{d}\mathbf{r}\ \Phi^*(\mathbf{r},t)\hat{\mathrm{H}}_{\mathrm{el}}(\mathbf{r},\mathbf{R})\Phi(\mathbf{r},t). \tag{2.49}$$

According to (2.49), the classical nuclei evolve on an ‘averaged PES’ due to $\Phi(\mathbf{r},t)$. The corresponding equation for the electronic degrees of freedom is obtained by writing $\chi(\mathbf{R},t) = \sum_\alpha \delta(\mathbf{R}(t) - \mathbf{R}_\alpha(t))$; it is simply the time-dependent Schrödinger equation for Φ, with a (re)introduced parametric dependence on $\mathbf{R}$:

$$i\hbar\frac{\partial\Phi(\mathbf{r};\mathbf{R}(t),t)}{\partial t} = \hat{\mathrm{H}}_{\mathrm{el}}(\mathbf{r},\mathbf{R})\Phi(\mathbf{r};\mathbf{R}(t),t). \tag{2.50}$$

Therefore, the method lends itself to be combined with time-dependent density functional theory (TDDFT),[65] where Φ in (2.49) can be expanded in terms of time-dependent Kohn-Sham noninteracting Slater determinants $\Phi(\mathbf{r},t) = \det|\psi_1(\mathbf{r},t),\ldots,\psi_N(\mathbf{r},t)|$, and $\mathbf{F}_\alpha$ is derived using the Hellmann-Feynman theorem. Eqs. (2.49) and (2.50) have to be solved simultaneously, which can be carried out on-the-fly *e.g.* by using a Runge-Kutta integrator.[65]

Due to the averaged nature of $\Phi(\mathbf{r},t)$, Ehrenfest dynamics is a suitable choice whenever the classical trajectories due to different electronic states do not differ considerably; such as when the relaxation of the electronic degrees of freedom is fast with respect to the nuclear motion. Otherwise, *e.g.* for molecular dissociations, the mean-field approximation may introduce large errors: After leaving regions of strong nonadiabaticity, the nuclei are unable to collapse on either of the PES, making their dynamics potentially unphysical.[61] Ehrenfest dynamics is therefore typically limited to ultrafast or instantaneous processes.

Semiclassical Methods: Trajectory Surface Hopping Trajectory Surface Hopping (TSH)[66] offers another possibility of including nonadiabatic effects by independently propagating a swarm of particles on different, single electronic states, and allowing them to 'hop' onto a different state according to some predefined hopping probability. As in the case of Ehrenfest dynamics, this propagation can be carried out on-the-fly.

Given a set of initial nuclear coordinates and velocities, trajectories are propagated on a single electronic state according to the Born-Oppenheimer scheme. Subsequently, the probability for a jump is calculated. In a first approximation, the Landau-Zener[67,68] transition probability may be used. Originally formulated in a diabatic framework, some simplifications allow for it to be reformulated based on adiabatic quantities[16]:

$$P(t) \approx \exp\left(-\frac{\pi}{2\hbar}\frac{\min\left\{|\Delta E_{01}(t)|^2\right\}}{\max\left\{\frac{\partial|\Delta E_{01}(t)|}{\partial t}\right\}}\right). \tag{2.51}$$

Here, P is the probability to perform a nonadiabatic transition at an avoided crossing for a two-state system and ΔE_{01} is the adiabatic electronic gap.

A more rigorous approach to the transition probabilities is given by Tully's 'Fewest Switches' formulation of TSH.[69] A set of complex amplitudes $\{C(t)\}$ is assigned to every trajectory to quantify the degree of nonadiabadicity during the propagation. The complex amplitudes themselves evolve along each trajectory according to:

$$i\hbar\frac{\partial C_j(t)}{\partial t} = \sum_i^\infty C_i(t)\left[E_i^{\mathrm{el}}(\mathbf{R})\delta_{ij} - i\hbar\sigma_{ij}(\mathbf{R},t)\right], \tag{2.52}$$

where $\sigma_{ij}(\mathbf{R},t) = \mathbf{d}_{ij}(\mathbf{R})\cdot\dot{\mathbf{R}}$ and $\mathbf{d}_{ij}(\mathbf{R})$ is the derivative coupling as defined in eq. (2.38). After integration, the probability of hopping from state j to state i within an infinitesimal time interval dt is computed:

$$g_{ij}(t, t+\mathrm{d}t) = 2\int_t^{t+\mathrm{d}t}\mathrm{d}\tau\frac{-\Re\left[C_i(\tau)C_j^*(\tau)\sigma_{ij}\right]}{C_j(\tau)C_j^*(\tau)}. \tag{2.53}$$

The hop is then accepted or rejected according to a Metropolis criterion by comparing the hopping probability to a random number $\zeta \in [0,1]$:

$$\sum_{k\leq i-1} g_{jk} < \zeta < \sum_{k\leq i} g_{jk}. \tag{2.54}$$

Total energy conservation is ensured by rescaling the nuclear velocities after an accepted hop. The trajectory γ is now further propagated along the new electronic state i in an adiabatic fashion, until the next hopping attempt.

The terms needed in trajectory surface hopping can be rigorously derived from linear response time-dependent density functional theory (LR-TDDFT).[70] The sum in eq. (2.52) may be truncated after N adiabatic states, and the (diagonal) electronic eigenvalues E_i^{el} may be replaced by a relative term, $\tilde{V}_{ij} = \left[E_i^{\mathrm{el}} - E_0^{\mathrm{el}}\right]\delta_{ij}$. $\tilde{V}_{ij}$ has an analogue in the TDDFT excitation energies ω_i, and eq. (2.52) may be recast in terms of a set of transformed coefficients $\left\{\tilde{C}(t)\right\}$ and the TDDFT excitation energies $\{\omega\}$:

$$i\hbar\frac{\partial\tilde{C}_i(t)}{\partial t} = \tilde{C}_i(t)\omega_i - i\hbar\sum_j^N \tilde{C}_j(t)\sigma_{ij}(\mathbf{R},t). \tag{2.55}$$

The σ_{ij} are accessible *via* finite differences:

$$\begin{aligned}\sigma_{ij}(\mathbf{R},t)\Big|_{t+\mathrm{d}t/2} =& \frac{1}{2\mathrm{d}t}\left[\int \mathrm{d}\mathbf{r}\ \Phi_i^*(\mathbf{r};\mathbf{R}(t))\Phi_j(\mathbf{r};\mathbf{R}(t+\mathrm{d}t))\right.\\ &\left. - \int \mathrm{d}\mathbf{r}\ \Phi_i^*(\mathbf{r};\mathbf{R}(t+\mathrm{d}t))\Phi_j(\mathbf{r};\mathbf{R}(t))\right].\end{aligned} \tag{2.56}$$

After integration of eq. (2.52), the switching probability is computed according to eq. (2.54), using $\tilde{C}_j$ rather than C_j and by linearly interpolating $\sigma_{ij}(\tau)$. This approach has been implemented in combination with a plane wave/pseudopotential formalism.[70, 71]

The nuclear dynamics in TSH are still governed by a single-state Born-Oppenheimer formalism, but the propagation of the complex amplitudes and the calculation of the hopping probability allow for a transition between states. Since TSH is carried out for a collection $\{\gamma\}$ of trajectories, this hopping can mimick the spread of the wavepacket in nonadiabatic regions. However, nuclear quantum effects such as tunnelling and the zero-point energy are not described, due to the classical nature of the individual trajectories. For a more extensive review of nonadiabatic dynamics, see *e.g.* Refs 16, 61, 72, 73.

Comments on Single-State Dynamics Single-state BO dynamics of classical point particles, with none of the aforementioned quantum corrections applied, remain reliable and sufficiently accurate for a vast majority of systems. The simplification may be taken even further by using a *classical* expression for the PES $E(\mathbf{R})$ on which the system evolves: Parametrised interactions are at the base of the highly successful *classical Molecular Dynamics* simulations that have provided valuable insight into dynamics and thermodynamics in both the solid state and condensed matter systems.

In this text, however, we wish to pursue the quantum path to the electronic hypersurface: By invoking the Born-Oppenheimer approximation and relying on classical nuclear dynamics, we are only left with one unknown, $|\Phi_l(\mathbf{r};\mathbf{R})\rangle$. But even with the explicit dependency on the nuclear coordinates $\mathbf{R}$ removed, we are left with an unknown, high-dimensional function depending on 3N Cartesian electron coordinates $\mathbf{r}$. This unknown, however, we may avoid: The single state dynamics in eq. 2.41 does not explicitly refer to the electronic wavefunction, but only to the electronic potential energy surface. It is in the next chapter that we will present an approach to find $E_l(\mathbf{R})$ without any explicit usage of the electronic wavefunction $|\Phi_l(\mathbf{r};\mathbf{R})\rangle$ and its untractable dependency on $\mathbf{r}$.

Chapter 3

The Electronic Structure Problem

In 1964, Hohenberg and Kohn published their two theorems for the nondegenerate ground state of an interacting electron gas in an external potential $v(\mathbf{r})$.[14] The first theorem states that the external potential $v(\mathbf{r})$ is, up to a constant, uniquely determined by the density $\rho(\mathbf{r})$. Since $v(\mathbf{r})$ fixes the Hamiltonian of the system, *'the many-particle ground state is a unique functional of the density'*. Therefore, a density functional $F[\rho]$ exists such that the energy E_v of a system subjected to the potential v is given by

$$E_v[\rho] = \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} + F[\rho]. \tag{3.1}$$

The second theorem proves the validity of a variational principle for $\rho(\mathbf{r})$, namely that *'the density minimising the total energy corresponds to the ground-state density'*.

$$E^0[v] = \min_{\rho}\left(E_v[\rho]\right). \tag{3.2}$$

The Hohenberg-Kohn theorems are the fundamental basis of a density functional theory (DFT), which has revolutionised Computational Chemistry. Instead of solving an eigenvalue problem for a high-dimensional wavefunction, it becomes possible to calculate the ground-state energy of a system - and, by extension, its properties - by resorting to a simple, 3-dimensional Cartesian observable: The electron density $\rho(\mathbf{r})$.

3.1 The Universal Density Functional

If the Hohenberg-Kohn (HK) theorems were revolutionary, there were some serious uncertainties and problems associated with the representability of

$\rho(\mathbf{r})$ and $v(\mathbf{r})$.[17,74–77] It is the work of Levy[74] and Lieb[75] that has put DFT on a rigorous mathematical basis and that has formally resolved these issues. In the following, we shall illustrate their proof of the existence of a truly *universal density functional* by recapitulating parts of Lieb's work.

3.1.1 Levy-Lieb Density Functional Theory

> *[The] exact theory of DFT is almost useless.*
> Walter Kohn, as reported by Kieron Burke during a lecture[78]

Almost. As intriguing as it is, HK-DFT is mathematically awkward - Lieb commented that *'the search space is filled with holes.'*[75] It all starts with the very basic problem of N-representability: If the minimisation in eq. 3.2 is fed with *some* $\rho(\mathbf{r})$, there is no guarantee that this $\rho(\mathbf{r})$ corresponds to a density of an N-electron system. Fractional electrons are chemically rather unacceptable (although of great conceptual importance). Furthermore, in eq. 3.2, the trial space in the $\min_\rho$ must imperatively be a convex space of *ground-state* densities in order for such a search to be performed at all.

The caveat is in the definition of the density functional itself. Hohenberg and Kohn defined the density functional according to eq. 3.1[14]:

$$F_{HK}[\rho] = E[v] - \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} \tag{3.3}$$

$$= \langle \Psi^v_{GS} | \, \hat{\mathrm{T}} + \hat{\mathrm{W}} \, | \Psi^v_{GS} \rangle \,, \tag{3.4}$$

where ${}^{v}_{GS}$ denotes a ground state due to the external potential v, $\hat{\mathrm{T}}$ denotes the kinetic energy operator and $\hat{\mathrm{W}} = \sum_{i<j} \hat{\mathrm{r}}_{ij}^{-1}$ is the Coulomb operator. $F_{HK}[\rho]$ is only defined for $\rho \in \mathscr{A}_N$ and for $v \in \mathscr{V}_N$:

$$\mathscr{A}_N = \{\rho \,|\, \rho \text{ comes from a ground state}\} \,, \tag{3.5}$$

$$\mathscr{V}_N = \left\{ v \,\middle|\, \hat{\mathrm{H}}[v] \text{ has a ground state} \right\} , \tag{3.6}$$

and we therefore have to search over all ground-state densities to find *our* ground-state density for a given v. More precisely, one would therefore have to write the original HK variational principle as:

$$E[v] = \min \left\{ F_{HK}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} \,\middle|\, \rho \in \mathscr{A}_N \right\} , \tag{3.7}$$

for an *unknown* $v \in \mathscr{V}_N$, and restricted to an *unknown* set $\mathscr{A}_N$. We especially do not know whether our search space is convex, and we are therefore left in the dark about whether or how we can take a functional derivative[77] - which we would need if we wish to find ρ^v_{GS}.

The criteria that account for N-representability are linked to the finiteness of T, with $\int \left|\nabla\rho^{1/2}(\mathbf{r})\right|^2 \mathrm{d}\mathbf{r} < T$ and $T < \infty$:

$$\rho(\mathbf{r}) \geq 0\,, \quad \int \rho(\mathbf{r})\mathrm{d}\mathbf{r}\,, \quad \text{and} \quad \int \left|\nabla\rho(\mathbf{r})^{1/2}\right|^2 \mathrm{d}\mathbf{r} < \infty. \tag{3.8}$$

It can also be shown that $\rho(\mathbf{r})^{1/2} \in H^1(\mathbb{R}^3)$, *i.e.* both $\rho(\mathbf{r})^{1/2}$ and $\nabla\rho(\mathbf{r})^{1/2}$ are in an L^2-space. Based on these conditions, one may now properly define a search space $\mathscr{I}_N$:

$$\mathscr{I}_N = \left\{\rho \,\middle|\, \rho(\mathbf{r}) \geq 0,\ \int \rho(\mathbf{r})\mathrm{d}\mathbf{r} = N,\ \rho(\mathbf{r})^{1/2} \in H^1\left(\mathbb{R}^3\right)\right\}, \tag{3.9}$$

which Lieb has demonstrated to be convex.[75,77] But there is no guarantee that all densities within $\mathscr{I}_N$ are indeed densities of any ground state. In order for the variational principle to hold soundly, the functional $F_{HK}[\rho]$ should be convex, and it should be defined on the well-defined search space $\mathscr{I}_N$ (or a convex subset thereof). But neither does the domain $\mathscr{A}_N$ of $F_{HK}[\rho]$ contain all of $\mathscr{I}_N$, nor is it convex!

We may, however, introduce another functional - *Levy's universal density functional* (DF).[74] We shall first have to confirm that $E[v]$ is a concave functional in the potential v, such that one can safely write:

$$E[v] = \inf\left\{\langle\Psi|\,\hat{\mathrm{H}}[v]\,|\Psi\rangle \,\middle|\, \Psi \in \mathscr{W}_N\right\}, \tag{3.10}$$

$$\mathscr{W}_N = \left\{\Psi \mid ||\Psi|| = 1, T(\Psi) < \infty\right\}, \tag{3.11}$$

which Lieb proved in theorem 3.1 of Ref. 75, and the appropriate search space is given by $\mathscr{W}_N$. $||\,||$ denotes an L^2 norm. In order for a trial wavefunction to be admissible, it must be normalised and its kinetic energy T must be finite. Any such minimising Ψ then satisfies the time-independent Schrödinger equation, $\hat{\mathrm{H}}[v]\,|\Psi\rangle = E[v]\,|\Psi\rangle$.

We now *define* Levy's universal DF to be:

$$\tilde{F}[\rho] = \inf_{\Psi}\left\{\langle\Psi|\,\hat{\mathrm{H}}[0]\,|\Psi\rangle \,\middle|\, \Psi \mapsto \rho,\ \Psi \in \mathscr{W}_N\right\} \tag{3.12}$$

$$= \inf_{\Psi}\left\{\langle\Psi|\,\hat{\mathrm{T}} + \hat{\mathrm{W}}\,|\Psi\rangle \,\middle|\, \Psi \mapsto \rho, \Psi \in \mathscr{W}_N\right\} = Q[\rho], \tag{3.13}$$

where we have introduced $\hat{\mathrm{H}}[0]$ as the Hamiltonian of a system in absence of an external potential. For a Coulomb system in the *absence* of an external potential, $\tilde{F}[\rho]$ searches the wavefunction space for some ground-state wavefunction $|\Psi\rangle$, the density of which yields ρ. Where the minimum as in eq. 3.2 does not exist, the greatest lower bound is obtained instead. From this, by combining eqs 3.12 and 3.10, we find that:

$$E[v] = \inf_{\rho} \left\{ \tilde{F}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} \;\middle|\; \rho \in \mathscr{I}_N \right\}, \tag{3.14}$$

and therefore, as a special case,

$$\tilde{F}[\rho] = F_{HK}[\rho] \quad \text{if} \quad \rho \in \mathscr{A}_N. \tag{3.15}$$

Lieb then proved that for each ρ in $\mathscr{I}_N$ (eq. 3.9) there is a $\Psi \in \mathscr{W}_N$ (eq. 3.11) such that $\tilde{F}[\rho] = \langle\Psi|\,\hat{\mathrm{H}}[0]\,|\Psi\rangle$, *i.e.* the infimum in eq. 3.12 is a minimum (Lieb's theorem 3.3). But this says nothing about Ψ being a ground-state Ψ, and unfortunately, Levy's $\tilde{F}[\rho]$ itself is not convex either; even though its search space is. For if $N > q =$ number of spin states, $\tilde{F}[\rho]$ is not a convex functional, and there exists a $\rho \in \mathscr{I}_N$ that does not come from a *ground-state* Ψ (Lieb's theorem 3.4), which is a rather subtle point.[77]

These limitations make it necessary to define yet another functional. Lieb showed that the construction of a convex, universal functional is possible by performing the *Lieb maximisation* over potentials v:

$$F[\rho] = \sup_{v} \left\{ E[v] - \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} \;\middle|\; v \in L^{3/2} + L^{\infty} \right\}. \tag{3.16}$$

By defining $F[\rho]$ on a set $X = L^3 \cap L^1$ of L^p-spaces, the restrictions $\rho(\mathbf{r}) \geq 0$ and $\int \rho(\mathbf{r})\mathrm{d}\mathbf{r} = N$ can be lifted, with $F[\rho] = \infty$ if $\rho \notin \mathscr{I}_N$, thus yielding a convex functional. It follows that $F[\rho] \leq \tilde{F}[\rho]$ for all $\rho \in \mathscr{I}_N$. We then find $E[v]$ as the Legendre-Fenchel transform of $F[\rho]$:

$$E[v] = \inf_{\rho} \left\{ F[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} \;\middle|\; \rho \in L^3 \cap L^1 \right\}, \tag{3.17}$$

which is a concave functional, and $F[\rho] = \mathrm{CE}\tilde{F}[\rho]$ for all $\rho \in L^3 \cap L^1$, *i.e.* F is the complex envelope of $\tilde{F}$. Consequently, equality holds when the search space is $\mathscr{I}_N \subset L^3 \cap L^1$. Lieb shows in his theorem 3.7 that F is the *Legendre-Fenchel transform* of $E[v]$, and that the Lieb maximisation can be performed for all concave $E[v]$. In approximate theories where $E[v]$ is not concave, $F[\rho]$ is still well-defined, but will yield an upper bound to $E[v]$.[75,77]

It is therefore possible to characterise a ground-state system based on its density and external potential only, since $E[v]$ and $F[\rho]$ are mutual Legendre-Fenchel transforms. The Hohenberg-Kohn variational principle has been extended to a well-defined, known search space of densities ρ, where the functional returns ∞ whenever the density under consideration does not come from $\mathscr{I}_N$ and is therefore not N-representable. If the potential does not admit a ground-state, the energy is obtained as an infimum instead.

Finally, we note that there is a complementary formulation for $F[\rho]$ in terms of the density operator on an ensemble density matrix:

$$F_{DM}[\rho] = \inf_{\hat{\gamma}\to\rho} \operatorname{Tr}\left(\hat{\mathrm{H}}[0]\hat{\gamma}\right). \tag{3.18}$$

This is the Levy-Lieb constrained search functional for *canonical ensembles* which, due to linearity $\hat{\gamma} \to \rho$ is convex on $\mathscr{I}_N$, and where a minimising $\hat{\gamma}$ always exists. Therefore,

$$E[v] = \inf_{\rho} \left\{ F_{DM}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} \,\middle|\, \rho \in \mathscr{I}_N \right\}. \tag{3.19}$$

Lieb[75]: *'The amusing fact is that:'*

$$F[\rho] = F_{DM}[\rho]\,, \quad \rho \in \mathscr{I}_N, \tag{3.20}$$

which is less obvious than it might seem at first glance.

A Few Notes on the Time-Dependent Case We cannot possibly conclude our discussion of the intricacies of ground-state DFT without mentioning its time-dependent analogue. In 1984, 20 years after the seminal paper by Hohenberg and Kohn, Runge and Gross[79] provided a (rather) general formulation of a time-dependent density functional theory (TDDFT) - a different theory with its own theorems, proofs and issues. Thanks to TDDFT, it has become possible to conveniently calculate excitation energies of molecules, which are available as the poles of the density response function. Many excited-state properties can then be calculated in the linear response regime. However, and just like ground-state DFT, N-representability of the time-dependent density remains an issue, and a rigorous generalisation has been provided by Cohen and Wasserman.[80]

3.2 Kohn-Sham Density Functional Theory

As appealing as it is, DFT carries its main problem in its name: The density functional. Although its existence has been put on a rigorous basis thanks to the work by Levy and Lieb, its functional form remains elusive. This does not come as a surprise: over 30 years before the seminal paper by Hohenberg and Kohn, the Thomas-Fermi-Dirac[81, 82] model of the homogeneous electron gas showed that it is all but trivial to describe the kinetic and correlation energy of some density $\rho(\mathbf{r})$. To date, an overwhelming majority[83] of DFT calculations is carried out within the Kohn-Sham DFT[15] (KS-DFT) framework, since it elegantly circumvents the problem of the description of the kinetic energy functional. We shall explore this approach in the following paragraphs.

3.2.1 The Kohn-Sham Formalism

Although the exact density functional exists, it is not known, and finding accurate approximations to it is quite a formidable task. In 1965, Kohn and Sham introduced a formalism that, in practice, considerably simplifies the determination of the functional $F[\rho]$. According to Hohenberg and Kohn,[14] one may separate the classical Coulomb energy of the interacting electrons J out of eq. 3.1:

$$E_v[\rho] = \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} + J[\rho] + G[\rho]\,, \quad \rho \in \mathscr{A}_N, \tag{3.21}$$

which leaves all the non-classical terms to $G[\rho]$. The key to the Kohn and Sham[15] picture of DFT lies in taking this decomposition one step further. This is possible thanks to the introduction of a simple way of constructing the electronic density and the resulting possibility of computing a large contribution to the kinetic energy.

We start by noting that the kinetic energy T_s of a *noninteracting* single[i] particle system is not only known, but that it also has an appealing, simple form. Such a system is represented by a single-determinantal wavefunction, and its corresponding kinetic energy is:

$$T_s = \sum_i^N \int \mathrm{d}\mathbf{r}\, \psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2\right) \psi_i(\mathbf{r})\,, \quad \psi \to \Psi\,, \quad \Psi \in \mathscr{W}_0, \tag{3.22}$$

where the single-determinantal wavefunction is constructed from single-particle orbitals $\{\psi_i\}$:

$$\mathscr{W}_0 = \{\Psi \mid \Psi \text{ is a single determinant}; \psi \to \Psi,\, \langle\psi_i|\psi_j\rangle = \delta_{ij}\}\,. \tag{3.23}$$

---

[i] Whence the s in T_s

We also note that one simple way of resolving the N-representability issue of ρ would lie in the construction of $\rho(\mathbf{r})$ from the N occupied, orthonormal orbitals that constitute the aforementioned determinant:

$$\rho(\mathbf{r}) = \sum_i^N \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \,, \quad \psi \to \Psi \,, \quad \Psi \in \mathscr{W}_0. \tag{3.24}$$

Any density constructed in such a way will be member of a set $\mathscr{A}_N^0$ of densities that come from a single-determinantal *noninteracting* ground state:

$$\mathscr{A}_N^0 = \left\{ \rho \;\middle|\; \sum_i |\psi_i|^2 = \rho \,,\ \psi \to \Psi \,,\ \Psi \in \mathscr{W}_0 \right\}, \tag{3.25}$$

and it will be N-representable, since $\mathscr{A}_N^0 \subset \mathscr{I}_N$ by construction.[77]

Kohn and Sham exploited those key concepts by inserting a density constructed from $\Psi \in \mathscr{W}_0$ into the Hohenberg-Kohn density functional. Since the ground-state energy is a functional of the density and the density alone, one may just as well construct the density from a set of auxiliary single-particle orbitals. Not only will this ensure N-representability, but it also makes the computation of the noninteracting part of T possible. This constitutes a rather natural approach since the noninteracting kinetic energy makes up the vast majority of the kinetic energy of a system, $T - T_s \ll T$. The energy difference should be easier to account for (and approximate) than the total kinetic energy.[15,17]

Kohn and Sham started from the splitting adopted in eq. 3.21: Given that T_s is known for any density in $\mathscr{A}_N^0$, one may take the given decomposition one step further and define

$$G[\rho] \equiv T_s[\rho] + E_{xc}[\rho]. \tag{3.26}$$

The remaining, unknown term - E_{xc} - is due to all non-classical *exchange-correlation* effects. The exchange-correlation energy is thus *defined* by the difference between the noninteracting kinetic energy T_s and the (true) interacting kinetic energy T of the system, as well as the difference between classical (J) and (true) quantum mechanical interaction (potential) energy (V_{ee}) of the electrons:

$$E_{xc}[\rho] \equiv \underbrace{T[\rho] - T_s[\rho]}_{\text{kinetic terms}} + \underbrace{V_{ee}[\rho] - J[\rho]}_{\text{potential terms}} \,, \quad \rho \in \mathscr{A}_N^0. \tag{3.27}$$

Since the exchange energy is orders of magnitudes larger than the correlation energy, it is customary and certainly more practical to write $E_{xc}[\rho] = E_x[\rho] + E_c[\rho]$; we shall make reference to this partitioning further on in the text.

By expressing the total energy in terms of the by now well-known functionals and the auxiliary orbitals $\{\psi_i\}$ that add up to ρ, Kohn and Sham found that:

$$\begin{aligned} E_v[\rho] = &\sum_i^N \int \psi_i^* (\mathbf{r}) \left(-\frac{1}{2}\nabla^2\right) \psi_i (\mathbf{r})\, d\mathbf{r} + J[\rho] + E_{xc}[\rho] \\ &+ \int v(\mathbf{r})\rho(\mathbf{r})\, d\mathbf{r}\,, \quad \sum_i |\psi_i|^2 = \rho,\ \psi \to \Psi,\ \Psi \in \mathscr{W}_0,\ v \in \mathscr{V}_N. \end{aligned} \tag{3.28}$$

The auxiliary single-particle orbitals are commonly referred to as the *Kohn-Sham orbitals. A priori*, their sole purpose lies in summing up to the correct density and in the straightforward computation of the kinetic energy T_s; the are intended to have no other physical interpretation.[15, 17, 84, 85] The key concept is that both the interacting system in eq. 3.28 and the noninteracting system that the $\{\psi_i\}$ belong to share the same N-electron density. The only quantities not shared by both systems are cast away in the functional $E_{xc}[\rho]$. The beauty of this set of equations lies in the mapping of an interacting, physical problem, onto an artificial, but easy-to-solve auxiliary noninteracting problem, where the only common element is the identical N-electron density. We note that T_s, although expressed in terms of auxiliary orbitals, is still an implicit functional of the density: The overall expression in eq. 3.28 remains a true density functional, since it can be shown that the Slater determinant Ψ_{KS} constructed from the Kohn-Sham orbitals fulfils

$$\Psi_{\mathrm{KS}}[\rho] = \underset{\Psi}{\operatorname{argmin}} \left\{ T_s[\rho] \,|\, \Psi \to \rho \right\}. \tag{3.29}$$

But not only does the Kohn-Sham approach make it possible to easily construct N-electron densities, but the result of inserting eq. 3.28 into the Euler-Lagrange equation for a density variation is most appealing, too.[ii] By subjecting eq. 3.28 to such a constrained minimisation, using the orthonormality of the auxiliary orbitals as boundary conditions, Kohn and Sham derived a set of *single particle equations* for the noninteracting ψ that yield the ground-state density:

[ii] The question whether the density variation in the Euler-Lagrange procedure can be properly carried out, *i.e.* whether the variation of the density and the functionals in eq. 3.32 is possible for every $\rho \in \mathscr{A}_N^0$ is a rather intricate one, and we refer the interested reader to the available, enlightening literature.[77]

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}), \tag{3.30}$$

where v_{eff} is an *effective potential* given by:

$$v_{\text{eff}} = v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta\rho\left(\mathbf{r}\right)} + \frac{\delta E_{xc}[\rho]}{\delta\rho\left(\mathbf{r}\right)}, \tag{3.31}$$

or, alternatively,

$$v_{\text{eff}} = v(\mathbf{r}) + \int \frac{\rho\left(\mathbf{r}'\right)}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}\left(\mathbf{r}\right), \tag{3.32}$$

where he have *defined*

$$v_{xc} \equiv \frac{\delta E_{xc}[\rho]}{\delta\rho\left(\mathbf{r}\right)} \tag{3.33}$$

as the *exchange-correlation potential*, which is the functional derivative of $E_{xc}[\rho]$ and the one and only part of the potential due to non-classical effects. This potential is local in space![iii]

The resulting expressions are neat, clean and (deceptively) simple: Given an expression for the exchange-correlation functional $E_{xc}[\rho]$, the solution to eq. 3.30, and therefore the solution of the HK variational principle, is obtained from a straightforward self-consistent field (SCF) approach.

The question is whether the word 'straightforward' is fully justified, since we have not discussed the search space $\mathscr{A}_N^0$ in any detail (and neither have we been concerned with any a practical functional form for $E_{xc}[\rho]$). $\mathscr{A}_N^0$ is clearly a subset of $\mathscr{A}_N$, but we do not know much about it, except how to construct its members. Kohn-Sham DFT carries all the burden of Hohenberg-Kohn DFT. This is no encouraging perspective. Moreover, in addition to the basic problems associated to the Hohenberg-Kohn theorems that we have discussed in the previous chapter, the Kohn-Sham formalism gives rise to an additional condition that has to be accounted for, which is noninteracting-v-representability[88] of the density. Much alike the v-representability issue,[75,76] this requirement is only met if a *noninteracting* ground-state density exists that is represented by the effective potential v_{eff}. In their discussion of v-representability, Levy[74] and Lieb[75] demonstrated that a *'perfectly smooth*

[iii] Burke referred to this as a semi-classical description of exchange and correlation effects.[86,87]

and innocent'[75] density constructed from degenerate ensembles is only *ensemble*-v-representable, rather than pure-state-v-representable. This directly implies that not every such ensemble-v-representable density can possibly be associated to a *single* noninteracting Kohn-Sham Slater determinant[77] ...

Is there any certainty that the Kohn-Sham equation yields some reasonable result? Luckily, this is the case, since if it were not, years of results based on KS-DFT calculations would be of more than questionable quality. Eschrig[77] pointed out that if a self-consistent solution is obtained from eq. 3.30, then the density obtained is forcibly noninteracting-$v-$representable. While mathematically ill defined, the questions evoked in the preceding paragraphs are of little practical importance: There may be systems for which no solution can be found, but where it *can* be found, we can be positive that $v \in \mathscr{V}_N$ and $\rho \in \mathscr{A}_N$, and since the density is in practice not decomposed into the $\{\psi_i\}$, but instead constructed from them, even $\rho \in \mathscr{A}_N^0$. We are therefore ready to tackle the next unknown: $E_{xc}[\rho]$.

On the Interpretation of Kohn-Sham Orbitals Kohn and Sham explicitly stressed the auxiliary nature of the noninteracting single particle orbitals.[15] However, Janak[89] later showed that the derivative of the energy of an N-electron system with respect to the occupation number n_i of a given orbital is equal to the orbital eigenvalue, $\partial E_N/\partial n_i = \epsilon_i$. A relation between an N and $N+1$ electron system can then be established by inserting $0 \leq n \leq 1$ (fractional) electrons in the lowest-lying unoccupied KS-orbital i of the N-electron system:

$$E_{N+1} - E_N = \int_0^1 \mathrm{d}n\, \epsilon_i(n). \tag{3.34}$$

Practical evidence also suggests that Kohn-Sham orbitals lend themselves to some straightforward physical interpretation within a one-electron picture.[17,84] Kohn-Sham gaps have been found to provide a good approximation of the optical gap in molecular systems, and the eigenvalues of the highest occupied KS orbitals agree with experimental ionisation potentials to about ≈ 0.1 eV.[85]

Yet Another Note on the Time-Dependent Case The Kohn-Sham scheme can also be applied to TDDFT, and Runge and Gross have developed the formalism in their seminal paper.[79] Such KS-TDDFT calculations will require the computation of the exchange-correlation kernel, $f_{xc}(\mathbf{r}t, \mathbf{r}'t')$, which describes the response of the exchange-correlation potential at $\mathbf{r}$ at time t to a density variation at $\mathbf{r}'$ at some other time t'. This expression is inserted into the Dyson equation of TDDFT, from which the excitation energies of the system can then be obtained. The search for the xc kernel is even more intricate than the search for the xc functional, and it is often approximated by using the adiabatic local density approximation (ALDA), where $f_{xc}^{\text{ALDA}}(\mathbf{r}t, \mathbf{r}'t') = \delta v_{xc}(\mathbf{r}, t)/\delta\rho(\mathbf{r}', t')\delta(\mathbf{r} - \mathbf{r}')\delta(t - t')$. Since the ALDA is the most common approximation, most functionals developed for ground-state KS-DFT can be (and are indeed) routinely applied in KS-TDDFT calculations. The considerations in the following pages will therefore not only be of importance for KS-DFT, but for KS-TDDFT alike (with the additional constraint that f_{xc}^{ALDA} be a sufficient approximation to f_{xc}).

3.2.2 The Exact Form of the Exchange-Correlation Functional

It is possible to reconstruct the exchange-correlation functional for every system.[90] However, this approach does not lead to any portable expression that would be general enough to be applied in routine calculations. Still, it has relevant implications in the development of new xc approximations and in the physical analysis of existing xc approximations.[91] We shall therefore quickly sketch how the exact KS xc functional is derived from Lieb's universal DF in the framework of the *adiabatic connection*, a formalism which has accompanied DFT development for the last few decades.[17,90–95]

Lieb proved that a minimising density matrix γ^ρ associated to $\hat{\text{H}}[0]$, the Hamiltonian in absence of an external potential, exists (turning the inf into a proper min)[75]:

$$F_\lambda[\rho] = \min_{\hat{\gamma}\to\rho} \left\{ \text{Tr}\, \hat{\text{H}}_\lambda[0]\hat{\gamma} \right\} = \text{Tr}\, \hat{\text{H}}_\lambda[0]\hat{\gamma}_\lambda^\rho, \tag{3.35}$$

where we have introduced λ as the *adiabatic connection* interaction parameter that 'switches on' the electron-electron interaction in a general Hamiltonian $\hat{\text{H}}[v]$:

$$\hat{\text{H}}_\lambda[v] = \hat{\text{T}} + \lambda\hat{\text{W}} + v(\mathbf{r}). \tag{3.36}$$

One can rewrite eq. 3.35 in terms of an integration[95]:

$$F_\lambda[\rho] = F_0[\rho] + \int_0^\lambda \mathrm{d}\lambda\, F'_\lambda[\rho]. \tag{3.37}$$

In the absence of an external potential, we simply have $T_s[\rho] = \text{Tr}(\hat{\text{H}}_0[0]\hat{\gamma}_0^\rho) = \text{Tr}(\hat{\text{T}}\hat{\gamma}_0^\rho)$ at $\lambda = 0$. We then recognise that, correspondingly, the kinetic contribution to $F_0[\rho]$ is simply $T_s[\rho]$. The integrand in 3.37 can be computed by calculating $F'_\lambda[\rho] = \mathrm{d}F_\lambda[\rho]/\mathrm{d}\lambda$ using the Hellman-Feynmann theorem, which leads us to:

$$F_\lambda[\rho] = T_s[\rho] + \int_0^\lambda \mathrm{d}\lambda\, \mathcal{W}_{J_{xc},\lambda}[\rho], \tag{3.38}$$

where we have introduced $\mathcal{W}_{J_{xc},\lambda}[\rho] = \mathrm{Tr}(\hat{\mathrm{W}}\hat{\gamma}^\rho_\lambda)$ in the integrand as the energy due to the Coulomb operator in a partially interacting system with a minimising density matrix $\hat{\gamma}^\rho_\lambda$ (thus the absence of a min). Note that the λ in front of $\hat{\mathrm{W}}$ has vanished due to the derivative $\partial\hat{\mathrm{H}}_\lambda[v]/\partial\lambda$, the λ dependency is only retained in the minimising ensemble-density matrix $\hat{\gamma}^\rho_\lambda$.[95] Decomposing this interaction energy as in the Kohn-Sham scheme, we have that:

$$\mathcal{W}_{J_{xc},\lambda}[\rho] = \lambda J[\rho] + \lambda E_x[\rho] + E_{c,\lambda}[\rho], \tag{3.39}$$

which we may integrate term-by-term within the interval $[0, 1]$. We then find for the terms that:

$$J[\rho] = \frac{1}{2}\iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}', \tag{3.40}$$

$$\tilde{E}_x[\rho] = \mathrm{Tr}\left(\hat{\mathrm{W}}\hat{\gamma}^\rho_0\right) - J[\rho], \tag{3.41}$$

$$\tilde{E}_c[\rho] = \int_0^1 \mathrm{d}\lambda\ \mathrm{Tr}\left(\hat{\mathrm{W}}\left(\hat{\gamma}^\rho_\lambda - \hat{\gamma}^\rho_0\right)\right). \tag{3.42}$$

We note that $\tilde{E}_x$ is a functional of the noninteracting density matrix $\hat{\gamma}^\rho_0$, and that $\tilde{E}_x$ is therefore a constant for all of λ: It is the exchange energy associated to a single, noninteracting Slater determinant[iv]. We note that this functional is explicitly known, and its form is very familiar, but we will not write it out explicitly just yet. The correlation energy, $\tilde{E}_c[\rho]$, on the other hand, is associated to the orbital relaxation upon switching on the interaction parameter λ, and is therefore given by the action of the Coulomb operator on the difference between the minimising density matrix at $\lambda = 1$ and the noninteracting minimising $\hat{\gamma}_0^\rho$. For the total exchange-correlation energy, one thus obtains the overall (but a tad less insightful) expression

---

[iv] Within the framework of the adiabatic connection, this actually gives rise to a formal *correlation* contribution for an energy computed from the Hartree-Fock method, even though any correlation energy is absent from Hartree-Fock theory by the very *wavefunction* definition of it[90, 91, 95]! The correlation energy in KS-DFT is therefore not identical to the correlation energy within wavefunction theory. The attribution of DFT-correlation to the Hartree-Fock method comes from the interpretation of the Hartree-Fock orbitals as the orbitals of an *interacting*, single Slater determinant, whereas the Kohn-Sham scheme imposes a purely determinantal exchange for a system of noninteracting particles. The difference between the energy of an interacting Hartree-Fock system and its noninteracting counterpart is minute, though.[91]

$$\tilde{E}_{xc}[\rho] = \int_0^1 \mathrm{d}\lambda \; \mathrm{Tr}\left(\hat{\mathrm{W}}\hat{\gamma}_\lambda\right) - \lambda J[\rho], \tag{3.43}$$

known as the *adiabatic connection* formulation of the xc functional. The stealthy appearance of a tilde in $\tilde{E}_{xc}$ is due to the Levy-Lieb constrained DF in eq. 3.35; only if a density under consideration is noninteracting pure-state v-representable - only if it belongs to $\mathscr{A}_N^0$ - is the AC xc functional $\tilde{E}_{xc}$ equal to the Kohn-Sham xc functional E_{xc}.

Admittedly, this is all rather abstract. The ‘unknown’ xc functional is exactly defined in terms of ensemble density matrices (!) - with all of the associated computational burden. And while this formulation of $E_{xc}[\rho]$ is more than just insightful for theoretical discussions of xc approximations, its use in the elucidation of chemical problems is as unfeasible as the computation of the complete $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$.

We still lack a reasonably accurate (and possibly elegant) *model* for $E_{xc}[\rho]$. We are set for a detour.

3.2.3 The Exchange-Correlation Hole

> *Ein Loch ist da, wo etwas nicht ist.*
> Kurt Tucholsky, ‘Zur soziologischen Psychologie der Löcher’[96]

If electrons were to behave purely classically,[17] the probability distribution function $\rho_2(\mathbf{r}, \mathbf{r}')$ of finding one electron at $\mathbf{r}$ and another one at $\mathbf{r}'$ would read:

$$\rho_2(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r})\rho(\mathbf{r}'), \tag{3.44}$$

and all of our previous considerations would become obsolete: Since the Coulomb operator $|\mathbf{r} - \mathbf{r}'|^{-1}$ is merely a two-body operator, the knowledge of ρ_2 would allow us to evaluate the complete electron-electron repulsion term[97,98]:

$$V_{ee} = \frac{1}{2} \iint \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}' \, \frac{\rho_2(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{3.45}$$

Arguably, electrons do not behave classically. Every electron creates around itself a hole in the probability distribution function, meaning that it is less likely that an electron be found at $\mathbf{r}$ if another one already sits around the same position. This hole is commonly referred to as the *exchange-correlation hole*.[17,93,99,100] So, if we wish to make proper use of eq. 3.45, we shall first have to be able to quantify the xc hole, which implies finding the corresponding expression for ρ_2 in a system of fermions.[98,101] This quest will make us encounter some formulations of density matrix functional theory.

Given the infamous $\hat{\gamma}$, we may carry out a straightforward integration over all of the $N-2$ coordinates that we do not need, yielding the *two-matrix* or the two-body reduced density matrix $\gamma_2(\mathbf{r}, \mathbf{r}')$[v] :

$$\gamma_2(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_1', \mathbf{r}_2') = N(N-1)\Bigg(\int \cdots \int \mathrm{d}\mathbf{r}_3 \ldots \mathrm{d}\mathbf{r}_N \int \cdots \int \mathrm{d}\mathbf{r}_3' \ldots \mathrm{d}\mathbf{r}_N' \\ \gamma(\mathbf{r}_1', \ldots, \mathbf{r}_N'; \mathbf{r}_1, \ldots, \mathbf{r}_N)\Bigg). \tag{3.46}$$

We are particularly interested in its diagonal elements, for which we shall adopt a short-hand notation:

$$\gamma_2(\mathbf{r}, \mathbf{r}') = \gamma_2(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_1, \mathbf{r}_2), \tag{3.47}$$

where the prime is now simply associated to the *second* coordinate in $\mathbf{r}$, rather than an off-diagonal element. This is the quantity we have been looking for, and $\rho_2(\mathbf{r}, \mathbf{r}') = \gamma_2(\mathbf{r}, \mathbf{r}')$, which we could now plug into eq. 3.45 - but we choose not to do it just yet.

Instead, let us focus on what is *not* there in a system of quantum fermions. From what we know, we may quantify the absence of things by defining the exchange-correlation hole $h_{xc}(\mathbf{r}, \mathbf{r}')$[17,93,102]:

$$h_{xc}(\mathbf{r}, \mathbf{r}')\rho(\mathbf{r}) = \gamma_2(\mathbf{r}, \mathbf{r}') - \rho(\mathbf{r}')\rho(\mathbf{r}), \tag{3.48}$$

which is nothing but the difference of conditional probability of finding an electron at $\mathbf{r}'$ at a fixed coordinate $\mathbf{r}$ between a system of interacting fermions and its classical counterpart.

This definition opens the path to an alternative approach to $\gamma_2(\mathbf{r}, \mathbf{r}')$. By isolating γ_2 in the above equation and inserting it into eq. 3.45, we arrive at an appealing use of h_{xc}:

---

[v] Yet another integration would lead to the one-matrix or one-body reduced density matrix γ_1, where $\mathrm{Tr}\,\gamma_1 = N$ and the diagonal elements of which are interpreted as our familiar $\rho(\mathbf{r})$!

$$V_{ee}[\rho] = \underbrace{\frac{1}{2}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,\frac{\rho(\mathbf{r}')\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}}_{J[\rho]} + \underbrace{\frac{1}{2}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,\frac{h_{xc}(\mathbf{r},\mathbf{r}')\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}}_{\tilde{E}_{xc}[\rho]}, \tag{3.49}$$

which can be used instead of γ_2 to calculate the expectation value of the two-body operator $\hat{\mathrm{V}}_{ee}$. We have found yet another definition of the exact xc functional; and although we would yet again have to make reference to an ensemble density matrix to find the exact $h_{xc}(\mathbf{r},\mathbf{r}')$, eq. 3.49 contains some valuable information.

First, let us consider what happens to eq. 3.49 if it is used as an explicit form for $\mathrm{Tr}\left(\hat{\mathrm{W}}\hat{\gamma}_\lambda\right)$ in eq. 3.43: The term due to the classical Coulomb energy $J[\rho]$ disappears, and we are left with:

$$\tilde{E}_{xc}[\rho] = \int_0^1 \mathrm{d}\lambda \iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,\frac{h^\lambda_{xc}(\mathbf{r},\mathbf{r}')\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}, \tag{3.50}$$

where h^λ_{xc}, through its dependency on the minimising ensemble density matrix, is still dependent on λ. Since ρ is fixed for all values of λ, we may just as well change the order of integration:

$$\tilde{E}_{xc}[\rho] = \iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,\frac{\rho(\mathbf{r})\int_0^1 \mathrm{d}\lambda\, h^\lambda_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{3.51}$$

and we find that $h_{xc}(\mathbf{r},\mathbf{r}') = \int_0^1 \mathrm{d}\lambda\, h^\lambda_{xc}(\mathbf{r},\mathbf{r}')$: The exchange-correlation hole introduced in eq. 3.48 and used in eq. 3.49 is actually a λ-*averaged xc hole*.[93,102]

We have noted before that the particularity of the $\lambda = 0$ limit is that it uniquely defines the exchange energy, which is simply the Coulomb energy due to a single Slater determinant, and we know its familiar form from Hartree-Fock theory.[4,90] It is thanks to this limit that we are now able to separate the exchange hole, commonly referred to as the *Fermi hole*, from the correlation or *Coulomb hole*[17,93,102]:

$$\begin{aligned} h_c(\mathbf{r},\mathbf{r}') &= h_{xc}(\mathbf{r},\mathbf{r}') - h_x(\mathbf{r},\mathbf{r}') && (3.52)\\ &= \int_0^1 \mathrm{d}\lambda\, h^\lambda_{xc}(\mathbf{r},\mathbf{r}') - h^0_{xc}(\mathbf{r},\mathbf{r}') && (3.53) \end{aligned}$$

It is rather straightforward to express the Hartree-Fock quantities in terms of density matrices, and after rearranging some terms, we find an analytical expression for the exchange hole in terms of single-particle orbitals[17]:

$$h_x(\mathbf{r},\mathbf{r}') = -\frac{1}{\rho(\mathbf{r})}\left(\sum_i |\psi_i(\mathbf{r})\psi_i(\mathbf{r}')|\right)^2 , \tag{3.54}$$

where $\{\psi_i\}$ may be a set of Hartree-Fock orbitals just as well as a set of Kohn-Sham orbitals[vi].

This hole has useful properties.[93,94,99,102–104] From eq. 3.54, it follows that the *sum rule* for the exchange hole reads:

$$\int \mathrm{d}\mathbf{r}'\, h_x(\mathbf{r},\mathbf{r}') = -1, \tag{3.55}$$

and

$$h_x(\mathbf{r},\mathbf{r}') \leq 0, \tag{3.56}$$

since the hole can remove at most one electron, for otherwise, by virtue of eq. 3.48 it would create some (anti)matter. Both equations have an important consequence for the correlation hole.

It is easy to find the corresponding sum rule for the λ-averaged h_{xc}: Since $\mathrm{Tr}\,\rho(\mathbf{r},\mathbf{r}') = N$ and $\mathrm{Tr}\,\gamma_2(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}'_1,\mathbf{r}'_2) = N(N-1)$, we must have from the very definition of h_{xc} in eq. 3.48 that

$$\int \mathrm{d}\mathbf{r}'\, h_{xc}(\mathbf{r},\mathbf{r}') = -1, \tag{3.57}$$

i.e. the exchange-correlation hole reflects that if an electron is observed at $\mathbf{r}$, only $N-1$ electrons can be observed elsewhere in space (*i.e.* at all $\mathbf{r}'$).[100,102] This, together with the sum rule for h_x, implies that

$$\int \mathrm{d}\mathbf{r}'\, h_c(\mathbf{r},\mathbf{r}') = 0. \tag{3.58}$$

These three sum rules impose important limits especially on the exchange hole[100,102]: The higher the value of the diagonal term $h_x(\mathbf{r},\mathbf{r})$ (known as the *on-top value*[105]), the more rapidly decaying the hole must be if it should integrate to -1 (since it is a negative definite kernel). We even know the overall behaviour of the on-top exchange hole, which for the spin-restricted case is:

[vi] We note, however, that the definition of the correlation energy will change accordingly.

$$h_x(\mathbf{r},\mathbf{r}) = -\frac{\rho(\mathbf{r})}{2}. \tag{3.59}$$

The correlation hole, unfortunately, is more evasive; much like its counterpart, the correlation energy, it just contains what one could blatantly describe as what we are unable to grasp with a more straightforward approach or - the whole rest of the array of quantum effects. This is reflected in its integral being naught.

These five conditions (eqs 3.55, 3.56, 3.57, 3.58 and 3.59) constitute a vital basis for both the development and assessment of new approximations to $E_{xc}[\rho]$. This task is substantially facilitated by the fact that in the computation of V_{ee}, it is sufficient to consider a *spherically averaged* xc hole[100]

$$\bar{h}_{xc}(\mathbf{r},u) = \int \frac{d\Omega_u}{4\pi} h_{xc}(\mathbf{r},\mathbf{r}+u), \tag{3.60}$$

which greatly simplifies the construction of models for $\bar{h}_{xc}$. Together with the known behaviour of $E_{xc}[\rho]$ under certain scaling conditions and in some particular limits, the spherically averaged xc hole may then serve as a guideline to what conditions a reasonable xc approximation should obey.[31,93,94,103–106] The importance of those conditions is, however, a matter of discussion.

It is in the following lines that we shall - finally - sketch the history of exchange-correlation approximations. Starting from the basics outlined in the seminal paper by Kohn and Sham, we shall encounter both pitfalls and small revolutions until we will, finally, cross paths again with the adiabatic connection for a last time.

3.2.4 Approximations to the Exchange-Correlation Functional

> *With four parameters I can fit an elephant,*
> *and with five I can make him wiggle his trunk.*
> John von Neumann, as attributed by Enrico Fermi.[107]

Over the years, a plethora of different approximations to E_{xc} have been developed.[83,108] Some of them share the same functional form, but differ in the constraints to which the parameters were fit; others are fitted to a similar set of physical or chemical constraints, but have a different functional form, and even others are a combination (a collection, so to say) of pre-existing functional forms, the panoply of parameters of which were parametrised based on an horrendously large database. It is a matter of vivid dispute which option is the most desirable, the most appropriate, the most justifiable, or simply the prettiest[vii].

In principle, the h_{xc} for every $E_x[\rho]$ and $E_c[\rho]$ should obey eqs 3.55 to 3.59, and the functionals should adopt a certain behaviour when fed with a density uniformly scaled by some constant γ[109–111]:

$$\rho_\gamma(\mathbf{r}) = \gamma^3 \rho(\gamma \mathbf{r}). \tag{3.61}$$

Specifically,

$$E_x[\rho_\gamma] = \gamma E_x[\rho], \tag{3.62}$$

$$E_c[\rho_\gamma] > \gamma E_c[\rho]. \tag{3.63}$$

The exchange energy is therefore homogeneous to degree one under a uniform scaling of the density. Many more general scaling relations exist, such that it is for instance possible to demonstrate that the noninteracting kinetic energy is homogeneous to degree two.[112,113]

Among the many limits, conditions and bounds the exact functional must fulfil, we shall not fail to mention one particularly important property. In principle, the overall E_{xc} is limited by the Lieb-Oxford bound[17]:

$$0 \geq E_{xc}[\rho] \geq -B \int \mathrm{d}\mathbf{r} \rho^{4/3}(\mathbf{r}), \tag{3.64}$$

where $1.67 < B < 2.273$ (and, following rigorous derivation, the optimum B *might* be either 2.273 or 2.215).[114]

In principle, that is. Not every functional necessarily obeys every single constraint we have evoked so far, and the question of which conditions have to be obeyed in order for a functional to be 'reasonable' are another source of never-ending ideological collisions.[115] Nevertheless, among the most common functionals used to this date, we find examples of every flavour. We shall start with the oldest of the xc approximations, and we will see that, despite (or rather for) its simplicity, it fulfils a surprisingly large amount of formal limits.

3.2.4$_\text{A}$ Local Density Approximation (LDA)

The concept of a locally approximated exchange energy makes us take a step back into the 1930ies: The Thomas-Fermi[81] model of the electron gas preceded DFT by about 40 years, and Dirac[82] proposed a functional form for its exchange energy soon after. Its exact form, $E_x = C_x \int \mathrm{d}\mathbf{r}\, \rho^{4/3}$, with C_x initially found to be $C_x = -3/(2\pi)(3\pi^2)^{1/3}$, was introduced by Slater in 1951 in his paper on *'a simplification of the Hartree-Fock model'*.[116] In Slater and Johnson's subsequent $X\alpha$ model,[117] the exchange energy is determined based on a locally averaged h_x, which fixes the prefactor at $C_x = -9\alpha(3/(4\pi))^{1/3}$, with a system-dependent parameter α.[108]

[vii] It is also a matter of dispute which one is the ugliest.

In order to establish an approximation for the unknown functional, Hohenberg and Kohn referred to the exchange energy of the uniform electron gas (UEG) even before the introduction of the xc functional by Kohn and Sham.[14, 15] The latter introduced an innocuous notation. In a general form, Kohn and Sham proposed to write:

$$E_x^{LDA} = \int \mathrm{d}\mathbf{r}\, \rho(\mathbf{r}) \epsilon_x^{\mathrm{UEG}}(\rho), \tag{3.65}$$

which is known as the *Local (Spin) Density Approximation* or L(S)DA to the Kohn and Sham xc functional. Kohn and Sham initially made rather pessimistic prognostics on its applicability in inhomogeneous systems such as molecules. However, calculations using the $X\alpha$ method soon revealed that a value of $\alpha = 2/3$[117] consistently yielded reasonable exchange energies for a variety of system - this was the α-value for a uniform electron gas.[108] By adopting this very value in the LDA, one finds that:

$$C_x^{\mathrm{UEG}} = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}, \tag{3.66}$$

$$\epsilon_x^{\mathrm{UEG}}(\rho) = -C_x^{\mathrm{UEG}} \rho^{1/3}, \tag{3.67}$$

$$E_x^{\mathrm{UEG}}[\rho] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int \rho^{4/3} \mathrm{d}\mathbf{r}, \tag{3.68}$$

which is used in solid state calculations until today, and it remains a valuable ingredient in almost all of the more elaborate approximations to $E_{xc}[\rho]$.

The assumption thus made is that the density of a non-homogeneous system can be *locally* described in terms of the exchange-hole due to a uniform electron gas. An analytical representation for the correlation functional is substantially more complicated to find, since its exact form is unknown and approximations have to be derived with the known limits for the interacting electron gas in mind.[15] This is especially cumbersome for the spin-dependent case: It usually involves different parametrisations for parallel and antiparallel correlation. Early examples include the work by Vosko, Wilk and Nusair (VWN)[118] or Perdew and Zunger (PZ).[119]The former parametrisation is based on data from the random-phase approximation (RPA),[118] whereas the latter is derived from exact Quantum Monte Carlo results.[119] Both VWN and PZ correlation have in principle been superseded by the Perdew-Wang functional form of 1992 (PW92),[120] which is simpler than its predecessors and remedies some minor fallacies associated to the latter. Nevertheless, both VWN and PZ are still in use today. The equations for $\epsilon_c^{\mathrm{UEG}}$ obtained from the aforementioned models are omitted here: Even if primordial to the development of modern day DFT, they are far less appealing than the sleek LDA for the exchange energy.

The LDA works surprisingly well both for molecules and in particular for solid state systems. What comes as a surprise bearing the original comments by Kohn and Sham (and many others) in mind becomes all the more expected as the approximation is compared to the conditions introduced on the previous pages. Not only does the LDA exchange functional obey all of the restrictions on the exchange hole, including its on-top derivative, but it also obeys the Lieb-Oxford bound and all of the scaling relations.[105, 114] Burke pointed out that an exact on-top derivative is almost impossible to obtain even in many wavefunction based methods.[100]

The LDA is still popular in solid state physics, but its use in chemistry has become relatively obsolete, due to a general overbinding tendency,[121] and since the computational overhead associated with more accurate, but slightly more complex approximations is negligible when using today's computers: The failures of LDA are especially detrimental in the prediction of hydrogen bonds and weakly bound complexes (where LDA tends to overbind ludicrously), and better functionals are available for most chemical problems.[83]

3.2.4$_{\mathrm{B}}$ Generalised Gradient Approximation (GGA)

Due to its very nature, the LDA performs reasonably well for systems with a slowly varying density: The closer the approximation of local homogeneity is to the physical truth[viii], the more accurate the LDA. But once densities exhibit considerable fluctuations, the LDA will break down to a - more or less - drastic extent.[86, 122, 123] The first proposition by Hohenberg, Kohn and Sham consisted in using the gradient expansion, *i.e.* a simple Taylor expansion of the density, the first term of which would be the LDA. What seemed so straightforward was of nightmarish practicability: the expansion with respect to the derivatives of the density diverges,[124] violating all (!)

[viii] For non-relativistic systems and only in so far as there is any physical 'truth'.

of the conditions on the exchange-correlation hole. Instead, a *generalised* gradient expansion can be derived, resulting in the *Generalised Gradient Approximation* (GGA)[125] to the xc functional. One of the first GGA regularly used in production runs was introduced by Perdew in 1986,[126] based on a rather cumbersome approach: Following a procedure originally proposed by Langreth[125] some years earlier, Perdew advocated the truncation of the divergent parts of the GEA,[127] so that it would obey the constraints on the xc hole. This was done numerically on a real space grid.[127] Perdew then 'simply' fit an analytical form to the numerical result.[126] Becke's approach to the problem was much simpler; he based his famous B88[122] ('Becke 1988') exchange functional on the GGA introduced by Langreth and Mehl, the free parameters of which he fit to the atomic atomisation energies of the noble gases He to Rn.

The general form of a GGA exchange functional is a simple generalisation of the LDA, and as such, it can be considered a leading correction to the locally homogeneous approximation:

$$E_x^{\mathrm{GGA}}[\rho] = \int H(\rho, \nabla\rho)\rho(\mathbf{r})\mathrm{d}\mathbf{r} = -\sum_\sigma \int K_\sigma(\rho, \nabla\rho)\rho_\sigma^{4/3}(\mathbf{r})\mathrm{d}\mathbf{r}. \tag{3.69}$$

$K(\rho, \nabla\rho)$ denotes the most common form of the spin-dependent exchange enhancement factor, and $H(\rho, \nabla\rho)$ is its spin-restricted analogue. The notational difference that comes with both forms is rather subtle; $H(\rho, \nabla\rho)$ is notationally closer to the proposition by Kohn and Sham, since it corresponds to an enhanced $\epsilon_x^{\mathrm{UEG}}(\rho)$. H and K are more or less easily converted from one to another thanks to the spin-scaling relation for the exchange functional:

$$E_x[\rho_\alpha, \rho_\beta] = \frac{1}{2}(E_x[2\rho_\alpha] + E_x[2\rho_\beta]), \tag{3.70}$$

from which conversion factors are easily derived. Usually, both forms and parametrisations for H or K are chosen such that in the limit of the uniform electron gas, eq. 3.68 is recovered.

It was the development of reliable GGA that constituted a real milestone in the ascent of KS-DFT[28, 87, 121]: Reasonable results could now be obtained for a variety of molecular systems. Structural predictions based on GGA are usually very reliable.[28] However, their performance for energies, and thus thermochemical properties, can be rather spurious at best (if not to say: mediocre, even though they already substantially improve over the LDA).[31] Reaction enthalpies can often not be reliably predicted, and barrier heights are almost consistently underestimated, sometimes to a drastic extent.[128]

However, there exist cases where GGA predictions seem to profit from some error compensation, making them surprisingly accurate.

3.2.4$_\mathrm{C}$ Kinetic Energy Density: Meta-GGA

In the non-trivial pursuit of accuracy, the logical sequel to the first order correction - the GGA - is a second-order leading correction - the *meta-GGA*.[127, 129–131] After the gradient comes the Laplacian, $\nabla^2\rho$. However, there exists a quantity related to the Laplacian of the density which is even more promising, the *kinetic energy density* τ,[129, 131]

$$\tau_\sigma(\mathbf{r}) = \sum_i \psi_i^{\sigma *}(\mathbf{r})\nabla^2\psi_i^\sigma(\mathbf{r}). \tag{3.71}$$

For if we consider a spin-dependent system, $\tau_\sigma(\mathbf{r})$ can serve as a probe of the one-electron character of the spin-density $\rho_\sigma(\mathbf{r})$: It will be equal to the von Weizsäcker kinetic energy density for real orbitals, $\tau_\sigma^W(\mathbf{r}) = |\nabla\rho_\sigma(\mathbf{r})|^2/8\rho_\sigma(\mathbf{r})$ if $\rho_\sigma(\mathbf{r})$ is a one-electron density. This behaviour also immediately lends itself to be used as a boundary condition in the derivation of the functional form. In a meta-GGA, the enhancement factor $H(\rho, \nabla\rho)$ in eq. 3.69 is simply replaced by a more flexible functional form $H(\rho, \nabla\rho, \nabla^2\rho, \tau)$ that may be designed to recover some underlying GGA approximation under certain limits for $\nabla^2\rho$ and τ, thereby ensuring that the limits to the LDA can be recovered *via* the GGA.[130, 132] Some meta-functionals completely omit the Laplacian term, $\nabla^2\rho$. Meta-functionals are more sensitive to long-range effects: Thanks to the presence of the higher-order derivatives, they seem to better probe the chemical environment. Due to their sensing the von Weizsäcker condition, they may substantially reduce the many-electron self-interaction error that plagues open-shell systems treated with LSDA and GGA.[129, 133, 134]

Unfortunately, the general, average performance of exchange-correlation functionals does not significantly improve upon inclusion of the Laplacian or the kinetic energy density[121][ix], and among the most successful functionals[83] that contain terms due to τ there are mostly hybrid functionals,[31] which combine the approximations described above with the exchange energy of a single Slater determinant, the theoretical base of which we shall now discuss at last.

[ix] Referring to functionals that are purely based on physical constraints.

3.2.4$_\mathrm{D}$ Hybrid Functionals from the Adiabatic Connection

The approximations to the exchange-correlation functional discussed so far all yield an orbital-independent, fully local exchange-correlation potential as the direct functional derivative of the exchange-correlation energy with respect to the density - just as it was proposed and intended by Kohn and Sham. A fundamental paradigm shift was suggested by Axel Becke in 1993[31] in a rather unusual - empirical - approach based on the $\lambda = 0$ limit in the adiabatic connection theorem.[31, 90, 92, 135, 136] Becke advocated the inclusion of only a percentage of orbital-dependent, nonlocal exchange as used in Hartree-Fock theory, along with the corresponding orbital-dependent, nonlocal potential, in KS-DFT: He provided a *generalised* KS-DFT with *hybrid functionals*.

Based on the shapes of h_{xc}, Becke noted that regions of different λ cannot be described equally well by LDA and GGA functionals. On one hand, an LDA xc-hole may be a suitable approximation in correlated systems where $\lambda \rightarrow 1$, but in the $\lambda \rightarrow 0$ limit of uncorrelated systems, it is qualitatively wrong (*e.g.*, the LDA xc hole considerably overestimates left-right correlation, which is absent in exchange-dominated systems).[31, 128] This comes at no surprise, since the LDA reference system is conceptually far from an exchange-only system constructed from a determinant of noninteracting orbitals. On the other hand, as we have seen, the exchange-only limit at $\lambda = 0$ is well known (and so is its xc hole). Based on the known ($\lambda = 0$) and at least somewhat hand-wavingly known ($\lambda = 1$) limits, Becke proposed to model the adiabatic connection integrand by interpolating between the two limits[31] E_x^0 and E_x^1:

$$E_x^0[\rho] = -\frac{1}{2}\sum_i\sum_j \iint \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\, \frac{\psi_i^*(\mathbf{r})\psi_j^*(\mathbf{r}')\psi_j(\mathbf{r})\psi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{3.72}$$

$$E_x^1[\rho] = \int \mathrm{d}\mathbf{r}\, K_x(\rho)\,\rho^{4/3}(\mathbf{r}). \tag{3.73}$$

Linear interpolation now yields the general form for a *hybrid exchange-correlation functional*:

$$E_{xc}^{\mathrm{hyb}}[\rho] = \lambda E_x^0[\rho] + (1-\lambda)\, E_x^1[\rho] + E_c^1[\rho]. \tag{3.74}$$

Implementations in Becke's spirit still use the orbital-dependent exchange potential from Hartree-Fock theory,[31, 121] which is why hybrid functionals are formally part of a *generalised*[137] Kohn-Sham (GKS) theory[x]. A local potential corresponding to the exact DFT exchange functional in eq. 3.72 can be derived, but it is computationally cumbersome to obtain[86,138] and while formally accurate, its practical benefits over the straightforwardly obtained nonlocal potential as used in most hybrid functionals often do not merit the much higher computational cost. Hybrid functionals in Becke's spirit have seen tremendous success,[83] and his idea still forms the basis of most of the xc functionals developed nowadays. With the introduction of hybrid functionals, various thermochemical properties and in particular barrier heights could be more accurately described. Just as importantly, the prediction of many molecular excitations within LR-TDDFT became possible due to an improved accuracy of the Kohn-Sham orbital energies and the inclusion of nonlocal effects: Since the orbital eigenvalues enter the response function, they have a direct influence on the spectra, and the nonlocality of the exact exchange functional can substantially improve the results obtained from frequency-independent kernels.

---

[x] Caveat ahead: Not all of the properties of KS-DFT orbital eigenvalues, scaling relations *etc.* are also valid in a GKS context, *cf.* also the discussion in Chapter 7.

Still, it is in the calculation of spectra that hybrid functionals have performed particularly disappointingly in some systems.[139,140] This problem will be the base of the last xc approximation we discuss here: The group of range-separated hybrids.

3.2.4$_{\mathrm{E}}$ Range-Separated Hybrid Functionals

Failures of an xc approximation can be, were and are still often attributed to an erroneous long-range decay of v_{xc}[xi]. Whereas the exact value is given by the presence of a Coulomb operator and should therefore be r^{-1}, GGA usually exhibit a different long-range dependency. One motivation for the use of hybrid functionals with a large percentage c of exact exchange is that the potential will then decay as cr^{-1},[31] which may improve the accuracy of the result while also allowing for some spatial nonlocality in the frequency-independent xc kernel $f_{xc}^{\mathrm{ALDA}}(\mathbf{r}t, \mathbf{r}'t)$ to be taken into account. However, molecular excitation energies can be particularly sensitive to both decay and nonlocality, and cr^{-1} may still not be accurate enough.[139,140] To this end, some functionals have been devised that combine 100% of exchange exchange with some correlation functional,[31] but due to the absence of typical error cancellation between approximate exchange and correlation functionals, such

---

[xi] We shall not hide that the situation is fare more involved than that: It has been shown that a GGA with a corrected asymptotic behaviour improves the description in some cases, but fails to perform up to initial expectations in many other cases.[139] This can be attributed to a lack of spatial nonlocality in the frequency-independent xc kernel $f(\mathbf{r}t, \mathbf{r}'t)$.

100%-HFX-hybrids are often less accurate than their truly mixed counterparts.

An alternative lies in splitting the Coulomb operator according to[38, 141, 142]:

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1 - f(r)}{r}, \tag{3.75}$$

which is referred to as the range-separation approach. The trick of introducing a resolution of identity in some appropriately chosen function $f(r)$ allows for the two terms to the right hand side of eq. 3.75 to be treated using different approximations.[38, 141, 143] If an appropriate $f(r)$ is selected such that the overall operator decays as r^{-1} for $r \to \infty$, it becomes possible to treat one part of the regime using the exact exchange functional, whereas the other part can be treated by a simple (meta-)GGA. This combines the advantages of a conventional hybrid functional with those of a functional with 100% HFX, but without the disadvantages of the latter (in particular, some error compensation is retained).[140, 144] The exact long-range decay of range-separated hybrids is especially beneficial for the description of Rydberg states and charge transfer states which exhibit a low overlap between the orbitals that are involved in the transition, and it is in these very transitions that both a proper r^{-1} decay and the inclusion of exact exchange become crucial (either ingredient alone will not suffice).

For now, we shall leave the discussion of range-separated functionals at this abstract level. A much more detailed account of the approach, the choice of $f(x)$ and its application, along with its implementation in a plane wave/pseudopotential framework, will be given in Chapter 5 of the next Part, where we will also describe the formal modifications to the GGA and exact exchange functional that are required upon introducing the resolution of identity in eq. 3.75.

Principles of Generalised Kohn-Sham Schemes By definition, the noninteracting orbitals of Kohn-Sham DFT are all subject to the same effective potential $v_{\text{eff}}(\mathbf{r})$, obtained as the functional derivative of $E_{\text{xc}}[\rho]$. Once $E_{\text{xc}}[\rho]$ depends explicitly on the noninteracting $\psi_i(\mathbf{r})$ - which, within KS-DFT, are still an implicit functional of $\rho(\mathbf{r})$ - it is this very functional derivative that becomes both notationally and computationally tedious. It may therefore appear appealing to resort to the orbital-dependent exchange (or Fock) potential of Hartree-Fock theory, but such a potential is *not* admissible within what is now commonly referred to as the Kohn-Sham scheme[xii].

[xii] Even though Kohn and Sham themselves were the first to propose the use of pure Hartree-Fock exchange (along with the corresponding orbital-dependent potential).[15]

A proper formalisation of this problem is due to Görling and Levy.[137,145] Consider a general AC Hamiltonian of the form of eq. 3.36 and a search space of determinantal wavefunctions $\Psi^{\mathrm{GKS}} \in \mathscr{W}_0$. We shall use a general two-electron operator $\hat{\mathrm{W}}$ rather than the classical Coulomb operator, thus pre-empting use of alternative forms for the electron-electron interaction. One then decomposes the Hohenberg-Kohn functional according to:

$$F_{\mathrm{HK}}[\rho] = F_{\lambda}^{\mathrm{GKS}}[\rho] + Q_{\lambda}^{\mathrm{GKS}}[\rho], \tag{3.76}$$

with

$$F_{\lambda}^{\mathrm{GKS}}[\rho] = \min_{\Psi} \left\{ \langle \Psi^{\mathrm{GKS}} | \, \hat{\mathrm{T}} + \lambda \hat{\mathrm{W}} \, | \Psi^{\mathrm{GKS}} \rangle \, \middle| \, \Psi^{\mathrm{GKS}} \to \rho, \Psi^{\mathrm{GKS}} \in \mathscr{W}_0 \right\} \tag{3.77}$$

and $Q_{\lambda}^{\mathrm{GKS}}[\rho]$ defined as $F_{\mathrm{HK}}[\rho] - F_{\lambda}^{\mathrm{HFX}}[\rho]$. The functional $F_{\lambda}^{\mathrm{HFX}}[\rho]$ searches over all single N-electron determinants that yield a given $\rho(\mathbf{r})$. For $\lambda \neq 0$, $\mathrm{T} \neq \mathrm{T}_s$ due to the presence of an orbital-dependent potential[xiii]. Solving the *generalised* Kohn-Sham (GKS) equations is then equivalent to solving the Hartree-Fock equations with a generalised two-electron operator and an effective potential $v_{\lambda}^{\mathrm{eff}}(\mathbf{r})$, which corresponds to the AC Hamiltonian evaluated at fixed ρ and a defined value of λ, but using a single-determinantal search space $\mathscr{W}_0$ only (in lieu of a search over density matrices). We are left with determining $v_{\lambda}^{\mathrm{eff}}(\mathbf{r})$ such that the minimising density fulfils $\rho \in \mathscr{A}_N$.

Subtracting the Euler-Lagrange equations for $F_{\mathrm{HK}}[\rho]$ and $F_{\lambda}^{\mathrm{GKS}}[\rho]$ evaluated at the ground-state density $\rho(\mathbf{r}) = \rho_0(\mathbf{r})$ leads to:

$$v_{\lambda}^{\mathrm{eff}}(\mathbf{r}) - v_{\mathrm{eff}}(\mathbf{r}) = \left. \frac{\delta \left(F_{\mathrm{HK}}[\rho] - F_{\lambda}^{\mathrm{GKS}}[\rho] \right)}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_0} = \left. \frac{\delta Q_{\lambda}^{\mathrm{GKS}}[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_0}, \tag{3.78}$$

and hence

$$\begin{aligned} v_{\lambda}^{\mathrm{eff}}(\mathbf{r}) &= v_{\mathrm{eff}}(\mathbf{r}) + \left. \frac{\delta Q_{\lambda}^{\mathrm{GKS}}[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho=\rho_0} && (3.79) \\ &= v_{\mathrm{eff}}(\mathbf{r}) + (1-\lambda) v_J(\mathbf{r}) + v_{\mathrm{xc},\lambda}^{\mathrm{GKS}}(\mathbf{r}), && (3.80) \end{aligned}$$

where $v_{\mathrm{xc},\lambda}^{\mathrm{GKS}}(\mathbf{r})$ is defined by eqs 3.79 and 3.80, and $v_J(\mathbf{r})$ is the classical Coulomb potential. We then decompose further:

$$v_{\mathrm{xc},\lambda}^{\mathrm{GKS}}(\mathbf{r}) = (1-\lambda) v_x(\mathbf{r}) + v_c^{\mathrm{GKS}}(\mathbf{r}), \tag{3.81}$$

[xiii] *i.e.* Hartree-Fock correlation

where $v_c^{\mathrm{GKS}}(\mathbf{r})$ has yet again been defined by definition.

Quintessentially, the effective GKS potential differs from the KS potential by the Coulomb and exact exchange contribution that are implicit to the use of $\hat{\mathrm{W}}$ and the resulting orbital-dependent potential, which gives rise to a scaling by $(1-\lambda)$ with respect to the full effective KS potential for exchange. However, $v_c^{\mathrm{GKS}}(\mathbf{r}) \neq v_c(\mathbf{r})$, since the kinetic contribution due to $F_\lambda^{\mathrm{GKS}}[\rho]$ is $T_\lambda - T_s \neq 0$, whereas the difference $T_{\lambda=1} - T_s = T_c$ is explicitly accounted for in the KS correlation functional! It is however customary to use the same correlation functionals in both GKS and KS calculations. The success of hybrid functionals can then not only be rationalised by improvements made to the exchange hole, but also by the ability of common approximated $E_c[\rho]$ to describe $T_c^{\mathrm{GKS}} = T_c - T_\lambda$ (which might be better described than T_c alone).[145,146] GKS theory provides a true hybridisation that goes beyond the mixing of exact exchange with semi-local functionals: The use of an orbital-dependent potential results in a set of Hartree-Fock like equations which are solved subject to an external potential v_λ^{eff}. Just as in the spirit of the original Kohn-Sham scheme - which covers the special case $\lambda = 0$ -, the determinant Ψ^{GKS} is *not* the best single-determinantal approximation to the full Ψ, but instead provides $\rho_0(\mathbf{r})$ through $v_\lambda^{\mathrm{GKS}}(\mathbf{r})$.

In the following, a certain notational leniency should be allowed. We will use ψ_i for single-particle orbitals, independent of their coming from a GKS or KS determinant. Similarly, for the exact exchange functional, we shall write $E_x[\rho]$ rather than $E_x[\{\psi\}]$ since in KS-DFT (with $\lambda = 0$ and a fully local $v_x(\mathbf{r})$) the Kohn-Sham orbitals are uniquely defined by eq. 3.29, whereas in GKS-DFT ($0 < \lambda \leq 1$, with a hybridised exchange-correlation potential), they follow by virtue of eq. 3.79. We finally note that, just like in KS-DFT, the (trial) GKS orbitals generated during an energy minimisation do not necessarily come from any ground-state $\rho(\mathbf{r})$, which we have assumed by the use of $F_{\mathrm{HK}}[\rho]$ in eq. 3.77.

We have now gathered the theoretical basis needed for the numerical simulation of chemicophysical processes. Still, it is the *basis* itself that we still lack: The basis vectors in which our systems of equations are expanded in practice have gone unmentioned so far. The next chapter will conclude this First Part with a discussion of one of the many possible choices of basis set, a choice that has proven to be particularly attractive for molecular dynamics: The plane wave/pseudopotential framework.

Chapter 4

A Numerical Solution

Sampling the quantum partition function in an approximate way is an art - solving such an approximation numerically is an art on its own.

Different communities have developed different approaches, and the following discussion shall be limited to a basis set that is of particular appeal in the study of condensed matter not at least due to its intrinsic periodicity: In the condensed phase, plane waves offer many advantages over Gaussian basis sets[147] that are most commonly used in Computational Chemistry. But every medal has two sides, and in the following sketch of the plane wave/ pseudopotential approach, we will see that these advantages can sometimes come at a cost.

For a discussion of the general numerics of Molecular Dynamics (periodic boundary conditions, integrators), the interested reader may refer to Refs 16 and 51.

4.1 An Expansion in Plane Waves

While atom-centred bases intuitively appeal to the theoretical chemist - after all, what is a molecule but a collection of atoms that interact[i] - there are certain pitfalls to such a picture once the chemical system begins to wiggle and move. Since an atom-centred basis moves along with its atoms, this gives rise to additional forces on the nuclei, the Pulay forces.[148] These have to be appropriately accounted for at a non-negligible computational cost. Once diffuse components are involved, in particular when molecules approach each other, the localised nature of the basis means that the flexibility of the basis set in the intermolecular region can change significantly, thereby altering the results. This is known as the basis set superposition error, and corrections to it are not always straightforward and remain somewhat *ad hoc*.[149] It seems desirable to have a basis set at hand that uniformly fills the space, that does not depend on the nuclear positions and that lends itself to a periodic description. One such basis set is given by the solution of the Schrödinger equation of a free particle: The plane wave basis.

---

[i] In Organic Chemistry, they interact by means of lines and arrows; in Inorganic Chemistry, the linear combination of atomic orbitals (LCAO) approach is very popular.

4.1.1 Infinitely Replicated Systems

Plane waves are very popular in solid state physics: plane wave based **k**-point sampling over unit cells has become a standard tool to elucidate the electronic structure of zero- and small-band gap materials.[16] However, due to their nonlocality, plane waves are also ideally suited for the description of the dynamics of condensed matter (implying finite band gaps) where only the Brillouin zone (the Γ-point) needs to be sampled. In the following, we shall therefore restrict our discussion to the Γ-point, at the benefit of a much sleeker notation.

In a most general form, the plane wave expansion of the KS orbitals takes the form:

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{G}}^{\mathbf{G}_{\max}} \bar{\psi}_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{4.1}$$

where $\bar{\psi}_i(\mathbf{G})$ is an expansion coefficient of the plane wave at a given $\mathbf{G}$ vector in reciprocal space. The sum is in practice finite; the vector at which it is truncated is commonly expressed in terms of the supercell-independent energy cutoff value E_{cut}:

$$E_{\text{cut}} = \frac{1}{2}\mathbf{G}_{\max}^2. \tag{4.2}$$

The number of $\mathbf{G}$-vectors below $\mathbf{G}_{\max}$ is restricted by the domain of the periodic supercell D_h and the periodicity of the plane wave $e^{i\mathbf{G}\cdot\mathbf{r}}$. If we denote by $\mathbf{h}$ the matrix obtained from the supercell Bravais vectors, $\mathbf{h} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$, the supercell restricts the number of plane waves `ngw` up to a given cutoff energy approximately by:

$$\texttt{ngw} = \frac{1}{2\pi^2} E_{\text{cut}}^{3/2} \det \mathbf{h}\,, \tag{4.3}$$

where $\Omega = \det \mathbf{h}$ is the volume of the supercell. The finite expansion of the orbitals has a direct consequence on the maximum periodicity of the $\psi_i(\mathbf{r})$.

Eq. 4.1 reveals that $\bar{\psi}_i(\mathbf{G})$ is nothing but the Fourier *series* equivalent of the momentum space representation of the continuous orbital $\psi_i(\mathbf{r})$, represented by the equally continuous $\psi_i(\mathbf{G})$ (note the absence of a bar):

$$\bar{\psi}_i(\mathbf{G}) = \sum_{\mathbf{r}}^{\mathbf{r}_{\max}} \psi_i(\mathbf{r}) e^{i\mathbf{G}\cdot\mathbf{r}} = \mathcal{S}\left[\psi_i(\mathbf{r})\right], \tag{4.4}$$

$$\psi_i(\mathbf{G}) = \int \mathrm{d}\mathbf{r}, \psi_i(\mathbf{r}) e^{i\mathbf{G}\cdot\mathbf{r}} = \mathcal{F}\left[\psi_i(\mathbf{r})\right]. \tag{4.5}$$

Whereas the latter is nonperiodic and is obtained from a Fourier transform $\mathcal{F}$ (and therefore requires integrals to be taken over the entire space and momentum domain), the Fourier series representation $\mathcal{S}$[ii] implies periodicity of $\psi_i(\mathbf{r})$ over the domain D_h over which $\mathbf{r}$ or, equivalently, $\mathbf{G}$ is sampled (hence the bar in $\bar{\psi}_i(\mathbf{G})$). This domain is defined as nothing but the periodic supercell in which the system of interest is enclosed. The use of a Fourier series representation results in a discrete real-space grid expansion of $\psi_i(\mathbf{r})$, with the grid spacing being determined by the dimension of the supercell and the cutoff energy E_{cut}.

Following eq. 4.1, it becomes possible to switch between direct and reciprocal space representations by means of numerically efficient Fast Fourier Transforms (FFT). This has an important implication in the calculation of the Hartree term $J[\rho]$: Whereas cumbersome and nonlocal in direct space[iii], the matrix elements of the Coulomb operator $\hat{\mathrm{W}}$ are diagonal in reciprocal space. The Fourier transform of the L^2 real-space Coulomb potential is:

$$\begin{aligned} \langle \mathbf{G}' | \, \hat{\mathrm{W}} \, | \mathbf{G} \rangle &= \mathcal{F}\left[\langle \mathbf{r}' | \, \hat{\mathrm{W}} \, | \mathbf{r} \rangle \right] & (4.6) \\ &= \frac{4\pi}{\mathbf{G}^2} \delta(\mathbf{G}' - \mathbf{G}), & (4.7) \end{aligned}$$

which implies that the Coulomb potential of a reciprocal space density $\rho(\mathbf{G})$ is simply:

$$\left\langle \hat{\mathrm{W}} \right\rangle = \iint \mathrm{d}\mathbf{G}\, \mathrm{d}\mathbf{G}' \, \langle \mathbf{G} | \, \hat{\mathrm{W}} \, | \mathbf{G}' \rangle \, \langle \mathbf{G}' | \rho \rangle = \int \mathrm{d}\mathbf{G} \, \frac{4\pi\rho(\mathbf{G})}{\mathbf{G}^2}. \tag{4.8}$$

[ii] Fourier series

[iii] Even though this can be somewhat alleviated in Gaussian bases by making extensive use of the Gaussian product theorem.

Some technicalities are involved in the discrete representation of $\rho(\mathbf{G})$. While most conveniently constructed in real space according to eq. 3.24, the occurrence of a product $\psi_i^*(\mathbf{r})\psi_i(\mathbf{r})$ has important implications on the truncation of the sum in eq. 4.1: More Fourier components are needed in order to properly sample the resulting $\bar{\rho}(\mathbf{G})$, which implies that the plane wave cutoff used to obtain the reciprocal space density from the real-space KS orbitals needs to be higher with respect to the one adopted in the expansion of the orbitals alone. The cutoff energy for discrete Fourier transformations associated to the density, E^{ρ}_{cut}, has to be 4 times higher than the one used to expand the orbitals $\psi_i(\mathbf{r})$, E^{ψ}_{cut}. The final procedure to compute a given Coulomb potential employing a set of discrete Fourier transforms at a given cutoff value E_{cut} can be formalised as:

$$v_J(\mathbf{r}) = \mathcal{S}^{-1}_{E^{\rho}_{\text{cut}}} \left[\frac{4\pi}{\mathbf{G}^2} \left(\mathcal{S}_{E^{\rho}_{\text{cut}}} \left[\sum_i^{N_b} \left(\mathcal{S}^{-1}_{E^{\psi}_{\text{cut}}} \left[\bar{\psi}_i(\mathbf{G}) \right] \right)^2 \right] \right) \right], \tag{4.9}$$

where

$$\sum_i^{N_b} \left(\mathcal{S}^{-1}_{E^{\psi}_{\text{cut}}} \left[\bar{\psi}_i(\mathbf{G}) \right] \right)^2 = \sum_i^{N_b} |\psi_i(\mathbf{r})|^2 = \rho(\mathbf{r}), \tag{4.10}$$

and N_b denotes the number of occupied states ('number of bands' in solid state terminology). Eq. 4.9 can be solved efficiently thanks to the Fast Fourier Transform (FFT) algorithm, which scales as $\mathcal{O}(\texttt{ngw} \log(\texttt{ngw}))$, implying an overall scaling of the order $3N_b[\texttt{ngw} \log(\texttt{ngw})]$ with respect to the system size.

Comments on the Nuclear Cusp The cusp condition requires that the derivative of the wavefunction be infinite on top of the nucleus. This would imply an infinite plane wave cutoff. Even if the cusp were only approximated (*cf.* Gaussian bases[147, 150, 151]), high-frequency components due to rapid oscillations of higher-angular momentum orbitals would still require numbers of grid points that make any FFT untractably expensive. A solution to the problem lies in pseudising the atomic core.[152–154] The effect of the core orbitals on the chemically relevant valence orbitals is mimicked by an effective potential, and the valence orbitals are further smoothened inside a core radius r_c. This parameter should be chosen small enough that the effects of chemical bonding are captured for all $R \geq r_c$. Most pseudopotentials used in state of the art calculations are norm-conserving,[152] *i.e.* we have that for a radial atomic pseudo-wavefunction $\psi(r)$ and its all-electron counterpart $\psi^{\mathrm{ae}}(r)$:

$$\int_{r \geq r_c} \mathrm{d}r \; |\psi(r)|^2 = \int_{r \geq r_c} \mathrm{d}r \; |\psi^{\mathrm{ae}}(r)|^2, \tag{4.11}$$

and the eigenvalues for a given configuration of valence orbitals are identical by construction,

$$\epsilon_i = \epsilon_i^{\mathrm{ae}} \quad \text{for} \quad i \in \text{used in pseudopotential generation}; \tag{4.12}$$

and, most importantly, $\psi(r) = \psi^{\mathrm{ae}}(r)$ at $r \geq r_c$ for all orbitals used in the pseudisation procedure, making atomic pseudo-wavefunctions and all-electron wavefunctions indistinguishable in the chemically relevant space outside the core radius. Different pseudisation approaches exist, an in-depth discussion of which is readily available in the literature.[16]

With an appropriate choice of valence space and r_c, many (if not most) of the molecular properties that are calculated based on this plane wave/ pseudopotential approach are virtually indistinguishable from all-electron treatments[iv].

4.1.2 First Image Periodicity and the Poisson Solver

As compelling as eq. 4.8 is, it binds our discussion to infinitely replicated periodic systems.[155] For if $\rho(\mathbf{r})$ is implicitly periodic in a plane wave expansion in the spirit of eq. 4.1, restricting sums or integrals to the domain D_h, the matrix elements of $\langle \mathbf{G}' | \, \hat{\mathrm{W}} \, | \mathbf{G} \rangle$ are obtained from a Fourier transformation (and not a Fourier series) of the Coulomb operator that has to be carried out over the *entire* space. $\hat{\mathrm{W}}$ will therefore incorporate long-range interactions between *all* the periodic images of $\rho(\mathbf{r})$, which are sizeable due to the slow decay of the Coulomb operator with respect to spatial distance.

The plane wave basis can still be a suitable choice for clusters (‘isolated systems’), since we can always assume that the cluster density simply vanishes at the boundaries of D_h, while nothing keeps it from having the periodicity of the domain implicit to the expansion in eq. 4.1.[155,156] We may then define any applied (general) potential Φ to have the periodicity of the domain, too, which is the *first* or *nearest image form.* The error due to a first image periodic description when describing a true isolated cluster can be controlled by the size of the simulation supercell. Written with respect to the expectation value of the general potential Φ^{155}:

[iv] This will be demonstrated later on at the example of the results presented in Part II.

$$\left\langle \Phi^{(\mathrm{c})} \right\rangle = \left\langle \Phi^{(1)} \right\rangle + \epsilon^{(1)}(\mathbf{h}). \tag{4.13}$$

In contrast to the true cluster limit that we shall denote by c, the first image form will take the superscript 1. The decay of $\epsilon^{(1)}(\mathbf{h})$, and thus the approach to the cluster limit, can be systematically analysed by performing convergence tests on simulation supercells of different sizes. For the Coulomb potential of an electron density, the cluster limit in eq. 4.13 is reached once the simulation supercell spans about twice the extent of the charge density under consideration (which is also the requirement for the validity of a multipole expansion, but does not apply to long-range interactions in charged or highly polar systems).[155]

While this gives a practical justification for the use of the nearest-image periodic form, we still lack an appropriate first-image Coulomb operator. Tuckerman and Martyna[155] suggested to obtain the proper Fourier series representation of $\hat{\mathrm{W}}$, denoted by $\hat{\mathrm{W}}^{(\mathrm{c})}$, by splitting both the periodic and the first image potential into long- (sr) and short-range (lr) components:

$$\hat{\mathrm{W}} = \hat{\mathrm{W}}^{\mathrm{lr}} + \hat{\mathrm{W}}^{\mathrm{sr}}. \tag{4.14}$$

The short range is then chosen such that the integral in the Fourier transform in eq. 4.8 taken over the whole space does not differ from the integral taken over the domain of the simulation supercell D_h (Fourier series representation) only:

$$\begin{aligned} \left\langle \mathbf{G}' \right| \hat{\mathrm{W}}^{\mathrm{sr}} \left| \mathbf{G} \right\rangle \big|_{\mathcal{S}_{D_h}} &= \mathcal{S}_{D_h}\left[\left\langle \mathbf{r}' \right| \hat{\mathrm{W}}^{\mathrm{sr}} \left| \mathbf{r} \right\rangle \right] && (4.15)\\ &= \mathcal{F}\left[\left\langle \mathbf{r}' \right| \hat{\mathrm{W}}^{\mathrm{sr}} \left| \mathbf{r} \right\rangle \right] + \epsilon && (4.16)\\ &= \left\langle \mathbf{G}' \right| \hat{\mathrm{W}}^{\mathrm{sr}} \left| \mathbf{G} \right\rangle \big|_{\mathcal{F}} + \epsilon, && (4.17) \end{aligned}$$

where we have introduced a vertical bar to represent matrix elements obtained as either discrete Fourier series coefficients ($|_{\mathcal{S}_{D_h}}$), or calculated from Fourier transforms and evaluated at quantised values of $\mathbf{G}$, ($|_{\mathcal{F}}$) respectively. The error ϵ is exponentially small if the range of $\hat{\mathrm{W}}^{\mathrm{sr}}$ is small enough with respect to D_h and can therefore be neglected provided that the domain of the real-space operator be suitably chosen. We may therefore write:

$$\begin{aligned} \left\langle \mathbf{G}' \right| \hat{\mathrm{W}} \left| \mathbf{G} \right\rangle \big|_{\mathcal{S}_{D_h}} &= \left\langle \mathbf{G}' \right| \hat{\mathrm{W}}^{\mathrm{lr}} \left| \mathbf{G} \right\rangle \big|_{\mathcal{S}_{D_h}} + \left\langle \mathbf{G}' \right| \hat{\mathrm{W}}^{\mathrm{sr}} \left| \mathbf{G} \right\rangle \big|_{\mathcal{F}} + \epsilon && (4.18)\\ &= \left\langle \mathbf{G}' \right| \hat{\mathrm{W}}^{\mathrm{sc}} \left| \mathbf{G} \right\rangle + \left\langle \mathbf{G}' \right| \hat{\mathrm{W}} \left| \mathbf{G} \right\rangle \big|_{\mathcal{F}} + \epsilon, && (4.19) \end{aligned}$$

where we have introduced matrix elements of a screening function *sc* that is easily seen to be:

$$\left\langle \mathbf{G}'\right| \hat{\mathrm{W}}^{\mathrm{sc}} \left|\mathbf{G}\right\rangle = \mathcal{S}_{D_h}\left[\left\langle \mathbf{r}'\right| \hat{\mathrm{W}}^{\mathrm{lr}} \left|\mathbf{r}\right\rangle\right] - \mathcal{F}\left[\left\langle \mathbf{r}'\right| \hat{\mathrm{W}}^{\mathrm{lr}} \left|\mathbf{r}\right\rangle\right]. \tag{4.20}$$

This function screens the interaction of the system with its infinite replicas and is nothing but the difference between the Fourier series coefficients of the Coulomb potential and its analytical Fourier transform for a given value of $\mathbf{G}$. This is rather important semantics. The Coulomb potential of a given density of an isolated system is then obtained from:

$$\left\langle \hat{\mathrm{W}}^{(\mathrm{c})}\right\rangle = \int \mathrm{d}\mathbf{G}' \left\langle \mathbf{G}\right| \hat{\mathrm{W}} + \hat{\mathrm{W}}^{\mathrm{sc}} \left|\mathbf{G}'\right\rangle \left\langle \mathbf{G}'|\rho\right\rangle = \left(\frac{4\pi}{\mathbf{G}^2} + \Phi^{\mathrm{sc}}(\mathbf{G})\right)\rho(\mathbf{G}), \tag{4.21}$$

where $\Phi^{\mathrm{sc}}(\mathbf{G}) = \int \mathrm{d}\mathbf{G}' \left\langle \mathbf{G}'\right| \hat{\mathrm{W}}^{\mathrm{sc}} \left|\mathbf{G}\right\rangle$. $\Phi^{\mathrm{sc}}(\mathbf{G}) = 0$ recovers the infinitely replicated interactions for full periodic boundary conditions. Tuckerman and Martyna also proposed an efficient scheme to calculate the matrix elements $\mathcal{S}_{D_h}[\langle \mathbf{r}'|\, \hat{\mathrm{W}}^{\mathrm{lr}} \,|\mathbf{r}\rangle]$ needed for the screening function, which can be calculated once at the beginning of the simulation. The *direct* computational overhead due to the method is therefore very low, since the matrix elements can be precomputed. However, due to the use of the nearest image convention, in order for the method to be valid, the simulation supercell must be chosen at least twice as large as the extent of the density of the system - given the cubic scaling of the FFT with increasing supercell volume, this renders calculations on large systems rather computationally expensive.

We have not discussed an alternative algorithm due to Hockney,[156] which requires additional Fourier transforms to be performed in every iteration of the density minimisation. The Hockney model only requires that the charge density vanish at the border of the simulation supercell, which can easily be accounted for by modestly enlarging the latter[v]. We should not fail to mention that this is beneficial for larger systems, since the overhead due to the computation of two additional FFTs in an overall smaller simulation supercell is lower than the one due to the conditions imposed by the Martyna-Tuckerman approach. It is, however, prohibitively expensive if exact exchange integrals need to be evaluated, as we shall see in the next few paragraphs. In those cases, the Martyna-Tuckerman algorithm as outlined above is the method of choice.

---

[v] However, *stricto sensu*, the Hockney algorithm is limited to spherical charge densities.

4.1.3 Exact Exchange Potentials in Plane Waves

While hybrid functionals have become a standard tool for many applications, their use in plane wave based *first principles* dynamics is hampered by the considerable computational cost associated to the evaluation of the integrals of the exact exchange functional.[157] While analytical expression in reciprocal space can be derived both for the energy and the potential, they are nonlocal in **G**, making them untractable for practical applications.[158] Instead, by recognising that orbital pairs $\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})$ of a single noninteracting Slater determinant form nothing but elements of the one-particle reduced density matrix, an algorithm similar to the computation of the Coulomb term of the N-electron density can be devised.[157] In particular, we can rewrite eq. 3.72 in terms of eq. 3.45 by introducing the pair densities $\rho_{ij}(\mathbf{r}) = \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})$, which gives:

$$E_x^0[\rho] = -\frac{1}{2}\sum_i^{N_b}\sum_j^{N_b}\int \mathrm{d}\mathbf{r}\,\rho_{ij}^*(\mathbf{r})\int \mathrm{d}\mathbf{r}'\,\frac{\rho_{ij}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{4.22}$$

where we have only used occupied orbitals in constructing the $\rho_{ij}(\mathbf{r})$, thus circumventing the use of occupation numbers. As for the N-electron density, we may exploit the reciprocal space Coulomb operator in eq. 4.8 to obtain E_x^0 *via* two discrete Fourier transforms. The reciprocal space representation of ρ_{ij} fulfils the properties of a correlation function, and after some algebra, we find that:

$$E_x^0[\rho] = -\frac{1}{2}\sum_i^{N_b}\sum_j^{N_b}\sum_{\mathbf{G}}^{\mathbf{G}_{\max}}\Phi(\mathbf{G})\left|\rho_{ij}(\mathbf{G})\right|^2, \tag{4.23}$$

where we have used a generalised Coulomb matrix element $\Phi(\mathbf{G})$. In a cluster, $\Phi(\mathbf{G}) = \frac{4\pi}{\mathbf{G}^2} + \Phi^{\mathrm{sc}}(\mathbf{G})$ as discussed earlier on in the text, whereas the $\mathbf{G} = 0$ term needs to be appropriately screened in a fully periodic setup.[158, 159] The necessity for this screening stems from the divergence of the *discrete* representation of **G**-vectors at $\mathbf{G} = 0$ and would not occur in a continuous momentum basis, since it is integrable. A discussion of this issue will be given in Chapter 5 of this text; for what follows, it suffices to note that $\Phi(\mathbf{G})$ can therefore never be just $\frac{4\pi}{\mathbf{G}^2}$. Now, just as for $\rho(\mathbf{r})$, $\rho_{ij}(\mathbf{r})$ can be obtained from the reciprocal-space expansion coefficients by a series of Fast Fourier Transforms:

$$\rho_{ij}(\mathbf{r}) = \left(\mathcal{S}^{-1}_{E^{\psi}_{\mathrm{cut}}}\left[\bar{\psi}_i(\mathbf{G})\right]\right)^* \mathcal{S}^{-1}_{E^{\psi}_{\mathrm{cut}}}\left[\bar{\psi}_j(\mathbf{G})\right], \tag{4.24}$$

$$\bar{\rho}_{ij}(\mathbf{G}) = \mathcal{S}_{E^{\rho}_{\mathrm{cut}}}\left[\rho_{ij}(\mathbf{r})\right]. \tag{4.25}$$

The calculation of the exact exchange energy therefore requires at least N_b inverse FFTs at the wavefunction cutoff E^{ψ}_{cut} (provided that the ψ can be stored) and N_b^2 FFTs on the density cutoff E^{ρ}_{cut}.[157,160] In practice, due to memory restrictions, inverse FFTs have to be recomputed, making the scaling on E^{ψ}_{cut} approach N_b^2, too. However, it is possible to decrease the effective number of inverse FFTs in eq. 8.13 by storing *blocks* of them. An efficient blocking scheme for the orbital pairs has been presented in Ref. 160. In a plane wave basis, there is no tractable equation comparable to the Roothaan-Hall equation used for atom-centred non-orthogonal functions, and the orbital coefficients have to be updated directly from the (functional) derivative of the energy with respect to an orbital, evaluated at a given discrete value of the basis $\mathbf{G}$. This quantity is obtained from a FFT of the real-space functional derivative, performed at the wavefunction cutoff.

$$\frac{\delta E_x^0[\rho]}{\delta \bar{\psi}_i^*(\mathbf{G})} = \mathcal{S}_{E^{\psi}_{\text{cut}}} \left[\frac{\delta E_x^0[\rho]}{\delta \psi_i^*(\mathbf{r})} \right], \tag{4.26}$$

where it is easy to show that

$$\frac{\delta E_x^0[\rho]}{\delta \psi_i^*(\mathbf{r})} = \sum_j^{N_b} \psi_j(\mathbf{r}) \int \mathrm{d}\mathbf{r}' \, \frac{\rho_{ij}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{4.27}$$

We now have from eq. 4.9 that:

$$\frac{\delta E_x^0[\rho]}{\delta \psi_i^*(\mathbf{r})} = \left[\sum_j^{N_b} \psi_j(\mathbf{r}) \left(\mathcal{S}^{-1}_{E^{\rho}_{\text{cut}}} \left[\Phi(\mathbf{G}) \bar{\rho}_{ij}(\mathbf{G}) \right] \right) \right], \tag{4.28}$$

which immediately implies that the calculation of the potential is substantially more expensive than that of the energy, since for every orbital pair ij, an additional inverse FFT on the density cutoff and one additional FFT on the wavefunction cutoff have to be performed. Hence, when compared to a (semi-)local density functional, calculations involving hybrid functionals carry an overhead of up to $2N_b^2(\text{FFT}[E^{\psi}_{\text{cut}}] + \text{FFT}[E^{\rho}_{\text{cut}}])$. It is therefore straightforward to see why hybrid xc functionals have often played a minor role only in plane wave based *first principles* MD: In the absence of vast computational resources (about a million of threads on a BG/Q), they are simply frustratingly expensive, greatly limiting the simulation time scale that can be reached within a reasonable project time.[160] The problem is even exacerbated for isolated systems, since the Hockney algorithm would require two additional FFTs for every *single* orbital pair, whereas the overall cost of the

FFT is artificially enlarged in the scheme by Tuckerman and Martyna due to the nearest image convention applied.

But not to despair. In Part III, we shall sketch a method which can partly alleviate the problems encountered in isolated systems, and which results in a considerable speedup of the exact exchange computation. However, before tackling this fascinating issue, we will take a step back to the functionals discussed in Chapter 3.2.4 at the example of two implementations of popular and rather accurate xc functionals in the plane wave/pseudopotential molecular dynamics package CPMD, which will be presented in Part II.

Part II

Climbing Jacob's Ladder in *first principles* Molecular Dynamics

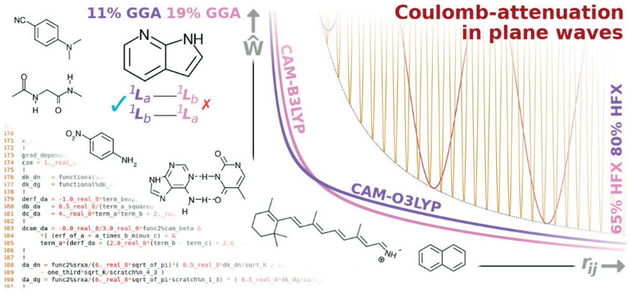
Preface

This part will be concerned with the implementation and validation of improved approximations to the xc functional in the plane wave/pseudopotential code CPMD. For one, we will present the implementation of a popular range separation scheme already outlined in Chapter $3.2.4_{\mathrm{E}}$. We will also discuss intricacies linked to a particular family of meta-GGA functionals first mentioned in Chapter $3.2.4_{\mathrm{C}}$, named the Minnesota family, when implemented in a highly flexible basis such as plane waves.

CHAPTER 5

Plane-Wave Implementation and Performance of à-la-carte Coulomb-Attenuated Exchange-Correlation Functionals for Predicting Optical Excitation Energies in Some Notorious Cases

Chapter 5 is a post-print version of an article published as:


11% GGA 19% GGA
^1L_a — ^1L_b ✗
^1L_b — ^1L_a ✓
ŵ
CAM-B3LYP
CAM-O3LYP
Coulomb-attenuation in plane waves
65% HFX 80% HFX
r_{ij}

Bircher, Martin P.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **2018**, *14*, 3184–3195

LINEAR-RESPONSE time-dependent density functional theory (LR-TDDFT) has become a valuable tool in the calculation of excited states of molecules of various sizes. However, standard generalised gradient approximation (GGA) and hybrid exchange-correlation (xc) functionals often fail to correctly predict charge-transfer (CT) excitations with low orbital overlap, thus limiting the scope of the method. The Coulomb-attenuation method (CAM) in the form of the CAM-B3LYP functional has been shown to reliably remedy this problem in many CT systems, making accurate predictions possible. However, in spite of a rather consistent performance across different orbital overlap regimes, some pitfalls remain. Here, we present a fully flexible and adaptable implementation of the CAM for Γ-point calculations within the plane-wave pseudopotential molecular dynamics package CPMD and explore how customised xc functionals can improve the optical spectra of some notorious cases. We find that results obtained using plane waves agree well with those from all-electron calculations employing atom-centred bases, and that it is possible to construct a new Coulomb-attenuated xc functional based on simple considerations. We show that such a functional is able to outperform CAM-B3LYP in some notorious cases, while retaining similar accuracy in systems where CAM-B3LYP performs well.

5.1 Introduction

Be it for the vital conversion of sunlight to chemical energy in a leaf, for photochemical reactions causing harmful DNA damage to skin, for the blue fluorescence of scorpions, or for energy conversion in man-made solar cells: electronically excited states are of crucial importance to fundamental processes in Nature, and in scientific fields ranging from biology over chemistry to solid state physics. The theoretical description of the excitations which are at the base of these phenomena makes it possible to ultimately gain an improved understanding of these key events.

A fully correlated description at the wave-function level is, unfortunately, prohibitively expensive for many, if not most systems of relevant size. It is thanks to the considerable progress in the field of Kohn-Sham time-dependent density functional theory (KS-TDDFT)[14, 15, 79] that such excited-state processes can these days be described at a comparably moderate computational cost. The application of the linear-response formulation of TDDFT[162] is routinely used by many a computational chemist, and the ever increasing availability of computational resources has made it possible to describe larger and larger systems. However, like ground-state DFT and even more so, the choice of an appropriate exchange-correlation (xc) functional, or respectively xc kernel, is crucial, and often decides between results in good agreement with high-level reference data, and spectra which are considerably red-shifted and exhibit an incorrect ordering of states.

The accuracy of the calculated transitions is mainly hampered by an inaccurate long-range decay of the xc potential, which leads to larger errors in the prediction of the excitation energies.[143] Hence, in analogy to their tendency to underestimate HOMO-LUMO gaps, generalised-gradient approximation (GGA)[124] xc functionals typically shift excitation energies by a considerable factor and may fail to recover the correct ordering of states.[139] Hybrid functionals,[31] which include a fixed proportion of exact exchange, may partly alleviate this problem for states in which there is an overlap between the orbitals involved in the transitions. However, even hybrid functionals will inevitably fail to describe low-overlap charge-transfer (CT) and Rydberg states, where the $1/r$ decay of the Coulomb operator is an important constituent in the correct description of the interaction between spatially distant orbitals.[140] But it is the inclusion of exact exchange that is most vital for these transitions: While an asymptotic correction of the (GGA) xc potential alone recovers the proper $1/r$ dependency and improves the description of Rydberg states, it cannot successfully capture the effects of pronounced charge separation.[139] A promising remedy to this problem has been found in an appropriate splitting of the Coulomb operator, making it possible to ensure a correct decay of the xc potential at long-range, while keeping the accuracy and simplicity of a standard local formulation for the short-range components. The long-range correction (LC) scheme[141, 143] and its generalisation, the Coulomb-attenuation method (CAM),[38] separate the Coulomb operator using an error function. While the short-range components are described using the GGA, the long-range components are taken into account using the exact exchange operator. This splitting captures the essentials of charge-transfer transitions and considerably increases the accuracy for such states. LC- and CAM-based xc functionals have therefore become a standard tool for the calculation of molecular excitation energies with pronounced charge-transfer character within TDDFT.[46] The correspondingly modified Coulomb operator is easily implemented in a Gaussian basis.

CAM-B3LYP is not only the first functional that was constructed using the Coulomb-attenuation method, it has also become the probably most prominent and abundantly used functional of this family. Peach *et al.*[140] have assessed its performance on a diverse test set of main-group molecules shortly after the original paper presenting CAM-B3LYP, and many successful applications of the functional followed: Among the substantial improvements documented with respect to B3LYP[31, 163] or PBE,[103] Peach *et al.* have found that CAM-B3LYP accurately describes excitations in (poly)acenes including naphthalene, yielding correct state ordering. They could also show that the spectra of model dipeptides improve substantially with respect to non-Coulomb attenuated functionals.[139] Subsequently, other studies have found that the use of CAM-B3LYP can predict spectra of AT nucleobase dimers[164] and indoles[165] of varying substitution patterns with increased accuracy. All these systems exhibit charge-transfer or Rydberg transitions between spatially separated states, which results in the inevitable failure of GGA and conventional hybrid functionals.

The orbital overlap in a transition can be quantitatively characterised by the parameter Λ introduced by Peach *et al.*[140]; a small value of Λ indicating a small overlap between the involved orbitals, a large value a substantial one. In their study covering excitations of different character, they have found that both GGA and conventional hybrids suffer from an inconsistent performance over the complete range of Λ values: For overlap values $\Lambda < 0.4$ in the case of PBE and $\Lambda < 0.3$ in the case of B3LYP, the errors in the excitation energies become substantially bigger. For CAM-B3LYP, no such correlation was found over the whole range $0 < \Lambda < 1$. CAM-B3LYP was shown to fare particularly well for charge transfer excitations, especially in the 'low overlap' regime, $\Lambda < 0.3$; although some cases with charge transfer character in the 'intermediate Λ regime' are accurately described, too (*e.g.* the retinal protonated Schiff base[166, 167]). However, in systems where there is significant overlap between the orbitals involved in the transition, conventional hybrid functionals such as PBE0[104, 168] often fare better, and CAM-B3LYP tends to red-shift the excitation energies. This is notably the case for the doubly fluorescent dye DMABN, for which CAM-B3LYP tends to overestimate the excitation energies of the S_2 CT state with $\Lambda = 0.72$.[140] Fully long-range corrected functionals[144] such as LC-BLYP[143] or LC-PBE0[169] were not evaluated in the study by Peach *et al.*, but they have since been shown to perform well in certain systems with very low overlap that cannot be accurately described with CAM-B3LYP.[164, 165, 170] This is attributed to the absence of any GGA exchange at longest range, which benefits the description of Rydberg states,[140, 144] but comes at the cost of an inferior average performance especially for local excitations.[38, 144, 165]

Despite the absence of correlation between the predicted excitation energies and their Λ-values, there exist some systems even in the low-overlap range ($\Lambda < 0.3$) for which CAM-B3LYP fails to deliver an accurate description. In the case of *p*-nitroaniline, the excitation energies are reasonably predicted, but solvatochromic shifts cannot be reproduced since the ground- and excited-state dipoles are grossly over- and underestimated, respectively.[171] Whereas CAM-B3LYP predicts a correct state ordering and reasonable energetics for the excitations in the AT base pair,[164] the HOMO is predicted to lie on the wrong base when compared to higher-level wavefunction methods and basic considerations based on the ionization potential of the isolated bases.[172] Seemingly reasonable results may therefore be obtained based on the wrong physical reason. Similarly, the ordering of close-lying, low energy excitations may be inversed in some systems; this is the case for the β-dipeptide model system introduced by Serrano-Andrés *et al.*[173] and subsequently popularised in the aforementioned benchmark set by Peach *et al.*[140] Other studies have found the same problem to occur in the case of 7-azaindole,[165] even though the state ordering for other substituted indoles could be correctly predicted. A wrong ordering of states may be especially detrimental for excited-state molecular dynamics,[174] where the forces exerted on the nuclei may substantially differ between the two swapped states, leading to a quantitatively as well as qualitatively wrong propagation of the system. For systems such as the β-dipeptide and 7-azaindole,[165] the use of an LC functional may yield a qualitatively correct ordering of the low-lying excitations, but the energies often remain too low. Alternatively, the range separation parameter μ in CAM-B3LYP may be tuned in order to ameliorate the performance of the functional.[175] This process known as γ-tuning adjusts the range separation parameter to a value that accurately reproduces ionization potentials. This approach often permits for an accurate description of the excitations of interest,[165] but it constitutes a rather system-specific remedy, lacking portability and thus predictive power in comparing various systems. All of these notorious systems are included in Fig. 5.1 and have been chosen here to test the performance of an *'à-la-carte'* combination of range-separated functionals.

In other applications, a splitting of the Coulomb operator opposite to the LC and CAM scheme may be beneficial. This has been proposed in screened hybrid functionals for solid state applications,[142] where the exact exchange is limited to short- and the GGA exchange to long range. Functionals such as HSE03[142] yield results superior to those obtained with the GGA for many systems. Screened hybrids are especially beneficial in combination with a plane wave/pseudopotential approach, since they conveniently eliminate the divergence of the Coulomb operator at the $\mathbf{G} = 0$ component of the plane wave basis in sufficiently large simulation cells. Since the splitting adopted in the LC and CAM schemes does not offer any particular advantage in solid state applications and does not eliminate this problematic divergence term, these methods have received much less attention in plane wave codes.

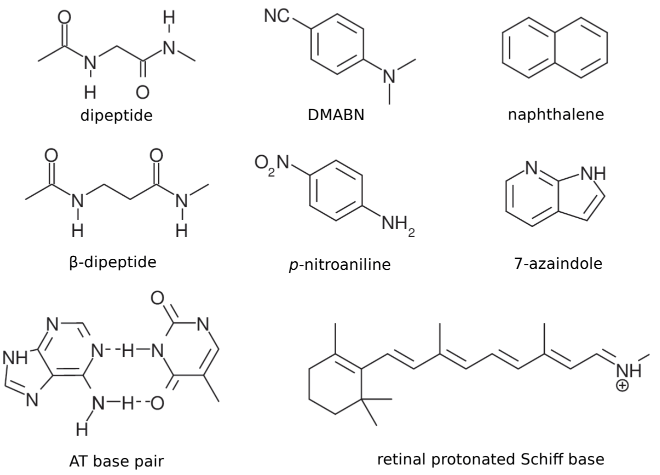


Figure 5.1: Molecules explicitly discussed in this study, all of which contain excitations that are difficult to describe in TDDFT when using a GGA or hybrid functional as the xc kernel.

In the following, we present the Coulomb-attenuation method applied to Γ-point calculations in a plane wave/pseudopotential framework. The implementation allows for the CAM to be combined with any exchange functional of choice, offering maximum flexibility. This makes the construction of customised *'à-la-carte'* Coulomb-attenuated xc functionals possible, which can be tailored to any system of choice, thus maximising the performance of the method. In order to gain maximum flexibility, all CAM parameters can be chosen freely. Our implementation of the CAM in the molecular dynamics package CPMD[176] targets applications where Γ-point sampling is routinely used, and makes the sampling of large systems possible via the fully Hamiltonian QM/MM-scheme implemented in the CPMD code. Simulations of charge-transfer systems in the gas phase as well as in condensed matter therefore become feasible using a plane wave/pseudopotential approach. To facilitate the calculation of the necessary terms, a new driver for the calculation of the exchange-correlation energy has been implemented in the CPMD code.

The paper is organised as follows: First, we give a short summary of the Coulomb-attenuation method, followed by a description of the implementation. We then give a more detailed account of the test systems used to benchmark both the implementation and a new Coulomb-attenuated xc functional. We discuss the performance of the CAM in plane waves with respect to the choice of pseudopotential, and in comparison to all-electron calculations with atom-centred basis sets. Finally, we will show how a flexible choice of the underlying GGA exchange-functional can improve accuracy in systems where standard functionals yield unsatisfactory results by comparing a customised *'à-la-carte'* CAM-xc functional constructed based on simple considerations to the well-established CAM-B3LYP.

5.2 Theory

5.2.1 The Coulomb-Attenuation Method

In the Coulomb-attenuation method, the Coulomb operator $\hat{\mathrm{W}}$ is split into two domains dominated by long-range (lr) and short-range (sr) components each:

$$\hat{\mathrm{W}} = \hat{\mathrm{W}}^{\mathrm{sr}} + \hat{\mathrm{W}}^{\mathrm{lr}}, \tag{5.1}$$

$$\left\langle \mathbf{r}' \right| \hat{\mathrm{W}}^{\mathrm{sr}} \left| \mathbf{r} \right\rangle = \frac{1 - [\alpha + \beta \,\mathrm{erf}(\mu \left| \mathbf{r} - \mathbf{r}' \right|)]}{\left| \mathbf{r} - \mathbf{r}' \right|}, \tag{5.2}$$

$$\left\langle \mathbf{r}' \right| \hat{\mathrm{W}}^{\mathrm{lr}} \left| \mathbf{r} \right\rangle = \frac{\alpha + \beta \,\mathrm{erf}\left(\mu \left| \mathbf{r} - \mathbf{r}' \right|\right)}{\left| \mathbf{r} - \mathbf{r}' \right|}, \tag{5.3}$$

where α, β and μ are adjustable parameters[38] and $\alpha = 0$ and $\beta = 1$ in the original LC method.[141,143] The first term is treated using a GGA expression for the exchange functional and becomes smaller for larger Coulomb distances, whereas the second term grows with increasing $|\mathbf{r} - \mathbf{r}'|$ and is treated using Fock's expression for the exchange energy.

The effective Coulomb operator in the exchange integrals then becomes:

$$E_x^{\mathrm{HFX}} = \frac{1}{2} \sum_{\sigma} \sum_{ij} \iint \psi_{i\sigma}^*(\mathbf{r}) \psi_{j\sigma}^*(\mathbf{r}') \frac{\alpha + \beta \,\mathrm{erf}(\mu \left| \mathbf{r} - \mathbf{r}' \right|)}{\left| \mathbf{r} - \mathbf{r}' \right|} \psi_{j\sigma}(\mathbf{r}) \psi_{i\sigma}(\mathbf{r}') \mathrm{d}\mathbf{r} \mathrm{d}\mathbf{r}'. \tag{5.4}$$

Correspondingly, the GGA enhancement factor has to be adapted to the screened Coulomb operator. The adaptation is based on the LDA for a short-range Coulomb operator and appropriately generalised[143]:

$$E_x^{\mathrm{GGA}} = \frac{1}{2} \sum_{\sigma} \rho_\sigma^{4/3} K_\sigma \left\{ 1 - \alpha - \beta \frac{8}{3} a_\sigma \left[\sqrt{\pi} \mathrm{erf} \left(\frac{1}{2a_\sigma} \right) + 2a_\sigma \left(b_\sigma - c_\sigma \right) \right] \right\}, \tag{5.5}$$

where K_σ is the spin-dependent formulation of the exchange enhancement factor and the terms due to the correction read:

$$a_\sigma = \frac{\mu K_\sigma^{1/2}}{6\sqrt{\pi}\rho_\sigma^{1/3}}, \tag{5.6}$$

$$b_\sigma = \exp\left(-\frac{1}{4a_\sigma^2}\right) - 1, \tag{5.7}$$

$$c_\sigma = 2a_\sigma^2 b_\sigma + \frac{1}{2}. \tag{5.8}$$

5.2.2 Singularity of the Coulomb Operator $\hat{\mathrm{W}}$ in Reciprocal Space

At the Γ-point, the exact exchange energy can be efficiently computed with a mixed real-space reciprocal-space algorithm[160] after introducing a resolution of identity in $\mathbf{G}$ and rearranging the terms due to the complex conjugate of the generalised Kohn-Sham (GKS) orbitals (where $\langle \mathbf{r}|i_\sigma\rangle = \langle i_\sigma|\mathbf{r}\rangle$ at the Γ-point):

$$E_x^{\mathrm{HFX}} = \sum_\sigma \sum_i \sum_{j\geq i} \int \mathrm{d}\mathbf{r}' \left\langle j_\sigma|\mathbf{r}'\right\rangle \left\langle \mathbf{r}'|i_\sigma\right\rangle \int \mathrm{d}\mathbf{r} \left\langle \mathbf{r}'\right| \hat{\mathrm{W}} \left|\mathbf{r}\right\rangle \left\langle i_\sigma|\mathbf{r}\right\rangle \left\langle \mathbf{r}|j_\sigma\right\rangle, \tag{5.9}$$

$$= \sum_\sigma \sum_i \sum_{j\geq i} \int \mathrm{d}\mathbf{G}' \int \mathrm{d}\mathbf{G} \left\langle \mathbf{G}'\right| \hat{\mathrm{W}} \left|\mathbf{G}\right\rangle \left(\int \mathrm{d}\mathbf{r} \left\langle \mathbf{G}|\mathbf{r}\right\rangle \left\langle i_\sigma|\mathbf{r}\right\rangle \left\langle \mathbf{r}|j_\sigma\right\rangle\right)^2. \tag{5.10}$$

Here, i, j index generalised Kohn-Sham orbitals. The matrix elements of the Coulomb operator in reciprocal space, $\langle \mathbf{G}'| \hat{\mathrm{W}} |\mathbf{G}\rangle = \frac{1}{\Omega}\frac{4\pi}{\mathbf{G}^2}\delta(\mathbf{G}-\mathbf{G}')$, exhibit an integrable divergence at $\mathbf{G} = 0$.[158] In practice, the plane wave/ pseudopotential formalism relies on a discrete representation of points in direct and reciprocal space, and the integrals become sums associated to discrete Fourier transforms. The divergence term becomes problematic in this discrete form, and the Coulomb operator has to be replaced by a suitable generalisation $\hat{\Phi}$:

$$E_x^{\mathrm{HFX}} = \sum_\sigma \sum_i \sum_{j\geq i} \int \mathrm{d}\mathbf{G}' \int \mathrm{d}\mathbf{G} \left\langle \mathbf{G}'\right| \hat{\Phi} \left|\mathbf{G}\right\rangle \left(\int \mathrm{d}\mathbf{r} \left\langle \mathbf{G}|\mathbf{r}\right\rangle \left\langle i_\sigma|\mathbf{r}\right\rangle \left\langle \mathbf{r}|j_\sigma\right\rangle\right)^2. \tag{5.11}$$

In the generalised matrix element $\langle \mathbf{G}'|\,\hat{\Phi}\,|\mathbf{G}\rangle$, the offending divergence is screened by a suitable function χ:

$$\langle \mathbf{G}'|\,\hat{\Phi}\,|\mathbf{G}\rangle = \begin{cases} \dfrac{1}{\Omega}\dfrac{4\pi}{\mathbf{G}^2}\delta(\mathbf{G}-\mathbf{G}') & \text{for } \mathbf{G} \neq 0 \\ \chi(0)\delta(\mathbf{G}') & \text{for } \mathbf{G} = 0, \end{cases} \tag{5.12}$$

where Ω denotes the supercell volume and $\hat{\Phi}$, like $\hat{\mathrm{W}}$, is diagonal in $\mathbf{G}$.

In the initial treatment proposed by Gygi and Baldereschi,[158] the screening term χ is obtained numerically. An auxiliary function which exhibits the same singularity as the problematic term is added to and subtracted from the Coulomb operator, and the screening is given by the difference of the discrete representation of the auxiliary function as a sum over $\mathbf{G}$ and its analytical integral over a continuous range $\mathbf{Q}$. Due to their particular choice of χ, the approach could not be applied to Γ-point sampling due to its poor convergence with respect to the number of k-points and simulation supercell size. In the following, we base our treatment on the scheme subsequently developed by Broqvist *et al*[159] which - in contrast to the initial approach by Gygi and Baldereschi - selects an auxiliary function $f(\mathbf{Q})$ that converges rapidly towards $1/\mathbf{Q}^2$:

$$f(\mathbf{Q}) = \frac{e^{-\gamma Q^2}}{Q^2}. \tag{5.13}$$

The $\mathbf{G} = 0$ term is then given by the residual difference between the integral and discrete sum over the auxiliary function:

$$\begin{aligned} \chi(0) &= \frac{1}{2\pi^2}\int_{\mathbf{Q}} f(\mathbf{Q})\mathrm{d}\mathbf{Q} - \frac{4\pi}{\Omega}\sum_{\mathbf{G}\neq 0} f(\mathbf{G}) & (5.14)\\ &= \lim_{\gamma\to 0}\left[\frac{1}{\sqrt{\pi\gamma}} - \frac{4\pi}{\Omega}\sum_{\mathbf{G}}\frac{e^{-\gamma\mathbf{G}^2}}{\mathbf{G}^2}\right], & (5.15) \end{aligned}$$

where the second term (the discrete sum) is obtained numerically. This results in a more rapid convergence with respect to the size of the supercell, and the scheme can therefore be applied in Γ-point only calculations.

Although a non-divergent analytical description for the $\mathbf{G} = 0$ component is found for screened hybrids (it is simply $\pi/\Omega^2\mu^2$), its accuracy depends on the size of the periodic supercell. If μ is small, the potential does not decay sufficiently rapidly with respect to the number of $\mathbf{G}$-vectors and a treatment analogous to the non-screened Coulomb operator has to be used.[159] Broqvist *et al* have therefore suggested to treat the short range exact exchange analogously to the full Coulomb operator:

$$\langle \mathbf{G}' | \, \hat{\Phi}_{\mathrm{sr}} \, |\mathbf{G}\rangle = \begin{cases} \dfrac{1}{\Omega} \dfrac{4\pi}{\mathbf{G}^2} \left[1 - e^{-G^2/4\mu^2} \right] \delta(\mathbf{G} - \mathbf{G}') & \text{for } \mathbf{G} \neq 0 \\ \tilde{\chi}(\mu) \delta(\mathbf{G}') & \text{for } \mathbf{G} = 0. \end{cases} \tag{5.16}$$

Using the same auxiliary function as for the full Coulomb potential, the resulting $\mathbf{G} = 0$ term for the short-range screened exchange then reads:

$$\tilde{\chi}(\mu) = \chi(0) - \chi\left(\frac{1}{4\mu^2}\right), \tag{5.17}$$

where μ is the range separation parameter.

Based on this treatment, the singularity correction for the Coulomb-attenuation method is easily found by using the identity $\mathrm{erf}(x) + \mathrm{erfc}(x) = 1$ and introducing the parameters α and β. We write for the long-range components:

$$\langle \mathbf{G}' | \, \hat{\Phi}_{\mathrm{lr}} \, |\mathbf{G}\rangle = \begin{cases} \dfrac{1}{\Omega} \dfrac{4\pi}{\mathbf{G}^2} \left[\alpha + \beta e^{-G^2/4\mu^2} \right] \delta(\mathbf{G} - \mathbf{G}') & \text{for } \mathbf{G} \neq 0 \\ \bar{\chi}(\mu, \alpha, \beta) \delta(\mathbf{G}') & \text{for } \mathbf{G} = 0, \end{cases} \tag{5.18}$$

where the $\mathbf{G} = 0$ term is simply a sum of the terms due to the full and the screened Coulomb potential, weighted by the attenuation parameters α and β:

$$\bar{\chi}(\mu, \alpha, \beta) = \alpha\chi(0) + \beta\chi\left(\frac{1}{4\mu^2}\right). \tag{5.19}$$

5.3 Implementation

Some exchange-correlation functionals make use of a linear combination of different GGA exchange contributions, such as XLYP[177] (using 72.2% B88[122] and 34.7% PW91[178] exchange) or the well-known B3LYP (using 80% LDA, 72% of the B88 gradient correction term and 20% exact exchange). Accordingly, the usage of the CAM does not have to be intrinsically limited to a single type of exchange functional (as it is the case for the most prominent CAM-B3LYP and LC-PBE0, where B88 and PBEx are attenuated, respectively). With respect to all possible combinations, our implementation achieves maximum flexibility in the choice of exchange functional by writing the exchange enhancement factor as a sum over individual contributions:

$$K_\sigma^x = \sum_f^N c_f K_\sigma^f, \tag{5.20}$$

where K_σ^f denotes any GGA exchange-functional and c_f are the corresponding weights of a total of N exchange functionals; $\sum_f^N c_f = 1$. The attenuation is then applied to K_σ^x after all N contributions have been added up. This makes it possible to consistently apply the CAM to any arbitrary combination of LDA and GGA exchange functionals, creating a custom *'à-la-carte'* xc functional.

Three terms are needed for the propagation of a Kohn-Sham wavefunction with a GGA description:

$$
\begin{aligned}
E[\rho] &= -\frac{1}{2}\sum_\sigma \int \eta[K_\sigma^x] K_\sigma^x[\rho_\sigma, \nabla\rho_\sigma]\rho_\sigma^{4/3}(\mathbf{r})\mathrm{d}\mathbf{r} && (5.21)\\
&= -\frac{1}{2}\sum_\sigma \int \epsilon[\rho_\sigma(\mathbf{r})]\mathrm{d}\mathbf{r}, && (5.22)\\
V_1^x &= \frac{\delta\epsilon[\rho]}{\delta\rho_\sigma}, && (5.23)\\
V_2^x &= \frac{\delta\epsilon[\rho]}{\delta\left|\nabla\rho_\sigma\right|}, && (5.24)
\end{aligned}
$$

where we have introduced η as the attenuation function. For functionals that are derived for the closed-shell case or use a different definition of the enhancement factor, *e.g.* $E[\rho] = \int \rho(\mathbf{r})F[\rho]\mathrm{d}\mathbf{r}$, the spin-dependent exchange enhancement factor K_σ^x is easily obtained from $F[\rho]$ or $K_{\alpha\beta}^x$ using the spin-scaling relations.

All derivatives can be efficiently calculated by making extensive use of the chain rule and by transiently storing frequently used terms (notably $\delta K_x/\delta\rho$, $\delta K_x/\delta\nabla\rho$) on the stack. In order to make further performance gains, certain powers of the density and the gradient $(\rho, \rho^{4/3}, \rho^{1/3}, |\nabla\rho|, \nabla\rho^2)$ are precomputed on a per-grid-point basis and reused in the calculation of K_σ^f and the attenuation function, thus avoiding repetitive, unnecessary operations. The implementation makes use of procedure pointers in order to facilitate the choice of functional. At the beginning of every run, the procedure pointer denoting K_σ^f is set to the selected exchange functional, no (explicit) `if`-construct is therefore necessary when looping over grid points. The new algorithm reaches an asymptotic speedup of 20% when calculating the LDA and GGA contributions with respect to the previous standard implementation, which obtained LDA and GGA contributions for each functional from separate drivers and used extensive `if`-constructs to select among the functionals at every individual gridpoint (Fig. 5.2).

```
SUBROUTINE calculate_xc_energy(rho,grad,E_x)

    real, dimension(:), &
          intent(in)   :: rho, grad
    real, intent(out)  :: E_xc
    !
    ! Stores rho, rho**1/3, rho**4/3, |\nabla rho|, |\nabla rho|**2
    !
    type(storage_t)    :: storage
    !
    ! Stores K, dK/drho, dK/d_nabla_rho, epsilon_x
    !
    type(functional_t) :: functional

    forall p in gridpts
        CALL storage%precalculate_reused_terms_from(rho(p),grad(p))
        forall f in K_x
            CALL functional%K(f)%calculate_energy_and_derivatives(storage,functional)
        if (functional_is_attenuated) CALL functional%attenuation(storage,functional)
        epsilon_x(p) = functional%epsilon_x

    E_x = sum(epsilon_x(:))

END SUBROUTINE calculate_xc_energy
```

Figure 5.2: Pseudocode describing the structure of the main routine of the new xc driver (`cp_xc_driver`) in the CPMD code.

5.4 Test Set and Computational Details

5.4.1 Description of the Test Set

Basic convergence tests with respect to basis set and supercell size were carried out on a single, isolated water molecule.

A test set for excitation energies was assembled comprising the molecules mentioned before and depicted in Fig. 5.1. The set includes the original test systems by Peach *et al.*,[140] augmented with other systems where range-separation is known to be of importance: *p*-nitroaniline, 7-azaindole, the AT nucleobase dimer and the retinal protonated Schiff base.

For the sake of comparison with localised basis sets, 5 of those systems were selected. The model dipeptide represents a typical system where CAM-B3LYP can successfully be applied: Peach *et al* found that a GGA (PBE) yields an inaccurate state ordering of the charge-transfer excitations, whereas B3LYP and CAM-B3LYP reproduce the ordering of the reference values. However, B3LYP is known to underestimate the energies of the $n_1 \rightarrow \pi_2^*$ and $\pi_1 \rightarrow \pi_2^*$ charge transfer excitations by up to 1.7eV, which are both reproduced by CAM-B3LYP with a reasonable accuracy of about 0.2eV. The spectrum of *p*-nitroaniline has also been reported to be reasonably predicted using CAM-B3LYP,[171] and although the charge separation in the first CT state was reported to be overestimated, we have used *p*-nitroaniline as another probe known to benefit from range-separation. Naphthalene served as the most simple example of an acene and another notable example of the influence of range separation: Whereas the $^1B_{3u}$ and $^1B_{3u}$ states are inverted when using both PBE and B3LYP, only CAM-B3LYP recovers the correct state ordering. However, the reported excitation energies for the two lowest optically allowed singlet transitions deviate from the reference by about $+0.21$eV for $^1B_{3u}$ and -0.16eV for $^1B_{2u}$, which results in a considerable underestimation of the state separation. The separation is reported to improve for larger acenes (anthracene, tetracene etc.); we have therefore chosen naphthalene as the most critical and sensitive compound to assess our implementation. An example where CAM-B3LYP (narrowly) fails to deliver a quantitatively correct description of states is given by 7-azaindole. Whereas a comprehensive study of various substituted indoles found CAM-B3LYP to be in good agreement with reference values, in the case of 7-azaindole,[165] the ^1L_a and ^1L_b states have been reported to be swapped. Analogously to naphthalene, the states also lie much too close in energy, but are now also incorrectly ordered. This molecule can therefore serve as a representative of excitations with CT character and low orbital overlap where CAM-B3LYP surprisingly fails. The final molecule used in the selected test set is DMABN, with a Λ value of 0.72 for the S_2 CT state it is a typical usage case for a non-range separated hybrid or even a simple GGA.[140] Indeed, whereas the errors for the ${}^1A$ and ${}^1B$ states are smaller than 0.2eV in conjecture with B3LYP, it is about doubled when using CAM-B3LYP. It serves as an example of an excitation with considerable overlap, thus completing the range of excitations covered here. The chosen test suite therefore includes both systems that are well described using CAM-B3LYP, as well as some notorious cases.

In order to assess a possible gain of accuracy by using a customised functional rather than the established CAM-B3LYP, three additional systems were studied. The retinal protonated Schiff base constitutes a system where CAM-B3LYP has been successfully used to predict both simple optical[166, 179] and two-photon absorption spectra,[180, 181] improving over conventional hybrids. In an investigation on the GC and AT nucleobase dimers,[164] it was found that CAM-B3LYP can predict accurate excitation energies for the AT base pair, but the functional localises the HOMO and the LUMO on the wrong moieties, respectively.[172] According to the ionization potential of the isolated base, the HOMO should be localized on adenine, but it is predicted to lie on thymine. The AT base pair can therefore serve as a probe for the correct orbital localization obtained with a given functional. Similarly as in 7-azaindole, in the β-dipeptide studied by Peach *et al*, all PBE, B3LYP and CAM-B3LYP fail to describe the ordering of the $\pi_1 \rightarrow \pi_2^*$ and $n_1 \rightarrow \pi_2^*$ transitions, with errors being larger than 0.75eV for CAM-B3LYP and reaching a maximum value of about 4.5eV when using PBE. The β-dipeptide therefore serves as yet another example in which CAM-B3LYP does not work even on a quantitative level, probably constituting the hardest test case for new range-separated functionals. Finally, in order to verify a possible overall gain of accuracy, the MAE of both CAM-B3LYP and CAM-O3LYP is compared for the complete test set by Peach *et al.*[140]

5.4.2 Computational Setup

Calculations using Gaussian basis sets were either carried out using DALTON 2016[182] (p-nitroaniline, the retinal protonated Schiff base and naphthalene) or the Gaussian09 suite of programs[183] (the remaining molecules) using Dunning's correlation-consistent basis sets.[151] All calculations employing a Slater basis set[184] were carried out using ADF.[185–187]

The structures for the dipeptides, DMABN and naphthalene were taken from the database published by Peach *et al.*. The structure from the retinal protonated Schiff base is an arbitrarily chosen snapshot extracted from molecular dynamics simulations at 300K, whereas the structure of p-nitroaniline was optimised using the aug-cc-pVTZ basis and the B3LYP xc functional. Structural optimisations for the remaining molecules were performed using the Gaussian suite of programs, following the published protocols of existing benchmarks for the AT base pair[164] and and 7-azaindole.[165] Excitation energies, where not quoted from the literature, were calculated using the Tamm-Dancoff approximation to TDDFT[188] and the cc-pVDZ, aug-cc-pVDZ and d-aug-cc-pVDZ basis sets, respectively.

The new exchange-correlation driver was implemented in a development version of CPMD[176] (successor of version 4.1). The GKS orbitals were expanded in plane waves contained in an orthorhombic supercell of varying dimensions and using either Martins-Troullier (MT)[153] or Goedecker-Teter-Hutter (GTH)[154] pseudisation of the atomic core orbitals (the respective values for the energy cutoff are given in the results section; the supercell size for every system is available in the supporting information). Following standard practice for hybrid functionals, BLYP pseudopotentials were used for CAM-B3LYP calculations, and OLYP pseudopotentials were used for calculations with CAM-O3LYP. The density was expanded with a 4 times greater cutoff value than the one adopted for the orbitals. The Poisson equations for the isolated systems were solved using the algorithm of Martyna and Tuckerman.[155] All calculations made use of an atomic wavefunction initialization using distributed Lanczos,[189] the new distributed linear algebra algorithm by Bekas and Curioni[190] and the 'new' exact exchange driver by Weber *et al.*[160]; the cutoff in the calculation of the Fock exchange energy were not changed with respect to the standard values for orbitals and density.

5.5 Results and Discussion

5.5.1 Convergence of Eigenvalues

5.5.1$_{\mathrm{A}}$ HOMO-LUMO Gaps in a Plane Wave/Pseudopotential Basis

When implementing the singularity correction for (screened) hybrid functionals, Broqvist *et al.*[159] have also assessed the convergence of the GKS-HOMO-LUMO gaps with respect to both the energy cutoff and the size of the supercell. In the following, we shall present a similar assessment on an isolated water molecule. Considering that the LUMO is very diffuse in this specific case, the HOMO-LUMO gap appears to be a sensitive measure of convergence.

Table 5.1: HOMO-LUMO gap of a water molecule calculated using CAM-B3LYP and GTH pseudopotentials at increasing energy cutoffs and varying cubic simulation cell lengths l using periodic or isolated system boundary conditions for solving the Poisson equation.

E_{cut} [Ry]	$\Delta\epsilon_i$ [eV], isolated system			$\Delta\epsilon_i$ [eV], periodic system		
	10 Å	20 Å	30 Å	10 Å	20 Å	30 Å
70	10.080	10.361	10.399	10.105	10.362	10.399
80	10.073	10.420	10.447	10.100	10.420	10.448
100	10.150	10.490	10.510	10.177	10.491	10.509
120	10.259	10.531	10.548	10.284	10.532	10.548
150	10.345	10.560	10.577	10.367	10.561	10.577
180	10.383	10.574	10.591	10.404	10.575	10.591

When assessing the convergence with respect to the size of the simulation supercell, two scenarios have to be distinguished: In a periodic setup, the gap for an isolated system can only be reproduced when the molecule at the centre of the cell is sufficiently far apart from its periodic images. When the Poisson equations of the replicas are decoupled (and the requirements of the used Poisson solver appropriately met[155]), the gap will converge with respect to the lowest **G**-vector components, which corresponds to increasingly longer-range components in real space as the supercell size increases. This may be especially important if the LUMO is very diffuse (we note that an unbound continuous state will only be appropriately described if the length of the simulation cell $l = \infty$). Hence, a change in cutoff value enhances the accuracy of the description by adding more rapidly oscillating, short-range components; the maximum 'diffuseness' allowed is essentially governed by the choice of l for the simulation cell.

Table 5.1 shows the gap obtained using the CAM-B3LYP xc functional for an isolated system contained in varying sizes of the simulation supercell using hard GTH pseudopotentials. Values for a fully periodic system are also given. The corresponding values for softer MT pseudopotentials are tabulated in the SI.

The gaps show convergence at 150Ry or all systems. For a small cubic simulation supercell (l = 10Å), choosing a lower cutoff value of 100Ry introduces a substantial error of 0.23eV. A notable error is still present at 100Ry even for the two larger simulation cells, but it becomes less relevant for practical purposes, since the maximal deviation of < 0.1eV lies below the typical accuracy of the functional itself. Whereas the values in the smallest of the supercells still have an error of about 0.1eV at a cutoff of 120Ry, the corresponding values have converged in the 20Å and 30Å supercells, with errors being lower than 0.05eV, and full convergence is reached at 150Ry for all of the three supercells considered.

The convergence behaviour is analogous to the one observed for GGA or standard hybrid functionals once the simulation cell is of sufficient size: Changes in the gap are still substantial when increasing the length l of the cubic simulation cell from 10 Å to 20 Å, with changes in the converged gap of 0.2eV. The change in gap is within the usual numerical tolerance (< 0.05eV) for a further extension to 30 Å, emphasising again the importance of an appropriately large cell for the correct description of the system's LUMO.

The results for the fully periodic system (where the Poisson equations are not decoupled) show the same trend, with the only relevant difference with respect to the isolated system occurring in the 10 Å box. Still, these differences are lower than those observed when enlarging the supercell. Given the trends observed for the isolated system, this is most likely attributed to spurious interactions between periodic images at this intermolecular distance. The influence of these interactions on the gap supports the view that the requests on the Tuckerman-Martyna Poisson solver are not yet met either, since the simulation cell must span at least twice the spatial extent of the charge density. The strong changes in gaps when increasing the cutoff within the small simulation cell is hence due to an insufficient cell size for both the isolated and periodic system, resulting in an incorrect description of the electron density.

The same considerations hold for the gaps obtained with the softer MT pseudopotentials (*cf.* the SI), albeit convergence is achieved at lower cutoff values. The maximum deviation in the converged gaps with respect to GTH pseudisation is very small, $\Delta\Delta\epsilon = 0.02$eV, illustrating that the influence of the pseudisation of the cusp on the gaps is negligibly low.

These results confirm that the use of the CAM in a plane wave/pseudopotential formalism does not introduce any convergence issues or additional restrictions, given that the size of the simulation supercell is chosen sufficiently large in order to accommodate the whole spectrum of **G**-vectors that are required to span the range of the μ-dependent switching function.

5.5.1$_{\mathrm{B}}$ Comparison with Atom-Centred Basis Sets

Plane waves inherently contain diffuse components even at a low cutoff energy, and the character of the diffuse functions is restricted only by the size of the supercell, and convergence is reached with respect to high-frequency components needed to describe the region around the pseudised core. In turn, the number of short-range components and the description of the orbitals around the nuclei can be systematically improved by using pseudopotentials of increasing hardness (which requires a simultaneous increase in the cutoff value). The complete basis set limit can therefore be systematically reached (once completeness holds with respect to the adopted pseudopotential).

The situation presents itself substantially different in atom-centred basis sets. In a Gaussian basis, the orbitals around the nuclei are well described, but for an accurate description of most molecular properties, the compact basis usually has to be enhanced by augmentation with long-range functions. When the GKS orbitals are expanded in Gaussians, a single set of diffuse functions is often sufficient for routine applications. (A more realistic description of the cusp and a correct decay of the basis at long range can be obtained by resorting to a Slater-type basis.)

A comparison between plane wave and atom-centred basis set calculations can therefore reveal the influence of the longest-range components (described well within plane waves) and the explicit description of the orbitals around the nuclear cusp (reproduced well using Gaussian functions). In the following, we will compare the HOMO-LUMO gaps of the preceding section with the corresponding results obtained from various Gaussian basis sets of increasing accuracy. Additional tests were performed using a Slater-type basis in order to obtain a systematic analysis with respect to the cusp condition. The results are shown in Table 5.2.

When increasing the size of the basis from double to hextuple zeta, changes in the gap are considerable for the non-augmented basis sets, spanning a range of a total of 0.9eV. Once a single diffuse function is included, the gaps become much more uniform, with a difference between aug-cc-pVDZ and aug-cc-pVTZ of only 25 meV. When increasing zeta to $\zeta = 6$, the gap fluctuates within a negligible range of 3 meV. Adding more diffuse functions never changes the gap by more than 5 meV. The aug-cc-pVDZ basis can therefore be considered sufficiently accurate when calculating HOMO-LUMO gaps with CAM-B3LYP.

The trend is very similar for a Slater-type basis, where we have only included single-zeta values for comparison. As for the Gaussian basis set, the omission of diffuse functions leads to an insufficient description of the gap. However, as soon as a single set of diffuse functions is included, the gap again converges rapidly. The change from augmented double to triple zeta is only about 8meV. Even an augmented single zeta basis seems to yield a surprisingly accurate gap, with the difference being only 0.08eV with respect to the augmented triple zeta basis.

While the difference between the converged gaps in plane waves and the

Table 5.2: HOMO-LUMO gap of a water molecule calculated using atom-centred basis sets augmented with a varying number of diffuse functions (aug. fcts).

Aug. fcts	0	1	2	3
Gaussian basis				
cc-pVDZ	9.832	10.679	10.667	10.666
cc-pVTZ	10.387	10.702	10.696	10.695
cc-pVQZ	10.576	10.705	10.670	10.700
cc-pV5Z	10.691	10.705	10.700	10.700
cc-pV6Z	10.714	10.704	10.700	10.700
Slater basis				
SZ	17.300	10.711		
DZ	12.318	10.631		
TZ	11.310	10.639		

corresponding values in a Gaussian basis is about 0.1eV, it is only 0.03eV when compared to a Slater type basis (augmented triple zeta *vs.* GTH/180 Ry/30 Å cell). Even though the orbitals are pseudised around the core, the plane wave/pseudopotential approach yields results which are virtually indistinguishable from all-electron calculations with an atom-centred basis. The slightly larger deviation with respect to the Gaussian basis may be attributed to differences in the long-range decay and the description of the cusp, but they still lie well within what is usually deemed chemical accuracy.

5.5.2 Excitation Energies

The most frequent use of Coulomb-attenuated functionals is the description of excited states (which is influenced by the accuracy of the GKS eigenvalues examined in the previous section through the linear response equations). We therefore conclude the assessment of our implementation of the CAM in plane waves by comparing the results to excited-state energies obtained using Gaussian bases. In line with the trends observed in the preceding section, only the cc-pVDZ basis was considered, which was augmented with a different number of diffuse functions. The results are depicted in Table 5.3.

The excitation energies were computed for the set of molecules introduced above, which apart from DMABN all contain transitions with charge-transfer or Rydberg character. Since the plane wave/pseudopotential implementation of LR-TDDFT in the CPMD code is limited to the Tamm-Dancoff approximation (TDA), the electronic spectra in a Gaussian basis were obtained within the same approximation for the sake of comparison. While it has been found that high-overlap singlet transitions can be substantially affected by the use of the TDA, it has also been reported that the results obtained from the TDA compare more favourably to high-level reference values than those obtained from full TDDFT.[191] This effect can be understood in terms of the triplet stability measure, the corresponding values for a subset of the molecules considered here have been reported and discussed in reference 191.

For the dipeptide, DMABN and 7-azaindole, the results of the singly-augmented Gaussian basis sets already exhibit a negligible difference to those obtained with the plane wave/pseudopotential approach. Neither the character nor the energetics of the transitions do change when using a doubly augmented Gaussian basis. Both Gaussians and plane waves yield the same ordering of states and differences in the excitation energies, which are smaller than 0.05eV, *i.e.* they lie within a range that we have previously considered as converged. The small deviations may be attributed to both the pseudisation of the orbitals in plane waves, as well as the limited spatial extent of a localised atom-centred basis, along with its predefined decay properties.

The situation is different in naphthalene and *p*-nitroaniline. In naphthalene, the S_1 state dominated by a HOMO-1 → LUMO transition is only predicted using plane waves or the singly augmented atom-centred Gaussian basis. However, the energetics of the S_2 state (mainly HOMO → LUMO) is consistent between the singly and doubly augmented Gaussian basis. The excitation energies agree well between our plane wave implementation and aug-cc-pVDZ.

The basis-set dependency issue becomes more involved in the case of *p*-nitroaniline. Whereas the S_2 $\pi \rightarrow \pi^*$ transition is predicted in all cases, the S_3 $n \rightarrow \pi^*$ state at 4.68eV only appears when one single diffuse function is used. Once a further set of diffuse functions is included, the S_3 $n \rightarrow \pi^*$ transition disappears again. It can therefore be concluded that in the case of *p*-nitroaniline, the description of the S_3 state using CAM-B3LYP is particularly sensitive with respect to the choice of basis. Since the S_3 transition disappears once the basis is enlarged, this indicates that the (quantitatively correct) prediction using aug-cc-pVDZ is a mere artefact, and that CAM-B3LYP is not properly able to describe the excitations in the limit of a complete basis. This view is further supported by the results obtained in plane waves, where the S_3 state is absent: The ordering of the states obtained at a cutoff of 120Ry coincides with the one obtained using a doubly augmented Gaussian basis. It has to be noted that the S_1 state has an oscillator strength $f = 0$, and that it is not predicted when using plane waves.

Overall, the results obtained using a plane wave pseudopotential ap-

Table 5.3: Excitation energies of the first two to four excited singlet states obtained for various systems and basis sets using LR-TDDFT and the CAM-B3LYP xc functional within the Tamm-Dancoff approximation. Values for a customised CAM-O3LYP functional are indicated, too. Literature reference values from high-level wavefunction calculations are given where available. In *p*-nitroaniline, the first forbidden transition is not predicted in a plane-wave basis when using CAM-B3LYP; its oscillator strength is $f = 0$ in the other calculations. The character of the transition is indicated whenever it is not consistent within different methods. Reference values for naphthalene, DMABN and the dipeptide correspond to those adopted by Peach *et al.*[140]

	CAM-B3LYP				CAM-O3LYP	
State	cc-pVDZ	aug-cc-pVDZ	d-aug-cc-pVDZ	MT/80 Ry	MT/80 Ry	Ref.
Naphthalene[192]						
S_1	–	4.66	–	4.65	4.63	4.46
S_2	4.73	4.86	4.86	4.79	4.88	4.88
DMABN[193]						
S_1	4.91	4.76	4.76	4.77	4.68	4.25
S_2	5.29	4.98	4.96	4.95	4.81	4.56
7−**azaindole**[194]						
S_1	5.19 ^1L_a	5.06 ^1L_a	5.05 ^1L_a	5.03 ^1L_a	5.05 ^1L_b	4.22 ^1L_b
S_2	5.24 ^1L_b	5.08 ^1L_b	5.08 ^1L_b	5.07 ^1L_b	5.09 ^1L_a	4.49 ^1L_a
S_3	5.36	5.32	5.32	5.33	5.35	5.27
Dipeptide[173]						
S_1	5.67	5.67	5.67	5.72	5.71	5.62
S_2	5.92	5.89	5.88	5.93	5.91	5.79
S_3	7.09	6.24	6.24	6.20	6.16	7.18
S_4	7.33	6.55	6.51	6.44	6.38	8.07
***p*-nitroaniline**[195]						
S_1	4.02	4.00	4.00	–	4.19	4.12
S_2	4.57	4.53	4.53	4.54	4.59	4.66
S_3	–	4.57	–	–	4.69	4.68
S_4	4.79	4.87	4.87	4.87	4.92	4.75

proach are in excellent agreement with the ones obtained with all-electron calculations in a Gaussian basis set. If any of the orbitals included in the transitions of interest is highly diffuse, plane waves fare better than a singly augmented Gaussian basis. There is no indication that the presence (or absence) of pseudisation has any (relevant) influence on the spectra, with the remaining differences between the doubly augmented Gaussian basis and the plane waves being vanishingly small. Although the excitation energies usually do not change considerably when approaching the diffuse limit, the character and number of states may, which can be important for applications such as excited-state dynamics. The observations made in *p*-nitroaniline further stress the importance of a sufficiently large basis, since seemingly correct predictions may be an artefact due to an incomplete basis. Only the use of a very diffuse Gaussian basis set or plane waves reveal that CAM-B3LYP does not properly predict one of the transitions. Given that most standard applications of Coulomb-attenuated functionals use only a single set of augmentation functions, plane waves hence offer the advantage of converging much more rapidly towards the basis set limit, thanks to their inherently diffuse character. This is further illustrated by the haphazard description of the S_1 state in naphthalene, which is easily recovered in plane waves.

5.5.3 Performance of New Customised *'à-la-carte'* Coulomb-Attenuated Functionals

Our generalised implementation of the Coulomb-attenuation method renders adjustments to the established xc functionals straightforward, for instance by using different range-separation parameters α, β, or by changing μ (commonly referred to as γ-tuning). The flexible form of K^x_σ also opens the distinct possibility of assembling new long-range corrected functionals based on simple physico-chemical considerations.

CAM-B3LYP has become a valuable tool in the calculation of excited states, as illustrated by the substantial improvements of the excitation energies with respect to a simple GGA published in the aforementioned studies. Still, the values in Table 5.3 reveal that not all transitions may be accurately captured quantitatively or qualitatively. The basis-set sensitivity of the excitations in *p*-nitroaniline are more of a practical issue, since reliable values may still be recovered in an accurately large basis, although at a substantial computational cost. However, the wrong ordering of states in 7-azaindole is one example that can only be overcome by resorting to generally less widely applicable xc functionals such as LC-BLYP, with a similar situation occurring in the β-dipeptide. The orbital localization problems in the AT base pair represent a further challenge. Such problems are often attributed to stem from an imbalance between short- and long-range exchange. The usual remedy to this weakness is the use of LC-BLYP, which fares worse for the systems for which CAM-B3LYP excels, but offers a more reasonable description of compounds where the ordering of CT or Rydberg states proves to be inaccurate. In the following, we have attempted to construct a functional that is sufficiently accurate for both systems where CAM-B3LYP is of satisfactory accuracy, as well as for the few charge-transfer systems where the functional has its known weaknesses.

A modified version of CAM-B3LYP with 80% exact exchange at long range was reported to yield eigenvalue differences closer to LC-BLYP, but performed worse for other properties where the 'standard' CAM-B3LYP yields accurate results.[38] LC-BLYP includes no exact exchange at shortest range, whereas CAM-B3LYP still includes 19%, a value close to the 20% used in standard B3LYP. A suitable compromise between LC-BLYP and CAM-B3LYP could therefore lie in attenuating an existing hybrid functional with more GGA exchange at short range, and less at long range. Following the CAM-B3LYP approach, the short-range contribution of the exact exchange should lie close to the value used in the conventional hybrid to ensure proper balance at short-range.

Handy's OPTX[196] functional in conjecture with Lee-Yang-Parr correlation[197] has on several occasions been shown to be superior to Becke's 1988 exchange functional,[198–201] and hybrids including OPTX such as O3LYP[196] include a lower percentage of exact exchange than the famous B3LYP while retaining comparable accuracy. We therefore assumed a Coulomb-attenuated version of O3LYP with 80% exact exchange at long range (as in the CAM-B3LYP declination in reference 38) and only 11% at short range (as in the O3LYP-hybrid) to offer the same benefits as LC-BLYP or CAM-B3LYP with 80% exact exchange at long range, but with an improved description of the short-range region due to the inclusion of the more accurate OPTX.

The performance of the CAM-O3LYP functional on the test set used in the previous chapter, including the problematic 7-azaindole, is summarised in Table 5.3. In the dipeptide, much alike the retinal protonated Schiff base, CAM-O3LYP yields virtually indistinguishable results from CAM-B3LYP, with a maximum deviation in S_4 of 0.06eV. The S_1 and S_2 states are therefore accurately described, whereas the deviations for S_3 and S_4 remain too large for practical applications in both of the functionals. We should note that the good agreement between CAM-B3LYP and reference values reported by Peach *et al.* were based on values obtained in a basis without diffuse functions and without the TDA; the higher energies of the CT states we re-

Table 5.4: Comparison of excitation energies between CAM-B3LYP and CAM-O3LYP for a structure of the retinal protonated Schiff base obtained from a molecular dynamics snapshot. CAM-B3LYP is known to give good agreement with respect to high-level wavefunction methods for this molecule.[166, 180]

State	CAM-B3LYP	CAM-O3LYP
Retinal protonated Schiff base		
S_1	2.66	2.63
S_2	4.13	4.25
S_3	4.97	4.95
S_4	5.12	5.08

port for the cc-pVDZ basis are in line with their results. With CAM-O3LYP, the state ordering in 7-azaindole is now correctly reproduced, although the absolute errors are still considerable, with the energetic difference between the ^1L_a and ^1L_b states being too low. (This is also observed for CAM-B3LYP and may be attributed to the use of the TDA, since in general, the energy difference between the two states becomes larger if the TDA is not employed.) In *p*-nitroaniline, the spectrum substantially improves with the use of CAM-O3LYP, due to an improved description of the S_3 state, which is now predicted even when approaching the basis set limit. For systems in which a classical hybrid functional is preferable to CAM-B3LYP, such as DMABN, the error due to CAM-O3LYP is comparable to the one of CAM-B3LYP, even slightly smaller in the case of the S_2 state.

Table 5.4 shows the excitation energies for a structure of the retinal protonated Schiff base from a molecular dynamics snapshot calculated with both CAM-B3LYP and CAM-O3LYP. Since CAM-B3LYP has been shown to yield very accurate excitation energies in this system, this can serve as an additional benchmark for the accuracy of our new functional. Indeed, CAM-O3LYP yields virtually indistinguishable excitation energies.

The remaining ‘problematic cases’, the AT base pair and the β-dipeptide are presented in table 5.5, where excitation energies obtained from CAM-O3LYP and CAM-B3LYP are compared to reference values.

For one of the two notorious cases, CAM-O3LYP outperforms CAM-B3LYP on a qualitative level: The spectrum of the AT base pair is qualitatively correctly reproduced, with the HOMO lying on adenine rather than thymine, in contrast to CAM-B3LYP results. Overall, both functionals capture the energetics of all but the S_1 state accurately, but CAM-O3LYP gives a correct theoretical description of the orbital localization. Only in the β-dipeptide there is no improvement with respect to CAM-B3LYP, but the

Table 5.5: Comparison of excitation energies between CAM-B3LYP, CAM-O3LYP and reference values for some typical usage cases of range-separated functionals known to be problematic when treated with CAM-B3LYP. The character of the transition is given where it differs with respect to the reference. The reference values for the β-dipeptide correspond to those adopted by Peach *et al.*[140]; remaining values for molecules included therein and not explicitly mentioned so far are tabulated in the Supporting Information.

State	CAM-B3LYP	CAM-O3LYP	Ref.
AT base pair[202]			
S_1	5.25	5.24	4.94
S_2	5.26	5.29	5.21
S_3	5.36	5.40	5.40
S_4	5.38	5.45	5.47
β-dipeptide[173]			
S_1	5.69 $n_1 \rightarrow \pi_1^*$	5.66 $n_1 \rightarrow \pi_1^*$	5.10 $n_2 \rightarrow \pi_2^*$
S_2	5.76 $n_2 \rightarrow \pi_2^*$	5.74 $n_2 \rightarrow \pi_2^*$	5.40 $n_1 \rightarrow \pi_1^*$
S_3	6.06 $\pi_1 \rightarrow \pi_2^*$	6.00 $\pi_1 \rightarrow \pi_2^*$	7.99 $\pi_1 \rightarrow \pi_2^*$
S_4	6.14 $n_1 \rightarrow \pi_2^*$	6.01 $n_1 \rightarrow \pi_2^*$	9.13 $n_1 \rightarrow \pi_2^*$

qualitative and quantitative behaviour is again very similar.

Overall, CAM-O3LYP shows identical or superior performance to CAM-B3LYP for all of the excitations; it appears to be more versatile in the description of systems that require a larger percentage of exact exchange at long range, where it is able to remedy some of the pitfalls encountered with CAM-B3LYP. These are systems where LC-BLYP has typically been used so far. While the mean absolute error (MAE) over all the molecules considered here is about MAE = 0.65eV for CAM-B3LYP, it improves for CAM-O3LYP, where MAE = 0.55eV.

This demonstrates that based on simple considerations, a new Coulomb-attenuated functional can be constructed. CAM-O3LYP predicts qualitatively correct state ordering where CAM-B3LYP fails, and in these systems yields excitation energies similar to the latter. With CAM-O3LYP, it is therefore possible to cover a larger range of systems than with CAM-B3LYP with slightly improved accuracy. This is further reflected by the MAE of both functionals, which is about 0.1eV lower for the CAM-O3LYP presented here. For other delicate systems and specific problems, our implementation offers the possibility of adapting existing xc functionals at hand, or to assemble entirely new Coulomb-attenuated functionals.

5.6 Conclusions

We have presented a new, efficient and fully flexible implementation of Coulomb-attenuated functionals in the plane wave/pseudopotential code CPMD which allows for a customised composition of exchange-correlation functionals. On the base of a comprehensive test suite, we could demonstrate that the results obtained within the plane wave/pseudopotential framework do not significantly deviate with respect to results obtained in all-electron calculations with Gaussian bases. The results indicate that the complete basis set limit is more easily reached in plane waves, and that the pseudisation of the nuclear cusp is of no relevant influence on HOMO-LUMO gaps or excitation energies. We have also shown that based on the same considerations that led to the construction of CAM-B3LYP, a new xc functional CAM-O3LYP can be constructed, which shows improved performance over CAM-B3LYP in systems where the latter fails, and yields comparable accuracy for systems where CAM-B3LYP typically fares well. This demonstrates that the flexibility of *'à-la-carte'* combinations of xc functionals can help in obtaining excitations of higher accuracy over a larger range of systems.

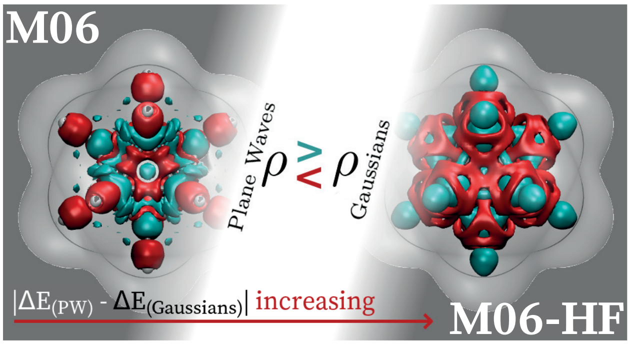
5.7 Acknowledgements

The authors thank J. Magnus Haugaard Olsen for providing a molecular dynamics snapshot of the retinal protonated Schiff base, Valéry Weber for comments on the implementation of the new xc driver, and they thank the CADMOS project for providing computing time on the BG/Q. UR gratefully acknowledges financial support from the Swiss National Science Foundation Grant No. 200020-146645 and the NCCR MUST.

CHAPTER 6

Shedding Light on the Basis Set Dependence of the Minnesota Functionals: Differences Between Plane Waves, Slater Functions and Gaussians

Chapter 6 is a pre-print version of an article submitted to:


M06
Plane Waves
$\rho \gtrless \rho$
Gaussians
$|\Delta E_{(PW)} - \Delta E_{(Gaussians)}|$ increasing
M06-HF

Bircher, Martin P.; Lopez-Tarífa, P.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **submitted** [i]

---

[i] An initial closed-shell-only implementation of the Minnesota Functional Module (MFM) in the old xc driver of the CPMD code is due to P.L. M.P.B. performed the definitive implementation in the new xc driver for both spin-restricted and -unrestricted systems, including the appropriate adaptation of definitions and the tests and data analysis presented in this chapter.

THE Minnesota family of xc functionals are among the most popular, accurate and abundantly used functionals available to date. However, their use in plane-wave based *first principles* MD has been limited by their sparse availability. Here, we present an implementation of the M05, M06 and M11 families of xc functionals within a plane wave/pseudopotential framework allowing for a comprehensive analysis of their basis set dependence. While it has been reported that in Gaussian bases, some members of the Minnesota family only converge slowly to the basis set limit,[204] we show that converged energies can be conveniently obtained from plane waves if sufficiently dense integration meshes are used. Based on the HC7/11 database, we assess the influence of basis set type on the calculation of reaction enthalpies and show that complete basis set values obtained in plane waves may occasionally differ notably from their atom-centred counterparts. We provide an analysis of the origin of these differences and discuss implications on practical usage.

6.1 Introduction

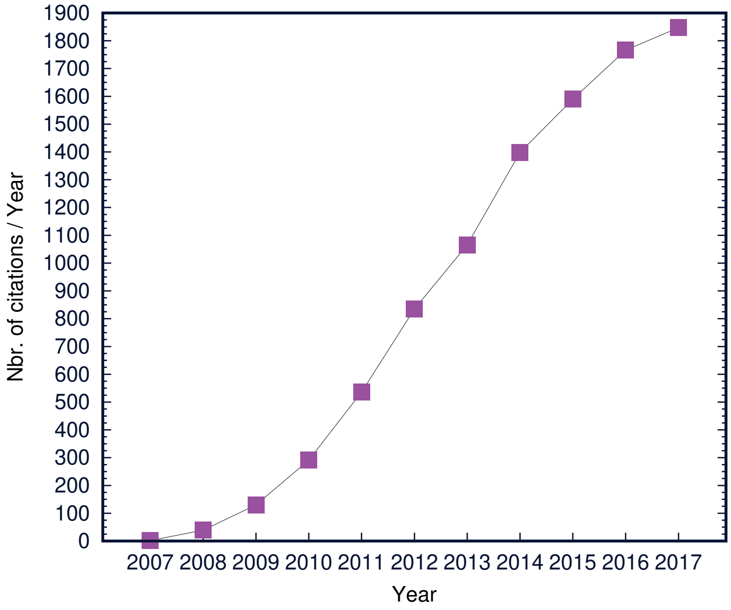
Density Functional Theory (DFT)[14] in its Kohn-Sham (KS) formulation[15] is one of the fundamental pillars of modern-day computational chemistry. Large systems of several hundreds of atoms can presently be treated. This is vital for the description of condensed matter systems with *first principles* Molecular Dynamics (MD) simulations, in which thermodynamic properties can be obtained as time-averages at finite temperature. Even larger systems can be routinely treated using mixed quantum mechanics/molecular mechanics (QM/MM)[72] approaches and with computational power ever increasing, the time scales that can be sampled are growing continuously, thus decreasing the statistical error. Powerful enhanced sampling methods have helped to reduce the time scales that have to be simulated, and at present, the error due to the sampling can become smaller than the error of the underlying potential energy surface (PES). The accuracy of the underlying exchange-correlation (xc) functional therefore becomes an increasingly dominant factor. The reliability of the underlying PES will not only be influenced by the choice of functional, but also by the basis set used to expand the density. Some functionals may be particularly sensitive to the choice and size of basis.[161, 204, 205] A plane wave expansion offers the intrinsic advantage of convergence control *via* a single parameter, the cutoff energy E_{cut}. A plane wave description is therefore ideally suited whenever a highly flexible, delocalised basis set is needed in order to obtain converged PES.

If KS-DFT has become abundantly used, then this is not at least due to the vast effort, persistence and creativity that have been invested in the development of approximate forms for the unknown xc functional. This fundamental problem can be approached in many different ways.[83,115] In some philosophies, functionals that contain as few free parameters as possible are generally preferred; the free parameters of such functionals are obtained from physical constraints and limits,[103] In another school of thought, a design that is directly fit to thermochemical data is considered nothing but an extension of basic physical constraints to the chemical world.[122,196] Such an approach may be taken even further by obtaining free parameters not from atomic data, but by resorting to larger databases of thermochemical and structural data.[43,46] Even if the foundations of the different approaches to the problem may not be the same, they have ultimately been guided by the very same goal of improved accuracy and reliability.

A very prominent example of the latter philosophy is given by the widely used family of the Minnesota functionals,[39–46] which have been developed by the Truhlar group. Many prominent functionals have been derived to offer one single functional that is as versatile as possible. Instead, the Minnesota philosophy aims to find the best possible performance for a given functional model and for a certain range of systems. Based on a given functional form, the free parameters are optimised according to different constraints: Certain parts of a test set may or may not be included in a given fit, and the functional form may or may not include an exact exchange contribution. Therefore, for every new generation of the Minnesota functionals, different variants have been developed, each with specific advantages in a certain regime, but sharing a basic functional form. This approach has proven to be highly successful for many chemical problems.[46,206] Even though the early members of the Minnesota family, the M05[39,40] and M06[41–43,207] group of functionals, have since been superseded by more accurate models, the M06 family is still vastly popular in computational chemistry, as shown in Fig. 6.1 by the rate of citations per year for the seminal M06 paper.

Along with the advent of empirically optimised functionals, several databases have been created which can be used both for the fitting of free parameters in functional development as well as for performance assessment.[46,83,208,209] The HC7/11 database (Fig. 6.2), for instance, gathers difficult hydrocarbon reaction enthalpies,[43,208–210] including a set of particularly challenging isodesmic reactions. Truhlar and coworkers have shown that all functionals of the M06 family (*i.e* M06-2X, M06-HF, M06-L and the parent M06 itself) outperform the other contemporary xc functionals, including the still very popular B3LYP[31,163]; and similar trends were found for other data sets.[46,206] In the case of the CT7/04 set of charge-transfer dimers,[208]

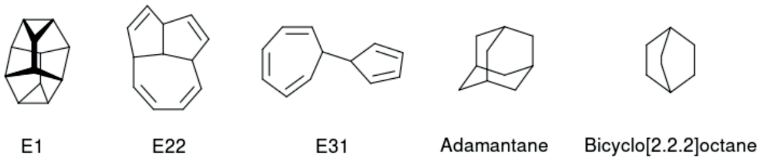
Figure 6.1: Citations per year for the seminal M06 paper[43] (source:webofknowledge). As of July 2018, the paper has accumulated a total of 10545 citations, making it the most cited of the Minnesota papers.



the performance of the Minnesota functionals was either superior (M06-2X, M06-HF, M05) or at least on par (M06) with the best non-Minnesota xc functionals considered in the study[43, 209] (B97-3).[211]

Despite such an encouraging performance, the underlying approach has also been met with scepticism.[115, 204] It has been argued that a fit that is not based on physical limits may deteriorate the formal qualities of a functional (however, all the Minnesota functionals are constrained to fulfil the UEG limit). Mardirossian *et al.*[204] have shown that in an atom centred basis of Gaussian functions, the energetics of some members of the Minnesota functional family converge remarkably slowly, and may at first appear not to converge at all. By analysing the inhomogeneity correction factors of the slowly converging functionals, they have shown that slow convergence correlates with either large correction factors for certain limits of the working function or with the occurrence of positive energy densities. The slow convergence of certain Minnesota functionals is also reflected in a high sensitivity towards basis-set superposition errors (BSSE).[204] Since the Minnesota func-

Figure 6.2: The HC7/11 set of difficult hydrocarbon reactions.[43]



1 E1 $\rightarrow$ E22

2 E1 $\rightarrow$ E31

3 $(CH_4)_4CC(CH_3)_3 \rightarrow n\text{-}C_8H_{18}$

4 $n\text{-}C_6H_{14} + 4CH_4 \rightarrow 5\ C_2H_6$

5 $n\text{-}C_8H_{18} + 6CH_4 \rightarrow 7\ C_2H_6$

6 Adamantane $\rightarrow 3\ C_2H_4 + 2\ C_2H_2$

7 Bicyclo[2.2.2]octane $\rightarrow 3\ C_2H_4 + C_2H_2$

tionals have been fit employing a specific basis set, the reference results can often only be obtained for the same (or a very similar) basis; while changing the 'balance' of the basis set - or even approaching the basis set limit - may also change the energetics by up to a few kcal mol^{-1} . However, in many practical applications, functional/basis set combinations are benchmarked against accurate reference data[46] and then used in that specific configuration. In these cases, a given functional/basis set combination may be found to be accurate, even though the basis set limit values might show larger deviations with respect to the reference data than the combination adopted. From a theoretical point of view, however, the possibility of reaching convergence systematically and assessing the performance of a functional with a fully converged basis set is certainly desirable, in particular concerning the transferability of the results.

The study by Mardirossian *et al.* was conducted by expanding the density in atom-centred Gaussian functions. Inspired by the concept of overlapping atomic orbitals, atom-centred functions - and in particular Gaussians - are probably the most prominent choice of a basis set. However, the electron density may as well be expanded in other functional forms that need not necessarily be localised in space. Delocalised bases such as plane waves offer particular advantages beyond a simple control of convergence: Neither do Pulay forces occur,[148] nor are Basis-set superposition effects of any concern. This comes at the price of having to pseudise[16] the effect of the core electrons; but the error due to this procedure has been shown to be negligible for the vast majority of chemically relevant properties of main group elements.[16,161]

These features make plane waves not only intrinsically suitable for *first principles* Molecular Dynamics, but also for the calculation of quantities that are sensitive to basis-set superposition or that require multiple long-range functions. In a plane wave basis, the convergence of some property is monitored by increasing the cutoff energy of the highest-frequency reciprocal space vector. This makes plane waves ideally suited for the assessment of xc functional convergence, since the basis can be systematically increased until the basis set limit is attained. These features enable systematic convergence studies of xc functionals that may be difficult to converge in atom-centred bases.

Despite their popularity, availability of the Minnesota functionals in plane wave codes is scarce, being mostly limited to the (semi-)local declinations of the functionals.[212,213] Here, we present a comprehensive numerical analysis of an implementation of the M05, M06, M08 and M11 members of the Minnesota family in a plane wave/pseudopotential framework. Our implementation in the CPMD code makes Minnesota functionals available for routine plane wave calculations without the computational overhead (up to 20%) due to external libraries.

The text is organised as follows: First, we give a short overview of the functional forms of the M05, M06, M08 and M11 families, followed by a brief description of the plane-wave specific details of our implementation. We then provide a comprehensive assessment of the basis set convergence in plane waves, which will be discussed at the example of the HF dimer. This member of the CT7/04 database had also been used by Mardirossian *et al.* in their convergence analysis in Gaussian bases,[83] making a direct comparison to their results possible. Particular emphasis will be put on the functionals that Mardirossian *et al.* have identified as slowly converging. We will then show that using standard ratios between density and orbital cutoff values, for a certain subset of the functionals considered here, no systematic convergence is reached in a plane wave basis and that it is possible to systematically resolve this convergence issue by increasing the ratio between density and wavefunction cutoff. By including this additional parameter in the convergence analysis, the energetics of all functionals can be analysed straightforwardly in a plane wave setup. Based on the possibility of obtaining converged energetics, we will then compare the influence of basis set type on the reaction enthalpies of hydrocarbons at the example of the HC7/11 database. Values will be compared between our plane wave implementation, two Gaussian basis sets commonly employed with the Minnesota family (aug-cc-pVTZ[151] and 6-311+G(2df,2p)),[150] as well as two polarized Slater bases (TZ2P and ATZ2P).[184] Slater functions exhibit an exact r^{-1} decay, which makes it possible to assess the effects of long-range decay and basis set delocalisation separately. This analysis will reveal that the range-separated functionals M11 and M11-L are particularly sensitive to the choice of basis and that converged energy differences obtained from plane waves or Slater functions may differ from their atom-centred equivalents, stressing the importance of a fully flexible basis. We will then see that results for M06-2X and M06-HF can deviate significantly from the values obtained in both Gaussian and Slater bases. We will show that these deviations can be tracked down by analysing the exchange energy density $\epsilon_x(\mathbf{r})$, illustrating that the analytical form of the M06 family can lead to particular electron density differences between plane waves and Gaussian functions. Finally, we will provide a short discussion of the implication of our findings on the run time of plane wave calculations using the Minnesota family, and suggest approaches to reduce the computational overhead.

6.2 Theory

6.2.1 The Minnesota Functionals

The oldest members of the Minnesota family are M05[39] and M05-2X.[40] The exchange part of the M05 family can be seen as a kinetic-energy dependent empirical extension of the PBE[103] exchange functional. Introducing the working variable:

$$f(\omega_\sigma) = \sum_{i=0}^{m} a_i \omega_\sigma^i, \tag{6.1}$$

where the $\{a_i\}$ are empirically determined weights and

$$\omega_\sigma(t_\sigma) = \frac{(t_\sigma - 1)}{(t_\sigma + 1)}, \tag{6.2}$$

$$t_\sigma(\rho_\sigma, \tau_\sigma) = \frac{3}{10}(6\pi^2)^{2/3}\frac{\rho_\sigma^{5/3}}{\tau_\sigma}, \tag{6.3}$$

where ρ_σ is the spin density and $\tau_\sigma = \sum_i^{occ} \left|\nabla\psi_i^{(\sigma)}\right|^2$ is the spin-dependent kinetic energy density of the (generalised) Kohn-Sham orbitals, $\{\psi_i^{(\sigma)}\}$, Truhlar and coworkers proposed to write:

$$E_x^{M05} = \mathrm{X}E_{HFX} + \sum_\sigma (1 - \mathrm{X}) \int \mathrm{d}\mathbf{r}\, F_x^{PBE}[\rho, \nabla\rho] f(\omega_\sigma). \tag{6.4}$$

Here, X denotes the contribution of exact exchange, E_{HFX}, and F_x^{PBE} is the PBE enhancement factor that depends on both the electron density ρ and the gradient $\nabla\rho$. The correlation functional is written as a sum over equal and opposite spin components using the Stoll ansatz:

$$E_c^{M05} = \int d\mathbf{r}\, e_{c,\alpha\beta}^{UEG} g_{\alpha\beta}(x_{\alpha\beta}, z_{\alpha\beta}) + \sum_\sigma \int d\mathbf{r}\, e_{c,\sigma\sigma}^{UEG} g_{\sigma\sigma}(x_\sigma, z_\sigma) D_\sigma(x_\sigma, z_\sigma), \tag{6.5}$$

where e_c^{UEG} is the correlation energy of the (spin-polarised) uniform electron gas and D_σ is the self-interaction correction factor:

$$D_\sigma(x_\sigma, z_\sigma) = 1 - \frac{x_\sigma^2}{4z_\sigma + C_F}. \tag{6.6}$$

D_σ is a function of the reduced gradient $x_\sigma = \rho_\sigma^{-4/3} |\nabla\rho_\sigma|$ and $z_\sigma = \frac{2\tau_\sigma}{\rho_\sigma^{5/3}} - C_F$ with the constant $C_F = \frac{3}{5}(6\pi^2)^{2/3}$. Finally, the $g_{\alpha\beta}$ and $g_{\sigma\sigma}$ are constructed from:

$$g_{\alpha\beta}(x_{\alpha\beta}, z_{\alpha\beta}) = \sum_{i=0}^{n} c_{\alpha\beta}^{(i)} \left(\frac{\gamma_{\alpha\beta}(x_\alpha^2 + x_\beta^2)}{1 + \gamma_{\alpha\beta}(x_\alpha^2 + x_\beta^2)} \right), \tag{6.7}$$

$$g_{\sigma\sigma}(x_\sigma, z_\sigma) = \sum_{i=0}^{n} c_{\sigma\sigma}^{(i)} \left(\frac{\gamma_{\sigma\sigma} x_\sigma^2}{1 + \gamma_{\sigma\sigma} x_\sigma^2} \right), \tag{6.8}$$

where $\gamma_{\sigma\sigma} = 0.060$ and $\gamma_{\alpha\beta} = 0.031$ are constants and the $\{c_{\sigma\sigma}\}$ and $\{c_{\alpha\beta}\}$ are free parameters. The free parameters of the exchange and correlation functional are obtained simultaneously. X and the number of expansion coefficients m, n are fixed arbitrarily and the $\{a_i\}$, $\{c_{\sigma\sigma}\}$ and $\{c_{\alpha\beta}\}$ are then determined by a fit to a database with the exception of $a_0 = 1$, which is kept fixed in order to recover the uniform electron gas (UEG) limit. Two fits with different values of X were performed, resulting in the M05 (X = 0.28) and M05-2X (X = 0.56) functionals.

Building on the initial success of the M05 family, the M06 functionals[41–43] constitute an extension of the M05 philosophy by adding additional degrees of freedom. This was achieved by mixing the M05 expression with a reparametrisation of the VS98[214] exchange correlation functional:

$$E_{xc}^{VS98} = E_x^{VS98} + E_c^{VS98}, \tag{6.9}$$

$$E_x^{VS98} = \sum_\sigma \int \mathrm{d}\mathbf{r}\, \rho_\sigma e_{x,\sigma}^{UEG} h_x(x_\sigma, z_\sigma), \tag{6.10}$$

$$E_c^{VS98} = \int \mathrm{d}\mathbf{r}\, e_{c,\alpha\beta}^{UEG} h_c(x_{\alpha\beta}, z_{\alpha\beta}) + \sum_\sigma \int \mathrm{d}\mathbf{r}\, e_{c,\sigma}^{UEG} h_c(x_\sigma, z_\sigma), \tag{6.11}$$

where $e_{x,\sigma}^{UEG}$ is the exchange energy density of a uniform electron gas and $x_{\alpha\beta}^2 \equiv x_\alpha^2 + x_\beta^2$ and $z_{\alpha\beta} \equiv z_\alpha + z_\beta$. $h(x, z)$ takes the general form:

$$h(x, z) = \frac{a}{\gamma(x, z)} + \frac{bx + cz}{\gamma(x, z)^2} + \frac{dx^2 + exz + fz^2}{\gamma(x, z)^3}, \tag{6.12}$$

with free parameters a to f. The functionals of the M06 family then take the form:

$$E_x^{M06} = \mathrm{X}E_{HFX} + \sum_\sigma \int \mathrm{d}\mathbf{r} \left[F_x^{PBE}[\rho, \nabla\rho] f(\omega_\sigma) + \rho_\sigma e_{x,\sigma}^{UEG} h_x(x_\sigma, z_\sigma)\right], \tag{6.13}$$

$$\begin{aligned} E_c^{M06} = &\int \mathrm{d}\mathbf{r}\, e_{c,\alpha\beta}^{UEG} \left[g_{\alpha\beta}(x_{\alpha\beta}, z_{\alpha\beta}) + h_{\alpha\beta}(x_{\alpha\beta}, z_{\alpha\beta})\right] \\ &+ \sum_\sigma \int \mathrm{d}\mathbf{r}\, e_{c,\sigma\sigma}^{UEG} \left[g_{\sigma\sigma}(x_\sigma, z_\sigma) D_\sigma(x_\sigma, z_\sigma) + h_{\sigma\sigma}(x_\sigma, z_\sigma)\right], \end{aligned} \tag{6.14}$$

where, contrary to the M05 family, no parameter is explicitly fixed. Instead, the UEG limit is enforced using appropriate constraints on the zeroth-order expansion coefficients. Different fitting databases were used to fit four different functions: A completely semi-local meta-GGA (X = 0), M06-L,[41] and three meta hybrid functionals: M06,[43] M06-2X[43] and M06-HF.[42] Of these functionals, M06-2X was fitted to a smaller database. Therefore, $d = 0 \rightarrow h_x = 0$ was imposed for E_x^{VS98}, *i.e* the functional has an exchange part that reduces to the M05 expression. A detailed recommendation concerning typical usage cases for every of the M06 family members is given in the seminal paper by Zhao *et al.*[43] as well as in Ref. 206 ; here, it should not go unmentioned that M06-L has proven to be the best semi-local functional for many systems, whereas M06-2X and M06 have probably been the most successful meta-hybrid functional of the M05/M06 group.[204]

Inspired by the encouraging results obtained with the M06 family - and in particular with M06-2X - Truhlar and coworkers investigated possible improvements by including even more flexible functional forms,[207] while retaining a large exact exchange contribution as in M05-2X and M06-2X. Contrary to previous work, the next generation of exchange functionals was defined using the spin-unpolarised system with density $\rho = \rho_\alpha + \rho_\beta$:

$$\begin{aligned} E_x^{M08} &= \mathrm{X} E_x^{HFX} \\ &+(1-\mathrm{X}) \int \mathrm{d}\mathbf{r}\, \rho e_x^{LDA} \left[f_1(\omega) F_x^{PBE}[\rho, \nabla\rho] + f_2(\omega) F_x^{RPBE}[\rho, \nabla\rho] \right], \end{aligned} \tag{6.15}$$

where F_x^{PBE} and F_x^{RPBE} are the PBE and RPBE enhancement factors, respectively, and the weights $f_i(\omega)$ are defined as in eq. 6.1. A generalisation to spin is trivial using the spin-scaling relations. The Stoll ansatz for the correlation functional is abandoned in favour of the expression:

$$E_c^{M08} = \int \mathrm{d}\mathbf{r}\, \rho e_C^{LSDA}(\rho, \zeta) f_3(\omega) + \int \mathrm{d}\mathbf{r}\, \rho H^{PBE}(\rho, \nabla\rho, \zeta, t) f_4(\omega), \tag{6.16}$$

where $\zeta = (\rho_\alpha - \rho_\beta)/\rho$ is the spin polarisation, e_C^{LSDA} is the correlation energy per unit density in the Perdew-Wang form[120] and H^{PBE} is the PBE correction to the correlation energy. Fitting these forms by only imposing the UEG limit yields the functional M08-HX. Furthermore, by imposing that the gradient expansion to second order be obeyed for both exchange and correlation, M08-SO was parametrised. While the former outperformed its predecessors M05 and M06 on many databases (and related properties), the latter was found to be particularly accurate in the prediction of main-group thermochemistry.

The next generation of functionals following the Minnesota philosophy were published in 2011[44,45] and made use of the promising results[140] obtained by using range-separated xc functionals in either full long-range correction (LC)[141,143] schemes or the Coulomb attenuation method (CAM).[38] The M11 family comprises two members; the parent range-separated hybrid M11[44] as well as a completely local - but still range-separated - meta-GGA M11-L.[45] Based on the M08 exchange functional, Peverati *et al.* applied the CAM to the M08 exchange energy:

$$E_x^{M11}[\rho] = E_{x,LR}^{HFX}[\rho] + E_{x,SR}^{loc}[\rho]. \tag{6.17}$$

For M11, the long-range exchange contribution is given by a Coulomb-attenuated exact exchange operator:

$$E_{x,LR}^{HFX}[\rho] = \frac{1}{2}\sum_i\sum_j \iint \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\,\psi_i^*(\mathbf{r})\psi_j^*(\mathbf{r}')\frac{\alpha+\beta\,\mathrm{erf}(\mu\,|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{r})\psi_i(\mathbf{r}'), \tag{6.18}$$

where α, β govern the contribution of exact exchange at short range (α) and long range ($\alpha+\beta$) respectively, and μ determines how rapidly the exact exchange contribution increases between short and long range. The (semi-)local contribution to exchange, $E_{x,SR}^{loc}$, is obtained by straightforwardly replacing e_x^{LDA} by a suitable range-separated integrand, e_x^{CAM}:

$$E_{x,SR}^{loc} = \int \mathrm{d}\mathbf{r}\,\rho e_x^{SR}\left[f_1^{SR}(\omega)F_x^{PBE}[\rho,\nabla\rho] + f_2^{SR}(\omega)F_x^{RPBE}[\rho,\nabla\rho]\right], \tag{6.19}$$

$$e_x^{SR} = -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\rho^{1/3}G(\rho,\alpha,\beta,\mu), \tag{6.20}$$

where the function $G(\rho,\alpha,\beta,\mu)$ is defined as:

$$\begin{aligned} G(\rho,\alpha,\beta,\mu) = 1-\alpha-\beta\Bigg\{\frac{8}{3}a\Bigg[\sqrt{\pi}\mathrm{erf}\left(\frac{1}{2a}\right) - 3a \\ +\ 4a^3 + (2a-4a^3)\exp\left(-\frac{1}{4a^2}\right)\Bigg]\Bigg\}, \end{aligned} \tag{6.21}$$

$$a = \frac{\mu}{2(6\pi^2\rho)^{1/3}}. \tag{6.22}$$

The correlation energy functional takes the same form as in eq. 6.16. The free parameters were then fit imposing the UEG limit, the second-order gradient expansion as well as two constraints concerning extrema of τ, which had previously been proposed by Becke.[129] In a completely local declination, M11-L, the long-range exchange component $E_{x,LR}^{HFX}$ is replaced by pure (local) DFT exchange, but using a different set of parameters f^{LR}:

$$E_{x,LR}^{loc} = \int \mathrm{d}\mathbf{r}\,\rho e_x^{LR}\left[f_1^{LR}(\omega)F_x^{PBE}[\rho,\nabla\rho] + f_2^{LR}(\omega)F_x^{RPBE}[\rho,\nabla\rho]\right], \tag{6.23}$$

$$e_x^{LR} = -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\rho^{1/3}\left[1-G(\rho,\alpha,\beta,\mu)\right]. \tag{6.24}$$

Like the M08 and M11 hybrids, M11-L has been fitted under appropriate constraints, such that the resulting parametrisation is correct to second order in both exchange and correlation. Due to the absence of single-determinantal exchange, it has been found to be particularly appropriate for multireference systems, with an average performance superior to M06-L.

6.2.2 Expanding the Electron Density in Plane Waves

At the Γ-point, the (real, generalised) Kohn-Sham orbitals can be expanded in a plane wave basis of the form[16]:

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{G}=0}^{\mathbf{G}_{max}} \psi_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{6.25}$$

where $\mathbf{r}, \mathbf{G}$ are real and reciprocal space vectors, respectively, and the $\{\psi_i(\mathbf{G})\}$ are the (Fourier) expansion coefficients of an orbital indexed by i. The length of the expansion is given by the highest-frequency $\mathbf{G}$-vector $\mathbf{G}_{max}$ and is usually indicated with respect to its energy, the cutoff energy E_{cut}. In a discretised Cartesian basis, the $\psi_i(\mathbf{r})$ and $\psi_i(\mathbf{G})$ are mutual Fourier transforms. It is therefore straightforward to switch between representations at an $\mathcal{O}(N(\log N))$ scaling using Fast Fourier Transforms (FFT). This is of particular advantage for the computation of the Coulomb potential: Since the Coulomb operator is diagonal in reciprocal space, $\Phi(\mathbf{G}) = \frac{4\pi}{\mathbf{G}^2}$, the corresponding potential is easily obtained from the reciprocal space density $\rho(\mathbf{G})$. The density itself is most conveniently constructed in real space:

$$\rho(\mathbf{r}) = \sum_i f_i \left|\psi_i(\mathbf{r})\right|^2, \tag{6.26}$$

where f_i are occupation numbers and $\rho(\mathbf{G})$ is obtained via a FFT. However, due to the square in eq. 6.26, the spectrum of the function ρ will extend to higher-frequency $\mathbf{G}$-vectors than the initial ψ. This requires for a second cutoff energy to be introduced, E^{ρ}_{cut}, with a value of 4 being analytically sufficient to guarantee correspondence between ψ and ρ. The ratio between E^{ρ}_{cut} and E_{cut} is commonly referred to as the dual:

$$\xi = \frac{E^{\rho}_{cut}}{E_{cut}}. \tag{6.27}$$

Choosing a value of ξ larger than the default value of r results in real-space representations of ρ which is Fourier-interpolated with respect to the default grid. This makes it possible to arbitrarily increase the resolution of the mesh without introducing higher-frequency components in the description of ψ, therefore minimising numerical noise.

6.2.3 Range-Separated and Screened Exchange in Plane Waves

In plane waves, the exact exchange of a generalised Kohn-Sham determinant is usually calculated in reciprocal space from[157]:

$$E_x^0[\rho] = -\frac{1}{2}\sum_i^{N_b}\sum_j^{N_b}\sum_{\mathbf{G}}^{\mathbf{G}_{\max}} \Phi(\mathbf{G})\left|\rho_{ij}(\mathbf{G})\right|^2, \tag{6.28}$$

where $\Phi(\mathbf{G})$ denotes the reciprocal space (*i.e.* Fourier series) representation of a (generic) Coulomb operator $\hat{\mathrm{W}}$ and $\rho_{ij}(\mathbf{G}) = \mathcal{F}\left[\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})\right]$ are the Fourier transforms of pair densities constructed from N_b occupied generalised Kohn-Sham orbitals. In a fully periodic setup and within a discrete representation of $\mathbf{G}$, $E_x^0[\rho]$ exhibits an (integrable) divergence at $\mathbf{G} = 0$. $\Phi(\mathbf{G})$ therefore needs to be appropriately modified.[158,159] For the conventional Coulomb operator $\hat{\mathrm{W}} = \sum_{ij}\frac{1}{r_{ij}}$, based on an initial procedure by Gygi and Baldereschi,[158] Broqvist *et al.*[159] have suggested to write:

$$\Phi(\mathbf{G}) = \begin{cases} \dfrac{1}{\Omega}\dfrac{4\pi}{\mathbf{G}^2} & \text{for } \mathbf{G} \neq 0 \\ \chi(0) & \text{for } \mathbf{G} = 0, \end{cases} \tag{6.29}$$

where Ω denotes the supercell volume and the screening function $\chi(0)$ is obtained as the $\lim_{\gamma\to 0}$ from:

$$\chi(\gamma) = \left[\frac{1}{\sqrt{\pi\gamma}} - \frac{4\pi}{\Omega}\sum_{\mathbf{G}}\frac{e^{-\gamma\mathbf{G}^2}}{\mathbf{G}^2}\right]. \tag{6.30}$$

It is straightforward to show that for a Coulomb attenuated operator $\hat{\mathrm{W}}_{\mathrm{lr}} = \frac{\alpha+\beta\mathrm{erf}(\mu r_{ij})}{r_{ij}}$, we have that[161]:

$$\Phi_{\mathrm{lr}}(\mathbf{G}) = \begin{cases} \dfrac{1}{\Omega}\dfrac{4\pi}{\mathbf{G}^2}\left[\alpha + \beta e^{-G^2/4\mu^2}\right] & \text{for } \mathbf{G} \neq 0 \\ \alpha\chi(0) + \beta\chi\left(\frac{1}{4\mu^2}\right) & \text{for } \mathbf{G} = 0. \end{cases} \tag{6.31}$$

Our implementation of the range-separated Minnesota family members make use of eq. 6.31, a derivation of which is available in the literature.[159,161]

6.3 Computational Details

Plane wave calculations were carried out using the CPMD code.[176] Hard Goedecker-Teter-Hutter (GTH)[154] pseudopotentials have been used in order to ensure maximum transferability. Hartree-Fock exchange energies,[160] where not otherwise stated, were obtained using a constant, default $\xi = 4$, whereas ξ used in the calculation of τ, $\nabla\rho$, ρ and the xc energy was set to different values. These are reported in the result sections, along with orbital cutoff energies E_{cut}. The Poisson equations of the periodic images were decoupled using the algorithm by Martyna and Tuckerman.[155] The orthorhombic supercell for the HF dimer spanned $20 \times 15 \times 10$ Å^3, whereas the supercells for the reactions of the HC7/11 database were cubic of dimension 30^3 Å^3. In order to rule out basis set convergence issues for the M06-2X functional for reactions **1** and **2**, reaction energies were also calculated in a 20^3 Å^3 supercell, using a cutoff energy of 275 Ry and $\xi = 12$ for both local and Hartree-Fock contributions.

Calculations employing a Slater basis were performed using the ADF[185–187] program package and the TZ2P and ATZ2P basis sets,[184] adding diffuse fitting functions for the RI calculations and eliminating linearly dependent basis functions using a threshold of 10^{-4}. The xc energy and potentials were calculated using libxc[215] on a very fine Becke grid[216] ('excellent quality' in ADF jargon).

Calculations using Gaussian basis sets were carried out using the Gaussian16[217] suite of programs, a tight convergence criterion on the (generalised) Kohn-Sham orbitals and a superfine integration grid.

6.4 Results and Discussion

6.4.1 The Total Energy of a HF Dimer

In the following, we will provide an analysis of the convergence of the M05 to M11 families of functionals in a plane wave basis. First, we will consider the conventional case at fixed $\xi = 4$ and increasing E_{cut}. We will then perform the same analysis at increasing values of ξ in order to assess the effect of a denser mesh.

6.4.1$_{\mathrm{A}}$ Convergence in Plane Waves

The convergence behaviour of the functionals considered here with respect to the energy cutoff E_{cut} and using a standard $\xi = 4$ is illustrated in Fig. 6.3 at the example of the binding energy of an HF dimer. A reference value for every functional, obtained at a cutoff of 300 Ry and $\xi = 12$ is illustrated as a straight line.

Panel 6.3a) shows the convergence of the PBE and TPSS[132] xc functionals, which will serve as an example for standard generalised gradient approximation (GGA) and meta-generalised gradient approximation (MGGA) functionals, respectively. Even at the lowest cutoff energy considered here, 125 Ry, the energy is converged to less than 0.3 kcal mol^{-1} , which is already within chemical accuracy. Values at 225 Ry are essentially converged in a sub 0.1 kcal mol^{-1} regime. In both cases, the energies at lower cutoff values are an upper bound to the reference value. The situation is fundamentally different for the Minnesota functionals studied here. In panel 6.3b), the oldest members of the group, M05 and M05-2X, are shown. In both cases, the reference energy can only be reached at a cutoff value of 300 Ry, with deviations at 125 Ry reaching the 3 kcal mol^{-1} range. In both cases, the unconverged energies do not serve as an upper bound for the converged values. A similar behaviour is observed in panel 6.3c) for the M06-L functional, although the maximum deviation is less than 1.5 kcal mol^{-1} . The spread of the values is much lower for M06, where deviations < 0.3 kcal mol^{-1} are reached at 200 Ry.

The situation improves for M06-HF in panel 6.3d), where deviations only narrowly exceed 1 kcal mol^{-1} and lie within a 0.3 kcal mol^{-1} range from 225 Ry onward. Values for M06-2X, on the other hand, are most erratic and vary by up to almost 3 kcal mol^{-1} for certain cutoff values. For the two members of the M08 family shown in panel 6.3e), both M08-SO and M08-HX converge around 225 Ry, even though deviations at 125 Ry are larger than for the M06 family. Deviations are again larger for M11 and M11-L shown in panel f), with maximum deviations of 2 kcal mol^{-1} , second only to the M05 family.

When discussing the basis set convergence in a Gaussian basis, several authors have stressed the importance of a fine integration grid[204, 218, 219] in order to obtain accurate values for most Minnesota functionals. In plane waves, while an increase in cutoff energy E_{cut} implies a finer mesh on which the xc energy is evaluated, the high frequency components in reciprocal space may also introduce further noise. A finer integration grid can be obtained by increasing the value of ξ. Performing the convergence analysis at higher values of the dual ξ therefore allows to assess the influence of grid granularity on the energy convergence.

Fig. 6.3 shows a clear improvement of the convergence behaviour of the Minnesota functionals for higher values of ξ. For the reference functionals in panel 6.3a), the influence of an increased mesh is far below chemical accuracy. Even for the more sensitive TPSS functional, deviations between different values of ξ do not exceed 0.2 kcal mol^{-1} . For both PBE and TPSS, convergence within a sub 0.1 kcal mol^{-1} interval can be reached at 175 Ry when using $\xi \geq 8$. Values obtained using ξ of 8 and 12 are virtually indistinguishable for both functionals. The convergence behaviour of the

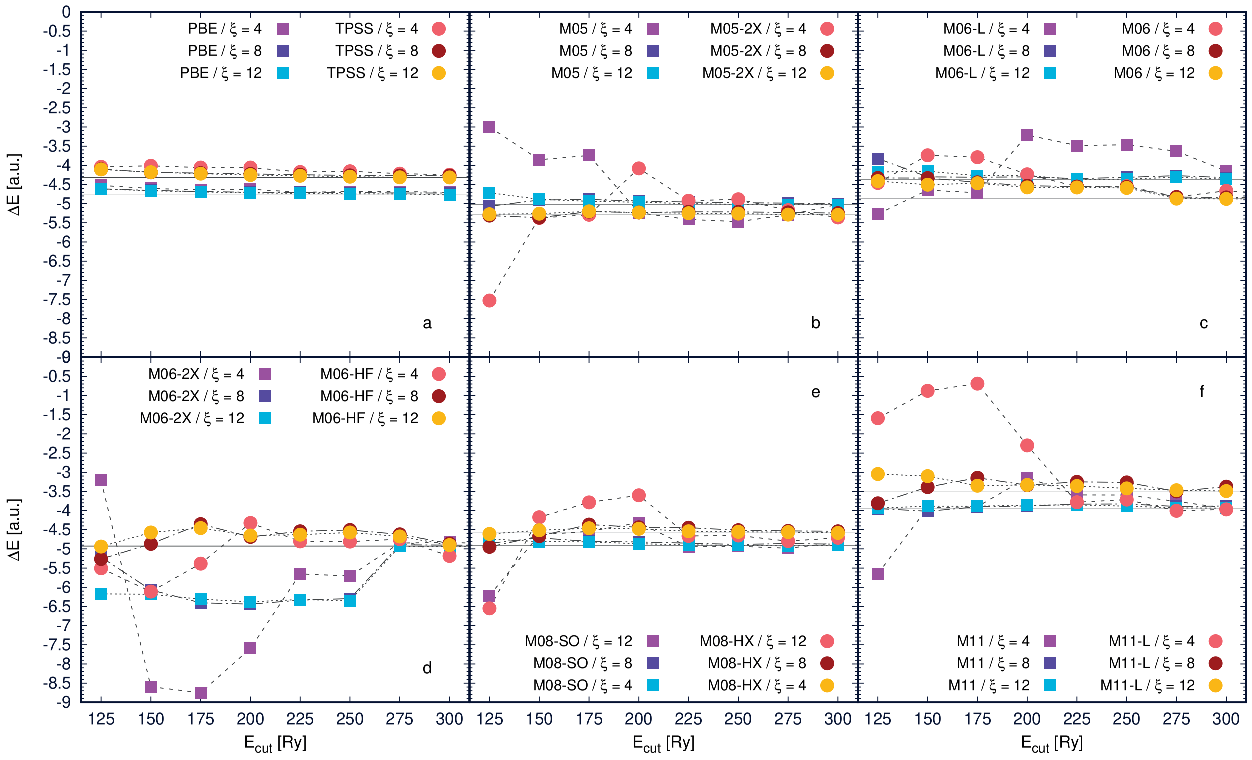


Figure 6.3: Convergence of the binding energy ΔE of a HF dimer, grouped according to functionals, with respect to cutoff energy E_{cut} and ξ. Converged binding energy values are illustrated with a line.

Minnesota functionals in panels 6.3b) to 6.3f) substantially improves upon increasing ζ. For both ξ of 8 and 12, all energies except M06-HF, M06-2X and M11-L lie within a < 0.3 kcal mol^{-1} interval starting from 150 to 175 Ry, and are close to within a sub 0.1 kcal mol^{-1} range from 200 Ry on. The convergence behaviour is therefore very similar to the PBE and TPSS functionals indicating that it is not the high-frequency components in the wavefunction expansion that are needed for accurate energetics, but rather a very fine mesh for the density. These observations are in line with the common requirement imposed on the integration grid for Minnesota functionals in atom-centred bases. Increasing ξ further from 8 to 12 results in changes of energetics that are negligible. The remaining outliers are M06-2X, M06-HF and M06-L. For M06-HF, values for ξ of 8 or 12 remain within a 0.3 kcal mol^{-1} range. For cutoff energies lower than 200 Ry, results obtained from ξ 8 deviate visibly from those obtained at $\xi = 12$. The only other example of this behaviour is M11-L, which overall behaves very similarly to M06-HF for higher values of ξ and is therefore slightly less well-behaved than many of the earlier-generation Minnesota functionals studied here. Still, due to the narrow spread of the energies once a higher ξ is used, this is not expected to result in practical problems. The least well-behaved outlier is M06-2X, where binding energies abruptly change at 275 Ry. Only from this value onward do the energies remain within a very narrow range and appear to be properly converged. We have verified the stability of the M06-2X results by increasing the value of $\xi = 20$, which did not alter the convergence behaviour, nor did it have any significant influence on the binding energies. It therefore appears that, in addition to a suitably fine integration grid, M06-2X needs an increased amount of high-frequency components in the wavefunction expansion in order to account for the binding of the HF dimer.

For practical applications, energy differences can therefore be converged straightforwardly, provided that the value of ξ is increased above the standard of 4. For chemical accuracy, a value of 8 has proven sufficient for the system considered here. This makes it possible to obtain reference values for the Minnesota functionals in a fully nonlocal basis, enabling studies free of basis-set superposition errors and independent of the balance of basis functions and integration grids employed. In the case of M06-2X, at least for the weakly bound HF dimer studied here, particular attention has to be paid to the choice of integration grid and cutoff.

The substantial improvement of convergence behaviour when increasing ξ from 4 to 8 illustrates yet once more the need for a dense enough mesh. Our observations suggest that the cutoff of the basis set is as straightforwardly controlled as it is the case of conventional functionals (with the exception of M06-2X, which requires particular attention). Instead, for a basis of a given size, it is the integration mesh of the xc term that is the most crucial ingredient. This is an unconventional observation in a plane wave basis, since commonly, and as reported here for PBE and TPSS, convergence is reached straightforwardly by increasing the energy of the highest-frequency plane wave in the expansion at a default $\xi = 4$. However, this exceptional behaviour of the Minnesota functionals is in line with the requirements for very dense integration grids that are recommended in Gaussian bases.[218] In contrast to commonly employed Gaussian bases and the observations by Mardirossian *et al.* we can rule out any influence of the balance of the exponents of the basis set on the final result, implying that for a given pseudopotential and given an appropriately dense mesh, the convergence behaviour of the Minnesota functionals does not differ from that of other meta-GGA functionals.

The flexibility of a plane-wave based description therefore makes it possible to obtain truly converged energetics, which can serve as reference values for calculations carried out in other (atom-centred) basis sets.

6.4.1$_{\mathrm{B}}$ Comparison to Atom-Centred Bases

Among the functionals studied, M05, M06-2X, M06-HF and, to a lesser extent, the M08 and M11 families are the most sensitive to changes in cutoff energies. Both M06-HF and M11-L are functionals that Mardirossian *et al.* have identified as particularly slowly convergent with respect to the Gaussian basis set size; unlike the M05 family and M06-2X, which were found to be comparably well-behaved. However, the convergence of M05 and M05-2X substantially improves once ξ is increased resulting in a well-behaved approach to the converged limit (which has also been documented in an atom-centred basis of Gaussian functions at the example of a sufficiently fine integration grid). M06-L, however, appears to be much more well-behaved in plane waves than what was reported for a Gaussian basis.

Table 6.1 provides more insight into the basis-set sensitivity of the Minnesota family of functionals. Values of the HF dimer binding energy are reported for all three choices of dual considered here (4,8,12). The converged values from plane waves at ξ 12 are compared to values obtained from the commonly used and popular aug-cc-pVTZ and 6-311++G(2df,2p) basis sets as well as the complete Gaussian basis set of Ref. 204 as a reference. In order to exclude possible effects due to the unphysical long-range decay and cusp behaviour of Gaussian functions, results for a larger Slater basis set (TZ2P) and an augmented Slater basis (ATZ2P) are also given.

Table 6.1: HF dimer binding energy calculated for the M05, M06, M08 and M11 family of Minnesota functionals in different basis sets. Results for aug-cc-pVTZ and a customised complete Gaussian basis (Limit) are reproduced from Ref. 204. Plane waves results were obtained at a wavefunction cutoff of 300 Ry. Atom-centred results are not Counterpoise corrected, in agreement with common procedures used in functional benchmarking and fitting.

	Plane Waves			Slater Functions		Gaussian Functions		
Functional	$\xi = 4$	$\xi = 8$	$\xi = 12$	TZ2P	ATZ2P	aug-cc-pVTZ	6-311++G(2df,2p)	Limit
M05	-5.02	-5.00	-5.03	-4.98	-4.95	-4.99	-5.36	-4.96
M05-2X	-5.36	-5.25	-5.30	-5.12	-5.08	-5.06	-5.18	-5.04
M06	-4.74	-4.84	-4.88	-4.20	-4.24	-4.40	-4.71	-4.44
M06-2X	-4.87	-4.90	-4.94	-4.90	-4.88	-4.95	-5.13	-4.87
M06-HF	-5.18	-4.86	-4.90	-4.43	-4.48	-4.71	-5.01	-5.24
M06-L	-4.15	-4.32	-4.37	-4.49	-4.52	-4.46	-4.55	-4.59
M08-HX	-4.71	-4.54	-4.58	-4.69	-4.70	-4.94	-5.14	-4.74
M11	-3.94	-3.88	-3.94	-4.45	-4.53	-4.58	-4.87	-4.47
M11-L	-3.97	-3.38	-3.50	-3.57	-3.73	-3.86	-4.10	-3.41

For both members of the M05 family, results across all bases agree well, with a maximum deviation of 0.26 kcal mol^{-1} between the converged plane wave result and the complete basis set limit obtained from Gaussian functions. Results obtained in a Slater basis do not deviate by more than 0.04 kcal mol^{-1} from aug-cc-pVTZ and the limit value from Ref. 83; changes upon inclusion of an augmentation function are lower than 0.04 kcal mol^{-1} . However, the binding energy obtained with 6-311++G(2df,2p), which is similar to the basis commonly used to benchmark the database that contains the HF dimer, deviates by about 0.4 kcal mol^{-1} from the reference value in the case of M05. As reported by Mardirossian *et al.* changes between the complete basis and aug-cc-pVTZ are negligible. For M05 and M05-2X, the overall agreement between the different bases is well within chemical accuracy.

The deviations are larger for some members of the M06 family. For the parent M06 itself, plane wave binding energies using $\xi = 12$ exceed the reference value by 0.44 kcal mol^{-1} . The values obtained in a Slater basis are by about 0.20 kcal mol^{-1} lower than the reference value and are again insensitive to the addition of augmentation functions. Results from 6-311++G(2df,2p) are closer to plane wave values than to the converged value in a Gaussian basis. For M06-2X, the agreement between plane waves, both Slater bases, aug-cc-pVTZ and the value from a complete Gaussian basis show excellent agreement, with a deviation of ≤ 0.10 kcal mol^{-1} . The largest error with respect to the limit occurs again for 6-311++G(2df,2p), with a difference of 0.25 kcal mol^{-1} . Given the rather erratic convergence behaviour of this functional, these are encouraging results; indicating that once the requirements on grid and cutoff are met, the plane wave implementation accurately reproduces results from atom-centred bases. In the case of M06-HF, converged plane wave values differ by 0.34 kcal mol^{-1} from the reference, which is smaller than the deviation that occurs for aug-cc-pVTZ (0.53 kcal mol^{-1}). The value for 6-311++G(2df,2p) is closer to the limit, while both augmented and non-augmented Slater bases give deviations of up to 0.81 kcal mol^{-1} . The situation is improved for the completely local M06-L, where the deviation between the converged plane wave value and the limit in a Gaussian basis is 0.22 kcal mol^{-1} . The value for aug-cc-pVTZ is slightly closer (error of 0.13 kcal mol^{-1}), whereas the two Slater bases and 6-311++G(2df,2p) are closest to the value of the limit.

The situation is similar for the more recent M08 and M11 families. For M08-HX, the converged plane wave result differs by only 0.16 kcal mol^{-1} , which is smaller than the deviations of aug-cc-pVTZ and 6-311++G(2df,2p) (0.20 and 0.40 kcal mol^{-1} , respectively). The binding energies obtained using TZ2P and ATZ2P show only a negligible deviation of $\leq$ 0.10 kcal mol^{-1} . Plane wave deviations are largest for M11, where the converged plane wave binding energy differs by 0.51 kcal mol^{-1} from the reference value. It should however be noted that this value is comparable to the deviations observed for Slater bases in the case of M06-HF, and for 6-311++G(2df,2p) in the case of M06. Values obtained within a Slater basis and aug-cc-pVTZ are again closer to the reference (deviations of maximum 0.11 kcal mol^{-1}), while 6-311++G(2df,2p) is again the least accurate Gaussian basis, with an error of 0.40 kcal mol^{-1} . Errors for the local M11-L are much lower for plane waves, with a deviation of 0.09 kcal mol^{-1} between converged values. The error in a Slater basis is about 0.32 kcal mol^{-1} for the augmented basis, but it is only 0.16 kcal mol^{-1} for TZ2P. M11-L is the only functional for which the influence of augmentation functions in a Slater basis exceeds the 0.10 kcal mol^{-1} regime. The largest deviation between aug-cc-pVTZ, 6-311++G(2df,2p) and the reference value are also observed for M11-L, ranging from 0.45 to 0.69 kcal mol^{-1} .

Overall, for all of the functionals but M06-HF, M11 and M11-L, the results between plane waves, Slater bases and different Gaussian bases show excellent agreement. The largest deviation is observed in the non-augmented Slater basis for M06-HF, even though it is still smaller than chemical accuracy (1.0 kcal mol^{-1}). The maximum deviation observed in a Gaussian basis occurs for M11-L/6-311++G(2df,2p), exceeding 0.60 kcal mol^{-1} . Errors of about 0.50 kcal mol^{-1} can be observed for all bases, but not necessarily for the same functional. Maximum deviations in plane waves are reached for M11, whereas the maximum deviation for aug-cc-pVTZ occurs for M11-L. These results highlight the importance of benchmarking the results not only when changing the type of basis - plane waves, Slater, Gaussians - but even when changing from one kind of Gaussian basis (aug-cc-pVTZ) to another (6-311++G(2df,2p)). This can be of particular importance when comparing to results from large benchmarking data bases, where sometimes, very specific basis sets are used. The deviations observed so far suggest that, for the Minnesota family of functionals, basis-set effects may exceed half a kcal mol^{-1} . While it is encouraging that these errors are still smaller than chemical accuracy, they can possibly affect the average performance of a functional, and conclusions based on benchmarks that cannot be carried out sufficiently close to the basis set limit should bear this source of error in mind. In view of common practice, a careful assessment of basis-set related errors will be of particular importance if a study makes use of a basis set different than the benchmark basis set.

Our plane-wave implementation offers the possibility of smoothly reaching a converged value, as exemplified in the usually small changes of binding energies when increasing ξ from 8 to 12. The residual differences between the converged values obtained from the custom Gaussian basis by Mardirossian *et al.* and our converged plane wave results suggest that the flexibility due to the fully delocalised plane waves leaves room for an improved description of the basis-set limit.

6.4.2 Implications for Reaction Enthalpies of the HC7/11 Database

Given the popularity and success of the Minnesota functionals in the description of organic molecules,[46, 209] we now consider their performance for a particularly difficult set of thermochemical data: The reaction enthalpies of the HC7/11 database.[208, 210] In particular for reactions **1** and **2**, the introduction of the M06 family constituted a significant improvement over its predecessor, M05.

In order to assess the influence of the choice of basis on the overall performance of the functionals, we have calculated reaction enthalpies at 0 K for the HC7/11 database of hydrocarbon reactions in a plane wave basis, aug-cc-pVTZ and the ATZ2P Slater basis for M05 as well as the complete M06 and M11 families. While the standard plane wave setup was a 30 Å^3 box with $E_{cut} = 125$ Ry and $\xi = 8$, values for M06, M06-2X and M06-HF were also calculated in a 20 Å^3 box with $E_{cut} = 275$ Ry and $\xi = 12$. These values, however, differ by less than 0.5 kcal mol^{-1} from those at $E_{cut} = 125$ and are therefore not reported.

6.4.2$_\text{A}$ The M05 family

The values for the related M05 and M06 families are given in Table 6.2. For M05, all values show excellent agreement between plane waves and the atom centred bases. The deviations with respect to the Slater basis range from a minimum of 0.05 kcal mol^{-1} for reaction **6** to 0.68 kcal mol^{-1} for reaction **4**, and from 0.30 kcal mol^{-1} for reaction **2** to 2.12 kcal mol^{-1} for reaction **6** (which is still in the sub-percent regime) aug-cc-pVTZ. However, reactions **5** and **6** show the largest spread over all bases, with results differing by about 0.22 kcal mol^{-1} between 6-311++G(2df,2p) and aug-cc-pVTZ for reaction **5**, and reaction **6** even showing the best agreement between Slater bases and plane waves among all the test set. The difference between plane waves and the Slater basis is only about 0.29 kcal mol^{-1} for reaction **5** and therefore much lower than the deviation between Slater basis and aug-cc-pVTZ (0.77 kcal mol^{-1}). Overall, best agreement is reached between the

Slater basis and plane waves, with deviation between Slater bases or plane waves and Gaussian bases being slightly larger. However, deviations between aug-cc-pVTZ and 6-311++G(2df,2p) can be equally sizeable for systems **6** and **7**. Finally it should be noted that for reactions **1** and **2**, the spread is about equal between all the bases considered here. Overall, the reaction enthalpies obtained from the different bases agree very well, making the results reproducible across different approaches.

6.4.2$_{\mathbf{B}}$ The M06 family

The average agreement is slightly inferior for the M06 family. With respect to M05, the spread of the values increases substantially between all the atom-centred bases. While the deviations remain in a sub-percent or sub kcal mol^{-1} range for reactions **3** and **4**, they can be larger for the remaining reactions. Slater basis results for reactions **1**, **2** and **7** differ between 1 and 4 kcal mol^{-1} from Gaussian results, which themselves exhibit a significant spread of up to ≈ 2 kcal mol^{-1} for reaction **7**.

For reactions **3** to **5**, M06 shows outstanding agreement between plane wave basis and Slater results, with a minimum deviation of 0.04 kcal mol^{-1} , and a maximum deviation of 0.17 kcal mol^{-1} . Deviations between plane waves and the Gaussian bases correspond to those observed between ATZ2P and Gaussian functions. The agreement for reactions **1**,**2**,**6** and **7** is less homogeneous, ranging from 1.5 to 2.07 kcal mol^{-1} . The relative errors for reactions **6** and **7**, however, do not exceed 2 % with respect to a Slater basis. For reactions **1** and **2**, plane wave and Slater basis values show a comparable absolute deviation with respect to the Gaussian bases, but are of opposite sign.

Plane wave results for the local M06-L are again closest to those obtained in a Slater basis, but consistently lower. In spite of a good overall agreement, deviations between ATZ2P and plane waves can occasionally exceed 1 kcal mol^{-1} . For the isomerisation reactions **1** and **2**, values show excellent agreement between the Slater ATZ2P and the Gaussian reference basis, whereas the differences between both plane waves and aug-cc-pVTZ with respect to ATZ2P are comparable, but of opposite sign. In both cases, the spread of values slightly exceeds chemical accuracy with respect to the reference value. Similarly, the octane isomerisation **3** exhibits a large spread, with the value from ATZ2P again lying in between plane waves and aug-cc-pVTZ. In this case, notably, relative errors are large, with the plane wave, Slater basis and aug-cc-pVTZ values being roughly one, two and three quarters of the value obtained in the reference basis. However, due to the small magnitude of

Table 6.2: Reaction enthalpies for the members of the HC7/11 database computed using the M05, M06, M06-2x, M06-HF and M06-L xc functional in plane waves (125 Ry, $\xi = 8$), an augmented Slater basis (ATZ2P) and the commonly used Gaussian aug-cc-pVTZ basis. Reference values correspond to 6-311+G(2df,2p), which is the recommended basis for benchmarking in Ref. 43.

System		$\xi = 8$	ATZ2P	aug-cc-pVTZ	Reference
M05					
1	E22-E1	28.59	28.36	29.89	30.17
2	E31-E1	35.60	35.46	37.30	37.68
	E31-E22	7.01	7.10	7.41	7.51
3	Octane isomerisation	7.73	7.64	7.05	7.14
4	Reaction (a)	5.12	5.84	5.58	5.49
5	Reaction (b)	0.16	0.45	1.22	1.00
6	Reaction (c)	191.17	191.12	188.55	190.75
7	Reaction (d)	125.11	124.81	123.61	124.99
M06					
1	E22-E1	20.96	18.46	19.65	19.26
2	E31-E1	27.05	24.98	26.61	26.11
	E31-E22	6.09	6.52	6.96	6.85
3	Octane isomerisation	2.24	2.20	2.89	3.00
4	Reaction (a)	7.32	7.48	7.77	7.89
5	Reaction (b)	13.20	13.37	14.52	14.78
6	Reaction (c)	190.81	192.12	196.69	198.49
7	Reaction (d)	124.46	125.56	129.04	130.08
M06-L					
1	E22-E1	15.71	16.45	17.59	16.28
2	E31-E1	18.04	19.71	21.50	19.91
	E31-E22	2.33	3.26	3.91	3.63
3	Octane isomerisation	0.25	0.61	0.99	1.23
4	Reaction (a)	5.70	6.07	6.08	6.23
5	Reaction (b)	8.78	9.68	10.07	10.51
6	Reaction (c)	179.41	185.29	191.53	190.41
7	Reaction (d)	115.98	120.44	124.85	124.01
M06-2X					
1	E22-E1	20.18	15.77	16.51	16.30
2	E31-E1	28.39	23.78	23.94	23.65
	E31-E22	8.21	8.01	7.43	7.35
3	Octane isomerisation	1.34	1.44	1.76	2.11
4	Reaction (a)	7.58	7.55	7.71	7.98
5	Reaction (b)	12.65	12.70	13.28	14.02
6	Reaction (c)	203.77	198.24	198.46	198.36
7	Reaction (d)	133.81	129.60	129.64	129.44
M06-HF					
1	E22-E1	21.38	13.30	13.98	14.36
2	E31-E1	32.52	22.78	23.22	23.90
	E31-E22	11.14	9.48	9.24	9.54
3	Octane isomerisation	1.34	2.22	3.85	3.62
4	Reaction (a)	7.32	7.97	8.66	8.86
5	Reaction (b)	12.30	14.10	16.79	16.92
6	Reaction (c)	214.00	201.76	202.74	201.17
7	Reaction (d)	139.92	131.26	131.99	130.61

the reaction enthalpy, the errors stay below chemical accuracy. Best agreement is reached for reaction **4**, with a deviation of 0.37 kcal mol^{-1} between ATZ2P and plane waves and 0.16 kcal mol^{-1} between ATZ2P and the reference basis. The spread is again larger for reactions **5**, **6** and **7**, but the difference between plane waves/ATZ2P and aug-cc-pVTZ/ATZ2P remain either within the sub-kcal mol^{-1} range (reaction **5**) or lie within 5% of the reference value. In all three cases, the deviations are again of very similar absolute values, but of opposite sign.

M06-2X exhibits a behaviour largely reminiscent of M06. Enthalpies for reactions **3** to **5** agree well between plane waves and ATZ2P, while the deviations for reactions **1**, **2**, **6** and **7** roughly double with respect to M06, although they still exhibit similar trends. M06-HF shows the worst agreement among the M06 family, with deviations up to one order of magnitude larger than for M06 itself. For reactions **3** to **5**, however, deviations with respect to a Slater basis remain comparable to those between the Slater basis and the Gaussian reference basis. In general, for the hybrids of the M06 family, the maximum deviation with respect to atom-centred bases increases with the percentage of exact exchange.

Overall, for M05, M06 and M06-L, the enthalpies obtained in plane waves show excellent agreement with the values obtained in an augmented Slater basis. Deviations with respect to Gaussian bases may be somewhat larger, with changes between Slater functions and plane waves often being comparable to a change from Slater functions to Gaussians. For M06-2X and M06-HF, the most important deviations are observed for reactions **1** and **2**. Absolute errors are larger for reactions **6** and **7**, but due to the large magnitude of the enthalpy, relative errors remain much lower than for **1** and **2**. However, the enthalpy of the conversion of reactant E22 to product E31 - which is not listed as an official database entry - reveals good agreement between the basis sets for all the functionals; for this system, only M06-HF shows a value that exceed chemical accuracy. This observation indicates that the problem might be linked to reactant E1. The source of this considerable deviations for reactions **1** and **2** will therefore be examined further on in the text.

6.4.2$_{\mathrm{C}}$ The M11 family

The basis-set dependence of the enthalpies obtained from M11 is comparably uniform for all reactions. For the isomerisations **1** and **2**, the values from the reference Gaussian basis (6-311+G(2df,2p)) lie between the values from ATZ2P/aug-cc-pVTZ and plane wave results, respectively, with deviations

Table 6.3: Reaction enthalpies for the members of the HC7/11 database computed using the M11 and M11-L xc functional in plane waves (125 Ry, $\xi = 8$), an augmented Slater basis (ATZ2P), aug-cc-pVTZ and the reference basis (6-311+G(2df,2p)) .

	System	$\xi = 8$	ATZ2P	aug-cc-pVTZ	Reference
	M11				
1	E22-E1	21.67	20.14	20.19	20.83
2	E31-E1	29.48	28.09	27.59	28.94
3	Octane isomerisation	3.52	3.00	4.10	4.52
4	Reaction (a)	6.60	6.94	7.38	7.47
5	Reaction (b)	13.40	13.35	15.16	15.68
6	Reaction (c)	193.15	196.20	193.97	198.68
7	Reaction (d)	125.61	128.15	126.46	129.52
	M11-L				
1	E22-E1	19.99	14.60	17.12	15.91
2	E31-E1	26.52	21.12	24.26	22.80
3	Octane isomerisation	0.01	1.29	1.27	1.14
4	Reaction (a)	6.51	8.26	8.17	7.88
5	Reaction (b)	9.69	13.60	13.56	12.82
6	Reaction (c)	195.12	194.15	197.70	198.55
7	Reaction (d)	126.83	126.52	129.47	129.74

ranging from -0.9 to $+0.9$ kcal mol^{-1} with respect to the reference. For the octane isomerisation **3**, plane wave results deviate by about 0.5 to 1.0 kcal mol^{-1} from results obtained using Gaussians, whereas the difference between Slater and Gaussian bases exceeds chemical accuracy, ranging from 1.1 to 1.5 kcal mol^{-1} . For reactions **4** and **5**, plane waves and ATZ2P show excellent agreement, whereas the values from aug-cc-pVTZ and the Gaussian reference basis lie up to 2 kcal mol^{-1} higher. Interestingly, the situation is different for reactions **6** and **7**, where plane waves and aug-cc-pVTZ agree well; the ATZP and 6-311+G(2df,2p) results are up to 4 kcal mol^{-1} higher in energy (which corresponds to a 5 % range).

Energetics vary more considerably in the case of M11-L, in analogy with the slow basis set convergence documented in Ref. 204 and the trends observed in the present work for the HF dimer. For the isomerisations, the enthalpies span a range of over 4 kcal mol^{-1} ; ATZ2P values are the lowest, plane wave values the largest, aug-cc-pVTZ and 6-311+G(2df,2p) lie in between and differ themselves by almost 2 kcal mol^{-1} . For reactions **3** to **5**, all atom-centred bases agree well, with plane wave results deviating by more than 1 kcal mol^{-1} with respect to the closest Gaussian value. Reactions **6** and **7** show again good agreement within about 1 kcal mol^{-1} between plane waves and Slater functions, whereas the enthalpies obtained for aug-cc-pVTZ and the reference basis are between $2 - 3$ kcal mol^{-1} higher. These results highlight the possible high sensitivity of range-separated density functionals to the underlying basis set.

Given the large spread of enthalpies obtained in atom-centred bases, plane wave results still agree well with the other basis sets. Consistent with Ref. 204, we find that M11-L is more sensitive to basis set effects than its hybrid equivalent, M11. This sensitivity is reflected in a considerable change of reaction enthalpies when changing the type of basis function. While similar trends hold for M11, they are exacerbated for M11-L. M11-L therefore appears to be exceptionally sensitive to the underlying functional form of the basis, which may be due to its dual-range functional form: Similar trends were observed for certain systems when comparing excited states obtained from range-separated functionals in plane waves and Gaussians.[161] Like for the M06 family, plane waves therefore allow for new insights on converged energetics to be obtained. In view of the improved accuracy of M11 over M06, the fact that converged quantities are readily available from plane waves is particularly favourable.

6.4.2$_{\text{D}}$ Summary

Overall, results obtained from plane waves and Slater bases show satisfactory agreement with Gaussian bases, even if there appears to be no systematic trend linking plane wave, Slater and Gaussian basis results among all of the functionals and systems. While the spread between the results may be considerable for a given reaction, the deviations between the different bases are rather evenly spread for many of the functionals considered here (M05, M06, M06-L, M11, M11-L). Large relative deviations are only observed for the isomerisation reactions **1** and **2**. Often, where larger deviations can be observed, Slater basis results lie between plane wave and aug-cc-pVTZ values. In many cases, the deviations of plane waves and aug-cc-pVTZ with respect to the Slater basis are of comparable absolute value, but are of opposite sign. In the case of M11, which has a different functional form, values from the Gaussian reference basis lie between those obtained in plane waves and those obtained with Slater functions, which may be attributed to effects due to range-separation. The only significant absolute deviations between plane waves and atom-centred bases occur for M06-2X, M06-HX and M11-L, which, with the exception of M06-2X, have both been identified as difficult to converge in Gaussian bases by Mardirossian *et al.*[204] M06-L, on the other hand, appears to be more well-behaved in plane waves than in Gaussians, exhibiting the lowest deviations within the M06 family along with M06 itself. In the case of M11-L, deviations are larger than for M11, but comparable to the differences observed within different atom-centred bases. This indicates that the functional simply calls for a very flexible basis. M05, M06 and M06-L show the best overall agreement across all bases considered here.

Our results show that a plane wave description can add valuable information on the convergence and basis set limit of reaction enthalpies. In particular, where energetics differ considerably between Slater functions and Gaussian basis sets, the fully delocalised plane wave basis allows for a more flexible description and a convenient approach to the basis set limit.

6.4.3 Energy Density and Exact Exchange Analysis for the M06 Family

For the hybrid functionals of the M06 family, agreement between plane wave and atom-centred basis results for the isomerisations **1** and **2** noticeably deteriorates with increasing Hartree-Fock exchange contribution: From M06 over M06-2X to M06-HF, the maximum errors increase from 2 over 5 to 14 kcal mol^{-1} . In contrast, the conversion of E31 to E22 is much more accurately described by all the functionals, with a maximum error of about 1.5 kcal mol^{-1} for M06-HF. The problem therefore appears to be related to the description of reactant E1.

We first consider the effect of exact exchange. The pure Hartree-Fock enthalpy for reaction **1** calculated using the same plane wave setup as used in Table 6.2 is -10.29 kcal mol^{-1} , whereas the corresponding value obtained using an atom-centred aug-cc-pVTZ basis is -12.30 kcal mol^{-1} . While this difference is not negligible, it is well within the spread of values that could be observed for several systems and functionals - even among the two Gaussian basis sets. Therefore, the exact exchange contribution cannot alone account for all of the sizeable differences between plane wave and Gaussian values that were obtained using M06-2X and M06-HF.

In order to assess the possible influence of pseudising the core electrons, it appears pertinent to analyse the exchange and correlation energy densities $\epsilon_x(\mathbf{r})$ and $\epsilon_c(\mathbf{r})$, *i.e.* the integrands of the xc functional

$$E_{xc} = \int \mathrm{d}\mathbf{r}\, \rho^{4/3}(\mathbf{r})\, (K_x(\mathbf{r}) + K_c(\mathbf{r})) \tag{6.32}$$

$$= \int \mathrm{d}\mathbf{r}\, \epsilon_x(\mathbf{r}) + \epsilon_c(\mathbf{r}), \tag{6.33}$$

for reactant and product of reaction **2** and all of the functionals of the M06 family. Table 6.4 lists values for ϵ summed over the nuclear core region, *i.e.* regions of real space that lie at points $\mathbf{r}_{\text{core}}$ within the cutoff radius r_c of the pseudopotentials employed for a nucleus α at nuclear coordinate $\mathbf{R}$:

$$E_{xc}^{(r_c)} = \sum_{\alpha} \int_0^{\mathbf{r}_c} \mathrm{d}(\mathbf{r} - \mathbf{R}_\alpha) \; (\epsilon_x(\mathbf{r}) + \epsilon_c(\mathbf{r}))\,. \tag{6.34}$$

Table 6.4 shows that irrespective of the differences between plane wave and atom-centred results in Table 6.3, the changes of the contribution to ϵ_{xc} due to the core are below 1%. Similarly, the magnitude of the energy density itself is no indicator of the deviation that is to be expected with respect to all-electron calculations. It is however notable that the exchange-correlation contribution from M06-HF takes on an (unphysical[204]) positive value which counterbalances the 100% exact exchange employed in this functional.

In order to assess the effect of pseudisation further, Fig. 6.4 shows the electron density along with the energy density ϵ for a cut through one of the $C - H$ bonds of the reaction product E31. Both exchange and correlation energy densities are smooth outside the core region r_c for both M06 and M06-L. Within r_c, they can oscillate considerably, which is due to a local maximum in $\tau(\mathbf{r})$ that occurs on top of the carbon nucleus (given that $\tau(\mathbf{r})$ can be viewed as a probe for chemical bonds, one might refer to this as a misdiagnosed chemical bond). Since the oscillations are limited to the pseudised region, no transferability issues are to be expected, which is further illustrated by the values in Table 6.4. M06-2X and M06-HF, however, show turning points in ϵ_{xc} even outside r_c. In particular, ϵ_{xc} exhibits a step-like pattern when approaching the cusps of the density, whereas the correlation functional is smoother and more well-behaved. It appears that while the parametrisation of M06-2X and M06-HF can be beneficial for weak and long-range interactions, it results in unphysical phenomena when combined with pseudised nuclear cusps. Due to the high flexibility of the plane wave basis, these effects may become particularly prominent. While our results suggests that in many cases, these effects do not considerably influence energy differences due to error compensation, this is not necessarily the case for reactions involving reactant E1.

A final quantification of the numerical behaviour of the M06 family in a pseudopotential framework can be obtained by comparing the integrated ϵ_{xc} between reactant and product, where we define:

$$E_{xc}^{\text{spher}} = \int_{r_c}^{r_c+\delta r} \mathrm{d}(\mathbf{r} - \mathbf{R}_\alpha) \; (\epsilon_x(\mathbf{r}) + \epsilon_c(\mathbf{r})) \tag{6.35}$$

Table 6.4: Values for $E_x^{(r_c)}$, $E_c^{(r_c)}$ and their sum, $E_{xc}^{(r_c)}$ around the nuclear core region, *i.e.* regions within the pseudopotential cutoff radius r_c, for reaction **2**.

	Reactant (E1)			**Product** (E31)		
Functional	$E_x^{(r_c)}$ [a.u.]	$E_c^{(r_c)}$ [a.u.]	$E_{xc}^{(r_c)}$ [a.u.]	$E_x^{(r_c)}$ [a.u.]	$E_c^{(r_c)}$ [a.u.]	$E_{xc}^{(r_c)}$ [a.u.]
M05	-0.2420	-0.0028	-0.2438	-0.2406	-0.0027	-0.2432
M06	-0.1431	-0.0589	-0.2020	-0.1443	-0.0576	-0.2019
M06-L	-0.2537	-0.0208	-0.2735	-0.2532	-0.0199	-0.2732
M06-2X	-0.1354	-0.0194	-0.1547	-0.1353	-0.0187	-0.1540
M06-HF	+0.0130	-0.0596	-0.0466	+0.0137	-0.0587	-0.0450

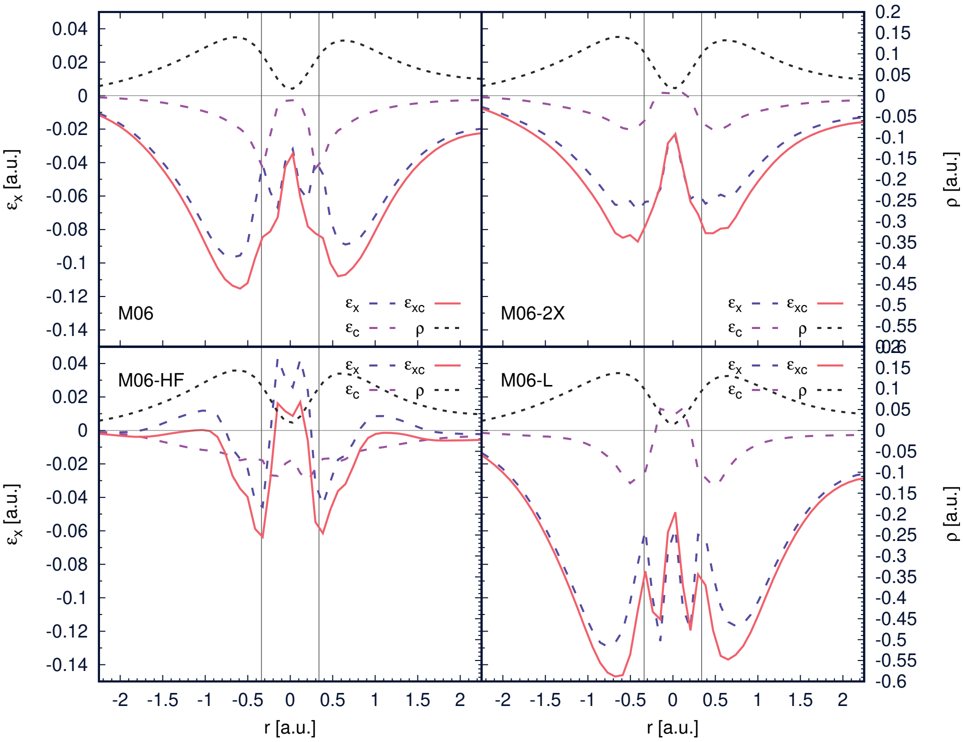


Figure 6.4: Zoom on the values of the energy density $\epsilon(\mathbf{r})$ for exchange x, correlation c and exchange-correlation xc for the M06, M06-L, M06-2X and M06-HF functional at the example of a cut through a carbon-hydrogen bond of product E31 of reaction **2**. The electron density is displayed on the secondary y axis.

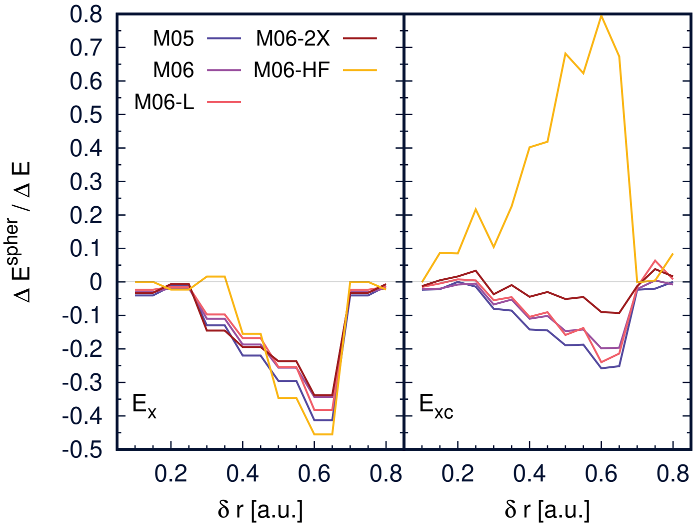


Figure 6.5: Values for the ratio of change in energy around the nuclei, ΔE^{spher}, and the total change in energy, ΔE, for exchange only x as well as exchange-correlation xc for the M05, M06, M06-L, M06-2X and M06-HF functional for the isomerisation reaction **2**.

as the integral of $\epsilon_{xc}(\mathbf{r})$ around nuclei α, starting from the pseudopotential cutoff radius r_c - outside of which the pseudo-orbitals are identical to their all-electron counterparts - and integrating up to an upper bound $r_{\max} = r_c + \delta r$. Here, we have chosen the upper bound such that $2\delta r + r_c^\alpha + r_c^\beta < d_{\alpha\beta}$ to ensure that there is no double counting between nuclei α and β separated by a bond of length $d_{\alpha\beta}$. We then plot the ratio of change of E_{xc}^{spher} between reactants and products, $\Delta E_{xc}^{\text{spher}}$, and the total change in exchange-correlation energy, ΔE_{xc}, as a function of $r_{\max}$, in Fig. 6.5. The same data is also provided for the exchange-only contribution. For comparison, the values for M05 are given, too.

For the latter, all functionals exhibit negative values for the ratio given by $\Delta E_x^{\text{spher}}/\Delta E_x$. This implies that the change of ϵ_x spherically integrated from r_c to $r_{\max}$ around each nucleus is opposite in sign to the overall change of the exchange-correlation energy. This behaviour is consistent for all functionals of the M06 family. However, for M06-2X, the ratio increases much slower and reaches lower values, whereas it increases more sharply than for

M06-HF. M05, M06 and M06-L show an almost identical behaviour. For $\delta r \geq 0.7$, the ratio decreases by almost one order of magnitude, indicating that there must be considerable areas where $\Delta e_x(\mathbf{r}) > 0$, which will compensate for the negative values at $\delta r < 0.7$, thus leading to a lower absolute value of the integral. Note that $\delta r \approx 0.7$ is close to the centre of a C-C or C-H bond. For the overall exchange-correlation energy, those trends are even more pronounced. While M06 and M06-L behave almost identically, and M05 very similarly, the ratio is consistently lower in magnitude for M06-2X, and reaches positive values for M06-HF. Given that the components of E_{xc} close to the nuclei counterbalance the overall change in E_{xc}, this can provide a rationale as for why the predicted reaction enthalpy increases in the series M06 → M06-2X → M06-HF. For M06-2X and M06-HF, the overall contribution close to the nuclei is considerably lower than for the other functionals (which is also illustrated by the values in Table 6.4 and the CH bond in Fig. 6.4), the functionals therefore lack some counterbalancing element. A flexible basis such as plane waves may amplify this tendency, resulting in an overestimation of the energy of the isomerisation. The different energetics of the Hartree-Fock exchange between all-electron bases and the plane wave/pseudopotential framework can further amplify this behaviour.

Fig. 6.6 shows the influence of different basis functions on the electron density at the example of reactant E1. Shown are electron density differences between a converged plane wave density and its analogue obtained using the aug-cc-pVTZ basis for M06 and M06-HF as well as the electron density contour with the same isovalue. In plane waves, for both functionals, the electron density is depleted around the nuclei due to pseudisation, but due to a tight cutoff radius r_c, those regions are not visible in Fig. 6.6. It can be seen from panels 6.6a) and 6.6c) that for M06, residual changes in the electron density occur along σ bonds. Their shape is a distinctive result of the higher flexibility in plane waves, leading to small changes of the electron density along bonds. These effects are expected to be highly transferable and to be compensated when comparing different systems. In contrast, panels 6.6b) and 6.6c) show that for M06-HF, plane waves show increased density along all bonds, whereas aug-cc-pVTZ yields more density further away from the nuclei. As can be seen from panels 6.6a) and 6.6b), this behaviour is amplified around the cyclopropane structure in the case of M06-HF. The large change of enthalpy between plane waves and atom-centred correlates with a less localised density difference on E1, in particular around the cyclopropane rings. Other reactions, where agreement between M06-2X, M06-HF and the atom-centred bases is much better, appear to be far less sensitive to those density differences; in particular, the enthalpies of the 'difficult' isode-

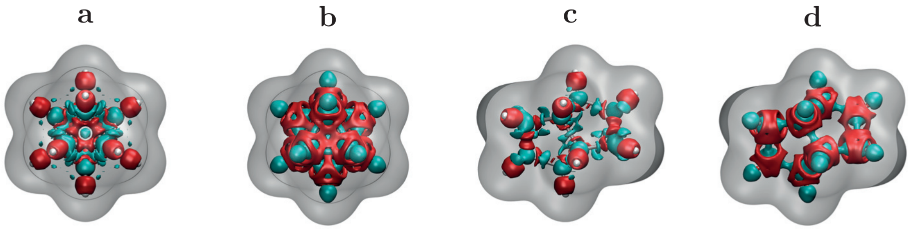


Figure 6.6: Isosurface (0.0035 a.u.) of the electron density (grey) and the electron density difference between a plane wave setup and aug-cc-pVTZ for compound E1. Regions depicted in red show more density in aug-cc-pVTZ, regions in turquoise have more density in a plane wave basis. Panel a) and c) show densities for M06, panels b) and d) for M06-HF.

smic reactions singled out by Grimme[210] are rather equally predicted by all functionals.

The differences with respect to atom-centred functions are due to the highly flexible nature of plane waves and can be amplified by the varying exact exchange energies predicted in different bases. The features shown in Figs 6.5 and 6.6 are a result of the functional form and parametrisation of the Minnesota functionals and may only manifest themselves in a highly flexible, delocalised basis such as plane waves.

Plane waves therefore offer a new perspective on the problem of functional convergence. Our results have illustrate that considerable differences with respect to atom-centred systems might occur for some specific functionals and in some particular cases. This is particularly important if results from benchmarks carried out in one basis are to be expected to be transferable to another basis. If the energetics differ considerably between bases, this is reflected in differences in the electron density around the bonding regions and by the ratio of the integral $\Delta E_{xc}^{\mathrm{spher}}/\Delta E_{xc}$. Instead, the influence of core-pseudisation has been shown to only negligibly contribute to the energy difference between all-electron and plane wave/pseudopotential calculations.

6.5 Computational Overhead

Fig. 6.7 shows the increase in computational time for one DIIS step using a fixed number of 384 CPU cores. While the scaling of the local quantities computed in plane wave codes is in principle linear with an increase in cutoff, here, for certain cutoffs, real space meshes may not be uniformly distributed

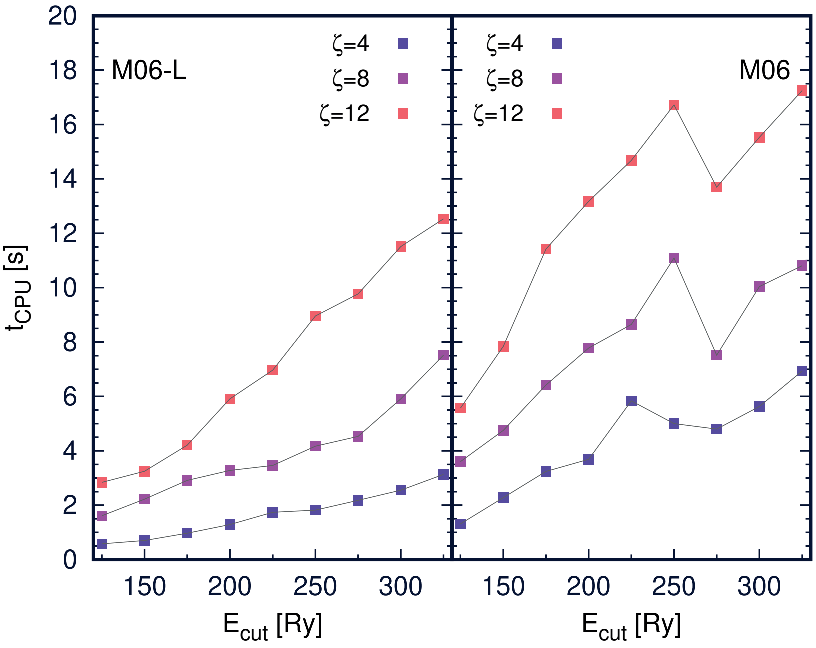


Figure 6.7: Time per DIIS step for the HF dimer using the M06 and M06-L xc functionals at different cutoff energies and using different values of ξ. All calculations were carried out in a $20 \times 15 \times 10$ Å box using 384 processors.

over processors, resulting in a computational overhead.[16] At the example of the M06-L functional, it is easily seen that the increase in runtime is proportional to the ratio between ξ used in the calculation and the baseline using $\xi = 4$. Since for any meta-GGA, the computational bottleneck is given by the 3D FFTs, an increase in ξ is directly reflected in the run time. The same holds for the hybrid functional M06, although typical load-balancing issues may occur, leading to a non-linear increase in execution time. This is especially prominent at cutoff energies above 250 Ry, and is more pronounced for higher values of ξ. Since the Hartree-Fock exchange potential does not have to be calculated on the density grid, an increase in ξ does not affect the run time in a non linear manner. Due to the small number of orbitals in the HF dimer, the overhead due to the computation of the exact exchange integrals is modest (about 2.5). The run time of the hybrid-meta GGA in an isolated system and a plane wave basis can further be decreased by resorting to recently proposed density scaling algorithms.[220]

As $\xi = 8$ has been found to be sufficient for both the HC7/11 database and the HF dimer, the computational overhead of the Minnesota functionals with respect to standard (hybrid) meta-GGA is therefore a factor of 2. Since an increase of ξ is directly reflected in a finer Fourier mesh, this overhead is very easily overcome by doubling the amount of CPUs.

6.6 Conclusions

We have presented a comprehensive analysis of the convergence behaviour of the M05 to M11 families in a plane wave basis, providing a complete basis set description within a pseudopotential framework. We could demonstrate that given a sufficiently fine integration grid, energy differences for a HF dimer converge rapidly for most functionals with the exception of M06-2X, which requires a particularly high cutoff energy. Results between atom-centred Gaussian and Slater bases agree favourably with those obtained in plane waves. Residual differences were attributed to the high flexibility of the plane wave basis.

At the example of reaction enthalpies of the HC7/11 database, we could subsequently show that for M05, M06, M06-L, M11 and M11-L, there is good agreement between values obtained in plane waves and those obtained from an augmented Slater basis, whereas deviations with respect to augmented Gaussian-type bases may be larger and can exceed chemical accuracy. For the M06-HF and M06-2X functionals, which both include a large percentage of exact exchange, we found that in the case of two isomerisations involving species E1, E22 and E31, changes with respect to atom-centred basis sets can be notable. We have shown that these differences cannot be attributed to core pseudisation. Instead, the difference between exact exchange calculated in a plane wave basis and its atom-centred equivalent can partially account for the differences. Most importantly, sizeable deviations between plane waves and atom-centred bases correlate with the ratio between the spherically integrated energy density around the nuclei and the overall exchange-correlation energy, which we have further illustrated at the example of electron density difference maps. For both M06-2X and M06-HF, this ratio has a higher value than in the case of the more well-behaved M06, M06-L and M05, which all show a more uniform behaviour. In particular, for M06, changes in electron density between plane waves and aug-cc-pVTZ were shown to be localised close to the nuclei and σ bonds, whereas the electron density differences for M06-HF are substantially more pronounced in regions far from the nuclear core. For the more recent M11 and M11-L, an unusual sensitivity with respect to the kind of basis functions was observed, which was reflected in considerable differences between results obtained with Slater functions and their Gaussian equivalents. We have attributed this to the underlying range-separation, for which the highly flexible plane waves offer a more versatile and complete description. These functionals are the most accurate considered in this study and, contrary to Gaussian bases, do not pose convergence problems when used in a plane wave/pseudopotential framework. Finally, it was shown that the need for a finer integration mesh introduces an overhead of a factor of 2, independent of the inclusion of exact exchange.

Hence, with this plane wave implementation, it becomes possible to obtain values at the complete basis set limit for the M05, M06 and M11 families. While residual changes with respect to converged Gaussian basis set calculations are small for most systems, they can be sizeable for specific reactions described using M06-2X, M06-HF or M11-L. While these features may remain hidden in a Gaussian basis, they will only surface once functions with an exact asymptotic decay (such as Slater functions) or a fully flexible and delocalised basis (such as plane waves) are used.

6.7 Acknowledgements

UR gratefully acknowledges financial support from the Swiss National Science Foundation Grant No. 200020-146645 and the NCCR MUST.

PART III

Accelerating the Evaluation of Exact Exchange Integrals in Plane Waves

Preface

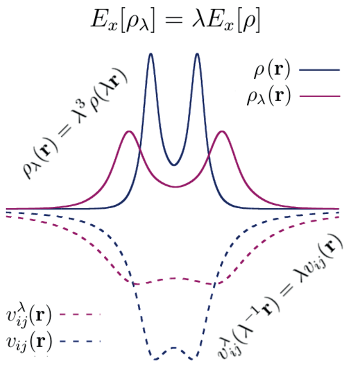
In the previous part, we have made use of xc functionals that for the most require inclusion of the exact exchange of a Kohn-Sham determinant. As we have seen in Chapter 4.1.3, this can be a costly endeavour in a plane wave basis. In this part, we present an approach that can substantially speed up the calculation of the exact exchange contribution in the case of clusters.

In the following, the use of the Fourier series representation for the reciprocal space forms of $\psi(\mathbf{G})$ and $\rho(\mathbf{G})$ shall be implicit. For notational simplicity, the bar for the Fourier series coefficients $\bar{\psi}(\mathbf{G})$ and $\bar{\rho}(\mathbf{G})$ will therefore be omitted.

CHAPTER 7

Exploiting Coordinate Scaling Relations to Accelerate Exact Exchange Calculations

Chapter 7 is a post-print version of a letter published as:



Bircher, Martin P.; Rothlisberger, U. *The Journal of Physical Chemistry Letters* **2018**, *9*, 3886–3890

For reasons of consistency, the notation used in the published article has been changed to reflect the conventions adopted throughout the present dissertation.

EXACT exchange is an important constituent in many state of the art approximations to the exchange-correlation (xc) functional of Kohn-Sham DFT. However, its evaluation can be computationally intensive, which can be particularly prohibitive in DFT-based Molecular Dynamics (MD) simulations, often restricted to semi-local functionals. Here, we derive a scheme based on the formal coordinate scaling properties of the exact xc functional that allows for a substantial reduction of the cost of the evaluation of both the exact exchange energy and potential. We show that within a plane wave/pseudopotential framework, excellent accuracy is retained, while speedups from up to ~ 6 can be reached. The coordinate-scaling thus accelerates hybrid-functional based *first-principles* MD simulations by nearly one order of magnitude.

The history of Kohn-Sham density functional theory (KS-DFT)[14,15] has been marked by the quest for increasingly accurate approximations to the exchange-correlation (xc) functional. While a computationally tractable form for the exact xc functional[75] remains, alas, elusive, a plethora of approximate forms have been developed and successfully applied over the past decades.[87]

Known properties of the exact xc functional have served as a valuable guide in the design of many an approximate functional,[114,221,222] but their use has mostly been limited to the theoretical realm of functional development[113,223–226] rather than the improvement of the computational performance in practical implementations. Here, we show how the use of a simple scaling relation[109–111] can substantially lower the computational overhead of the evaluation of exact exchange in plane waves for isolated systems. This is achieved without loss of accuracy, enabling studies of systems that have hitherto been untractable.

Among the conditions an exact functional must fulfil, the coordinate scaling relation[112,113] for the exact exchange energy is of particular simplicity. Fig. 7.1 illustrates the concept of density scaling on a hydrogen molecule. Given a (L^1-norm-conserving) electron density scaled in the coordinates by a constant λ:

$$\rho_\lambda(\mathbf{r}) = \lambda^3 \rho(\lambda \mathbf{r}), \tag{7.1}$$

where $\lambda > 1$ contracts the density and $\lambda < 1$ stretches it out, the exchange energy is homogeneous to degree one:

$$E_x[\rho_\lambda] = \lambda E_x[\rho]. \tag{7.2}$$

While a similar relation only gives an upper bound for the correlation energy,

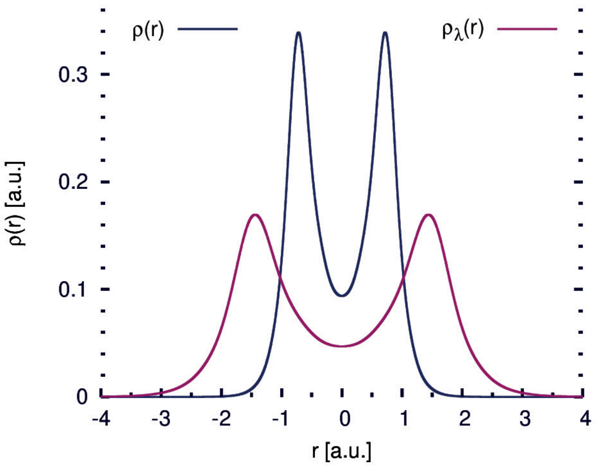


Figure 7.1: Illustration of the coordinate scaling at the example of the hydrogen molecule. Shown are a cut of the N-electron density $\rho(\mathbf{r})$ along the bonding axis, as well as its one-dimensional coordinate-scaled analogue $\rho_\lambda(\mathbf{r})$ for $\lambda = 0.5$.

in the case of the exchange functional, eq. 7.2 imposes strict boundaries on functional forms.

This scaling relation is trivially derived from the exact exchange functional, the familiar functional form of which is given as the exchange energy (or Fock exchange) of the single Slater determinant of the N_b occupied (generalised) Kohn-Sham orbitals that sum up to $\rho(\mathbf{r}) = \sum_i^{N_b} |\psi_i(\mathbf{r})|^2$:

$$E_x[\rho] = \frac{1}{2} \sum_{ij} \iint \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \; \frac{\psi_i^*(\mathbf{r})\psi_j^*(\mathbf{r}')\psi_j(\mathbf{r})\psi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{7.3}$$

Eq. 7.3 is the primordial ingredient in the family of hybrid exchange-correlation functionals[31] which combine a fraction of exact exchange with semilocal functional forms and which have seen tremendous success over the last decades.[83,227] The advent of hybrid functionals extended the applicability of density functional theory (DFT) to many chemical systems for which a description at the generalised gradient approximation (GGA) level was hampered by insufficient accuracy, substantially improving properties ranging from ionisation potentials over excitation energies to reaction enthalpies.[83]

The improved accuracy of energetics and excitation energies obtained from hybrid functionals is of particular appeal to *first principles* ground- and excited-state molecular dynamics (MD) simulations. However, due to the large computational overhead associated with the evaluation of the integral in eq. 7.3, applications using hybrid functionals are often limited by the time scale that can be simulated at an achievable computational cost.[30] The overhead due to the evaluation of the Fock integral can be particularly expensive in the plane wave/pseudopotential framework that is frequently used for *first principles* MD.[157]

Hybrid functionals are most commonly used within the generalised Kohn-Sham (GKS) scheme,[137] where the (local) Kohn-Sham optimised effective potential can be replaced by the familiar orbital-dependent form known from Hartree-Fock exchange. In a plane wave basis at the Γ-point, the gradient of the exact exchange energy with respect to every one out of N_b orbital expansion coefficients $\bar{\psi}_i(\mathbf{G})$ is then obtained from a discrete Fourier transform $\mathcal{S}$[160, 228]:

$$\frac{\partial E_x}{\partial \bar{\psi}_i^*(\mathbf{G})} = \mathcal{S}_{D^S_{\Delta\mathbf{R}}} \left[\sum_j \psi_j(\mathbf{R}) v_{ij}(\mathbf{R}) \right], \tag{7.4}$$

where we have introduced capital symbols $(\mathbf{G}, \mathbf{R})$ to denote discrete representations of continuous functions. The domain $D^S_{\Delta\mathbf{R}}$ of the discrete Fourier transform are the real space mesh points $\mathbf{R}$ spaced by $\Delta\mathbf{R}$ that are contained within the entire simulation supercell S. For notational simplicity, we have restricted $D^S_{\Delta R}$ to the one-dimensional case:

$$D^S_{\Delta R} = \{R \mid (0 \leq R < l) \wedge (R = n\Delta R,\ 0 \leq n < n_{\text{max}},\ n \in \mathbb{Z}_{\geq 0})\}, \tag{7.5}$$

where l is the edge length of the simulation supercell. A generalisation to 3D Cartesian space is straightforward. For a given supercell, the spacing of the real space grid $\Delta\mathbf{R}$ and the number of grid points n_{max} is defined by the value of the cutoff energy in reciprocal space, E_{cut}. The potential $v_{ij}(\mathbf{R})$ is determined by two discrete Fourier transforms, exploiting the fact that the Coulomb operator is diagonal in reciprocal space:

$$v_{ij}(\mathbf{R}) = \mathcal{S}^{-1}_{D^S_{\Delta\mathbf{R}}} \left[\Phi(\mathbf{G}) \left(\mathcal{S}_{D^S_{\Delta\mathbf{R}}} \left[\psi_i^*(\mathbf{R}) \psi_j(\mathbf{R}) \right] \right) \right], \tag{7.6}$$

where $\Phi(\mathbf{G})$ is an appropriately defined form of the Coulomb operator in reciprocal space $\mathbf{G}$ that eliminates the divergence at $\mathbf{G} = 0$.[158, 159] Orbital pair products can therefore be conveniently obtained in real space, whereas the associated Coulomb potential is straightforwardly computed from the diagonal reciprocal-space Coulomb operator. The computational cost for the evaluation of a given pair is only determined by the granularity of the real space mesh; a resolution $\mathbf{R}'$ halved in every Cartesian component with respect to a reference mesh $\mathbf{R}$ decreases the cost for every 3D FFT by a factor of about 8.

While for *first principles* MD, plane waves offer many advantages over localised, atom-centred bases (such as intrinsic periodicity, absence of Pulay forces and basis set superposition effects, and scaling as $N \log N$ due to the extensive use of fast Fourier transforms in the calculation of $\mathcal{S}$), following eqs 7.3, 7.4 and 7.6, practical calculations employing hybrid functionals for a system with N_b occupied KS orbitals require at least $(3N_b)^2$ additional fast Fourier transforms (FFT) with respect to a GGA. Although the use of blocking techniques can improve the effective scaling for large N_b on a large number of processors,[160] the computational cost can remain prohibitively expensive in the absence of specialised computational hardware that provides multiples of thousands of threads.

This problem is exacerbated in the case of isolated systems, which are usually required for nonperiodic systems in a full QM or QM/MM context. In order to recover the first/nearest image interaction, rather than the infinitely periodic description which is intrinsic to a plane wave expansion, the Poisson equation has to be solved under appropriate boundary conditions. A solution to this problem has, among others, been proposed by Tuckerman and Martyna[155] (TM), using an appropriate Fourier series representation for $\Phi(\mathbf{G})$. This approach requires that the simulation supercell spans twice the range of the charge density; since the cost of an FFT scales approximately cubically $(N \log N)^3$ with respect to the supercell volume, this approach carries a significant computational overhead due to large regions of the simulation supercell where the orbitals are effectively zero. For a system under the TM decoupling scheme, we will distinguish the domain of non-negligible electron density $D^{\rho}_{\Delta\mathbf{R}}$ within $D^{S}_{\Delta\mathbf{R}}$ by the superscript ρ, and its one-dimensional analogue reads:

$$D^{\rho}_{\Delta R} = \{R \mid (l/4 \leq R < 3l/4) \ \wedge \ (R = n\Delta R, \\ n_{\max}/4 \leq n < 3n_{\max}/4, \ n \in \mathbb{Z}_{\geq 0})\}. \tag{7.7}$$

In combination with the $(3N_b)^2$ FFTs required in the calculation of the exact exchange energy, this can make calculations on isolated systems prohibitively expensive for a typical number of electronic states N_b that would still be tractable in an infinitely periodic system.

In the following, based on the extension of the coordinate scaling relations to the Kohn-Sham orbitals,[17, 112] we shall demonstrate that the exact exchange energy and potential for isolated systems described by a plane wave expansion of the GKS orbitals can be straightforwardly obtained from the coordinate-scaled quantities at a substantially lower computational cost.

The homogeneous coordinate scaling of the density by λ can be generalised to the KS orbitals according to[17]:

$$\psi_i^\lambda(\mathbf{r}) = \lambda^{3/2}\psi_i(\lambda\mathbf{r}), \tag{7.8}$$

where $\lambda^{3/2}$ ensures that the norm of the N-electron density be conserved. The corresponding Coulomb potential $v_{ij}^\lambda(\mathbf{r})$ due to a scaled orbital pair $\psi_i^{\lambda*}(\mathbf{r})\psi_j^\lambda(\mathbf{r})$ expressed in terms of the unscaled generalised Kohn-Sham orbitals reads:

$$v_{ij}^\lambda(\mathbf{r}) = \int \mathrm{d}\mathbf{r}'\, \frac{\psi_j^{\lambda*}(\mathbf{r}')\psi_i^\lambda(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = \int \mathrm{d}\mathbf{r}'\, \frac{\lambda^3 {\psi_j}^*(\lambda\mathbf{r}')\psi_i(\lambda\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{7.9}$$

Introducing the coordinate transformation $\mathbf{q} = \lambda\mathbf{r}$, where the infinitesimal volume element changes according to the Jacobian of the transformation matrix, $\mathrm{d}\mathbf{r} = \lambda^{-3}\mathrm{d}\mathbf{q}$, one finds:

$$v_{ij}^\lambda(\lambda^{-1}\mathbf{q}) = \lambda \int \mathrm{d}\mathbf{q}'\, \frac{{\psi_j}^*(\mathbf{q}')\psi_i(\mathbf{q}')}{|\mathbf{q}-\mathbf{q}'|}. \tag{7.10}$$

By comparing with the Coulomb potential $v_{ij}(\mathbf{r})$ of an unscaled orbital pair $\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})$, it follows that the coordinate scaled Coulomb potential and its unscaled counterpart are related according to:

$$v_{ij}^\lambda(\lambda^{-1}\mathbf{r}) = \lambda v_{ij}(\mathbf{r}). \tag{7.11}$$

While the Kohn-Sham $\psi_i^\lambda(\mathbf{r})$ minimise T_s for a given scaled density, no such proof exists for orbitals obtained from a GKS-type potential. In GKS theory, the $\psi_i^\lambda(\mathbf{r})$ are therefore not necessarily the minimising orbitals associated to $\rho_\lambda(\mathbf{r})$. By extending eqs 7.9 through 7.11 to the exact energy functional, it is readily seen that any set of orbitals obtained from eq. 7.8 obey a relation of the type $E_x[\{\psi^\lambda\}] = \lambda E_x[\{\psi\}]$, independent of them coming from the minimising determinant of $\rho_\lambda(\mathbf{r})$. Since here, the Euler-Lagrange equation is solved on the initial, unscaled orbital space of $\psi_i(\mathbf{r})$, the $\psi_i^\lambda(\mathbf{r})$ of GKS theory are not constrained to be the minimising orbitals for a given $\rho_\lambda(\mathbf{r})$. Instead, they can be defined to be the set of orbitals that fulfil eq. 7.2, which allows for the same scaling relations to be applied to both GKS and KS orbitals. T_s itself is always calculated using the original, unscaled orbitals.

We now seek to lower the resolution of the real space grid $\mathbf{R}$ for the discrete Fourier transforms in eq. 7.6 without loss of accuracy. This is possible by exploiting the considerable volume in the supercell where the density is zero, which is imposed by the necessity for a decoupling of the Poisson equation.

We set $\lambda = 0.5$ and define the set of stretched, coordinate-scaled generalised Kohn-Sham orbitals:

$$\psi_i^\lambda(\mathbf{R}') = \lambda^{3/2}\psi_i(\lambda\mathbf{R}' + \mathbf{T}_D(\lambda)), \tag{7.12}$$

where the translation $\mathbf{T}_D(\lambda) = \frac{1}{2}(\mathbf{l} - \lambda\mathbf{l})$ ensures that the orbitals remain centred within the periodic supercell of length $\mathbf{l}$, thus preserving translational invariance. Since the representation of $\psi_i(\mathbf{R})$ is not continuous, the coordinates $\mathbf{R}'$ obtained from eq. 7.12 have to be discretised on a grid with $\Delta\mathbf{R}' = 2\Delta\mathbf{R}$. Only the non-zero domain $D^\rho_{\Delta\mathbf{R}}$ of the unscaled orbitals is now used in constructing $\psi_i^\lambda(\mathbf{R}')$. The spatial extent of the simulation supercell remains the same. The resulting coordinate-scaled generalised Kohn-Sham orbitals are now defined on the domain $D^S_{\Delta\mathbf{R}'}$ containing the points of a coarser mesh $\mathbf{R}'$ on the entire supercell S. For the one-dimensional case:

$$D^S_{\Delta R'} = \{R' \mid (0 \le R' < l) \wedge (R' = m\Delta R',\ 0 \le m < \lambda n_{\max},\ m \in \mathbb{Z}_{\ge 0})\}. \tag{7.13}$$

$v_{ij}^\lambda(\mathbf{R}')$ is then easily obtained from eq. 7.6 by performing two discrete Fourier transforms on the domain $D^S_{\Delta\mathbf{R}'}$,

$$v_{ij}^\lambda(\mathbf{R}') = \mathcal{S}^{-1}_{D^S_{\Delta\mathbf{R}'}} \left\{ \Phi(\mathbf{G}) \left[\mathcal{S}_{D^S_{\Delta\mathbf{R}'}} \left(\psi_i^{\lambda *}(\mathbf{R}')\psi_j^\lambda(\mathbf{R}') \right) \right] \right\}, \tag{7.14}$$

which reduces the energy cutoff in $\mathbf{G}$ by a factor of 8 compared to the one used for $v_{ij}(\mathbf{R})$. Based on eq. 7.11, one obtains $v_{ij}(\mathbf{R})$ from the low-resolution, scaled Coulomb potential $v_{ij}^\lambda(\mathbf{R}')$ according to:

$$v_{ij}(\mathbf{R}) = \begin{cases} \lambda^{-1} v_{ij}^\lambda(\lambda^{-1}\mathbf{R} + \mathbf{T}_D(\lambda^{-1})) & \text{for} \quad \mathbf{R} \in D^\rho_{\Delta\mathbf{R}} \\ 0 \quad \text{for} \quad \mathbf{R} \in D^S_{\Delta\mathbf{R}} \setminus D^\rho_{\Delta\mathbf{R}}, \end{cases} \tag{7.15}$$

where $\mathbf{T}_D(\lambda^{-1})$ again ensures proper centring of $v_{ij}(\mathbf{R})$ within the periodic supercell. After multiplication by $\psi_j(\mathbf{R})$, $v_{ij}(\mathbf{R})$ can be used to update the expansion coefficients $\bar{\psi}_i(\mathbf{G})$ via eq. 7.4 from the full $v_{ij}(\mathbf{R})$, which is again defined for all points of $D^S_{\Delta\mathbf{R}}$. We have implemented this approach in the CPMD code.[176]

Table 7.1 compares total energy, HOMO-LUMO gaps and the first 5 excited states obtained from the standard $v_{ij}(\mathbf{r})$ to those obtained from eq. 7.15 at the example of the organic dye DMABN. Total energies between the approaches differ by only $0.3 \cdot 10^{-4}$ kcal/mol. The forces on the atoms span a range from 10^{-2} to 10^{-5} a.u. and differ at maximum by 1%. However, 80% of the deviations are $< 0.1\%$. The HOMO-LUMO gap is identical within $2 \cdot 10^{-4}$ eV, and good accuracy is retained in the excitation energies for the first five singlet states, where the maximum deviation is observed for the S_4 state. Still, the difference between the approaches is below 0.1 eV, *i.e.* lower than 2%. The differences are not sensitive to changes in the geometry that are smaller than the grid spacing (*cf.* Supporting Information).

Practically, the performance of the algorithm described here is not only limited by the last Fourier transform in eq. 7.4, which is carried out on the full reference mesh, but also by inter-processor communication in a distributed memory framework. Table 7.2 compares execution times for the calculation of the first 5 excited states of both DMABN and a DMABN trimer at the example of the three different routines in which the exact exchange contribution has to be calculated: ODIIS[229] wavefunction optimisation to obtain the occupied (but non-diagonal) ground-state KS orbitals, the Davidson diagonalisation that yields the canonical (occupied and non-occupied) KS orbitals, and the final Davidson diagonalisation of the TDDFT matrix. In the CPMD code, the fast index of the three-dimensional real space mesh, x, is distributed over MPI tasks.[160] In order to resolve load-balancing issues, a second parallelisation layer is available, where electronic states are distributed over processor groups. All timings are given for the most efficient parallel setup (x planes *vs.* bands) for a given number of CPU cores, such that they are representative of practical applications.

For a single DMABN molecule ($N_b = 28$), an average step of DIIS optimisation is sped up by a factor of about $\eta = 2.5$, whereas the calculation of the orbital eigenvalues (diagonalisation of the non-canonical KS orbitals) is accelerated by about $\eta = 4.5$. The solution of the Tamm-Dancoff equation for the 5 first singlet states is $\eta = 3.3$ times faster when using the scaled densities. Speedups are particularly sizeable for a DMABN trimer ($N_b = 84$), where an average DIIS step is accelerated by a factor of $\eta = 4.5$. The speedup of $\eta = 4.9$ in the calculation of the canonical KS orbitals remains comparable to the monomer, whereas the diagonalisation of the TDDFT matrix is faster by a factor of $\eta = 5.7$. These considerable speedups make the density scaling approach particularly beneficial for applications where many independent trajectories have to be gathered in parallel, such as thermodynamic integration or excited-state applications like surface hopping, thus making

Table 7.1: Total energy, HOMO-LUMO gap $\Delta\epsilon_{\text{gap}}$ and singlet excitation energies for the organic dye DMABN obtained from the standard $v_{ij}(\mathbf{r})$ and its scaled counterpart presented in this work (eq. 7.15; $\lambda = 0.5$). Calculations were carried out in an orthorhombic $37.5 \times 18.75 \times 18.75$ Å^3 supercell at a plane wave cutoff of 75 Ry using the B3LYPxc functional.

	E_{tot} [a.u.]	$\Delta\epsilon_{\text{gap}}$ [eV]	S_1 [eV]	S_2 [eV]	S_3 [eV]	S_4 [eV]	S_5 [eV]
Full $v_{ij}(\mathbf{r})$	-76.689581	4.7690	4.410	4.465	4.752	4.992	5.080
Scaled $v_{ij}(\mathbf{r})$	-76.689586	4.7588	4.408	4.504	4.753	5.080	5.140
Difference	$5.0 \cdot 10^{-6}$	0.0002	0.002	0.039	0.001	0.088	0.060

Table 7.2: Execution time for optimisation (ODIIS) and diagonalisation (Davidson) of the ground-state KS wavefunction, as well as the diagonalisation of the TDDFT matrix (LR) for both a monomer and a trimer of DMABN (28 and 84 occupied states, respectively). Values averaged over 10 iterations are given for both the full $v_{ij}(\mathbf{r})$ (Ref) and its scaled counterpart (λ) and have been used to calculate the speedup η. Calculations were carried out in an orthorhombic $25.0 \times 12.5 \times 12.5$ Å^3 supercell at a plane wave cutoff of 75 Ry using the B3LYPxc functional on 256 CPU cores. Optimal parallel configurations are compared. The values for the full $v_{ij}(\mathbf{r})$ were obtained by using 4 CP groups and a block size of 4 in the calculation of the exact exchange. The calculations using the scaled $v_{ij}(\mathbf{r})$ were carried out using 4 CP groups and a block size of 14. For details on the parallelisation, see Ref. 160.

	DMABN			$(\text{DMABN})_3$		
Routine	t_{ref} [s]	t_λ [s]	η	t_{ref} [s]	t_λ [s]	η
ODIIS	3.94	1.59	2.50	34.28	7.62	4.50
Davidson	13.63	3.02	4.52	110.00	22.08	4.98
LR	85.00	25.43	3.34	976.30	170.22	5.74

optimal use of the computing resources available.

Here, we have shown how the simple coordinate-scaling relations of Kohn-Sham DFT can be applied in practical calculations, offering a substantial performance benefit over conventional approaches. The calculation of the orbital-pair Coulomb potential on the scaled orbitals allows for a much lower cutoff energy in the discrete Fourier transforms, thereby significantly reducing the computational cost. Sizeable speedups can be achieved, and the coordinate-scaled calculation of exact exchange can increase the available timescale or the number of independent trajectories that can be run within a given time by a factor of up to ~ 6, which significantly increases both sampling accuracy and efficiency with respect to current state of the art plane wave/pseudopotential calculations.

Acknowledgements

MPB is indebted to Simone Meloni for all of the fruitful discussions on plane wave algorithms and exact exchange approximations. UR gratefully acknowledges financial support from the Swiss National Science Foundation Grant No. 200020-146645 and the NCCR MUST.

Chapter 8

From a Week to Less Than a Day: Speedup and Scaling of Coordinate-Scaled Exact Exchange Calculations in Plane Waves

Chapter 8 is a pre-print version of an article submitted to:

Bircher, Martin P.; Rothlisberger, U. *Computer Physics Communications* **submitted**

EXACT exchange is a primordial ingredient in Kohn-Sham Density Functional Theory based Molecular Dynamics (MD) simulations whenever thermodynamic properties, kinetics, barrier heights or excitation energies have to be predicted with high accuracy. However, the cost of such calculations is often prohibitive, restricting the use of *first principles* MD to (semi-)local density functionals, in particular in a plane wave basis. We have recently proposed the use of coordinate-scaled orbitals to reduce the cost of the most expensive Fourier transforms during the calculation of the exact exchange potential of isolated systems. Here, we present the implementation and parallelisation of this coordinate scaling approach in the `CPMD` code and analyse its performance under different parallelisation schemes. We show that speedups gather with system size and that with an optimal configuration, speedups of up to one order of magnitude are possible with respect to conventional calculations. Simulations that have previously taken one week can therefore be finished within less than a day.

8.1 Introduction

Chemical processes take place at finite temperature and, to a large majority, in the condensed phase. The typical temperature ranges of chemical and biological processes imply that entropy can have a substantial influence on the outcome and time-scale of a chemical reaction. In solution, the behaviour of the reactants is influenced by their local environment and the specific reaction conditions. Taking all these effects into account calls for a statistical approach that goes beyond the picture of a few molecules interacting in the gas phase.

Following the ergodic hypothesis, the phase-space average in a partition function can be replaced by a time-average that can be conveniently obtained by simulating the time-evolution of the system in a molecular dynamics (MD) scheme. Since the very first calculations in the 1950ies,[57] the simulation of the dynamics of condensed matter has become one of the pillars of modern-day computational chemistry.[21,22,72,231] The introduction of the extended electronic Lagrangian by Car and Parrinello[232] in 1985 made it possible to simulate (reactive[22,33]) (bio-)chemical events from *first principles* by exploiting on-the-fly calculated potential energy surfaces (PES) obtained from Kohn-Sham density functional theory (DFT).[14,15] Ever since, the increased performance and availability of computational resources have made it possible to go beyond ground-state simulations through TDDFT[79] based excited-state approaches such as Tully's trajectory surface hopping

(TSH),[27,66] where a swarm of trajectories are propagated on excited-state surfaces. Such calculations do call for appropriate underlying computational methods and suitable algorithms.[16]

In Kohn-Sham DFT, the accuracy of a PES for a given system is determined by the model for the exchange-correlation (xc) functional that is adopted. While xc approximations have become more and more accurate, and their predictive power has considerably increased over the last decades,[83] the fundamental quantities used to model the quantum nature of the electrons have mostly been the same since the end of the 1990ies.[87] Most modern day exchange-correlation functionals $E_{xc}[\rho]$ are hybrid[31] functionals that consist of a (semi-)local part which directly depends on the density, as well as a nonlocal contribution due to the determinantal exchange of the set of (generalised) Kohn-Sham orbitals $\{\psi_i\}$ which is commonly referred to as exact or Hartree-Fock exchange (HFX):

$$E_{\mathrm{xc}}[\rho] = E_{\mathrm{xc}}^{\mathrm{loc}}[\rho] + E_{\mathrm{xc}}^{\mathrm{exact}}[\{\psi_i\}], \tag{8.1}$$

where, for a spin-restricted system, $E_{\mathrm{xc}}^{\mathrm{loc}}[\rho]$ takes the general form:

$$E_{\mathrm{xc}}^{\mathrm{loc}}[\rho] = \int \mathrm{d}\mathbf{r}\; e_x(\rho(\mathbf{r}), |\nabla\rho(\mathbf{r})|, \tau(\mathbf{r})) + \int \mathrm{d}\mathbf{r}\; e_c(\rho(\mathbf{r}), |\nabla\rho(\mathbf{r})|, \tau(\mathbf{r})). \tag{8.2}$$

e_x and e_c are the exchange and correlation energy densities, respectively, that can depend on the density $\rho(\mathbf{r})$ alone (local density approximation or LDA[15]), the density and its gradient $\nabla\rho$ (generalised gradient approximation or GGA[126]) or the density, its gradient and the kinetic energy density of the i Kohn-Sham orbitals $\tau(\mathbf{r}) = \sum_i |\nabla\psi_i(\mathbf{r}|^2$ (meta-GGA[127,129–131]). These quantities are all easily expanded and evaluated as a sum on a real space grid. The exact exchange contribution, on the other hand, involves the computation of nonlocal integrals over generalised Kohn-Sham (GKS) orbitals. It reads:

$$E_{\mathrm{x}}^{\mathrm{exact}}[\{\psi_i\}] = \frac{c}{2}\sum_{ij} \iint \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\; \psi_i^*(\mathbf{r})\psi_j^*(\mathbf{r}')\hat{\mathrm{W}}(\mathbf{r},\mathbf{r}')\psi_j(\mathbf{r})\psi_i(\mathbf{r}'); \tag{8.3}$$

$\hat{\mathrm{W}}(\mathbf{r},\mathbf{r}')$ is a general[38,141] two-electron operator. In the case of conventional hybrid functionals, it takes the form of the classical Coulomb operator $\hat{\mathrm{W}}(\mathbf{r},\mathbf{r}') = |\mathbf{r}-\mathbf{r}'|^{-1}$, where c denotes the fraction of exact exchange in the functional that has to be suitably counter-balanced in the functional form of $e_x(\mathbf{r})$. The inclusion of $E_{\mathrm{xc}}^{\mathrm{exact}}$ has been shown to be primordial for the accurate prediction of excited states,[140] and it can also substantially improve the description of ground-state properties.[31,83,163] However, the evaluation of the exact exchange energy and potential is cumbersome and expensive in particular in the plane wave basis commonly employed in condensed matter computer codes.[157,158,160] We have recently proposed a scheme that substantially reduces the computational cost of this procedure for isolated molecules and clusters[220] by resorting to the exact coordinate-scaling properties of the exchange functional.

The *first principles* MD program `CPMD`[176] is a direct successor of the seminal code developed by Roberto Car and Michele Parrinello. It has been developed at IBM Zurich Research Laboratory and by the `CPMD` consortium since the early 1990ies and has been considerably refactored before the release of the current version, 4.1. In the following, we present an efficient implementation of our recently devised coordinate-scaling scheme within the existing exact exchange driver[160] and give a detailed assessment of its performance and speedups with respect to reference algorithms. With the extended driver, it becomes possible to treat systems with hundreds of orbitals at a significantly lowered computational cost.

The paper is organised as follows: First, we give an account of the algorithmics of plane wave Kohn-Sham DFT. We will pay particular attention to the use of Fast Fourier Transforms (FFT) and their parallelisation in the `CPMD` code. We shall then describe the algorithm specific to the existing exact-exchange driver implemented by Weber *et al.*[160] and discuss typical parallelisation issues, before giving an account of our new, parallel implementation of the coordinate-scaled exact exchange. We will finally provide an analysis of the performance of both full and coordinate-scaled exact exchange at different levels of parallelisation and we will show that for hundreds of orbitals, sizeable reductions of CPU time by almost one order of magnitude can be achieved by optimally tuning the parallelisation of the coordinate-scaled approach.

8.2 Fast Fourier Transforms in Plane-Wave Kohn-Sham DFT

The ascent of *first principles* MD is not at least due to the development of highly efficient algorithms that scale beyond thousands of processors.[189,190] Since in MD, time-averages replace the ensemble averages over the partition function, long enough simulations have to be performed in order to ensure that the observables converge properly.[20,30,72] This can only be achieved if the execution time spent in the generation of the on-the-fly PES is sufficiently low to allow for sufficiently long simulations to be performed. Such a goal can be achieved in three ways: By ensuring that the algorithms scale well with system size, by employing algorithms with the lowest possible execution time, and by exploiting algorithms that parallelise well over a large number of processors.

In this respect, a particularly appealing approach lies in the use of a plane wave basis, where the (generalised) Kohn-Sham orbitals are expanded according to[16]

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{G}}^{\mathbf{G}_{\max}} \psi_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{8.4}$$

where $\psi_i(\mathbf{G})$ is the expansion (Fourier series) coefficient of the orbital i at a given reciprocal space vector $\mathbf{G}$. $\mathbf{G}_{\max}$ is commonly expressed in terms of the orbital cutoff energy $E_{\text{cut}}^{\psi} = 0.5\mathbf{G}_{\max}^{-2}$. As it is easily seen from eq. 8.4, it is straightforward to switch representation from a 3D array $\psi_i(\mathbf{r})$ to its reciprocal space counterpart $\psi_i(\mathbf{G})$ by means of Fast Fourier Transforms (FFTs), $\psi_i(\mathbf{r}) = \text{FFT}^{-1}[\psi_i(\mathbf{G})]$. In practice, 3D FFTs are resolved as a series of one-dimensional FFTs, which results in a scaling of $\mathcal{O}(N\log N)$ where N denotes the number of 3D vectors $\mathbf{G}$ used in eq. 8.4. This quantity depends both on the cutoff energy E_{cut} as well as on the size of the supercell in which the periodic Fourier series is expanded. Many highly efficient implementations of the FFT algorithm exists.

With the possibility of treating both real- and reciprocal space forms of ψ and, by extension, ρ, implementations of (semi-local) KS-DFT where the performance determining steps scale linearly with system size (or, equivalently, orbital number N_b) become straightforward. In particular, both the gradient of the electron density as well as its classical Coulomb potential $v_J(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \; |\mathbf{r}'|^{-1}\rho(\mathbf{r}-\mathbf{r}')$ are easily obtained from the reciprocal space density $\rho(\mathbf{G})$:

$$v_J(\mathbf{r}) = \text{FFT}^{-1}\left[\frac{4\pi}{\mathbf{G}^2}\rho(\mathbf{G})\right], \tag{8.5}$$

$$\nabla\rho(\mathbf{r}) = \text{FFT}^{-1}\left[i\mathbf{G}\rho(\mathbf{G})\right]. \tag{8.6}$$

The cutoff E_{cut}^{ρ} at which the density is expanded is commonly $\zeta = 4$ times bigger than E_{cut}^{ψ} itself, since the orbital products in $\rho(\mathbf{r}) = \sum_i^{N_b} |\psi_i(\mathbf{r})|^2$ generate higher frequency components that cannot be accurately described if $E_{\text{cut}}^{\rho} = E_{\text{cut}}^{\psi}$ (the case $\zeta = 1$). In practice, the reciprocal space density used in eqs 8.5 and 8.6 is constructed from the arrays of wavefunction coefficients

$\psi(\mathbf{G})$ by a series of FFTs:

$$\rho(\mathbf{G}) = \mathrm{FFT}_{E^{\rho}_{\mathrm{cut}}} \left[\sum_{i}^{N_b} \left| \mathrm{FFT}^{-1}_{E^{\psi}_{\mathrm{cut}}} \left[\psi_i(\mathbf{G}) \right] \right|^2 \right], \tag{8.7}$$

which, for a given size of the periodic supercell, results in a linear scaling with the number $N_{E^{\psi}_{\mathrm{cut}}}$ of electrons for the Fourier transforms at the wavefunction cutoff E^{ψ}_{cut}, whereas the number $N_{E^{\rho}_{\mathrm{cut}}}$ of (more expensive) Fourier transforms at the density cutoff E^{ρ}_{cut} is constant. The overall scaling due to the FFTs is therefore of the order $\mathcal{O}(N_b(N_{E^{\psi}_{\mathrm{cut}}} \log N_{E^{\psi}_{\mathrm{cut}}}) + c(N_{E^{\rho}_{\mathrm{cut}}} \log N_{E^{\rho}_{\mathrm{cut}}}))$ where c is a constant. For sufficiently large N_b, the algorithm therefore scales linearly with the number of Kohn-Sham orbitals and exhibits the typical FFT scaling for the number of basis vectors. Below the orders of thousands of atoms, where the cubic scaling of the wavefunction orthogonalisation may become dominant, the FFTs constitutes the main bottleneck of the calculation. Therefore, even if FFTs scale favourably with system size, their efficient parallelisation is primordial if low execution times are to be achieved.

In the `CPMD` code, distributed memory (MPI) parallelisation of the FFTs is straightforward, since the 3D FFTs are in practice decomposed into a series of 1D FFTs. Every task holds all y and z components of the real space mesh in the periodic supercell, but only a part of the total of $i_x^{\max}$ planes. In the following, and in the style of Fortran, we will use colons between delimiters to denote a range of elements, while a colon alone will denote the entire extent of a matrix (or an array). The array holding the density is effectively distributed over N_P MPI tasks m:

$$\rho(i_x^{\mathrm{low}} : i_x^{\mathrm{up}}, i_y^{\min} : i_y^{\max}, i_z^{\min} : i_z^{\max}) = \sum_{\mathrm{p}}^{N_P} \rho(i_{x_p}^{\mathrm{low}} : i_{x_p}^{\mathrm{up}}, i_y^{\min} : i_y^{\max}, i_z^{\min} : i_z^{\max}), \tag{8.8}$$

i.e. every MPI task holds $i_{x_p}^{\mathrm{up}} - i_{x_p}^{\mathrm{low}} + 1$ FFTs out of the total $i_x^{\max}$ yz planes (called x-pencils) in real space.[16] Optimal load balancing is recovered if $i_{x_p}^{\mathrm{up}} - i_{x_p}^{\mathrm{low}}$ is constant for every single task m. This restricts the numbers of processors N_P that can be used for a given real space mesh in to $x = N_P(i_{x_m}^{\mathrm{up}} - i_{x_m}^{\mathrm{low}} + 1)$. Consequently, for $N_P > i_x^{\max}$ or for a number of available CPU that is not a divisor of $i_x^{\max}$, this straightforward parallelisation is quickly exhausted (a typical mesh in modern-day *first principles* MD may comprise between 256 and 712 mesh points per direction, depending on E^{ρ}_{cut}. Additionally, speedups may be limited at large numbers of processors due to the overhead of the all-to-all communication that is due to the decomposition of the 3D FFT into a 1D series.[160]

Instead, Weber *et al.*[160] have implemented a processor group strategy called CP groups. For a total of N_{CPU} processor cores, every single c out of N_C CP groups unites N_P processors that replicate the parallelisation strategy of eq. 8.8:

$$\rho(i_x^{\text{low}} : x^{\text{up}}, i_y^{\text{min}} : i_y^{\text{max}}, i_z^{\text{min}} : i_z^{\text{max}}) = \sum_{\text{c}}^{N_C} \sum_{\text{p}}^{N_P} \rho_g(i_{x_p}^{\text{low}} : i_{x_p}^{\text{up}}, i_y^{\text{min}} : i_y^{\text{max}}, i_z^{\text{min}} : i_z^{\text{max}}). \tag{8.9}$$

Each CP group with index c still holds N_P x-pencils, *i.e.* the delimiters $i_{x_p}^{\text{low}} : i_{x_p}^{\text{up}}$ remain the same for every pth member across all N_C groups, as well as all N_b orbitals. Instead, only a part of the electron density is computed by each CP group, corresponding to a partial sum over (generalised) Kohn-Sham orbitals:

$$\rho_c(\mathbf{G}) = \text{FFT}_{E_{\text{cut}}^{\rho}} \left[\sum_{i=i_{\text{low}}}^{i_{\text{up}}} \left| \text{FFT}_{E_{\text{cut}}^{\psi}}^{-1} \left[\psi_i(\mathbf{G}) \right] \right|^2 \right], \tag{8.10}$$

$$\rho(\mathbf{G}) = \sum_{c}^{N_C} \rho_c(\mathbf{G}), \tag{8.11}$$

where $N_C(i_{\text{up}} - i_{\text{low}} + 1) = N_b$. For (semi-)local functionals, inter-group communication is only necessary once the total $\rho(\mathbf{r})$ is assembled, which - as a reduction over all processors - only carries a small overhead This allows for an optimal, balanced distribution over the real-space grid due to the FFTs. While a distribution of gridpoints alone can result in load-balancing issues and is exhausted once $x_{\text{max}} = N_{\text{CPU}}$, a distribution of the FFTs for the N_b electronic states allows for an optimal trade off between the numbers of x-pencils per FFT and the cost of communication. In contrast to earlier parallelisation schemes,[233] the reciprocal space representations $\{\psi(\mathbf{G})\}$ of all N_b states remain replicated across all processors, but are not used for the FFTs of a given group.

It is particularly important that load balancing issues be avoided if exact exchange energies and potentials have to be calculated. The calculation of the exact exchange contribution requires the computation of a correlation function and is most conveniently carried out in a mixed real/reciprocal space approach similar to the computation of the Coulomb potential.[157] The exact exchange energy in eq. 8.3 is then given by:

$$E_x^{\text{exact}}[\{\psi\}] = -\frac{1}{2}\sum_i^{N_b}\sum_j^{N_b}\sum_{\mathbf{G}}^{\mathbf{G}_{\max}} \Phi(\mathbf{G})\,|\rho_{ij}(\mathbf{G})|^2\,, \tag{8.12}$$

where $\Phi(\mathbf{G})$ is an appropriately generalised reciprocal space form of a two-electron operator $\hat{\mathrm{W}}$[158,159] and we have introduced the pair densities ρ_{ij} as products of GKS orbitals:

$$\rho_{ij}(\mathbf{r}) = \left(\mathrm{FFT}^{-1}_{E^{\psi}_{\text{cut}}}\left[\psi_i(\mathbf{G})\right]\right)^* \mathrm{FFT}^{-1}_{E^{\psi}_{\text{cut}}}\left[\psi_j(\mathbf{G})\right], \tag{8.13}$$

$$\rho_{ij}(\mathbf{G}) = \mathrm{FFT}_{E^{\rho}_{\text{cut}}}\left[\rho_{ij}(\mathbf{r})\right]. \tag{8.14}$$

The potential acting on an expansion coefficient $\psi(\mathbf{G})$ is obtained from:

$$\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_i^*(\mathbf{G})} = \sum_j^{N_b} \mathrm{FFT}_{E^{\psi}_{\text{cut}}}\left[\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_{i(j)}^*(\mathbf{r})}\right], \tag{8.15}$$

and

$$\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_{i(j)}^*(\mathbf{r})} = \psi_j(\mathbf{r})\left(\mathrm{FFT}^{-1}_{E^{\rho}_{\text{cut}}}\left[\Phi(\mathbf{G})\rho_{ij}(\mathbf{G})\right]\right). \tag{8.16}$$

From eqs 8.12 to 8.16 it is easily seen that the computation of the exact exchange energy and potential is the most expensive part of a calculation involving a hybrid functional. In a straightforward implementation, where the real-space GKS orbitals are not stored, N_b^2 FFTs at the wavefunction cutoff (eq. 8.13) and N_b^2 FFTs at the density cutoff (eq. 8.12) are required. In order to determine the potential, an additional N_b^2 FFTs at the density cutoff (eq. 8.16) and N_b^2 FFTs at the wavefunction cutoff (eq. 8.15) are necessary. According to the $\mathcal{O}N\log N$ scaling of the FFTs and by exploiting symmetry in the sum of eq. 8.12 that effectively reduces the number of pairs by one half, the bottleneck is given by a total of N_b^2 FFTs at the density cutoff.

Due to the large number of additional FFTs at the density cutoff, load-balancing issues due to all-to-all communication can become particularly prohibitive. The CP group approach by Weber *et al.* can remedy this problem if the parameters controlling the parallelisation N_C and N_P are properly balanced. Further performance gains can then be made by blocking the orbitals in the calculation of eq. 8.12 and storing the inverse Fourier transforms of the GKS orbitals: By measuring the load-balancing on the fly, blocks of

$b^2 = (i_{\text{low}} : i_{\text{up}}) \times (j_{\text{low}} : j_{\text{up}})$ orbitals are distributed over CP groups, and a subset $\{\psi(\mathbf{r}\}$ with indices i_{low} to i_{up} and j_{low} to j_{up} is precomputed on every block and stored. Then, the associated b^2 FFTs for all pairs ρ_{ij} constructed for a given subset are performed, followed by $i_{\text{low}} : i_{\text{up}}$ forward FFTs that yield the potential in eq. 8.15. All these operations can be carried out independently for different subsets on every CP group. The total potential is then given by a sum over the contributions of all CP groups:

$$\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_i^*(\mathbf{G})} = \sum_c^{N_C}\sum_j^{N_b} \text{FFT}_{E_{\text{cut}}^{\psi}}\left[\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_{i(j)}^*(\mathbf{r})}\right]. \tag{8.17}$$

Weber *et al.* have shown that the resulting algorithm scales up to thousands of processors and yields unprecedented performance gains.[160] For sufficiently large numbers of states N_b, they have reported linear scaling with respect to the FFTs at E_{cut}^{ρ} by the use of thresholding algorithms: By exploiting Wannier centres in systems with a finite band gap, it becomes possible to screen orbital pairs that have a negligible overlap.

8.3 Coordinate-Scaled Exact Exchange

For isolated systems such as clusters and quantum mechanical systems treated within a classical environment (QM/MM),[72] we have recently reported a scheme that can substantially lower the computational cost of the exact exchange computation.[220] For a (pair) density stretched by a scaling factor $\lambda < 1$, $\rho_\lambda(\mathbf{r}) = \lambda^3\rho(\lambda\mathbf{r})$, the coordinate scaling relations for the exact exchange functional impose that:

$$E_x^{\text{exact}}[\rho_\lambda] = \lambda E_x^{\text{exact}}[\rho]. \tag{8.18}$$

Using the decoupling scheme of Martyna *et al.* to isolate the system from its periodic replicas requires that the length of the simulation supercell span twice the extent of the charge density,[155] resulting in large regions of space where the density is effectively naught. We have advocated to exploit this empty space by applying the coordinate scaling of the exact exchange functional to the computation of the exact exchange energy and the potential. If orbitals and potential are stretched by a scaling factor $\lambda < 1$, this results in smoothened functions that can be described on a real space mesh much coarser than the unscaled reference. Choosing $\lambda = 0.5$ allows to optimally exploit the size of a supercell in a cluster, while the Fourier transforms needed to compute the Coulomb potential of pairs ρ_{ij}^{λ} can be performed at an auxiliary cutoff energy $E_{\text{cut}}^{\rho(\lambda)} = E_{\text{cut}}^{\rho}/4$. The reduction of the cutoff energy by a factor of 4 allows for the real-space mesh sampling Δx to be halved, $\Delta X = 2\Delta x$. Applying coordinate-scaling, we have demonstrated that total ground-state energies, generalised Kohn-Sham orbital eigenvalues and excitation energies obtained from linear-response TDDFT are in excellent agreement with calculations that employ the full, standard exact exchange, while preliminary speedups of up to ≈ 6 have been achieved.

8.4 Implementation of Coordinate-Scaled Exact Exchange

The use of the coordinate scaling relations requires a set of stretched GKS orbitals[17]:

$$\psi_i^{\lambda}(\mathbf{r}) = \lambda^{3/2}\psi_i(\lambda\mathbf{r}), \tag{8.19}$$

which have to be computed in real space and therefore require at least N FFT at the usual wavefunction cutoff E_{cut}^{ψ} (*cf.* eq. 8.13. The coordinate-scaled pair densities $\rho_{ij}^{\lambda}(\mathbf{r}) = \psi_i^{\lambda}(\mathbf{r})\psi_j^{\lambda}(\mathbf{r})$ are then used in eqs 8.14 and 8.16, which allows for the FFTs to be computed on a coarse grid with $E_{\text{cut}}^{\rho(\lambda)}$ for the scaled potential:

$$v_{ij}^{\lambda}(\mathbf{r}) = \lambda v_{ij}(\lambda\mathbf{r}). \tag{8.20}$$

The full exact exchange potential has to be used to obtain accurate derivatives of the expansion coefficients in eq. 8.15, *i.e.* the last sets of Fourier transforms have to be performed at the standard wavefunction cutoff again. To this end, the full potential can be obtained by inversing the relation given in eq. 8.20. The coordinate scaling scheme can be implemented based on the existing exact exchange driver, which makes it possible to exploit its high level of efficiency. All extensions to the existing exact exchange driver were written in accordance with the Fortran2003 standard.

In the following, we shall denote the set of discrete Cartesian coordinates by x, y, z in the case of the standard (fine) mesh associated to E_{cut}^{ρ} and use X, Y, Z for the corresponding value on the coarse mesh of $E_{\text{cut}}^{\rho(\lambda)}$. The grid spacing of both meshes is related by $\Delta X = 2\Delta x$, and the maximum numbers of elements are therefore $i_X^{\max} = i_x^{\max}/2$ *etc.* . The grid index of a specific discrete coordinate will be $i_x = x/\Delta x + 1$ and $i_X = X/\Delta X + 1$, *i.e.* the lower boundary for the grid indices is 1, as it is common in Fortran. Both meshes have their origin at $X = x = 0$ with $i_x = i_X = 1$. Therefore, all points of the coarse grid have an equivalent on the fine grid, $X = x$, only if $x = 2\Delta x + 1$. Points on the fine mesh with $x' = 2\Delta x$ are not described on the coarse mesh. Choosing a scaling parameter of $\lambda = 0.5$ doubles the spatial extent of a given (pair) density, which allows for a straightforward mapping between the grids, as outlined in figure 8.1.

After computation of a block b of $i_{\text{low}} : i_{\text{up}}$ and $j_{\text{low}} : j_{\text{up}}$ orbitals $\{\psi(\mathbf{r})\}$ from their reciprocal space form $\{\psi(\mathbf{G})\}$ using the standard inverse FFT on the wavefunction grid, the scaled orbitals are created. This is done by first reducing the y and z dimension of all xyz planes that a given MPI task holds into an auxiliary orbital ψ_{scr} of dimension $\psi(i_x^{\max}, i_Y^{\max}, i_Z^{\max})$:

$$\psi_{\text{scr}}(i_x, i_y, i_z) = \lambda^{3/2}\psi(i_x, i_y + c_z, z + c_z), \tag{8.21}$$

where c_y and c_z denote shifts of indices (*cf.* Fig. 8.1), $c_y = l_y/(4\Delta y)$, based on the length of the supercell edge l_y. Only after this step are the x_p-components of the mesh redistributed within the N_P tasks of a given group. In order to minimise communication in this step, we have mapped the existing distributed memory parallelisation of the `CPMD` code to an easily accessible structure that suits the needs of grid-redistribution as shown in Fig. 8.1.

All N_P MPI tasks associated to a given CP group are organised in *receiver* and *sender* groups, in ascending order of the running indices i_X, i_x. Given two equidistributed grids with a total of $x_{\max}$ and $X_{\max}$ elements, respectively, every MPI task will hold $n_x = N_C(x_{\max}/N_P)$ and $n_X = N_C(X_{\max}/N_P)$ yz-planes. For every CP group, the first sender/receiver pair carries the ID of the processor that will hold the element $i_X^r = 1$ ($r$eceiver) as well as the ID of the processor that sends the element $i_x = c_x^s$ ($s$ender).

$$\psi_\lambda(i_X^r, :, :) = \psi_{\text{scr}}(i_x^s, :, :). \tag{8.22}$$

Communication is managed by the standard MPI routines `mpi_send` and `mpi_recv`. At the example of Fig. 8.1, the first sender/receiver group $sr = (r, s)$ is given by the MPI sender III and the receiver I; the second group will be formed of sender III and receiver II, *etc.*İf possible, one send/receive call does not send single yz-pencils, but instead sends the maximum number of X components that can be transferred between a given sender/receiver pair at once. In the example of the mesh of Fig. 8.1, this corresponds to a pair of two yz planes that can be sent together, which can substantially reduce the communication overhead:

$$\psi_\lambda(i_{X_{\text{low}}}^r : i_{X_{\text{up}}}^r, :, :) = \psi_{\text{scr}}(i_{x_{\text{low}}}^s : i_{x_{\text{up}}}^s, :, :). \tag{8.23}$$

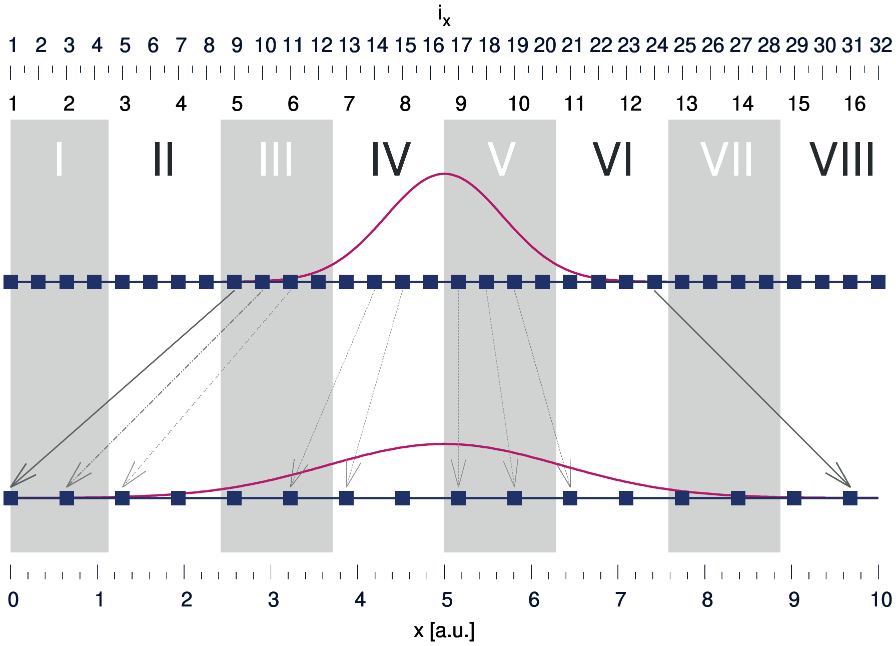


Figure 8.1: Illustration of the coordinate scaling approach at the example of the x-pencils over which the FFTs are parallelised. Every index i_x in x corresponds to an entire yz plane of the periodic supercell. An example of some (pair) density is shown in the upper part, its coordinate-scaled analogue is depicted below. Points correspond to discretised real-space mesh points indexed by i_x. The choice of $\lambda = 0.5$ allows for a straightforward mapping between the original density and its coordinate-scaled counterpart. The MPI parallelisation is visualised in the form of N_P white and grey blocks, which correspond to one MPI task each. In the case of the reference mesh, every task holds 4 yz planes, whereas there are only 2 yz planes for the coarse, coordinate-scaled mesh. If the non-vanishing parts of the density are distributed on the coordinate-scaled, coarser mesh shown in the lower half of the figure, different tasks have to communicate with each other. By moving from left to right, it can be seen that the first two tasks (I and II) send no data, whereas they receive density from task III. Task III in turn sends its data to both tasks I (first two gridpoints) and II (subsequent two gridpoints). Tasks IV and V will retain parts of their gridpoints; all the others completely swap their data. Once the full potential $v_{ij}(x, y, z)$ has to be obtained from its scaled counterpart $v_{ij}^{\lambda}(X, Y, Z)$, the mapping is inverted. If CP groups are used, this mapping can be carried out independently on all N_C groups without any inter-group communication.

As soon as either the ID of the sender or the receiver changes, the index of the pair sr is incremented, and the next pair accordingly treated. The mapping is structured in an incremental fashion with respect to the mesh, rather than in an ascending order of processor IDs (their ordering coincides in Fig. 8.1, but does not necessarily have to). Therefore, the manipulations can be performed on the same arrays without running the risk of overwriting data that is needed in the construction of the scaled orbitals (all data that is overwritten consists of either points where the density is naught, or points that have already been transferred to a receiving process), since the indices accessed obey $i_x \geq i_X$. The structure that holds the mapping can be generated once at the beginning of the calculation, after the two Fourier pools at E_{cut}^{ρ} and $E_{\text{cut}}^{\rho(\lambda)}$ have been initialised; this operation therefore comes at no overhead with respect to the total execution time of the program.

All $\rho_{ij}^{\lambda}(\mathbf{r})$ of a given block can then be constructed from the orbitals $\psi_{i,j}^{\lambda}(\mathbf{r})$. After the forward FFT at $E_{\text{cut}}^{\rho(\lambda)}$, multiplication with the reciprocal space Coulomb operator and an according inverse FFT, the scaled Coulomb potential is first multiplied with the scaled orbital $\psi_j^{\lambda}(\mathbf{r})$. In the original algorithm, a forward FFT on E_{cut}^{ψ} then follows for every pair, and the sum over all pairs ij for a given state i is carried out on the reciprocal space form of $\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_{i(j)}^*(\mathbf{G})}$. Since the forward FFTs at the wavefunction cutoff are not the bottleneck of the calculation, this helps to save memory.

If employed with the coordinate scaling approach, however, this would create a large communication overhead, as the computation of $\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_{i(j)}^*(\mathbf{G})}$ has to be carried out at the full wavefunction cutoff E_{cut}^{ψ} for reasons of accuracy. Instead, following the generation of $\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_{i(j)}^*(\mathbf{r})}$ on the coarse grid, the quantity for a given state i is saved, which results in the storage of a total of N_b real-space potentials to which new pairs ij are added. Instead of eqs 8.15 and 8.16, the quantity

$$\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_{i(c)}^*(\mathbf{r})} = \sum_{j\ ;\ ij\in c}^{N_b} \psi_j(\mathbf{r}) \left(\text{FFT}_{E_{\text{cut}}^{\rho}}^{-1}\left[\Phi(\mathbf{G})\rho_{ij}(\mathbf{G})\right]\right) \tag{8.24}$$

is stored on every CP group c. Since this quantity is still expanded on the coarse mesh, it requires a factor of 2^3 less storage than its full counterpart, making the additional memory requirements tractable. The intermediate storage allows for the density to be mapped back onto the initial, fine mesh only after all pair potentials have been obtained in real space; the term $\frac{\delta E_x^{\text{exact}}[\rho]}{\delta\psi_i^*(\mathbf{r})}$ is updated on-the-fly by adding freshly computed pair potentials ij

to the existing sum. This reduces the computational overhead due to processor communication considerably. The points of the fine grid are generated by reversing the mapping loop used for the generation of the scaled densities; all points for which there is no data on the coarse grid are set to zero. Then, the next block of orbitals with indices $i'_{\text{low}} : i'_{\text{up}}$ and $j'_{\text{low}} : j'_{\text{up}}$ is treated in the same manner, until all b blocks have been treated.

After the iteration over all blocks is completed, the N_b real-space arrays of eq. 8.24 are subjected to a final FFT on every CP group. Finally, just like in the conventional algorithm the total $\frac{\delta E_x^{\text{exact}}[\rho]}{\delta \psi_i^*(\mathbf{G})}$ is obtained by summing over all N_C groups:

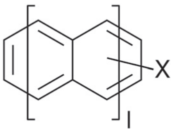
$$\frac{\delta E_x^{\text{exact}}[\rho]}{\delta \psi_{i(c)}^*(\mathbf{G})} = \text{FFT}_{E_{\text{cut}}^{\psi}} \left[\frac{\delta E_x^{\text{exact}}[\rho]}{\delta \psi_{i(c)}^*(\mathbf{r})} \right], \tag{8.25}$$

$$\frac{\delta E_x^{\text{exact}}[\rho]}{\delta \psi_i^*(\mathbf{G})} = \sum_{c}^{N_C} \frac{\delta E_x^{\text{exact}}[\rho]}{\delta \psi_{i(c)}^*(\mathbf{G})}. \tag{8.26}$$

8.5 Results and Discussion

We have assessed the performance of our algorithm with respect to the standard, full exchange by comparing average execution times for hybrid functional calculations with the B3LYP[31, 122, 163, 197] xc functional at the example of an ODIIS[229] wavefunction optimisation for N_b orbitals and the calculation of the first four singlet excited states using linear-response TDDFT for a series of polycyclic aromatic molecules. Performance gains are therefore representative both for ground-state Born-Oppenheimer or Car-Parrinello MD as well as for excited-state dynamics such as Tully's fewest switches trajectory surface hopping (TSH). All calculations were carried out using Martins-Troullier pseudopotentials[153] and a wavefunction cutoff of $E_{\text{cut}}^{\psi} = 60$ Ry and $E_{\text{cut}}^{\rho} = 240$ Ry. The Fourier transforms on the scaled orbitals were carried out with a cutoff energy of $E_{\text{cut}}^{\rho(\lambda)} = 60$ Ry. The (semi-)local contribution to the *xc* energy was evaluated using the new `XC_driver` of `CPMD` described in Ref. 161. The periodic images were decoupled using the Poisson solver by Tuckerman and Martyna[155]; the dimensions of the periodic supercell are tabulated in Table 8.1.

In order to assess the scaling with the number of electronic states, we considered acenes of increasing chain length l:



(8.27)

with X = H for $l = 1, 3, 5$. We have also included a stacked dimer for $X = \mathrm{H}, l = (5)_2$ and its fully chlorinated derivative $X = \mathrm{Cl}, l = (5)_2$, both at a distance of $d = 6.4$ a.u.

The reference implementation of exact exchange scales well on thousands of processors and is therefore suitable for very large numbers of electronic states. However, such resources may not always be available and here, we will therefore focus on medium-size systems that can in principle still be run on typical server infrastructure with a few dozen racks of 16-core nodes, even though execution times may be sizeable if the full exact exchange potential is computed. In this light, when assessing the performance gain due to our coordinate-scaled implementation, we will focus on the overhead generated by the redistribution of planes during the coordinate rescaling.

8.5.1 Box Size: Execution Times and Speedups

In order to assess the influence of yz-plane redistribution in the coordinate scaling scheme, we have performed calculations on naphthalene ($l = 1$, $N_b = 24$) at varying sizes of the simulation supercell and using a completely distributed Fourier grid. The number of processors is chosen such that every MPI task carries $n_x = 1$ yz-planes at $E_{\mathrm{cut}}^{\rho(\lambda)}$ and $n_X = 2$ yz-planes at E_{cut}^{ρ}. An increase in the size of the simulation supercell is therefore accompanied by an increase in the numbers of processors; therefore, the only parameter

Table 8.1: Number of CPU (or MPI tasks) for a given size of the simulation supercell spanned by vectors of length l_x, l_y, l_z. All calculations were carried out on identical processors on 16-node racks.

N_P	l_x [a.u.]	l_y [a.u.]	l_z [a.u.]
64	25.0	25.0	12.5
96	37.5	25.0	12.5
128	50.0	25.0	12.5
160	62.5	25.0	12.5
192	75.0	25.0	12.5

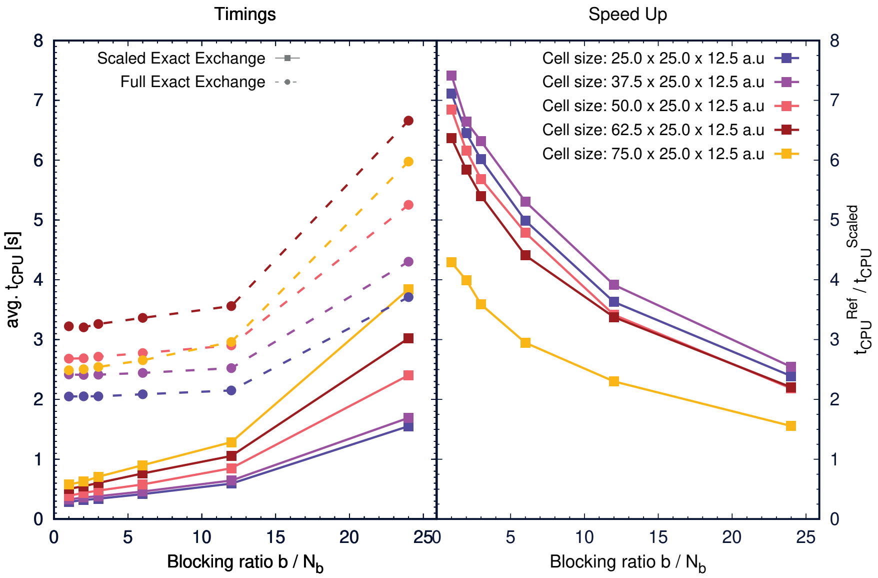


Figure 8.2: Average execution time and speedups for an average step of wavefunction optimisation for naphthalene using the ODIIS algorithm at different blocking ratios b/N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. Numbers of processors N_P for every simulation supercell size are given in Table 8.1; only one CP group is used.

varied is the communication during the FFTs and in grid redistribution. The number of processors for every supercell is summarised in Table 8.1.

The resulting timings and speedups are shown in Fig. 8.2 for the ground-state wavefunction optimisation, and Fig. 8.3 depicts timings for the calculation of the first four singlet states from a converged ground-state wavefunction. The data is given as a function of the blocking ratio b/N_b, where b is the block size. Both in the case of the coordinate-scaled exchange and the full exact exchange, best performance is obtained if $b = N_b$, *i.e.* a blocking ratio of 1, and efficiency decreases monotonously as b approaches 1 and the blocking ratio tends to the number of states. For the coordinate-scaled exchange, an increase in system and processor size consistently increases the execution time. For a constant number of blocks, this increase is almost linear. In the reference implementation, there is no clear trend in this respect; while the smallest supercell has the lowest execution times, the largest box

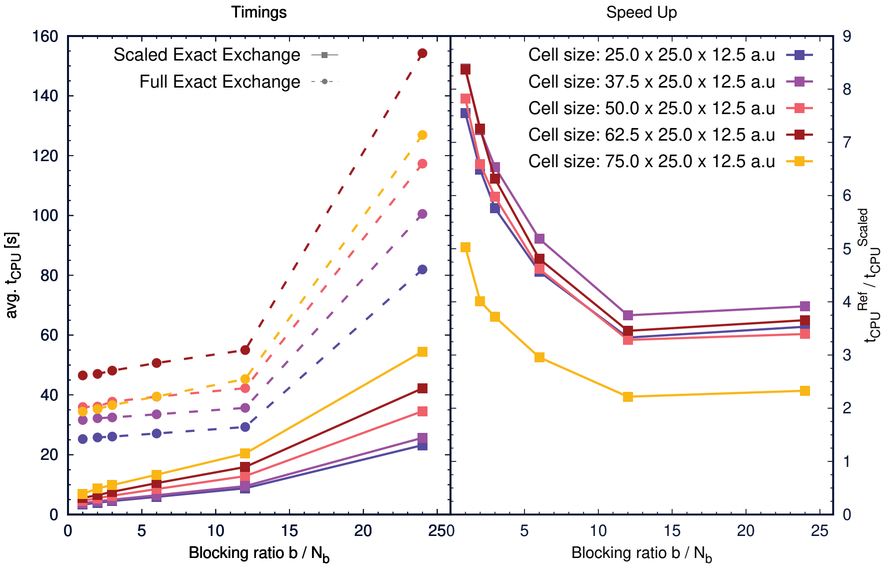


Figure 8.3: Average execution time and speedups for the calculation of the first four excited singlet states of naphthalene at different blocking ratios b/N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. Numbers of processors N_P for every simulation supercell size are given in Table 8.1; only one CP group is used.

is only slightly more expensive, but considerably cheaper to treat than the second-largest one.

Hence, the speedups that can be obtained vary between systems: The largest speedups are all observed for an optimal blocking ratio of 1 and range from about 7.5 for $x_l = 37.5$ a.u. to about 4 for $x_l = 75.0$ a.u. for an average DIIS step; the speedups obtained in a TDDFT calculation reach factors of 9.5 and 5, respectively. The efficiency of the wavefunction optimisation at $b = 1$ is severely hampered by load-balancing issues both if the full exact exchange and its coordinate-scaled analogue are computed.

8.5.2 Distribution of Orbital Pairs vs. Planes for various N_b

We will now assess the influence of varying the number N_C of CP groups on the efficiency for systems with varying numbers of orbitals, but at a

Table 8.2: Number of yz-planes, $x_{\mathrm{up}} - x_{\mathrm{low}} + 1$, held by ever MPI tasks, for a $75.0 \times 25.0 \times 12.5$ Å^3 supercell and different numbers of CP groups N_C at constant E_{cut}^{ρ} and $E_{\mathrm{cut}}^{\rho(\lambda)}$. The mesh for E_{cut}^{ρ} comprises a total of $x = 384$, $y = 128$ and $z = 64$ gridpoints; its counterpart for $E_{\mathrm{cut}}^{\rho(\lambda)}$ contains only half of the gridpoints per direction. Calculations were carried out on $N_P = 192$ cores.

$N_P = 192$		$x_{\mathrm{up}} - x_{\mathrm{low}} + 1$	
N_C	N_P/N_C	E_{cut}^{ρ}	$E_{\mathrm{cut}}^{\rho(\lambda)}$
1	192	2	1
2	96	4	2
3	64	6	3
6	32	12	6

constant size of the simulation supercell. Table 8.2 lists the number of yz-planes per MPI task for varying numbers of CP groups. Fig. 8.4 and 8.5 show execution times for acenes with $l = 1, 3, 5, (5)_2$ and $N_b = 24, 42, 60, 120$ orbitals, respectively. For ease of readability, execution times have been divided by N_b in all of the following plots.

Best performance for the scaled exchange is achieved using a blocking ratio of 1 with only one CP group, $N_C = 1$ for $N_b \leq 120$. Only for the largest system studied here, the chlorinated acene dimer with $N_b = 216$ does communication become significant, and the best performance is achieved by using $N_C = 3$. For the full exact exchange, load balancing issues are more important, and $N_C = 1$ is only optimal if $N_b \leq 24$. For higher numbers of electrons, maximum efficiency is obtained using $N_b = 3$. While the speedups increase from $3 - 4$ to $5 - 6$ from $N_b = 24$ to $N_b = 120$, a speedup of more than a factor of 8 can be achieved for the largest system studied here. This speedup is higher than the theoretical gain from the use of a lower FFT cutoff and must therefore be due to improved load balancing. The trends are the same for the TDDFT calculations performed with the scaled exchange. While the speedups show a larger spread, they start at a higher value for $N_b = 24$ (in a range of $3 - 5$) and reach only slightly smaller maximal values than their ODIIS counterpart; speedups are about $6 - 7.5$ for $N_b = 216$. The only notable difference between full and scaled exchange lies in the best configuration for $N_b > 24$, which is given by $N_C = 2$ rather than $N_C = 3$ for the full exchange, even though the differences between $N_C = 2$ and $N_C = 3$ are of the order of $\ll 10\%$.

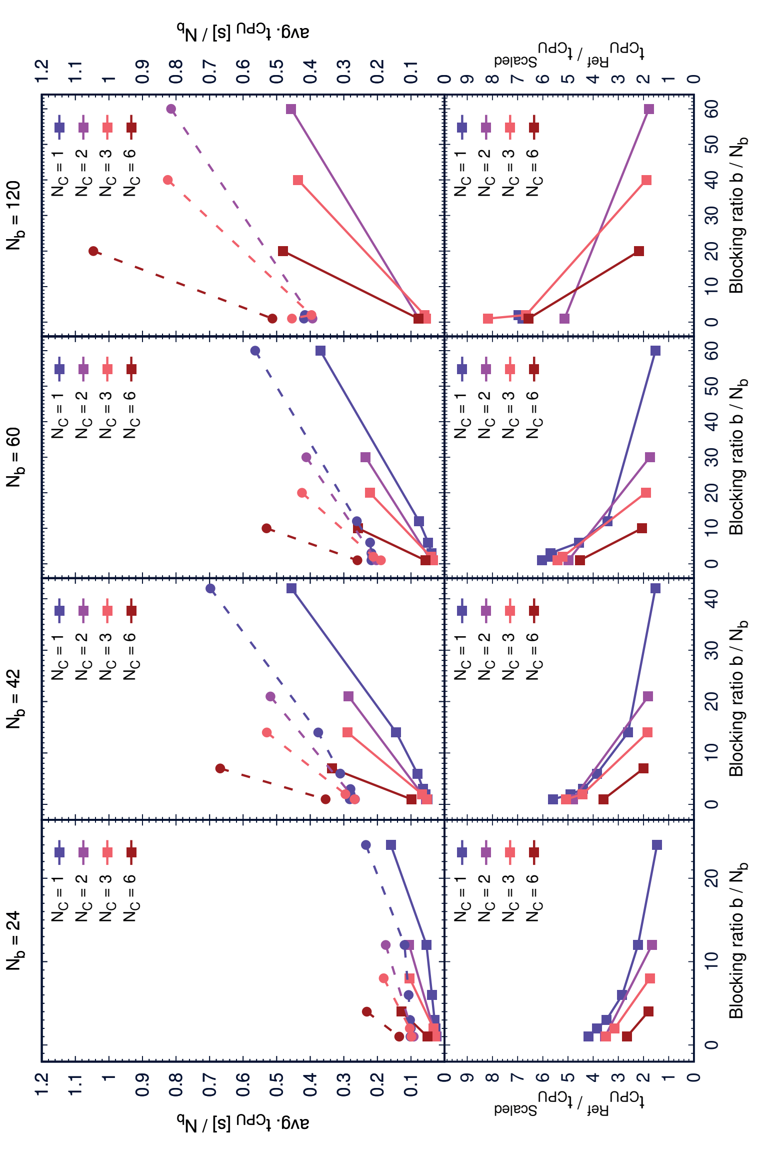


Figure 8.4: Average execution time and speedups for one step of wavefunction optimisation using the ODIIS algorithm at fixed supercell size for systems with different numbers of orbitals N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12 × 16 processor cores.

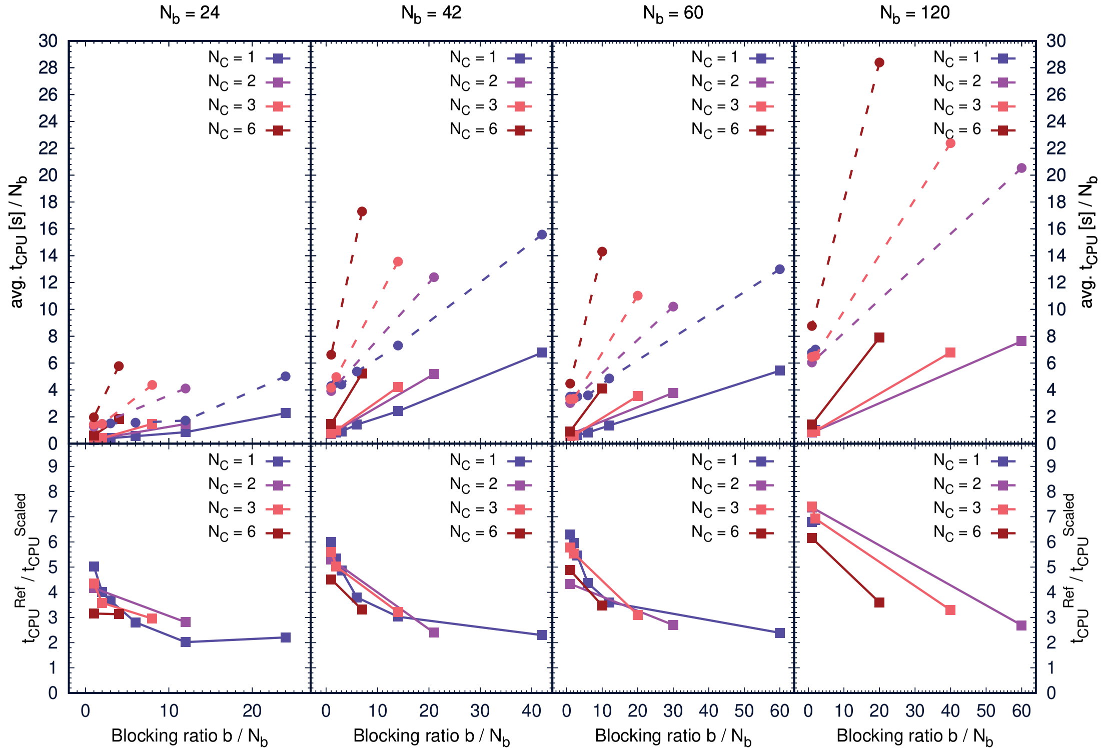


Figure 8.5: Average execution time and speedups for one step of optimisation of the first four Singlet states with linear-response TDDFT. All calculations were carried out at fixed supercell size for systems with different numbers of occupied orbitals N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12 × 16 processor cores.

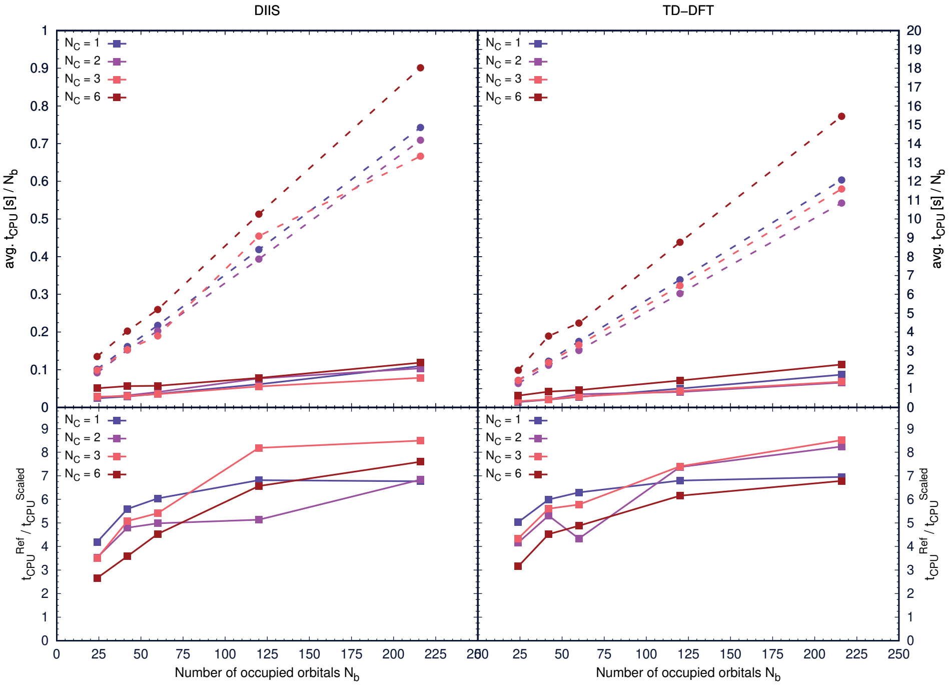


Figure 8.6: Average execution time and speedups for one step of wavefunction optimisation as well as for a TDDFT calculation involving the first four singlet states at fixed supercell size for systems with different numbers of orbitals N_b and different numbers of CP groups N_C. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12×16 processor cores.

8.5.3 Computational Cost as a Function of N_b and Effective Speedup

Fig. 8.6 compares execution times and speedups for the optimised blocking ratios with $b = N_b/N_C$ for different numbers of occupied orbitals N_b. Both execution times for the scaled and the full exchange nicely illustrate the on average cubic scaling with the number of orbitals (note that t_{CPU} is normalised by the number of states), apart from some outliers which can be attributed to load-balancing issues in the distribution of the orbital pairs. For both wavefunction optimisation and TDDFT calculations, speedups reach a range of up to 6.5−8.5 for $N_b = 216$. While the $N_C = 1$ and $N_C = 6$ speedups

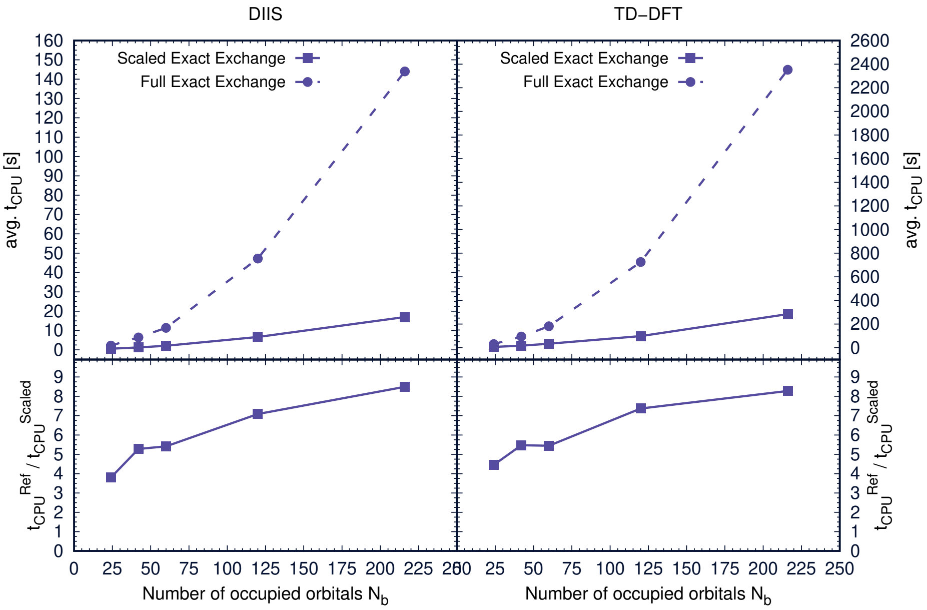


Figure 8.7: Average execution time and speedups for one step of wavefunction optimisation at N_b as well as for a TDDFT calculation for the first four singlet states at fixed supercell size for systems with different numbers of orbitals N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12 × 16 processor cores; values are compared between optimal parallelisation with different N_C, N_P and $b = N_b/N_C$.

evolve most smoothly, those for $N_C = 2$ and $N_C = 3$ are slightly more irregular for $N_b < 60$. This is also correlated with a more irregular increase of the execution times for the full exchange potential and is therefore likely to be due to load- and communication imbalances. Notably, the highest speedups are recorded for a large number of orbitals, reaching values of almost 9 for $N_C = 3$. The smaller increase of the speedup when increasing N_b from 120 to 216 suggests that the speedups reach a plateau region for a large number of states. Yet again, optimal configurations differ between the scaled and the full exact exchange; $N_C = 3$ has proven most beneficial for the former, whereas $N_C = 2$ is the most efficient for the latter. Overall, the coordinate-scaled exchange scales excellently with system size, allowing for sizeable speedups with respect to the full exact exchange. This alleviates the need for a large number of processors.

Fig. 8.7 summarises the shortest possible CPU times for both algorithms at every N_b and the corresponding speedup; Table 8.3 lists the corresponding values. While those values do not necessarily compare identical distributions comprising an identical number of N_C CP groups and N_P MPI tasks per CP group, they give an accurate estimate of the speedup that can be achieved with respect to an optimally set up calculation. While the best possible speedup is about 3.75 at $N_b = 24$, it increases to a virtually constant 5.5 for $N_b = 42$ and $N_b = 60$, which already constitutes a considerable decrease of execution time. The speedup is even larger for the most expensive systems that are treated in this study; at $N_b = 120$, it is larger than 7.0, and for $N_b = 216$, it reaches a factor of 8.5. This is almost one order of magnitude and more than the theoretical gain due to the decrease of E^{ρ}_{cut}. Given that there is additional cost due to the communication during the grid redistribution, this indicates that our coordinate-scaling approach can not only substantially reduce the cost of the intermediate FFTs, but that it also improves load-balancing when used within the existing exact exchange driver. Overall, for larger systems and independent of the approach, the optimal numbers of CP groups suggest that the best trade-off between distribution of FFTs and the distribution of orbital pairs is achieved once the number of yz planes per processors lies between $N_C = 3$ and $N_C = 4$.

Table 8.3: Lowest average execution times t_{CPU} for a wavefunction optimisation (ODIIS) and the optimisation of the first four excited states using linear-response TDDFT with increasing numbers of orbitals N_b for both coordinate-scaled (Scaled) and full (Ref) hybrid functional calculations performed within a $75.0 \times 25.0 \times 12.5$ Å^3 supercell on 192 CPU cores using an optimal block size b and group distribution N_C, N_P.

	DIIS		TDDFT	
N_b	$t^{\mathrm{Scaled}}_{\mathrm{CPU}}$ [s]	$t^{\mathrm{Ref}}_{\mathrm{CPU}}$ [s]	$t^{\mathrm{Scaled}}_{\mathrm{CPU}}$ [s]	$t^{\mathrm{Ref}}_{\mathrm{CPU}}$ [s]
24	0.6	2.2	6.8	30.5
42	1.2	6.4	17.1	94.0
60	2.1	11.4	33.4	181.5
120	6.7	47.2	98.4	725.3
216	17.0	144.0	284.3	2352.6

8.6 Summary

We have shown that our recently reported coordinate-scaling scheme is easily implemented in a plane wave code based on existing parallelisation strategies. After proposing a grid-redistribution scheme for the generation of the coordinate-scaled orbitals, we have shown that the approach scales well with increasing size of the simulation supercell for a constant number of orbitals by providing timings for both a ground-state wavefunction optimisation as well as the optimisation of the first four excited singlet states. We have subsequently identified the optimal blocking size b in the calculation of both the coordinate-scaled and the full exchange potential to be $b = N_b/N_C$, where N_C denotes the number of CP groups. We have compared execution times and speedups at constant numbers of processors, but different values of N_C and found that optimal load-balancing is achieved when every one out of the N_P processors holds $i_x^{\max} = 3$ to 4 yz pencils. A comparison of the fastest execution times showed that the speedup obtained with the coordinate scaling approach can reach ≈ 8.5 for a large number ($N_b = 216$) of generalised Kohn-Sham orbitals.

Our coordinate-scaling scheme has therefore been show to scale excellently with increasing system size and to provide unprecedented speedups of almost one order of magnitude. This will allow for hybrid functional calculations on systems with hundreds of orbitals to be carried out with considerably sleeker computational infrastructure. This applies to both ground- and excited-state calculations. With the coordinate-scaled exact exchange, much longer time scales can be reached and the (statistical) convergence of approaches that require many trajectories to be run in parallel (TSH, thermodynamic integration) can be sensibly improved - in practice, simulations that have previously taken one week to finish will be complete within less than one day.

8.7 Acknowledgements

MPB thanks Simone Meloni for many constructive discussions on plane wave algorithms and distributed memory parallelisation. UR gratefully acknowledges financial support from the Swiss National Science Foundation Grant No. 200020-146645 and the NCCR MUST.

PART IV

Retrospectives and Perspectives

CHAPTER 9

Further Work and Projects

WE have dedicated most of our discussion to the accuracy of the PES that underlies a putative *first principles* MD run. This chapter will conclude this perspective by summarising additional work that has been concerned with the improvement of dispersion forces in GGA based DFT. We will then shift our focus from improving the PES itself to improving its sampling; a description of a new, biased MD scheme tailored to structural elucidation based on electron densities only will complete the review of development carried out in the present work.

9.1 Dispersion Correction for Plane Wave DFT

Many xc functionals provide a description of dispersion forces that is either spurious at best, or lacking completely. In 2004, the Röthlisberger group have proposed to alleviate this issue for GGA xc functionals by resorting to an effective atom-centred potential formalism (DCACP) which will be sketched over the next paragraphs. The present work extends the availability of DCACP to parts of rows 5 and 6 of the periodic table and provides a case study for metals at the example of gold.

9.1.1 Atom-Centred Potentials for Molecular Properties

Placing basis functions on nuclei seems strikingly natural, even though it can be associated to certain drawbacks such as those discussed in Chapter 4. The concept of atom-centred properties (electronegativity, electrophilicity, etc.) is in itself abundant in chemistry, and it may therefore provide an intuitive basis for computational corrections that are atom-dependent. We shall outline two such approaches over the next few lines.

Historically, the implementation of a fully Hamiltonian quantum mechanics/molecular mechanics (QM/MM) coupling scheme[234] in the plane-wave pseudopotential code CPMD required the introduction of suitable capping techniques for QM/MM boundaries crossing covalent bonds, in order to assure a smooth transition between the QM and MM part. Standard approaches are either based on capping of dangling bonds with dummy hydrogens or on the inclusion of link pseudopotentials with an adjusted valence charge.[234, 235] However, in the absence of additional modifications to the pseudopotential, the latter is a rather crude approximation and can result in energy drifts.[234] Instead of a simple modification of the valence charge, the Röthlisberger group proposed a set of modified pseudopotentials[i] which, in their most general form, are designed to yield certain molecular properties.[235] One property relevant to QM/MM MD is the electronic density at the capping site. By minimising a penalty function that is constructed from the target property, analytical capping pseudopotentials were derived that mimic the density of a reference system, which then substantially improves on the quality of dipoles and ESP charges around the capping site. The penalty function $P(h_l, r_l)$ that guides the optimisation of the set of parameters $\{h_l, r_l\}$ of the analytical pseudopotential is given by a sum over real-space mesh points $\mathbf{R}$:

$$\{h_l, r_l\} = \operatorname{argmin} P(h_l, r_l) = \operatorname{argmin} \sum_{\mathbf{R}} (\rho_{\text{ref}}(\mathbf{R}) - \rho_{\text{cap}}(\mathbf{R}))^2, \tag{9.1}$$

where ρ_{ref} denotes the full QM reference electronic density in the region of interest, ρ_{cap} is the corresponding density for the capped system, and h_l, r_l are pseudopotential parameters that are optimised during the fit.

For reasons of simplicity, the fit was carried out on Goedecker-Teter-Hutter (GTH) pseudopotentials[154] which have a simple analytical form. In the GTH approach, the nonlocal part of the Hamiltonian is completely separable and defined as follows:

$$v_{\text{GTH}}(\mathbf{r}, \mathbf{r}') = \sum_{i}^{2} \sum_{l} \sum_{m} Y_{l,m}(\hat{\mathbf{r}}) p_i^l(r) h_i^l p_i^l(r') Y_{l,m}^*(\hat{\mathbf{r}}'). \tag{9.2}$$

$\hat{\mathbf{r}}$ is a unit vector in the direction of $\mathbf{r}$ and r is given by the distance $|\mathbf{r} - \mathbf{R}_I|$, where $\mathbf{R}_I$ is the nuclear centre on which the projectors are positioned. The radial projectors $p_i^l(r)$ are normalised and read

---

[i] The first use of such numerical potentials in the Röthlisberger group dates back to the late 1990ies

$$p_i^l(r) = \sqrt{2}\frac{r^{l+2(i-1)}e^{-\frac{1}{2}\left(\frac{r}{r_l}\right)^2}}{r_l^{l+2i-0.5}\sqrt{\Gamma(l+2i-0.5)}}, \tag{9.3}$$

where Γ denotes Euler's gamma function. In the original pseudopotential optimisation, h_i^l and r_l are optimised with respect to the all-electron atom; in the approach by Röthlisberger and coworkers, they are fit to best describe the more general target quantities instead (such as in eq. 9.1).

9.1.2 Dispersion-Corrected Atom-Centred Potentials

Inspired by the liberty of fitting pseudopotential parameters to *any* property, the Röthlisberger group proceeded with trials correcting qualitatively wrong behaviour of GGA exchange-correlation functionals. GGA are infamous for their failure of taking into account truly nonlocal correlation effects,[47,87,121,236] such as the formation of van der Waals complexes due to electron dispersion. Whilst some functionals exhibit spurious attraction in certain systems (PBE[94]), others show a completely repulsive interaction curve (BP86,[122,237] BLYP[122,197]). Although some of the Minnesota functionals[39–46] presented in Chapter 6 are able to describe (most) equilibrium van der Waals complexes, they may predict qualitatively wrong results in particular off equilibrium. This is due to the underlying parameterisation: In the Minnesota family, dispersive interactions are included by *training* the functional, *i.e.* the training data base includes a certain amount of van der Waals complexes without physically accounting for those interactions in the underlying functional form.

The approach by von Lilienfeld *et al.*[47,238] is an extension of the capping potential scheme outlined above and hence based on the electron density, while generating a very low computational overhead. In contrast to later popular approaches such as Grimme's DFT-D2[210] and DFT-D3,[239] the effect of electron dispersion is therefore self-consistently included during the optimisation of the Kohn-Sham orbitals, rather than by superimposing a force-field like potential term based on the molecular structure. While the approach was inspired by the capping potentials, the concept of the dispersion-corrected atom-centred potentials (DCACP) differs in one fundamental aspect.

Due to the encouraging results obtained with the capping potentials and the ease of implementation connected to the exploitation of an existing pseudopotential, the DCACP were calibrated by the use of modified pseudopotential files. However, and in contrast to the capping potentials, all of the angular momentum channels of the underlying pseudopotential were kept fixed. Instead, an additional channel (commonly f) was fit on top of the existing pseudopotential. Those additional DCACP parameters are orders of magnitude smaller than the pseudopotential parameters and, by construction, orthogonal to them. Since the atomic pseudopotential itself remains unmodified, this amounts to including a correction term in the Hamiltonian that is of the form:

$$E_{xc}[\rho] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) \left(H_{xc}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}' v_{\mathrm{GTH}}^{(f)}(\mathbf{r}, \mathbf{r}') \right), \tag{9.4}$$

where $H_{xc}(\mathbf{r})$ is the conventional GGA enhancement factor for exchange and correlation:

$$H_{xc}(\mathbf{r}) = K_{xc}(\mathbf{r})\rho^{1/3}(\mathbf{r}), \tag{9.5}$$

and $v_{\mathrm{GTH}}^{(f)}$ is a potential with the functional form of eq. 9.2 at angular momentum f. The DCACP parameters $\sigma_1 = h_l$ and $\sigma_2 = r_l$ are then calibrated at the equilibrium geometry $\mathbf{r}_{AB}$ of a van der Waals complex using the penalty function:

$$\{\sigma_1, \sigma_2\} = \operatorname{argmin} P(\sigma_1, \sigma_2), \tag{9.6}$$

$$\begin{aligned} P(\sigma_1, \sigma_2) = & \sum_{\mathbf{R}_\alpha}^{\text{config.}} \left(\Delta E^{\mathrm{DCACP}}(\mathbf{R}_\alpha, \sigma_1, \sigma_2) - \Delta E^{\mathrm{ref}}(\mathbf{R}_\alpha) \right)^2 \\ & + \sum_{\alpha}^{N_\alpha} 10^{-\omega} \left| \mathbf{F}_\alpha^{\mathrm{DCACP}}(\mathbf{R}_\alpha^{\mathrm{eq}}, \sigma_1, \sigma_2) - \mathbf{F}_\alpha^{\mathrm{ref}}(\mathbf{R}_\alpha^{\mathrm{eq}}) \right|^2 , \end{aligned} \tag{9.7}$$

$$\Delta E = \Delta E_{AB}(\mathbf{r}_{AB}) - \Delta E_{AB}(\infty), \tag{9.8}$$

where E^{ref} is the interaction energy of the van der Waals complex at a set of dimer nuclear coordinates (or distances) $\mathbf{R}_\alpha$, E^{DCACP} is its counterpart obtained from the use of DCACP, and the $\mathbf{F}_\alpha$ are the forces acting on the N_α nuclei α at the equilibrium geometry characterised by $\mathbf{R}_\alpha^{\mathrm{eq}}$. ∞ denotes the limit of a dissociated complex. ω is a scaling factor that is used if the forces fall below a certain threshold:

$$\omega = \begin{cases} 0 & \text{for} \quad \left| \mathbf{F}_\alpha^{\mathrm{DCACP}}(\mathbf{R}_\alpha^{\mathrm{eq}}, \sigma_1, \sigma_2) - \mathbf{F}_\alpha^{\mathrm{ref}}(\mathbf{R}_\alpha^{\mathrm{eq}}) \right| \geq 10^{-4} \\ 4 & \text{for} \quad \left| \mathbf{F}_\alpha^{\mathrm{DCACP}}(\mathbf{R}_\alpha^{\mathrm{eq}}, \sigma_1, \sigma_2) - \mathbf{F}_\alpha^{\mathrm{ref}}(\mathbf{R}_\alpha^{\mathrm{eq}}) \right| < 10^{-4} \end{cases} . \tag{9.9}$$

The penalty function may be applied to the minimum of the interaction curve at $\mathbf{r}_{eq}$ only, but improved results were obtained for certain elements if

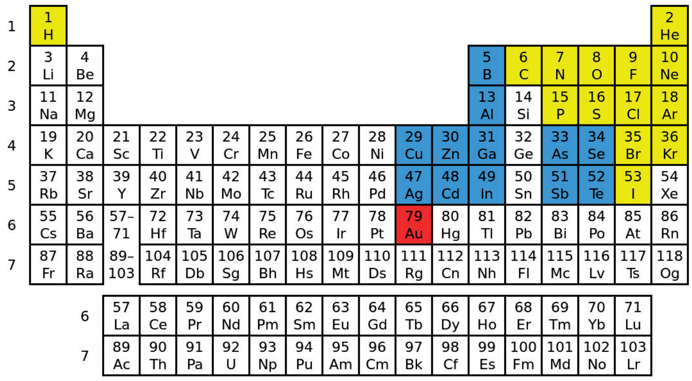

1
2
3
4
5
6
7
1 H
2 He
3 Li
4 Be
5 B
6 C
7 N
8 O
9 F
10 Ne
11 Na
12 Mg
13 Al
14 Si
15 P
16 S
17 Cl
18 Ar
19 K
20 Ca
21 Sc
22 Ti
23 V
24 Cr
25 Mn
26 Fe
27 Co
28 Ni
29 Cu
30 Zn
31 Ga
32 Ge
33 As
34 Se
35 Br
36 Kr
37 Rb
38 Sr
39 Y
40 Zr
41 Nb
42 Mo
43 Tc
44 Ru
45 Rh
46 Pd
47 Ag
48 Cd
49 In
50 Sn
51 Sb
52 Te
53 I
54 Xe
55 Cs
56 Ba
57–71
72 Hf
73 Ta
74 W
75 Re
76 Os
77 Ir
78 Pt
79 Au
80 Hg
81 Tl
82 Pb
83 Bi
84 Po
85 At
86 Rn
87 Fr
88 Ra
89–103
104 Rf
105 Db
106 Sg
107 Bh
108 Hs
109 Mt
110 Ds
111 Rg
112 Cn
113 Nh
114 Fl
115 Mc
116 Lv
117 Ts
118 Og
6
57 La
58 Ce
59 Pr
60 Nd
61 Pm
62 Sm
63 Eu
64 Gd
65 Tb
66 Dy
67 Ho
68 Er
69 Tm
70 Yb
71 Lu
7
89 Ac
90 Th
91 Pa
92 U
93 Np
94 Pu
95 Am
96 Cm
97 Bk
98 Cf
99 Es
100 Fm
101 Md
102 No
103 Lr

Figure 9.1: Availability of DCACP for the BLYP xc functional: Currently published and tested (yellow), unpublished and tested in the current work (red), calibrated and tested in the current work (blue). This table is a derivative of Pekka Pyykkö's 172-element (extended) periodic table by Pekka Pyykkö and WhiteTimberwolf, used under CC BY-SA 3.0, initially reproduced from Ref. 240 with permission of the author.

a midpoint $\mathbf{r}_{mid}$, which corresponds to a point where the interaction energy is half of its maximal value, is included in the penalty function.[238] Initially, all reference data was obtained from high-level wavefunction calculations. Naturally, the parameters σ differ substantially from those used in the generation of a pseudo- or capping potential: $\sigma_1 = h_l$ is orders of magnitudes smaller than for pseudowavefunction projectors, and $\sigma_2 = r_l$ is located further away from the nuclear centre, which effectively achieves a complete spatial separation of the DCACP and the pseudopotential[47]; the pseudowavefunctions themselves are not modified, which maintains their chemical transferability.

In its essence, the DCACP scheme is a perturbative approach to the electron density.[47, 238] The success of DCACP was linked to Bader's atoms-in-molecules theory,[238] but a derivation based on analytical expressions is still missing. DCACP have successfully been used in a broad range of applications. Notable examples include first-principles molecular dynamics[236] of water using the BLYP functional, where the use of DCACP significantly improved structural properties over the uncorrected BLYP, thus pinpointing the significance of van der Waals interactions in hydrogen bonding of water. DCACP accurately predict stacking of organic molecules[238, 241] and biomolecules,[242] including base pairs.[243] They are also able to take halogen-halogen interactions into account, a phenomenon where many other disper-

Table 9.1: Calibration systems for DCACP as used in the present work. Values for σ_1, σ_2 are given for the species calibrated in the present work, denoted in bold face.

	ΔE [kcal mol^{-1}]		r_{eq} [Å]			
System	DCACP	Ref.	DCACP	Ref.	σ_1 [a.u.]	σ_2 [a.u.]
CuAr[247]	-0.254	-0.255	4.00	4.00	2.821	$-3.533\cdot10^{-3}$
AgAr[248]	-0.255	-0.257	3.99	4.00	2.786	$-4.062\cdot10^{-3}$
ZnAr[249]	-0.240	-0.240	4.40	4.38	2.825	$-2.840\cdot10^{-3}$
CdAr[250]	-0.417	-0.320	4.27	4.28	3.662	$-1.662\cdot10^{-3}$
$(\mathbf{B}_3N_3H_6)_2$[ii;251]	-3.57	-3.3	3.57	3.42	3.235	$-0.530\cdot10^{-3}$
AlAr[252]	-0.428	-0.431	3.5	3.5	2.320	$-4.404\cdot10^{-3}$
GaAr[253]	-0.100	-0.100	4.73	4.76	3.753	$-2.657\cdot10^{-3}$
InAr[254]	-0.492	-0.512	3.6	3.6	2.516	$-4.863\cdot10^{-3}$
$(\mathbf{As}_4)_2$[255]	-5.42	-5.49	3.72	3.70	3.221	$-1.939\cdot10^{-3}$
$(\mathbf{Sb}_4)_2$[255]	-6.87	-6.89	3.91	3.92	3.971	$-2.979\cdot10^{-3}$
$(\mathbf{Se}_6)_2$[255]	-5.70	-5.72	4.43	4.43[iii]	3.113	$-1.887\cdot10^{-3}$
$(\mathbf{Te}_6)_2$[255]	-9.75	-8.83	3.39	3.40	3.344	$2.372\cdot10^{-3}$

sion correction schemes fail.[244]

9.1.3 Extension of the DCACP Library

Although portable and of high accuracy, the limited availability of DCACP for large parts of the periodic table may discourage potential users. Fig. 9.1 shows the DCACP calibrated in previous work[47,238,244–246] and the currently unpublished extensions to it.

In this work, the DCACP library has been extended to comprise the elements Cu, Ag, Zn, Cd, B, Al, Ga, In, As, Sb, Se and Te for the BLYP functional. In contrast to the existing DCACP, the additions to the library made use of literature reference data for complexation energies and geometries (both high-level wavefunction methods and experimental values were included). Table 9.1 summarises the calibration systems and the origin of the reference data. All calibrations were carried out using a modular Nelder-Mead-Simplex or Particle Swarm (PS) algorithm written in Fortran2008. The implementation introduces the possibility to relax molecular structures in short geometry optimisations, which allows for a self-consistent optimisation of the DCACP parameters wherever equilibrium structures may be more severely impacted by electron dispersion. Calibrations for Se and Te were

---

[ii] Sandwich

[iii] **r** between Se atoms with identical xy coordinates

Table 9.2: Test systems for DCACP calibrated in the present work. The species calibrated in the present work are indicated in bold face.

	ΔE [kcal mol^{-1}]		r_{eq} [Å]	
System	DCACP	Ref.	DCACP	Ref.
CuHe[256]	-0.059	-0.012	4.00	4.85
CuKr[247]	-0.443	-0.427	3.92	3.92
AgHe[256]	-0.021	-0.017	4.63	3.75
ZnHe[257]	-0.031	-0.022	4.49	4.46
ZnKr[249]	-0.362	-0.329	4.28	4.20
CdHe[250]	-0.143	-0.043	4.3	4.5
CdKr[250]	-0.556	-0.383	4.4	4.4
$(\mathbf{B}_3N_3H_6)(C_6H_6)$[251]	-2.52	-2.1	3.80	3.62
$(\mathbf{B}_3N_3H_6)_2$[iv;251]	-2.70	-2.6	2.49	2.67
AlKr[254]	-0.610	-0.571	3.6	3.7
GaKr[258]	-0.728	-0.945	3.93	3.78
InKr[254]	-0.705	-0.512	3.7	3.9

carried out by creating a dimer on-the-fly, based on monomer structures optimised using the updated DCACP parameters. In contrast, for As and Sb, the monomers were kept fixed, while the dimer was allowed to relax upon changing the DCACP parameters. All the other systems did not require structural relaxation. Preliminary investigations on alkali and earth-alkali dimers revealed spurious attraction due to the BLYP functional, indicating that the bonding was significantly influenced by other than dispersion interactions; calibrations for groups I and II were therefore not pursued any further. A previously calibrated DCACP for Au[167] is currently undergoing testing at the example of physisorption of benzene on a gold layer, *cf.* Chapter 9.1.4.

The performance assessment for the metal DCACP in Table 9.2 shows good energetic agreement for all complexes; while the relative error for very weakly bound systems that involve helium is large, the absolute error is still very low. Given the weak interaction, the DCACP can still reproduce the energetics of those systems very well, while the equilibrium geometry shows larger errors due to the extreme flatness of the binding interaction curves. Due to the low magnitude of the interaction energy, calibrations and validations have been carried out on a particularly fine mesh. This can introduce oscillations in the PES in the 10^{-2} kcal mol^{-1} range that generate various spurious

---

[iv] T-shaped

minima with respect to the forces. However, those effects are only limited to complexes with helium, the krypton complexes are of satisfactory accuracy throughout. It has to be stressed that errors might be slightly larger for the present calibrations since they rely on mixed dimers. Existing DCACP have all been calibrated using pure compounds that contain no other species apart hydrogen, eliminating contributions from other heavier elements. Still, the mixed calibration leads to DCACP of good accuracy, which is further reflected in the data for boron. Both a mixed borazine-benzene dimer as well as a T-shaped borazine-borazine dimer are well described with the new boron DCACP, in notable contrast to the popular DFT-D3[239] correction that severely overestimates the binding energy of the borazine reference dimer by almost 100% ($\Delta E_{\text{BLYP-D3}} = 5.2$ [kcal mol^{-1}]).

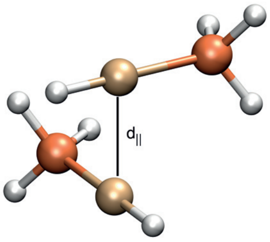
9.1.4 Gold: A Case Study

Gold is a particularly interesting element not only for its socioeconomic implications[v]. Both the particular physics and chemistry of gold are governed by strong relativistic effects[261] which are not only at the root of the noble metal's characteristic colour and its lattice constant, but which also give rise to an aurophilic effect: London dispersion forces are not only crucial for the formation of surface-adsorbate complexes (just as it is the case for many metals), but they are also assumed to play an essential role in the formation of the comparably strongly bound aurophilic complexes such as $(HAuPH_3)_2$.[259] Table 9.3 shows the performance of an unpublished DCACP

---

[v] Or its glistening shininess.

Table 9.3: Equilibrium distance $d_{||}$ and binding energy $\Delta E_{||}$ for the aurophilic complex $(HAuPH_3)_2$. DCACP-BLYP predicts a minimum close to the CCSD(T) reference.[259] The energetic error of DFT-D2,[210] however, considerably exceeds the range of chemical accuracy.



Method	$d_{||}$ [Å]	$\Delta E_{||}$ [kcal mol^{-1}]
CCSD(T) [4]	3.00	-6.13
DCACP-BLYP	3.15	-6.46
BLYP-D2	3.65	-9.79

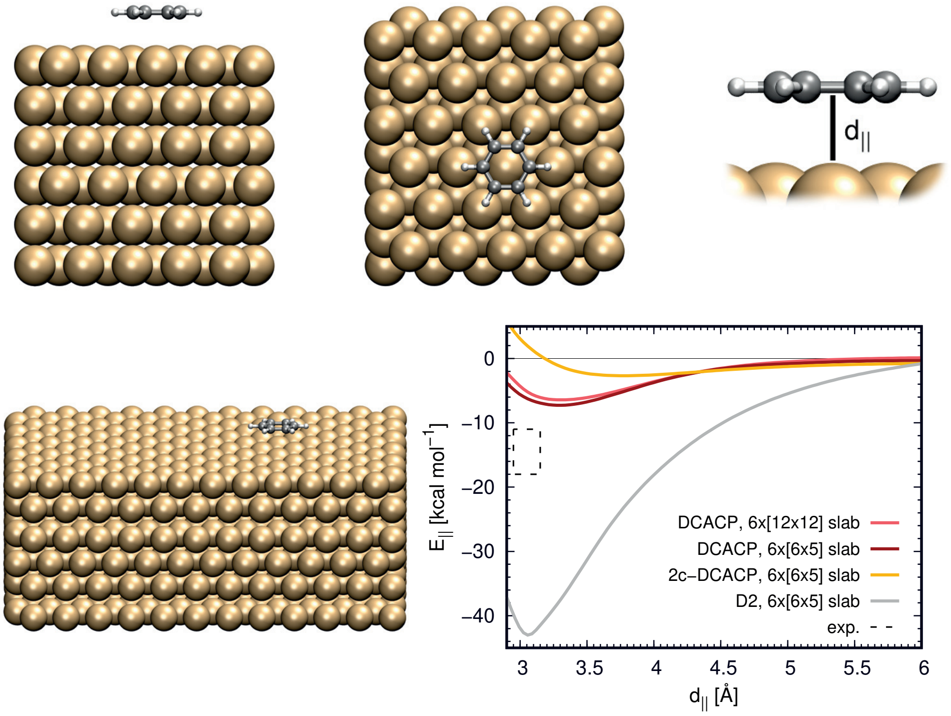


Figure 9.2: Performance of a one- and two-channel (2c) DCACP for the physisorption of benzene on a Au(111) surface on a $6 \times [12 \times 12]$ (bottom) and $6 \times [5 \times 6]$ (top) slab. Values obtained using DFT-D2 are given for comparison. The reference data is interpolated from experiment; it has been shown that by altering the collision factor, the values could change by up to 20 kcal mol^{-1}.[260]

for gold[167] in the description of said aurophilic complex. DCACP-BLYP shows good performance for both energy and equilibrium distance of the aurophilic dimer and is in both cases more accurate than Grimme's DFT-D2[210]; while the latter is reasonably accurate for the equilibrium distance, the complexation energy is off by over 50%, which by far exceeds chemical accuracy.

Usually, the fit of a set of parameters σ_1, σ_2 that corresponds to one projector is sufficient,[47,238] but an improved r^{-6} asymptotic decay of the potential could be obtained by including two angular momentum channels (d and f instead of f only) in a two-channel DCACP.[238] However, this was never verified for interactions between a surface and an adsorbate, where the potential

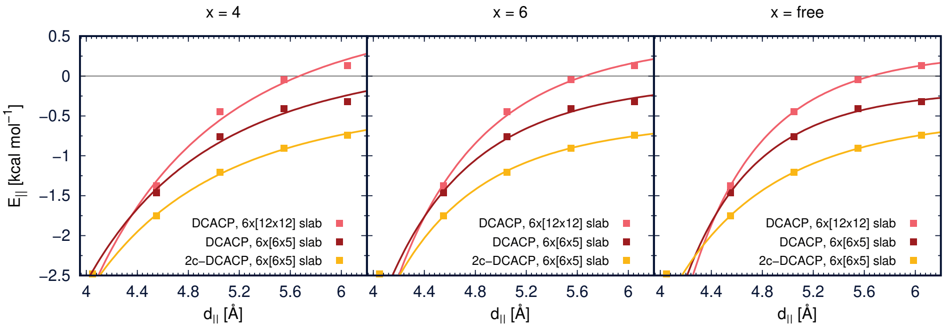


Figure 9.3: Decay of the surface-adsorbate interaction for the graphs of Fig. 9.2 fit on $ad_{||}^{-x} + b$ for fixed $x = 4$, $x = 6$ and using x as a variable (free). Optimised parameters are $x = 7.55$ for DCACP, $6 \times [12 \times 12]$ slab, $x = 7.51$ for DCACP, $6 \times [6 \times 5]$ slab and $x = 5.16$ for 2c-DCACP, $6 \times [6 \times 5]$ slab.

decays as r^{-4}. Based on both the DCACP tested for $(HAuPH_3)_2$ and its unpublished two-channel extension, the accuracy of two-channel DCACPs to describe the long-range decay of a physisorbed molecule on a metal surface was examined at the example of benzene on a Au(111) surface and compared to results from Grimme's DFT-D2, the corresponding potential energy profiles are depicted in Fig. 9.2. For the one-channel DCACP, the minimum of adsorption is reasonably well reproduced. The two-channel DCACP shifts the minimum by about 0.5 Å and predicts only about half of the one-channel DCACP binding energy. Comparison of the interaction energy to reference data[260] is made difficult by the uncertainty in the latter. Since the best (experimental) reference data available requires that a typical collision factor be assumed in order to reconstruct the adsorption energy,[260] the energy thus obtained can vary by almost two orders of magnitude as a function of the collision factor. This can change the absorption energy by up to over 20 kcal mol^{-1}. The drastic deviation of the DFT-D2 curve from the reference range can be attributed to a general overbinding tendency of DFT-D2, multiply reported in the literature.[262–264] Due to the large spread of the experimental data, DCACP and DFT-D2, it is at present not possible to draw a definitive conclusion with respect to the energetics.

Fig. 9.3 shows the long-range decay of the potential for one- and two-channel DCACP. The r^{-4} long-range decay expected for the interaction is not reproduced in any of the systems. Instead, by fitting the exponent of the long-range decay, one finds that the one-channel DCACP decays roughly as r^{-8}, while its two-channel counterpart exhibits a r^{-5}-like decay. Notably, if a one-channel DCACP is used to describe the $6 \times [12 \times 12]$ slab, there appears a small artificial barrier at large $d_{||}$. This is a behaviour that has previously been observed to occur in certain systems and which is remedied by the two-channel DCACP.[238]

9.1.5 Direct Implementation of the DCACP Library in the CPMD Code

In order to facilitate the use of DCACP and to avoid the manual modification of pseudopotential files, the DCACP library was directly implemented in the CPMD code and can therefore be accessed in an intuitive keyword based way. For ease of implementation, the internal library still makes use of empty pseudopotential projector channels, but their generation is automatised and does not require user interaction.

9.2 Density-Difference Driven MD: d3MD

We have discussed various aspects of *first principles* MD accuracy without actually having performed any dynamics. We should not want to border on this insolence for any longer; and so in this very last section, atoms will finally be made to wiggle.

The Röthlisberger group has recently demonstrated that *first principles* Molecular Dynamics can be used as a predictive tool in the refinement of time-resolved crystal structures.[265] Such experiments provide density differences maps between a state at time t and the initial state $t = 0$, but the resolution may not be sufficient in order for conventional refining algorithms to predict a sensible structure without imposing too many (potentially non-physical) constraints, or the structures of the intermediates may not be described by conventional models used for the refinement. A *first principles* refinement is therefore particularly beneficial for such systems. The electron density difference maps from experiment can then be compared to those computed on-the-fly during the MD, therefore quantifying the agreement between structures obtained from simulation and experiment. In particular, a dynamic approach does not only allow for the characterisation of stable and meta-stable intermediates, but makes transient structures accessible as well.

However, success relies on the possibility of describing the (rare) event that links states 0 to t. But what if that event cannot be simulated (at least within reasonable computational time)? One obvious possibility would lie in combining experimental and *first principles* density difference maps in order to guide the time evolution of the system at hand. It is this very approach that will be explored over the next few paragraphs: Density-Difference Driven Molecular Dynamics (d3MD).

9.2.1 Biasing Potentials from Electron Density Differences

Instead of relying on sets of collective variables or nuclear constraints, d3MD is based on an electronic biasing potential which, in a general form, reads:

$$\hat{\mathrm{H}}_{\mathrm{KS}}^{\mathrm{d3MD}}(t) = \hat{\mathrm{H}}_{\mathrm{KS}} + \hat{\mathrm{V}}_{\mathrm{d3MD}}(t), \tag{9.10}$$

where $\hat{\mathrm{H}}_{\mathrm{KS}}$ is the familiar time-independent Kohn-Sham Hamiltonian and $\hat{\mathrm{V}}_{\mathrm{d3MD}}(t)$ is a time-dependent biasing potential constructed from differences-of-differences between a set of reference data (Ref) and the MD run:

$$\begin{aligned} \Delta\Delta\rho(\mathbf{r},t) &= \Delta\rho_{\mathrm{MD}}(\mathbf{r},t) - \Delta\rho_{\mathrm{Ref}}(\mathbf{r},t) & (9.11)\\ &= \rho_{\mathrm{MD}}(\mathbf{r},t) - \rho_{\mathrm{MD}}(\mathbf{r},0) - \left(\rho_{\mathrm{Ref}}(\mathbf{r},t) - \rho_{\mathrm{Ref}}(\mathbf{r},0)\right). & (9.12) \end{aligned}$$

Here, the point t identifies an *on-the-fly* density of the MD run. The biasing potential $v_{\mathrm{d3MD}}(\mathbf{r},t)$ is then constructed from:

$$v_{\mathrm{d3MD}}(\mathbf{r},t) = \kappa(t)\Delta\Delta\rho(\mathbf{r},t), \tag{9.13}$$

where $\kappa(t)$ is some time-dependent function that controls the magnitude of the bias, and the biasing potential energy and its functional derivative evaluated at a fixed time step τ are given by:

$$V_{\mathrm{d3MD}}(t) = \int \mathrm{d}\mathbf{r}\; v_{\mathrm{d3MD}}(\mathbf{r},t)\rho(\mathbf{r},t), \tag{9.14}$$

$$\left.\frac{\partial V^{\mathrm{d3MD}}\left[\rho, \Delta\rho_{\mathrm{Ref}}\right]}{\partial\rho(\mathbf{r},t)}\right|_{t=\tau} = \kappa(t)\left\{\rho_{\mathrm{MD}}(\mathbf{r},t) + \Delta\Delta\rho(\mathbf{r},t)\right\}. \tag{9.15}$$

This scheme can be straightforwardly implemented in existing MD codes at little computational overhead. Depending on the sensitivity of the problem at hand with respect to the applied bias, $\kappa(t)$ can either be a constant $\kappa(t) = c$, or it can be chosen to grow linearly $\kappa(t) = (c_{\mathrm{End}} - c_0)t$. Any other functional form is in principle possible, such as *e.g.* an oscillating potential.

9.2.2 Application to Test Systems

We have implemented d3MD in the molecular dynamics code CPMD[176] and carried out preliminary studies using theoretical density differences as a reference, *i.e.* density differences calculated from KS-DFT density differences maps between an initial and final structure[vi]. We consider both inter- and in-

---

[vi] Which are both known, since the aim is to provide a proof of principle.

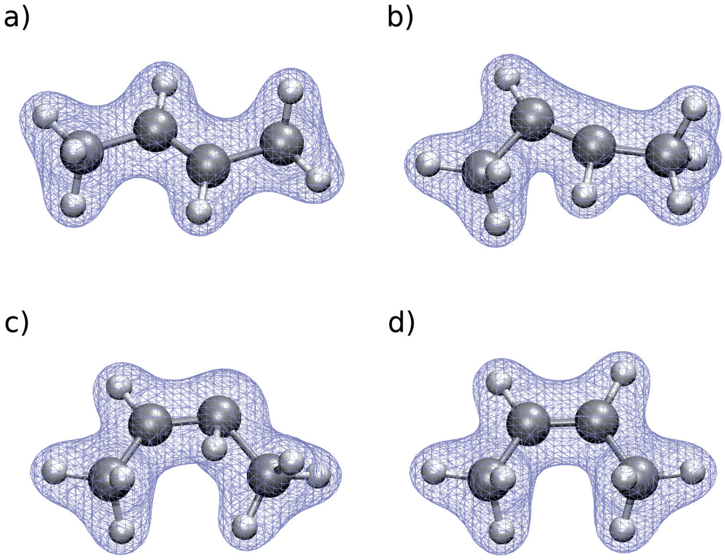


Figure 9.4: Isomerisation of *trans*-2-butene to its *cis* form by performing d3MD at 400 K and using $\kappa(t) = 0.50$ a.u. The N-electron density is depicted in lilac. Snapshots were taken from 0 ps (panel a) to 15 ps (panel d) and are spaced by 5 ps.

tramolecular reactions, as well as different forms for $\kappa(t)$. All simulations are carried out at finite temperature using a simple velocity rescaling algorithm.

9.2.2$_{\text{A}}$ Pushing Against the Barrier: *trans*-2-Butene to *cis*-2-Butene

The isomerisation of 2-butene can take place at high temperatures, with the equilibrium being shifted to the more stable isomer, *trans*-2-butene. Here, we perform a forced isomerisation of *trans*-2-butene to *cis*-2-butene in the gas phase at 400 K by performing d3MD using a scaling factor of $\kappa(t) = 0.50$ a.u. Since this reaction does not readily occur in the ground state and strict single state dynamics are enforced, a large barrier is to be expected; the system can therefore serve as a test case for the general potential of the method even if the underlying lowest-energy mechanism does not occur in the ground state. Such transformations - which do not necessarily have to correspond to the minimum energy transition path - are of particular relevance with regard to

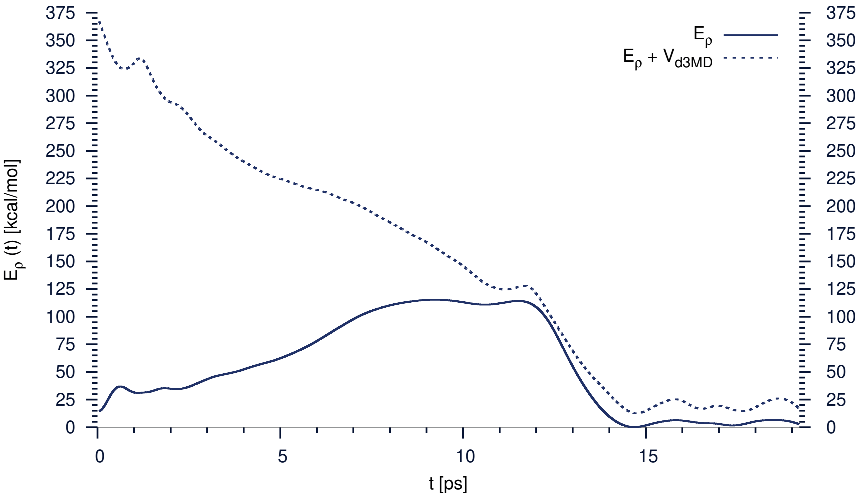


Figure 9.5: Biased ($E_\rho + V_{\mathrm{d3MD}}$) and unbiased ($E_\rho$) potential energy profile for the isomerisation of *trans*-2-butene to its *cis* form. The unbiased potential energy profile shows that an effective barrier of about 100 kcal mol^{-1} has to be overcome.

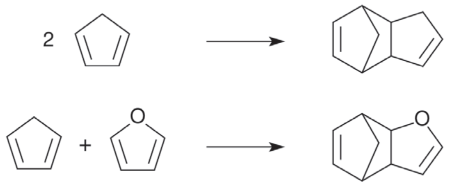
the possibility of carrying out structural refinement based on electron density difference maps.

Fig. 9.4 shows various snapshots collected along the trajectory, the corresponding biased and unbiased potential energy profile together with the biasing potential V_{d3MD} of eq. 9.14 are depicted in Fig. 9.5. At constant value of κ, the biasing potential continuously decreases in magnitude along the trajectory, whereas the biased potential energy is only slightly growing until the reactive event at 12 ps. The biased potential energy decreases almost monotonically during the reaction and does not appear to exhibit any barrier; the slight bump at 12 ps is well within the range of thermal fluctuations. The biasing potential itself is almost zero at the top of the barrier on the underlying PES. In agreement with the isomerisation not readily taking place in the ground state, the unbiased potential energy is substantial, reaching an order of 10^2 kcal mol^{-1} . Still, due to the lack of a barrier on the biased potential energy profile, the system readily isomerises into the target structure. As it is evident from Fig. 9.4, the path from panel a) to d) bears no resemblance to an intuitive minimum energy path (note the delayed isomerisation of the hydrogen atom). Nonetheless, the reaction takes place without disrupting the molecular structure. The method therefore appears suitable for structural refinements where there is no principal interest in the minimum energy path and, hence, no need of more involved enhanced sampling techniques.

The rather abrupt isomerisation (in particular of the hydrogen atom) raises the question of adiabaticity and hysteresis effects. The applied potential may be too strong for the nuclei to follow the electron density adiabatically, which would require a biasing potential of lower magnitude. In the next section, we will first consider a monotonically growing biasing function at the example of two other systems, before revisiting 2-butene at the very end of this chapter.

9.2.2$_{\mathrm{B}}$ Intermolecular Reactions: Diels-Alder Cyclo-Additions

Encouraged by the results obtained for the isomerisation of 2-butene, and with the hope of characterising more than just reactants and products, we will now examine a simple intermolecular reaction: The Diels-Alder cycloaddition of cyclopentadiene with itself and with furan, the latter being disfavoured with respect to the former[266]:



The reactant state consists of a snapshot of a short molecular dynamics run of the van der Waals complex of the monomers. Both reactions are simulated at a typical temperature of 400 K using a scaling factor of the form $\kappa(t) = (c_{\mathrm{End}} - c_0)t$, with $c_0 = 0.010$ a.u. and $c_{\mathrm{End}} = 0.105$ a.u. Snapshots of the molecular structure and density of the dimerisation of cyclopentadiene are depicted in panels a) to h) of Fig. 9.6. The reactants and products of this reaction are shown in Fig. 9.7, along with the reference density difference $\Delta\rho_{\mathrm{Ref}}(\mathbf{r})$.

The evolution of the carbon-carbon (CC) bonds that are formed in the product are shown in Fig. 9.8. While the profile for both CC-bonds is virtually identical for the dimerisation to dicyclopentadiene, it is asymmetric for the formation of the cyclopentadiene-furan adduct. For the latter, the CC distance involving the carbon adjacent to the oxygen is smaller upon approach of the reactants and then stays at a plateau while the other CC distance shortens. Fig. 9.10 compares the strength of the biasing potentials as a function of time (or, equivalently, of κ) for the formation of both the

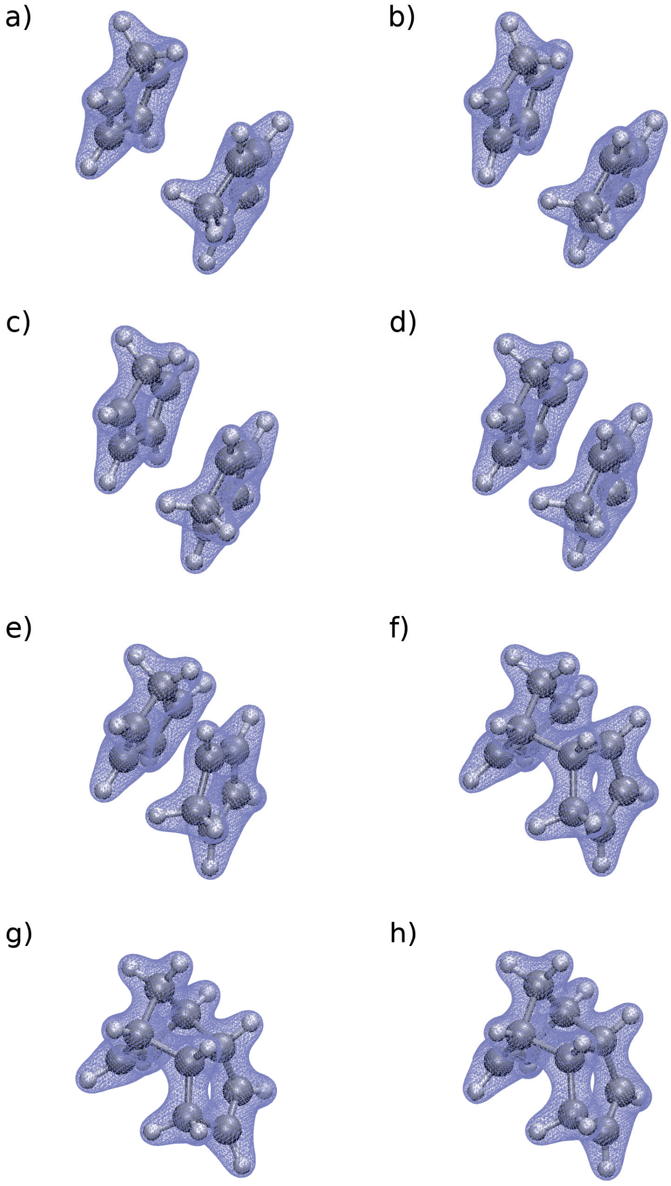


Figure 9.6: Dimerisation of two molecules of cyclopentadiene to yield dicyclopentadiene. Shown is the N-electron density in lilac. Snapshots were taken from 3 ps (panel a) to 24 ps (panel h) and are spaced by 3 ps.

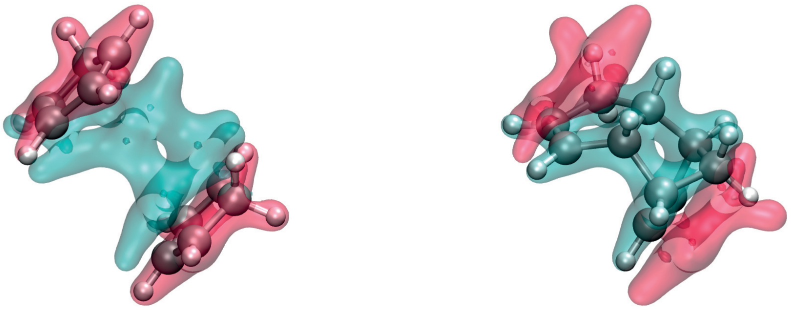


Figure 9.7: Reactants and product for the formation of dicyclopentadiene from cyclopentadiene along with the reference density difference map $\Delta\rho_{\mathrm{Ref}}(\mathbf{r})$. Regions of density depletion (negative density difference) are depicted in red, regions of accumulation (positive density difference) in turquoise.

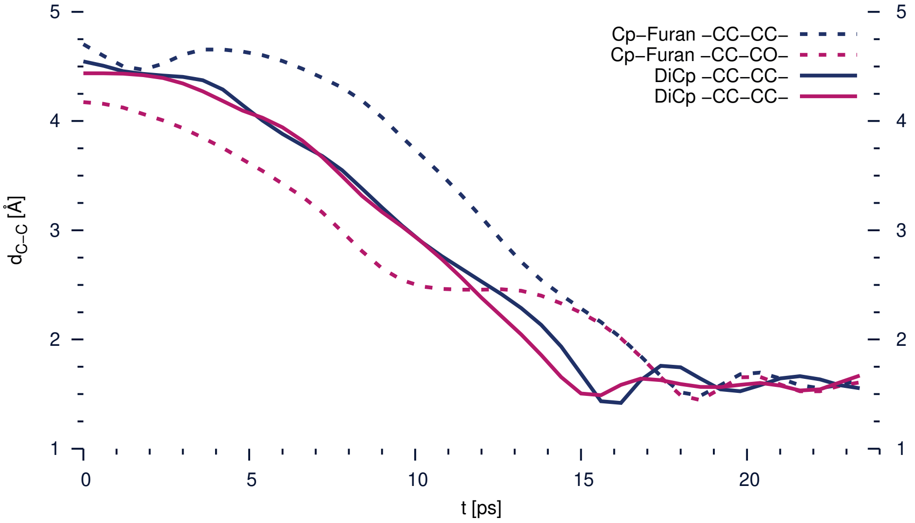


Figure 9.8: Evolution of the carbon-carbon (CC) distances that constitute the future bonds for the formation of dicyclopentadiene (DiCp) and the cyclopentadiene-furan adduct (Cp-Furan). Adjacent atoms are indicated in the labels.

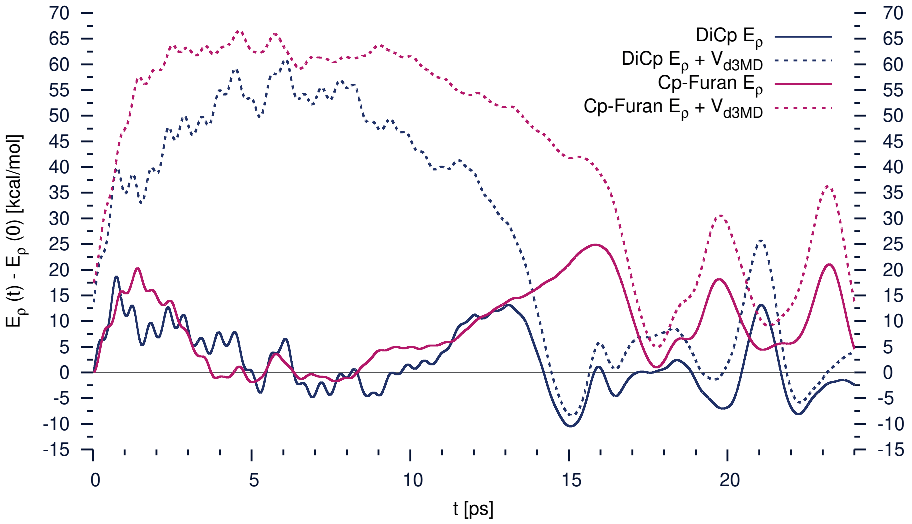


Figure 9.9: Biased (E_ρ+V_{d3MD}) and unbiased (E_ρ) potential energy profile for the reaction of cyclopentadiene to dicyclopentadiene (DiCp) and the reaction of cyclopentadiene with furan (Cp-Furan). Clearly, the potential energy at the putative 'transition state' is lower for the favoured reaction, *e.g.* the formation of DiCp.

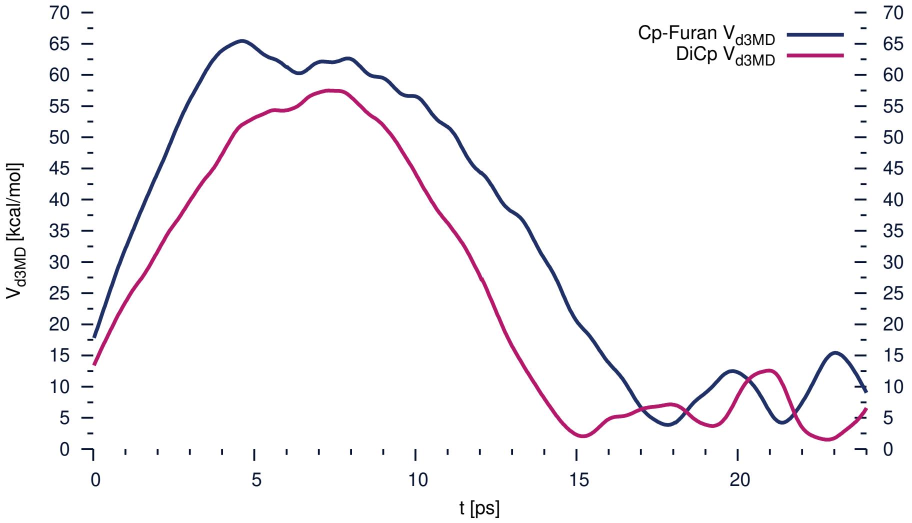


Figure 9.10: Strength of the biasing potential $V_{\mathrm{d3MD}}(t)$ as a function of t for the dimerisation of cyclopentadiene to dicyclopentadiene (DiCp) and the reaction of cyclopentadiene with furan (Cp-Furan); the latter being kinetically disfavoured over the former.

dicyclopentadiene (DiCp) as well as the cyclopentadiene-furan adduct (Cp-Furan). It can be seen that the biasing potential is stronger for a given value of $\kappa(t)$ in the case of the less favoured reaction (Cp-Furan); interestingly, this reflects chemical intuition (the disfavoured, sluggish reaction should have a higher barrier that needs to be overcome).[266]

Fig. 9.9 shows the underlying unbiased potential energy (PE) as well as the biased d3MD-PE (dashed lines). Note that contrary to the PE, the biasing potential itself (Fig. 9.10) is much smoother, and that it is comparably low close to the barrier on the PE profile. Again, the reluctance of Cp-Furan to form is reflected in a higher barrier on the PE profile, which also occurs later through the trajectory. It should be noted that the unbiased PE exhibit some features that can be linked to concepts from organic chemistry: Fast reactions have an early transition state, *i.e.* a transition state that resembles the reactants rather than the product, whereas slow processes tend to have a late transition state that bears more resemblance to the product instead.[266] This can be linked to the CC bonds in Fig. 9.8; the larger potential energy contribution for the formation of the Cp-Furan adduct appears to result from the asymmetric approach of the reactant, *i.e.* the earlier approach of the carbon adjacent to the oxygen and the corresponding delayed approach of the other reactive carbon. Hence, in the case of DiCp, the configuration of the approaching reactants does not deviate significantly from the reactive configuration, indicating an early transition state. In contrast, for the Cp-Furan adduct, the asymmetric configuration of the approaching reactants implies that the transition state bears more resemblance to the product than the reactant, corresponding to a late transition state. These observations are encouraging with regard to a possible predictive power of the method, which would extend its applicability beyond simple structure elucidation. However, the physical origin of those observations is as of yet unclear; it has yet to be demonstrated that the chemically intuitive phenomena observed so far are rooted in physics in order for them to be more than just anecdotal evidence.

9.2.3 Electron Transfer in d3MD

A preliminary analysis of the effect of the biasing potential has been performed based on the d3MD-electron densities. Fig. 9.11 visualises the effect of the d3MD potential $v_{\mathrm{d3MD}}(\mathbf{r})$ on the electron density; compared is the N-electron density of the perturbed and the unperturbed system. Clearly, in a d3MD simulation, a small amount of electron density is transferred from the reactant to the product state with respect to a reference density $\rho(\mathbf{r})$. The effective absolute values of electron transfer have been quantified by

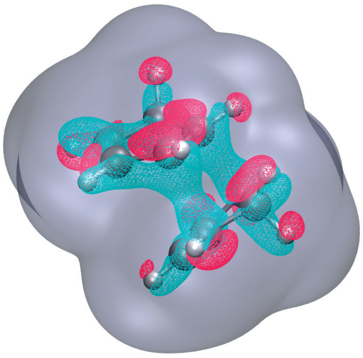


Figure 9.11: Perturbation introduced by the d3MD potential. A density contour cloud is shown in grey, along with the density difference of the same magnitude. Density differences were calculated between the d3MD density and an unperturbed density. Regions of depletion are shown in red, density accumulations in turquoise.

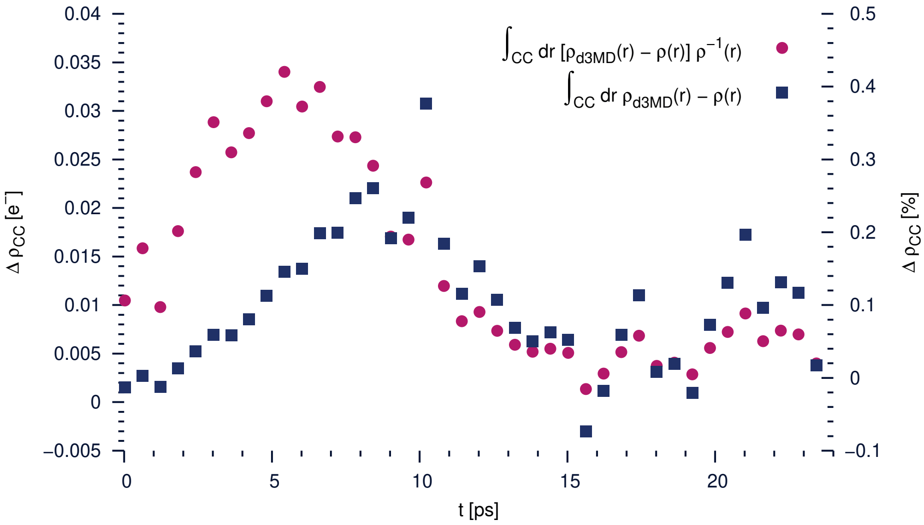


Figure 9.12: Integral over the difference between a biased (d3MD) and unbiased electron density for structures along the formation of DiCp from its monomers. Relative values are given on the secondary y-axis and refer to the differences normalised by the value of the reference density.

integrating the electron density within 1.57 Å spheres centred in the middle of the basis vectors connecting the bond-forming carbons:

$$\Delta\rho_{\mathrm{CC}} = \int_{\mathrm{CC}} \mathrm{d}\mathbf{r}\ \left(\rho_{\mathrm{d3MD}}(\mathbf{r}) - \rho(\mathbf{r})\right), \tag{9.16}$$

where CC denotes the integration sphere around the centres of the carbon-carbon bonds. As can be seen from Fig. 9.12, only a minute amount of electron density is shifted to the new bond. Relative values have been obtained by dividing by the integral of the unbiased electron density within the same spheres, and as can be seen from Fig. 9.12, maximum transfer rates do not even reach 0.5 %. The d3MD methodology therefore only introduces slight perturbations on the electron density, which appear nonetheless significant enough to influence the dynamics of the nuclei and ultimately lead to a reactive event.

9.2.4 Versatile Grid Interpolation for Electron Density Maps

In view of future applications of d3MD which may resort to experimental density difference maps, a suitable interpolation between the experimental mesh and the real-space mesh used in a plane wave code has to be performed. In order to ensure that the interpolated density difference be as smooth as possible, a Gaussian interpolation algorithm has been devised that accepts any orthogonal grid, irrespective of the orientation of the grid vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ (which are general 3-tuples of Cartesian coordinates) in Cartesian coordinate space.

The output basis vectors have to be diagonal in Cartesian space and will be denoted $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Based on the input and output basis vectors, a set of auxiliary output vectors is computed:

$$\mathbf{x}' = \begin{pmatrix} ||\mathbf{a}||^{-1}\mathbf{a}\cdot\mathbf{x} \\ ||\mathbf{b}||^{-1}\mathbf{b}\cdot\mathbf{x} \\ ||\mathbf{c}||^{-1}\mathbf{c}\cdot\mathbf{x} \end{pmatrix}, \quad \mathbf{y}' = \begin{pmatrix} ||\mathbf{a}||^{-1}\mathbf{a}\cdot\mathbf{y} \\ ||\mathbf{b}||^{-1}\mathbf{b}\cdot\mathbf{y} \\ ||\mathbf{c}||^{-1}\mathbf{c}\cdot\mathbf{y} \end{pmatrix}, \quad \mathbf{z}' = \begin{pmatrix} ||\mathbf{a}||^{-1}\mathbf{a}\cdot\mathbf{z} \\ ||\mathbf{b}||^{-1}\mathbf{b}\cdot\mathbf{z} \\ ||\mathbf{c}||^{-1}\mathbf{c}\cdot\mathbf{z} \end{pmatrix}, \tag{9.17}$$

where $\mathbf{a}\cdot\mathbf{x}$ denotes a dot product. The auxiliary vectors $\mathbf{x}', \mathbf{y}', \mathbf{z}'$ are nothing but the diagonal (conventional) output vectors expressed within the rotated basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$. After translating the origins of the coordinate systems to coincide with each other, it then becomes easy to calculate distances between discrete points in both coordinate systems. We shall denote discretised grid points in the output mesh as $\mathbf{R}$, use $\mathbf{Q}$ for the input mesh and express both of them in terms of the auxiliary output coordinates $\mathbf{x}', \mathbf{y}', \mathbf{z}'$. We then write the discrete elements of the output density difference $\Delta\rho(\mathbf{R})$ as a linear combination of elements of the input density difference $\Delta\rho(\mathbf{Q})$, weighted by a Gaussian $g(\mathbf{R}, \mathbf{Q})$:

$$\Delta\rho(\mathbf{R}) = N \sum_{\mathbf{Q}} \Delta\rho(\mathbf{Q}) g(\mathbf{R}, \mathbf{Q}), \tag{9.18}$$

$$g(\mathbf{R}, \mathbf{Q}) = e^{-\sigma_x^{-2} ||\mathbf{e}_{x'} \cdot (\mathbf{R}-\mathbf{Q})||^2} e^{-\sigma_y^{-2} ||\mathbf{e}_{y'} \cdot (\mathbf{R}-\mathbf{Q})||^2} e^{-\sigma_z^{-2} ||\mathbf{e}_{z'} \cdot (\mathbf{R}-\mathbf{Q})||^2}, \tag{9.19}$$

$$N = \frac{\sum_{\mathbf{Q}} \Delta\rho^2(\mathbf{Q})}{\sum_{\mathbf{R}} \Delta\rho^2(\mathbf{R})}. \tag{9.20}$$

$\mathbf{e}_{\{x,y,z\}}$ denote unit vectors in the direction of $\mathbf{x}, \mathbf{y}, \mathbf{z}$, $\sigma_{\{x,y,z\}}$ are Gaussian widths and N is a normalisation constant that preserves the L^2 norm of the input and output density difference. In practice, to save computational resources, the sum in eq. 9.18 is only taken over non-vanishing elements of $g(\mathbf{G}, \mathbf{Q}) > \min$, which is easily determined from σ. This algorithm has been directly implemented in the d3MD driver of the CPMD code and requires no user input; the $\sigma_{\{x,y,z\}}$ are determined from the input grid dimensions such that the distance matrix calculated for eqs 9.18 and 9.19 can be restricted to contain at maximum 5^3 elements with no loss of accuracy.

We have verified that the use of a coarse input grid does not affect the ability of a test system to isomerise by revisiting the isomerisation of 2-butene. Progressively coarser reference density differences have been generated comprising grid spacings of up to 2.4 a.u.; Fig. 9.13 shows unbiased and biased potential energies for reference density differences of varying resolution. For all simulations but the $27 \times 27 \times 10$ mesh, isomerisation could be observed. While the granularity of the reference density difference seems to slightly influence the time at which the reaction occurs (at identical initial conditions), the potential energy of the putative transition state structure does not change, and the evolution of the biased potential energy is very similar for the different meshes. When employing a very coarse $27 \times 27 \times 10$ mesh, no reaction takes place, since the biasing potential in this simulation is very low and can take negative values even for the reactant.

It seems that a minimal resolution for the input grid still has to be maintained since the reference density difference from which the potential is constructed will otherwise be too coarse. Still, for grid spacings of up to ~ 1 Å, coarse reference data can reliably be treated thanks to the interpolation algorithm that has been described here. Given that today's experiments can achieve resolutions lower than 1 Å, the feasibility of simulations based on

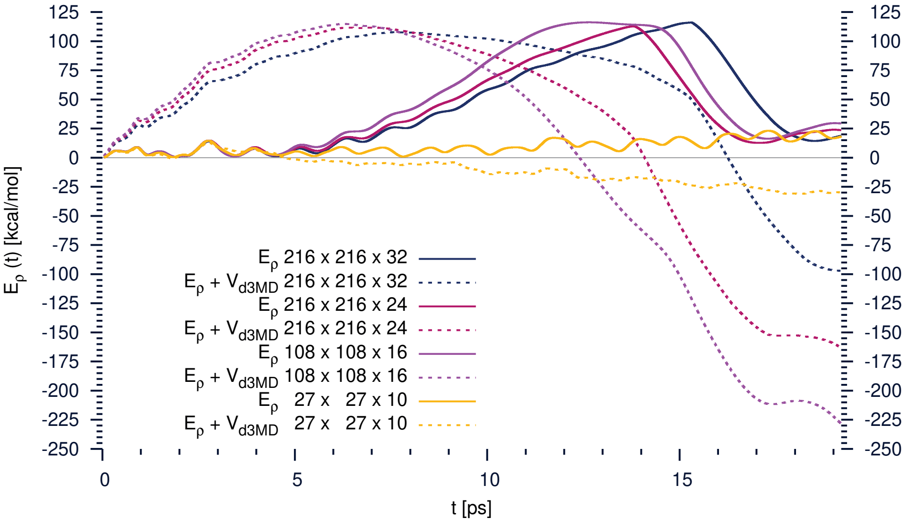


Figure 9.13: d3MD of the isomerisation of *trans*-2-butene to *cis*-2-butene performed with reference density differences of differing resolution. During the simulations, the electronic density is resolved on a $216 \times 216 \times 160$ mesh in a $20 \times 20 \times 15$ Å^3 supercell. Apart from the $27 \times 27 \times 10$ reference density difference grid, all simulations led to an isomerised structure.

---

experimental maps will have to be more carefully assessed in a case-by-case evaluation.

Chapter 10

Conclusions and Outlook

What is the ultimate goal of Computational Chemistry? Opinions will differ, as do priorities. While certain events can only be observed at very long time scales and can be difficult to treat using conventional enhanced sampling algorithms, others may lend themselves to shorter time scale studies but suffer from the inaccuracy of (one single) PES. So: does the key to an accurately sampled free energy surface lie in improved sampling or in improved accuracy of the PES?

Eventually, both. The issue boils down to the computational tractability of the underlying methods. Yes, ever increasing computational power and readily available resources do indeed allow for far more accurate simulations to be performed than one or two decades ago. But while more CPU can provide more answers, they can also pose more questions, in particular regarding most efficient usage. Modern quantum chemical codes are highly parallel, efficient and hence complex; many years of development and modifications by different authors may hamper new implementations, often requiring that a substantial amount of time be dedicated to the understanding of their inner workings, long before it can be invested into their improvement[i].

In the field of DFT-based *first principles* MD, the main issues are well known; it is the improbable combination of tractability and accuracy of the underlying density functional that ultimately constitute a prohibitive bottleneck for efficient sampling. While $\mathcal{O}(N)$ simulations are straightforwardly and most elegantly implemented in a plane wave basis, they are limited to

---

[i] Fellow CPMD developers might remember tongue-twisting eye-burning finger-twiddling classics such as `tpiba2`, `lrxpl`, `inzf` and `nzfs`. And, of course, the chameleon `rhoe` that almost never contains the quantity that one would expect it to hold: $\rho(\mathbf{r})$.

(semi-)local functional forms. As we have seen in Chapters 5 and 6, such methods can perform satisfactorily in many cases - in particular if they combine highly flexible functional forms with fits on large reference databases[ii]. However, we also had to acknowledge that the average performance of higher-level functionals which include a fraction of the exact exchange of the KS determinant is all too often much superior compared to semi-local forms. It can at present be very difficult to avoid hybrids - or their range-separated derivatives - if not utterly impossible for certain systems (CT complexes and Rydberg states). This need for hybrid functionals in turn exposes one principal weakness of the plane wave/pseudopotential approach, which is the high computational overhead linked to the evaluation of the FFTs that appear with the exchange integrals. While for insulators, efficient algorithms have been described that can achieve linear scaling for large systems,[160] the prefactor is still sizeable and has prevented hybrids from being abundantly used in plane wave MD.

The developments sketched in this text are an attempt at alleviating those problems by either providing an existing methodology within a plane wave framework to obtain more accurate PES for on-the-fly *first principles* MD, or by designing new algorithms that can ultimately save precious computational time and thereby allow to improve sampling. We have started in the excited state by demonstrating that the Coulomb-attenuation method can be straightforwardly implemented in a plane wave/pseudopotential code and that the results obtained from Coulomb-attenuated functionals generally show excellent agreement with commonly employed atom-centred basis sets. Nonetheless, we have also seen that in the case of excitation energies, a plane wave basis can provide a more robust reference than Gaussian functions, as many commonly employed Gaussian basis sets do either not predict excitations that occur in the limit of larger, more diffuse (and computationally expensive) bases or give different overall results. This now makes it possible to study the excited-state dynamics and properties in a completely delocalised basis and, in combination with the QM/MM engine of the CPMD code, to study such processes in a realistic (be it solvent or protein) environment.

We have then focused our attention on ground-state thermodynamics at the example of the widely used family of Minnesota functionals, hitherto unavailable in plane waves. By providing a comprehensive convergence analysis of the M05 to M11 families of xc functionals, we identified the dual ζ to be a primordial ingredient if converged results are to be obtained, and we have linked this observation with the commonly stated need for very fine integration grids in atom-centred bases. While we were able to show that plane wave results compare very well with those obtained in a Slater basis, deviations with respect to Gaussian bases were often larger and inconsistent in nature. In the case of the two largest deviations, we could link this behaviour to the exceptionally high basis-set sensitivity of the affected functionals and we have shown that this can be related to the integral of the exchange energy density outside of the pseudisation radius. While caution is advised when Minnesota functionals are used with plane waves due to their general basis set sensitivity, it is important to note that the highly flexible plane waves only expose existing fallacies in the construction of the functionals that went unnoticed in the (more restricted) atom-centred basis sets. There is no counterindication to the use of the Minnesota family for plane wave MD, as long as the reference results from literature are carefully compared (and transferred) to the completely flexible and delocalised basis that are plane waves. Due to the encouraging results that have been obtained with Minnesota functionals, especially for barrier heights and certain transition metal complexes, the availability of these functionals can prove valuable in particular for the simulation of biological processes (which often involve transition metals) or reactive events in general (be it a chemical process in solution or a biological system).

[ii] And even if such an approach is often highly disputed.

With both of the implementations heavily relying on exact exchange, we have suggested an approach that can substantially reduce the computational cost associated to the evaluation of the determinantal exchange potential in a plane wave basis. This method is based on the use of coordinate-scaled pair densities and potentials in the computation of the exact exchange terms. In doing so, we could show how concepts from theoretical and conceptual DFT can inspire the development of new, more efficient computational algorithms, and we have also discussed what formal conditions have to be met in order for those relations to be valid in a generalised context. Using our coordinate scaling scheme, we could document speedups of up to a factor of 8.5, which is almost one order of magnitude. This substantially accelerates the simulation of nonperiodic systems with more accurate hybrid functionals. Those clusters can themselves be embedded in much larger systems (such as in a QM/MM context), which finally makes it possible to sample sufficiently large time scales with hybrid functionals on conventional server infrastructure, without having to resort to an exceptionally large number of CPU on specialised architecture.

In our last attempt at improving the PES in the perspectives of Chapter 9, we have shown how the accuracy of semi-local functionals can be improved by calibrating new DCACP for some metals and metalloids. We have presented a fit uniquely based on literature data rather than tailored high-level wavefunction calculations without losing accuracy of the resulting potentials. This approach allows for a rapid generation of new DCACP, which, as we have shown at the examples of gold and boron, can outperform popular dispersion correction schemes such as DFT-D2 and DFT-D3. With the DCACP of this work being the first reported for metals, we were able to show that despite the weak magnitude of the dispersion correction, rare gas complexes with metals can be reliably described. At last, we have assessed the performance of the gold DCACP for surface-adsorbate interactions. While the metal DCACP works very well for aurophilic complexes and yields acceptable structural minima for the interaction of benzene with a Au(111) surface, the use of DCACP results in an unphysical long-range decay of the potential. This has been documented before for systems in the gas phase, with the notable difference that the use of a two-channel DCACP that usually remedies this pitfall only slightly improves the long-range behaviour. This raises questions on the origin of the long-range decay of DCACP-mediated van der Waals interactions. Future work might hold the answer to this question. In the long term, obtaining a correct long-range decay would allow metal DCACP to be calibrated and routinely used for the study of surface-adsorbate interactions.

The final approach outlined in this text offers a complementary perspective of the work dedicated to the improvement of the PES: Our Density-Difference Driven MD (d3MD) scheme allows to overcome time-scale restrictions by the use of electronic biasing potentials. By exploiting electron density difference maps between products and reactants, we were able to isomerise *trans*- to *cis*-2-butene and to initiate Diels-Alder cycloadditions between two molecules of cyclopentadiene as well as cyclopentadiene and furan, respectively. We have shown that those reactions are mediated by a low-magnitude density transfer between product and reactant state and that chemical trends are reflected in the evolution of the d3MD biasing potential under otherwise identical conditions. We have also introduced a general interpolation algorithm which allows for the use of coarser reference density difference maps, and we have successfully demonstrated that grid resolutions of up to 1 Å are sufficient in order for a reactive event to take place. With modern experimental time-resolved techniques reaching resolutions of 2 Å, this leaves open the possibility of using experimental density difference maps for the structural refinement of transient, meta-stable or stable biological structures. This constitutes a promising possibility for all cases where excessive assumptions have to be made in conventional refinement algorithms, which can insert an unphysical bias into the final result. And not at least, the results obtained for the pericyclic reactions suggest that, once properly formalised, it might be possible to extract some physical (equilibrium) behaviour out of the d3MD simulations, which could be combined with nonequilibrium approaches (*cf.* Jarzynski's theorem[267–269]). This is an exciting perspective in particular with respect to the absence of (arbitrary) explicit collective variables (we may consider d3MD to implicitly modify the chemical potential by perturbing the electron density itself), but also in view of the low computational cost of the method. Failing that, the method can still provide an excellent, chemically intuitive and computationally efficient guess for a transition path from A to B, which can then be the basis for more refined enhanced sampling techniques.[33,270]

The methods presented here can ultimately benefit the simulation of condensed matter by providing more accurate tools at a reduced computational cost and by potentially overcoming the time-scale limit due to rare events. But our field is still riddled with open questions. Ruminations on the ultimate limits of semi-local density functional approximations,[105,106] advances in wavefunction methods and their computational tractability,[5] progress in orbital-free DFT,[271,272] density matrix functionals,[97] the increasing importance of complex fitting procedures,[273–276] the need to incorporate nuclear quantum effects into dynamics[25–27] and not least the new possibilities and issues that will accompany the ascent of quantum computing[277,278] have the potential of haunting and inspiring many a theoretician for years to come.

- and very luckily so! For how frustratingly boring would the world of chemistry be if the answers to those questions were already and trivially known.

PART V

Postscript

CHAPTER A

Supporting Information

A.1 Supporting Information for Chapter 5

A.1.1 HOMO-LUMO Gaps in a Plane Wave/Pseudopotential Basis

Gaps obtained from a soft pseudopotential show convergence at 120 Ry for all systems. For a small simulation supercell ($l = 10$ Å), choosing a lower cutoff value of 70 Ry introduces a substantial error of 0.3 eV. A notable error is still present at 70 Ry even for the two larger simulation cells, but it becomes less relevant for practical purposes, since the maximal error of 0.1 eV lies below the typical accuracy of the functional itself. Values at 100 Ry are essentially converged in all supercells, with errors being lower than 0.05 eV, and full convergence is reached at 120 Ry.

Table A.1: HOMO-LUMO gap of a water molecule calculated using CAM-B3LYP and MT pseudopotentials at increasing energy cutoffs and varying simulation cell lengths l representing an isolated and periodic system, respectively.

E_{cut} [Ry]	$\Delta\epsilon_i$ [eV], isolated system			$\Delta\epsilon_i$ [eV], periodic system		
	10 Å	20 Å	30 Å	10 Å	20 Å	30 Å
70	10.082	10.457	10.479	10.111	10.458	10.479
80	10.234	10.510	10.527	10.259	10.511	10.527
100	10.360	10.546	10.562	10.381	10.547	10.562
120	10.375	10.549	10.565	10.396	10.550	10.565
150	10.377	10.550	10.566	10.397	10.551	10.566
180	10.379	10.551	10.567	10.399	10.552	10.567

The convergence behaviour is analogous to the one observed for GGA or standard hybrid functionals once the simulation cell is of sufficient size: Changes in the gap are still substantial when increasing the side l of the cubic simulation cell from 10 Å to 20 Å, with changes in the converged gap of 0.2 eV. The change in gap is within the usual numerical tolerance (< 0.5 eV) for a further extension to 30 Å, emphasising again the importance of an appropriately large cell for the correct description of the system's LUMO.

The results for the truly periodic system (where the Poisson equations are not decoupled) show the same trend, with the only relevant difference with respect to the isolated system occurring in the 10 Å box. Given the trends observed for the isolated system, this is most likely attributed to spurious interactions between periodic images at this intermolecular distance. The

Table A.2: Size l of the orthorhombic periodic supercell for all systems considered in this study. For structures and abbreviations, see Ref. 140.

System	l_x [a.u.]	l_y/l_x	l_z/l_x
7-azaindole	34.9555	1.00	1.00
AT nucleobase pair	71.8096	0.40	0.40
4-(N,N-dimethylamino)benzonitrile (DMABN)	34.9599	1.00	1.00
p-nitroaniline	34.9599	1.00	1.00
Acene (n=1 / naphthalene)	34.9556	1.00	1.00
Acene (n=2)	45.0000	1.00	0.50
Acene (n=3)	45.0000	1.00	0.50
Acene (n=4)	45.0000	0.75	0.50
Acene (n=5)	60.0000	0.50	0.33
Polyacetylene (PA) oligomer (n=2)	45.0000	0.66	0.50
Polyacetylene (PA) oligomer (n=3)	30.0000	1.00	0.50
Polyacetylene (PA) oligomer (n=4)	30.0000	1.00	0.50
Polyacetylene (PA) oligomer (n=5)	45.0000	0.50	0.50
N-phenylpyrrole (PP)	45.0000	0.66	0.50
Dipeptide	34.9599	1.00	1.00
Tripeptide	60.0000	0.50	0.25
β-dipeptide	69.9199	0.50	0.50
Retinal protonated Schiff base	71.8096	0.40	0.40
CO	18.9036	1.00	1.00
H_2CO	18.9036	1.00	1.00
HCl	20.0000	1.00	1.00
N_2	10.0000	1.00	1.00

considerable influence of these interactions on the gap supports the view that the requests on the Tuckerman-Martyna Poisson solver are not met yet either, since the simulation cell must span at least twice the spatial extent of the charge density.[155] The strong changes in gaps when increasing the cutoff within the small simulation cell is hence most probably due to an insufficient cell size for both the isolated and periodic system, resulting in an incomplete description of the electron density.

A.1.2 Supercell Size for Plane Wave Calculations

Plane wave calculations were carried out using the supercell sizes tabulated in Table A.2.

A.1.3 Excitation Energies for the Test Set by Peach *et al.*

Table A.3 summarises the excitation energies for the molecules contained in the test set by Peach *et al.*[140] that have not been included in the main text. All excitation energies were calculated with both CAM-B3LYP and CAM-O3LYP in a plane wave/pseudopotential framework as described in the Computational Setup.

A.2 Supporting Information for Chapter 7

A.2.1 Translational Invariance

Table A.4 compares total energy, HOMO-LUMO gaps and the first 5 excited states obtained from the standard $v_{ij}(\mathbf{r})$ to those obtained from the procedure outlined in the main text for DMABN. The configuration of DMABN has been distorted by 0.05 Å in the Cartesian x and z directions. The resolution of the full density grid is $\Delta R_x = \Delta R_z = 0.096\,\text{Å}$. Comparison to Table 1 in the main text clearly indicates that changes in the geometry smaller than the grid spacing do not influence the error of the method. Forces are of the same accuracy as mentioned in the main text. The method is therefore suitable for geometry optimisations and molecular dynamics.

Table A.3: Excitation energies for the molecules contained in the test set by Peach *et al.*[140] that have not been tabulated in the main text. Calculations used a cutoff value for the KS orbitals of 75 Ry. For structures and abbreviations, see Ref. 140.

State	CAM-B3LYP	CAM-O3LYP	Ref.	State	CAM-B3LYP	CAM-O3LYP	Ref.
Acene (n=2)				Acene (n=3)			
S_1	3.71	3.83	4.46	S_1	2.98	3.10	2.90
S_2	4.05	4.08	4.88	S_2	3.68	3.72	3.52
Acene (n=4)				Acene (n=5)			
S_1	2.47	2.60	2.35	S_1	2.16	2.30	1.95
S_2	3.43	3.48	3.27	S_2	3.20	3.33	3.09
CO				H_2CO			
S_1	8.89	8.86	8.51	S_1	3.93	3.85	3.94
S_2	8.89	8.86	9.88	S_2	6.27	6.18	7.09
S_3	9.77	9.80	10.23	S_3	7.44	7.34	7.97
S_4	9.86	9.81	10.87	S_4	7.66	7.55	8.12
HCl				N_2			
S_1	7.43	7.43	8.23	S_1	9.33	9.48	9.31
				S_2	9.54	9.53	9.92
				S_3	9.54	9.53	10.27
				S_4	9.60	11.01	12.20
Polyacetylene (PA) oligomer (n=2)				Polyacetylene (PA) oligomer (n=3)			
S_1	4.10	4.20	5.92	S_1	5.55	5.77	4.95
Polyacetylene (PA) oligomer (n=4)				Polyacetylene (PA) oligomer (n=5)			
S_1	3.21	3.21	4.49	S_1	4.54	4.63	4.49
N-phenylpyrrole (PP)				Tripeptide			
S_1	5.03	5.08	4.85	S_1	3.82	5.74	5.74
S_2	5.06	5.16	5.13	S_2	4.16	5.95	5.61
S_3	5.22	5.37	5.47	S_3	4.68	5.98	5.91
S_4	5.25	5.40	5.94	S_4	5.02	6.00	7.01

Table A.4: Total energy, HOMO-LUMO gap $\Delta\epsilon_{\mathrm{HOMO}}^{\mathrm{LUMO}}$ and singlet excitation energies for the organic dye DMABN obtained from the standard $v_{ij}(\mathbf{r})$ and its scaled counterpart presented in this work ($\lambda = 0.5$) on a geometry of DMABN distorted by $\Delta r = 0.05\,\text{Å}$ in the x and z direction. Calculations were carried out in an orthorhombic $37.5 \times 18.75 \times 18.75\ \text{Å}^3$ supercell at a plane wave cutoff of 75 Ry using the B3LYP[31,163] xc functional.

	E_{tot} [a.u.]	$\Delta\epsilon_{\mathrm{HOMO}}^{\mathrm{LUMO}}$ [eV]	S_1 [eV]	S_2 [eV]	S_3 [eV]	S_4 [eV]	S_5 [eV]
Full $v_{ij}(\mathbf{r})$	-76.689521	4.7589	4.408	4.465	4.752	4.992	5.078
Scaled $v_{ij}(\mathbf{r})$	-76.689527	4.7590	4.408	4.450	4.753	5.080	5.140
Difference	$6.0 \cdot 10^{-6}$	0.0001	0.000	0.015	0.001	0.088	0.062

List of Figures

5.1 Molecules explicitly discussed in this study, all of which contain excitations that are difficult to describe in TDDFT when using a GGA or hybrid functional as the xc kernel. 84
5.2 Pseudocode describing the structure of the main routine of the new xc driver (`cp_xc_driver`) in the CPMD code. 90

6.1 Citations per year for the seminal M06 paper[43] (source:webofknowledge). As of July 2018, the paper has accumulated a total of 10545 citations, making it the most cited of the Minnesota papers. 108
6.2 The HC7/11 set of difficult hydrocarbon reactions.[43] 109
6.3 Convergence of the binding energy ΔE of a HF dimer, grouped according to functionals, with respect to cutoff energy E_{cut} and ξ. Converged binding energy values are illustrated with a line. 120
6.4 Zoom on the values of the energy density $\epsilon(\mathbf{r})$ for exchange x, correlation c and exchange-correlation xc for the M06, M06-L, M06-2X and M06-HF functional at the example of a cut through a carbon-hydrogen bond of product E31 of reaction **2**. The electron density is displayed on the secondary y axis. 135
6.5 Values for the ratio of change in energy around the nuclei, ΔE^{spher}, and the total change in energy, ΔE, for exchange only x as well as exchange-correlation xc for the M05, M06, M06-L, M06-2X and M06-HF functional for the isomerisation reaction **2**. 136
6.6 Isosurface (0.0035 a.u.) of the electron density (grey) and the electron density difference between a plane wave setup and aug-cc-pVTZ for compound E1. Regions depicted in red show more density in aug-cc-pVTZ, regions in turquoise have more density in a plane wave basis. Panel a) and c) show densities for M06, panels b) and d) for M06-HF. 138

6.7 Time per DIIS step for the HF dimer using the M06 and M06-L xc functionals at different cutoff energies and using different values of ξ. All calculations were carried out in a $20 \times 15 \times 10$ Å box using 384 processors. 139

7.1 Illustration of the coordinate scaling at the example of the hydrogen molecule. Shown are a cut of the N-electron density $\rho(\mathbf{r})$ along the bonding axis, as well as its one-dimensional coordinate-scaled analogue $\rho_\lambda(\mathbf{r})$ for $\lambda = 0.5$. 149

8.1 Illustration of the coordinate scaling approach at the example of the x-pencils over which the FFTs are parallelised. Every index i_x in x corresponds to an entire yz plane of the periodic supercell. An example of some (pair) density is shown in the upper part, its coordinate-scaled analogue is depicted below. Points correspond to discretised real-space mesh points indexed by i_x. The choice of $\lambda = 0.5$ allows for a straightforward mapping between the original density and its coordinate-scaled counterpart. The MPI parallelisation is visualised in the form of N_P white and grey blocks, which correspond to one MPI task each. In the case of the reference mesh, every task holds 4 yz planes, whereas there are only 2 yz planes for the coarse, coordinate-scaled mesh. If the non-vanishing parts of the density are distributed on the coordinate-scaled, coarser mesh shown in the lower half of the figure, different tasks have to communicate with each other. By moving from left to right, it can be seen that the first two tasks (I and II) send no data, whereas they receive density from task III. Task III in turn sends its data to both tasks I (first two gridpoints) and II (subsequent two gridpoints). Tasks IV and V will retain parts of their gridpoints; all the others completely swap their data. Once the full potential $v_{ij}(x, y, z)$ has to be obtained from its scaled counterpart $v_{ij}^{\lambda}(X, Y, Z)$, the mapping is inverted. If CP groups are used, this mapping can be carried out independently on all N_C groups without any inter-group communication. 168

8.2 Average execution time and speedups for an average step of wave-function optimisation for naphthalene using the ODIIS algorithm at different blocking ratios b/N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. Numbers of processors N_P for every simulation supercell size are given in Table 8.1; only one CP group is used. 172

8.3 Average execution time and speedups for the calculation of the first four excited singlet states of naphthalene at different blocking ratios b/N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. Numbers of processors N_P for every simulation supercell size are given in Table 8.1; only one CP group is used. 173

8.4 Average execution time and speedups for one step of wavefunction optimisation using the ODIIS algorithm at fixed supercell size for systems with different numbers of orbitals N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12×16 processor cores. 175

8.5 Average execution time and speedups for one step of optimisation of the first four Singlet states with linear-response TDDFT. All calculations were carried out at fixed supercell size for systems with different numbers of occupied orbitals N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12×16 processor cores. 176

8.6 Average execution time and speedups for one step of wavefunction optimisation as well as for a TDDFT calculation involving the first four singlet states at fixed supercell size for systems with different numbers of orbitals N_b and different numbers of CP groups N_C. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12×16 processor cores. 177

8.7 Average execution time and speedups for one step of wavefunction optimisation at N_b as well as for a TDDFT calculation for the first four singlet states at fixed supercell size for systems with different numbers of orbitals N_b. Dashed lines with dots denote a full reference calculation of the exact exchange energy and potential, straight lines with squares the coordinate-scaled exchange. All calculations were carried out on the same 12×16 processor cores; values are compared between optimal parallelisation with different N_C, N_P and $b = N_b/N_C$. 178

9.1 Availability of DCACP for the BLYP xc functional: Currently published and tested (yellow), unpublished and tested in the current work (red), calibrated and tested in the current work (blue). This table is a derivative of Pekka Pyykkö's 172-element (extended) periodic table by Pekka Pyykkö and WhiteTimberwolf, used under CC BY-SA 3.0, initially reproduced from Ref. 240 with permission of the author. 187

9.2 Performance of a one- and two-channel (2c) DCACP for the physisorption of benzene on a Au(111) surface on a $6 \times [12 \times 12]$ (bottom) and $6 \times [5 \times 6]$ (top) slab. Values obtained using DFT-D2 are given for comparison. The reference data is interpolated from experiment; it has been shown that by altering the collision factor, the values could change by up to 20 kcal mol^{-1}.[260] 191

9.3 Decay of the surface-adsorbate interaction for the graphs of Fig. 9.2 fit on $ad_{||}^{-x} + b$ for fixed $x = 4$, $x = 6$ and using x as a variable (free). Optimised parameters are $x = 7.55$ for DCACP, $6 \times [12 \times 12]$ slab, $x = 7.51$ for DCACP, $6 \times [6 \times 5]$ slab and $x = 5.16$ for 2c-DCACP, $6 \times [6 \times 5]$ slab. 192

9.4 Isomerisation of *trans*-2-butene to its *cis* form by performing d3MD at 400 K and using $\kappa(t) = 0.50$ a.u. The N-electron density is depicted in lilac. Snapshots were taken from 0 ps (panel a) to 15 ps (panel d) and are spaced by 5 ps. 195

9.5 Biased ($E_\rho + V_{\text{d3MD}}$) and unbiased (E_ρ) potential energy profile for the isomerisation of *trans*-2-butene to its *cis* form. The unbiased potential energy profile shows that an effective barrier of about 100 kcal mol^{-1} has to be overcome. 196

9.6 Dimerisation of two molecules of cyclopentadiene to yield dicyclopentadiene. Shown is the N-electron density in lilac. Snapshots were taken from 3 ps (panel a) to 24 ps (panel h) and are spaced by 3 ps. 198

9.7 Reactants and product for the formation of dicyclopentadiene from cyclopentadiene along with the reference density difference map $\Delta\rho_{\text{Ref}}(\mathbf{r})$. Regions of density depletion (negative density difference) are depicted in red, regions of accumulation (positive density difference) in turquoise. 199

9.8 Evolution of the carbon-carbon (CC) distances that constitute the future bonds for the formation of dicyclopentadiene (DiCp) and the cyclopentadiene-furan adduct (Cp-Furan). Adjacent atoms are indicated in the labels. 199

9.9 Biased ($E_\rho + V_{\text{d3MD}}$) and unbiased (E_ρ) potential energy profile for the reaction of cyclopentadiene to dicyclopentadiene (DiCp) and the reaction of cyclopentadiene with furan (Cp-Furan). Clearly, the potential energy at the putative 'transition state' is lower for the favoured reaction, *e.g.* the formation of DiCp. 200

9.10 Strength of the biasing potential $V_{\text{d3MD}}(t)$ as a function of t for the dimerisation of cyclopentadiene to dicyclopentadiene (DiCp) and the reaction of cyclopentadiene with furan (Cp-Furan); the latter being kinetically disfavoured over the former. 200

9.11 Perturbation introduced by the d3MD potential. A density contour cloud is shown in grey, along with the density difference of the same magnitude. Density differences were calculated between the d3MD density and an unperturbed density. Regions of depletion are shown in red, density accumulations in turquoise. 202
9.12 Integral over the difference between a biased (d3MD) and unbiased electron density for structures along the formation of DiCp from its monomers. Relative values are given on the secondary y-axis and refer to the differences normalised by the value of the reference density. 202
9.13 d3MD of the isomerisation of *trans*-2-butene to *cis*-2-butene performed with reference density differences of differing resolution. During the simulations, the electronic density is resolved on a $216 \times 216 \times 160$ mesh in a $20 \times 20 \times 15$ Å^3 supercell. Apart from the $27 \times 27 \times 10$ reference density difference grid, all simulations led to an isomerised structure. 205

List of Tables

5.1 HOMO-LUMO gap of a water molecule calculated using CAM-B3LYP and GTH pseudopotentials at increasing energy cutoffs and varying cubic simulation cell lengths l using periodic or isolated system boundary conditions for solving the Poisson equation. 94

5.2 HOMO-LUMO gap of a water molecule calculated using atom-centred basis sets augmented with a varying number of diffuse functions (aug. fcts). 97

5.3 Excitation energies of the first two to four excited singlet states obtained for various systems and basis sets using LR-TDDFT and the CAM-B3LYP xc functional within the Tamm-Dancoff approximation. Values for a customised CAM-O3LYP functional are indicated, too. Literature reference values from high-level wavefunction calculations are given where available. In *p*-nitroaniline, the first forbidden transition is not predicted in a plane-wave basis when using CAM-B3LYP; its oscillator strength is $f = 0$ in the other calculations. The character of the transition is indicated whenever it is not consistent within different methods. Reference values for naphthalene, DMABN and the dipeptide correspond to those adopted by Peach *et al.*[140] 99

5.4 Comparison of excitation energies between CAM-B3LYP and CAM-O3LYP for a structure of the retinal protonated Schiff base obtained from a molecular dynamics snapshot. CAM-B3LYP is known to give good agreement with respect to high-level wavefunction methods for this molecule.[166,180] 102

5.5 Comparison of excitation energies between CAM-B3LYP, CAM-O3LYP and reference values for some typical usage cases of range-separated functionals known to be problematic when treated with CAM-B3LYP. The character of the transition is given where it differs with respect to the reference. The reference values for the β-dipeptide correspond to those adopted by Peach *et al.*[140]; remaining values for molecules included therein and not explicitly mentioned so far are tabulated in the Supporting Information. 103

6.1 HF dimer binding energy calculated for the M05, M06, M08 and M11 family of Minnesota functionals in different basis sets. Results for aug-cc-pVTZ and a customised complete Gaussian basis (Limit) are reproduced from Ref. 204. Plane waves results were obtained at a wavefunction cutoff of 300 Ry. Atom-centred results are not Counterpoise corrected, in agreement with common procedures used in functional benchmarking and fitting. 123

6.2 Reaction enthalpies for the members of the HC7/11 database computed using the M05, M06, M06-2x, M06-HF and M06-L xc functional in plane waves (125 Ry, $\xi = 8$), an augmented Slater basis (ATZ2P) and the commonly used Gaussian aug-cc-pVTZ basis. Reference values correspond to 6-311+G(2df,2p), which is the recommended basis for benchmarking in Ref. 43. 128

6.3 Reaction enthalpies for the members of the HC7/11 database computed using the M11 and M11-L xc functional in plane waves (125 Ry, $\xi = 8$), an augmented Slater basis (ATZ2P), aug-cc-pVTZ and the reference basis (6-311+G(2df,2p)) 130

6.4 Values for $E_x^{(r_c)}$, $E_c^{(r_c)}$ and their sum, $E_{xc}^{(r_c)}$ around the nuclear core region, *i.e.* regions within the pseudopotential cutoff radius r_c, for reaction **2**. 134

7.1 Total energy, HOMO-LUMO gap $\Delta\epsilon_{\text{gap}}$ and singlet excitation energies for the organic dye DMABN obtained from the standard $v_{ij}(\mathbf{r})$ and its scaled counterpart presented in this work (eq. 7.15; $\lambda = 0.5$). Calculations were carried out in an orthorhombic $37.5 \times 18.75 \times 18.75$ Å^3 supercell at a plane wave cutoff of 75 Ry using the B3LYPxc functional. 155

7.2 Execution time for optimisation (ODIIS) and diagonalisation (Davidson) of the ground-state KS wavefunction, as well as the diagonalisation of the TDDFT matrix (LR) for both a monomer and a trimer of DMABN (28 and 84 occupied states, respectively). Values averaged over 10 iterations are given for both the full $v_{ij}(\mathbf{r})$ (Ref) and its scaled counterpart (λ) and have been used to calculate the speedup η. Calculations were carried out in an orthorhombic $25.0\times12.5\times12.5$ Å^3 supercell at a plane wave cutoff of 75 Ry using the B3LYPxc functional on 256 CPU cores. Optimal parallel configurations are compared. The values for the full $v_{ij}(\mathbf{r})$ were obtained by using 4 CP groups and a block size of 4 in the calculation of the exact exchange. The calculations using the scaled $v_{ij}(\mathbf{r})$ were carried out using 4 CP groups and a block size of 14. For details on the parallelisation, see Ref. 160. 155

8.1 Number of CPU (or MPI tasks) for a given size of the simulation supercell spanned by vectors of length l_x, l_y, l_z. All calculations were carried out on identical processors on 16-node racks. 171

8.2 Number of yz-planes, $x_{\text{up}} - x_{\text{low}} + 1$, held by ever MPI tasks, for a $75.0 \times 25.0 \times 12.5$ Å^3 supercell and different numbers of CP groups N_C at constant E^{ρ}_{cut} and $E^{\rho(\lambda)}_{\text{cut}}$. The mesh for E^{ρ}_{cut} comprises a total of $x = 384$, $y = 128$ and $z = 64$ gridpoints; its counterpart for $E^{\rho(\lambda)}_{\text{cut}}$ contains only half of the gridpoints per direction. Calculations were carried out on $N_P = 192$ cores. 174

8.3 Lowest average execution times t_{CPU} for a wavefunction optimisation (ODIIS) and the optimisation of the first four excited states using linear-response TDDFT with increasing numbers of orbitals N_b for both coordinate-scaled (Scaled) and full (Ref) hybrid functional calculations performed within a $75.0 \times 25.0 \times 12.5$ Å^3 supercell on 192 CPU cores using an optimal block size b and group distribution N_C, N_P. 179

9.1 Calibration systems for DCACP as used in the present work. Values for σ_1, σ_2 are given for the species calibrated in the present work, denoted in bold face. 188

9.2 Test systems for DCACP calibrated in the present work. The species calibrated in the present work are indicated in bold face. 189

9.3 Equilibrium distance $d_{||}$ and binding energy $\Delta E_{||}$ for the aurophilic complex $(HAuPH_3)_2$. DCACP-BLYP predicts a minimum close to the CCSD(T) reference.[259] The energetic error of DFT-D2,[210] however, considerably exceeds the range of chemical accuracy. 190

A.1 HOMO-LUMO gap of a water molecule calculated using CAM-B3LYP and MT pseudopotentials at increasing energy cutoffs and varying simulation cell lengths l representing an isolated and periodic system, respectively. 215
A.2 Size l of the orthorhombic periodic supercell for all systems considered in this study. For structures and abbreviations, see Ref. 140. 216
A.3 Excitation energies for the molecules contained in the test set by Peach *et al.*[140] that have not been tabulated in the main text. Calculations used a cutoff value for the KS orbitals of 75 Ry. For structures and abbreviations, see Ref. 140. 218
A.4 Total energy, HOMO-LUMO gap $\Delta\epsilon_{\mathrm{HOMO}}^{\mathrm{LUMO}}$ and singlet excitation energies for the organic dye DMABN obtained from the standard $v_{ij}(\mathbf{r})$ and its scaled counterpart presented in this work ($\lambda = 0.5$) on a geometry of DMABN distorted by $\Delta r = 0.05\,\text{Å}$ in the x and z direction. Calculations were carried out in an orthorhombic $37.5 \times 18.75 \times 18.75\ \text{Å}^3$ supercell at a plane wave cutoff of 75 Ry using the B3LYP[31,163] xc functional. 219

Symbols and Nomenclature

β	Thermodynamic beta, $(k_B T)^{-1}$
χ	Nuclear component in the expansion of the total wavefunction
ϵ	Energy of a single particle orbital
ϵ_c	Correlation energy density
ϵ_x	Exchange energy density
Γ	Phase space
$\hbar$	Planck constant
$\mathbf{G}$	Reciprocal space vector
$\mathbf{P}$	Nuclear momenta in continuous basis
$\mathbf{p}$	Electronic or general momenta in continuous basis
$\mathbf{q}$	Generalised coordinates in continuous basis
$\mathbf{R}$	Nuclear coordinates in continuous Cartesian basis *or* electronic coordinates discretised on a grid (context)
$\mathbf{r}$	Electronic or general coordinates in continuous Cartesian basis
$\mathcal{F}$	Fourier transformation ($\mathcal{F}^{-1}$ being its inverse)
$\mathcal{H}$	H is a classical function
$\mathcal{O}$	Order

$\mathcal{S}$	Fourier series representation or discrete Fourier transformation ($\mathcal{S}^{-1}$ being its inverse)
A	Helmholtz free energy
E_{xc}	Exchange-correlation energy, exchange-correlation functional
E	Energy
H	Hamiltonian
J	Classical coulomb energy
O	Generic operator
S	Entropy
T_{c}	Correlation contribution to kinetic energy
T_{s}	Noninteracting kinetic energy
T	Kinetic energy (full)
V	Potential energy
W	Two-electron operator (Coulomb or some appropriate generalisation)
Z	Partition function (Zustandsfunktion)
$\hat{\gamma}$	Density operator, density matrix
$\hat{\mathrm{H}}$	H is an operator
Φ	Generalised Coulomb operator *or* electronic wavefunction (context)
Ψ	Electronic *or* total wavefunction (context)
ψ_i	Single-particle orbital with index i
ρ	N-electron density
ρ_0	N-electron density associated to some ground state
det	Determinant
Tr	Trace
h_{xc}	Exchange-correlation hole

k_B	Boltzmann constant
v	Any potential (suitably suffixed); external potential (no suffix)
F	Density Functional
N	Number of particles
N_b	Number of (occupied) bands/orbitals
S	Singlet
V, Ω	Volume
`ngw`	Number of **G**-vectors in the expansion of the wavefunction
AC	Adiabatic Connection
BO	Born-Oppenheimer
CAM	Coulomb-Attenuation Method
CPU	Core Processing Unit
d3MD	Density Difference-Driven Molecular Dynamics
DF	Density Functional
DFT	Density Functional Theory
eff	Effective
FFT	Fast Fourier Transform
FPMD	*First principles* Molecular Dynamics
GGA	Generalised Gradient Approximation
GKS	Generalised Kohn-Sham, such as in GKS-DFT
GS	Ground State
HFX	Hartree-Fock (exact) exchange
HK	Hohenberg-Kohn
KS	Kohn-Sham, such as in KS-DFT

LC	Long range-Correction
LDA	Local Density Approximation
LR	Linear Response
lr	Long range
LSDA	Local Spin Density Approximation
MD	Molecular Dynamics
MPI	Message Passing Interface
PES	Potential Energy Surface
sc	Screened
sr	Short range
TDA	Tamm-Dancoff Approximation
TDDFT	Time-Dependent Density Functional Theory
UEG	Uniform Electron Gas
xc	exchange-correlation

Bibliography

[1] Brecht, B. *Die Dreigroschenoper: Nach John Gays 'The Beggar's Opera'*; Suhrkamp Verlag, 2013.

[2] Barks, K. *Die Geldquelle*; Aus dem Englischen von Erika Fuchs, Ed.; Egmont Ehapa: Berlin, 2010.

[3] Hall, G. G. *Reports on Progress in Physics* **1959**, *22*, 1.

[4] Szabo, A.; Ostlund, N. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Books on Chemistry Series; Dover Publications, 1996.

[5] Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. *Chemical Reviews* **2012**, *112*, 543–631.

[6] Roux, B.; Simonson, T. *Biophysical Chemistry* **1999**, *78*, 1–20.

[7] Kleinjung, J.; Fraternali, F. *Current Opinion in Structural Biology* **2014**, *25*, 126–134.

[8] Guerard, J. J.; Arey, J. S. *Journal of Chemical Theory and Computation* **2013**, *9*, 5046–5058.

[9] Piccini, G.; Alessio, M.; Sauer, J. *Angewandte Chemie International Edition* **2016**, *55*, 5235–5237.

[10] Kazemi, M.; Åqvist, J. *Nature Communications* **2015**, *6*, 7293.

[11] Gómez-Zaleta, B.; Gómez-Balderas, R.; Hernández-Trujillo, J. *Physical Chemistry Chemical Physics* **2010**, *12*, 4783–4790.

[12] Povie, G.; Suravarapu, S. R.; Bircher, Martin P.; Mojzes, M. M.; Rieder, S.; Renaud, P. *Science Advances* **2018**, *4*.

[13] Feynman, R. *Statistical Mechanics: A Set Of Lectures*; Advanced Books Classics; Avalon Publishing, 1998.

[14] Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864–B871.

[15] Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133–A1138.

[16] Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press, 2009.

[17] Parr, R.; Weitao, Y. *Density-Functional Theory of Atoms and Molecules*; International Series of Monographs on Chemistry; Oxford University Press, 1994.

[18] Bhoorasingh, P. L.; West, R. H. *Physical Chemistry Chemical Physics* **2015**, *17*, 32173–32182.

[19] Karplus, M.; McCammon, J. A. *Nature Structural Biology* **2002**, *9*, 646–652.

[20] Dror, R. O.; Dirks, R. M.; Grossman, J.; Xu, H.; Shaw, D. E. *Annual Review of Biophysics* **2012**, *41*, 429–452.

[21] De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. *Journal of Medicinal Chemistry* **2016**, *59*, 4035–4061.

[22] Bernardi, R. C.; Melo, M. C.; Schulten, K. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2015**, *1850*, 872–877.

[23] Jorgensen, W. L.; Tirado-Rives, J. *The Journal of Physical Chemistry* **1996**, *100*, 14508–14513.

[24] Morrone, J. A.; Car, R. *Physical Review Letters* **2008**, *101*, 017801.

[25] Antipov, S. V.; Bhattacharyya, S.; El Hage, K.; Xu, Z.-H.; Meuwly, M.; Rothlisberger, U.; Vaníček, J. *Structural Dynamics* **2017**, *4*, 061509.

[26] Bircher, Martin P.; Liberatore, E.; Browning, N. J.; Brickel, S.; Hofmann, C.; Patoz, A.; Unke, O. T.; Zimmermann, T.; Chergui, M.; Hamm, P.; Keller, U.; Meuwly, M.; Woerner, H.-J.; Vaníček, J.; Rothlisberger, U. *Structural Dynamics* **2017**, *4*, 061510.

[27] Curchod, B. F. E.; Martínez, T. J. *Chemical Reviews* **2018**, *118*, 3305–3336.

[28] Kohn, W. *Reviews of Modern Physics* **1999**, *71*, 1253–1266.

[29] Bartlett, R. J.; Musiał, M. *Reviews of Modern Physics* **2007**, *79*, 291–352.

[30] Liberatore, E.; Meli, R.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **2018**, *14*, 2834–2842.

[31] Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 1372–1377.

[32] Jacquemin, D.; Adamo, C. *Journal of Chemical Theory and Computation* **2011**, *7*, 369–376.

[33] Barducci, A.; Bonomi, M.; Parrinello, M. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 826–843.

[34] Dellago, C.; Bolhuis, P. G.; Geissler, P. L. In *Advances in Chemical Physics*; Wiley-Blackwell, 2003; Chapter 1, pp 1–78.

[35] Dellago, C.; Bolhuis, P. G.; Chandler, D. *The Journal of Chemical Physics* **1999**, *110*, 6617–6625.

[36] Borrero, E.; Dellago, C. *The European Physical Journal Special Topics* **2016**, *225*, 1609–1620.

[37] Bolhuis, P. G.; Csányi, G. *Physical Review Letters* **2018**, *120*, 250601.

[38] Yanai, T.; Tew, D. P.; Handy, N. C. *Chemical Physics Letters* **2004**, *393*, 51–57.

[39] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *The Journal of Chemical Physics* **2005**, *123*, 161103.

[40] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2006**, *2*, 364–382.

[41] Zhao, Y.; Truhlar, D. G. *The Journal of Chemical Physics* **2006**, *125*, 194101.

[42] Zhao, Y.; Truhlar, D. G. *The Journal of Physical Chemistry A* **2006**, *110*, 13126–13130.

[43] Zhao, Y.; Truhlar, D. G. *Theoretical Chemistry Accounts* **2008**, *120*, 215–241.

[44] Peverati, R.; Truhlar, D. G. *The Journal of Physical Chemistry Letters* **2011**, *2*, 2810–2817.

[45] Peverati, R.; Truhlar, D. G. *The Journal of Physical Chemistry Letters* **2012**, *3*, 117–124.

[46] Peverati, R.; Truhlar, D. G. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **2014**, *372*.

[47] von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *Physical Review Letters* **2004**, *93*, 153004–153007.

[48] Schrödinger, E. *Annalen der Physik* **1926**, *384*, 361–376.

[49] Boltzmann, L. *Vorlesungen über Gastheorie: Th. Theorie van der Waals'; Gase mit zusammengesetzten Molekülen; Gasdissociation; Schlussbemerkungen*; Vorlesungen über Gastheorie; J. A. Barth, 1898.

[50] Gibbs, J. W. *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*; Cambridge Library Collection - Mathematics; Cambridge University Press, 1902.

[51] Tuckerman, M. *Statistical Mechanics: Theory and Molecular Simulation*; Oxford graduate texts; Oxford University Press, 2011.

[52] Bellac, M.; Mortessagne, F.; Batrouni, G. *Equilibrium and Non-equilibrium Statistical Thermodynamics*; Cambridge University Press, 2014.

[53] von Helmholtz, H.; Hittorf, J.; Waals, J.; Puluj, J. *Physical Memoirs Selected and Translated from Foreign Sources*; Physical Memoirs Selected and Translated from Foreign Sources; Taylor & Francis, 1888.

[54] Cohen-Tannoudji, C.; Diu, B.; Laloë, F. *Quantum mechanics*; Quantum Mechanics; Wiley, 1977.

[55] Bohm, A. *Quantum Mechanics: Foundations and Applications*; Theoretical and Mathematical Physics; Springer New York, 1994.

[56] Neumann, J. v. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1927**, *1927*, 245–272.

[57] Alder, B. J.; Wainwright, T. E. *The Journal of Chemical Physics* **1959**, *31*, 459–466.

[58] Born, M.; Oppenheimer, R. *Annalen der Physik* **1927**, *389*, 457–484.

[59] Born, M.; Huang, K. *Dynamical Theory of Crystal Lattices*; International series of monographs on physics; Clarendon Press, 1954.

[60] Bohm, D. *Physical Review* **1952**, *85*, 166–179.

[61] Curchod, B. F. E.; Rothlisberger, U.; Tavernelli, I. *ChemPhysChem* **2013**, *14*, 1314–1340.

[62] Feynman, R.; Hibbs, A.; Styer, D. *Quantum Mechanics and Path Integrals*; Dover Books on Physics; Dover Publications, 2010.

[63] Domcke, W.; Yarkony, D.; Köppel, H. *Conical Intersections: Theory, Computation and Experiment*; Advanced series in physical chemistry; World Scientific Publishing Company, 2011.

[64] C. Tully, J. *Faraday Discuss.* **1998**, *110*, 407–419.

[65] Tavernelli, I.; Röhrig, U. F.; Rothlisberger, U. *Molecular Physics* **2005**, *103*, 963–981.

[66] Tully, J. C.; Preston, R. K. *The Journal of Chemical Physics* **1971**, *55*, 562–572.

[67] Zener, C. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **1932**, *137*, 696–702.

[68] Landau, L. *Physikalische Zeitschrift der Sowjetunion* **1932**, *2*, 46–51.

[69] Tully, J. C. *The Journal of Chemical Physics* **1990**, *93*, 1061–1071.

[70] Tapavicza, E.; Tavernelli, I.; Rothlisberger, U. *Physical Review Letters* **2007**, *98*, 023001.

[71] Tavernelli, I.; Tapavicza, E.; Rothlisberger, U. *The Journal of Chemical Physics* **2009**, *130*, 124107.

[72] Brunk, E.; Rothlisberger, U. *Chemical Reviews* **2015**, *115*, 6217–6263.

[73] Malhado, J. a. P.; Bearpark, M. J.; Hynes, J. T. *Frontiers in Chemistry* **2014**, *2*, 97.

[74] Levy, M. *Physical Review A* **1982**, *26*, 1200–1208.

[75] Lieb, E. H. *International Journal of Quantum Chemistry* **1983**, *24*, 243–277.

[76] Englisch, H.; Englisch, R. *Physica A: Statistical Mechanics and its Applications* **1983**, *121*, 253–268.

[77] Eschrig, H. *The Fundamentals of Density Functional Theory*; Teubner-Texte zur Physik; Springer Fachmedien: Wiesbaden, 1996.

[78] Kohn, W. **2018**; as quoted by Kieron Burke during a lecture at EPFL.

[79] Runge, E.; Gross, E. K. U. *Physical Review Letters* **1984**, *52*, 997–1000.

[80] Cohen, M. H.; Wasserman, A. *Physical Review A* **2005**, *71*, 032515.

[81] Thomas, L. H. *Mathematical Proceedings of the Cambridge Philosophical Society* **1927**, *23*, 542–548.

[82] Dirac, P. A. M. *Mathematical Proceedings of the Cambridge Philosophical Society* **1930**, *26*, 376–385.

[83] Mardirossian, N.; Head-Gordon, M. *Molecular Physics* **2017**, *115*, 2315–2372.

[84] Stowasser, R.; Hoffmann, R. *Journal of the American Chemical Society* **1999**, *121*, 3414–3420.

[85] van Meer, R.; Gritsenko, O. V.; Baerends, E. J. *Journal of Chemical Theory and Computation* **2014**, *10*, 4432–4441.

[86] Elliott, P.; Lee, D.; Cangi, A.; Burke, K. *Physical Review Letters* **2008**, *100*, 256406.

[87] Burke, K. *The Journal of Chemical Physics* **2012**, *136*, 150901.

[88] van Leeuwen, R. *Physical Review Letters* **1999**, *82*, 3863–3866.

[89] Janak, J. F. *Physical Review B* **1978**, *18*, 7165–7168.

[90] Harris, J.; Jones, R. O. *Journal of Physics F: Metal Physics* **1974**, *4*, 1170.

[91] Teale, A. M.; Coriani, S.; Helgaker, T. *The Journal of Chemical Physics* **2010**, *132*, 164115.

[92] Harris, J. *Physical Review A* **1984**, *29*, 1648–1659.

[93] Becke, A. D. *The Journal of Chemical Physics* **1988**, *88*, 1053–1062.

[94] Burke, K.; Ernzerhof, M.; Perdew, J. P. *Chemical Physics Letters* **1997**, *265*, 115–120.

[95] Teale, A. M.; Coriani, S.; Helgaker, T. *The Journal of Chemical Physics* **2009**, *130*, –.

[96] Tucholsky, K. *Die Weltbühne* **1931**, *11/389*; published under the name of Kaspar Hauser.

[97] Coleman, A. J. In *Many-Electron Densities and Reduced Density Matrices*; Cioslowski, J., Ed.; Springer US: Boston, MA, 2000; pp 1–17.

[98] Goedecker, S.; Umrigar, C. J. In *Many-Electron Densities and Reduced Density Matrices*; Cioslowski, J., Ed.; Springer US: Boston, MA, 2000; pp 165–181.

[99] Burke, K. In *Electronic Density Functional Theory: Recent Progress and New Directions*; Dobson, J. F.; Vignale, G.; Das, M. P., Eds.; Springer US: Boston, MA, 1998; pp 19–29.

[100] Maitra, N. T.; Burke, K. In *Many-Electron Densities and Reduced Density Matrices*; Cioslowski, J., Ed.; Springer US: Boston, MA, 2000; pp 183–208.

[101] Koga, T. In *Many-Electron Densities and Reduced Density Matrices*; Cioslowski, J., Ed.; Springer US: Boston, MA, 2000; pp 267–298.

[102] Perdew, J. P.; Kurth, S. In *A Primer in Density Functional Theory*; Fiolhais, C.; Nogueira, F.; Marques, M. A. L., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2003; pp 1–55.

[103] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1996**, *77*, 3865–3868.

[104] Perdew, J. P.; Ernzerhof, M.; Burke, K. *The Journal of Chemical Physics* **1996**, *105*, 9982–9985.

[105] Burke, K.; Perdew, J. P.; Ernzerhof, M. *The Journal of Chemical Physics* **1998**, *109*, 3760–3771.

[106] Sun, J.; Ruzsinszky, A.; Perdew, J. P. *Physical Review Letters* **2015**, *115*, 036402.

[107] Dyson, F. *Nature* **2004**, *427*, 297.

[108] Jones, R. O. *Rev. Mod. Phys.* **2015**, *87*, 897–923.

[109] Sham, L. J. *Physical Review A* **1970**, *1*, 969–970.

[110] Levy, M.; Perdew, J. P. *Physical Review A* **1985**, *32*, 2010–2021.

[111] Liu, S.; Parr, R. G. *Physical Review A* **1996**, *53*, 2211–2219.

[112] Calderín, L. *Physical Review A* **2012**, *86*, 032510.

[113] Borgoo, A.; Teale, A. M.; Tozer, D. J. *The Journal of Chemical Physics* **2012**, *136*, 034101.

[114] Perdew, J. P.; Ruzsinszky, A.; Sun, J.; Burke, K. *The Journal of Chemical Physics* **2014**, *140*, 18A533.

[115] Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. *Science* **2017**, *355*, 49–52.

[116] Slater, J. C. *Physical Review* **1951**, *81*, 385–390.

[117] Slater, J. C.; Johnson, K. H. *Phys. Rev. B* **1972**, *5*, 844–853.

[118] Vosko, S. H.; Wilk, L.; Nusair, M. *Canadian Journal of Physics* **1980**, *58*, 1200–1211.

[119] Perdew, J. P.; Zunger, A. *Physical Review B* **1981**, *23*, 5048–5079.

[120] Perdew, J. P.; Wang, Y. *Physical Review B* **1992**, *45*, 13244–13249.

[121] Becke, A. D. *The Journal of Chemical Physics* **2014**, *140*, 18A301.

[122] Becke, A. D. *Physical Review A* **1988**, *38*, 3098–3100.

[123] Elliott, P.; Burke, K. *Canadian Journal of Chemistry* **2009**, *87*, 1485–1491.

[124] Langreth, D. C.; Perdew, J. P. *Physical Review B* **1980**, *21*, 5469–5493.

[125] Langreth, D. C.; Mehl, M. J. *Physical Review B* **1983**, *28*, 1809–1834.

[126] Perdew, J. P.; Yue, W. *Physical Review B* **1986**, *33*, 8800–8802.

[127] Perdew, J. P. *Physical Review Letters* **1985**, *55*, 1665–1668.

[128] Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

[129] Becke, A. D. *The Journal of Chemical Physics* **1998**, *109*, 2092–2098.

[130] Perdew, J. P.; Kurth, S.; Zupan, A. c. v.; Blaha, P. *Physical Review Letters* **1999**, *82*, 2544–2547.

[131] Kurth, S.; Perdew, J. P.; Blaha, P. *International Journal of Quantum Chemistry* **1999**, *75*, 889–909.

[132] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Physical Review Letters* **2003**, *91*, 146401.

[133] Staroverov, V. N. In *A Matter of Density*; Sukumar, N., Ed.; John Wiley & Sons, Inc.: New York, 2012; pp 125–126.

[134] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Physical Review Letters* **2003**, *91*, 146401.

[135] Gunnarsson, O.; Lundqvist, B. I. *Physical Review B* **1976**, *13*, 4274–4298.

[136] Langreth, D. C.; Perdew, J. P. *Physical Review B* **1977**, *15*, 2884–2901.

[137] Seidl, A.; Görling, A.; Vogl, P.; Majewski, J. A.; Levy, M. *Physical Review B* **1996**, *53*, 3764–3774.

[138] Städele, M.; Moukara, M.; Majewski, J. A.; Vogl, P.; Görling, A. *Physical Review B* **1999**, *59*, 10031–10043.

[139] Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andres, L. *Molecular Physics* **1999**, *97*, 859–868.

[140] Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *The Journal of Chemical Physics* **2008**, *128*, 044118.

[141] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *The Journal of Chemical Physics* **2001**, *115*, 3540–3544.

[142] Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *The Journal of Chemical Physics* **2003**, *118*, 8207–8215.

[143] Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *The Journal of Chemical Physics* **2004**, *120*, 8425–8433.

[144] Peach, M. J. G.; Cohen, A. J.; Tozer, D. J. *Physical Chemistry Chemical Physics* **2006**, *8*, 4543–4549.

[145] Görling, A.; Levy, M. *The Journal of Chemical Physics* **1997**, *106*, 2675–2680.

[146] Ivanov, S.; Levy, M. In *Density Functional Theory*; Löwdin, P.-O., Ed.; Advances in Quantum Chemistry, Vol. 33; Academic Press, 1998; pp 11 – 29.

[147] Nagy, B.; Jensen, F. In *Reviews in Computational Chemistry*; Wiley-Blackwell, 2017; Chapter 3, pp 93–149.

[148] Pulay, P. *Molecular Physics* **1969**, *17*, 197–204.

[149] Boys, S.; Bernardi, F. *Molecular Physics* **1970**, *19*, 553–566.

[150] Ditchfield, R.; Hehre, W. J.; Pople, J. A. *The Journal of Chemical Physics* **1971**, *54*, 724–728.

[151] Dunning, T. H. *The Journal of Chemical Physics* **1989**, *90*, 1007–1023.

[152] Bachelet, G. B.; Hamann, D. R.; Schlüter, M. *Physical Review B* **1982**, *26*, 4199–4228.

[153] Troullier, N.; Martins, J. L. *Physical Review B* **1991**, *43*, 1993–2006.

[154] Goedecker, S.; Teter, M.; Hutter, J. *Physical Review B* **1996**, *54*, 1703–1710.

[155] Martyna, G. J.; Tuckerman, M. E. *The Journal of Chemical Physics* **1999**, *110*, 2810–2821.

[156] Hockney, R. *Methods of Computational Physics* **1970**, *9*, 135–211.

[157] Chawla, S.; Voth, G. A. *The Journal of Chemical Physics* **1998**, *108*, 4697–4700.

[158] Gygi, F.; Baldereschi, A. *Physical Review B* **1986**, *34*, 4405–4408.

[159] Broqvist, P.; Alkauskas, A.; Pasquarello, A. *Physical Review B* **2009**, *80*, 085114.

[160] Weber, V.; Bekas, C.; Laino, T.; Curioni, A.; Bertsch, A.; Futral, S. In *2014 IEEE 28th International Parallel and Distributed Processing Symposium*; pp 735–744.

[161] Bircher, Martin P.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **2018**, *14*, 3184–3195.

[162] Casida, M.; Ipatov, A.; Cordova, F. In *Time-Dependent Density Functional Theory*; Marques, M. A.; Ullrich, C. A.; Nogueira, F.; Rubio, A.; Burke, K.; Gross, E. K. U., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2006; pp 243–257.

[163] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *The Journal of Physical Chemistry* **1994**, *98*, 11623–11627.

[164] Jensen, L.; Govind, N. *The Journal of Physical Chemistry A* **2009**, *113*, 9761–9765.

[165] Arulmozhiraja, S.; Coote, M. L. *Journal of Chemical Theory and Computation* **2012**, *8*, 575–584.

[166] Valsson, O.; Campomanes, P.; Tavernelli, I.; Rothlisberger, U.; Filippi, C. *Journal of Chemical Theory and Computation* **2013**, *9*, 2441–2454.

[167] Laktionov, A.; Rothlisberger, U. **2014**; unpublished.

[168] Adamo, C.; Barone, V. *The Journal of Chemical Physics* **1999**, *110*, 6158–6170.

[169] Rohrdanz, M. A.; Herbert, J. M. *The Journal of Chemical Physics* **2008**, *129*, 034107.

[170] Riffet, V.; Jacquemin, D.; Cauët, E.; Frison, G. *Journal of Chemical Theory and Computation* **2014**, *10*, 3308–3318.

[171] Eriksen, J. J.; Sauer, S. P.; Mikkelsen, K. V.; Christiansen, O.; Jensen, H. J. A.; Kongsted, J. *Molecular Physics* **2013**, *111*, 1235–1248.

[172] Sevilla, M. D. *The Journal of Physical Chemistry A* **2009**, *113*, 11093–11094.

[173] Serrano-Andrés, L.; Fülscher, M. P. *Journal of the American Chemical Society* **1998**, *120*, 10912–10920.

[174] Odelius, M.; Laikov, D.; Hutter, J. *Journal of Molecular Structure: THEOCHEM* **2003**, *630*, 163–175; WATOC '02 Special Issue.

[175] Kronik, L.; Stein, T.; Refaely-Abramson, S.; Baer, R. *Journal of Chemical Theory and Computation* **2012**, *8*, 1515–1531.

[176] www.cpmd.org; *CPMD*; Copyright IBM Corp 1990-2008, Copyright MPI für Festkörperforschung Stuttgart 1997-2001.

[177] Xu, X.; Goddard, W. A. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 2673–2677.

[178] Perdew, J. P. *Electronic structure of solids*; Akademie Verlag, Berlin, 1991; Vol. 11.

[179] Rostov, I. V.; Amos, R. D.; Kobayashi, R.; Scalmani, G.; Frisch, M. J. *The Journal of Physical Chemistry B* **2010**, *114*, 5547–5555.

[180] Sneskov, K.; Olsen, J. M. H.; Schwabe, T.; Hattig, C.; Christiansen, O.; Kongsted, J. *Physical Chemistry Chemical Physics* **2013**, *15*, 7567–7576.

[181] Palczewska, G.; Vinberg, F.; Stremplewski, P.; Bircher, Martin P.; Salom, D.; Komar, K.; Zhang, J.; Cascella, M.; Wojtkowski, M.; Kefalov, V. J.; Palczewski, K. *Proceedings of the National Academy of Sciences* **2014**, *111*, E5445–E5454.

[182] Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Jensen, H. J. Aa.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawlowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Sałek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. *WIREs Computational Molecular Science* **2014**, *4*, 269–284.

[183] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov,

A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ã.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; *Gaussian09 Revision D.01*; Gaussian Inc. Wallingford CT 2009.

[184] Van Lenthe, E.; Baerends, E. J. *Journal of Computational Chemistry* **2003**, *24*, 1142–1156.

[185] te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931–967.

[186] Baerends, E. J.; Ziegler, T.; Atkins, A. J.; Autschbach, J.; Bashford, D.; Baseggio, O.; Bérces, A.; Bickelhaupt, F. M.; Bo, C.; Boerritger, P. M.; Cavallo, L.; Daul, C.; Chong, D. P.; Chulhai, D. V.; Deng, L.; Dickson, R. M.; Dieterich, J. M.; Ellis, D. E.; van Faassen, M.; Ghysels, A.; Giammona, A.; van Gisbergen, S. J. A.; Goez, A.; Götz, A. W.; Gusarov, S.; Harris, F. E.; van den Hoek, P.; Hu, Z.; Jacob, C. R.; Jacobsen, H.; Jensen, L.; Joubert, L.; Kaminski, J. W.; van Kessel, G.; König, C.; Kootstra, F.; Kovalenko, A.; Krykunov, M.; van Lenthe, E.; McCormack, D. A.; Michalak, A.; Mitoraj, M.; Morton, S. M.; Neugebauer, J.; Nicu, V. P.; Noodleman, L.; Osinga, V. P.; Patchkovskii, S.; Pavanello, M.; Peeples, C. A.; Philipsen, P. H. T.; Post, D.; Pye, C. C.; Ramanantoanina, H.; Ramos, P.; Ravenek, W.; Rodríguez, J. I.; Ros, P.; Rüger, R.; Schipper, P. R. T.; Schlüns, D.; van Schoot, H.; Schreckenbach, G.; Seldenthuis, J. S.; Seth, M.; Snijders, J. G.; Solà, M.; M., S.; Swart, M.; Swerhone, D.; te Velde, G.; Tognetti, V.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T. A.; van Wezenbeek, E. M.; Wiesenekker, G.; Wolff, S. K.; Woo, T. K.; Yakovlev, A. L.; *ADF2017, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, https://www.scm.com.*

[187] Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. *Theoretical Chemistry Accounts* **1998**, *99*, 391–403.

[188] Hirata, S.; Head-Gordon, M. *Chemical Physics Letters* **1999**, *314*, 291–299.

[189] Bekas, C.; Curioni, A.; Andreoni, W. *Parallel Computing* **2008**, *34*, 441–450.

[190] Bekas, C.; Curioni, A. *Computer Physics Communications* **2010**, *181*, 1057–1068.

[191] Peach, M. J. G.; Tozer, D. J. *The Journal of Physical Chemistry A* **2012**, *116*, 9783–9789.

[192] Grimme, S.; Parac, M. *ChemPhysChem* **2003**, *4*, 292–295.

[193] Bulliard, C.; Allan, M.; Wirtz, G.; Haselbach, E.; Zachariasse, K. A.; Detzer, N.; Grimme, S. *The Journal of Physical Chemistry A* **1999**, *103*, 7766–7772.

[194] Svartsov, Y. N.; Schmitt, M. *The Journal of Chemical Physics* **2008**, *128*, 214310.

[195] Kosenkov, D.; Slipchenko, L. V. *The Journal of Physical Chemistry A* **2011**, *115*, 392–401.

[196] Handy, N. C.; Cohen, A. J. *Molecular Physics* **2001**, *99*, 403–412.

[197] Lee, C.; Yang, W.; Parr, R. G. *Physical Review B* **1988**, *37*, 785–789.

[198] Hoe, W.-M.; Cohen, A. J.; Handy, N. C. *Chemical Physics Letters* **2001**, *341*, 319–328.

[199] Baker, J.; Pulay, P. *The Journal of Chemical Physics* **2002**, *117*, 1441–1449.

[200] Baker, J.; Pulay, P. *Journal of Computational Chemistry* **2003**, *24*, 1184–1191.

[201] Xu, X.; Goddard, W. A. *The Journal of Physical Chemistry A* **2004**, *108*, 8495–8504.

[202] Lange, A. W.; Herbert, J. M. *Journal of the American Chemical Society* **2009**, *131*, 3913–3922.

[203] Bircher, Martin P.; Lopez-Tarífa, P.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **submitted**.

[204] Mardirossian, N.; Head-Gordon, M. *Journal of Chemical Theory and Computation* **2013**, *9*, 4453–4461.

[205] Martell, J. M.; Goddard, J. D.; Eriksson, L. A. *The Journal of Physical Chemistry A* **1997**, *101*, 1927–1934.

[206] Zhao, Y.; Truhlar, D. G. *Accounts of Chemical Research* **2008**, *41*, 157–167.

[207] Zhao, Y.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2008**, *4*, 1849–1868.

[208] Zhao, Y.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2005**, *1*, 415–432.

[209] Zhao, Y.; Truhlar, D. G. *Organic Letters* **2006**, *8*, 5753–5755.

[210] Grimme, S. *The Journal of Chemical Physics* **2006**, *124*, 034108.

[211] Keal, T. W.; Tozer, D. J. *The Journal of Chemical Physics* **2005**, *123*, 121103.

[212] Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Dam, H. V.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.; de Jong, W. *Computer Physics Communications* **2010**, *181*, 1477–1489.

[213] Gonze, X.; Jollet, F.; Araujo, F. A.; Adams, D.; Amadon, B.; Applencourt, T.; Audouze, C.; Beuken, J.-M.; Bieder, J.; Bokhanchuk, A.; Bousquet, E.; Bruneval, F.; Caliste, D.; Côté, M.; Dahm, F.; Pieve, F. D.; Delaveau, M.; Gennaro, M. D.; Dorado, B.; Espejo, C.; Geneste, G.; Genovese, L.; Gerossier, A.; Giantomassi, M.; Gillet, Y.; Hamann, D.; He, L.; Jomard, G.; Janssen, J. L.; Roux, S. L.; Levitt, A.; Lherbier, A.; Liu, F.; Lukačević, I.; Martin, A.; Martins, C.; Oliveira, M.; Poncć, S.; Pouillon, Y.; Rangel, T.; Rignanese, G.-M.; Romero, A.; Rousseau, B.; Rubel, O.; Shukri, A.; Stankovski, M.; Torrent, M.; Setten, M. V.; Troeye, B. V.; Verstraete, M.; Waroquiers, D.; Wiktor, J.; Xu, B.; Zhou, A.; Zwanziger, J. *Computer Physics Communications* **2016**, *205*, 106–131.

[214] Van Voorhis, T.; Scuseria, G. E. *The Journal of Chemical Physics* **1998**, *109*, 400–410.

[215] Marques, M. A.; Oliveira, M. J.; Burnus, T. *Computer Physics Communications* **2012**, *183*, 2272–2281.

[216] Becke, A. D. *The Journal of Chemical Physics* **1988**, *88*, 2547–2553.

[217] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J.; *Gaussian16 Revision B.01*; 2016; Gaussian Inc. Wallingford CT.

[218] Gräfenstein, J.; Izotov, D.; Cremer, D. *The Journal of Chemical Physics* **2007**, *127*, 214103.

[219] Dasgupta, S.; Herbert, J. M. *Journal of Computational Chemistry* **2017**, *38*, 869–882.

[220] Bircher, Martin P.; Rothlisberger, U. *The Journal of Physical Chemistry Letters* **2018**, *9*, 3886–3890.

[221] Levy, M. In *Density Functional Theory of Many-Fermion Systems*; Löwdin, P.-O., Ed.; Advances in Quantum Chemistry, Vol. 21; Academic Press, 1990; pp 69–95.

[222] Levy, M. In *Density Functional Theory*; Gross, E. K. U.; Dreizler, R. M., Eds.; Springer US: Boston, MA, 1995; pp 11–31.

[223] Levy, M.; Ou-Yang, H. *Physical Review A* **1990**, *42*, 651–652.

[224] Borgoo, A.; Teale, A. M.; Tozer, D. J. *Physical Chemistry Chemical Physics* **2014**, *16*, 14578–14583.

[225] Gledhill, J. D.; Tozer, D. J. *The Journal of Chemical Physics* **2015**, *143*, 024104.

[226] Sharpe, D. J.; Levy, M.; Tozer, D. J. *Journal of Chemical Theory and Computation* **2018**, *14*, 684–692.

[227] Csonka, G. I.; Perdew, J. P.; Ruzsinszky, A. *Journal of Chemical Theory and Computation* **2010**, *6*, 3688–3703.

[228] Wu, X.; Selloni, A.; Car, R. *Physical Review B* **2009**, *79*, 085102.

[229] Hutter, J.; Lüthi, H. P.; Parrinello, M. *Computational Materials Science* **1994**, *2*, 244–248.

[230] Bircher, Martin P.; Rothlisberger, U. *Computer Physics Communications* **submitted**.

[231] van Gunsteren, W. F.; Berendsen, H. J. C. *Angewandte Chemie International Edition in English* **1990**, *29*, 992–1023.

[232] Car, R.; Parrinello, M. *Physical Review Letters* **1985**, *55*, 2471–2474.

[233] Hutter, J.; Curioni, A. *ChemPhysChem* **2005**, *6*, 1788–1793.

[234] Laio, A.; VandeVondele, J.; Rothlisberger, U. *The Journal of Chemical Physics* **2002**, *116*, 6941–6947.

[235] von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *The Journal of Chemical Physics* **2005**, *122*, 014113.

[236] Tavernelli, I.; Lin, I.-C.; Rothlisberger, U. *Physical Review B* **2009**, *79*, 045106–045110.

[237] Perdew, J. P. *Physical Review B* **1986**, *33*, 8822–8824.

[238] Lin, I.-C.; Coutinho-Neto, M.; Felsenheimer, C.; von Lilienfeld, O.; Tavernelli, I.; Rothlisberger, U. *Physical Review B* **2007**, *75*, 205131.

[239] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *The Journal of Chemical Physics* **2010**, *132*, 154104.

[240] Pyykkö, P. *Physical Chemistry Chemical Physics* **2011**, *13*, 161–168.

[241] Tapavicza, E.; Lin, I.-C.; von Lilienfeld, O. A.; Tavernelli, I.; Coutinho-Neto, M. D.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **2007**, *3*, 1673–1679.

[242] Lin, I.-C.; Rothlisberger, U. *Physical Chemistry Chemical Physics* **2008**, *10*, 2730–2734.

[243] Lin, I.-C.; von Lilienfeld, O. A.; Coutinho-Neto, M. D.; Tavernelli, I.; Rothlisberger, U. *The Journal of Physical Chemistry B* **2007**, *111*, 14346–14354.

[244] Doemer, M.; Tavernelli, I.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **2013**, *9*, 955–964.

[245] Aeberhard, P. C.; Arey, J. S.; Lin, I.-C.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **2009**, *5*, 23–28.

[246] Cascella, M.; Lin, I.-C.; Tavernelli, I.; Rothlisberger, U. *Journal of Chemical Theory and Computation* **2009**, *5*, 2930–2934.

[247] Shen, Y.; BelBruno, J. J. *The Journal of Physical Chemistry A* **2005**, *109*, 10077–10083.

[248] Jouvet, C.; Lardeux-Dedonder, C.; Martrenchard, S.; Solgadi, D. *The Journal of Chemical Physics* **1991**, *94*, 1759–1764.

[249] Zhao, Y.; Truhlar, D. G. *The Journal of Physical Chemistry A* **2006**, *110*, 5121–5129.

[250] Czuchaj, E.; Stoll, H. *Chemical Physics* **1999**, *248*, 1–16.

[251] Bettinger, H. F.; Kar, T.; Sánchez-García, E. *The Journal of Physical Chemistry A* **2009**, *113*, 3353–3359.

[252] Callender, C. L.; Mitchell, S. A.; Hackett, P. A. *The Journal of Chemical Physics* **1989**, *90*, 5252–5261.

[253] Park, S. J.; Kim, M. C.; Lee, Y. S.; Jeung, G.-H. *The Journal of Chemical Physics* **1997**, *107*, 2481–2487.

[254] Callender, C. L.; Mitchell, S. A.; Hackett, P. A. *The Journal of Chemical Physics* **1989**, *90*, 2535–2543.

[255] Brndiar, J.; Štich, I. *Journal of Chemical Theory and Computation* **2012**, *8*, 2301–2309.

[256] Cargnoni, F.; Kuś, T.; Mella, M.; Bartlett, R. J. *The Journal of Chemical Physics* **2008**, *129*, 204307.

[257] Partridge, H.; Stallcop, J. R.; Levin, E. *The Journal of Chemical Physics* **2001**, *115*, 6471–6488.

[258] Dunning, T. H.; Valley, M.; Taylor, H. S. *The Journal of Chemical Physics* **1978**, *69*, 2672–2681.

[259] Koppen, J. V.; Hapka, M.; Modrzejewski, M.; Szczcesniak, M. M.; Chalasinski, G. *The Journal of Chemical Physics* **2014**, *140*, 244313.

[260] Syomin, D.; Kim, J.; Koel, B. E.; Ellison, G. B. *The Journal of Physical Chemistry B* **2001**, *105*, 8387–8394.

[261] Pyykkö, P. *Angewandte Chemie International Edition* **2004**, *43*, 4412–4456.

[262] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *The Journal of Chemical Physics* **2010**, *132*, 154104.

[263] Schiavo, E.; Muñoz García, A. B.; Barone, V.; Vittadini, A.; Casarin, M.; Forrer, D.; Pavone, M. *Chemical Physics Letters* **2018**, *693*, 28 – 33.

[264] Ilawe, N. V.; Zimmerman, J. A.; Wong, B. M. *Journal of Chemical Theory and Computation* **2015**, *11*, 5426–5435.

[265] van Keulen, S. C.; Liberatore, E.; Bircher, Martin P.; Bhattacharyya, S.; Rothlisberger, U. **2018**; in perparation.

[266] Smith, M.; March, J. *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*; MARCH'S ADVANCED ORGANIC CHEMISTRY; Wiley, 2001.

[267] Jarzynski, C. *Physical Review C* **1997**, *56*, 5018–5035.

[268] Crooks, G. E. *Physical Review C* **2000**, *61*, 2361–2366.

[269] Dellago, C.; Hummer, G. *Entropy* **2014**, *16*, 41–61.

[270] Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annual Review of Physical Chemistry* **2002**, *53*, 291–318.

[271] Wesolowski, T. A.; Wang, Y. A. *Recent Progress in Orbital-free Density Functional Theory*; WORLD SCIENTIFIC, 2013.

[272] Borgoo, A.; Green, J. A.; Tozer, D. J. *Journal of Chemical Theory and Computation* **2014**, *10*, 5338–5345.

[273] Goldman, B. B.; Walters, W. P.; Spellmeyer, D. C., Ed.; Annual Reports in Computational Chemistry, Vol. 2; Elsevier, 2006; pp 127–140.

[274] Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. *Journal of Chemical Theory and Computation* **2015**, *11*, 2087–2096.

[275] Hansen, K.; Biegler, F.; Ramakrishnan, R.; Pronobis, W.; von Lilienfeld, O. A.; Müller, K.-R.; Tkatchenko, A. *The Journal of Physical Chemistry Letters* **2015**, *6*, 2326–2331.

[276] Brockherde, F.; Vogt, L.; Li, L.; Tuckerman, M. E.; Burke, K.; Müller, K.-R. *Nature Communications* **2017**, *8*, 872.

[277] Kandala, A.; Mezzacapo, A.; Temme, K.; Takita, M.; Brink, M.; Chow, J. M.; Gambetta, J. M. *Nature* **2017**, *549*, 242.

[278] Colless, J. I.; Ramasesh, V. V.; Dahlen, D.; Blok, M. S.; Kimchi-Schwartz, M. E.; McClean, J. R.; Carter, J.; de Jong, W. A.; Siddiqi, I. *Physical Review X* **2018**, *8*, 011021.

[279] Brecht, B. *Der gute Mensch von Sezuan: Ein Parabelstück*; Suhrkamp Verlag, 2013.

Martin P. Bircher

*08.07.1990

Computational Chemist

Contact

Martin P. Bircher
EPFL SB ISIC LCBC
Batochime UNIL
Av. F.-A. Forel 2
CH-1015 Lausanne
Switzerland

bircher.ma@gmail.com
ORCID 0000-0002-6905-3130
Google Scholar

Languages

German	native
English	C2
French	C1/C2
Serbo-Croatian	B1/B2

Programming

Object-Oriented
Fortran2003/08
C++ (basic)

Procedural
Fortran90/95
FORTRAN77
C (basic)

Scripting

bash
Python

Development

CPMD (plane-wave MD)

Programs

CPMD
CP2K
DALTON
Gaussian
Amber
ADF
GAMESS US
VMD

Education

from 02.14 **Ph.D. Candidate (Computational Chemistry)**

Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
with **Prof. U. Röthlisberger**

Thesis: *Improving Performance and Accuracy of Hybrid-Functional Based Molecular Dynamics in Plane Waves*

Development and implementation of new techniques that accelerate the calculation of exact exchange in plane waves; contribution of a new, object-oriented (Fortran2008) exchange-correlation driver to the CPMD code, along with the implementation of several xc functionals and features (CAM family, Minnesota functionals). Principal developer of a new scheme for structural elucidation based on experimental density difference maps, d3MD, combining experimental data with DFT-based *first principles* MD; main author of `blobby`, a code to convert crystallographic data to arbitrary cube files (Fortran2008/C).

07.12-01.14 **MSc (Chemistry and Molecular Sciences)**

summa cum laude: 6.0/6.0 Universität Bern, Berne, Switzerland
with **Prof. M. Cascella**

Thesis: *Molecular Mechanisms of Reduction of an α-Ester Radical by Catechol Investigated by Molecular Dynamics Simulations*

Application of enhanced sampling techniques in *first principles* MD to elucidate the reaction mechanism of catechol with a radical. Elucidation of the two-photon adsorption of rhodopsin using LR-TDDFT, and calculation of spectra of organic dyes in collaboration with experimental groups.

09.09-07.12 **BSc (Chemistry and Molecular Sciences)**

insigni cum laude: 5.5/6.0 Universität Bern, Berne, Switzerland
with **Prof. P. Renaud**

Thesis: *Triethylborane Mediated Reductive Radical Coupling of α-Substituted α-Iodoesters with Mono- and Disubstituted Olefins*

Experimental work on the hydroalkylation of unactivated alkenes using catechol derivatives.

08.15-07.09 **Matura**

Final grade: 5.9/6.0 Gymnasium Interlaken/Gstaad

Equivalent of a secondary school (UK) or high school degree (US).

Teaching Activities

09.15-12.15
09.14-12.14 **Introduction to Electronic Structure Methods**

Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Course taught by **Prof. U. Röthlisberger**

Design, typesetting and preparation of a complete set of exercises for a 3rd year Bachelor's course in computational chemistry, covering the basics of both wavefunction methods and DFT. Exercises involved a repetition of the theoretical concepts of the lectures, as well as practical calculations ranging from geometry optimizations to rigid PES-scans and transition state searches. Supervision of the exercises and design of the report-grading system.

02.16-06.16
02.15-06.15 **Molecular Dynamics and Monte Carlo Simulations**

Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Course taught by **Prof. U. Röthlisberger**

Co-authoring and -design of a complete set of exercises for a 3rd year Bachelor's course on MD and Monte Carlo simulations. Exercises covered theoretical subjects from basic statistical mechanics to practical applications using an MD code.

03.15-05.15 **Semester Project in Computational Chemistry**

Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Supervision of Jérôme Bonnet and Valentin Queloz — Semester project

Design and supervision of a semester project for 3rd year Bachelor students, examining the influence of dispersion correction schemes on the spectra of metallocomplexes by performing *first principles* MD, including a small introduction to programming in Fortran.

10.14-12.14 **Computational Methods in Molecular Quantum Mechanics**

Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Supervision of Angelina Drahi and Marco Sutti — 7-week project

Design and supervision of a 7-week computational project for two 1st year Master students which involved assessing the influence of dispersion interactions on the physico-chemical properties of metal complexes.

05.10-06.10 **Substitute Teacher**

Gymnasium Interlaken/Gstaad

Teaching high-school level organic chemistry and kinetics.

Scientific Interests

CHEMISTRY - PHYSICS - MATHEMATICS: The fundamental, quantum mechanical principles that govern the chemistry of condensed matter: From the fundamental mathematics and approximate physics of Density Functional Theory (DFT) to their application in statistical mechanics, in particular using *first principles* Molecular Dynamics (MD). Efficient algorithms rooted in physics. Still passionate about organic chemistry and in particular reaction mechanisms. Enthusiastic about scientific programming and very fond of juggling equations.

Miscellaneous

OTHER INTERESTS AND HOBBIES: A friend of (not only foreign) languages, literature, poetry and theatre; listening to and making music (piano). Fascinated by technology and scale models. Loves animals (reptiles and canines), flowers (orchids, indigenous or not), hiking. Passionate gardener, enjoys cooking and baking.

Publications in Peer-Reviewed Scientific Journals

9 **Radical chain repair: The hydroalkylation of polysubstituted unactivated alkenes** G. Povie, S. R. Suravapu, M. P. Bircher, M. Meyer Mosjev, S. Rieder and P. Renaud: *Science Advances*, **2018**, 4 (7), eaat6031

8 **Exploiting coordinate scaling relations to accelerate exact exchange calculations** M. P. Bircher and U. Rothlisberger: *J. Phys. Chem. Lett.* , **2018**, 9 (14), 3886-3890

7 **Plane-wave implementation and performance of à-la-carte Coulomb-attenuated exchange-correlation functionals for predicting optical excitation energies in some notorious cases** M. P. Bircher and U. Rothlisberger: *J. Chem. Theory Comput.* , **2018**, 14 (6), 3184-3195

6 **Nonadiabatic effects in electronic and nuclear dynamics** M. P. Bircher, E. Liberatore, N. J. Browning, S. Brickel, C. Hofmann, A. Patoz, O. Unke, T. Zimmermann, M. Chergui, P. Hamm, U. Keller, M. Meuwly, H.-J. Woerner, J. Vaníček and U. Rothlisberger: *Structural Dynamics*, **2017**, 4, 061510

5 **Anthanthrene dye-sensitized solar cells: influence of the number of anchoring groups and substitution motif** Y. Geng, c. Yi, M. P. Bircher, S. Decurtins, M. Cascella, M. Grätzel and S.-X. Liu: *RSC Adv.* , **2015**, 5, 98643-98652

4 **Human infrared vision is triggered by two-photon chromophore isomerization** G. Palczewska, F. Vinberg, P. Stremplewski, M. P. Bircher, D. Salom, K. Komar, J. Zhang, M. Cascella, M. Wojtkowski, V. J. Kefalov, K. Palczewski: *PNAS*, **2014**, 111 (50), E5445-E5454

3 **A quinoxaline-fused tetrathiafulvalene-based sensitizer for efficient dye-sensitized solar cells** A. Amacher, C. Yi, J. Yang, M. P. Bircher, Y. Fu, M. Cascella, M. Grätzel, S. Decurtins and S.-X. Liu: *Chem. Commun.* , **2014**, 50, 6540-6542

2 **Electronic tuning effects *via* cyano substitution of a fused tetrathiafulvalenebenzothiadiazole dyad for ambipolar transport properties** A. Amacher, H. Luo, Z. Liu, M. Bircher, M. Cascella, J. Hauser, S. Decurtins, D. Zhang and S.X. Liu: *RSC Adv.* , **2014**, 4, 2873-2878

1 **On the acceleration of Cu electrodeposition by TBPS (3,3-thiobis-1-propanesulfonic acid): A combined electrochemical, STM, NMR, ESI-MS and DFT study.** N.T.M. Hai, J. Furrer, I. Giuroski, M.P. Bircher, M. Cascella and P. Broekmann: *J. Electrochem. Soc*, **2013**, 16 (12), D3158-D3164

Submitted Manuscripts

11 **From a week to less than a day: Speedup and scaling of coordinate-scaled exact exchange calculations in plane waves** M. P. Bircher and U. Rothlisberger: *Comp. Phys. Comm.* , **2018**, submitted

10 **Shedding light on the basis set dependence of the Minnesota functionals: Differences between plane waves, Slater functions and Gaussians** M. P. Bircher, P. Lopez-Tarífa and U. Rothlisberger: *J. Chem. Theory Comput.* , **2018**, submitted

Unpublished Work

15 **Density-difference driven molecular dynamics: Structural resolution based on electron density difference maps** M. P. Bircher, S. Bhattacharyya and U. Rothlisberger: , **2018**, in preparation

14 **Blobby: A tool to convert, interpolate and compare electron densities on arbitrary grids** M. P. Bircher, S. Bhattacharyya and U. Rothlisberger: , **2018**, in preparation

13 **An extension of the dispersion-corrected atom-centred potential library to metals and metalloids** M. P. Bircher, A. Laktionov and U. Rothlisberger: , **2018**, in preparation

12 **Capping potentials for accurate excitation energies without ghost states in polarisable QM/MM calculations** J. M. Haugaard Olsen, M. P. Bircher and U. Rothlisberger: , **2018**, in preparation

Software Contributions

3 **`xopt`: A modular minimiser for penalty functions**

2 **`blobby.x`: A program to convert, interpolate and compare electron density maps**

1 **Implementation of references (7-8,10-11,13 and 15), refactoring, bug fixes**
The `CPMD` code: www.cpmd.org

Talks, Seminars and Posters

Talks and Seminars

3 **Refinement of time-resolved experimental electron - density data through QM/MM MD simulations: the 10ps intermediate of bacteriorhodopsin**
S.C. van Keulen, M.P. Bircher, E. Liberatore, S. Bhattacharyya and U. Rothlisberger. ‘7th Annual Meeting of the NCCR MUST’, January 10-11 2017, Grindelwald, Switzerland

2 **A fast scheme for approximated Fock exchange potentials in plane wave implementations of Kohn-Sham density functional theory**
M.P. Bircher, S. Meloni and U. Rothlisberger. ‘SCS Fall Meeting 2016’, September 15 2016, University of Zürich, Zurich, Switzerland’

1 **The role of dispersion-correction for the description of metal compounds in density functional theory**
M.P. Bircher, A. Laktionov and U. Rothlisberger. ‘Modeling activity vs. selectivity in metalloproteins’, June 29 - July 1 2015, CECAM-FR-IDF, Paris, France

Posters

1 **The role of dispersion-correction in the description of metal-ligand bonds in density functional theory**
M.P. Bircher and U. Rothlisberger. 'SCS Fall Meeting 2015', September 4 2015, EPFL Lausanne, Switzerland'

Index of Chapters

Abstract (French) **v**

Abstract (English) **vii**

Abstract (German) **ix**

0 Preface **3**
- 0.1 Theoretical Chemistry - No Dry Matter 4
  - 0.1.1 Schrödinger 4
  - 0.1.2 Boltzmann 6

1 Introduction **9**
- 1.1 *In Silico*: From Work Bench to Core Processing Unit 9
  - 1.1.1 Computational Chemistry 10
- 1.2 Time-Scale *vs.* Accuracy and the Aim of this Thesis 11
  - 1.2.1 Accuracy: A Three Fold Problem 11
  - 1.2.2 Structure and Aim of the Present Work 12

I A Quantum Picture of Chemistry **15**

2 The Foundations of Molecular Dynamics **19**
- 2.1 Classical Statistical Mechanics 20
  - 2.1.1 The Fundamental Postulate in the Microcanonical Ensemble 20
  - 2.1.2 The Canonical Partition Function 21
- 2.2 Quantum Statistical Mechanics 23
  - 2.2.1 About Pure and Mixed States: The Density Operator 23

2.2.2 Time-evolution and Ensemble Averages: The Ergodic Hypothesis 26
2.3 Molecular Dynamics in the Born-Oppenheimer Formulation . 28
2.3.1 Separating the Hamiltonian: The Born-Oppenheimer Approximation 28
2.3.2 Nuclear Dynamics with the Born-Huang Expansion . . 30
Semiclassical Methods: Ehrenfest Dynamics 32
Semiclassical Methods: Trajectory Surface Hopping . . 34
Comments on Single-State Dynamics 35

3 The Electronic Structure Problem 37
3.1 The Universal Density Functional 37
3.1.1 Levy-Lieb Density Functional Theory 38
A Few Notes on the Time-Dependent Case 41
3.2 Kohn-Sham Density Functional Theory 41
3.2.1 The Kohn-Sham Formalism 42
On the Interpretation of Kohn-Sham Orbitals 46
Yet Another Note on the Time-Dependent Case 46
3.2.2 The Exact Form of the Exchange-Correlation Functional 47
3.2.3 The Exchange-Correlation Hole 49
3.2.4 Approximations to the Exchange-Correlation Functional 53
3.2.4$_{\mathrm{A}}$ Local Density Approximation (LDA) 54
3.2.4$_{\mathrm{B}}$ Generalised Gradient Approximation (GGA) 56
3.2.4$_{\mathrm{C}}$ Kinetic Energy Density: Meta-GGA 58
3.2.4$_{\mathrm{D}}$ Hybrid Functionals from the Adiabatic Connection . . 59
3.2.4$_{\mathrm{E}}$ Range-Separated Hybrid Functionals 60
Principles of Generalised Kohn-Sham Schemes 61

4 A Numerical Solution 65
4.1 An Expansion in Plane Waves 65
4.1.1 Infinitely Replicated Systems 66
Comments on the Nuclear Cusp 68
4.1.2 First Image Periodicity and the Poisson Solver 69
4.1.3 Exact Exchange Potentials in Plane Waves 72

II Climbing Jacob's Ladder in *first principles* Molecular Dynamics 75

5 À-la-carte Coulomb-Attenuation Method for Plane Waves 79

5.1 Introduction 80
5.2 Theory 85
5.2.1 The Coulomb-Attenuation Method 85
5.2.2 Singularity of the Coulomb Operator $\hat{\mathrm{W}}$ in Reciprocal Space 86
5.3 Implementation 88
5.4 Test Set and Computational Details 91
5.4.1 Description of the Test Set 91
5.4.2 Computational Setup 92
5.5 Results and Discussion 93
5.5.1 Convergence of Eigenvalues 93
$5.5.1_{\mathrm{A}}$ HOMO-LUMO Gaps in a Plane Wave/Pseudopotential Basis 93
$5.5.1_{\mathrm{B}}$ Comparison with Atom-Centred Basis Sets 95
5.5.2 Excitation Energies 97
5.5.3 Performance of New Customised *‘à-la-carte’* Coulomb-Attenuated Functionals 100
5.6 Conclusions 104

6 Basis Set Dependence of Minnesota Functionals 105
6.1 Introduction 106
6.2 Theory 111
6.2.1 The Minnesota Functionals 111
6.2.2 Expanding the Electron Density in Plane Waves . . . 116
6.2.3 Range-Separated and Screened Exchange in Plane Waves 117
6.3 Computational Details 117
6.4 Results and Discussion 118
6.4.1 The Total Energy of a HF Dimer 118
$6.4.1_{\mathrm{A}}$ Convergence in Plane Waves 118
$6.4.1_{\mathrm{B}}$ Comparison to Atom-Centred Bases 122
6.4.2 Implications for Reaction Enthalpies of the HC7/11 Database 126
$6.4.2_{\mathrm{A}}$ The M05 family 126
$6.4.2_{\mathrm{B}}$ The M06 family 127
$6.4.2_{\mathrm{C}}$ The M11 family 129
$6.4.2_{\mathrm{D}}$ Summary 131
6.4.3 Energy Density and Exact Exchange Analysis for the M06 Family 132
6.5 Computational Overhead 138
6.6 Conclusions 140

III Accelerating the Evaluation of Exact Exchange Integrals in Plane Waves 143

7 Coordinate Scaling Relations for Exact Exchange 147

8 Implementation of Coordinate-Scaled Exact Exchange 157
- 8.1 Introduction 158
- 8.2 Fast Fourier Transforms in Plane-Wave Kohn-Sham DFT . . 160
- 8.3 Coordinate-Scaled Exact Exchange 165
- 8.4 Implementation of Coordinate-Scaled Exact Exchange 166
- 8.5 Results and Discussion 170
  - 8.5.1 Box Size: Execution Times and Speedups 171
  - 8.5.2 Distribution of Orbital Pairs vs. Planes for various N_b 173
  - 8.5.3 Computational Cost as a Function of N_b and Effective Speedup 177
- 8.6 Summary 180

IV Retrospectives and Perspectives 181

9 Further Work and Projects 183
- 9.1 Dispersion Correction for Plane Wave DFT 183
  - 9.1.1 Atom-Centred Potentials for Molecular Properties . . 183
  - 9.1.2 Dispersion-Corrected Atom-Centred Potentials 185
  - 9.1.3 Extension of the DCACP Library 188
  - 9.1.4 Gold: A Case Study 190
  - 9.1.5 Direct Implementation of the DCACP Library in the CPMD Code 193
- 9.2 Density-Difference Driven MD: d3MD 193
  - 9.2.1 Biasing Potentials from Electron Density Differences . 194
  - 9.2.2 Application to Test Systems 194
  - 9.2.2$_\mathrm{A}$ Pushing Against the Barrier: *trans*-2-Butene to *cis*-2-Butene 195
  - 9.2.2$_\mathrm{B}$ Intermolecular Reactions: Diels-Alder Cyclo-Additions 197
  - 9.2.3 Electron Transfer in d3MD 201
  - 9.2.4 Versatile Grid Interpolation for Electron Density Maps 203

10 Conclusions and Outlook 207

V Postscript 213

A Supporting Information 215
- A.1 Supporting Information for Chapter 5 215
  - A.1.1 HOMO-LUMO Gaps in a Plane Wave/Pseudopotential Basis 215
  - A.1.2 Supercell Size for Plane Wave Calculations 217
  - A.1.3 Excitation Energies for the Test Set by Peach *et al.* . . 217
- A.2 Supporting Information for Chapter 7 217
  - A.2.1 Translational Invariance 217

List of Figures 221

List of Tables 227

Abbreviations and Nomenclature 234

Bibliography 235

Curriculum Vitae 255

Index of Chapters 261

Wir stehen selbst enttäuscht und seh'n betroffen
Den Vorhang zu und alle Fragen offen
Bertolt Brecht, 'Der gute Mensch von Sezuan'[279]