Efficient Implementation of Double-Hybrid Functionals for Condensed Phase Systems

---

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich
von
Frederick Stein
aus
Deutschland

Promotionskommission
Prof. Dr. Jürg Hutter (Vorsitz)
Prof. Dr. Titus Neupert
Prof. Dr. Markus Reiher

Zürich, 2022

Contents

1 Introduction **7**

2 Theoretical Background **9**
2.1 Fundamentals 9
2.2 The Hartree-Fock approximation 11
2.3 The MP2 approximation 15
2.4 Resolution-of-the-Identity Approximation 19
2.5 Density Functional Theory 20
2.6 Random-Phase Approximation 22
2.7 Double-Hybrid Density Functionals 24
2.8 CP2K and the GPW approach 28

3 Implementation of Double-Hybrid Functionals with Range-Separated Correlation Functional **29**
3.1 Introduction 29
3.2 Theoretical Background 30
3.3 Implementation of Integrals within the GPW approach 34
3.3.1 Motivation and Theoretical Background 34
3.3.2 Performance and Accuracy 36
3.4 Pseudopotentials 40
3.5 Application to Rare Gas Dimers 44
3.6 Conclusions 46
3.7 Appendix: Pseudopotential Parameters and Basis Sets 46

4 Double-Hybrid DFT Functionals for the Condensed Phase: Gaussian and Plane Waves Implementation and Evaluation **55**
4.1 Introduction 55
4.2 Theoretical Background 57

4.3 Computational Details 60
4.3.1 Gaussian and plane waves method (GPW) and integral evaluation 60
4.3.2 Test systems 61
4.3.3 Parameters of the calculations 61
4.3.4 Choice of functionals and implementation 62
4.3.5 Basis sets and pseudopotentials 63
4.3.6 Cohesive energies and basis set superposition error 63
4.4 Results 64
4.4.1 General Remarks 64
4.4.2 Convergence with respect to super cell size 67
4.4.3 Convergence of the BSSE 67
4.4.4 Convergence with respect to basis set size 70
4.5 Discussion 72
4.6 Conclusions 72

5 Double-Hybrid Density Functionals for the Condensed Phase: Gradients, Stress Tensor, and Auxiliary-Density Matrix Method Acceleration 74
5.1 Introduction 74
5.2 Theoretical Background 76
5.3 Computational Details 83
5.4 Accuracy 84
5.5 Performance 84
5.6 Large Benchmark systems 92
5.7 Discussion 94
5.8 Conclusions 95

6 Conclusions and Outlook 96

For my family

Abstract

Density-functional-based methods represent the working horse for condensed matter calculations. Despite of their efficiency and acceptable accuracy, they describe weakly interacting systems only unsatisfactorily and a systematic improvement of the underlying models is not in sight. Wave function methods, in contrast, are very accurate, systematically improvable methods. Because of their poor scalability, their application is restricted to molecular systems. Furthermore, the lower levels of wave function theory tend to perform worse than many density functionals due to the missing short-ranged correlation effects. Double-hybrid density functionals (DHDFs) mediate between both approaches by combining them. The advantage of this methodology is the significant improvement upon both, density functional theory and wave function theory. However, DHDFs are restricted to mostly molecular systems which is related to the high computational costs of the underlying wave function correlation method.

In the present thesis, we consider different possibilities to broaden the applicability of DHDFs to periodic systems. After a review of the fundamental theoretical aspects, we discuss the implementation of DHDFs with range-separated exchange and correlation functionals in CP2K to improve the bad basis set convergence. Afterwards, we test different DHDFs in condensed phase and analyze the different kinds of errors of these functionals. Finally, we describe an implementation of analytical gradients for DHDFs in condensed phase. The application of the auxiliary density matrix method reduces the computational costs of the required Hartree-Fock exchange contribution. Therewith, we provide the tools for advanced studies like the calculation of training data set for machine-learning models or simulations on the accuracy level of DHDFs.

Zusammenfassung

Dichtefunktionalbasierte Methoden (DFM) stellen aufgrund ihrer Effizienz und akzeptablen Genauigkeit das Arbeitspferd für Rechnungen an kondensierten Systemen dar. Allerdings beschreiben sie schwach wechselwirkende Systeme aber oft nur mit unbefriedigender Genauigkeit und eine systematische Verbesserung der zugrundeliegenden Modelle ist nicht in Sicht. Dagegen stehen mit Wellenfunktionsmethoden (WFM) sehr genaue, systematisch verbesserbare Methoden zur Verfügung, deren Anwendung allerdings aufgrund der schlechten Skalierbarkeit auf molekulare Systeme beschränkt bleibt. Ausserdem beschreiben die niedrigeren Theorielevel kurzreichweitige Korrelationseffekte nur unzureichend, welche in ungenaueren Ergebnissen als bei typischen DFM resultiert. Einen Kompromiss zwischen beiden Ansätzen stellen Doppelhybridfunktionale (DHF) dar, welche DFM mit korrelierten WFM verbinden und dabei die Genauigkeit gegenüber beiden Methodenklassen signifikant verbessern können. Allerdings bleibt deren Anwendbarkeit wegen der hohen Kosten der zugrundeliegenden WFM im Wesentlichen auf molekulare Systeme beschränkt.
In der vorliegenden Dissertationsschrift werden verschiedene Möglichkeiten betrachtet, um die Anwendbarkeit von DHF bei periodischen Systemen zu erhöhen. Nach einer Wiederholung der wesentlichen theoretischen Aspekte wird die Implementierung von DHF mit reichweitenseparierten Austausch- und Korrelationsfunktional in CP2K diskutiert, um die schlechte Konvergenz bezüglich der Basissatzgrösse zu verbessern. Im Anschluss testen wir unterschiedliche DHF in kondensierter Phase und analysieren verschiedene auftretende Fehler. Zuletzt beschreiben wir eine Implementierung von Gradienten für DHF für kondensierte Systeme und reduzieren die Rechenkosten des notwendigen Hartree-Fock-Austauschterms mit der Auxiliardichtemethode. Wir stellen damit Werkzeuge für weiterführende Studien wie die Berechnung von Trainingsdatensätzen für Modelle maschinellen Lernens oder Simulationen auf dem Genauigkeitsniveau von DHF zur Verfügung.

Acknowledgement

I thank Prof. Dr. Jürg Hutter for giving me the opportunity of a PhD position in his group and the access to his wide knowledge in quantum chemistry and algorithms.
I thank Prof. Dr. Markus Reiher and Prof. Dr. Titus Neupert for their membership in my PhD commitee. I thank Prof. Dr. Luuk Visscher to thankfully prepare the external expert report.
I thank all former and current members of Hutter group for the great time at UZH, lunch/ coffee/ tea breaks, barbecues etc. Especially I thank Tiziano Müller for his patience with technical support, Vladimir Rybkin for valuable comments on my work and music related discussions, Jan Wilhelm for discussions about the code, Patrick Seewald for many fruitful discussion on DBCSR tensors, RI methods and integral calculation, Anna-Sophie Hehn for the discussions on correlated and explicitly-correlated methods and the few times she joined for lunch (do it more often!), Augustin Bussy for discussions on RI methods and Alfio Lazzaro, Ole Schütt and Mathieu Taillefumier for their technical support with GPUs. Especially, I thank Anna and Augustin for proof-reading this thesis.
Mutti, Vati und Oma Hannelore, ich danke euch für eure mentale Unterstützung während meiner Promotion, insbesondere während der Corona-Krise und meiner privaten Umorientierungsphase. Ich hätte in meiner Einsamkeit fast die Nerven verloren und nur wegen euch habe ich durchhalten können.
Mahal, thank you for your love, your support and the wonderful time, we have already spent together. Let us stay together for many more years and move together as soon as possible. Mahal na Mahal kita.

Chapter 1

Introduction

Although the concept of the classical chemical bond enables the description of bond formation, it is the intermolecular interactions that are required to describe enzyme-substrate complexes, molecular crystals, liquids or adsorption processes on surfaces.[1–4]
The working horse of calculations in condensed phase is density functional theory (DFT) within the Kohn-Sham approach with a large number of linear-scaling implementations available.[5–23] Common density functionals (DFs) provide a reasonable accuracy for a variety of chemically bonded systems and can be applied to systems containing several thousands as even up to millions of atoms.[11, 24, 25] Despite of its success, a systematic approach for improved DFs is missing. Furthermore, its drawback to describe weakly bounded systems has to be overcome by special van-der-Waals DFs or different empirical or theoretical correction schemes.[26–33]
Wave function methods provide a different approach by solving the Schrödinger equation, the fundamental equation of quantum mechanics, approximately.[34–36] They can be systematically improved and treat dispersion interactions with ease.[37–40] Their significantly higher computational costs and unfavourable scaling of at least fourth order with respect to the system size decreases the number of potential applications compared to cubically scaling canonical DFT implementations. Low-scaling approaches are capable of reducing the tremendous scaling down to linear scaling, however they are often developed with molecular systems in mind. As a consequence, these implementations are not suitable for periodic systems due to less sparse tensors or missing consideration of high parallelizability required for large super-computing devices.[41–52] In addition, the simplest wave function approaches like the second-order Møller-Plesset perturbation the-

ory (MP2) perform even worse than many conventional DFs leading to the development of reparametrized methods like SCS-MP2, simplified schemes like SOS-MP2 or corrected versions like RPA+AXK or RPA+SOSEX.[38, 53–57] The slow convergence of the obtained quantities with respect to the basis set size has to be dealt with extrapolation schemes, explicitly-correlated methods or special range-separated methods.[58–65]

Double-hybrid DFs (DHDFs) are modern quantum chemical approaches combining conventional DFs with simple wave function correlation methods and can supposedly perform better than the underlying methods.[66, 67] Their excellent accuracy for molecular systems encourages the application of these methods to condensed phase systems.[24, 25] Due to the high computational costs of correlated methods in condensed phase, actual applications of DHFs in condensed phase are limited, benchmark studies scarce and dynamics simulations usually unfeasible.[68–70]

The thesis presented here will tackle the high computational costs and provide more studies of DHFs for condensed matter. We will start with a discussion of the most important theoretical aspects and the GPW approach as implemented in the CP2K simulation package. We will use CP2K for all DHF implementations and benchmarks. Subsequently, we will discuss the implementation of DHFs with a long-range Coulomb operator which are supposed to remove the slow basis set convergence of the total energies.[64] Next, we will present a benchmark study of different classes of DHFs discussing further sources of errors related to this class of approaches.[71] Finally, we will present a gradient implementation of MP2-based DHFs for large systems and discuss a method to reduce the overhead of the Hartree-Fock calculations in the context of DHFs and present large benchmark systems which can be treated with the given methodology.[72]

Chapter 2

Theoretical Background

The following chapter aims to provide a general understanding of the methodologies discussed in the subsequent chapters. Most information can be found in standard textbooks about quantum chemistry. If nothing else is stated, the information was taken from references [73–75]. Integrals over spatial coordinates are carried out over the whole three-dimensional space, summations over all possible values of an index, summations with lower index $i < j$ run over all unique pairs i, j.

2.1 Fundamentals

The basis of non-relativistic quantum mechanics is the Schrödinger equation (SE). The time-independent SE is the eigenvalue equation

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle \tag{2.1}$$

with the constant energy of the system E, the corresponding eigenstate $|\Psi\rangle$ and the time-independent Hamilton operator (Hamiltonian) or total energy operator $\hat{H}$. $|\Psi\rangle$ represents the many-particle quantum state and underlies the normalization constraint

$$\langle\Psi|\Psi\rangle = 1. \tag{2.2}$$

To perform actual calculations and apply the SE, we have to work in a given representation. Here, we will consider the real space representation only in which the state $|\Psi\rangle$ is represented as a wave function (WF) depending on the coordinates

of all particles in the system. The square of its modulus is interpreted as the probability of finding the particles in space.
Chemically relevant systems contain K nuclei, described as point charges with charge Z_A at position $\vec{R}_A$ and mass M_A, $A = 1...K$ and N electrons, at positions $\vec{r}_n$, $n = 1...N$, respectively. The Hamilton operator in real space is given in atomic coordinates as

$$\hat{H} = -\frac{1}{2}\sum_A \frac{\Delta_A}{M_A} - \frac{1}{2}\sum_n \Delta_n + \sum_{A<B} \frac{Z_A Z_B}{\left|\vec{R}_A - \vec{R}_B\right|} - \sum_A \sum_n \frac{Z_A}{\left|\vec{R}_A - \vec{r}_n\right|} + \sum_{n<m} \frac{1}{|\vec{r}_n - \vec{r}_m|} \tag{2.3}$$

with Δ_A and Δ_n denoting the Laplace operators defined with respect to the coordinates of nucleus A and the electronic coordinates of electron n, respectively. In the given order, the different terms represent the kinetic energy operator of the nuclei, the kinetic energy of the electrons, the nuclei-nuclei repulsion energy, the nuclear-electron attraction energy and the electron-electron interaction energy.
The SE can in general not be solved analytically such that we have to rely on approximations. We assume that the many-particle WF is separable according to

$$\Psi\left(\vec{R}_1,\ldots,\vec{R}_K,\vec{r}_1,\ldots,\vec{r}_N\right) = \chi\left(\vec{R}_1,\ldots,\vec{R}_K\right)\psi\left(\vec{r}_1,\ldots,\vec{r}_N\right), \tag{2.4}$$

with the nuclear WF $\chi\left(\vec{R}_1,\ldots,\vec{R}_K\right)$ and the electronic WF $\psi\left(\vec{r}_1,\ldots,\vec{r}_N\right)$. The Born-Oppenheimer approximation further assumes that the movement of the nuclei is static for the electrons. This approximation is justified by the large ratio of the masses of nuclei and electrons. Then, the nuclear and the electronic WFs obey their respective SEs

$$\hat{H}^{\text{nuc}}\chi\left(\vec{R}_1,\ldots,\vec{R}_K\right) = E^{\text{nuc}}\chi\left(\vec{R}_1,\ldots,\vec{R}_K\right), \tag{2.5}$$

$$\hat{H}^{\text{nuc}} = -\frac{1}{2}\sum_A \frac{\Delta_A}{M_A} + E^{\text{el}}\left(\vec{R}_1,\ldots,\vec{R}_K\right) \tag{2.6}$$

$$\hat{H}^{\text{el}}\psi\left(\vec{r}_1,\ldots,\vec{r}_N\right) = E^{\text{el}}\left(\vec{R}_1,\ldots,\vec{R}_K\right)\psi\left(\vec{r}_1,\ldots,\vec{r}_N\right) \tag{2.7}$$

$$\hat{H}^{\text{el}} = -\frac{1}{2}\sum_n \Delta_n - \sum_A \sum_n \frac{Z_A}{\left|\vec{R}_A - \vec{r}_n\right|} + \sum_{n<m} \frac{1}{|\vec{r}_n - \vec{r}_m|}. \tag{2.8}$$

We note that the electronic WF depends parametrically on the positions of the nuclei and the nuclear movement depends on the electronic energy E_{el}. The total energy of the system E_{tot} with fixed nuclei is then given as

$$E^{\mathrm{tot}} = E^{\mathrm{el}} + \sum_{A<B} \frac{Z_A Z_B}{\left|\vec{R}_A - \vec{R}_B\right|}. \tag{2.9}$$

We will refer to the quantum state with the smallest energy E^{el} as the ground state $|\Psi_0\rangle$, the corresponding WF as ground state WF, its corresponding energy as the ground state energy E_0. If there is more than one WF to any given energy, the energy level is called degenerate.
Most (electronic) SEs cannot be solved analytically implying the necessity of further approximations, but we can establish a variational principle. Let $|\Psi_0\rangle$ be the only ground state of a SE with energy E_0 and be $\left|\tilde{\Psi}_0\right\rangle = |\Psi_0\rangle + |\delta\rangle$ an approximate ground state. Its energy expectation value $\tilde{E}$ is given by

$$\tilde{E}_0 = \frac{\left\langle\tilde{\Psi}_0\left|\hat{H}\right|\tilde{\Psi}_0\right\rangle}{\left\langle\tilde{\Psi}_0|\tilde{\Psi}_0\right\rangle}. \tag{2.10}$$

The variational principle states that $\tilde{E}_0 \geq E_0$ with equality if and only if $|\delta\rangle = 0$ and provides us with an approximate approach, the variational method. Let $\left|\tilde{\Psi}_0(C)\right\rangle$ with free parameter(s) C be a test quantum state. Then, we obtain the best approximation with respect to the energy by optimizing its energy expectation value with respect to C. Due to the variational principle, the obtained energy is then an upper bound to the correct energy.
In the following, we will consider only many-electron systems and thus ignore the index el. As such, we will also not explicitly add the nuclei-nuclei interaction term. In addition, we will work in real space only and represent quantum states by their WF.

2.2 The Hartree-Fock approximation

Electrons are identical indistinguishable fermions with total spin $s = \frac{1}{2}$. This property implies the Pauli principle stating that the electronic WF has to be antisymmetric with respect to the exchange of two electronic coordinates. Here, we will write the N-electron WF $\Psi(x_1, ..., x_N)$ with the variable $x_i = (\vec{r}_i, \sigma_i)$ which combines the spatial coordinate $\vec{r}_i$ and the spin variable σ_i. The antisymmetry con-

straint can be fulfilled by representing the WF as a Slater determinant

$$\Phi(x_1,\ldots,x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(x_1) & \ldots & \psi_N(x_1) \\ \vdots & \ddots & \vdots \\ \psi_1(x_N) & \ldots & \psi_N(x_N) \end{vmatrix} \tag{2.11}$$

where the spin orbitals $\psi_I(x) = \phi_I(\vec{r})\chi_I(\sigma)$ form an orthonormal basis. The spin variable of an electron has one of two values $\pm\frac{1}{2}$ often denoted α,β. The spin functions $\chi_i(\sigma)$ can be only one of the two functions $A(\sigma)$ or $B(\sigma)$ defined by their function values $A(\alpha) = 1, A(\beta) = 0, B(\alpha) = 0, B(\beta) = 1$.
Due to the properties of the spin orbitals, the real-space parts of the spin orbitals are divided into two groups of functions depending on the spin state of the corresponding orbital. We will write for the index $I = i_\sigma$ with σ being the spin state of orbital I and i being an index enumerating the real-space functions to each spin-channel such that for spin channel σ we have $i = 1...N_\sigma$ and $N = N_\alpha + N_\beta$ with the number of electrons N_σ in spin channel σ. Each set of functions $\phi_{i\sigma}(\vec{r})$ by themselves form an orthonormal basis in real space implying that the integral of two different spin orbitals are zero because of different spin states or else because of different real-space functions. Furthermore, we will consider Γ-point calculations only such that the real-space functions can be chosen real.
The simplest approximation to solve the SE is given by the Hartree-Fock (HF) approximation which applies the variational principle to a single Slater determinant by optimizing the energy expectation value with respect to the spin orbitals. The slater determinant automatically fulfills the antisymmetry property of the WF due to the antisymmetry of the determinant with respect to the interchange of two rows equivalent to the interchange of two coordinates.
With the HF approximation, one obtains an equation to determine the spin orbitals

of each spin channel σ (HF eigenvalue equation)

$$\hat{h}^{\mathrm{HF}}\psi_{i_\sigma}(\vec{r}) = \varepsilon_{i_\sigma}\psi_{i_\sigma}(\vec{r}) \tag{2.12}$$

$$\begin{aligned} &= -\frac{\Delta}{2}\psi_{i_\sigma}(\vec{r}) - \sum_A \frac{Z_A}{\left|\vec{R}_A - \vec{r}\right|}\psi_{i_\sigma}(\vec{r}) \\ &\quad + \sum_{\sigma'}\sum_{j_{\sigma'}} \int d^3r' \frac{\left|\psi_{j_{\sigma'}}(\vec{r}')\right|^2}{|\vec{r}-\vec{r}'|}\psi_{i_\sigma}(\vec{r}) \\ &\quad - \sum_{j_\sigma} \int d^3r' \frac{\psi^*_{j_\sigma}(\vec{r}')\psi_{i_\sigma}(\vec{r}')}{|\vec{r}-\vec{r}'|}\psi_{j_\sigma}(\vec{r}) \end{aligned} \tag{2.13}$$

with the orbital energy ε_{i_σ} of orbital i_σ. The first two terms of equation (2.13) describe the kinetic energy and the electric potential of the nuclei. The third term is called Hartree term and describes the classical electrostatic potential of the electrons caused by the electronic density

$$n(\vec{r}) = \sum_\sigma \sum_{i_\sigma} |\phi_{i_\sigma}(\vec{r})|^2 . \tag{2.14}$$

The last term in equation (2.13) is the Fock exchange term and is related to the anti-symmetry property of the electrons. In contrast to the Hartree potential, it is a non-local potential. All terms of the HF Hamiltonian except from the kinetic energy term describe an effective potential in which a single electron is moving. Because the one-particle HF Hamiltonian $\hat{h}^{\mathrm{HF}}$ depends on all orbitals, an analytical solution is in general not possible but starting from an initial guess for the orbitals, one solves the HF equations, obtains a new set of orbitals, reconstructs the Hamiltonian and repeats this procedure until self-consistency (self-consistent field method, SCF method).

The electronic HF energy is then given by

$$E^{\mathrm{HF}} = \sum_{\sigma}\sum_{i_\sigma}\int d^3r\,\phi^*_{i_\sigma}(\vec{r})\left(-\frac{\Delta}{2}-\sum_A\frac{Z_A}{\left|\vec{R}_A-\vec{r}\right|}\right)\phi_{i_\sigma}(\vec{r})$$
$$+\frac{1}{2}\sum_{\sigma\sigma'}\sum_{i_\sigma j_{\sigma'}}\left((i_\sigma i_\sigma|j_{\sigma'}j_{\sigma'})-\delta_{\sigma\sigma'}(i_\sigma j_\sigma|j_\sigma i_\sigma)\right) \tag{2.15}$$
$$=\sum_{\sigma}\sum_{i_\sigma}\varepsilon_{i_\sigma}-\frac{1}{2}\sum_{\sigma\sigma'}\sum_{i_\sigma j_{\sigma'}}\left((i_\sigma i_\sigma|j_{\sigma'}j_{\sigma'})-\delta_{\sigma\sigma'}(i_\sigma j_\sigma|j_\sigma i_\sigma)\right) \tag{2.16}$$

with the electron repulsion integral (ERI)

$$(i_\sigma j_\sigma|k_{\sigma'}l_{\sigma'}) = \int d^3r\,d^3r'\frac{\phi^*_{i_\sigma}(\vec{r})\phi_{j_\sigma}(\vec{r})\phi^*_{k_{\sigma'}}(\vec{r}')\phi_{l_{\sigma'}}(\vec{r}')}{|\vec{r}-\vec{r}'|}. \tag{2.17}$$

An important special case is the one-electron system. Here, we have only a single orbital with real-space part $\phi(\vec{r})$ and the expressions of the electronic energy and the HF equation simplify to

$$E^{\mathrm{HF}} = -\int d^3r\,\phi^*(\vec{r})\hat{h}^{\mathrm{HF}}\phi(\vec{r}), \tag{2.18}$$
$$\hat{h}^{\mathrm{HF}}\phi(\vec{r}) = -\frac{\Delta}{2}\phi(\vec{r})-\sum_{A=1}^{N}\frac{Z_A}{\left|\vec{R}_A-\vec{r}\right|}\phi(\vec{r}). \tag{2.19}$$

We note that the HF Hamiltonian and the Hamiltonian of the SE 2.8 coincide implying that the HF theory is exact for one-electron systems.
Most quantum chemistry codes solve the HF equations by expanding the real-space parts of the spin orbitals into a finite basis of functions $f_\mu(\vec{r})$, $\mu = 1...S$, the so-called basis set (of size S). The expansion coefficients of the real-space functions of each spin channel σ are denoted by $C_{i\sigma\mu}$. Considering the orthonormality constraint on the orbitals, the expansion coefficients of each spin-channel are solutions of the Roothaan-Hall equations

$$\sum_\nu F^{\mathrm{HF}}_{\mu\nu\sigma}C_{i\sigma\nu} = \varepsilon_{i\sigma}\sum_\nu S_{\mu\nu}C_{i\sigma\nu} \tag{2.20}$$

with the elements of the HF matrix of spin channel σ

$$F^{\mathrm{HF}}_{\mu\nu\sigma} = \int d^3r\, f^*_\mu(\vec{r})\hat{h}^{\mathrm{HF}} f_\nu(\vec{r}) \tag{2.21}$$

and the elements of the overlap matrix

$$S_{\mu\nu} = \int d^3r\, f^*_\mu(\vec{r}) f_\nu(\vec{r}). \tag{2.22}$$

Although the Roothaan-Hall equations appear like a general eigenvalue problem, these equations have to be solved iteratively because the HF matrix depends on the orbitals themselves.

2.3 The MP2 approximation

A single Slater determinant is in general not a good approximation of the correct many-electron ground state. Due to the variational principle, the correct ground state energy must be lower than the energy obtained from the HF approximation. The error between the HF ground state energy and the correct energy of the ground state is called correlation energy.
A systematic way to improve upon the HF approximation is to employ a superposition of Slater determinants obtained by replacing the spin orbitals of the HF ground state Slater determinant (so-called occupied orbitals) with further solutions of the HF equations orthogonal to the occupied orbitals (so-called unoccupied orbitals) and optimize the expectation value of the energy with respect to the expansion coefficients, an approach called Configuration Interaction (CI). The minimization condition results in an eigenvalue problem. Because its size scales exponentially with respect to the number of electrons, its solution becomes unfeasible for systems with more than a few electrons.
In order to circumvent the exponential growth of the method, we can truncate the expansion to a subset of the possible Slater determinants. Although the scaling of these methods is more favorable than the original CI scheme, we lose the size extensivity of the method. Size extensivity reflects that the energy of a system containing two non-interacting subsystems is the sum of the energy of the two subsystems. It turns out that the truncated CI expansion is not size-extensive which gives rise to alternative WF correlation schemes like the Quadratic Configuration Interaction (QCI) methods, the Coupled Cluster (CC) Methods, the Møller-Plesset perturbation theory (MPn) or the different Multi-Reference Methods and the Den-

sity Matrix Renormalization Group (DMRG).[38–40, 76–79]
In the following, we will focus on Møller-Plesset perturbation theory which is based on the following exact decomposition of the Hamiltonian

$$\hat{H}\Psi = \hat{H}^{\mathrm{HF}}\Psi + \left(\hat{H} - \hat{H}^{\mathrm{HF}}\right)\Psi = E\Psi \tag{2.23}$$

with the HF operator $\hat{H}^{\mathrm{HF}} = \sum_{i=1}^{N} \hat{h}_i^{\mathrm{HF}}$ with $\hat{h}_i^{\mathrm{HF}}$ being the HF operator from equation (2.13) operating on particle i. The assumption is that the operator $\hat{H} - \hat{H}^{\mathrm{HF}}$ is small enough permitting the application of perturbation theory. We introduce a formal perturbation parameter λ before the perturbation term which is afterwards set to one according to

$$\hat{H}^{\mathrm{HF}}\Psi + \lambda\left(\hat{H} - \hat{H}^{\mathrm{HF}}\right)\Psi = E\Psi. \tag{2.24}$$

Using Rayleigh-Schrödinger perturbation theory, we write the energy and the WF as a series of the perturbation parameter according to[80]

$$E = E_0 + \lambda E_1 + \lambda^2 E_2 + \ldots, \tag{2.25}$$

$$\psi = \psi_0 + \lambda\,\psi_1 + \lambda^2\psi_2 + \ldots. \tag{2.26}$$

Plugging this expression into the above decomposition and comparing the coefficients of the different powers of λ and we obtain up to the second order in λ

$$\hat{H}^{\mathrm{HF}}\Psi_0 = E_0\Psi_0, \tag{2.27}$$

$$\hat{H}^{\mathrm{HF}}\Psi_1 + \left(\hat{H} - \hat{H}^{\mathrm{HF}}\right)\Psi_0 = E_0\Psi_1 + E_1\Psi_0, \tag{2.28}$$

$$\hat{H}^{\mathrm{HF}}\Psi_2 + \left(\hat{H} - \hat{H}^{\mathrm{HF}}\right)\Psi_1 = E_0\Psi_2 + E_1\Psi_1 + E_2\Psi_0. \tag{2.29}$$

The solution to the zeroth-order equation (2.27) is the HF WF, the zeroth-order energy is however only the sum of the HF orbital energies of the occupied orbitals

$$E_0 = \sum_{I=1}^{N} \varepsilon_I, \tag{2.30}$$

not the HF energy. To obtain the HF energy, we have to consider the first-order correction, too, which is given by

$$E_1 = -\sum_{IJ} \left((II|JJ) - (IJ|JI)\right). \tag{2.31}$$

If we compare the zeroth- and first-order energy corrections with the HF energy, we obtain

$$E^{\mathrm{HF}} = E_0 + E_1. \tag{2.32}$$

Thus, we preserve the HF theory with the perturbation ansatz.
The determining equation for the first-order WF correction is the solution to

$$\left(\hat{H}^{\mathrm{HF}} - E_0\right)\Psi_1 = \left(E^{\mathrm{HF}} - \hat{H}\right)\Psi_0. \tag{2.33}$$

We try to obtain an expansion of the first-order expansion in terms of all possible Slater determinants. Let k be the number of the occupied orbitals replaced by the same amount of different unoccupied orbitals. Then, we call the resulting Slater determinant to be a k-times excited Slater determinant. Because the HF Hamiltonian and the the exact Hamiltonian are one- and two-electron operators, respectively, the contributions from more than doubly excited Slater determinants to the first-order WF correction are zero.
Due to the Brillouin theorem, the coupling between the HF WF and the singly excited determinants is zero, and we are left with contributions from the doubly excited states. We denote the doubly excited states with Ψ_{IJ}^{AB} being the Slater determinant obtained by replacing the occupied orbitals I,J with the unoccupied orbitals A,B in the HF WF. Imposing the orthogonality of the first-order correction and the HF WF, the first-order WF correction is simply

$$\Psi_1 = \sum_{I<J}\sum_{A<B} t_{IJ}^{AB}\Psi_{IJ}^{AB}. \tag{2.34}$$

The first-order amplitudes t_{IJ}^{AB} are obtained after multiplication of equation (2.33) with Ψ_{IJ}^{AB} and integration

$$t_{IJ}^{AB} = -\frac{(AI|BJ) - (AJ|BI)}{\varepsilon_A + \varepsilon_B - \varepsilon_I - \varepsilon_J}. \tag{2.35}$$

With the first-order WF correction, we determine the second order energy correction by multiplying equation (2.29) with Ψ_0, subtracting the product of equation (2.27) and Ψ_2 and integrating. We obtain the second-order energy

$$E_2 = \sum_{I<J}\sum_{A<B} t_{IJ}^{AB}\left((IA|JB) - (JA|IB)\right). \tag{2.36}$$

The sum of the HF energy and the second-order energy defines the second-order Møller-Plesset perturbation energy (MP2).
The indices I, J, A, B refer to states with any spin state. Because of the properties of the spin orbitals, an ERI (2.17) can only be non-zero if both spin orbitals with the same electronic coordinate have the same spin state. To account for this behavior in the equations, we will write a spin state I (and analogously for the unoccupied orbitals) as i_σ with $\sigma = \alpha, \beta$. Then, the second order energy is

$$E_2 = -\sum_{\sigma} \sum_{i_\sigma < j_\sigma} \sum_{a_\sigma < b_\sigma} \frac{((i_\sigma a_\sigma | j_\sigma b_\sigma) - (j_\sigma a_\sigma | i_\sigma b_\sigma))^2}{\varepsilon_{a_\sigma} + \varepsilon_{b_\sigma} - \varepsilon_{i_\sigma} - \varepsilon_{j_\sigma}} - \sum_{i_\alpha j_\beta} \sum_{a_\alpha b_\beta} \frac{(i_\alpha a_\alpha | j_\beta b_\beta)^2}{\varepsilon_{a_\alpha} + \varepsilon_{b_\beta} - \varepsilon_{i_\alpha} - \varepsilon_{j_\beta}}. \tag{2.37}$$

The first term in equation (2.37) is called same-spin or triplet component, the second term is the opposite-spin or singulet component. To increase the accuracy of MP2, both components can be rescaled (SCS-MP2, SNS-MP2) or the same-spin term is omitted enabling a lower-scaling approach (SOS-MP2).[54, 55, 81]
An important special case represent closed-shell systems with $i_\alpha = i_\beta = i$ such that

$$E_2 = -2 \sum_{i<j} \sum_{a<b} \frac{((ia|jb) - (ja|ib))^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} - \sum_{ij} \sum_{ab} \frac{(ia|jb)^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}. \tag{2.38}$$

After lifting the restrictions $i < j$ and $a < b$, we obtain the more common form

$$E_2 = -\sum_{ij} \sum_{ab} \frac{(ia|jb)^2 - (ia|jb)(ja|ib)}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} - \sum_{ij} \sum_{ab} \frac{(ia|jb)^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{2.39}$$

$$= -\frac{1}{2} \sum_{ij} \sum_{ab} \frac{2(ia|jb)^2 - (ia|jb)(ja|ib)}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}. \tag{2.40}$$

We note that the quality of the MP2 energy will depend on the number of employed unoccupied orbitals. In practice, the convergence with respect to the number of unoccupied orbitals is slow but smooth and the slow convergence is tackled with extrapolation schemes or improved with a different representation of the first-order WF correction within explicitly correlated MP2.[58–60]
Finally, Møller-Plesset perturbation theory provides self-correlation free meth-

ods meaning that the correlation energy of a one-electron system is correctly described as zero because in this case, the HF WF provides the correct solution of the Schrödinger equation.

2.4 Resolution-of-the-Identity Approximation

A bottleneck for calculations within WFT is the evaluation of the ERIs. An important approach to reduce the computational efforts of their calculation is the Resolution-of-the-Identity (RI) approximation.[82–84] The RI approximation is related to the problem of density fitting, i.e. approximating a density ρ by a suitable linear combination of auxiliary basis functions f_i according to

$$\rho \approx \tilde{\rho} = \sum_i a_i f_i. \tag{2.41}$$

We will minimize the interaction energy due to a potential operator $g(r) \geq 0$ in a least square sense which means, find the components a_i of the vector $\mathbf{a}$ such that

$$\int d^3r_1 d^3r_2 \left(\rho(\vec{r}_1) - \tilde{\rho}(\vec{r}_1)\right) g\left(|\vec{r}_1 - \vec{r}_2|\right) \left(\rho(\vec{r}_2) - \tilde{\rho}(\vec{r}_2)\right) \tag{2.42}$$

is minimized. The coefficients are then given by

$$\mathbf{a} = \mathbf{S}^{-1}\mathbf{t}, \tag{2.43}$$

$$S_{ij} = \int d^3r_1 d^3r_2 f_i(\vec{r}_1) g\left(|\vec{r}_1 - \vec{r}_2|\right) f_j(\vec{r}_2), \tag{2.44}$$

$$t_i = \int d^3r_1 d^3r_2 \rho(\vec{r}_1) g\left(|\vec{r}_1 - \vec{r}_2|\right) f_i(\vec{r}_2). \tag{2.45}$$

Please note that the matrix $\mathbf{S}$ defined here is different from the one used in the HF section.
Calculating the interaction energy due to a second operator $v(r)$, we obtain the expression

$$\int d^3r_1 d^3r_2 \rho(\vec{r}_1) v\left(|\vec{r}_1 - \vec{r}_2|\right) \rho(\vec{r}_2) = \mathbf{t}^T \mathbf{S}^{-1} \mathbf{V} \mathbf{S}^{-1} \mathbf{t} \tag{2.46}$$

$$V_{ij} = \int d^3r_1 d^3r_2 f_i(\vec{r}_1) v\left(|\vec{r}_1 - \vec{r}_2|\right) f_j(\vec{r}_2). \tag{2.47}$$

In the context of approximating Coulomb integrals, $g(r)$ is usually referred to as metric. With the trivial choice $g(r) = v(r)$, the interaction energy simplifies to

$$\int d^3r_1 d^3r_2 \rho(\vec{r}_1) v(|\vec{r}_1 - \vec{r}_2|) \rho(\vec{r}_2) = \mathbf{t}^T \mathbf{V}^{-1} \mathbf{t}. \tag{2.48}$$

Despite this choice being seemly convenient, because there is the need to calculate only one set of integrals, it inhibits the development of low-scaling methods in case of the Coulomb operator or might cause numerical instabilities if the matrix $\mathbf{V}$ is ill-conditioned like in the case of a long-range Coulomb operator.[41, 85]

2.5 Density Functional Theory

The high computational costs of WF correlation methods are related to the calculation of ERIs and the effort to parametrise the WF in an affordable manner. A drawback of the WF is its non-measurability whereas the electron density,

$$n(\vec{r}) = N \sum_{\sigma} \int dx_2 \dots dx_N \left|\Psi((\vec{r}, \sigma), x_2, \dots, x_N)\right|^2, \tag{2.49}$$

is measurable. Furthermore, the electron density is a function of three spatial coordinates only whereas the many-particle WF is a function of $3N$ spatial coordinates. But both functions depend only on the local external potential $v^{ext}(\vec{r})$ which is a function of three spatial coordinates only. Thus, why do we actually have to employ a complex object of much more spatial variables like the WF if the electron density has the same complexity like the determining external potential and both actually carry information about the molecule?
This question lead to the development of density functional theory (DFT) with the electron density as its central quantity. Hohenberg and Kohn proved the one-to-one correspondence between the electronic ground-state density and the external potential.[6] Because of the SE, there is a mapping of the external potentials to the ground-state WFs. Due to equation (2.49), there is a mapping of the ground-state WFs to the ground state densities and as such a mapping of external potentials to the ground state densities. The Hohenberg-Kohn theorem now states the mapping of the external with the important result that there is a map of the ground state densities to the ground-state WF. Because we obtain all relevant physical quantities by applying the corresponding operator to the ground state WF, all relevant physical quantities of the ground state of a system are functionals of the ground

state density. This has to be especially true for the ground state energy. Hohenberg and Kohn further established the variational principle of the energy DF with respect to the space of trial densities.
Albeit its framework provides an exact theory, approximations are required for actual calculations with DFT, most commonly based on the Kohn-Sham (KS) approach.[5] Within this approach, we consider a reference system of non-interacting electrons with the same energy and the same density. Due to the Hohenberg-Kohn theorem, there is a corresponding external potential, the Kohn-Sham potential $v^{\mathrm{KS}}(\vec{r})$. The DFT energy is then written as

$$E[n] = T^0[n] + U[n] + E^{\mathrm{H}}[n] + E^{\mathrm{XC}}[n] \tag{2.50}$$

with the kinetic energy of non-interacting electrons

$$T^0[n] = -\frac{1}{2}\sum_i \int d^3r\, \psi_i^*(\vec{r})\Delta\psi_i(\vec{r}), \tag{2.51}$$

the energy interaction energy of the electrons $U[n]$ with the external potential from above, the Hartree energy $E^{\mathrm{H}}[n] = \frac{1}{2}\int d^3r d^3r' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|}$ describing the classical self-interaction energy of the electron density and the yet unknown eXchange-Correlation (XC) energy DF $E^{\mathrm{XC}}[n]$. The XC energy DF includes the quantum mechanical contributions to the electron-electron interaction energy and a correction to the non-interacting kinetic energy. Because of the picture of non-interacting electrons, the density is given in terms of single-electron WFs (orbitals) $\psi_i(\vec{r})$ employed as auxiliary quantities to assemble the electronic ground state density $n(\vec{r}) = \sum_i |\psi_i(\vec{r})|$. The KS potential is the functional derivative of $U[n] + E^{\mathrm{H}}[n] + E^{\mathrm{XC}}[n]$ with respect to the density

$$v^{\mathrm{KS}}(\vec{r}) = v^{\mathrm{ext}}(\vec{r}) + v^{\mathrm{H}}(\vec{r}) + v^{\mathrm{XC}}(\vec{r}), \tag{2.52}$$

$$v^{\mathrm{H}}(\vec{r}) = \int d^3r' \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}, \tag{2.53}$$

$$v^{\mathrm{XC}}(\vec{r}) = \frac{\delta E^{\mathrm{XC}}}{\delta n(\vec{r})}. \tag{2.54}$$

The still unknown XC DF is the starting point for approximations within DFT, providing DF Approximations (DFAs). In the literature, the distinction between DFAs and density functionals (DFs) is often ignored and DFAs are just referred to as DFs implicitly considering the approximate character of the actually employed

DF. In the following, we will use the same terminology and refer to DFAs as DFs. Most DFs are local XC DFs

$$E^{\mathrm{XC}}[n] = \int d^3r\, e^{\mathrm{XC}}[n](\vec{r})n(\vec{r}) \tag{2.55}$$

with the XC energy per electron $e^{XC}[n](\vec{r})$, a function of different directly available quantities in terms of the electron density.
Depending on the employed quantities, there is a hierarchy of DFAs with supposedly increasing accuracy but also higher computational costs to the "heaven of chemical accuracy" (Jacob's ladder analogy).[86] On the first rung, the Local Density Approximation (LDA), e^{XC} is a function of the local value of the electron density only.[87–89] On the second rung (Generalized Gradient Approximation, GGA), there is an additional dependence on the local density gradient.[90, 91] meta-GGA DFs augment with the kinetic energy density $\tau(\vec{r}) = -\frac{1}{2}\sum_i \psi_i(\vec{r})\Delta\psi_i(\vec{r})$ and/or the Laplacian of the density.[92, 93] Fourth-rung DFs add an explicit dependence with respect to the occupied orbitals by mixing DFT exchange with HF exchange, fifth-rung DFs even add a dependence with respect to the unoccupied orbitals.[66, 94–97]
In analogy to the WFT theory, the XC DF is commonly split into an exchange DF and a correlation DF. Most DFAs are local DFs in the sense that they are expressed as an integral of a given function depending on several auxiliary quantities of the electron density. The terminology of the Jacob's ladder analogy is applied accordingly.

2.6 Random-Phase Approximation

In the last years, the Random-Phase approximation (RPA) became popular among quantum chemists.[98–100] It is based on the adiabatic connection fluctuation-dissipation theorem stating that the correlation energy E^C of a system is given by

$$E^{\mathrm{C}} = -\frac{1}{2}\mathrm{Im}\int_0^1 d\alpha \int_0^\infty \frac{d\omega}{\pi}\int dx_1 dx_2 \frac{\chi_\alpha(\omega, x_1, x_2) - \chi_0(\omega, x_1, x_2)}{|\vec{r}_1 - \vec{r}_2|} \tag{2.56}$$

with the coupling constant of the electron-electron interaction α and the frequency-dependent density-density response function $\chi_\alpha(\omega, x_1, x_2)$. $\alpha = 0$ is equivalent to the KS ground state, $\alpha = 1$ to the fully interacting system. Both density-density

response functions are related to each other via a Dyson-like equation

$$\chi_\alpha(\omega,x_1,x_2) = \chi_0(\omega,x_1,x_2) + \int dx_1' dx_2' \chi_\alpha(\omega,x_1,x_1') f_\alpha^{\mathrm{HXC}}(\omega,x_1',x_2') \chi_0(\omega,x_2',x_2) \quad (2.57)$$

with the unknown Hartree, exchange, correlation kernel $f_\alpha^{HXC}(\omega,x_1,x_2)$.[101] RPA neglects exchange and correlation effects and sets $f_\alpha^{HXC}(\omega,x_1,x_2) = \frac{\alpha}{|\vec{r}_1-\vec{r}_2|}$. Furche proved that the RPA correlation energy is given by the expression[102]

$$E^{\mathrm{C,RPA}} = \frac{1}{2}\mathrm{Tr}(\Omega - \mathbf{A}) \quad (2.58)$$

with Ω^2 being the eigenvalues of the generalized Eigenvalue problem

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B} & -\mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y} & \mathbf{X} \end{pmatrix} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y} & \mathbf{X} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \Omega^2, \quad (2.59)$$

$$\mathbf{X}\mathbf{X}^T - \mathbf{Y}\mathbf{Y}^T = 1, \quad (2.60)$$

$$A_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + (ia|jb) - (ib|ja), \quad (2.61)$$

$$B_{ia,jb} = (ia|jb) - (ib|ja). \quad (2.62)$$

This method has a theoretical scaling of $\mathscr{O}\left(N^6\right)$ with N being a measure of system size.
If we neglect the integral $(ij|ab)$, the eigenvalue problem simplifies to a smaller sized Hermitian eigenvalue problem

$$E^{\mathrm{C,RPA}} = \frac{1}{2}\mathrm{Tr}\left(\mathbf{M}^{1/2} - \mathbf{A}\right), \quad (2.63)$$

$$\mathbf{M}\mathbf{Z} = \mathbf{Z}\Omega^2, \quad (2.64)$$

$$\mathbf{Z}\mathbf{Z}^T = 1, \quad (2.65)$$

$$\mathbf{M} = (\mathbf{A}-\mathbf{B})^{1/2}(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B})^{1/2} \quad (2.66)$$

reducing the prefactor significantly, iterative approaches reduce the scaling to $\mathscr{O}\left(N^5\right)$.[102] This simplified RPA method is called direct RPA (dRPA). In the following, we will ignore this distinction and will refer to direct RPA simply as RPA whereas RPA including the exchange-like integrals is referred to as RPA with exchange.
Rewriting the matrix root as an integral and introducing the RI approximation

employing a Coulomb metric (see chapter 2.4), the final result is[103]

$$E^{\mathrm{C,RPA}} = \int_0^\infty \frac{d\omega}{2\pi} \mathrm{Tr}\left(\ln\left(1+\mathbf{Q}(\omega)\right) - \mathbf{Q}(\omega)\right), \tag{2.67}$$

$$\mathbf{Q}(\omega) = 2\mathbf{S}^T\mathbf{G}(\omega)\mathbf{S}, \tag{2.68}$$

$$S_{iaP} = \sum_Q (ia|Q)(Q|P)^{-1/2}, \tag{2.69}$$

$$G_{ia,jb} = \delta_{ij}\delta_{ab}\frac{\varepsilon_a - \varepsilon_i}{(\varepsilon_a - \varepsilon_i)^2 + \omega^2} \tag{2.70}$$

Despite of the efficiency of RPA compared to RPA with exchange, RPA suffers from self-correlation effects implying non-zero correlation energies for one-electron systems in contrast to common WF methods.

2.7 Double-Hybrid Density Functionals

Double-hybrid DFs (DHDFs) are given by the XC DF[67]

$$\begin{aligned} E^{\mathrm{XC}}[n] = {} & a_{\mathrm{X,DFT}}E^{\mathrm{X,DFT}}[n] + a_{\mathrm{X,HF}}E^{\mathrm{X,HF}}[n] \\ & + a_{\mathrm{C,DFT}}E^{\mathrm{C,DFT}}[n] + a_{\mathrm{C,WFT}}E^{\mathrm{C,WFT}}[n] \end{aligned} \tag{2.71}$$

with the DFT-like exchange DF $E^{\mathrm{X,DFT}}[n]$, the HF exchange DF $E^{\mathrm{X,HF}}[n]$, the DFT-like correlation DF $E^{\mathrm{C,DFT}}[n]$ and the WFT-based correlation DF $E^{\mathrm{C,WFT}}[n]$ with the corresponding scaling factors $a_{\mathrm{X,DFT}}$, $a_{\mathrm{X,HF}}$, $a_{\mathrm{C,DFT}}$ and $a_{\mathrm{X,WFT}}$. Note that the HF exchange DF and the WFT correlation DF depend explicitly on the orbitals which are both implicitly given as DFs via their dependence on the external potential.

Before the first occurrence of the terms double-hybrid DF and doubly-hybrid DF in literature by Zhao *et al.* and more prominently by Grimme, there had already been similar approaches mixing DFT and WFT based on the Coupled Cluster or the Configuration Interaction method, sometimes referred to as "Multi-Configurational DFT".[65, 66, 104–106]

Although the HF exchange functional and the WFT-based correlation methods are commonly given in terms of orbitals, they are implicit DFs due to their dependence on the external potential. Nevertheless, the KS potential being the derivative with respect to the density is usually taken as the derivative with respect to the orbitals implying that the HF exchange potential in DFT is a non-local DF in

analogy to HF theory. Within the method of Optimized Effective Potentials, the derivatives of the HF and the WFT energies is obtained as a local potential.[107–109] Due to the higher computational costs, this approach is rarely applied in applications.
Because the evaluation of the WFT correlation functional is commonly the highest scaling step of DHDF calculations, it is considered to be a post-SCF correction, thus the Kohn-Sham potential is the derivative of the first three terms in equation (2.71) only. Depending on the underlying hybrid DF, there is a large variety of flavours of DHDFs. DHDFs are classified according to

- the rung of the underlying DF (LDA, GGA, meta-GGA),
- the underlying WFT correlation method (MP2, SCS-MP2, SOS-MP2, RPA etc.),
- the operator of the HF kernel (Coulomb, range-separated),
- the operator of the WFT functional (Coulomb, long-range Coulomb),
- the way the parameters of the DF are determined (fitted to experimental data or object to theoretical constraints),
- the inclusion of dispersion correction (kind of dispersion correction, parameters are optimized together with other DHDF parameters),
- DF for orbitals, DF for energy (different or the same).

Table 2.1: Overview of the different flavors of MP2-, SCS-MP2, SOS-MP2- and RPA-based DHDFs. "NRS" refers to non-range-separated DHDFS, "RS" to range-separated DHDFs.

		MP2	SCS-MP2	SOS-MP2	RPA
NRS	LDA		DSD-SVWN[110]	DOD-SVWN[111]	
	GGA	MC3MPW[104] B2PLYP[66] mPW2-PLYP[112] B2KPLYP[113] PBE0-DH[114]	DSD-BLYP[110]	SOS0-PBE0-2[115] DOD-BLYP[111]	RPA++[116]
	meta-GGA	MC3BB1[104] DS1DH-TPSS[117]		PTPSS[24] PWPB95[24]	PWRB95[118]
RS	LDA	RSHLDA+MP2[65]	RSHLDA+SCS[70]		RSHLDA+RPA[119, 120]
	GGA	RSX-QIDH[121] RSHPBE+MP2[65] B2-P3LYP[122]	ωB97X-2[123] RSHPBE+SCS[70]	B2-OS3LYP[122]	RSHPBE+RPA[119, 120]
	meta-GGA		ωB97M-2[124]		

Examples for some groups are provided in table 2.1. In principle, one could employ other WFT correlation methods than RPA or MP2 like Quadratic Configuration Interaction or Coupled Cluster but the higher computational costs of these approaches reduce their usability without actually improving the results. The high accuracy of DHDFs is related to several drawbacks of WFT and common DFAs. Most DFAs are neither self-interaction free nor self-correlation free which results in an incorrect description of charged species, charge transfer processes and weakly interacting systems.[88, 125–127] Inclusion of HF exchange reduces the self-interaction error, WFT correlation reduces the self-correlation error. The usually larger amount of HF exchange compared to most hybrid DFs reduces the self-interaction error significantly. The improvement over the underlying WFT method is related to the missing higher excitations required for a reasonable description of short-range interactions. These interactions are again well-described in DFAs.
A further improvement is given for the description of long-range interactions. With perturbation theory, one obtains a R^{-6}-dependence for the long-range interaction energy of two uncharged atoms of distance R whereas common DFAs predict an exponential decay. This behavior is crucial for the description of interactions of uncharged molecules like in molecular crystals, molecular liquids but also for molecule-surface interactions and the interaction of biomolecules. For common DFAs, there are several empirical and non-empirical corrections available.[26–29, 31–33] DHDFs (partially) describe these interactions with MP2 or RPA, thus they represent a non-empirical correction to common DFAs. Most DHDFs do not employ full MP2 or RPA such that they still require a partial correction for dispersion corrections.
MP2 and RPA were derived assuming HF orbitals. These are not employed for DH calculations. Thus, the derivation for these methods is invalid and actually requires different equations. We will discuss this issue for the common MP2-based DHDFs.
Görling and Levy derived the leading term of the correlation energy DF[128, 129]

$$E^{\mathrm{C}}[n] \approx \sum_k \frac{\left|\left\langle \Psi_0 \left| \hat{V}_{\mathrm{ee}} - \sum_{i=1}^{N} \left(v^{\mathrm{H}}[n](\vec{r}_i) + v^{\mathrm{X}}[n](\vec{r}_i) \right) \right| \Psi_k \right\rangle\right|^2}{E_0 - E_k}. \tag{2.72}$$

If we compare this expression with the energy expression from MP2 in equation (2.36), there is a difference in the exchange potential. In MP2, the exchange potential is non-local whereas in DFT, the exchange potential is local. Furthermore,

the excited reference states need not to be the same. Thus, Görling-Levy perturbation theory and MP2 theory do not provide the same results but MP2 is rather considered to be an approximation to Görling-Levy perturbation theory.[66]

2.8 CP2K and the GPW approach

CP2K is an electronic-structure code tackling condensed matters on large super computers.[130] Its efficiency relies on the Quickstep algorithm and sparse matrix multiplications employing the DBCSR library.[131–133] The basis of the Quickstep algorithm is the GPW method, a dual-basis approach for orbitals and densities.[134] While orbitals are expanded in Gaussian basis sets enabling an efficient linear-scaling integration of integrals, densities are represented on a regular grid equivalent to a plane wave basis set. The plane wave basis allows a linear-scaling determination of the Hartree potential in reciprocal space being backtransformed to real space by a Fast Fourier Transform (FFT). The plane wave basis requires the use of pseudopotentials of which CP2K employs the Goedecker-Teter-Hutter type.[135] All-electron calculations are not possible within the GPW approach but are enabled with the GAPW method.[136]

Chapter 3

Implementation of Double-Hybrid Functionals with Range-Separated Correlation Functional

3.1 Introduction

Double-hybrid density functionals (DHDFs) belong to the fifth rung of the Jacob's ladder of density functional theory (DFT).[86] They augment common local density functionals with Hartree-Fock (HF) exchange and wave function correlation (WFC) providing the highest accuracy possible within the framework of density functional theory (DFT).[25]
Albeit their excellent performance across a large variety of systems, the underlying WFC method does not only increase the computational costs but also increases the errors with respect to the basis set incompleteness and basis set superposition error.[64, 70, 71] The required computing resources limit the application of DHDFs to molecular systems or require super computing hardware for calculations in condensed phase, a possibility to treat the large basis set dependence is given by explicitly correlated methods.[60–63]
It was proved that the introduction of range-separated correlation functionals turns the slow cubically scaling convergence with respect to the size of the basis set of conventional DHDFs into the exponential decay of standard non-DHDFs.[64] Exchange and correlation functional are separately divided into short-range and long-range contributions. Then, the short-range contributions are described by a local density functional and the respective long-range contributions by a wave

function method. Therefore, different correlated methods were studied in the literature.[65, 105, 119, 137–139]
The whole methodology correctly predicts a $\frac{1}{R^6}$-decay for dispersion interactions. Their application to very weakly interacting systems like rare gas dimers and rare-gas-alkaline-metal dimers and molecular crystals proves the accuracy of this scheme.[65, 105, 106] The required short-range correlation energy functionals or long-range electron repulsion integrals for correlated methods are missing in many electronic structure codes limiting a wide availability of these functionals.
In the following, we describe the implementation of long-range Coulomb integrals into the GPW approach for RI-MP2 and RI-RPA and long-range HF exchange for pseudopotential (PP) calculations into the CP2K package. Benchmarking both ingredients allows to discuss benefits and drawbacks of the implementation of DH functionals with range-separated DH functionals within CP2K.

3.2 Theoretical Background

In order to understand the basis set convergence of correlated methods, we will examine a class of simple toy systems, the Helium-like atoms. These are the most simple two-electron systems and in our discussion, we follow the derivation of Schwartz.[58]
The Schrödinger equation for a Helium-like atom with nuclear charge Z is

$$-\frac{\Delta_1 \psi(\vec{r}_1,\vec{r}_2)}{2} - \frac{\Delta_2 \psi(\vec{r}_1,\vec{r}_2)}{2} - \left(\frac{Z}{r_1}+\frac{Z}{r_2}-\frac{1}{r_{12}}\right)\psi(\vec{r}_1,\vec{r}_2) = E\,\psi(\vec{r}_1,\vec{r}_2) \quad (3.1)$$

with $\vec{r}_i$ being the coordinate of electron i, Δ_i the corresponding Laplace operator, and $r_i = |\vec{r}_i|$ and $r_{12} = |\vec{r}_1 - \vec{r}_2|$ denote the distance vectors.
Dividing by Z^2 and scaling the vectors according to $\vec{r}_i = \frac{1}{Z}\vec{r}_i\,'$ and the energy according to $E = Z^2\tilde{E}$, we obtain

$$-\frac{\tilde{\Delta}_1 \tilde{\psi}(\vec{r}_1\,',\vec{r}_2\,')}{2} - \frac{\tilde{\Delta}_2 \tilde{\psi}(\vec{r}_1\,',\vec{r}_2\,')}{2} - \left(\frac{1}{r_1'}+\frac{1}{r_2'}-\frac{1}{Z}\frac{1}{r_{12}'}\right)\tilde{\psi}(\vec{r}_1\,',\vec{r}_2\,') = \tilde{E}\,\tilde{\psi}(\vec{r}_1\,',\vec{r}_2\,') \quad (3.2)$$

To simplify the notation, we will ignore the tildes and primes and introduce the one-electron hamiltonian operators

$$h_i = -\frac{\Delta_i}{2} - \frac{1}{r_i} \tag{3.3}$$

and the interaction hamiltonian operator

$$h_{12} = \frac{1}{r_{12}}. \tag{3.4}$$

We observe that both operators are hermitian.
Then, we apply the perturbation ansatz $E = E_0 + \frac{1}{Z}E_1 + \frac{1}{Z^2}E_2 + ...$ and $\psi(\vec{r}_1, \vec{r}_2) = \psi_0 + \frac{1}{Z}\psi_1 + ...$ with $h_1 + h_2$ being the unperturbed hamiltonian and h_{12} as perturbation with perturbation parameter $\frac{1}{Z}$. Applying theory, we obtain the equation of the unperturbed system (system of noninteracting electrons)

$$(h_1 + h_2)\psi_0 = E_0\psi_0 \tag{3.5}$$

and the first-order equation

$$(h_1 + h_2)\psi_1 + h_{12}\psi_0 = E_0\psi_1 + E_1\psi_0. \tag{3.6}$$

The zeroth-order equation has the known solution

$$\psi_0(\vec{r}_1, \vec{r}_2) = \phi_{1s}(\vec{r}_1)\phi_{1s}(\vec{r}_2), \tag{3.7}$$

$$E_0 = -1, \tag{3.8}$$

with the 1s-orbital $\phi_{1s}(\vec{r})$.
Multiplicating the first-order equation with ψ_0 from the left and integrating, we remove the terms with ψ_1 and we obtain the first-order energy correction $E_1 = \frac{5}{8}$. Then, the equation for the first-order correction of the wave function (WF) reads

$$(h_1 + h_2 + 1)\psi_1 = \left(\frac{5}{8} - h_{12}\right)\psi_0. \tag{3.9}$$

For small interelectronic distances, the equation is approximately solved by

$$\psi_1 = \frac{r_{12}}{2}\psi_0. \tag{3.10}$$

From the known partial wave expansion (PWE) of the Coulomb potential, we derive the PWE of the solution

$$\psi_1 = \frac{\psi_0}{2} \sum_l P_l(\cos\theta) \frac{r_<^l}{r_>^{l+1}} \left(\frac{r_<^2}{2l+3} - \frac{r_>^2}{2l-1} \right), \tag{3.11}$$

which becomes in the limit of small differences δ in the radii and large l

$$\psi_1 = -\frac{r\psi_0}{2} \sum_l P_l(\cos\theta) \frac{1}{l^2} e^{-\delta l}(1+\delta l), \tag{3.12}$$

With the variational principle for the second order energy correction, we obtain an upper bound for the second order correction for each angular momentum quantum number $-\frac{45}{256}\frac{1}{l^4}$. Summing all terms from a given l to infinity, we obtain an upper bound for the error in the second order energy to be $-\frac{15}{256}\frac{1}{l^3}$. To illustrate the decay, a decrease of the error by one order of magnitude requires at least twice the number of considered angular momentum quantum numbers. The size of the usually employed correlation basis sets increases cubically with the highest considered angular momentum quantum number. For a given system, the computational cost of MP2 increases quartically with respect to the basis set size. In total, the error decays with the fourth root of the computational timing implying that an improvement of one order of magnitude for the error in the total energy requires 10,000 times higher computational costs. Fortunately, the decay of the error is very smooth which allows the application of an extrapolation scheme to improve upon already available results. The general behavior was mathematically rigorously proved.[140, 141]
For the long-range potential $\frac{\mathrm{erf}(\omega r_{12})}{r_{12}}$, Gori-Giorgi *et al.* derived an expression in the limit of small inter-electronic distances given by[142]

$$\psi_1 = r_{12} p_1(\omega r_{12}) \psi_0 \tag{3.13}$$

$$p_1(y) = \frac{\exp(-y^2) - 2}{2\sqrt{\pi} y} + \left(\frac{1}{2} + \frac{1}{4y^2} \right) \mathrm{erf}(y). \tag{3.14}$$

With a fixed value of ω and sufficiently small r_{12}, we find

$$\psi_1 = \frac{\omega}{3\sqrt{\pi}} r_{12}^2 \psi_0 \tag{3.15}$$

whereas the condition by Schwarz (3.10) is recovered in the limit of infinitely large ω. In spite of analytical expressions of the PWE of $p_1(\omega r_{12})$ are not available, Franck *et al.* gave numerical proof for the fast decay of the PWE and consequently also a fast decay of the error with respect to the basis set size.[64]
In order to exploit the improved convergence, Savin *et al.* constructed a range-separated DH functional (rsDH functional) given the energy functional

$$E^{\mathrm{rsDH}} = T^0 + U + E^{\mathrm{H}} + E^{\mathrm{X,srDFT}} + E^{\mathrm{X,lrHF}} + E^{\mathrm{C,srDFT}} + E^{\mathrm{C,lrWFT}} \tag{3.16}$$

with the kinetic energy of non-interacting electrons T^0, the electron-nucleus interaction energy U, the Hartree-energy E^{H}, the short-range exchange DFT functional $E^{\mathrm{X,srDFT}}$, the long-range HF functional $E^{\mathrm{X,lrHF}}$, the short-range correlation DFT functional $E^{\mathrm{C,srDFT}}$ and the long-range WF correlation (WFC) energy $E^{\mathrm{C,lrWFT}}$.
Ignoring the correlation energy terms, the theory describes a range-separated hybrid functional without correlation. Most range-separated hybrid functionals like HSE06 or ωB97M-V describe the correlation energy with a non-range-separated density functional whereas range-separated DHDFs employ a mixture of density functional correlation and WFC with the Coulomb operator.[94, 143] In the theory discussed here, the correlation energy term is split into long-range and short-range contributions. The long-range correlation energy term is described by a WFC method whereas the short-range correlation energy is described by a density functional. It exploits the fact that common density functionals accurately describe short-range interactions due to their local character whereas WFC methods easily describe long-range interactions providing the correct R^{-6}-decay of interaction energies of uncharged systems thus describing each part of the interaction with the most suitable method.
Because the exchange-energy depends linearly on the interaction potential, the exchange-energy expression for the given potential is uniquely defined given a separation of the Coulomb potential into long-range and short-range contributions. This is not true for the correlation energy because of its non-linear dependence of the interaction potential. Thus, one defines a long-range correlation energy functional first and defines the short-range correlation functional as the compliment of the long-range correlation functional. Alternatively, one may provide a short-range correlation functional first and defines the long-range correlation functionals as its compliment.
Range-separated exchange energy functionals are widely available and are based on different Coulomb-based exchange functionals and a variety of operators and employed by several functionals.[94, 143–147] The same availability does not

hold for range-separated correlation energy functionals. Until today, most of these functionals were developed by the groups of Savin and Gori-Giorgi. Until now, only one of them is available in the LibXC library interfaced by several quantum chemistry codes like CP2K.[130, 148–151]

3.3 Implementation of Integrals within the GPW approach

3.3.1 Motivation and Theoretical Background

The GPW approach within the CP2K software package is a scheme to efficiently calculate matrix elements of Gaussians.[130, 134] A common problem of quantum-chemistry codes is the calculation of the Hartree potential related to a given density. In the GPW approach, densities are represented in two different manners: as a sum of products of two Gaussian basis functions or in a plane-wave basis set. This approach allows to calculate the potential in reciprocal space efficiently from the density in reciprocal space, back-transform the potential into real space, and calculate the required matrix-elements. The transition between both spaces is given by Fast Fourier Transforms (FFTs) for which efficient libraries exist.[152–158]
In order to implement the long-range Coulomb operator $\frac{\mathrm{erf}(\omega r)}{r}$ into the GPW approach, we will rely on the Fourier transform of the potential given by

$$\int d^3r \frac{\mathrm{erf}(\omega r)}{r} e^{-\vec{k}\cdot\vec{r}} = \frac{1}{(2\pi)^3}\frac{4\pi}{r^2}\exp\left(-\left(\frac{k}{2\omega}\right)^2\right) \tag{3.17}$$

which differs just by the exponential factor from the Fourier transform of the Coulomb potential

$$\int d^3r \frac{1}{r} e^{-\vec{k}\cdot\vec{r}} = \frac{1}{(2\pi)^3}\frac{4\pi}{r^2} \tag{3.18}$$

providing an easy way to calculate the matrix elements of the long-range Coulomb operator using the machinery for the Coulomb potential.
Let $\mathbf{A}$ be the real hermitian matrix of size n representing the Coulomb operator in the RI basis set. Then, the required matrix inversion and matrix root of $\mathbf{A}$ is

usually determined by a Cholesky decomposition[159]

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T \tag{3.19}$$

with a lower triangular matrix $\mathbf{L}$ and its final inversion, both efficiently implemented in the libraries Lapack and ScaLapack.[160, 161] Although $\mathbf{A}$ ought to be positive definite with the long-range Coulomb operator because of $\text{erf}(x) \geq 0$, it is numerically only positive semi-definite and the Cholesky decomposition becomes numerically unstable. Instead, we employ its eigenvalue decomposition

$$\mathbf{A} = \mathbf{C}\Lambda\mathbf{C}^T \tag{3.20}$$

with an orthogonal matrix $\mathbf{C}$ containing the eigenvectors and a diagonal matrix $\Lambda = \text{diag}(\lambda_1, ...\lambda_n)$ and the eigenvalues λ_i.[162, pp.3-10, 24–30, 265] Because $\mathbf{A}$ is positive semi-definite, we have $\lambda_i \geq 0$. Further, we assume the sequence $\lambda_1, ...\lambda_n$ to be decreasing. Eigenvalue decompositions can be calculated efficiently with libraries like ELPA.[163, 164] With these ingredients, we define an alternative symmetric decomposition

$$\mathbf{A} = \mathbf{B}\mathbf{B}^T \tag{3.21}$$

$$\mathbf{B} = \mathbf{C}\mathbf{D}^{1/2} \tag{3.22}$$

$$\mathbf{D}^{1/2} = \text{diag}(\lambda_1^{1/2}, ..., \lambda_n^{1/2}) \tag{3.23}$$

and its Moore-Penrose-pseudoinverse

$$\mathbf{A}^+ = \mathbf{B}^+\mathbf{B}^{+T} \tag{3.24}$$

$$\mathbf{B}^+ = \mathbf{C}\mathbf{D}^{-1/2} \tag{3.25}$$

where $\mathbf{D}^{-1/2}$ is the diagonal matrix with its i-th diagonal element being zero if the λ_i is non-positive and $\lambda_i^{-1/2}$ else. Note that the two matrices $\mathbf{L}$ and $\mathbf{B}$ are equivalent to each other, i.e. there is an orthogonal matrix $\mathbf{O}$ such that $\mathbf{L} = \mathbf{B}\mathbf{O}$.
Because $r = \text{rank}(\mathbf{A}) \leq n$, we can save memory and floating point operations in subsequent steps with the low-rank decomposition by removing the last $n - r$ eigenpairs.
With the eigenvalue-decomposition, we have to calculate a decomposition of the inverse of a numerically singular matrix. We can increase the numerical stability by employing a different metric (see chapter 2.4) as suggested by Mezei *et al.*[85]

3.3.2 Performance and Accuracy

We implemented the low-rank decomposition into CP2K. For all calculations, we considered 64 water molecules in a box at ambient density, correlation-consistent primary basis sets (PBSs) and suitable RI auxiliary basis sets (ABSs).[68] In any case, we employed PBE orbitals and considered two values of the range-separation parameter $\omega = 0.3, 0.5$. PBSs and ABSs were optimized for the use with the Coulomb operator. MP2 and RPA denote the MP2 and RPA methods using the Coulomb operator, lrMP2 and lrRPA denote the MP2 and RPA methods using the long-range Coulomb operator.

First, we compared our GPW implementation against the MME implementation by Patrick Seewald employing a Coulomb metric.[165–167] Our implementation shows an excellent agreement with the MME integral approach with errors of the order of 0.1 μE_h per water molecule. This error is comparable to the RI error with the Coulomb operator. But we found that the long-range Coulomb metric is numerically unstable with the MME method such that the MME method had to be used with the Coulomb metric. Furthermore, the GPW implementation is one order of magnitude faster than the MME implementation. Consequently, we recommend the GPW method for long-range Coulomb integrals because of its better stability and performance.

We investigated the convergence of the lrMP2 and the lrRPA energy with respect to the size of the primary basis set. In figure 3.1, we report the errors relative to the energies extrapolated to the complete basis set limit of the total energies per molecule and the total energies per molecule relative to the single water molecule at the MP2-, RPA-, lrMP2- and lrRPA-levels of theory. The MP2 and RPA energies were extrapolated with a cubic function from the results of the triple- and quadruple zeta energies, the lrMP2 and lrRPA energies with an exponential fit using all three points as suggested by Franck *et al.*[64] Obviously, the general behavior and the magnitude of the errors is independent of the correlation method. Please note that the obvious exponential decay of the energy errors is a result of the exponential three-point extrapolation function. Nevertheless, the errors of the range-separated functionals decay fast with respect to the basis set size. We observe a basis set incompleteness error of 0.14 mE_h per water molecule with a quadruple zeta-quality basis set for lrMP2 whereas the error of MP2 is two orders of magnitude larger with an error of 24 mE_h. For the energy differences, the remaining errors with a quadruple zeta basis set are 0.26 mE_h (lrMP2) and 0.53 mE_h (MP2) and comparable errors for RPA and lrRPA. Because of the expected methodological error, a basis set of triple zeta quality provides already a

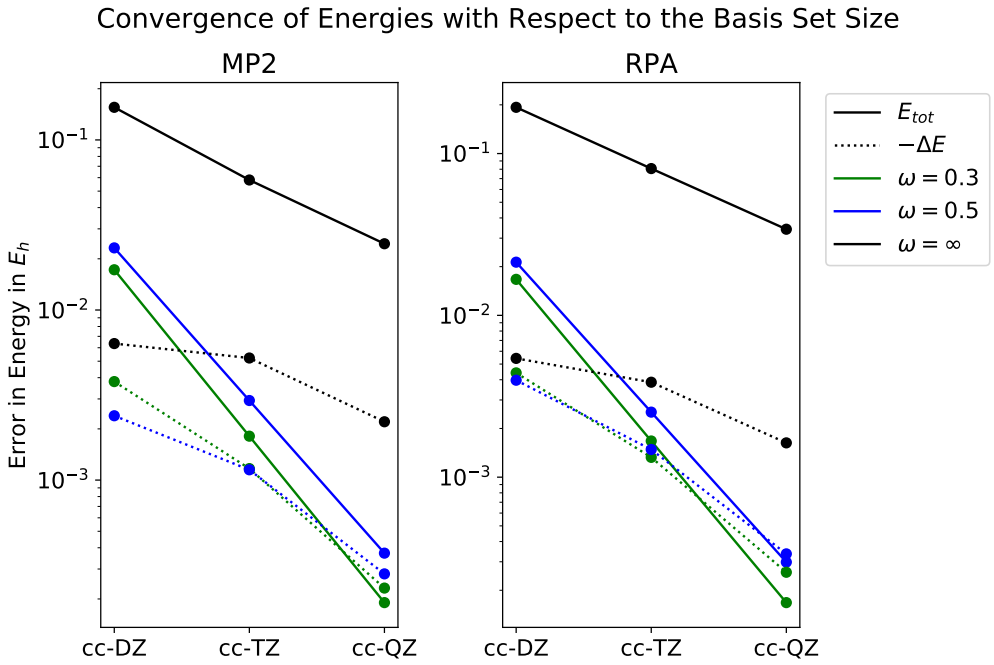

Figure 3.1: Convergence of total energies per molecule and the difference to the energy of a single molecule for MP2/lrMP2 (left panel) and RPA/lrRPA (right panel). The RI basis sets were optimized for the given primary basis set employing a Coulomb potential.

sufficient accuracy whereas the original MP2 and RPA methods should be employed with a quadruple zeta basis set.
In figure 3.2, we report the numerical ranks of matrix $\mathbf{A}$ for the RI basis sets corresponding to correlation-consistent basis sets of double to quadruple zeta quality for MP2 and lrMP2 and two different values of the range-separation parameter ω. An obvious result is the absence of a rank reduction for MP2 and lrMP2 with $\omega = 0.5$ whereas we observe a rank reduction of up to 20 % with lrMP2 depending on the employed value of ω. We can justify this observation by considering that the short-range contributions of the Coulomb potential allow a distinction of the different basis functions whereas the long-range field has the same behavior for all RI basis functions of the same angular quantum number. This effect is more pronounced for larger basis sets. With lower values of ω, we remove more short-range contributions reducing the numerical ranks of the two-center matrix

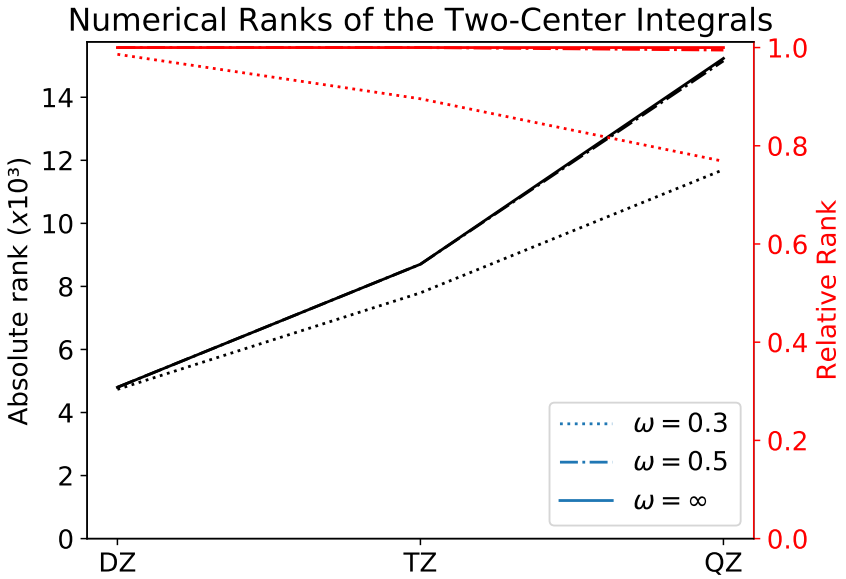

Figure 3.2: Numerical ranks of the two-center matrices of MP2 and lrMP2 for two different values of the range-separation parameter ω (left axis: absolute ranks (black lines), right axis: relative ranks (red lines)).

A further.
Because of the reduction in the RI basis set size, we examined the related reduction in the computational costs of the different available WFC methods in CP2K using the same setting as above. The Laplace-RI-SOS-MP2 method is expected to show a similar behavior as the RI-RPA method due to the same rate-determining step and is discarded. The cubically scaling RI-RPA implementation will not benefit from the rank reduction for the given test system because its performance is mostly determined by the low-scaling tensor contraction steps of the three-center integrals employing the unreduced RI basis set for systems with up to 1,000 atoms and is discarded, too.[41, 167]
In figure 3.3, we report the computational timings for the MP2, lrMP2, RPA and lrRPA calculations with different primary basis sets and suitable RI basis sets. Exploitation of long-range Coulomb integrals increases the costs of the integral calculation step in all cases because of the in general higher computational costs of the eigenvalue decomposition compared to the Cholesky decomposition. Furthermore, we observe a decrease in the computational timing of the higher-scaling

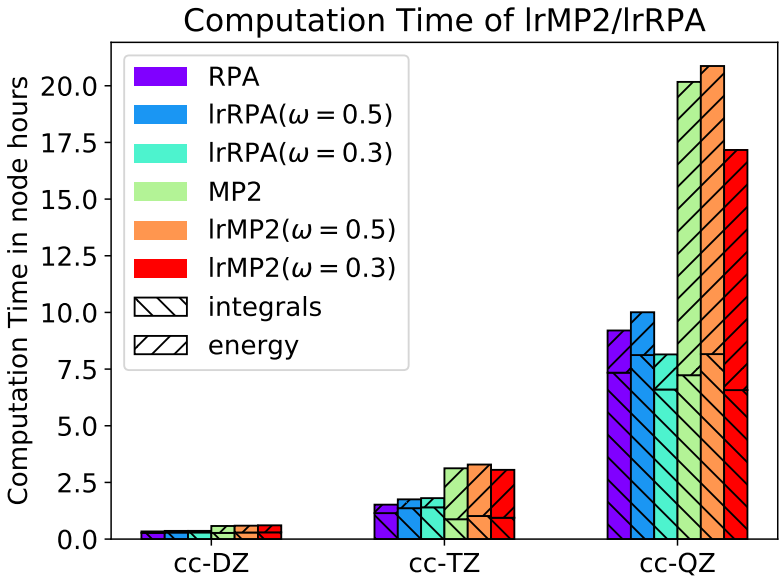

Figure 3.3: Computational costs RPA, MP2 and lrRPA and lrMP2 with two different values of the range-separation parameter ω. "integrals" and "energy" refers to the timing of the routines "mp2_ri_gpw_compute_in" and "mp2_ri_gpw_compute_en"/"rpa_ri_compute_en" (for MP2/RPA respectively). Calculations with PBSs of DZ and TZ quality were run on 16 nodes with 12 ranks per node, calculations with PBSs of QZ quality on 36 nodes with 4 ranks per node.

Table 3.1: Errors of total energies per molecule of the different metrics with lrMP2 and lrRPA, each with two different values of ω with respect to the long-range Coulomb metric in mE_h. The calculations were carried out with a primary basis set of triple zeta quality, suitable RI basis sets. The numerical integration of the RPA method employed a 8-point minimax grid.[168] For the short-range (sr.) Coulomb metric, we chose $\omega = 0.5$, for the truncated (trunc.) Coulomb metric $R_{\text{cut}} = 6$Å which is roughly half of the cell size.

Method	**Coulomb**	**Sr. Coulomb**	**Overlap**	**Trunc. Coulomb**
lrMP2($\omega = 0.3$)	0.008	0.019	0.024	0.003
lrMP2($\omega = 0.5$)	0.011	0.061	0.018	-0.085
lrRPA($\omega = 0.3$)	0.007	0.016	0.015	0.003
lrRPA($\omega = 0.5$)	0.015	0.050	0.002	0.067

energy calculation if we reduce ω. The total costs decrease by 10-15 % in the case $\omega = 0.3$. For larger systems, the quintically-scaling energy calculation step will dominate the costs whereas for smaller systems, the quartically-scaling determination and contraction of the three-center integrals dominate the costs. Thus, we will not only expect a reduction of the computational costs for larger systems in the case of MP2 but also in the case of RPA. The reduction of the computational costs of RPA is not too pronounced because of the in general smaller computational costs of the quartically-scaling RPA method compared with the quintically-scaling MP2 method.
As an alternative to the long-range Coulomb metric, we tested the different available RI metrics (see table 3.1). We find an excellent accuracy with all metrics of the order of 10 μE_h per molecule. Consequently, all metrics are recommended to use with long-range integrals.

3.4 Pseudopotentials

The pseudopotentials (PPs) are another important ingredient for the GPW approach in CP2K. Because the functionals described here contain 100 % of a local density functional and 100 % of non-local long-range HF, we expect that the required PPs are a mixture of the underlying full-range density functional and HF. Especially for transition-metals, a proper PP is crucial for the accuracy of the results. Thus, for a given value of the range-separation parameter ω, we have to optimize suitable PPs using the atom code of CP2K. Because the integral calcu-

lation routines within the atom code of CP2K are independent of those employed for the molecules and condensed-phase systems, we have to implement the long-range integrals into the HF atom code.
The task is to determine the integral

$$I^{n_2 l_2 m_2, n_2' l_2' m_2'}_{n_1 l_1 m_1, n_1' l_1' m_1'} = \int d^3 r \chi^*_{n_1 l_1 m_1}(\vec{r}_1) \chi_{n_1' l_1' m_1'}(\vec{r}_1) g(r_{12}) \chi^*_{n_2 l_2 m_2}(\vec{r}_2) \chi_{n_2' l_2' m_2'}(\vec{r}_2) \quad (3.26)$$

for a given interaction operator $g(r_{12})$ and basis functions $\chi_{nlm}(\vec{r}) = R_{nl}(r) Y_{lm}(\hat{r})$ with given real-valued radial basis functions $R_{nl}(r)$ and the unit-vector $\hat{r}$ in direction of $\vec{r}$. In the case of potentials representable as

$$g(r_{12}) = \sum_{NL} f_{NL}(r_1, r_2) \sum_{M=-L}^{L} Y^*_{LM}(\hat{r}_1) Y_{LM}(\hat{r}_2), \quad (3.27)$$

the above integral factorizes according to

$$I^{n_2 l_2 m_2, n_2' l_2' m_2'}_{n_1 l_1 m_1, n_1' l_1' m_1'} = \sum_{NLM} I^{R,NL}_{n_1 l_1, n_1' l_1', n_2 l_2, n_2' l_2'} I^{\Omega,LM}_{l_1 m_1, l_1' m_1', l_2 m_2, l_2' m_2'}, \quad (3.28)$$

$$I^{R,NL}_{n_1 l_1, n_1' l_1', n_2 l_2, n_2' l_2'} = \int dr_1 dr_2 r_1^2 r_2^2 R^*_{n_1 l_1}(r_1) R_{n_1' l_1'}(r_1) f_{NL}(r_1, r_2) R^*_{n_2 l_2}(r_2) R_{n_2 l_2}(r_2), \quad (3.29)$$

$$I^{\Omega,LM}_{l_1 m_1, l_1' m_1', l_2 m_2, l_2' m_2'} = \left(\tilde{I}^{\Omega,LM}_{l_1 m_1, l_1' m_1'} \right)^* \tilde{I}^{\Omega,LM}_{l_2 m_2, l_2' m_2'}, \quad (3.30)$$

$$\tilde{I}^{\Omega,LM}_{lm, l'm'} = \int d\theta d\phi Y_{lm}(\hat{r}) Y_{l'm'}(\hat{r}) Y^*_{LM}(\hat{r}) \quad (3.31)$$

considering that the unit vectors can be expressed by the spherical angles θ_i and ϕ_i. The angular integrals $\tilde{I}^{\Omega,LM}_{lm,l'm'}$ are related to the Clebsch-Gordan coefficients whereas the radial integrals $I^{R,NL}_{n_1 l_1, n_1' l_1', n_2 l_2, n_2' l_2'}$ depend on the involved potential.
For the Coulomb operator, we have the well known identity

$$\frac{1}{r_{12}} = 4\pi \sum_l \frac{r_<^l}{r_>^{l+1}} \sum_{m=-l}^{l} Y^*_{lm}(\hat{r}_1) Y_{lm}(\hat{r}_2) \quad (3.32)$$

with $r_< = \min(r_1, r_2), r_> = \max(r_1, r_2)$ and the spherical harmonics Y_{lm}. The derivation of an analogous relation for the long-range Coulomb operator is more involving. Limpanuparb *et al.* suggested the application of a Gauss-Hermite

quadrature to the integral

$$\frac{\mathrm{erf}(\omega r_{12})}{r_{12}} = \frac{2\omega}{\pi}\int_{-\infty}^{\infty} d\beta\, j_0(2\beta\omega r_{12})\exp\left(-\beta^2\right) \tag{3.33}$$

$$= 8\omega\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\int_{-\infty}^{\infty} d\beta\, j_l(2\beta\omega r_1)\, j_l(2\beta\omega r_2) Y_{lm}^*(\vec{r}_1) Y_{lm}(\vec{r}_2)\exp\left(-\beta^2\right) \tag{3.34}$$

with the spherical Bessel functions $j_l(x)$.[169] Application of a $2N+1$-point Gauss-Hermite quadrature results in

$$\frac{\mathrm{erf}(\omega r_{12})}{r_{12}} = 16\omega\sum_{n=0}^{N}\sum_{l=0}^{\infty}\sum_{m=-l}^{l} b_n j_l(2\beta_n\omega r_1)\, j_l(2\beta_n\omega r_2) Y_{lm}^*(\vec{r}_1) Y_{lm}(\vec{r}_2). \tag{3.35}$$

As noticed by Limpanuparb *et al.*, the number N' of significant quadrature points from an N-point quadrature rule grows like $\mathscr{O}\left(\sqrt{N}\right)$ because of the exponential decay of the quadrature weights b_n.[169]
For the long-range Coulomb potential and real-valued basis functions, the radial integral factorizes according to

$$I^{R,NL}_{n_1 l_1, n_1' l_1', n_2 l_2, n_2' l_2'} = \tilde{I}^{NL}_{n_1 l_1, n_1' l_1'} \tilde{I}^{NL}_{n_2 l_2, n_2' l_2'} \tag{3.36}$$

$$\tilde{I}^{NL}_{nl,n'l'} = 4\sqrt{\omega b_N}\int_0^{\infty} dr\, r^2\, j_L(2\beta_N\omega r) R_{nl}(r) R_{n'l'}(r). \tag{3.37}$$

If the radial basis functions are given on a grid with weights a_k and abscissa α_k, the integrals can be recasted as a matrix-vector product $\tilde{\mathbf{I}} = \mathbf{M}\boldsymbol{\rho}$ with the product density vector with elements $\rho^k_{nl,n'l'} = R_{nl}(\alpha_k) R_{n'l'}(\alpha_k)$ and the integration matrix with elements $M^{Nk}_{L\omega} = 2r\sqrt{\omega b_N a_k}\, j_L(2\beta_N\omega\alpha_k)$. The long-range potential on the grid is then $\mathbf{I} = \mathbf{M}^T\mathbf{M}\boldsymbol{\rho}$. This matrix-vector multiplication is cheaper to implement as a two-step process because of the in practice much smaller number of Hermite-grid points than primary grid points.
In figure 3.4, we report the total energies and numerical errors of the Helium atom with different numbers of integration points and range-separation parameters ω. The larger ω, the more we approach the Coulomb limit. Because we increase the interaction potential, the exchange energy has to decrease (increase in terms of magnitude) until the Coulomb limit is reached. The too high exchange energies of the larger values of ω show that the numerical integration scheme has not con-

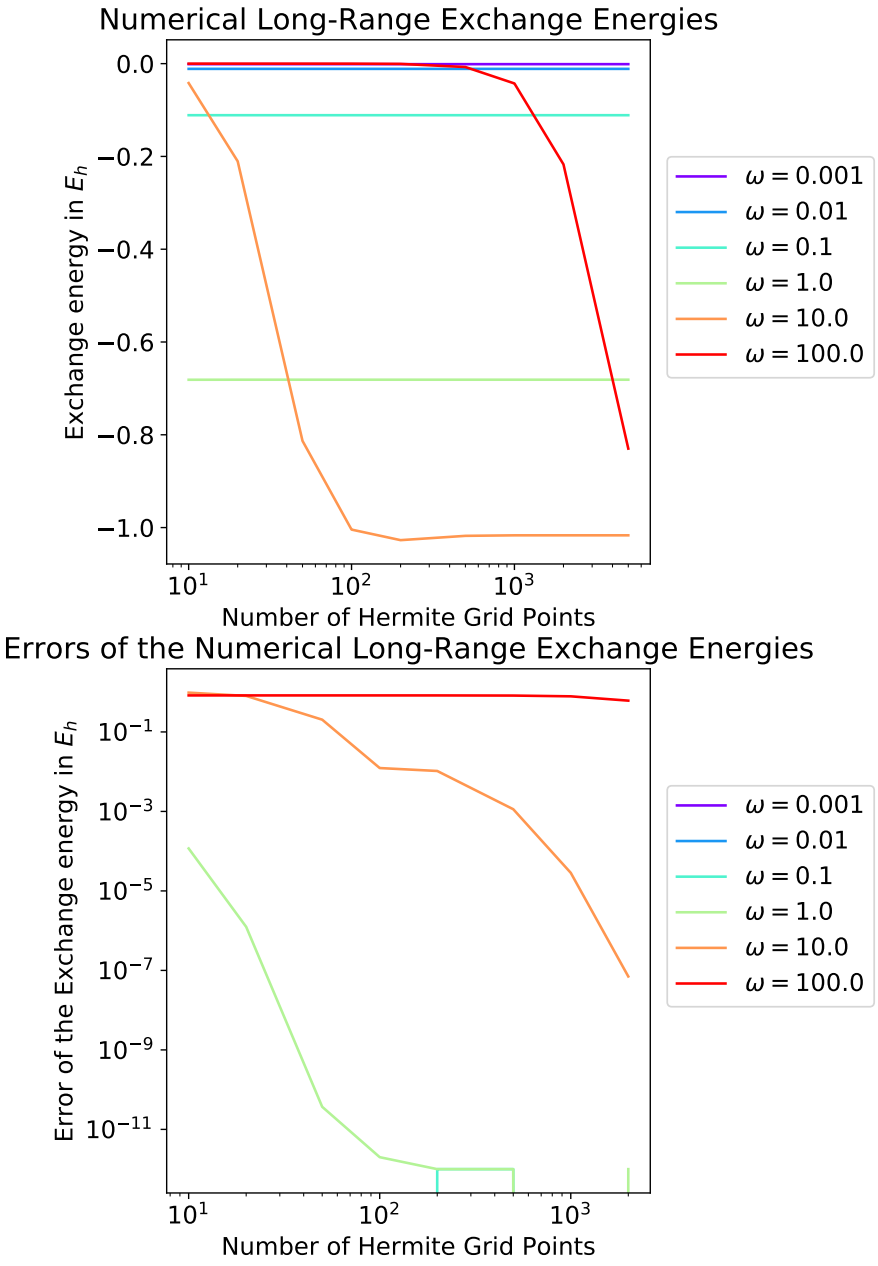

Figure 3.4: Convergence of the exchange energy (in E_h) of a Helium atom for different range-separation parameters ω and different numbers of grid points (upper panel: energies, lower panel: absolute errors with respect to the last grid point).

verged yet and a larger number of grid points is required. Using the actual errors, the energies of the smaller values of ω are converged with 10 grid points whereas we require at most 50 grid points when $\omega = 1$. When $\omega = 10$, we require at most 1,000 grid points for reasonable results. The bad convergence in case of high values of ω implies that our integration scheme is not suitable in these cases as pointed out by Limpanuparb *et al*.[169] Because we approach the Coulomb potential in the limit of large ω, we suggest to employ PPs optimized for the Coulomb potential instead.

3.5 Application to Rare Gas Dimers

In table 3.2, we report the dissociation energies, binding energies, vibrational frequencies and the basis set superposition errors of some rare gas dimers calculated with an LDA-based and MP2-based range-separated double-hybrid functional and with MP2 in comparison to the literature.[65, 105] All calculations were carried out with the RI-MP2 implementation in CP2K 8.1, large cutoffs (1,200 Ry for the DFT grids, 300 Ry for the MP2 grids) and low thresholds (EPS_DEFAULT of 1.0E-15 and EPS_SCHWARZ of 1.0E-15). PPs were optimized in accordance to reference [170] and are compiled in the appendix of this chapter. RI basis sets were automatically generated within CP2K (large basis sets). We employed the implementation of the short-range LDA-based functional by Gori-Giorgi *et al*.[148, 171] We extrapolated the results to the complete basis set limit according to Franck *et al*. in the case of the range-separated functional and according to Halkier *et al*. with the results from the triple- and quadruple-zeta basis sets in the case of MP2.[59, 64]
If we increase the basis set size, we observe a systematic reduction of the bond lengths and a systematic increase in the binding energies and the vibrational frequencies for both methods. This proves that our basis sets for each element systematically improve the results. If we compare our results with the quadruple-zeta basis sets and extrapolated values with the results obtained with a pentuple-zeta basis sets by Angyan *et al*. and our results with a triple-zeta basis sets with the comparable results by Goll *et al*., we find an excellent agreement between our results and the literature which proves the accuracy of our implementation and our basis sets in general.

Table 3.2: Errors in the binding energies of rare gas dimers obtained from calculations with different augmented correlation-consistent basis sets with MP2 and an MP2-based range-separated DHDF (RSH-MP2). aXZ refers to an augmented correlation-consistent basis set of X-zeta quality, d-a5Z refers to a doubly augmented pentuple-zeta basis set. Distances are given in Å, energies in μE_h, wave-numbers in cm^{-1}. [a] This work. [b] Reference Goll *et al.*[106] [c] Reference Angyan *et al.*[65] [d] Experimental values.[172]

		r(MP2)	E(MP2)	ν(MP2)	BSSE(MP2)	r(RSH-MP2)	E(RSH-MP2)	ν(RSH-MP2)	BSSE(RSH-MP2)
	aDZ[a]	3.239	12.7	19.6	132.0	3.269	16.1	21.1	6.4
	aTZ[a]	3.140	17.6	24.0	23.8	3.233	18.7	22.6	10.8
	aTZ[b]	-	-	-	-	3.218	15.4	18.7	-
He2	aQZ[a]	3.104	19.5	24.7	6.0	3.217	19.8	23.2	10.5
	d-a5Z[c]	3.069	21.9	-	5.6	3.203	21.4	-	1.1
	CBS[a]	3.073	21.0	26.6	-	3.212	20.2	23.4	-
	exp.[d]	2.970	34.7	33.2	-	2.970	34.7	33.2	-
	aDZ[a]	3.890	3.93	19.6	5.4	3.837	16.1	4.1	5.7
	aTZ[a]	3.306	60.5	23.8	19.1	3.226	85.7	23.2	10.8
	aTZ[b]	-	-	-	-	3.074	84.1	23.3	-
Ne2	aQZ[a]	3.265	70.9	24.7	21.1	3.228	92.4	23.0	10.5
	a5Z[c]	3.223	78.9	-	20.1	3.202	100.8	-	3.4
	CBS[a]	3.241	79.0	22.0	-	3.232	92.5	22.7	-
	exp.[d]	3.091	133.8	28.5	-	3.091	133.8	28.5	-
	aDZ[a]	3.969	257.4	22.9	221.3	3.874	318.4	26.3	48.8
	aTZ[a]	3.829	397.2	28.7	412.0	3.814	427.3	29.3	30.3
	aTZ[b]	-	-	-	-	3.762	400.6	29.1	-
Ar2	aQZ[a]	3.754	474.1	31.7	42.7	3.780	464.3	30.9	7.4
	a5Z[c]	3.750	482.7	-	62.7	3.783	472.7	-	10.0
	CBS[a]	3.701	539.1	33.5	-	3.772	474.6	31.4	-
	exp.[d]	3.756	453.6	30.9	-	3.756	453.6	30.9	-

The observed differences arise from the different implementations. Whereas the literature values were obtained from all-electron calculations, we used valence-only calculations with PPs and suitable basis sets such that we miss the core-core and core-valence correlation contributions. Furthermore, we exploited the RI approximation whereas the literature values did not.[173, 174] The excellent agreement suggests that our approximations are valid even for systems with very low binding energies.

3.6 Conclusions

We achieved the implementation of DH functionals with long-range separated correlation functionals. With the GPW approach, there is an efficient machinery for the calculation of long-range Coulomb integrals for RI-MP2 and RI-RPA calculations available. We are able to confirm the convergence with respect to the PBS size. We can employ the already used RI BSs optimized for the Coulomb potential for calculations with long-range Coulomb operator. The eigenvalue decomposition allows to slightly reduce the computational costs of the WFC method via a low-rank decomposition of the two-center integral matrix. Furthermore, other metrics are available for improved numerical accuracy. With the implementation of long-range HF integrals, we enabled the optimization of PPs for these DH functionals within the GPW approach.
The drawback of the implementation in CP2K is the lack of short-range DFT correlation functionals of which there is, starting with LibXC, version 5.0, only a LDA-based functional available.[70, 106, 142, 149, 151] Further benchmarks with respect to energies and gradients are required to test the performance of these functionals in condensed phase.

3.7 Appendix: Pseudopotential Parameters and Basis Sets

The HF-based PPs were taken from reference [170]. The PPs for range-separated and LDA-based calculations (rsLDA) were optimized starting from their corresponding HF PPs and a range separation parameter $\omega = 0.5$. The augmented double-zeta and triple-zeta primary basis sets of Ar and Ne were published by Stein *et al.*[175]. The basis sets of H and O were taken from Del Ben *et al.*[68].

Table 3.3: HF PP parameters.

H	1	1	
0.20049539759096	-4.17780338804233	0.72403926676805	
He	2	2	
0.20014385954257	-9.12199233589607	1.70294694788987	
C	6	4	
0.34816792458406	-8.54312820557867	1.33276540541946	
0.30230247000627	9.59710582360109		
N	7	5	
0.28300476743411	-12.39840200798251	1.86939057420079	
0.25539202567537	13.64483766978610		
O	8	6	
0.24676969870316	-16.66528269564613	2.52030687064467	
0.22121058101998	18.39425181647437		
Ne	10	8	
0.19050265092574	-27.39590696172339	4.41958869715540	0.01834396326683
0.17637388496062	28.18533818441574	0.83365996989179	-1.04842942962620
0.19585379054851	-0.27609661906079		
Ar	18	8	
0.39771927261258	-7.21348927487361	0.01323122557817	
0.31872450490949	17.20921819285275	-5.58549109340678	7.19978913165534
0.35357441343299	4.98951929408379		

Table 3.4: rsLDA ($\omega = 0.5$) PP parameters.

He	2	2	
0.20015287180513	-9.12150476631057	1.70260640996690	
Ne	10	8	
0.19050348250454	-27.14656173242935	4.43610130368390	0.00369847342065
0.17613915425891	28.18518742516878	0.83365363804867	-1.06358276339357
0.21738121484921	-0.23703124847681		
Ar	18	8	
0.40001306693831	-7.12841297015224	-0.04210995291428	
0.31892117677935	17.24158292202911	-5.58548506017518	7.24052949085434
0.35338638000839	4.97390476308691		

Table 3.5: Parameters of the cc-DZVP basis set of Hydrogen.

Shell type	Exponents	Contraction Coefficients
s	8.3744350009	-0.0283380461
	1.8058681460	-0.1333810052
	0.4852528328	-0.3995676063
s	0.1658236932	1.0
p	0.7270000000	1.0

Table 3.6: Parameters of the cc-TZVP basis set of Hydrogen.

Shell type	Exponents	Contraction Coefficients
s	10.8827241585	-0.0167058885
	3.0968750876	-0.0627538300
	0.9874518162	-0.1917521975
s	0.3450687533	1.0
s	0.1492693554	1.0
p	1.4070000000	1.0
p	0.3880000000	1.0
d	1.0570000000	1.0

Table 3.7: Parameters of the cc-QZVP basis set of Hydrogen.

Shell type	Exponents	Contraction Coefficients
s	12.5350945530	0.0114321687
	4.5491539489	0.0344048283
	1.6685193606	0.1007746979
s	0.6523849556	1.0
s	0.2733011244	1.0
s	0.1390608890	1.0
p	2.2920000000	1.0
p	0.8380000000	1.0
p	0.2920000000	1.0
d	2.0620000000	1.0
d	0.6620000000	1.0
f	1.3970000000	1.0

Table 3.8: Parameters of the aug-cc-DZVP basis set of Helium.

Shell type	Exponents	Contraction Coefficients
s	13.1305278312	-0.0500802904
	4.1977275150	-0.1474339352
	1.3647725570	-0.3245281495
	0.4549715461	-0.4365284986
s	0.1513197845	1.0
p	1.27500000	1.0
s	0.07255000	1.0
p	0.24730000	1.0

Table 3.9: Parameters of the aug-cc-TZVP basis set of Helium.

Shell type	Exponents	Contraction Coefficients
s	14.9231368362	0.0343142688
	5.9479610723	0.0933152472
	2.1932485067	0.2143451850
	0.8304881034	0.3500759421
s	0.3182586587	1.0
s	0.1197839846	1.0
p	3.04400000	1.0
p	0.75800000	1.0
d	1.96500000	1.0
s	0.05138000	1.0
p	0.19930000	1.0
d	0.45920000	1.0

Table 3.10: Parameters of the aug-cc-QZVP basis set of Helium.

Shell type	Exponents	Contraction Coefficients
s	16.7962354569	-0.0217345796
	8.0784265748	-0.0637981790
	3.2397448253	-0.1438730374
	1.3463173459	-0.2567923947
s	0.5669193043	1.0
s	0.2404041556	1.0
s	0.0990547078	1.0
p	5.99400000	1.0
p	1.74500000	1.0
p	0.56000000	1.0
d	4.29900000	1.0
d	1.22300000	1.0
f	2.68000000	1.0
s	0.04819000	1.0
p	0.16260000	1.0
d	0.35100000	1.0
f	0.69060000	1.0

Table 3.11: Parameters of the cc-DZVP basis set of Oxygen.

Shell type	Exponents	Contraction Coefficients	
sp	8.3043855492	0.1510165999	-0.0995679273
	2.4579484191	-0.0393195364	-0.3011422449
	0.7597373434	-0.6971724029	-0.4750857083
sp	0.2136388632	1.0	1.0
d	1.1850000000	1.0	

Table 3.12: Parameters of the cc-TZVP basis set of Oxygen.

Shell type	Exponents	Contraction Coefficients	
sp	10.2674419938	0.0989598460	-0.0595856940
	3.7480495696	0.1041178339	-0.1875649045
	1.3308337704	-0.3808255700	-0.3700707718
sp	0.4556802254	1.0	1.0
sp	0.1462920596	1.0	1.0
d	2.3140000000	1.0	
d	0.6450000000	1.0	
f	1.4280000000	1.0	

Table 3.13: Parameters of the cc-QZVP basis set of Oxygen.

Shell type	Exponents	Contraction Coefficients	
sp	11.8298677444	-0.0658681959	0.0389990698
	4.9326307395	-0.1327117781	0.1329824319
	1.9312997298	0.1529660213	0.2737974007
sp	0.7724003899	1.0	1.0
sp	0.2976334392	1.0	1.0
sp	0.1070061063	1.0	1.0
d	3.7750000000	1.0	
d	1.3000000000	1.0	
d	0.4440000000	1.0	
f	2.6660000000	1.0	
f	0.8590000000	1.0	
g	1.8460000000	1.0	

Table 3.14: Parameters of the aug-cc-DZVP basis set of Neon.

Shell type	Exponents	Contraction Coefficients	
sp	13.8523672900	0.1501498200	0.10214300
	4.0685498000	-0.0314908700	0.3058092400
	1.2730584300	-0.7070497300	0.4766050400
sp	0.3565013600	1.0	1.0
d	2.2020000000	1.0	
s	0.1230000	1.0	
p	0.1064000	1.0	
d	0.6310000	1.0	

Table 3.15: Parameters of the aug-cc-TZVP basis set of Neon.

Shell type	Exponents	Contraction Coefficients	
sp	17.4276488400	0.073686700	0.0702714400
	6.3439264100	0.0969132500	0.2167849700
	2.2823205800	-0.3010470300	0.4317763300
sp	0.7945993700	1.0	1.0
sp	0.2560537300	1.0	1.0
d	4.0140000000	1.0	
d	1.0960000000	1.0	
f	2.5440000000	1.0	
s	0.1133000	1.0	
p	0.0917500	1.0	
d	0.3860000	1.0	
f	1.0840000	1.0	

Table 3.16: Parameters of the aug-cc-QZVP basis set of Neon.

Shell type	Exponents	Contraction Coefficients	
sp	19.5499411153	-0.0632766758	0.0422167837
	8.0769601838	-0.1370919085	0.1383304483
	3.1678758504	0.1578325934	0.2813677061
sp	1.2535625634	1.0	1.0
sp	0.4749693396	1.0	1.0
sp	0.1666682167	1.0	1.0
d	6.47100000	1.0	
d	2.21300000	1.0	
d	0.74700000	1.0	
f	4.65700000	1.0	
f	1.52400000	1.0	
g	2.98300000	1.0	
s	0.10540000	1.0	
p	0.08178000	1.0	
d	0.27300000	1.0	
f	0.68900000	1.0	
g	1.22400000	1.0	

Table 3.17: Parameters of the aug-cc-DZVP basis set of Argon.

Shell type	Exponents	Contraction Coefficients	
sp	2.6724631600	0.1547491900	0.2663267700
	1.5750569800	-0.1300613000	-1.0821938600
	0.5528926600	-0.1247859600	0.1177549000
sp	0.1720724500	1.0	1.0
d	0.7380000000	1.0	
s	0.0709000	1.0	
p	0.0533000	1.0	
d	0.2400000	1.0	

Table 3.18: Parameters of the aug-cc-TZVP basis set of Argon.

Shell type	Exponents	Contraction Coefficients	
sp	3.5650652500	-0.03560400	-0.0341601300
	2.8711385000	0.107453000	0.0274003200
	0.928908200	-0.070935800	0.1084604500
sp	0.3762992800	1.0	1.0
sp	0.1388133000	1.0	1.0
d	1.2540000000	1.0	
d	0.4100000000	1.0	
f	0.8900000000	1.0	
s	0.0685000	1.0	
p	0.0487000	1.0	
d	0.1690000	1.0	
f	0.4060000	1.0	

Table 3.19: Parameters of the aug-cc-QZVP basis set of Argon.

Shell type	Exponents	Contraction Coefficients	
sp	5.2342384840	0.0594091942	-0.0218235164
	3.0293732626	-0.3808584814	-0.0397609427
	1.3121914980	0.0527214554	0.2096390562
sp	0.5914580696	1.0	1.0
sp	0.2478147218	1.0	1.0
sp	0.0969644987	1.0	1.0
d	0.31100000	1.0	
d	0.76300000	1.0	
d	1.87300000	1.0	
f	0.54300000	1.0	
f	1.32500000	1.0	
g	1.00700000	1.0	
s	0.06100000	1.0	
p	0.04350000	1.0	
d	0.11600000	1.0	
f	0.29400000	1.0	
g	0.45900000	1.0	

Chapter 4

Double-Hybrid DFT Functionals for the Condensed Phase: Gaussian and Plane Waves Implementation and Evaluation

This chapter is a reprint of reference [71]. The introduction was originally written by Vladimir Rybkin but is kept for completeness. We benchmark cohesive energies of molecular crystals obtained from different flavours of hybrid and double-hybrid density-functionals and analyze convergence with respect to the super cell size, basis set superposition errors, convergence with respect to basis set size and the general quality of the results.

4.1 Introduction

Electronic structure calculations for realistic condensed-phase systems are generally more involved than those for molecules. The former include more atoms and are performed under periodic boundary conditions (PBC), implying interactions between periodic images. Therefore, condensed-phase electronic structure modelling often relies on simple approximations. Tight-binding approaches - semiempirical methods, density functional based tight-binding (DFTB) - used to be the work horse in the field. With increased computational power Kohn-Sham density functional theory (KS DFT) [5] became a standard approach. Recently, implementations of wave function theories (WFT) became available, although their

application is far from routine.

In DFT, the energy is given as unique functional of electron density alone (Hohenberg-Kohn theorem) [6]. Although the exact functional is unknown, several approximation levels are available, often classified as rungs of the Jacob's ladder of accuracy [86]. The most simple approximation includes only local information on the density (local density approximation, LDA) [87, 89, 176–178]. More eleborate theories take more properties of the density into account. Including the density gradient yields generalized-gradient (GGA) approximations (LYP [179], PBE functionals [91]), whereas including the kinetic energy density gives meta-GGA functionals [92, 180, 181] (*e.g.* TPSS functional [93]).

Incorporating a portion of exact exchange (non-local) leads to hybrid functionals (*e.g.* PBE0 [96], B3LYP [95]). Exact exchange energy is not density-dependent, but is rather a non-local quantity (dependent on the density matrix) borrowed from WFT, *viz.* from Hartree-Fock (HF) theory [37, 182]. Hence, the term "hybrid" functional means including quantities from WFT, *i.e.* the Hartree-Fock exchange energy, into DFT functionals. Further examples of this approach are range-separated methods (HSE [94], ωB97X [183]) and double-hybrid functionals, the latter can also involve range separation. Whereas hybrid functionals depend on the occupied KS orbitals, double-hybrid functionals include additionally virtual orbitals. They account for electron correlation in both DFT fashion via exchange-correlation functional and WFT fashion via excited determinants. We will refer to WFT methods which include correlation energy as Wave-function correlation (WFC) method.

Double-hybrid functionals [66, 67, 114, 129, 184, 185] can potentially take "the best of the two worlds". GGA-, meta-GGA- and hybrid DFT functionals are relatively fast and accurate for covalently and ionically bound systems. However, they intrinsically fail to describe long-range dispersion interactions (which is often coped with by explicit dispersion corrections [26–29, 186, 187] and non-local functionals [33, 188]) and strong correlations. WFC methods, on the other hand, inherently include the correct asymptotic R^{-6} behaviour. Their significant disadvantage is the high computational cost: N^4 scaling and higher in the canonical formulations. Reduced-cost methods allow decreasing the scaling, although with high prefactors. Consequently, the cost of a double-hybrid DFT calculation is defined by the cost of its WFT part. The question may then arise: why not use pure WFT instead of double-hybrid functionals? The answer is that the $N^4 - N^5$ scaling of WFT methods used for double-hybrid functionals (second-order Møller-Plesset perturbation theory, MP2 [38]; random phase approximation, RPA [189]) are relatively crude approximations, and despite capturing long-range interactions they

can be outperformed by DFT functionals. Thus, inclusion of electron correlation in WFT and DFT fashion may lead to the improved accuracy of both at moderate price as compared to highly precise WFT approaches, such as coupled-cluster methods [39] scaling as N^6 and higher.

Most condensed-phase implementations of electronic structure methods are based either on the use of plane waves (PW) or Gaussian basis sets. Plane waves constitute a basis in a strict mathematical sense: they are orthogonal and complete. In PW basis DFT and correlation energies converge systematically with basis size [190]. However, due to the fact that PW do not reflect the character of chemical bonding, a larger number of basis functions is needed for accurate calculations, which is detrimental for calculations with WFC methods as virtual space becomes huge. Since atom-centered Gaussian functions reasonably approximate atomic orbitals, good accuracy can be achieved with compact basis sets, *i.e.* at a lower computational price, especially for DFT. WFC energies are more sensitive to basis set size and exhibit slow convergence with basis set size [58], especially for long-range dispersion interactions [64]. WFC methods and DHDFs are available for PW basis sets in VASP[190, 191], for Gaussian basis sets in CP2K[130], CRYSTAL[192] and GAMESS (US)[193] and for Slater type basis sets in ADF[194].

4.2 Theoretical Background

In the following, $a, b, ...$ are virtual orbital indices, $i, j, ...$ occupied orbital indices, $p, q, ...$ general orbital indices, and $P, Q, ...$ auxiliary function indices. In DFT, the total energy is given as a functional of the total ground-state density $n(\vec{r})$:

$$E_{DFT}[n] = T_0[n] + E_{ne}[n] + E_H[n] + E_{XC}[n], \tag{4.1}$$

where $E_{DFT}[n]$ is the total energy functional, $T_0[n]$ is the kinetic energy of a reference system of non-interacting electrons, $E_{ne}[n]$ is the nuclei-electron interaction energy, $E_H[n]$ is the Hartree energy describing the classical electron-electron interaction energy, and $E_{XC}[n]$ is the exchange-correlation energy describing the quantum mechanical contributions of the electron-electron interaction. The ground-state density is expressed in terms of orbital functions $\psi_i(\vec{r})$

$$n(\vec{r}) = \sum_i |\psi_i(\vec{r})|^2 \tag{4.2}$$

where i runs over all occupied orbitals. The orbital functions fulfill the orthonormality constraint:

$$\int d^3r \psi_i^*(\vec{r})\psi_j(\vec{r}) = \delta_{ij} \tag{4.3}$$

with the Kronecker delta δ_{ij}. The orbitals are solutions of the Kohn-Sham (KS) equation:

$$\left(-\frac{\Delta}{2} + v_{ne}(\vec{r}) + v_H[n](\vec{r}) + v_{XC}[n](\vec{r})\right)\psi_i(\vec{r}) = \varepsilon_i\psi_i(\vec{r}) \tag{4.4}$$

with the potential arising from the nuclei $v_{ne}(\vec{r})$, the Hartree potential $v_H[n](\vec{r})$, the exchange-correlation (XC) potential $v_{XC}[n](\vec{r})$, and the orbital energy ε_i of orbital i. In this article, we will consider Gaussian functions centered at the atoms only.
Because the total energy functional is not known explicitly in terms of the ground-state density, we rely on approximations of the XC functional. These approximate energy functionals are given as integrals of a function explicitly depending on the ground-state density, its gradient and its Laplacian. For convenience, the XC functional is split into an exchange functional $E_X[n]$ and a correlation functional $E_C[n]$.
The more complex hybrid density functionals (HDFs)[95] include explicit information of the occupied orbitals. They modify the exchange functional by including a certain amount $\alpha_{X,HF}$ of Hartree-Fock (HF) exchange $E_{X,HF}[n]$ providing the exchange functional

$$E_{X,hybrid} = \alpha_{X,HF}E_{X,HF}[n] + \alpha_{X,DFT}E_{X,DFT}[n]. \tag{4.5}$$

We introduce the amount $\alpha_{X,DFT}$ of DFT exchange $E_{X,DFT}[n]$ to reflect that the DFT exchange functional is an already known GGA or meta-GGA functional (compare [95, 96]). Using the Mulliken notation (chemists' notation) for electron repulsion integrals

$$(pq|rs) = \int d^3r \int d^3r' \frac{\phi_p^*(\vec{r})\phi_q(\vec{r})\phi_r^*(\vec{r}\,')\phi_s(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} \tag{4.6}$$

the HF exchange energy can be written as

$$E_{X,HF} = -\frac{1}{2}\sum_{ij}(ij|ji). \tag{4.7}$$

Non-HDFs suffer from self-interaction errors[88]. These are reduced in HDFs but usually not fully cancelled since $\alpha_{X,HF} \neq 1$ in general case. This self-interaction error results in erroneous description of charge-separation processes and transition states. But even hybrid methods and HF lack a reasonable description of dispersion interactions decaying like R^{-6} with R being a measure of charge separation. For increased flexibility, we can further split the exchange functional in a long-range and a short-range functional and describe both with a given mixture of HF theory and DFT resulting in range-separated HDFs[94].
The highest flexibility is achieved by including virtual orbitals $\psi_a(\vec{r})$. Double-hybrid density functionals (DHDFs) are HDFs in which the correlation functional is composed of a mixture of a DFT correlation functional $E_{C,DFT}[n]$ with ratio $\alpha_{C,DFT}$ and correlation energy $E_{C,WFT}[n]$ of a WFC method with ratio $\alpha_{C,WFT}$ providing a functional

$$E_{C,double-hybrid} = \alpha_{C,WFT} E_{C,WFT}[n] + \alpha_{C,DFT} E_{C,DFT}[n]. \tag{4.8}$$

Because WFC methods are computationally more demanding than HDFs or standard DFT functionals, most DHDFs exploit the MP2 theory, the SOS-MP2 theory, or the RPA method.
The correlation energy within the MP2 theory for closed-shell systems is

$$E_{C,MP2} = \sum_{ijab} \frac{(ia|jb)\left[2(ia|jb) - (ib|ja)\right]}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}. \tag{4.9}$$

The computationally most expensive step of the MP2 method is given by the transformation of the electron interaction integrals from atom orbital basis to molecular orbital basis leading to a $\mathcal{O}(N^5)$ scaling with N being a measure of system size. The prefactor can be reduced by the resolution-of-the-identity (RI) approach introducing an auxiliary basis in which densities are expanded giving the equation

$$(pq|rs) = \sum_P B_P^{pq} B_P^{rs} \tag{4.10}$$

with

$$B_P^{ia} = \sum_Q (pq|Q)(Q|P)^{-1/2} \tag{4.11}$$

$$(pq|P) = \int d^3r \int d^3r' \frac{\phi_p^*(\vec{r})\phi_q(\vec{r})\phi_P(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} \tag{4.12}$$

$$(P|Q) = \int d^3r \int d^3r' \frac{\phi_P(\vec{r})\phi_Q(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}. \tag{4.13}$$

This method is called RI-MP2[195, 196].
A simplified version of the RI-MP2 method is the Scaled-Opposite-Spin(SOS)-MP2 method[55] given by

$$E_{C,SOS-MP2} = -\int_0^\infty d\tau \mathrm{Tr}\left(\mathbf{Q}_{SOS-MP2}(\tau)\mathbf{Q}^T_{SOS-MP2}(\tau)\right) \tag{4.14}$$

with

$$(\mathbf{Q}_{SOS-MP2}(\tau))_{PQ} = \sum_{ia} B^{ia}_P e^{\tau(\varepsilon_i-\varepsilon_a)} B^{ia}_Q. \tag{4.15}$$

The integration is carried out numerically using a Minimax quadrature. The RI-SOS-MP2 method scales like $\mathscr{O}(N^4)$.
Another correlation method with increasing popularity is the Random Phase Approximation (RPA) method[102, 103] within the RI approximation

$$E_{C,RPA} = \frac{1}{2}\int_0^\infty \frac{d\omega}{2\pi} \mathrm{Tr}\left(\ln(1+\mathbf{Q}_{RPA}(\omega)) - \mathbf{Q}_{RPA}(\omega)\right) \tag{4.16}$$

with

$$\mathbf{Q}_{RPA}(\omega) = 2\sum_{ia} B^{ia}_P \frac{\varepsilon_a-\varepsilon_i}{\omega^2+(\varepsilon_a-\varepsilon_i)^2} B^{ia}_Q. \tag{4.17}$$

RI-RPA scales like $\mathscr{O}(N^4)$. As with the RI-SOS-MP2 method, the integration is carried out numerically using a Clenshaw-Curtis grid[197] or a Minimax grid[165, 198].
All WFC methods and all DHDFs correctly reproduce the R^{-6} energy behaviour of long-range interactions. Comparable to range-separated HDFs, there are DHDFs with range-separated exchange functionals like the ωB97X-2 functional[123]. Further, there are DHDFs with range-separated correlation functionals[65]. In this article, we will not focus on DHDFs with range-separated correlation functionals and refer to the literature[64, 85, 120, 121, 199] for more details.

4.3 Computational Details

4.3.1 Gaussian and plane waves method (GPW) and integral evaluation

The Gaussian and plane waves method (GPW) [134] allows for efficient periodic calculations with Gaussian basis sets using a dual representation of the electronic

Table 4.1: Structural information about the bulk structures used in this study. n_{fu} is the number of formula units per unit cell. References for the geometrical information of the respective system are provided in the last column. Please note that there was a mistake in the cell parameters of CO_2 provided in reference [70].

System	a;b;c **(Å)**	n_{fu}	**References**
NH_3	5.048	4	[201]
HCN	4.13; 4.85; 4.34	2	[202]
Ne	4.464	4	[203–205]
Ar	5.300	4	[206, 207]

density and molecular orbitals. It assumes the use of a primary Gaussian basis for the expansion of matrix quantities (density matrix, KS matrix) and an auxiliary plane waves (PW) basis for the evaluation of the Hartree potential and the numerical integration of density functionals. To converge GPW calculations, one has to pay attention to both the size and quality of the Gaussian basis and the energy cutoff for the PWs. In the current implementation, GPW is used for the calculation of the Hartree potential, XC functionals, and two and three center integrals necessary for the RI-MP2 and RI-RPA methods. Exchange integrals are computed analytically using a truncated Coulomb potential [200].

4.3.2 Test systems

Because we are interested in the description of intermolecular interactions, we are testing the functionals on molecular crystals (NH3, HCN) and rare-gase crystals (Ar, Ne) which have been studied by Sansone et al[70]. Structural information of the unit cells are summarized in table 4.1. In case of molecular solids, we were using structures reoptimized at the B3LYP-D* level[208].
Molecular crystals represent systems with a mixture of covalent bonding and dispersion interactions. NH_3 and HCN crystals additionally contain hydrogen bonds which are crucial for the discription of proteins. In contrast to that, there are only dispersion interactions within the rare-gas crystals. This results in low cohesive energies and the need for well-balanced functionals.

4.3.3 Parameters of the calculations

All calculations have been carried using a development version 8.0 of CP2K[130]. To ensure convergence with respect to the density cutoff, we were using high

cutoffs of $1,500$ *Ry* for all RPA and MP2 calculations, $4,000$ *Ry* for remaining calculations of the molecular crystals NH_3 and HCN, and 10000 *Ry* for the rare-gas crystals Ar and Ne (see chapter 4.4.1 for more details) and a relative cutoff of 50 *Ry*. For the rare-gas crystals, we set the parameters *EPS_DEFAULT*, *EPS_PGF_ORB*, *EPS_SCF*, and *EPS_SCHWARZ* in the HF section to 10^{-30}, 10^{-50}, 10^{-5}, and 10^{-10}, respectively, for the molecular crystals, we were using for the same parameters 10^{-20}, 10^{-40}, 10^{-5}, and 10^{-9}, respectively (see the CP2K manual for the meaning of these parameters). HF calculations for the bulk systems were using a truncated Coulomb potential with a cutoff radius of roughly half the super cell size. All densities have been smoothed using the *NN*10 method.
RI-MP2, RI-SOS-MP2 and RI-RPA calculations have been carried out using the GPW method to determine all integrals with a primary cutoff of 300 *Ry* and a relative cutoff of 50 *Ry*. We have exploited an 8-point minimax grid for all RI-RPA and RI-SOS-MP2 calculations.

4.3.4 Choice of functionals and implementation

We carried out calculations at the PBE[91], ωB97M-V[143], ωB97X-2[123], PW6B95[209], PWRB95[118], SOS-PBE0-2[115], RI-MP2[38, 195] and RI-RPA[103] levels of theory. PBE and RI-MP2 are used to compare differences between valence-only calculations of our valence-only calculations and the all-electron calculations of Sansone et al[70]. PW6B95 is a meta-hybrid functional which performed best for weakly interacting systems with more pronounced dispersion interactions. PWRB95 is its RPA-based DHDF. ωB97M-V is a dispersion-corrected range-separated meta-hybrid functional. ωB97X-2 is an MP2-based DHDF with range-separated exchange functional. SOS-PBE0-2 is a RI-SOS-MP2-based DHDF. With this choice, we cover a large variety of different flavours of meta-hybrid and DHDF theories. Due to very high computational cost, we have restricted ourselves to this small, but representative set of functionals: one for each flavour of DHDFs and a corresponding HDF.
PBE calculations have been carried out using the CP2K implementation of PBE. For the ωB97M-V, and the PW6B95 functionals, we exploited the implementations of the LibXC library[151], version 4.3.4. Since the VV10 dispersion correction is not available in CP2K, we relied on the rVV10 correction and the parametrization suggested by Mardirossian et al[210]. For the ωB97X-2 and PWRB95 functionals, we implemented the required parameter sets into the LibXC library.

4.3.5 Basis sets and pseudopotentials

The MP2 and RPA implementations within CP2K rely on a pseudopotential (PP) approach with Goedecker-Teter-Hutter PPs[135]. For the PBE functional, we used PPs optimized for PBE, for RPA and MP2 calculations, we were using PPs optimized for HF whereas for both HDFs and all DHDFs, we utilized PPs optimized for the PBE0 functional. All PPs have been taken from the Github repository of Jürg Hutter[170].
Correlation-consistent primary basis sets and suitable auxiliary basis sets of double zeta (DZ) and triple zeta (TZ) quality for the elements C, H, N and O have been taken from Del Ben et al.[68] We have optimized appropriate correlation-consistent primary and auxiliary basis sets of the same qualities for Ne and Ar using the polarization functions of the respective Dunning basis sets[196, 211, 212]. All PPs and primary and auxiliary basis sets are compiled in the Supporting Information.

4.3.6 Cohesive energies and basis set superposition error

To determine total energies per formula unit, we carried out calculations of 2x2x2, 3x3x3 and 4x4x4 supercells of all given unit cells and used a linear fit of the total energy per formula unit against the inverse of the cell volume.
Because calculations of cohesive energies usually suffer from basis set superposition errors (BSSE), we perform a counterpoise correction[213]. The BSSE-free cohesive energies E_{coh} are calculated according to

$$E_{coh} = E_{bulk} - E_{mol+ghost,bulk} + E_{mol,bulk} - E_{mol,gas} \tag{4.18}$$

with the total bulk energy per formula unit extrapolated to infinite cell volume E_{bulk}, the energy of the molecule with ghost atoms $E_{mol+ghost,bulk}$, the energy of the molecule using the bulk geometry $E_{mol,bulk}$, and the total energy of the molecule using an optimized gas phase structure $E_{mol,gas}$. For Ar and Ne, we trivially have $E_{mol,bulk} = E_{mol,gas}$.
The corresponding BSSE is given by

$$\Delta E_{BSSE} = E_{mol+ghost,bulk} - E_{mol,bulk}. \tag{4.19}$$

For the BSSE calculations, we took the crystalline structures, chose one molecule (or atom for Ar and Ne) surrounded by all ghost atoms within a 3x3x3 supercell.

4.4 Results

4.4.1 General Remarks

We found the convergence of total energies of meta-HDFs PW6B95 and PWRB95 requires very tight energy cutoffs for the auxiliary PW basis of at least 4000 *Ry*. In contrast to that, calculations with the other meta-HDFs in our benchmark study, ωB97M-V, provided reasonable results with a cutoff of only 1200 *Ry*. Because the basis functions for the elements argon and neon are more localized than those for hydrogen, carbon and nitrogen, higher cutoffs for the noble gases were needed for an adequate representation of the basis functions of these elements on the grid. It is well-known that GGA functionals and especially meta-GGA functionals require very tight integration grids for convergence and thus accurate results. Such cutoffs reflect numerical issues and the need for very fine integration grids when using the PW6B95 and PWRB95 functionals. Such grids are not necessary for the ωB97M-V functional which has been optimized with coarser integration grids in mind[143]. Thus, energy differences converged faster with ωB97M-V and PBE. Nevertheless, the total energies were not converged. To remove any possible problems due to incomplete convergence with respect to cutoffs, we utilized unusually high cutoffs for all density functionals.
Furthermore, we have found convergence problems with the PW6B95 and PWRB95 functionals, which can be resolved with density smoothing. Unfortunately, in some cases an increase of the energy cutoff for the density resulted in SCF convergence issues which could not be resolved with tighter filter thresholds. Nevertheless, we were able to achieve convergence by restarting the calculations with a higher cutoff starting from the converged SCF results with a lower cutoff. This was not possible for argon, where we exploited a cutoff of 4000 *Ry* for the PW6B95 and PWRB95 functional. Thus, some numbers for the PW6B95 and PWRB95 functionals are not fully converged with respect to the density cutoff.
Due to the higher computational costs, we have not carried out calculations of the 4x4x4 supercells on the TZ level.
All cohesive energies are compiled in tables 4.2 and 4.3.

Table 4.2: Cohesive energies E_{coh} and absolute relative error with respect to experimental results for all considered methods and systems in kJ/mol exploiting basis sets of DZ quality. The statistical indicators are the mean absolute error (MAE) and the mean absolute relative error (MARE). [a] This work. [b] Values by Sansonse et al.[70]. [c] Experimental values[204, 207, 214], corrected for zero-point energy (ZPE) and thermal effects at 298 K[70, 215].

Functional	$E_{coh}(Ne)$	**%**	$E_{coh}(Ar)$	**%**	$E_{coh}(NH_3)$	**%**	$E_{coh}(HCN)$	**%**	**MAE**	**MARE**
PBE[a]	-2.92	48	2.47	131	-39.9	9	-29.0	32	7.09	0.55
PBE[b]	-0.27	86	2.36	130	-26.9	25	-28.0	34	8.95	0.69
RPA[a]	0.82	141	4.61	159	-19.0	47	-28.1	33	11.7	0.96
MP2[a]	0.83	141	2.95	138	-27.6	23	-35.2	17	7.40	0.80
MP2[b]	0.22	111	3.13	140	-24.2	33	-31.7	25	9.01	0.78
ωB97M-V[a]	-2.65	34	-8.30	7	-41.2	13	-58.0	36	5.37	0.23
PW6B95[a]	-1.99	1	-0.03	99	-27.1	25	-33.8	20	6.43	0.37
PWRB95[a]	-2.29	16	-4.15	46	-34.8	4	-40.3	5	1.91	0.18
ωB97X-2[a]	-0.43	78	-3.60	53	-35.2	3	-42.3	0	1.77	0.34
SOS-PBE0-2[a]	0.57	128	2.76	135	-25.0	31	-31.0	27	8.98	0.81
Expt.[c]	-1.97		-7.73		-36.3		-42.6			

Table 4.3: Same as table 4.2, but with basis sets of TZ quality. [b] exploits basis sets of augmented DZ quality.

Functional	$E_{coh}(Ne)$	**%**	$E_{coh}(Ar)$	**%**	$E_{coh}(NH_3)$	**%**	$E_{coh}(HCN)$	**%**	**MAE**	**MARE**
PBE[a]	-1.22	37	0.49	106	-28.5	21	-29.5	30	7.47	0.49
PBE[b]	-0.40	79	0.42	105	-26.2	27	-29.7	30	8.18	0.61
RPA[a]	0.32	116	0.44	105	-24.1	33	-34.9	17	7.58	0.68
MP2[a]	0.05	102	-4.48	42	-34.4	5	-43.9	3	2.13	0.38
MP2[b]	-1.10	44	-6.45	16	-31.8	12	-41.4	2	1.96	0.19
ωB97M-V[a]	-3.51	78	-9.28	20	-39.5	8	-50.9	19	3.65	0.32
PW6B95[a]	-2.31	17	-1.95	74	-25.6	29	-32.6	23	6.72	0.36
PWRB95[a]	-2.67	35	-3.10	59	-36.1	0	-44.0	3	1.73	0.25
ωB97X-2[a]	-2.59	31	-8.15	5	-38.6	6	-43.0	0	0.93	0.11
SOS-PBE0-2[a]	0.10	105	-0.12	98	-26.4	27	-33.0	22	7.28	0.63
Expt.[c]	-1.97		-7.73		-36.3		-42.6			

4.4.2 Convergence with respect to super cell size

In figure 4.1, we compiled the differences in total energies per formula unit relative to the extrapolated total energies. In general, we expect the total energies to decrease with increasing supercell size and the extrapolated value is a lower bound for the total energies of the super cells. Our results show exactly this behaviour.
An important question is for which supercell size the error becomes negligible. A useful magnitude is given by the chemical accuracy of $4kJ \cdot mol^{-1}$. For weakly-interacting systems such as rare-gas crystals with cohesive energy of less than chemical accuracy, the order of magnitude is set by the cohesive energy itself. As the error of a method should be not larger than chemical accuracy, the allowed error of the supercell method must be at least one order of magnitude smaller then the methodological error, *i. e.* not larger than $0.4kJ \cdot mol^{-1}$. We find that a 3x3x3 super cell provides sufficient accuracy for all functionals and test systems. This behaviour is in agreement with the literature[68]. Sometimes, the total energy per formula unit of the 4x4x4 super cell has a higher magnitude than this of the 3x3x3 supercell, which may be due to numerical issues. For PBE, a cubic fit does not seem to be appropriate, and an exponential fit should be used instead.

4.4.3 Convergence of the BSSE

The BSSEs for the different test systems are compiled in figure 4.2. First, we would like to point out that the BSSE is significantly larger for the molecular crystals than for the rare-gas crystals. This might be related to the larger number of atoms per molecule and to the spread of the basis functions. Since the effective core charge of rare-gas atoms is larger than for carbon or nitrogen, the basis functions are more localized which results in weaker overlap with neighbouring atoms. This is supported by the smaller reduction in BSSE for Ar and Ne when we exploit larger basis sets. Thus, augmentation of basis sets must significantly reduce BSSEs of Ar and Ne. Indeed, diffuse basis functions actually improve cohesive energies as shown by Sansone et al.[70]
Molecular crystals are thus more suitable objects to study BSSE than rare-gas crystals. For both molecular crystals in the test set, the non-DHDFs PBE, PW6B95 and ωB97M-V, provide the smallest BSSEs whereas the two WFC methods MP2 and RPA have the largest BSSEs, as expected. The DHDFs have a BSSE between both classes of methods because they employ a mixture of DFT and WFC methods.
Furthermore, we note that the WFC methods in CP2K are implemented within

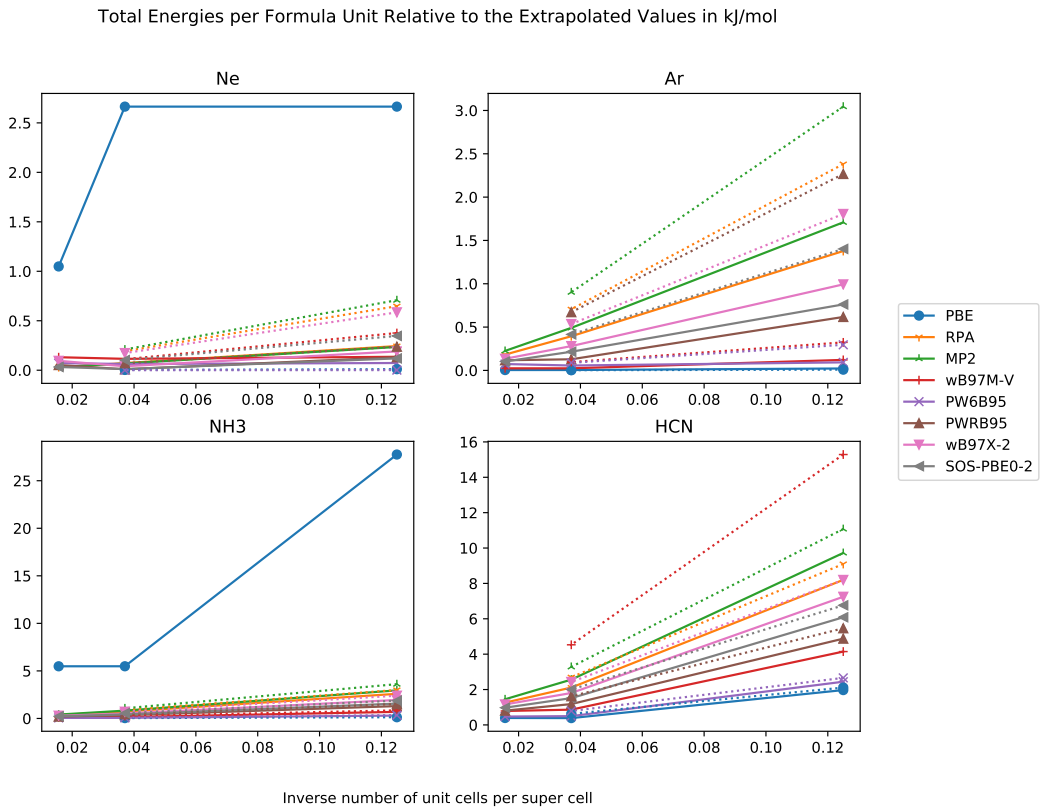

Figure 4.1: Total energies per formula unit relative to extrapolated total energy in $kJ \cdot mol^{-1}$ against inverse number of unit cells in supercell with basis sets of DZ and TZ quality for the systems a) Ne, b) Ar, c) NH_3, d) HCN.

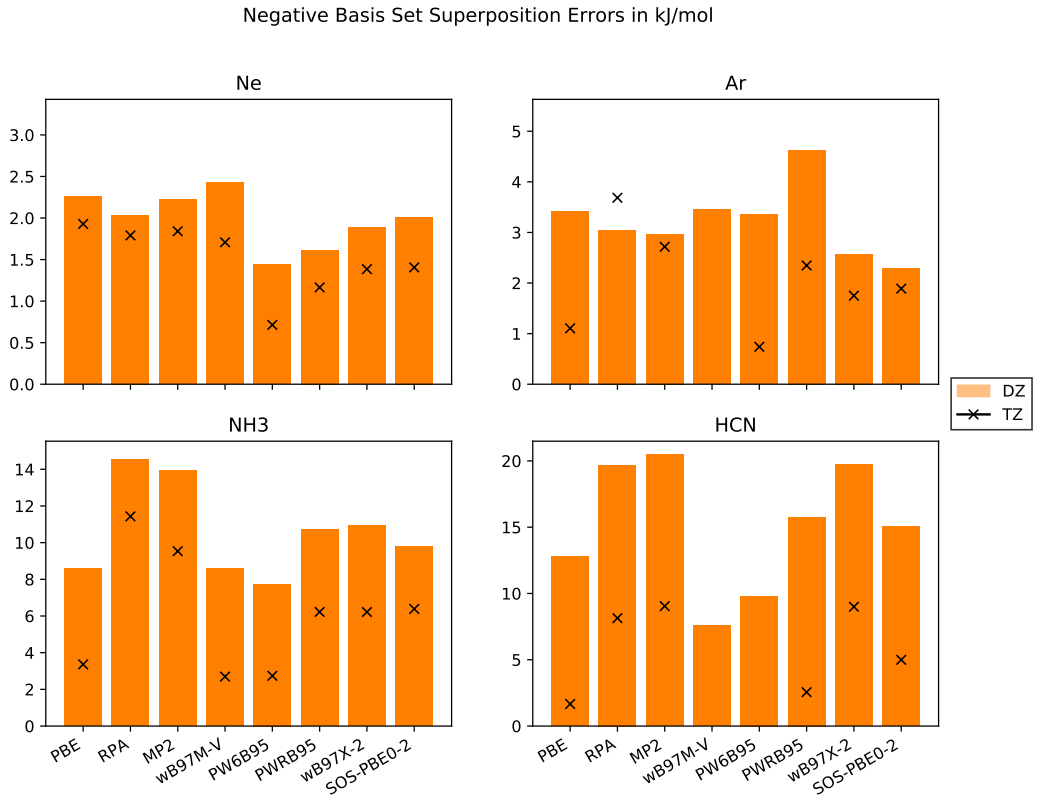

Figure 4.2: Negative Basis set superposition errors in $kJ \cdot mol^{-1}$ with basis sets of DZ and TZ quality for the systems a) Ne, b) Ar, c) NH_3, d) HCN.

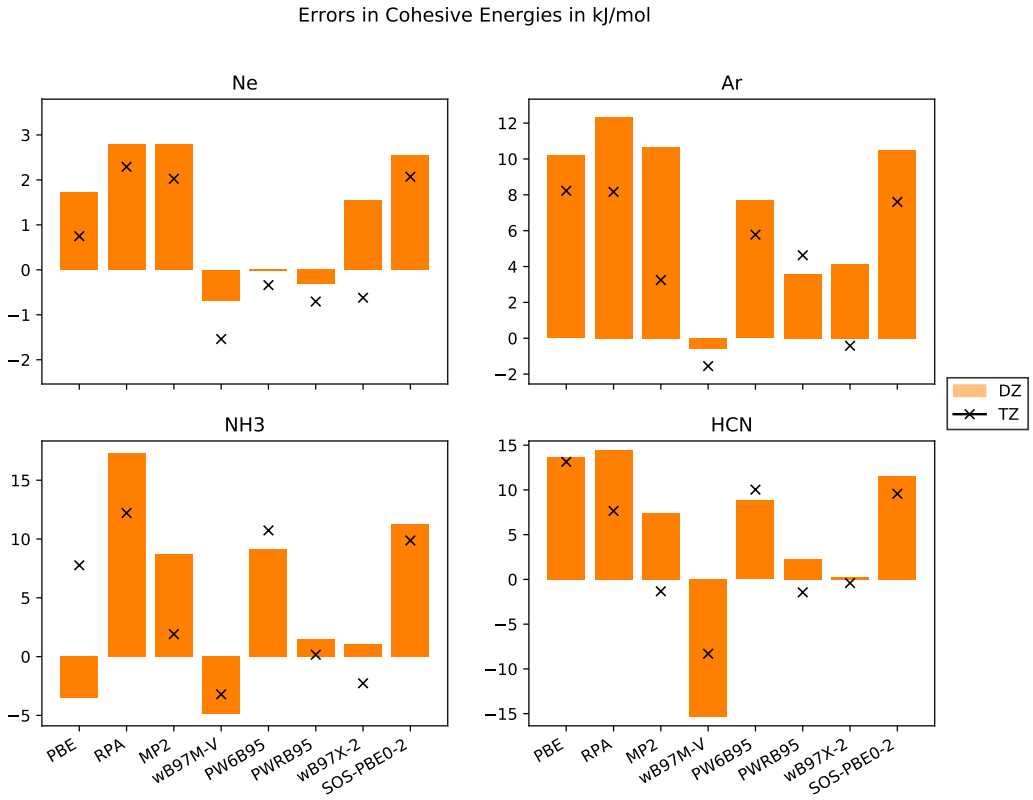

Figure 4.3: Errors in cohesive energies in $kJ \cdot mol^{-1}$ with respect to the experimental values with basis sets of DZ and TZ quality for the systems a) Ne, b) Ar, c) NH_3, d) HCN.

the RI approximation employing an auxiliary basis set. This leads to an additional source of BSSE for RPA, MP2 and all the DHDFs because the addition of the auxiliary functions of the ghost atoms increases the overall accuracy.

4.4.4 Convergence with respect to basis set size

In numerous studies, it was shown that total energies from DFT calculations converge exponentially with respect to basis set size. In contrast to that, total energies from WFC methods converge cubically with respect to basis set size when employing correlation-consistent basis sets. Thus, most DHDFs are expected to have a cubic convergence with respect to basis set size but with a smaller prefactor.

DHDFs employing a long-ranged Coulomb operator only and describing short-ranged interactions with a density functional, converge exponentially[64]. This behaviour is confirmed with our data compiled in figure 4.3.
Since larger basis sets systematically reduce total energies, cohesive energies increase. We observe this behaviour for the WFC methods and almost all DHDFs. The slight difference for PWRB95 in case of Ar may be due to not full convergence with respect to super cell size. For the other functionals - PBE, ωB97M-V and PW6B95 - the cohesive energy from the TZ basis set is sometimes higher, *i.e.* the system is weaker bound. One problem might be that the 2x2x2 super cells are not yet fully converged or the extrapolation scheme using a linear fit of the total energies versus the inverse of the volume is not appropriate and an exponential fit might be more suitable.
Next, we would like to discuss the results obtained for the molecular crystals NH_3 and HCN. They are bound together by covalent bonds, dipole-dipole interactions, and dispersion interactions. For both systems, the results with the RPA and MP2 methods significantly improve the results over GGA DFT functionals, MP2 even achieving chemical accuracy. The ωB97M-V functional also provides very accurate numbers. The PW6B95 functional, as PBE, systematically underestimates the cohesive energies with errors compatible to PBE. The PWRB95 functional significantly improves upon the results of its relative PW6B95, bringing them within 1 $kJ \cdot mol^{-1}$ from the experiment. The same holds for the ωB97X-2 functional compared with the ωB97M-V functional, although the DHDF is based on the non-meta-GGA HDF ωB97X[216]. One of the worst performing functionals is SOS-PBE0-2.
For the rare-gas crystals, the picture is more complicated because the absolute values of the cohesive energies are of the order of the chemical accuracy. As pointed out by Sansone et al.[70], augmented basis sets are required for these systems. Our cohesive energies from MP2 with a TZ basis set are only slightly lower than their result with a DZ basis set but still much worse than those with an augmented basis set for both Ne and Ar. Consequently, our results for Ne do not allow for an evaluation of the performance of these functionals and further studies employing either quadruple or augmented basis sets (which are to be constructed) are needed. Nevertheless, our results for Ar show that the ωB97X-2 functional provides a good description. The same holds for MP2, PW6B95 and PWRB95, although one needs further investigations with augmented basis sets.
This issue does not apply to the molecular crystals. Indeed, our cohesive energies with a TZ basis set are even lower than those with an augmented basis set. Thus, the use of augmented basis sets is not necessary for the molecular crystals.

This result is important for reducing computational costs of HF calculations and low-scaling WFC methods.

4.5 Discussion

Because DHDFs can be considered to be a mixture of DFT and WFC methods, the flexibility of DHDF parametrizations can yield approaches more accurate than the parent DFT and WFC functionals. At the same time, they inherit the shortcomings of both classes. Due to the dependence on the grid parameters, the functionals PW6B95 and PWRB95 are more difficult to use than others: care must be taken to check whether the results are converged with respect to the grid parameters, in CP2K, the density cutoff.
As expected, PBE can only provide the order of magnitude for weakly interacting systems, although it converges fast with respect to basis set size and has a low BSSE. MP2 and RPA are more sensitive to the basis set size and exhibit large BSSEs. These methods provide a moderate accuracy for different systems with small basis sets.
Non-DHDFs benefit from lower BSSEs. The PW6B95 functional has high demands on integration grids. Both considered functionals also provide a moderate accuracy and should be favourable over MP2 and RPA with their higher computational costs.
The double-hybrid functionals PWRB95 and ωB97X-2 show excellent performance with moderate BSSEs and lower basis set incompleteness errors. Both have computational costs compatible to full MP2 or RPA calculations and inherit the need of fine integration grids for accurate results, especially for PWRB95.
The non-empirical SOS-MP2 based DHDF, SOS-PBE0-2, does not provide any advantage as compared to the original methods. It was pointed out by different authors[25, 217] that non-empirical DHDFs usually perform worse than empirical DHDFs.

4.6 Conclusions

In this study, we examined a selection of different HDFs and DHDFs by computing cohesive energies in four different crystal structures. Our results show that DHDFs inherit the shortcomings of the underlying DFT functional (integration grids) and the underlying WFC method (computational costs, BSSE, basis

set dependence). We were able to show that the PWRB95 and the ωB97X-2 functionals provide excellent accuracy for molecular and rare-gas crystals. The HDFs ωB97M-V and PW6B95 also provide reasonable accuracy for these systems, whereas the SOS-PBE0-2 functional underperforms and can not be recommended.
The exploited basis sets allow a good description of molecular crystals. For the rare-gas crystals, we showed that non-augmented basis sets are not sufficient to achieve energy convergence with respect to the basis set size. Due to the high computational costs, we leave studies with augmented basis sets (and the construction of those basis sets) as well as benchmarking more range-separated DHDFs for future prospect.

Chapter 5

Double-Hybrid Density Functionals for the Condensed Phase: Gradients, Stress Tensor, and Auxiliary-Density Matrix Method Acceleration

This chapter is a reprint of a still unpublished but accepted article.[72] We present an efficient implementation of gradients of MP2-based double-hybrid functionals. We accelerate Hartree-Fock calculations with the Auxiliary-Density Matrix Method and analyse the performance and accuracy and discuss suitable large-scale applications. We show that our implementation performs efficiently for systems containing a few hundred atoms on 1,000 GPUs.

5.1 Introduction

Most applications of electronic structure methods are currently based on Density Functional Theory (DFT).[5, 6] DFT methods are widely available in standard quantum chemical software, are computationally efficient and a range of low scaling approaches was developed.[7–23] The different density functionals are often classified within a hierarchical scheme (Jacob's ladder analogy) with respect to their explicit dependence on information of the electronic density.[86] On the first rung, we find functionals based on the Local Density Approximation (LDA) employing the local density only, on the second rung are functionals taking additionally the density gradient into account (Generalized Gradient Approxima-

tion, GGA) while meta-GGA functionals augment them with an additional dependence on the kinetic energy density and/or the Laplacian of the density.[89, 91–93, 178, 179, 181] Up to the third rung, density functionals usually suffer from self-interaction errors which are, at least partially, removed on the fourth rung by the (partial) inclusion of exact (Hartree-Fock, HF) exchange in hybrid functionals.[37, 88, 94–96, 182, 183] Dispersion interactions, crucial for the description of biomolecules, molecular crystals, liquids or interactions on surfaces, are still absent on the hybrid functional level. Its inclusion is achieved with either empirical correction schemes or explicitly non-local functionals.[26, 27, 29, 33, 55, 187, 188]

Especially for weakly interacting systems, common functionals have severe shortcomings and a systematic approach which allows an easy construction of DFT functionals is missing. In contrast to that, correlated methods derived from Wave Function Theory (WFT) provide systematic approaches with increasing accuracy, e.g. methods based on HF theory, Moller-Plesset perturbation theory, Coupled-Cluster theory or the Random-Phase Approximation (RPA).[38, 39, 102, 103, 189, 196] The advanced WFT methods, having formally at least quartic scaling behavior, are usually more accurate but also more expensive than hybrid density functionals. Though, recent low-scaling variants of these methods allow the treatment of larger systems.[41–52] As a drawback, correlated methods require much larger basis sets than DFT methods. This basis set dependence can be tackled with range-separated approaches or explicitly correlated methods.[58, 64] Dispersion interactions are, unlike in local density functionals, mostly well described, however, some systems require higher levels of theory for improved accuracy.

Double-Hybrid (DH) functionals try to combine the best of both worlds: DFT is augmented with the description of long-ranged interactions from a correlated method like MP2 or RPA.[66, 115, 118] DH functionals can be routinely applied to molecular systems, but large computational resources are necessary for condensed phase calculations. Therefore the number of applications to condensed phase systems is rather limited. Nevertheless, in these applications the increased accuracy of DH functionals over (meta)-GGA functionals and hybrid functionals for the description of condensed phase systems has been established.[68, 71, 218, 219] A variety of DH functionals, relying on various underlying hybrid density functionals and the correlated methods is described in the literature.[65, 69, 70, 85, 118, 121, 123] For practical applications the availability of analytical gradients for these methods is important, enabling geometry optimizations and molecular dynamics simulations.

Even for smaller condensed phase systems the calculation of HF exchange

still requires large computational resources. The total number of electron repulsion integrals of atom centered basis functions increases at most quartically with respect to system size. This scaling can be reduced by different techniques up to linear scaling.[220–227] Despite of linear scaling approaches, HF calculations may suffer from large prefactors due to the high operation count of the integral calculation step, load-balancing problems in parallel applications or their large memory footprint. Modern approaches make use of specialized libraries with optimized hardware-specific kernels, the Resolution-of-the-Identity (RI) approach or the Auxiliary Density Matrix Method (ADMM), and a few target modern GPU hardware.[228–235] HF implementations for periodic systems need special care due to the singularity at the Gamma point and the slow decay of the Coulomb operator. These problems are addressed by the use of short-ranged operators or by direct summation approaches.[236–242] Especially HF and hybrid functional calculations with augmented basis sets for weakly interacting systems are computationally demanding because of the large extent of the additional diffuse basis functions and the resulting larger number of significant integrals.[130, 243–246]
In the following sections, we will present an implementation of DH functionals based on MP2 with nuclear gradients and stress tensors suitable for periodic calculations at the Γ-point. We will further show the extension to the ADMM approximation for HF exchange. Extensive tests for accuracy and efficiency are included.

5.2 Theoretical Background

In the following discussion, a,b,c represent an unoccupied (virtual) molecular orbital (MO), i,j,k occupied MOs, p,q denote any MO, Greek letters $\mu,\nu,\ldots$ denote primary basis set (PBS) indices, capital Latin letters $P,Q,R,\ldots$ denote RI basis function indices. A hat indicates quantities and indices of the Auxiliary Density Matrix Method (ADMM). We will consider Γ-point calculations only such that all quantities can be chosen real. All basis functions and orbitals are periodically repeated, integrals are over the computational cell, and Coulomb integrals are calculated using Ewald techniques.
The total energy E^{tot} of a density functional can be written as

$$E^{\mathrm{tot}}[\rho] = \sum_{\mu\nu} h_{\mu\nu} P_{\mu\nu} + \int d^3r \int d^3r' \frac{\rho(\vec{r})\rho(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} + E^{\mathrm{XC}}[\rho] \tag{5.1}$$

with the one-particle hamiltonian $h_{\mu\nu}$, the one-particle density matrix for closed-shell systems $P_{\mu\nu} = 2\sum_i c_{i\mu} c_{i\nu}$, coefficients of the canonical MOs $c_{p\mu}$, the one-particle density $\rho(\vec{r}) = \sum_{\mu\nu} P_{\mu\nu} \phi_\mu(\vec{r}) \phi_\nu(\vec{r})$ and the eXchange-Correlation (XC) functional $E^{\mathrm{XC}}[\rho]$. The MO coefficients are eigenvectors of the generalized eigenvalue problem (Kohn-Sham equation)[5]

$$\left(h_{\mu\nu} + V^{\mathrm{H}}_{\mu\nu}[\rho] + V^{\mathrm{XC}}_{\mu\nu}[\rho]\right) c_{i\mu} = \varepsilon_i S_{\mu\nu} c_{i\mu} \tag{5.2}$$

with the matrix elements of the Hartree potential

$$V^{\mathrm{H}}_{\mu\nu}[\rho] = \int d^3r \phi_\mu(\vec{r}) \phi_\nu(\vec{r}) \int d^3r' \frac{\rho(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} \tag{5.3}$$

and the XC potential $V^{\mathrm{XC}}_{\mu\nu}[\rho] = \int d^3r \phi_\mu(\vec{r}) \phi_\nu(\vec{r}) \frac{\delta E^{\mathrm{XC}}[n]}{\delta n(\vec{r})}$, the elements of the overlap matrix $S_{\mu\nu} = \int d^3r \phi_\mu(\vec{r}) \phi_\nu(\vec{r})$ and the MO energy ε_i of MO i. For DH functionals, the XC functional has the form

$$\begin{aligned} E^{\mathrm{XC}}[\rho] = {} & \alpha_{\mathrm{DFT,X}} E^{\mathrm{DFT,X}}[\rho] + \alpha_{\mathrm{HF,X}} E^{\mathrm{HF,X}}[\rho] \\ & + \alpha_{\mathrm{DFT,C}} E^{\mathrm{DFT,C}}[\rho] + \alpha_{\mathrm{WFT,C}} E^{\mathrm{WFT,C}}[\rho] \end{aligned} \tag{5.4}$$

with the DFT exchange functional $E^{\mathrm{DFT,X}}$, the Hartree-Fock exchange functional $E^{\mathrm{HF,X}}$, the DFT correlation functional $E^{\mathrm{DFT,C}}$, the WFT correlation functional $E^{WFT,C}$ and the corresponding scaling factors $\alpha_{\mathrm{DFT,X}}$,$\alpha_{\mathrm{HF,X}}$,$\alpha_{\mathrm{DFT,C}}$ and $\alpha_{\mathrm{WFT,C}}$, respectively. DFT XC functionals are usually local functionals written as integrals of a given function of the density, its gradient etc. depending on the complexity of the DFT part. The HF exchange functional is given by

$$E^{\mathrm{HF,X}}[\rho] = -\frac{1}{4} \sum_{\mu\nu\kappa\lambda} P_{\mu\nu} P_{\kappa\lambda} (\mu\kappa|\nu\lambda) \tag{5.5}$$

with the electron repulsion integral (ERI) in Mulliken notation

$$(\mu\kappa|\nu\lambda) = \int d^3r \int d^3r' \phi_\mu(\vec{r}) \phi_\kappa(\vec{r}) v(\left|\vec{r} - \vec{r}\,'\right|) \phi_\nu(\vec{r}\,') \phi_\lambda(\vec{r}\,') \tag{5.6}$$

with the interaction potential $v(r)$. Usually, the Coulomb potential $v(r) = \frac{1}{r}$ is employed, but for range-separated functionals or certain implementations of HF for periodic systems, it might be of a different kind.[94, 123, 144, 183, 216, 236,

241, 242]
Because the potentials in the Kohn-Sham equation depend on the unknown density, the Kohn-Sham equation has to be solved iteratively starting from an appropriate initial guess while recalculating densities and potentials subsequently until self-consistency (SCF approach). Under these conditions, the total energy functional is variational with respect to the density. Due to the high computational costs of WFT correlation methods, most DH functionals consider the wavefunction correlation (WFC) functional to be only a post-SCF correction such that the total energy functional is not variational.
The WFC functional is the energy expression of MP2, RPA or any other correlated method. Here, we will focus on the MP2 method given by

$$E^{\mathrm{MP2}}[\rho] = \sum_{ijab} (ia|jb) t_{ij}^{ab} \tag{5.7}$$

with the ERI in MO basis $(ia|jb) = \sum_{\mu\nu\kappa\lambda} c_{i\mu} c_{a\nu} c_{j\kappa} c_{b\lambda} (\mu\nu|\kappa\lambda)$ and the MP2 transition amplitude $t_{ij}^{ab} = \frac{2(ia|jb)-(ib|ja)}{\varepsilon_i+\varepsilon_j-\varepsilon_a-\varepsilon_b}$.
The Auxiliary Density Matrix Method (ADMM) for the acceleration of HF calculations relies on the relationship

$$E^{\mathrm{HF,X}}[\rho] = E^{\mathrm{HF,X}}[\hat{\rho}] + \left(E^{\mathrm{HF,X}}[\rho] - E^{\mathrm{HF,X}}[\hat{\rho}]\right) \tag{5.8}$$

$$\approx E^{\mathrm{HF,X}}[\hat{\rho}] + \left(E^{\mathrm{ADMM}}[\rho] - E^{\mathrm{ADMM}}[\hat{\rho}]\right). \tag{5.9}$$

with the auxiliary density $\hat{\rho}(\vec{r}) = \sum_{\hat{\mu}\hat{\nu}} \hat{P}_{\hat{\mu}\hat{\nu}} \hat{\phi}_{\hat{\mu}}(\vec{r}) \hat{\phi}_{\hat{\nu}}(\vec{r})$, the auxiliary density matrix $\hat{P}_{\hat{\mu}\hat{\nu}} = 2\sum_i \hat{c}_{i\hat{\mu}} \hat{c}_{i\hat{\nu}}$, the MO coefficients in the auxiliary basis set (ABS) $\hat{c}_{i\hat{\mu}}$ and the auxiliary basis functions $\hat{\phi}_{\hat{\mu}}(\vec{r})$. The ABS contains a smaller number of primitive basis functions and more localized basis functions to reduce the computational costs. The error in the density is accounted for by the ADMM exchange correction functional E^{ADMM}. Common choices are the PBE, the OPTX or the B88 exchange functionals.[90, 91, 247]
There are several constraints for the auxiliary density matrix and the auxiliary MO coefficients possible. For an overview about the different flavours of ADMM, we refer to the literature.[231, 248] In the following, we will consider the ADMM2 method only. The ADMM2 model requires the similarity of the MOs represented in both basis sets in a least-square sense such that the auxiliary MO coefficients

$\hat{c}_{i\hat{\mu}}$ are given by the expression

$$\hat{c}_{i\hat{\mu}} = \arg\min_{\tilde{c}_{i\hat{\mu}}} \int d^3r \left(\sum_{\hat{\mu}} \tilde{c}_{i\hat{\mu}} \hat{\phi}_{i\hat{\mu}}(\vec{r}) - \sum_{\mu} c_{i\mu} \phi_{\mu}(\vec{r}) \right)^2 . \tag{5.10}$$

Let $\mathbf{C}$ and $\hat{\mathbf{C}}$ be the matrix of MO coefficients in the PBS and the ABS, respectively. Both are related by $\hat{\mathbf{C}} = \mathbf{AC}$ with the projector matrix $\mathbf{A} = \hat{\mathbf{S}}^{-1}\mathbf{Q}$, the overlap matrix in the auxiliary basis $\hat{\mathbf{S}}$ and the overlap matrix between ABS and PBS $\mathbf{Q}$. The analogous relation for the density matrices is given by $\hat{\mathbf{P}} = \mathbf{APA}^T$.
A very successful approach to accelerate the calculation of ERIs is the Resolution-of-the-Identity (RI) approximation.[82, 83] It is based on the idea of fitting densities given by a product basis with a suitable ABS. In the following, we will focus on approximating ERIs with the Coulomb potential using a Coulomb metric, i.e. by fitting the Coulomb potential obtaining

$$(\mu\nu|\kappa\lambda) = \sum_{PQ} (\mu\nu|P)(P|Q)^{-1}(Q|\kappa\lambda) \tag{5.11}$$

with $(P|Q)^{-1}$ being the matrix element of the inverse of the matrix with elements $(P|Q)$. We will apply the RI approximation to the MP2 method resulting in a significant speed-up and reduction of memory requirements.[196, 249] The four-fold transformation of the ERIs in atomic orbital basis to MO basis is simplified because of

$$(ia|jb) = \sum_{P} B_{Pia} B_{Pjb} \tag{5.12}$$

with the intermediate

$$B_{Pia} = \sum_{Q} (ia|Q)(Q|P)^{-1/2}. \tag{5.13}$$

Because the total energy functional of DH functionals is not variational with respect to the density, we have to explicitly include the convergence constraints into the gradient calculation

$$c_{p\mu} S_{\mu\nu} c_{q\nu} = \delta_{pq} \tag{5.14}$$

$$c_{p\mu} F_{\mu\nu} c_{q\nu} = \varepsilon_p \delta_{pq}. \tag{5.15}$$

With the method of Lagrangian multipliers, we derive similarly to the RI-MP2 method the following expression for the gradients

$$E^{\mathrm{tot},x} = E^{\mathrm{SCF},x} + \sum_{Q\mu\nu} \Gamma_{Q\mu\nu}(\mu\nu|Q)^x + \sum_{PQ} \Gamma_{PQ}(P|Q)^x + \sum_{\mu\nu} \left(D_{\mu\nu} F^x_{\mu\nu} - W_{\mu\nu} S^x_{\mu\nu} \right), \tag{5.16}$$

with the MP2-relevant intermediates

$$\Gamma_{Q\mu\nu} = \sum_{ia} c_{i\mu} c_{a\nu} \Gamma_{Qia} \tag{5.17}$$

$$\Gamma_{PQ} = \sum_{iaR} B_{Pia} \Gamma_{Ria} (R|Q)^{-1/2} \tag{5.18}$$

$$\Gamma_{Qia} = \sum_{Pjb} (Q|P)^{-1/2} t^{ab}_{ij} B_{Pjb}, \tag{5.19}$$

the density matrix

$$D_{\mu\nu} = P_{\mu\nu} + P^{\mathrm{MP2}}_{\mu\nu} \tag{5.20}$$

$$P^{MP2}_{\mu\nu} = \sum_{pq} c_{p\mu} c_{q\nu} P^{\mathrm{MO,MP2}}_{pq} \tag{5.21}$$

$$P^{\mathrm{MO,MP2}}_{ij} = -\sum_{kab} t^{ab}_{ik} \frac{(ja|kb)}{\varepsilon_j + \varepsilon_k - \varepsilon_a - \varepsilon_b} \tag{5.22}$$

$$P^{\mathrm{MO,MP2}}_{ab} = \sum_{ijc} t^{ac}_{ij} \frac{(ib|jc)}{\varepsilon_i + \varepsilon_j - \varepsilon_b - \varepsilon_c} \tag{5.23}$$

$$P^{\mathrm{MO,MP2}}_{ia} = 0 \tag{5.24}$$

$$P^{MO,MP2}_{ai} = Z_{ia}, \tag{5.25}$$

the energy-weighted density matrix

$$W_{\mu\nu} = \sum_{ij} c_{i\mu} \varepsilon_i c_{i\nu} + \sum_{pq} c_{p\mu} c_{q\nu} W^{\mathrm{MO,MP2}}_{pq} \tag{5.26}$$

$$W^{\mathrm{MO,MP2}}_{ij} = \sum_{aQ} (ja|Q) \Gamma_{Qia} + \frac{1}{2}(\varepsilon_i + \varepsilon_j) P^{\mathrm{MO,MP2}}_{ij} + \frac{1}{2} \sum_{pq} P^{MO,MP2}_{pq} E^{(2)}_{pqij} \tag{5.27}$$

$$W^{\mathrm{MO,MP2}}_{ab} = \sum_{iQ} (ib|Q) \Gamma_{Qia} + \frac{1}{2}(\varepsilon_a + \varepsilon_b) P^{\mathrm{MO,MP2}}_{ab} \tag{5.28}$$

$$W_{ai}^{\mathrm{MO,MP2}} = \sum_{jQ}(ji|Q)\Gamma_{Qja} + \varepsilon_i P_{ai}^{\mathrm{MO,MP2}} \tag{5.29}$$

$$W_{ia}^{\mathrm{MO,MP2}} = 0, \tag{5.30}$$

the derivative of the Fock matrix

$$\begin{aligned} F_{\mu\nu}^x = h_{\mu\nu}^x + \frac{1}{2}P_{\kappa\lambda}(\mu\nu|\kappa\lambda)^x - \frac{\alpha_{HF,X}}{4}\hat{P}_{\hat{\kappa}\hat{\lambda}}(\hat{\mu}\hat{\lambda}|\hat{\kappa}\hat{\nu})^x + V_{\mu\nu}^{\mathrm{XC,DFT,x}}[\rho] \\ + V_{\mu\nu}^{\mathrm{ADMM,x}}[\rho] - V_{\hat{\mu}\hat{\nu}}^{\mathrm{ADMM,x}}[\hat{\rho}] + \left(V_{\hat{\mu}\hat{\nu}}^{\mathrm{HF,X}} - V_{\hat{\mu}\hat{\nu}}^{\mathrm{ADMM}}\right) A_{\hat{\mu}\mu} A_{\hat{\nu}\nu}^x P_{\mu\nu} \end{aligned} \tag{5.31}$$

and the derivative of the overlap matrix $S_{\mu\nu}^x$. Note that the second line of the derivative of the Fock matrix vanishes without ADMM. The matrix Z_{ia} is the solution of the Z-vector equation

$$\begin{aligned} \sum_{ia}\left(\delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + E_{aibj}^{(2)}\right) Z_{ia} = -2\sum_{aQ}(ba|Q)\Gamma_{Qja} + 2\sum_{iQ}(ij|Q)\Gamma_{Qib} \\ - \sum_{ac} P_{ac}^{\mathrm{MO,MP2}} E_{acbj}^{(2)} - \sum_{ik} P_{ik}^{\mathrm{MO,MP2}} E_{ikbj}^{(2)} \end{aligned} \tag{5.32}$$

with the second order energy kernel

$$E_{qprs}^{(2)} = 2\sum_{\mu\nu\kappa\lambda} c_{p\mu}c_{q\nu}c_{r\kappa}c_{s\lambda}\left(\frac{\partial^2 E^{\mathrm{SCF}}}{\partial P_{\mu\nu}\partial P_{\kappa\lambda}} + \frac{\partial^2 E^{\mathrm{SCF}}}{\partial P_{\nu\mu}\partial P_{\kappa\lambda}}\right) \tag{5.33}$$

$$\begin{aligned} \frac{\partial^2 E^{\mathrm{SCF}}}{\partial P_{\mu\nu}\partial P_{\kappa\lambda}} = 2(\mu\nu|\kappa\lambda) + \frac{\partial^2\left(E^{\mathrm{DFT,XC}} + E^{\mathrm{ADMM}}\right)}{\partial P_{\mu\nu}\partial P_{\kappa\lambda}} \\ - \left(\frac{\partial^2 E^{\mathrm{ADMM}}}{\partial \hat{P}_{\hat{\mu}\hat{\nu}}\partial P_{\hat{\kappa}\hat{\lambda}}} + \frac{1}{2}(\hat{\mu}\hat{\kappa}|\hat{\nu}\hat{\lambda})\right) A_{\mu\hat{\mu}}A_{\nu\hat{\nu}}A_{\kappa\hat{\kappa}}A_{\lambda\hat{\lambda}}. \end{aligned} \tag{5.34}$$

The Z-vector equations form a system of linear equations allowing a direct solution, however, these equations are commonly solved iteratively because the computational efforts of the exact solution scales as $\mathscr{O}\left(N^6\right)$ with N being a measure of system size.

We briefly describe the implementation in CP2K, where the integral calculation is performed separately before the contraction steps. First, the integrals are calculated batch-wise in the atomic orbital basis in small subgroups, each responsible for all integrals with a given subset of auxiliary functions. Contraction of two-

center and three-center integrals with the MO coefficients provides quantity B_{Pia} distributed among all processes with the RI index distributed among the subgroups and the virtual index distributed among the processes of a subgroup. In the following contraction step, all unique pairs of occupied orbitals are assigned to one of the aforementioned subgroups, each determining subsequently all four-center ERIs with the different orbital pairs according to equation (5.12). From the ERIs, the MP2 amplitudes and the energy contributions, and for a gradient calculation additionally the intermediate Γ_{Pia} and the diagonal elements of the density matrices can be calculated by exchanging the required data between the subgroups. In that fashion, only a subset of the MP2 amplitudes has to be available at the same time lowering the memory requirements. Communication costs are reduced by replicating parts of the integrals B_{Pia} within subsets of the subgroups. The non-diagonal elements of the density matrix are determined within a canonical reformulation. For further details, we refer the reader to the respective publications.[249–251]

Without ADMM, the projectors become the unity matrix, the ABS is equal to the PBS and the exchange correction functional is zero. In this case, we obtain the already known equations for DH gradients.[252] Without a DFT functional, we are left with HF MOs and the equations reduce to their MP2 equivalents.[253, 254] Thus, the implemented methodology is a generalization of the common DH gradient theory. Due to the consideration of the ADMM approximation, the equations are related to those by Kuman et al. for linear response theory.[255]

The open-source CP2K software package is based on the Quickstep algorithm.[130, 131] Its GPW approach employs a dual representation of orbitals and densities in a mixed Gaussian and plane-wave basis. The local nature of the Gaussian basis allows a linear-scaling implementation of the integral evaluation whereas the plane-wave basis enables a linear scaling evaluation of the Hartree potential with respect to system size. The plane-wave basis requires the use of pseudo potentials for which CP2K employs the Goedecker-Teter-Hutter form.[135]

In contrast to other implementations of correlated methods with localized basis sets like the periodic CRYSCOR project which extends CRYSTAL to correlated methods, CP2K does not exploit any symmetry within the system, does not localize the orbitals and the gradient implementation does not allow for all-electron calculations making it more suitable for dynamics simulations which inhibit symmetry at all.[192, 256]

5.3 Computational Details

The DFT and ADMM kernels were implemented into CP2K and are available with version 9.1.[130] We employed Libint 2.6 for the calculation of ERIs.[228] Calls to the routines of the BLAS/LAPACK/PBLAS/ScaLAPACK libraries were accelerated with cray-libsci_acc, version 20.10.1. Calls to PDGEMM exploited the COSMA library, version 2.5.1.[257] All calculations were run on the Piz Daint supercomputer with 12 MPI processes per node. Each node is equipped with a NVIDIA Tesla P100 GPU.
All periodic HF calculations were carried out with a truncated Coulomb operator. The cutoff radius is chosen half of the least distance between two equivalent atoms of neighboring cells (minimum image convention).[144]
For the present study, we will consider the DSD-PBEP86 functional which is a GGA-based DH functional with separately scaled same-spin and opposite-spin MP2 contributions and directly optimized dispersion correction.[258] In the GMTKN55 study, it was one of the best performing functionals with an excellent performance for all different use cases.[25] Because its HF part requires only the Coulomb potential and the required PBE exchange and P86 correlation functional are widely available, it is easily implemented in most electronic structure software.
We employed the following test systems: three molecular crystals NH_3, HCN and Ar, three systems with 64 water molecules of different density, the benzene crystal and the anatase crystal. The molecular crystals are the same as the ones used in previous studies.[70, 71] The three water-containing systems were taken from a training set of 1593 structures for a machine learning set of water.[259] We propose the benzene and the anatase crystals as benchmark systems for double hybrid and MP2 gradients and potentially also for other wave function based methods to be run on large scale parallel computers with thousands of compute nodes.[260]
The PBSs are correlation-consistent valence-only basis sets of double-zeta (DZ) and triple-zeta (TZ) quality.[71, 249] For Ar, we added the diffuse basis functions of the Dunning basis sets.[211] These are not required for the other systems as discussed in a preceding study.[71] The basis sets of Ti were taken from a previous study.[261] For the ADMM calculations, we used ABSs of DZ quality for all elements. Suitable RI basis sets of the elements H, C, N and O for the MP2 calculations were taken from Del Ben et al., those for Ar were optimized according to Weigend et al. and the non-augmented basis sets published in Stein et al.[71, 196, 249] The pseudopotentials were optimized for the PBE0 method. All DH calculations used PBE orbitals as initial guess. The Schwarz screening parameter was set to values of 10^{-7} without ADMM and 10^{-10} with ADMM, and derivative

integrals use a corresponding 10^{-7} screening parameter for all calculations.

5.4 Accuracy

In the left panel of figure 5.1, we present the errors introduced by the ADMM approximation for specific energy differences of the test systems. For the three water systems, we use the differences in total energies. For the molecular crystals, we determined the cohesive energies corrected by the basis set superposition error (BSSE).[213] With a TZ quality PBS, we find an error of at most 1.5 $mE_h \approx 1\ kcal/mol$ (grey lines in the diagram). Thus, the general accuracy of the double-hybrid functional is preserved with ADMM.
On the right panel of figure 5.1, the corresponding absolute relative errors are shown. We find that the energy differences of the water systems have errors of 20-100 % whereas the cohesive energies of the molecular crystals have a much lower relative error. The rather small energy differences between the medium- and low-density structures cause high relative errors but the absolute error is still small. The cohesive energies of the molecular crystals are much larger with Ar having the smallest cohesive energy of 3.8 mE_h with DSD-PBEP86-D3 (experimental value: 2.9 mE_h).[207]
In figure 5.2, we report the mean absolute ADMM errors of the forces and stress tensor components. For the forces, we find an error of at most 0.1 mE_h/a_0. The average force component of the water system at medium density is 29.3 mE_h/a_0 providing an average relative error of 0.3 %. For the stress tensor, we found an error of usually not more than 0.1 *GPa* corresponding to a relative error of at most 10 % for the diagonal elements. Due to symmetry reasons, the forces on the argon atoms are zero in all cases.

5.5 Performance

Figure 5.3 reports the computational timing of all methods and test systems. Except for Ar, we find for the given setting no reduction in the computational costs with PBSs of DZ quality, whereas ADMM is able to reduce the timings by 20-70 % for TZ PBS. For Ar, we observe an improvement of about 50 % and 80 % with PBSs of DZ and TZ quality, respectively. These findings suggest for all systems except from Ar with a DZ basis set that the similar quality of the ABSs and the PBSs and the tighter threshold of the Schwarz criterion of the ADMM method

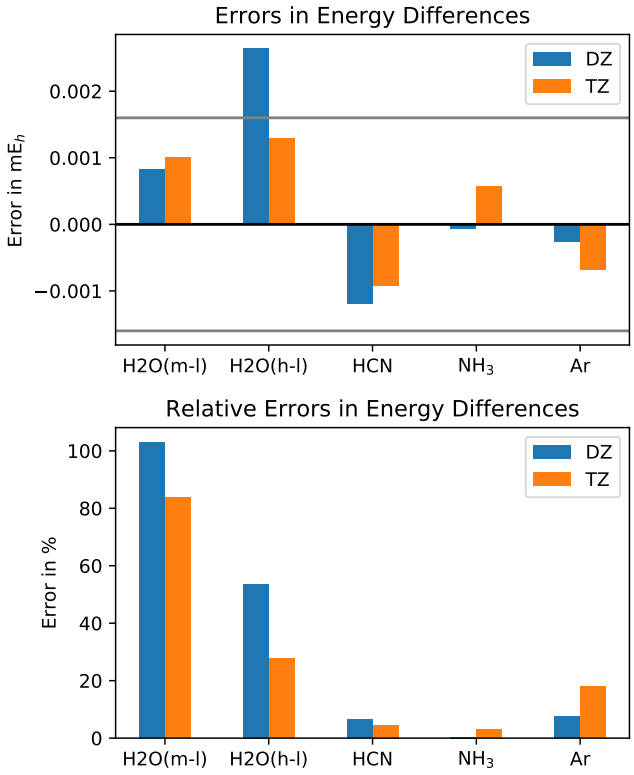

Figure 5.1: ADMM errors for energy differences. For the water systems, we employ the difference in total energies of medium and low density structure (H2O(m-l)) and high and low density structure (H2O(h-l)) per water molecule. For the molecular crystals, we report the cohesive energies corrected by the basis set superposition error (BSSE). Upper panel: signed errors. Lower panel: absolute relative errors.

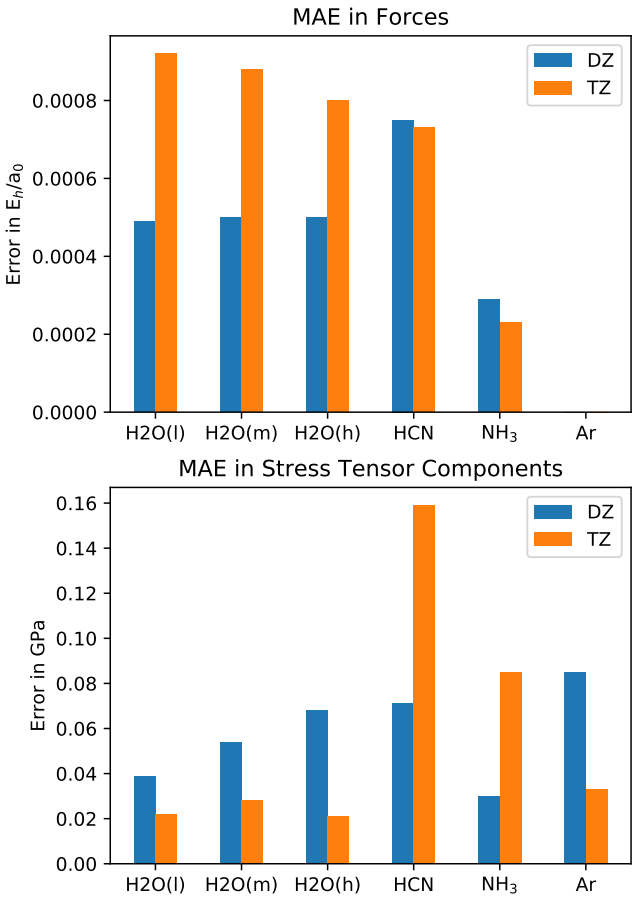

Figure 5.2: Mean absolute ADMM errors for forces and stress tensor components. "H2O(l)", "H2O(m)" and "H2O(h)" refer to the water-containing systems of low, medium and high density, respectively.

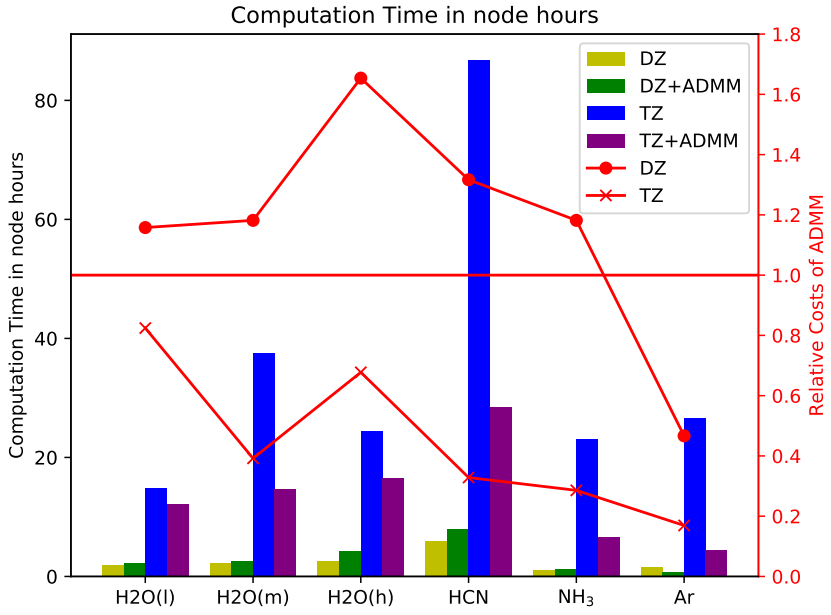

Figure 5.3: Computational time in node hours for different systems, PBSs and methods. For better comparison, we report also relative timings of ADMM calculations. (Augmented) basis sets of double-zeta (DZ) and triple-zeta (TZ) quality were employed. "+ADMM" refers to the additional application of the ADMM approximation.

result in an increase in the computational timings because of an increased number of significant integrals. In contrast to that, we employed augmented PBSs for Ar which result in a significantly larger number of electron repulsion integrals than for the non-augmented ABS.
In the left panel of figure 5.4, we present a breakdown of the computational timings of the different methods. We find that the computational costs of MP2 are independent of the presence of the ADMM approximation. This has to be expected because the actual MP2 energy calculation and the assembly of the quantities of the MP2 force calculation are not directly affected by the ADMM approximation. Thus, the reduction or increase in the computational costs is caused by the solution of the Kohn-Sham equation and the Z-vector equations. Excluding the Ar system and concentrating on the not directly to MP2 related costs, we find the following behavior. For DZ PBS the relative costs are 20-44 % without ADMM and 31-63 % with ADMM. However, for the TZ PBS the relative costs are getting smaller when using ADMM, namely, dropping from 19-75 % to 4-12 %. The computational costs of the MP2 energy calculation are at most as high as the computational costs of the integral calculation. We conclude that the quintical scaling limit of the MP2 method was not yet achieved in the examples and further optimizations of the integral calculation step should be targeted.
If we consider the different contributions to the computational costs not directly related to MP2, as presented in the right panel of figure 5.4, we observe that the non-MP2 part is dominated by the calculation of the four center integrals and their derivatives. Thus, the costs of that part is reduced due to ADMM as anticipated. The overhead of the ADMM projector matrix and the calculation of the exchange correction is negligible in all calculations.
In figure 5.5, we report the total number of calculated integrals. Apart from the DZ basis sets, ADMM reduces the number of significant integrals by at least one order of magnitude, in the case of Ar and the TZ basis set by two orders of magnitude.
In figure 5.6, the GPU usage according to the Slurm reference is shown. Slurm is the job scheduling system employed by the CSCS supercomputing centre hosting the supercomputer on which all calculations were run (see its manual for further information).[262] According to the Slurm manual, it provides the percentage of computational time which is spent on GPU. We observe an increase in GPU usage in accordance with a decrease in computational time. For the ADMM calculations with a DZ basis set, the GPU usage drops from 40-50 % to 25-45 %. For the TZ PBS the GPU usage increase from 20-60 % to 60-75 % when employing ADMM. Finally, for Ar with a TZ PBS we even find an increase from 14 % to 86 %. These

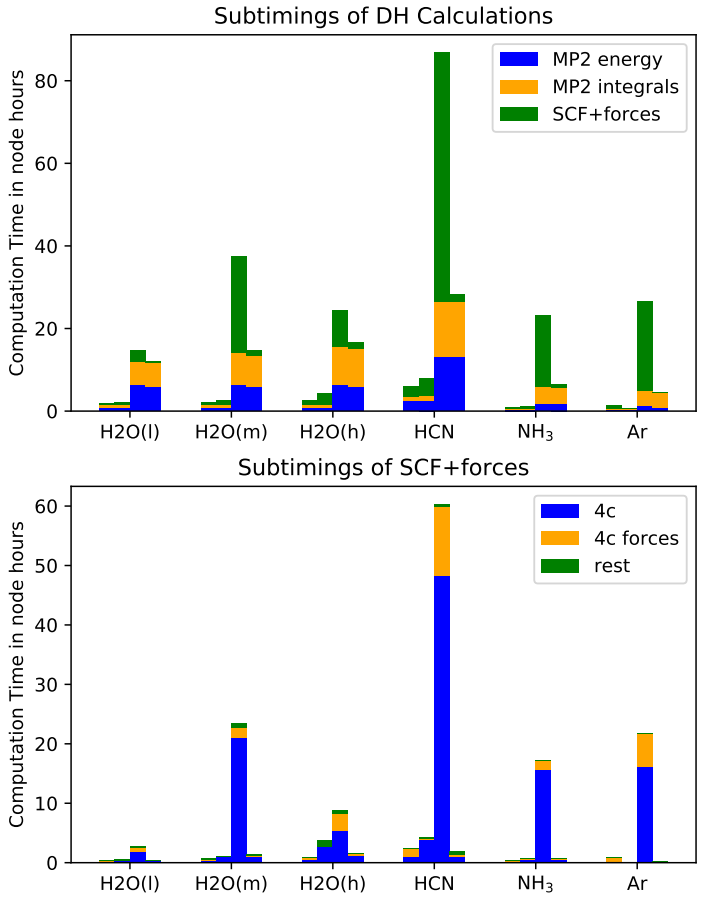

Figure 5.4: Subdivision of the total computation time (left panel) and of the SCF+forces part only (right panel). "MP2 energy" refers to the quintically scaling computation steps ("mp2_ri_gpw_compute_en" of the CP2K timing report). "MP2 integrals" refers to the computation of the integrals and their derivatives (sum of "mp2_ri_gpw_compute_in" and "calc_ri_mp2_nonsep" of the CP2K timing report). "SCF+forces" refers to the difference of the total CP2K run time and the aforementioned parts. "4c" and "4c forces" refer to the timings of the routines "integrate_four_center" and "derivatives_four_center", respectively, of the CP2K timing report. "rest" is the difference of "SCF+forces" and the compute time of the HF integrals. Each group of bars represents in the given order the timing information of DZ without ADMM, DZ with ADMM, TZ without ADMM and TZ with ADMM.

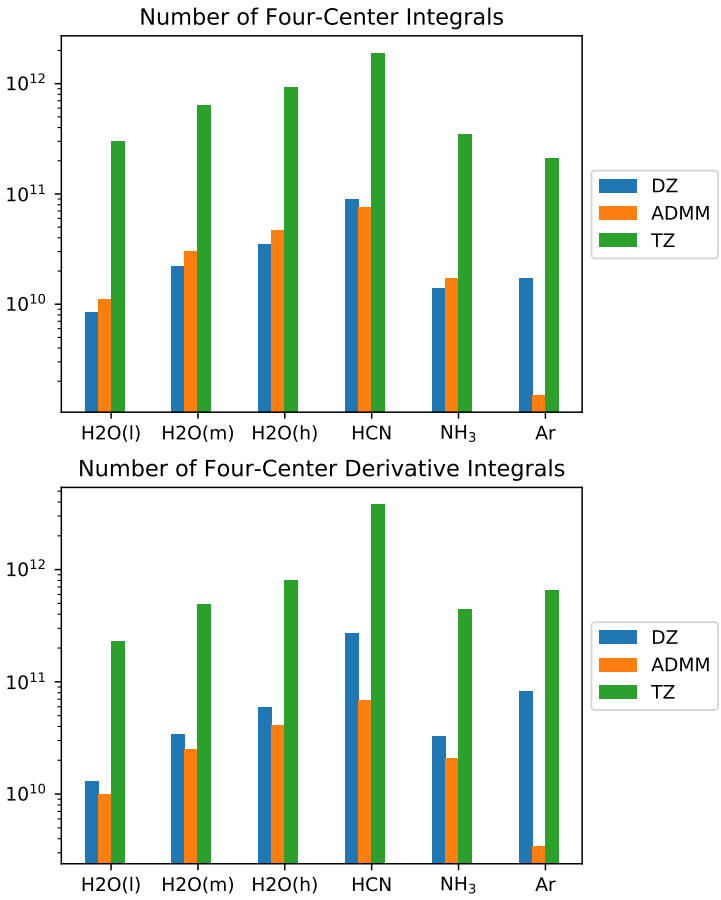

Figure 5.5: Number of four-center integrals (left panel) and four-center derivative integrals (right panel) calculated for different systems, PBSs and methods. The number of integrals calculated with the ADMM is independent of the PBS.

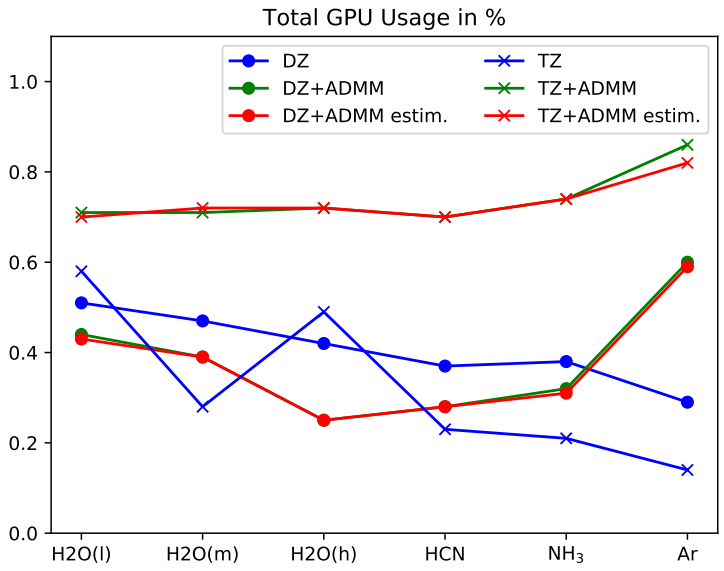

Figure 5.6: GPU usage averaged over all GPUs and estimates according to the Slurm reports for different systems and basis sets. Estimates based on the GPU usage without ADMM and the total computational timings with and without ADMM.

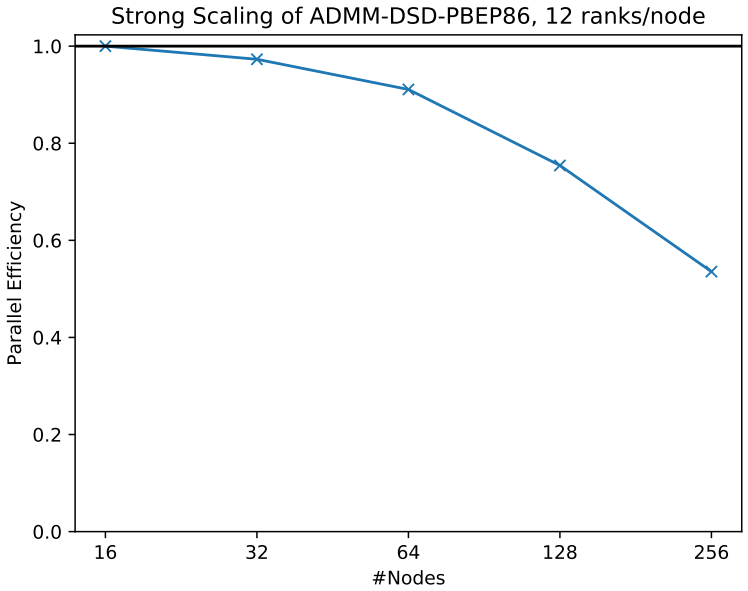

Figure 5.7: Strong scaling plot for the ADMM-DSD-PBEP86-D3 functional on Piz Daint with 12 MPI processes per node.

numbers are in agreement with the fact that the exchange Fock matrix calculation is not GPU accelerated.
Finally, we show in figure 5.7 the strong scaling behaviour of the ADMM-accelerated DH calculations. We used 64 water molecules at ambient density and 12 ranks per node. The smallest possible number of nodes is determined by the memory requirements of MP2 being 115 GB. Due to the parallel memory overhead of the Quickstep implementation, we observe a peak memory per process of 3878 MB resulting in total memory costs of about 727 GB, which corresponds to at least 12 nodes of the Piz Daint supercomputer. With 256 nodes, we observe a parallel efficiency of 55 % which provides a reasonable scaling of approximately one order of magnitude with respect to the amount of computational resources.

5.6 Large Benchmark systems

The exponential increase in computational power of modern computers make calculations with DH functionals for large systems more and more feasible.[263] We

Table 5.1: Geometrical parameters of the benchmark systems. See the input files in the Materialscloud for further information.[175]

System	super cell size	$a;b;c$ (Å)	$\alpha;\beta;\gamma$ (°)	References
benzene	2x2x2	7.398; 9.435; 6.778	90; 90; 90	[260]
anatase	1x1x1	9.674; 9.826; 15.125	74.9; 75.9; 78.3	

Table 5.2: Information about the performance of the benchmark systems on the Piz Daint supercomputer. N, o, n, v, n_{ADMM}, n_{RI} refer to the number of occupied orbitals, unoccupied orbitals, basis functions, ADMM auxiliary basis functions, RI basis functions, respectively. *mem* provides the minimal required memory for the MP2 gradient calculation according to [250]. $FLOPs$ refers to the number of floating point operations of the quintical scaling steps. t provides the computational time in node hours.

System	N	o	n	v	n_{ADMM}	n_{RI}	*mem*	$FLOPs$	t
benzene	384	480	8256	7776	3456	20352	1703 GB	$1.7 \cdot 10^{18}$	515
anatase	144	576	5424	4848	2928	15360	961 GB	$7.2 \cdot 10^{17}$	354

will demonstrate the current state with calculations carried out for large benchmark systems performed efficiently on compute nodes with 1024 GPUs. We present two different benchmark systems for which we calculated energies and forces: the benzene crystal and the anatase crystal. The geometrical parameters, the number of basis functions and the number of floating point operations of the quintically scaling steps for each test system are summarized in table 5.1. The computational timings and some performance related parameters are summarized in table 5.2.

Anatase is a highly investigated material in the context of photocatalysis.[264–268] Due to its large band gap, it can be tackled with perturbation theory. From a computational point of view, the inclusion of semicore states and high-angular quantum number polarization functions and the large RI basis set (in our case up to $l = 6$) provide a computationally demanding system.

Table 5.3: Relative timings and GPU usage of the benchmark systems on 1024 nodes of the Piz Daint supercomputer. See figure 5.4 for definitions.

System	HF	MP2 integrals	MP2 energy	GPU usage
benzene	9.3 %	19 %	59 %	38 %
anatase	10 %	17 %	62 %	29 %

Finally, we discuss the importance of the different calculation parts with respect to the total computational time. In table 5.3, we report the relative computational costs of the two calculations. The quintical scaling steps dominate the computational cost with a relative contribution of roughly 60 %. The HF calculation still contributes to 10 % of the costs which is mostly related to the limited scaling of the underlying communication algorithm.[144]

5.7 Discussion

Because the calculation of the HF kernel still consumes a significant amount of computation time of the DH gradient calculation, a reduction of its costs is of general interest. The ADMM approximation was developed to accelerate HF calculations while conserving the accuracy. We find that ADMM provides a significant reduction of the total computational time of DH gradient calculations with TZ PBSs. Besides the size of the ADMM basis, the most important parameter determining the performance of HF calculations is the Schwarz screening threshold. We increased this threshold for standard calculations as much as possible without significant loss of accuracy whereas we kept a tighter value for ADMM. This means that the performance benefit of ADMM is actually the lowest possible value because of the unoptimized Schwarz threshold. For water systems, we found that a threshold of order 10^{-7} can be used for ADMM. In practice, however, an optimized Schwarz screening threshold could be used to push ADMM accuracy without increasing the overall costs of the DH calculations.
The accuracy of ADMM determined here is sufficient for Machine Learning (ML) approaches. The ML fitting procedure removes potential numerical noise from the training data making it more robust. The applicability to other kinds of applications depends on the magnitude of the expected energy differences. For a higher accuracy, we recommend a slightly larger ABS, the actual size has to be adjusted to the available resources and the required accuracy. For molecular crystals with sufficiently large cohesive energies, the accuracy is considered sufficient because the employed functional introduces a larger uncertainty. Thus, for molecular dynamic simulations with the same kinds of systems, the presented methodology will provide suitable accuracy.
ADMM could be combined with RI-HF exchange or other methods like the COS approach to accelerate HF exchange calculations further.[230, 269, 270] This combination of approaches has the potential to further reduce the time for HF and response calculations. However, as the HF part of the DH gradient calcula-

tions has already been reduced significantly by ADMM, no major further relative improvements can be expected. The main advantage of such a combination might be in two other areas: RI-HFX shifts the computational most demanding part from the calculation of ERIs to matrix multiplications and tensor contractions. This allows efficient algorithms with a much smaller memory footprint and makes GPU acceleration accessible. We will investigate this approach and will report results in the future.

5.8 Conclusions

We presented three achievements: an efficient implementation of forces and stress tensors of DH functionals in the CP2K software package, application of ADMM HF and response calculations to reduce the computational costs of HF exchange in the context of DH functionals energy and gradients and establishment of benchmark systems for MP2-based functionals in condensed phase. Our approach exploits modern GPUs efficiently and reduces the prefactors of the lower scaling steps significantly. The methods allow to efficiently use large and diffuse basis functions in the condensed phase.
With the provided methodology, we open the way to perform geometry optimizations, molecular dynamics simulations and to calculate training data for machine learning approaches on the double hybrid functional level of systems with a finite band gap. The main advantage of ADMM is its universal applicability and its scaling behavior for condensed systems. Further, it allows an efficient treatment of diffuse basis functions which would otherwise increase the computational costs of HF significantly. For the calculation of DH gradients, ADMM allows to reduce the computational costs of the HF kernel to at most 10 % while conserving chemical accuracy.
The main drawback of the current implementation is the still quintical scaling of the underlying MP2 method. To overcome this issue, we will have to employ either low-scaling variants of MP2 or a computationally lower scaling method like RPA or OS-MP2.

Chapter 6

Conclusions and Outlook

This thesis dealt with efficient DHF calculations in condensed phase. The underlying correlation method poses a challenge because of its high computational costs and its slow convergence of properties with respect to the basis set size.
We started with the implementation of DHFs with range-separated correlation DFs. The removal of the singularity at the origin of the Coulomb potential significantly improves the basis set convergence of the DHFs. We implemented the necessary integrals into the GPW approach of RI-MP2 and RI-RPA to enable efficient integral calculations. The possibility to optimize pseudopotentials allows accurate calculations with range-separated functionals including any element of the periodic table. Albeit our methodology proves to be efficient with the ability to reduce the computational overhead of the correlated calculations, there is only one LDA-based short-range correlation DF available. This functional might provide a reasonable performance, there is still significant room for improvement with a GGA-correction. The implementation of the GGA-based DF is tedious and we will wait for their availability in the LibXC library.
The second part was concerned with benchmarking DHFs in condensed phase. Because most DHFs are fitted to empirical data from molecular systems, it raised the question whether these DFs are actually suitable for condensed phase systems. Furthermore, the amount of studies employing DHFs in periodic systems is rather scarce compared to molecular systems such that we provided a study discussing these questions in more detail. Our study suggests that the excellent performance of DHFs for molecular systems is indeed transferable to periodic systems. Benchmark studies with a larger variety of DFs and systems or geometries optimized by the employed DF are possible next steps.
In the third part, we discussed an efficient implementation of gradients with MP2-

based DHFs. The ADMM approximation allows a significant reduction of the HF-related computational costs and shows an excellent performance on supercomputing devices with 1,000 GPUs. It enables the computationally efficient calculation of training data for machine-learning approaches or dynamics simulations on the DHF level.
We suggest further improvements to the collocation and integration routine of the GPW approach. Moreover, the recently implemented gradients of the RI-HF approach as an alternative to the in this context newly established ADMM approximation should be investigated with respect to performance and accuracy.
The given methodology can be extended easily to the already available quartically scaling RI-RPA and Laplace-RI-SOS-MP2 methods which allow more cost-efficient DHF calculations and reduce the costs by one order of magnitude. Especially RPA-based DHDFs are of interest due to their inclusion of higher order excitations required for more accurate results in condensed matter.
Furthermore, an implementation of a low-scaling RI-MP2 method to increase the feasible system size of MP2 and MP2-based DHF calculations within CP2K comparable to the cubically-scaling RPA approach. An extension of the machinery to meta-GGA-based DHFs enables more accurate DHFs. Testing of the already available RPA-AXK method and implementation of the RPA-SOSEX method into CP2K is another possible next step.
A common drawback of MP2-based DHFs are metallic and strongly correlated systems because of their (nearly) vanishing band gap. Due to its foundation in perturbation theory, MP2 is not capable of describing these systems and the available implementation of RPA in CP2K is also restricted to systems with a non-zero band gap. Further developments can tackle these issues with Coupled Cluster (and corresponding DHFs), transcorrelated methods, an advanced RPA implementation or multi-reference methods.

Bibliography

(1) London, F. *Z. Phys.* **1930**, *63*, 245–279.

(2) Stone, A., *The Theory of Intermolecular Forces*; Oxford University Press: 2013.

(3) Piela, L. In *Ideas of Quantum Chemistry (Second Edition)*, Piela, L., Ed., Second Edition; Elsevier: Oxford, 2014, pp 793–882.

(4) Rantih, N. K.; Mulyani, S.; Widhiyanti, T. *Journal of Physics: Conference Series* **2019**, *1157*, 042029.

(5) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(6) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(7) Goedecker, S. *Rev. Mod. Phys.* **1999**, *71*, 1085–1123.

(8) Bowler, D. R.; Miyazaki, T. *Rep. Prog. Phys.* **2012**, *75*, 036503.

(9) Ordejón, P.; Drabold, D. A.; Martin, R. M.; Grumbach, M. P. *Phys. Rev. B* **1995**, *51*, 1456–1476.

(10) Shang, H.; Xiang, H.; Li, Z.; Yang, J. *Int. Rev. Phys. Chem.* **2010**, *29*, 665–691.

(11) VandeVondele, J.; Borštnik, U.; Hutter, J. *J. Chem. Theory Comput.* **2012**, *8*, 3565–3573.

(12) Niklasson, A. M. N.; Tymczak, C. J.; Challacombe, M. *J. Chem. Phys.* **2003**, *118*, 8611–8620.

(13) Yang, W.; Lee, T.-S. *J. Chem. Phys.* **1995**, *103*, 5674–5678.

(14) Daw, M. S. *Phys. Rev. B* **1993**, *47*, 10895–10898.

(15) Liang, W.; Saravanan, C.; Shao, Y.; Baer, R.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 4117–4125.

(16) Niklasson, A. M. N. *Phys. Rev. B* **2002**, *66*, 155115.

(17) Li, X.-P.; Nunes, R. W.; Vanderbilt, D. *Phys. Rev. B* **1993**, *47*, 10891–10894.

(18) Nunes, R. W.; Vanderbilt, D. *Phys. Rev. B* **1994**, *50*, 17611–17614.

(19) Galli, G.; Parrinello, M. *Phys. Rev. Lett.* **1992**, *69*, 3547–3550.

(20) Mauri, F.; Galli, G. *Phys. Rev. B* **1994**, *50*, 4316–4326.

(21) Kim, J.; Mauri, F.; Galli, G. *Phys. Rev. B* **1995**, *52*, 1640–1648.

(22) Palser, A. H. R.; Manolopoulos, D. E. *Phys. Rev. B* **1998**, *58*, 12704–12711.

(23) Kim, J.; Jung, Y. *Int. J. Quantum Chem.* **2016**, *116*, 563–568.

(24) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.

(25) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

(26) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

(27) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

(28) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(29) Caldeweyher, E.; Bannwarth, C.; Grimme, S. *J. Chem. Phys.* **2017**, *147*, 034112.

(30) Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.

(31) Vydrov, O. A.; Van Voorhis, T. *Phys. Rev. Lett.* **2009**, *103*, 063004.

(32) Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2009**, *130*, 104105.

(33) Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.

(34) Schrödinger, E. *Ann. Phys.* **1926**, *384*, 361–376.

(35) Schrödinger, E. *Ann. Phys.* **1926**, *386*, 109–139.

(36) Schrödinger, E. *Ann. Phys.* **1926**, *384*, 489–527.

(37) Fock, V. *Z. Phys.* **1930**, *61*, 126–148.

(38) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

(39) Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

(40) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968–5975.

(41) Wilhelm, J.; Seewald, P.; Del Ben, M.; Hutter, J. *J. Chem. Theory Comput.* **2016**, *12*, 5851–5859.

(42) Kaltak, M.; Klimeš, J.; Kresse, G. *J. Chem. Theory Comput.* **2014**, *10*, 2498–2507.

(43) Moussa, J. E. *J. Chem. Phys.* **2014**, *140*, 014107.

(44) Kállay, M. *J. Chem. Phys.* **2015**, *142*, 204105.

(45) Schurkus, H. F.; Ochsenfeld, C. *J. Chem. Phys.* **2016**, *144*, 031101.

(46) Kaltak, M.; Klime š, J. ř.; Kresse, G. *Phys. Rev. B* **2014**, *90*, 054115.

(47) Neuhauser, D.; Rabani, E.; Baer, R. *J. Comput. Chem. Lett.* **2013**, *4*, 1172–1176.

(48) Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **2001**, *114*, 661–681.

(49) Schütz, M.; Manby, F. R. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3349–3358.

(50) Saitow, M.; Becker, U.; Riplinger, C.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2017**, *146*, 164105–164105.

(51) Bangerter, F. H.; Glasbrenner, M.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2021**, *17*, 211–221.

(52) Baudin, P.; Ettenhuber, P.; Reine, S.; Kristensen, K.; Kjærgaard, T. *J. Chem. Phys.* **2016**, *144*, 054102.

(53) Freeman, D. L. *Phys. Rev. A* **1977**, *15*, 5512–5521.

(54) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(55) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

(56) Grüneis, A.; Marsman, M.; Harl, J.; Schimka, L.; Kresse, G. *J. Chem. Phys.* **2009**, *131*, 154115.

(57) Bates, J. E.; Furche, F. *J. Chem. Phys.* **2013**, *139*, 171103.

(58) Schwartz, C. *Phys. Rev.* **1962**, *126*, 1015–1019.

(59) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437–446.

(60) Klopper, W.; Kutzelnigg, W. *Chem. Phys. Lett.* **1987**, *134*, 17–22.

(61) Klopper, W.; Röhse, R.; Kutzelnigg, W. *Chem. Phys. Lett.* **1991**, *178*, 455–461.

(62) Noga, J.; Kutzelnigg, W.; Klopper, W. *Chem. Phys. Lett.* **1992**, *199*, 497–504.

(63) Hehn, A.-S.; Klopper, W. *J. Chem. Phys.* **2013**, *138*, 181104.

(64) Franck, O.; Mussard, B.; Luppi, E.; Toulouse, J. *J. Chem. Phys.* **2015**, *142*, 074107.

(65) Ángyán, J. G.; Gerber, I. C.; Savin, A.; Toulouse, J. *Phys. Rev. A* **2005**, *72*, 012510.

(66) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

(67) Goerigk, L.; Grimme, S. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2014**, *4*, 576–600.

(68) Del Ben, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2012**, *8*, 4177–4188.

(69) Sharkas, K.; Toulouse, J.; Maschio, L.; Civalleri, B. *J. Chem. Phys.* **2014**, *141*, 044105.

(70) Sansone, G.; Civalleri, B.; Usvyat, D.; Toulouse, J.; Sharkas, K.; Maschio, L. *J. Chem. Phys.* **2015**, *143*, 102811.

(71) Stein, F.; Hutter, J.; Rybkin, V. V. *Molecules* **2020**, *25*.

(72) Stein, F.; Hutter, J. *J. Chem. Phys.* **2022**, accepted.

(73) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry - Introduction to Advanced Electronic Structure Theory*; Courier Corporation: New York, 1996.

(74) Atkins, P. W.; Friedman, R. S., *Molecular Quantum Mechanics*; Oxford University Press: New York, 1996.

(75) Helgaker, T., *Molecular Electronic-Structure Theory*; Wiley: Chichester New York, 2000.

(76) Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. *Chem. Rev.* **2012**, *112*, 108–181.

(77) Park, J. W.; Al-Saadon, R.; MacLeod, M. K.; Shiozaki, T.; Vlaisavljevich, B. *Chem. Rev.* **2020**, *120*, 5878–5909.

(78) Khedkar, A.; Roemelt, M. *Phys. Chem. Chem. Phys.* **2021**, *23*, 17097–17112.

(79) White, S. R.; Martin, R. L. *J. Chem. Phys.* **1999**, *110*.

(80) Schrödinger, E. *Ann. Phys.* **1926**, *385*, 437–490.

(81) McGibbon, R. T.; Taube, A. G.; Donchev, A. G.; Siva, K.; Hernández, F.; Hargus, C.; Law, K.-H.; Klepeis, J. L.; Shaw, D. E. *J. Chem. Phys.* **2017**, *147*, 161725.

(82) Whitten, J. L. *J. Chem. Phys.* **1973**, *58*, 4496–4501.

(83) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 3396–3402.

(84) Vahtras, O.; Almlöf, J.; Feyereisen, M. *Chem. Phys. Lett.* **1993**, *213*, 514–518.

(85) Mezei, P. D.; Kállay, M. *J. Chem. Theory Comput.* **2019**, *15*, 6678–6687.

(86) Perdew, J. P. and Schmidt, K. and Van Doren, V. *AIP Conf. Proc.* **2001**, *577*, 1–20.

(87) Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.

(88) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048–5079.

(89) Perdew, J. P.; Wang, Y. *Phys. Rev. A* **1992**, *45*, 13244–13249.

(90) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(91) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(92) Van Voorhis, T.; Scuseria, G. E. *J. Chem. Phys.* **1998**, *109*, 400–410.

(93) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(94) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207–8215.

(95) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(96) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(97) Karton, A.; Tarnopolsky, A.; Lamère, J.-F.; Schatz, G. C.; Martin, J. M. L. *J. Comput. Chem. A* **2008**, *112*, 12868–12886.

(98) Langreth, D.; Perdew, J. *Solid State Commun.* **1975**, *17*, 1425–1429.

(99) 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 43–53.

(100) Furche, F.; Van Voorhis, T. *J. Chem. Phys.* **2005**, *122*, 164106.

(101) Furche, F. *J. Chem. Phys.* **2001**, *114*, 5982–5992.

(102) Furche, F. *J. Chem. Phys.* **2008**, *129*, 114105.

(103) Eshuis, H.; Yarkony, J.; Furche, F. *J. Chem. Phys.* **2010**, *132*, 234114.

(104) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Comput. Chem. A* **2004**, *108*, 4786–4791.

(105) Goll, E.; Werner, H.-J.; Stoll, H. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3917–3923.

(106) Goll, E.; Werner, H.-J.; Stoll, H.; Leininger, T.; Gori-Giorgi, P.; Savin, A. *Chem. Phys.* **2006**, *329*, 276–282.

(107) Sharp, R. T.; Horton, G. K. *Phys. Rev.* **1953**, *90*, 317–317.

(108) Talman, J. D.; Shadwick, W. F. *Phys. Rev. A* **1976**, *14*, 36–40.

(109) Krieger, J. B.; Li, Y.; Iafrate, G. J. *Phys. Rev. A* **1992**, *45*, 101–126.

(110) Kozuch, S.; Gruzman, D.; Martin, J. M. L. *J. Comput. Chem. C* **2010**, *114*, 20801–20808.

(111) Kozuch, S.; Martin, J. M. L. *J. Comput. Chem.* **2013**, *34*, 2327–2344.

(112) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.

(113) Tarnopolsky, A.; Karton, A.; Sertchook, R.; Vuzman, D.; Martin, J. M. L. *J. Comput. Chem. A* **2008**, *112*, 3–8.

(114) Brémond, E.; Ciofini, I.; Sancho-García, J. C.; Adamo, C. *Acc. Chem. Res.* **2016**, *49*, 1503–1513.

(115) Alipour, M. *Chem. Phys. Lett.* **2017**, *684*, 423–426.

(116) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I. *J. Chem. Theory Comput.* **2010**, *6*, 127–134.

(117) Souvi, S. M. O.; Sharkas, K.; Toulouse, J. *J. Chem. Phys.* **2014**, *140*, 084107.

(118) Grimme, S.; Steinmetz, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20926–20937.

(119) Toulouse, J.; Gerber, I. C.; Jansen, G.; Savin, A.; Ángyán, J. G. *Phys. Rev. Lett.* **2009**, *102*, 096404.

(120) Toulouse, J.; Zhu, W.; Ángyán, J. G.; Savin, A. *Phys. Rev. A* **2010**, *82*, 032502.

(121) Brémond, É.; Savarese, M.; Pérez-Jiménez, Á. J.; Sancho-García, J. C.; Adamo, C. *J. Chem. Theory Comput.* **2018**, *14*, 4052–4062.

(122) Benighaus, T.; DiStasio, R. A.; Lochan, R. C.; Chai, J.-D.; Head-Gordon, M. *J. Comput. Chem. A* **2008**, *112*, 2702–2712.

(123) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2009**, *131*, 174105.

(124) Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2018**, *148*, 241736.

(125) Ciofini, I.; Adamo, C.; Chermette, H. *Chem. Phys.* **2005**, *309*, 67–76.

(126) Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2014**, *140*, 18A513.

(127) Bao, J. L.; Gagliardi, L.; Truhlar, D. G. *J. Comput. Chem. Lett.* **2018**, *9*, 2353–2358.

(128) Görling, A.; Levy, M. *Phys. Rev. B* **1993**, *47*, 13105–13113.

(129) Görling, A.; Levy, M. *Phys. Rev. A* **1994**, *50*, 196–204.

(130) Kühne, T. D. et al. *J. Chem. Phys.* **2020**, *152*, 194103.

(131) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.

(132) Borstnik, U.; VandeVondele, J.; Weber, V.; Hutter, J. *Parallel Comput.* **2014**, *40*.

(133) The CP2K Developers Group DBCSR: Distributed Block Compressed Sparse Row matrix library https://github.com/cp2k/dbcsr (accessed 12/15/2022).

(134) Lippert, G.; Hutter, J.; Parrinello, M. *Mol. Phys.* **1997**, *92*, 477–488.

(135) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703–1710.

(136) Lippert, G.; Hutter, J.; Parrinello, M. *Theor. Chem. Acc.* **1999**, *103*, 124–140.

(137) Leininger, T.; Stoll, H.; Werner, H.-J.; Savin, A. *Chem. Phys. Lett.* **1997**, *275*, 151–160.

(138) Mussard, B.; Reinhardt, P.; Ángyán, J. G.; Toulouse, J. *J. Chem. Phys.* **2015**, *142*, 154123.

(139) Hedegård, E. D.; Knecht, S.; Kielberg, J. S.; Jensen, H. J. A.; Reiher, M. *J. Chem. Phys.* **2015**, *142*, 224108.

(140) Hill, R. N. *J. Chem. Phys.* **1985**, *83*, 1173–1196.

(141) Goddard, B. D. *SIAM J. Numer. Anal.* **2009**, *41*, 77–116.

(142) Gori-Giorgi, P.; Savin, A. *Phys. Rev. A* **2006**, *73*, 032506.

(143) Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2016**, *144*, 214110.

(144) Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2009**, *5*, 3010–3021.

(145) Savin, A.; Flad, H.-J. *Int. J. Quantum Chem.* **1995**, *56*, 327–332.

(146) Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. *J. Chem. Phys.* **2008**, *128*, 194105.

(147) Swart, M. *Chem. Phys. Lett.* **2013**, *580*, 166–171.

(148) Paziani, S.; Moroni, S.; Gori-Giorgi, P.; Bachelet, G. *Phys. Rev. B* **2006**, *73*, 155111.

(149) Toulouse, J.; Savin, A.; Flad, H.-J. *Int. J. Quantum Chem.* **2004**, *100*, 1047–1056.

(150) Toulouse, J.; Colonna, F.; Savin, A. *J. Chem. Phys.* **2005**, *122*, 014110.

(151) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. *SoftwareX* **2018**, *7*, 1–5.

(152) Frigo, M.; Johnson, S. G. *Proceedings of the IEEE* **2005**, *93*, Special issue on "Program Generation, Optimization, and Platform Adaptation", 216–231.

(153) Johnson, S. G.; Frigo, M. In *Fast Fourier Transforms*, Burrus, C. S., Ed.; Connexions: Rice University, Houston TX, 2008; Chapter 11.

(154) Frigo, M. In *Proceedings of the 1999 ACM SIGPLAN Conference on Programming Language Design and Implementation*, ACM: 1999; Vol. 34, pp 169–180.

(155) Frigo, M.; Johnson, S. G. In *Proceeding of the 1998 IEEE International Conference Acoustics Speech and Signal Processing*, IEEE: 1998; Vol. 3, pp 1381–1384.

(156) Frigo, M.; Johnson, S. G. *The Fastest Fourier Transform in the West*; tech. rep. MIT-LCS-TR-728; Massachusetts Institute of Technology, 1997.

(157) Frigo, M.; Leiserson, C. E.; Prokop, H.; Ramachandran, S. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, IEEE Comput. Soc.: 1999, pp 285–297.

(158) Johnson, S. G.; Frigo, M. *IEEE Trans. Signal Processing* **2007**, *55*, 111–119.

(159) Higham, N. J. *Wiley Interdiscip. Rev. Comput. Stat.* **2009**, *1*, 251–254.

(160) Anderson, E.; Bai, Z.; Bischof, C.; Blackford, S.; Demmel, J.; Dongarra, J.; Du Croz, J.; Greenbaum, A.; Hammarling, S.; McKenney, A.; Sorensen, D., *LAPACK Users' Guide*, Third; Society for Industrial and Applied Mathematics: Philadelphia, PA, 1999.

(161) Blackford, L. S.; Choi, J.; Cleary, A.; D'Azevedo, E.; Demmel, J.; Dhillon, I.; Dongarra, J.; Hammarling, S.; Henry, G.; Petitet, A.; Stanley, K.; Walker, D.; Whaley, R. C., *ScaLAPACK Users' Guide*; Society for Industrial and Applied Mathematics: Philadelphia, PA, 1997.

(162) Wilkinson, J. H., *The Algebraic Eigenvalue Problem*; Clarendon Press: 1988.

(163) Auckenthaler, T.; Blum, V.; Bungartz, H.-J.; Huckle, T.; Johanni, R.; Krämer, L.; Lang, B.; Lederer, H.; Willems, P. *Parallel Comput.* **2011**, *37*, 6th International Workshop on Parallel Matrix Algorithms and Applications (PMAA'10), 783–794.

(164) Marek, A.; Blum, V.; Johanni, R.; Havu, V.; Lang, B.; Auckenthaler, T.; Heinecke, A.; Bungartz, H.-J.; Lederer, H. *J. Phys. Condens. Matter* **2014**, *26*, 213201.

(165) Braess, D.; Hackbusch, W. *IMA J. Numer. Anal.* **2005**, *25*, 685–697.

(166) Ewald, P. P. *Ann. Phys.* **1921**, *369*, 253–287.

(167) Seewald, P. Low-scaling Electronic Structure Methods Based on Sparse Tensor Contraction, Ph.D. Thesis, Universität Zürich, 2021.

(168) Del Ben, M.; Schütt, O.; Wentz, T.; Messmer, P.; Hutter, J.; VandeVondele, J. *Comput. Phys. Commun.* **2015**, *187*, 120–129.

(169) Limpanuparb, T.; Gill, P. M. W. *J. Chem. Theory Comput.* **2011**, *7*, 2353–2357.

(170) Hutter, J. New optimization of GTH pseudopotentials for PBE, SCAN, PBE0 functionals. GTH pseudopotentials for Hartree-Fock. NLCC pseudopotentials for PBE. Latest commit from 8 August 2019. https://github.com/juerghutter/GTH (accessed 09/30/2020).

(171) Gori-Giorgi, P. Multideterminant Exchange-Correlation LSD Functional and Potentials (All Spin Polarizations, all Densities). https://www.paolagorigiorgi.org/s/LSD_SR.f (accessed 12/14/2021).

(172) Ogilvie, J.; Wang, F. Y. *J. Mol. Struct.* **1992**, *273*, 277–290.

(173) Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **2002**, *117*, 10548–10560.

(174) Ranasinghe, D. S.; Frisch, M. J.; Petersson, G. A. *J. Chem. Phys.* **2015**, *143*, 214110.

(175) Stein, F.; Hutter, J. Double-Hybrid Density functionals for the condensed phase: gradients, stress tensor, and Auxiliary-Density Matrix Method acceleration https://archive.materialscloud.org/record/2021.216 (accessed 12/14/2021).

(176) Gell-Mann, M.; Brueckner, K. A. *Phys. Rev.* **1957**, *106*, 364–368.

(177) Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* **1980**, *45*, 566–569.

(178) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(179) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. A* **1988**, *37*, 785–789.

(180) Kurth, Stefan and Perdew, John P. and Blaha, Peter *Int. J. Quantum Chem.* **1999**, *75*, 889–909.

(181) Perdew, J. P.; Kurth, S.; Zupan, A.; Blaha, P. *Phys. Rev. Lett.* **1999**, *82*, 2544–2547.

(182) Hartree, D. R. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 111–132.

(183) Mardirossian, N.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

(184) Zhang, I. Y.; Xu, X. *Int. Rev. Phys. Chem.* **2011**, *30*, 115–160.

(185) Martin, J. M. L.; Santra, G. *Isr. J. Chem.* **2020**, *60*, 787–804.

(186) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149–5155.

(187) Jurecka, P.; Cerny, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555–569.

(188) Sabatini, R.; Gorni, T.; de Gironcoli, S. *Phys. Rev. A* **2013**, *87*, 041108.

(189) Scuseria, G. E.; Henderson, T. M.; Sorensen, D. C. *J. Chem. Phys.* **2008**, *129*, 231101.

(190) Marsman, M.; Grüneis, A.; Paier, J.; Kresse, G. *J. Chem. Phys.* **2009**, *130*, 184103.

(191) Grüneis, A.; Marsman, M.; Kresse, G. *J. Chem. Phys.* **2010**, *133*, 074107.

(192) Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2018**, *8*, e1360.

(193) Barca, G. M. J. et al. *J. Chem. Phys.* **2020**, *152*, 154102.

(194) Förster, A.; Franchini, M.; van Lenthe, E.; Visscher, L. *J. Chem. Theory Comput.* **2020**, *16*, 875–891.

(195) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(196) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(197) Boyd, J. P. *J. Sci. Comput.* **1987**, *2*, 99–109.

(198) Braess, D., *Nonlinear Approximation Theory*; Springer Science & Business Media: Berlin Heidelberg, 2012.

(199) Bruneval, F. *Phys. Rev. Letters* **2012**, *108*, 256403.

(200) Guidon, M.; Schiffmann, F.; Hutter, J.; VandeVondele, J. *J. Chem. Phys.* **2008**, *128*, 214104.

(201) Kiefte, H.; Penney, R.; Breckon, S. W.; Clouter, M. J. *J. Chem. Phys.* **1987**, *86*, 662–665.

(202) Dulmage, W. J.; Lipscomb, W. N. *Acta Crystallogr.* **1951**, *4*, 330–334.

(203) Batchelder, D. N.; Losee, D. L.; Simmons, R. O. *Phys. Rev.* **1967**, *162*, 767–775.

(204) McConville, G. T. *J. Chem. Phys.* **1974**, *60*, 4093–4093.

(205) Endoh, Y.; Shirane, G.; Skalyo, J. *Phys. Rev. B* **1975**, *11*, 1681–1688.

(206) Peterson, O. G.; Batchelder, D. N.; Simmons, R. O. *Phys. Rev.* **1966**, *150*, 703–711.

(207) Schwalbe, L. A.; Crawford, R. K.; Chen, H. H.; Aziz, R. A. *J. Chem. Phys.* **1977**, *66*, 4493–4502.

(208) Civalleri, B.; Zicovich-Wilson, C.; Valenzano, L.; Ugliengo, P. *Crystengcomm* **2008**, *10*, 405–410.

(209) Zhao, Y.; Truhlar, D. G. *J. Comput. Chem. A* **2005**, *109*, 5656–5667.

(210) Mardirossian, N.; Ruiz Pestana, L.; Womack, J. C.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. *J. Comput. Chem. Lett.* **2017**, *8*, 35–40.

(211) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(212) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(213) Bernardi, F. *J. Phys. France* **1973**, *34*, 373–379.

(214) Chickos, J. S.; Acree, W. E. *J. Phys. Chem. Ref. Data* **2002**, *31*, 537–698.

(215) Gavezzotti, A. *Model. Simul. Mater. Sci. Eng.* **2002**, *10*, R1–R29.

(216) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.

(217) Santra, G.; Sylvetsky, N.; Martin, J. M. L. *J. Comput. Chem. A* **2019**, *123*, 5129–5143.

(218) Cheng, J.; VandeVondele, J. *Phys. Rev. Lett.* **2016**, *116*.

(219) Wang, Y.; Li, Y.; Chen, J.; Zhang, I. Y.; Xu, X. *J. Am. Chem. Soc.* **2021**, *1*, 543–549.

(220) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

(221) Obara, S.; Saika, A. *J. Chem. Phys.* **1986**, *84*, 3963–3974.

(222) McMurchie, L. E.; Davidson, E. R. *J. Comput. Phys.* **1978**, *26*, 218–231.

(223) Dupuis, M.; Rys, J.; King, H. F. *J. Chem. Phys.* **1976**, *65*, 111–116.

(224) Gill, P. M. In *Molecular integrals Over Gaussian Basis Functions*, Sabin, J. R., Zerner, M. C., Eds.; Advances in Quantum Chemistry, Vol. 25; Academic Press: 1994, pp 141–205.

(225) Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *Int. J. Quantum Chem.* **1989**, *36*, 269–280.

(226) Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **1990**, *94*, 5564–5572.

(227) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. *Int. J. Quantum Chem.* **1991**, *40*, 745–752.

(228) Valeev, E. F. Libint: A library for the evaluation of molecular integrals of many-body operators over Gaussian functions http://libint.valeyev.net/ (accessed 12/14/2021).

(229) Früchtl, H. A.; Kendall, R. A.; Harrison, R. J.; Dyall, K. G. *Int. J. Quantum Chem.* **1997**, *64*, 63–69.

(230) Sodt, A.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 104106.

(231) Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2010**, *6*, 2348–2364.

(232) Asadchev, A.; Gordon, M. S. *J. Chem. Theory Comput.* **2012**, *8*, 4166–4176.

(233) Wilkinson, K. A.; Sherwood, P.; Guest, M. F.; Naidoo, K. J. *J. Comput. Chem.* **2011**, *32*, 2313–2318.

(234) Ufimtsev, I. S.; Martínez, T. J. *J. Chem. Theory Comput.* **2008**, *4*, 222–231.

(235) Tornai, G. J.; Ladjánszki, I.; Rák, A.; Kis, G.; Cserey, G. *J. Chem. Theory Comput.* **2019**, *15*, 5319–5331.

(236) Spencer, J.; Alavi, A. *Phys. Rev. B* **2008**, *77*, 193110.

(237) Carrier, P.; Rohra, S.; Görling, A. *Phys. Rev. B* **2007**, *75*, 205126.

(238) Gygi, F.; Baldereschi, A. *Phys. Rev. B* **1986**, *34*, 4405–4408.

(239) Strout, D. L.; Scuseria, G. E. *J. Chem. Phys.* **1995**, *102*, 8448–8452.

(240) Ochsenfeld, C.; White, C. A.; Head-Gordon, M. *J. Chem. Phys.* **1998**, *109*, 1663–1669.

(241) Izmaylov, A. F.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 104103.

(242) Tran, F.; Blaha, P. *Phys. Rev. B* **2011**, *83*, 235118.

(243) Giannozzi, P. et al. *J. Phys. Condens. Matter* **2009**, *21*, 395502.

(244) Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(245) Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, *59*, 1758–1775.

(246) Blaha, P.; Schwarz, K.; Tran, F.; Laskowski, R.; Madsen, G. K. H.; Marks, L. D. *J. Chem. Phys.* **2020**, *152*, 074101.

(247) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403–412.

(248) Merlot, P.; Izsák, R.; Borgoo, A.; Kjærgaard, T.; Helgaker, T.; Reine, S. *J. Chem. Phys.* **2014**, *141*, 094104.

(249) Del Ben, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2013**, *9*, 2654–2671.

(250) Del Ben, M.; Hutter, J.; VandeVondele, J. *J. Chem. Phys.* **2015**, *143*, 102803.

(251) Rybkin, V. V.; VandeVondele, J. *J. Chem. Theory Comput.* **2016**, *12*, 2214–2223.

(252) Neese, F.; Schwabe, T.; Grimme, S. *J. Chem. Phys.* **2007**, *126*, 124115.

(253) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

(254) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem.* **1979**, *16*, 225–241.

(255) Kumar, C.; Fliegl, H.; Jensen, F.; Teale, A. M.; Reine, S.; Kjærgaard, T. *Int. J. Quantum Chem.* **2018**, *118*, e25639.

(256) Pisani, C.; Schütz, M.; Casassa, S.; Usvyat, D.; Maschio, L.; Lorenz, M.; Erba, A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7615.

(257) Kwasniewski, G.; Kabić, M.; Besta, M.; VandeVondele, J.; Solcà, R.; Hoefler, T. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2019, pp 1–22.

(258) Kozuch, S.; Martin, J. M. L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.

(259) Cheng, B.; Engel, E. A.; Behler, J.; Dellago, C.; Ceriotti, M. *PNAS* **2019**, *116*, 1110–1115.

(260) Jeffrey, G. A.; Ruble, J. R.; McMullan, R. K.; Pople, J. A.; Cox, E. G. *Proc. R. Soc. Lond., A Math. phys. sci.* **1987**, *414*, 47–57.

(261) Spreafico, C.; VandeVondele, J. *Phys. Chem. Chem. Phys.* **2014**, *16*, 26144–26152.

(262) team., T. S. Slurm Workload Manager. https://slurm.schedmd.com/overview.html (accessed 01/02/2022).

(263) Moore, G. E. *Electronics* **1965**, *38*.

(264) Kim, B. S. Y.; Minohara, M.; Hikita, Y.; Bell, C.; Hwang, H. Y. *Appl. Phys. Lett.* **2018**, *112*, 133506.

(265) Zhang, J.; Zhou, P.; Liu, J.; Yu, J. *Phys. Chem. Chem. Phys.* **2014**, *16*, 20382–20386.

(266) Zhang, C.; Hutter, J.; Sprik, M. *J. Phys. Chem. Lett.* **2019**, *10*, 3871–3876.

(267) Warren, D.; Shapira, Y.; Kisch, H.; Mcquillan, A. *J. Phys. Chem. C* **2007**, *111*, 14286–14289.

(268) O'Regan, B.; Grätzel, M. *Nature* **1991**, *353*, 737–740.

(269) Izsák, R.; Neese, F. *J. Chem. Phys.* **2011**, *135*, 144105.

(270) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. *Chem. Phys.* **2009**, *356*, Moving Frontiers in Quantum Chemistry: 98–109.