TECHNISCHE UNIVERSITÄT WIEN

INSTITUTE FOR THEORETICAL PHYSICS

DIPLOMARBEIT

Efficient Equation of Motion Coupled Cluster Theory using Low-Rank Factorization

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Technische Physik

eingereicht von

Thomas Plaikner, BSc

Matrikelnummer 11771063

ausgeführt am Institut für Theoretische Physik
der Fakultät für Physik der Technischen Universität Wien

Betreuer: Univ.Prof. Mag.rer.nat. Dr.rer.nat. Andreas Grüneis

Wien, 14.08.2023

Thomas Plaikner

Andreas Grüneis

Abstract

Density functional theory in the Kohn-Sham framework is the most common and widespread method in computational material science. However, the necessity for highly accurate electronic structure theory calculations gives rise to the continued exploration of alternative techniques. Within these approaches, quantum chemical wave function based methods such as the equation of motion coupled cluster theories show remarkable potential. In conjunction with these methods, the biggest drawback stems from the significantly larger computational cost, which scales polynomially with the system size.
The aim of this work is the implementation of several electronic structure algorithms culminating in the improvement of the computational efficiency of the equation of motion coupled cluster singles and doubles theory by applying low-rank factorization techniques to the Coulomb integrals. As proof of principle, we employ a model of two electrons in one dimension confined within a harmonic potential. We can validate the successful implementation of this approach and we are able to verify that the overall computational expenses are reduced through the approximation of the Coulomb integrals via six small matrices using tensor contraction. The low-rank factorized Coulomb integrals prove effective in calculating excited state energies without significant loss of accuracy. We find that adjusting two parameters of our low-rank approximation techniques allows for a systematic trade-off between accurate results and computational expenses.
The results of our work make us optimistic that, building on this proof of principle, we can go a step beyond and develop efficient methods for calculating linear absorption spectra of real materials.

Kurzfassung

Dichtefunktionaltheorie im Kohn-Sham Formalismus ist die am häufigsten verwendete Methode in der computer-gestützten Materialwissenschaft. Die Notwendigkeit hochgenauer Berechnungen elektronischer Strukturen öffnet jedoch die Tür für alternative Methoden. Eine Familie vielversprechender Vertreter sind quantenchemische Wellenfunktions-basierte Methoden unter welche auch die sogenannte „equation of motion coupled cluster“ Theorien fallen. Einer der größten Nachteile dieser Algorithmen ist der signifikant höhere Rechenaufwand, welcher polynomiell mit der Systemgröße skaliert.
Diese Arbeit beschreibt zunächst verschiedene Algorithmen für elektronische Strukturberechnungen, mit dem finalen Ziel der Reduzierung des Rechenaufwands der „equation of motion coupled cluster singles and doubles“ Theorie durch Faktorisierungtechniken. Wir verwenden ein Modell von zwei Elektronen in einem eindimensionalen harmonischen Potential. Die numerische Implementierung ist detailliert beschrieben und es ist gezeigt, dass sich die Gesamtrechenkosten durch die Approximation der Coulomb-Integrale mittels sechs Matrizen durch Tensor-Kontraktion reduzieren lassen. Weiters erlauben die faktorisierten Coulomb-Integrale die effektive Berechnung von Anregungsenergien ohne signifikante Genauigkeitsverluste. Die systematische Variation zweier Parameter, die in unseren Faktorisierungsmethoden vorkommen, ermöglichen einen Kompromiss zwischen genauen Ergebnissen und Rechenaufwand.
Die Ergebnisse unserer Arbeit stimmen uns zuversichtlich, dass wir einen Schritt weiter gehen und effiziente Methoden zur Berechnung von linearen Absorptionsspektren realer Materialien entwickeln können.

Contents

1 Introduction

In computational material science, a primary objective revolves around solving the many-electron Schrödinger equation. However, exact solutions are only feasible for the smallest systems, necessitating the utilization of numerical approximation methods. Density functional theory thereby represents without a doubt the method of choice when studying *ab initio* problems [1–3]. The good trade-off between accuracy and computational cost in the Kohn-Sham framework of approximate exchange and correlation energy functionals stands as a fundamental benchmark in contemporary material science. Despite the significant achievements of density functional theory over the past decades, there are still challenges in systematically improving the accuracy of existing approximate density functionals. Amongst various aspects, systems involving strong electronic correlation effects, non-local van der Waals interactions or density-driven errors may even exhibit qualitative failures [4]. This opens the door for other techniques such as wave function-based methods, which have the property of capturing electronic exchange and correlation effects in a systematically improvable manner. The drawbacks of such theories involve significantly larger numerical costs restricting the computation to relatively small system sizes only. Nevertheless, highly accurate electronic structure theories are much needed to get important benchmark results as well as to predict material properties without relying on non-improvable approximations.

The one theory to which we will devote ourselves in the course of this work is the so-called coupled cluster theory which has already had great success in predicting various material properties [5–11]. Initially proposed in the 1950s by Fritz Coester and Herman Kümmel for the field of nuclear physics [12,13], coupled cluster theory nowadays is also widespread in computational chemistry. Jiri Cizek and Josef Paldus were the first to introduce the coupled cluster method explicitly for the description of electron correlations in the 1960s [14, 15]. The underlying principle is the description of the many-electron wave function by an exponential ansatz of cluster operators acting on a single reference Slater determinant. Further advancement is introduced by the equation of motion coupled cluster theories enabling the calculation of electron addition and removal energies in solids as well as excited state properties [16–18]. In this case, the ansatz for the coupled cluster wave function is extended by incorporating an additional linear excitation operator.

For both coupled cluster and equation of motion coupled cluster theories, the achievable accuracy substantially depends on the employed level of truncation in the wave function ansatz. Consequently, the high dimensionality of the many-electron wave functions is one of the most limiting factors for such electronic structure theories. Over the past decade great efforts have been made to improve numerical efficiency through tensor rank decomposition and low-rank tensor approximation [19–23]. These techniques not only reduce computational costs but also minimize memory footprints, enabling the storage of the approximated many-electron wave functions.

The main goal of this thesis is to employ the aforementioned techniques to improve the computational efficiency of the equation of motion coupled cluster theory, focusing on a model of two electrons confined in a one-dimensional harmonic potential. To achieve this, we introduce in the first chapter the theoretical foundations necessary for the course of this thesis. The second chapter is dedicated to the Hartree-Fock method, which is needed as a prerequisite for all subsequent post-Hartree-Fock methods. These methods, namely Møller-Plesset perturbation theory, coupled cluster and equation of motion coupled cluster theory are discussed in detail in the third chapter. Following the implementation of these methods, the final chapter focuses on their computational optimization in terms of tensor rank reduction of the Coulomb integrals. In the concluding part of the thesis we present the most significant findings and their interpretation. Finally, a brief summary of the key outcomes is provided, along with an outlook on potential future extensions considering the equation of motion coupled cluster theory and low-rank approximation methods.

2 Many Electron Theory

In this chapter we provide an introductory overview of the fundamental principles essential for the subsequent work. The derivations are primarily based on the electronic Hamiltonian which is commonly employed for realistic materials. However, we also introduce the Hamiltonian of our model and, if necessary, we refer to it during the discussion. The main topics covered are the Schrödinger equation, the concepts of spin orbitals and Slater determinants. Based on this, we discuss one- and two-electron integrals and their evaluation outlining the main results of the Slater-Condon rules. Finally, we introduce the concept of second quantization, which is especially useful for the description of the coupled cluster theory.

2.1 Schrödinger Equation

The calculation of properties of a non-relativistic quantum system commonly requires the solution of the Schrödinger equation. It is a partial differential equation that describes the change in time of a quantum mechanical state of a system. This state is represented by a wave function, and the time evolution is given by the action of the Hamiltonian operator on this wave function. The most general form of the Schrödinger equation is

$$i\hbar\frac{\partial}{\partial t}\left|\Psi(t)\right\rangle = \hat{H}\left|\Psi(t)\right\rangle,$$

where $|\Psi(t)\rangle$ is the time-dependent wave function in Dirac notation and $\hat{H}$ the Hamiltonian operator. The constants i and $\hbar$ denote the imaginary unit and the reduced Planck constant, respectively. In the case of a time-independent Hamiltonian, the wave function can be factorized into a time-dependent and a time-independent part. Consequently, the time evolution of the wave function is given by:

$$\left|\Psi(t)\right\rangle = e^{-iHt/\hbar}\left|\Psi(t=0)\right\rangle = e^{-iHt/\hbar}\left|\Psi\right\rangle.$$

The decoupling of the time-dependent Schrödinger equation leads to a time-independent eigenvalue problem, known as the time-independent Schrödinger equation:

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle. \tag{2.1}$$

The explicit form of the Schrödinger equation is determined by the definition of the Hamiltonian operator which depends on the system under consideration. In molecular electronic structure theory an arbitrary quantum mechanical system with N_e electrons and N_n nuclei is often described by the many-body (Coulomb) Hamiltonian:

$$\hat{H} = -\sum_{i}^{N_e}\frac{\hbar^2}{2m}\vec{\Delta}_i - \sum_{k}^{N_n}\frac{\hbar^2}{2M_k}\vec{\Delta}_k + \frac{e^2}{4\pi\epsilon_0}\left[\frac{1}{2}\sum_{i,j=1;i\neq j}^{N_e}\frac{1}{|\vec{r}_i-\vec{r}_j|} - \sum_{i=1}^{N_e}\sum_{k=1}^{N_n}\frac{Z_k}{|\vec{r}_i-\vec{R}_k|} + \frac{1}{2}\sum_{k,l=1;k\neq l}^{N_n}\frac{Z_kZ_l}{|\vec{R}_k-\vec{R}_l|}\right]. \tag{2.2}$$

Here m and M_k stand for the electron mass and the k-th nucleus mass, respectively. Z_k is the atomic number, ϵ_0 corresponds to the vacuum permittivity, e denotes the elementary charge, $\vec{r}_i$ and $\vec{R}_k$ are the

coordinates of the electrons and the nuclei, respectively.
In order to get the electronic properties of such systems the so-called Born-Oppenheimer approximation is applied. This approximation assumes that the wave functions of atomic nuclei and electrons in a molecule can be treated independently. Since the relative mass of a nucleus is much larger than the one of the electron, it moves much slower. This allows a separation of the Hamiltonian (s. Eq. 2.2) into an electronic and a nuclear part, where cross terms between electrons and nuclei are neglected. The wave function associated with the Coulomb Hamiltonian in the Born-Oppenheimer approximation is then given by:

$$\Psi_{coul} = \Phi_e \cdot \xi_n.$$

Φ_e denotes the electronic and ξ_n the nuclear wave function. The Coulomb Hamiltonian applied to this wave function thus yields the Schrödinger equation for the motion of the electrons:

$$\hat{H}_e \Phi_e(\vec{r}; \vec{R}_n) = E_e(\vec{R}_n)\Phi_e(\vec{r}; \vec{R}_n), \tag{2.3}$$

with the electronic Hamiltonian:

$$\hat{H}_e = \sum_i^{N_e} -\frac{\hbar^2}{2m}\vec{\Delta}_i + \frac{e^2}{4\pi\epsilon_0}\left[\frac{1}{2}\sum_{i,j=1;i\neq j}^{N_e}\frac{1}{|\vec{r}_i - \vec{r}_j|} - \sum_{i=1}^{N_e}\sum_{k=1}^{N_n}\frac{Z_k}{|\vec{r}_i - \vec{R}_k|}\right], \tag{2.4}$$

and another Schrödinger equation for the motion of the nuclei:

$$\hat{H}_n \xi_n(\vec{R}_n) = E_n \xi_n(\vec{R}_n),$$

with the nuclear Hamiltonian:

$$\hat{H}_n = \sum_k^{N_n} -\frac{\hbar^2}{2M_k}\vec{\Delta}_k + \frac{e^2}{4\pi\epsilon_0}\frac{1}{2}\sum_{k,l=1;k\neq l}^{N_n}\frac{Z_k Z_l}{|\vec{R}_k - \vec{R}_l|}.$$

In order to obtain, for example, potential energy surfaces, the electronic Schrödinger equation (s. Eq. 2.3) is solved successively for different nuclear distances.
Many numerical algorithms, such as the methods discussed later, strive to solve the electronic Schrödinger equation by relying on the Born-Oppenheimer approximation. For further reading and more details, see Ref. [24].

At this point, we want to introduce our model Hamiltonian. Since this thesis represents a so-called proof of principle, it is reasonable to work with a rather simple model. This offers us two advantages. On one hand, we are able to qualitatively analyze the obtained results without exceeding the given computational resources. This means that all implemented codes can be executed on standard computer devices. On the other hand, we are able to focus more on the main goal of the work, the low-rank approximation of the two-electron integrals. Accordingly, our Hamiltonian shall describe two electrons confined in a harmonic potential in one dimension, i.e.

$$\hat{H} = \sum_{i=1}^{2} \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_i^2} + \frac{1}{2}m\omega^2 x_i^2 \right] + \frac{e^2}{4\pi\epsilon_0}\frac{1}{|x_1 - x_2| + \sigma}. \tag{2.5}$$

Here, ω is the oscillator frequency and $\sigma > 0$ a real number. The first term describes the kinetic energy of an electron while the second term represents the harmonic potential. The third expression gives a regularized Coulomb potential between the two electrons.
Further, the Schrödinger equation is transformed into atomic units ($\hbar = e = m = (4\pi\epsilon_0)^{-1} = 1\,a.u.$) making Eq. (2.5) read:

$$\hat{H} = \sum_{i=1}^{2} \left[-\frac{1}{2}\frac{\partial^2}{\partial x_i^2} + \frac{1}{2}\omega^2 x_i^2 \right] + \frac{1}{|x_1 - x_2| + \sigma}. \tag{2.6}$$

Due to the fact that our model is set up in one dimension only, the Coulomb interaction is not very intuitive since the electrons are not able to pass each other. For the present work, we stick to this designation and introduce σ as a regularization parameter to prevent singularities arising in the involved integrations (s. Sec. 2.4). This allows us quick testing and prototyping of different ideas for the further course of the thesis.

2.2 Spin Orbitals

We introduce the term orbital in the context of a single electron wave function. This wave function is obtained by the solution of the electronic Schrödinger equation (s. Eq. 2.1). A spatial orbital $\Phi(\vec{r})$ describes the spatial distribution of an electron with the probability $|\Phi(\vec{r})|^2$ of finding that electron in a small volume $d\vec{r}$ around $\vec{r}$. These orbitals have the property of orthonormality

$$\langle \Phi_i | \Phi_j \rangle = \int d\vec{r}\, \Phi_i^*(\vec{r})\Phi_j(\vec{r}) = \delta_{ij}.$$

Further, an arbitrary function $f(\vec{r})$ can be expanded in terms of the complete set of spatial orbitals $\{\Phi_i\}$

$$f(\vec{r}) = \sum_{i=1}^{\infty} a_i \Phi_i(\vec{r}).$$

However, the full description of an electron also requires the consideration of its spin. Electrons are fermions with spin $S = 1/2$, which leads to spin up ($M_S = 1/2$) and spin down ($M_S = -1/2$) states. In this context, we introduce the spin functions $\alpha(\bar{\omega})$ and $\beta(\bar{\omega})$ corresponding to spin up and down, respectively. We define the common spin variable with a horizontal line to avoid confusion with the frequency ω from the harmonic potential. The two spin functions have the following properties:

$$\begin{aligned} \int d\bar{\omega}\ \alpha^*(\bar{\omega})\alpha(\bar{\omega}) &= \int d\bar{\omega}\ \beta^*(\bar{\omega})\beta(\bar{\omega}) = 1, \\ \int d\bar{\omega}\ \alpha^*(\bar{\omega})\beta(\bar{\omega}) &= \int d\bar{\omega}\ \beta^*(\bar{\omega})\alpha(\bar{\omega}) = 0. \end{aligned} \tag{2.7}$$

Thus, $\alpha(\bar{\omega})$ and $\beta(\bar{\omega})$ are complete and orthonormal and an electron can be fully described by the three spatial coordinates $\vec{r}$ and the spin coordinate $\bar{\omega}$. We define a new variable that contains both the spatial and spin coordinates:

$$\vec{x} = \{\vec{r}, \bar{\omega}\}.$$

The spin orbital $\chi(\vec{x})$ is defined as the wave function that describes the spatial distribution as well as the spin of the electron. Due to the fact, that we have the possibility of finding an electron in spin up or down, we can form two spin orbitals from each spatial orbital:

$$\chi(\vec{x}) = \begin{cases} \Phi(\vec{r})\alpha(\bar{\omega}) \\ or \\ \Phi(\vec{r})\beta(\bar{\omega}) \end{cases}. \tag{2.8}$$

2.3 Slater Determinants

After introducing the spin orbitals for a single electron, we look at the generalization of a system with N electrons. We start with a historically important description of the many-electron wave function, the so-called Hartree product. For a system of N electrons, the Hartree product is

$$\Psi^{Hartree}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N) = \chi(\vec{x}_1)\chi(\vec{x}_2)\ldots\chi(\vec{x}_N). \tag{2.9}$$

This is a rather intuitive solution, where the N-electron wave function is represented as a product of one-electron spin orbitals. Indeed, for the special case where the total Hamiltonian is defined as the sum of one-electron Hamiltonians $\hat{h}(i)$, whose eigenstates are the spin orbitals, the Hartree product is the eigenstate of the total Hamiltonian operator. It can be shown that it holds:

$$\left(\sum_{i=1}^{N} \hat{h}_i\right) \Psi^{Hartree} = \left(\sum_{i=1}^{N} \epsilon_i\right) \Psi^{Hartree}.$$

However, the Hartree product does not cover the general case of a many-electron wave function. Moreover, it has two major problems that cannot be overlooked. Due to the definition of the Hartree product (s. Eq. 2.9), the probability of finding an electron (e.g. labeled 1) around the position $\vec{r}_1$ in the volume element $d\vec{r}_1$ is independent of the positions of the other electrons in the systems. It is well known that between the electrons there is a repulsive force, i.e. the electrons avoid the regions occupied by other electrons. Thus, the Hartree product represents an uncorrelated wave function that does not reflect the actual behavior of a many-electron system. The second deficiency is that the Pauli principle is not fulfilled. Electrons cannot occupy the same quantum state in a system. In other words, electrons are indistinguishable, but the Hartree product assigns each electron its own spin orbital. These inconveniences are overcome by introducing the antisymmetry principle, i.e. the wave function must change sign with respect to the interchange of the space and spin coordinates of any two electrons.
The mentioned drawbacks are overcome by the formalism of Slater determinants defined as:

$$\Psi^{SD}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\vec{x}_1) & \chi_2(\vec{x}_1) & \dots & \chi_N(\vec{x}_1) \\ \chi_1(\vec{x}_2) & \chi_2(\vec{x}_2) & \dots & \chi_N(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\vec{x}_N) & \chi_2(\vec{x}_N) & \dots & \chi_N(\vec{x}_N) \end{vmatrix}. \tag{2.10}$$

The factor $(N!)^{-1/2}$ stems from the normalization. The Slater determinant in Eq. 2.10 describes N electrons occupying N spin orbitals $\{\chi_1, \chi_2, \dots \chi_N\}$. The rows correspond to electrons and the columns are labeled by the spin orbitals. Interchanging the coordinates of two electrons means switching two rows of the determinant, which changes its sign. In the case of two equal columns, this corresponds to two electrons in the same spin orbital, making the Slater determinant zero (Pauli exclusion principle). Another remarkable property of the Slater determinants is the consideration of exchange effects. The motion of electrons with parallel spin is correlated, while the one for anti-parallel electrons is not affected. These properties of the Slater determinant wave function can be displayed as:

$$\Psi^{SD}(\dots, \vec{x}_i, \dots, \vec{x}_j, \dots) = -\Psi^{SD}(\dots, \vec{x}_j, \dots, \vec{x}_i, \dots).$$

Furthermore, we introduce a short notation, which simplifies our work later on. First, we rewrite the expression in Eq. 2.10 by showing only the diagonal elements of the determinant

$$\Psi^{SD}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = |\chi_1(\vec{x}_1)\chi_2(\vec{x}_2) \dots \chi_N(\vec{x}_N)\rangle .$$

In this notation, we drop the superscript SD due to the fact that in the course of this work, we only deal with Slater determinant like wave functions. The normalization factor is included implicitly. If we choose an ordered sequence of the variables $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N$ we get a further simplification:

$$|\Psi\rangle = |\chi_1 \chi_2 \dots \chi_N\rangle = |1\, 2 \dots N\rangle . \tag{2.11}$$

Here, we introduce the shorthand notation $\chi_i = i$ which is convenient to use for more tedious expressions. The simplest wave function describing a set of $2K$ spin orbitals and N electrons ($2K \geq N$) is the single Slater determinant. Within the Slater determinant we distinguish between occupied and unoccupied spin orbitals. The former are labeled with the indices $(i, j, k, \dots)$ while the latter, also called virtual spin orbitals, are labeled with the indices $(a, b, c, \dots)$. We use the indices $(p, q, r, \dots)$ for the generic orbitals, where the occupation is not specified. The most intuitive way to construct a single Slater determinant is where first the occupied and then the virtual orbitals are labeled:

$$|\Psi^{HF}\rangle = |1\, 2 \dots i\, j \dots N\rangle . \tag{2.12}$$

Here, we introduce the Hartree-Fock (ground state) wave function $|\Psi^{HF}\rangle$ which will be needed in all upcoming chapters. More details about Hartree-Fock are provided in chapter 3.
Further, we can take one Slater determinant, e.g. the Hartree-Fock determinant, as a reference and describe

others by their deviation from it. We name these different determinants excited Slater determinants, with the singly and doubly excited states being of most interest to us. A single excitation is given when an electron in an occupied orbital χ_i transitions to a virtual orbital χ_a:

$$|\Psi_i^a\rangle = |1\,2\ldots a\,j\ldots N\rangle\,.$$

Analogously, a doubly excited determinant corresponds to two electrons that are lifted from occupied to virtual orbitals

$$|\Psi_{ij}^{ab}\rangle = |1\,2\ldots a\,b\ldots N\rangle\,.$$

2.4 One-Electron and Two-Electron Operators

In this section, we look at the individual contributions from an arbitrary Hamiltonian operator. The electronic Hamiltonian defined in Eq. 2.4 and also our model Hamiltonian in Eq. 2.6 have in common that both can be split up into terms acting on only one electron and terms acting on two electrons. For N electrons, the one-electron terms are described by the one-electron operator $\hat{O}_1$:

$$\hat{O}_1 = \sum_{i=1}^{N} \hat{h}_i \quad ; \qquad \hat{h}_i = -\frac{1}{2}\Delta_i + \sum_{i=1}^{N} V_i, \tag{2.13}$$

where we don't specify the form of the potential V_i. For the electronic Hamiltonian, V is the electron-nucleus potential while for our model it represents the harmonic potential. On the other hand, the two-electron operator $\hat{O}_2$ is given by:

$$\hat{O}_2 = v_{ij}, \tag{2.14}$$

where, again, v_{ij} is an arbitrary two-electron potential. In the case of the electronic Hamiltonian, in atomic units, it represents the Coulomb potential $v_{ij} = \frac{1}{2}\sum_{i\neq j}^{N} \vec{r}_{ij}^{\,-1}$ where the factor $1/2$ is included to avoid double counting. In our model v has a similar form characterizing the same potential but only in one dimension (s. Eq. 2.6).

Next, we want to investigate the matrix elements of these one- and two-electron operators in the basis of the spin orbitals. The evaluation of these matrix elements corresponds to the solution of integrals involving the spin orbitals in the respective Slater determinants. The one-electron integrals are defined as

$$\langle\chi_i|\hat{h}|\chi_j\rangle = \langle i|\hat{h}|j\rangle = h_{ij} = \int d\vec{x}_1\; \chi_i(\vec{x}_1)\hat{h}(\vec{r}_1)\chi_j(\vec{x}_1),$$

where the property $\langle i|\hat{h}|j\rangle = \langle j|\hat{h}|i\rangle^*$ holds. Further, the operator $\hat{h}$ does not depend on spin coordinates, while the spin orbital variable $\vec{x}_i$ contain both spin and spatial coordinates. Here we assume that every orbital is occupied with two electrons which is referred to as the restricted spin orbital formalism, discussed in the next chapter. Similarly, we define the two-electron integrals as

$$\langle \chi_i \chi_j | v | \chi_k \chi_l \rangle = \langle ij|v|kl\rangle = \langle ij|kl\rangle = \int d\vec{x}_1 \int d\vec{x}_2 \; \chi_i^*(\vec{x}_1)\chi_j^*(\vec{x}_2) v_{ij}(\vec{r}_1,\vec{r}_2)\chi_k(\vec{x}_1)\chi_l(\vec{x}_2). \tag{2.15}$$

Also in this case the condition $\langle ij|kl\rangle = \langle kl|ij\rangle^*$ is valid.
Additionally, we briefly mention further notations regarding the one- and two-electron integrals which are commonly used by the community. The introduced Dirac bra-ket formalism in the integrals is also called the physicist's notation. However, there exists also the so-called chemist's notation, which is equal in the case of the one-electron integrals, but essentially different in the two-electron integrals. For the one-electron case, it holds

$$\langle i|\hat{h}|j\rangle = [i|\hat{h}|j],$$

where the former is the physicist's and the latter the chemist's notation. For the two-electron case, we get

$$\langle ij|kl\rangle = [ik|jl] = \int d\vec{x}_1 \int d\vec{x}_2 \; \chi_i^*(\vec{x}_1)\chi_k^*(\vec{x}_1) v_{ij}(\vec{r}_1,\vec{r}_2)\chi_j(\vec{x}_2)\chi_l(\vec{x}_2).$$

Chemists prefer to put the spin orbitals with the same coordinate together, while physicist's like to first write the complex conjugate spin orbitals. Moreover, we introduce the antisymmetric two-electron integral as:

$$\langle ij||kl\rangle = \langle ij|kl\rangle - \langle ij|lk\rangle = [ik|jl] - [il|jk].$$

So far, we discussed integrals concerning the spin orbitals. However, in the Hartree-Fock method, we mostly discuss the restricted closed-shell case (s. Sec. 3.3), where the spin part can be eliminated by integration yielding one- and two-electron integrals just over the spatial orbitals. Our convention for these spatial integrals is

$$(ij|kl) = \int d\vec{r}_1 \int d\vec{r}_2 \; \Phi_i^*(\vec{r}_1)\Phi_j^*(\vec{r}_2) v_{ij}(\vec{r}_1,\vec{r}_2)\Phi_k(\vec{r}_1)\Phi_l(\vec{r}_2), \tag{2.16}$$

where we designate the spatial integral with round brackets and refer to Φ_i as spatial orbitals.

2.5 Slater-Condon Rules

To evaluate the one- and two-electron integrals we exploit the Slater-Condon rules. These rules determine whether certain matrix elements contribute to the whole Hamiltonian matrix or vanish. The derivation of the Slater-Condon rules is beyond the scope of this work. Nevertheless, we present the main results and refer to Refs. [25, 26] for further reading. The approach of the Slater-Condon rules is to look at the one- and two-electron integrals between the Hartree-Fock determinant and excited determinants. Thereby, each determinant is expressed in the spin orbital basis:

$$|\chi_i\,\chi_j\dots\chi_k\rangle = (N!)^{-1/2}\sum_{n=1}^{N!}(-1)^{p_n}\hat{P}_n\{\chi_i(1)\,\chi_j(2)\dots\chi_k(N)\}, \tag{2.17}$$

where $\hat{P}_n$ is the permutation operator and p_n the number of transpositions required to obtain that permutation. The factor $N!$ in the sum accounts for all possible distinct permutations.
In the following, the Hartree-Fock ground state wave function $|\Psi^{HF}\rangle$ is taken as the reference determinant. We begin with the case of the identity operator $\mathbb{1}$, i.e. the scalar product between two determinants:

$$\langle\Psi^{HF}|\Psi^{HF}\rangle = 1 \quad ; \qquad \langle\Psi^{HF}|\Psi^{a\dots}_{i\dots}\rangle = 0. \tag{2.18}$$

We see that every scalar product vanishes except the one between two equal determinants. This means that every electron occupies the same spin orbital in the permutations of both determinants according to Eq. 2.17. For a one-electron operator $\hat{O}_1$ (s. Eq. 2.13) it holds:

$$\begin{aligned}
\langle\Psi^{HF}|\hat{O}_1|\Psi^{HF}\rangle &= \sum_i^N \langle i|\hat{h}|i\rangle = 2\sum_i^{N/2}(i|\hat{h}|i),\\
\langle\Psi^{HF}|\hat{O}_1|\Psi^{a}_{i}\rangle &= \langle i|\hat{h}|a\rangle = (i|\hat{h}|a),\\
\langle\Psi^{HF}|\hat{O}_1|\Psi^{ab\dots}_{ij\dots}\rangle &= 0.
\end{aligned} \tag{2.19}$$

The expectation value with the same determinant gives the sum over the spin orbitals, which can be rewritten as twice the sum over the spatial orbitals assuming two electrons occupy one orbital. According to Eq. 2.17 there are $(N-1)!$ ways that one electron (we label it 1) occupies each spin orbital. On the other hand, there exists only one possibility regarding the integration with the Hartree-Fock and the single excited determinant. This is due to the fact that the spin orbital in the first permutation is orthogonal to any other spin orbital in the second permutation unless electron 1 occupies it and thus yields an association with $\hat{h(1)}$. The integration with respect to higher excited determinants vanishes since there is no possible way electron 1 can occupy two spin orbitals and be associated with $\hat{h}(1)$. The Slater-Condon rules for the two-electron operator (s. Eq. 2.14) read:

$$\begin{aligned}
\langle\Psi^{HF}|\hat{O}_2|\Psi^{HF}\rangle &= \frac{1}{2}\sum_{ij}^{N}\langle ij||ij\rangle = \sum_{ij}^{N/2} 2(ij|ji) - (ij|ji),\\
\langle\Psi^{HF}|\hat{O}_2|\Psi^{a}_{i}\rangle &= \sum_{j}^{N}\langle ij||aj\rangle = \sum_{i}^{N/2} 2(ij|aj) - (ij|ja),\\
\langle\Psi^{HF}|\hat{O}_2|\Psi^{ab}_{ij}\rangle &= \langle ij||ab\rangle\,,\\
\langle\Psi^{HF}|\hat{O}_2|\Psi^{abc\dots}_{ijk\dots}\rangle &= 0.
\end{aligned} \tag{2.20}$$

Similar to the one-electron integrals, we find $(N-2)!$ possibilities of permuting $N-2$ electrons amongst the remaining spin orbitals. The factor $1/2$ again accounts for double counting. For the case with the singly

occupied determinant, electron 1 can be assigned to a spin orbital χ_i. Then electron 2 can be in any of the remaining $N-1$ spin orbitals. Concerning the doubly excited determinant, electron 1 can occupy the spin orbital χ_i and electron 2 χ_j, or vice versa. All other permutations don't give a contribution. Ultimately, matrix elements of two-electron operators that differ by more than two spin orbitals are zero. The second equality in the first and second line is a result of the spin integration, which we discuss in more detail in the next chapter of the Hartree-Fock method.

2.6 Second Quantization

To this point, we described many-body wave functions with the formalism of Slater determinants. We further constructed them to satisfy the antisymmetry and Pauli principle. Thereby, Slater determinants provide a rather intuitive approach to the many-electron theory. In the following, we want to introduce a more elegant way to describe wave functions, the second quantization. Here, the properties of determinants are transferred onto the algebraic properties of operators.

2.6.1 Annihilation and Creation Operators

We start with the introduction of the annihilation and creation operators by their action on an arbitrary Slater determinant. The creation operator is defined as:

$$\hat{a}_i^\dagger \left|\chi_j\, \chi_k \cdots \chi_N\right\rangle = \left|\chi_i\, \chi_j\, \chi_k \cdots \chi_N\right\rangle ,$$

while the annihilation operator reads:

$$\hat{a}_i \left|\chi_i\, \chi_j \cdots \chi_N\right\rangle = \left|\chi_j \cdots \chi_N\right\rangle .$$

On one hand $a_i^\dagger$ creates an electron in spin orbital χ_i. On the other hand, a_i annihilates an electron from spin orbital χ_i. We further introduce the occupation number operator

$$\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i .$$

With this, it is possible to express the Slater determinants in the occupation number basis

$$\left|\Psi\right\rangle = \left|n_i, n_j, \ldots, n_N\right\rangle ,$$

where the n's are called the occupation numbers being one if the respective spin orbital is occupied, zero otherwise. In this way, a variable number of electrons can be described. Another common representation of a Slater determinant in terms of second quantization is the Fermi vacuum formalism

$$\left|\Psi\right\rangle = \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k^\dagger \ldots \left|\,\right\rangle ,$$

where $\left|\,\right\rangle$ represents the orthonormalized ($\left\langle\,\middle|\,\right\rangle = 1$) vacuum state. Any determinant can now be expressed in terms of the action of creation operators on that vacuum state. The formalism of the second quantization with the creation and annihilation operators also satisfies the fermionic wave function properties specified for the Slater determinants. The action of the annihilation operator on the vacuum state yields a zero result

$$\hat{a}_i \,|\,\rangle = 0.$$

The vacuum state does not contain any electrons and therefore no electrons can be removed. Also, the Pauli exclusion principle is implicitly satisfied such that

$$\hat{a}_i^\dagger \,|\chi_k \dots \chi_l\rangle = 0 \qquad \text{if } \; i \in \{k, \dots, l\},$$

i.e. an electron cannot be created in a spin orbital that is already occupied. The antisymmetry is given by the fact that exchanging the order of two operators results in a sign change:

$$\hat{a}_i^\dagger \hat{a}_j = -\hat{a}_j \hat{a}_i^\dagger \qquad \text{for} \quad i \neq j.$$

The last two properties are equal to exchanging two columns or rows in a Slater determinant (s. Eq. 2.10), respectively. Further, the fermionic anti-commutation relations for the creation and annihilation operator read:

$$\begin{aligned} [\hat{a}_i^\dagger, \hat{a}_j^\dagger]_+ &= [\hat{a}_i, \hat{a}_j]_+ = 0, \\ [\hat{a}_i^\dagger, \hat{a}_j]_+ &= [\hat{a}_i, \hat{a}_j^\dagger]_+ = \delta_{ij}. \end{aligned}$$

Finally, the creation and annihilation operators are connected via adjunction:

$$\left(\hat{a}_i^\dagger\right)^\dagger = \hat{a}_i.$$

In summary, we are able to use the creation and annihilation operators to describe fermionic wave functions in the same way as the Slater determinant formalism. We exploit this fact in the next section as well as in the chapters 4.2 and 4.3.2. For more information about second quantization we recommend Ref. [27].

2.6.2 Wick's Theorem

For the entire theory of many-electron systems, we also need to express the one- and two-particle operators in terms of annihilation and creation operators. The expressions for $\hat{O}_1$ (s. Eq. 2.13) and $\hat{O}_2$ (s. Eq. 2.14) read:

$$\begin{aligned} \hat{O}_1 &= \sum_{pq} \langle p|\hat{h}|q\rangle \, \hat{a}_p^\dagger \hat{a}_q, \\ \hat{O}_2 &= \frac{1}{2} \sum_{pqrs} \langle pq|rs\rangle \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r. \end{aligned} \tag{2.21}$$

In the former, one electron in spin orbital χ_q gets annihilated, and another is created in spin orbital χ_p. For diagonal elements, i.e. for $p = q$, the occupation number operator $\hat{n}_p$ is present. For $\hat{O}_2$ we see the same pattern, but now with twice the number of operators in order to annihilate and create two electrons,

respectively. Note that in this section we consider a set of $2K \geq N$ spin orbitals with N being the number of electrons. The sums with the generic indices are limited to the total number of spin orbitals. For brevity, however, only the restriction of the sums involving occupied spin orbitals only is explicitly stated.
We want to introduce the normal order of second quantized operators. Normal order means that the annihilation operators are on the right of all creation operators. The benefit of normal-ordering is the better bookkeeping of non-zero matrix elements. Considering an arbitrary string of creation and annihilation operators $a_p a_q^\dagger a_r a_s^\dagger$, we can exploit the anti-commutation relations to get normal-ordering:

$$\begin{aligned} a_p a_q^\dagger a_r a_s^\dagger &= \delta_{pq} a_r a_s^\dagger - a_q^\dagger a_p a_r a_s^\dagger \\ &= \delta_{pq}\delta_{rs} - \delta_{pq} a_s^\dagger a_r - \delta_{rs} a_q^\dagger a_p + a_q^\dagger a_p a_s^\dagger a_r \\ &= \delta_{pq}\delta_{rs} - \delta_{pq} a_s^\dagger a_r - \delta_{rs} a_q^\dagger a_p + \delta_{ps} a_q^\dagger a_r - a_q^\dagger a_s^\dagger a_p a_r. \end{aligned}$$

In addition, we introduce the contraction between annihilation and creation operators. In general, the definition of two arbitrary contracted operators $\hat{\alpha}$ and $\hat{\gamma}$ reads:

$$\hat{\alpha}\hat{\gamma} \equiv \alpha\gamma - \{\alpha\gamma\},$$

where the notation $\{\alpha\gamma\}$ represents the normal-ordered form of the pair. For the annihilation and creation operator follows

$$\begin{aligned} \hat{a}_p\hat{a}_q &= \hat{a}_p\hat{a}_q - \{\hat{a}_p\hat{a}_q\} = \hat{a}_p\hat{a}_q - \hat{a}_p\hat{a}_q = 0, \\ \hat{a}_p^\dagger\hat{a}_q^\dagger &= \hat{a}_p^\dagger\hat{a}_q^\dagger - \{\hat{a}_p^\dagger\hat{a}_q^\dagger\} = \hat{a}_p^\dagger\hat{a}_q^\dagger - \hat{a}_p^\dagger\hat{a}_q^\dagger = 0, \\ \hat{a}_p^\dagger\hat{a}_q &= \hat{a}_p^\dagger\hat{a}_q - \{\hat{a}_p^\dagger\hat{a}_q\} = \hat{a}_p^\dagger\hat{a}_q - \hat{a}_p^\dagger\hat{a}_q = 0, \\ \hat{a}_p\hat{a}_q^\dagger &= \hat{a}_p\hat{a}_q^\dagger - \{\hat{a}_p\hat{a}_q^\dagger\} = \hat{a}_p\hat{a}_q^\dagger + \hat{a}_q^\dagger\hat{a}_p = \delta_{pq}. \end{aligned} \tag{2.22}$$

Note that in the last line we get a negative sign when evaluating the normal order of $\{\hat{a}_p\hat{a}_q^\dagger\}$.
Now, the so-called Wick theorem states that any string of annihilation and creation operators may be written as a linear combination of normal-ordered strings. For the sake of compactness, we look at a string of operators $ABCDE\ldots$ without the operator symbol "$\hat{}$" :

$$ABCDE\cdots = \{ABCDE\ldots\} + \sum_{singles} \{\overbrace{A}BCDE\ldots\} + \sum_{doubles} \{\overbrace{A}B\overbrace{C}DE\ldots\} + \ldots. \tag{2.23}$$

The lower limit of the sums indicates the number of pairwise contractions included in the respective summation. Due to the anti-commutation rules, where a sign change arises from switching two operators, the contraction introduces also a sign change $(-1)^N$ with N being the number of permutations required to bring the operators into adjacency. Given a large number of contractions, the sign of a fully contracted term can then be determined by the number of crossings in the contraction lines. If this number is odd, the sign of that term is negative, for an even number it's positive.

At this point, we are able to evaluate matrix elements using Wick's theorem. However, it is inconvenient and tedious to create every wave function from the true vacuum state $|\,\rangle$, especially when dealing with a large number of electrons. A better approach is to define normal-ordering relative to a given reference state, e.g. the Hartree-Fock ground state $|\Psi^{HF}\rangle$. Still, we need to redefine our nomenclature accordingly which leads to the particle-hole formalism of the Wick theorem. Here, the one-electron states occupied in $|\Psi^{HF}\rangle$ are called hole states, and the unoccupied ones are called particle states. This can be understood in the following way. A hole is created when an originally occupied state is affected by an annihilation operator. Similarly, a particle is created when an originally unoccupied state is affected by a creation operator. Now we find operators acting on hole states $(i, j \dots)$ and operators acting on particle states $(a, b \dots)$. This redefinition leads to two non-zero contractions compared to the original definition in Eq. 2.22:

$$\begin{aligned}
\hat{a}_i^\dagger \hat{a}_j &= \hat{a}_i^\dagger \hat{a}_j - \{\hat{a}_i^\dagger \hat{a}_j\} = \hat{a}_i^\dagger \hat{a}_j + \hat{a}_j \hat{a}_i^\dagger = \delta_{ij}, \\
\hat{a}_a \hat{a}_b^\dagger &= \hat{a}_a \hat{a}_b^\dagger - \{\hat{a}_a \hat{a}_b^\dagger\} = \hat{a}_a \hat{a}_b^\dagger + \hat{a}_b^\dagger \hat{a}_a = \delta_{ab}, \\
\hat{a}_a^\dagger \hat{a}_b &= \hat{a}_i \hat{a}_j^\dagger = 0.
\end{aligned}$$

Now we may rewrite the one- and two-electron operators in terms of the redefined normal-ordering. In contrast to Eq. 2.21, we look at a different but equivalent form of the electronic Hamiltonian. The motivation behind this is that we later start from this expression to derive the coupled cluster method (s. Sec. 4.2). We first deduce the new form of the electronic Hamiltonian and then look at the individual expressions for the $\hat{O}_1$ and $\hat{O}_2$ operators:

$$\begin{aligned}
\hat{H} &= \sum_{pq} \langle p|\hat{h}|q\rangle \, \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2} \sum_{pqrs} \langle pq|rs\rangle \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \\
&= \sum_{pq} \langle p|\hat{h}|q\rangle \, \hat{a}_p^\dagger \hat{a}_q + \frac{1}{4} \sum_{pqrs} \langle pq|rs\rangle \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \overset{r\leftrightarrow s}{+} \frac{1}{4} \sum_{pqrs} \langle pq|sr\rangle \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s \\
&= \sum_{pq} \langle p|\hat{h}|q\rangle \, \hat{a}_p^\dagger \hat{a}_q + \frac{1}{4} \sum_{pqrs} \langle pq|rs\rangle \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r - \frac{1}{4} \sum_{pqrs} \langle pq|sr\rangle \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \\
&= \sum_{pq} \langle p|\hat{h}|q\rangle \, \hat{a}_p^\dagger \hat{a}_q + \frac{1}{4} \sum_{pqrs} \langle pq||rs\rangle \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r.
\end{aligned} \tag{2.24}$$

Here, the antisymmetrized two-electron integrals $\langle pq||rs\rangle$ introduced in Sec. 2.4 appear again. The advantage of this representation will be clear in chapter 4 concerning post-Hartree-Fock theory. With this, the one-electron operator reads:

$$\hat{O}_1 = \sum_{pq} \langle p|\hat{h}|q\rangle \, \hat{a}_p^\dagger \hat{a}_q = \sum_{pq} \langle p|\hat{h}|q\rangle \, \{\hat{a}_p^\dagger \hat{a}_q\} + \sum_{pq} \langle p|\hat{h}|q\rangle \, \delta_{pq} = \sum_{pq} \langle p|\hat{h}|q\rangle \, \{\hat{a}_p^\dagger \hat{a}_q\} + \sum_{i}^{N} \langle i|\hat{h}|i\rangle \, .$$

Although we look at a slightly different two-electron operator $\hat{O}_2$ compared to Eq. 2.21, the algebra is the same:

$$\hat{O}_2 = \frac{1}{4}\sum_{pqrs}\langle pq||rs\rangle\,\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r = \frac{1}{4}\sum_{pqrs}\langle pq||rs\rangle\,\{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\} - \frac{1}{4}\sum_{qr}\sum_{i}^{N}\langle iq||ri\rangle\,\{\hat{a}_q^\dagger\hat{a}_r\} + \frac{1}{4}\sum_{pr}\sum_{i}^{N}\langle pi||ri\rangle\,\{\hat{a}_p^\dagger\hat{a}_r\}$$
$$+ \frac{1}{4}\sum_{qs}\sum_{i}^{N}\langle iq||is\rangle\,\{\hat{a}_q^\dagger\hat{a}_s\} - \frac{1}{4}\sum_{ps}\sum_{i}^{N}\langle pi||is\rangle\,\{\hat{a}_p^\dagger\hat{a}_s\} - \frac{1}{4}\sum_{ij}^{N}\langle ij||ji\rangle + \frac{1}{4}\sum_{ij}^{N}\langle ij||ij\rangle$$
$$= \frac{1}{4}\sum_{pqrs}\langle pq||rs\rangle\,\{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\} + \sum_{pr}\sum_{i}^{N}\langle pi||ri\rangle\,\{\hat{a}_p^\dagger\hat{a}_r\} + \frac{1}{2}\sum_{ij}^{N}\langle ij||ij\rangle\,.$$

In the last line we renamed several indices and used the fact that $\langle pq||rs\rangle = -\langle pq||sr\rangle$. In the end, the electronic Hamiltonian in normal-ordered second quantization reads:

$$\hat{H} = \frac{1}{4}\sum_{pqrs}\langle pq||rs\rangle\,\{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\} + \sum_{pq}\langle p|\hat{h}|q\rangle\,\{\hat{a}_p^\dagger\hat{a}_q\} + \sum_{pr}\sum_{i}^{N}\langle pi||ri\rangle\,\{\hat{a}_p^\dagger\hat{a}_r\} + \sum_{i}^{N}\langle i|\hat{h}|i\rangle + \frac{1}{2}\sum_{ij}^{N}\langle ij||ij\rangle\,. \tag{2.25}$$

In summary, second quantization is a very elegant and powerful formalism that is especially advantageous in solid-state physics, since an arbitrarily large and varying number of particles can be described.

3 Hartree-Fock Approximation

One of the first attempts to find an approximate solution to the Schrödinger equation was given by Douglas Rayner Hartree and Wladimir Alexandrowitsch Fock [28]. Thereby, the interactions of particles with each other are simplified in such a way that they no longer interact in pairs, but with a field generated on average by all other particles. Therefore, the Hartree-Fock method belongs to the group of mean-field methods. Not only does it provide reasonably accurate results, which is especially important to get a basic understanding of underlying systems, but it is also essential nowadays because many modern electronic structure calculations are based on its solution.

In the following sections we derive the Hartree-Fock equations from a general point of view using the standard electronic Hamiltonian for molecular systems (s. Eq. 2.4). Moreover, we explain the resulting Koopman and Brillouin theorem. After that, we discuss in detail the restricted closed-shell Hartree-Fock formalism which will be used for the rest of this work. In the penultimate section, we outline the necessary steps to solve the Hartree-Fock equations for the electronic Hamiltonian in a general manner. For the implementation, we modify these general considerations according to our model and explain the required steps to find a solution. We want to point out that our derivations in this chapter closely follow the book of Szabo and Ostlund [24].

3.1 Hartree-Fock Equations

We begin with the time-independent electronic Schrödinger equation

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\,. \tag{3.1}$$

Since Hartree-Fock is a variational method, we seek to find the best approximation to the ground state energy. In order to do that, we briefly summarize the variational principle. The energy of the system can be expressed as a functional of an arbitrary wave function

$$E[\Psi] = \langle\Psi|\hat{H}|\Psi\rangle\,.$$

Variation means modifying the parameters on which Ψ depends by a small amount, resulting in $\Psi \to \Psi+\delta\Psi$. Accordingly, we write the variation of the energy functional as follows:

$$\delta E = \delta\,\langle\Psi|\hat{H}|\Psi\rangle = \langle\delta\Psi|\hat{H}|\Psi\rangle + \langle\Psi|\hat{H}|\delta\Psi\rangle\,.$$

Here, δ works just like a differential operator. To find the best possible approximation of Ψ we use the method of Lagrange multipliers where we minimize the total energy such that the variation of the functional disappears. Further, we set the additional constraint that the wave function is normalized, i.e. $\langle\Psi|\Psi\rangle = 1$. The Lagrangian is given by:

$$\mathcal{L} = \langle\Psi|\hat{H}|\Psi\rangle - E\,(\langle\Psi|\Psi\rangle - 1)\,,$$

where E has the function of the Lagrange multiplier. By introducing a basis $|\Psi\rangle = \sum_i c_i\,|\tilde{\Psi}_i\rangle$ with the coefficients c_i and by setting $\delta\mathcal{L} = 0$ it can be shown that a generalized eigenvalue equation $Hc = ESc$

with the overlap matrix $S_{ij} = \langle\tilde{\Psi}_i|\tilde{\Psi}_j\rangle$ is obtained. The lowest eigenvalue provides the best approximation and simultaneously an upper limit of the true lowest Hamiltonian eigenvalue ϵ_0

$$E = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq \epsilon_0.$$

In the case of orthonormal basis functions $|\tilde{\Psi}_i\rangle$ the overlap matrix is just δ_{ij}. Then we get a simple eigenvalue equation $Hc = Ec$.

Now, we want to apply these general considerations to the Hartree-Fock method. Here, the energy is a functional of the occupied spin orbitals $\{\chi_i\}$ from the single determinant $|\Psi^{HF}\rangle = |\chi_1\chi_2\dots\chi_i\chi_j\dots\chi_N\rangle$. With this the Langrangian reads:

$$\mathcal{L}[\{\chi_i\}] = E_0[\{\chi_i\}] - \sum_i^N \sum_j^N \epsilon_{ji}\Big(\langle\chi_i|\chi_j\rangle - \delta_{ij}\Big),$$

where E_0 is the determinantal energy and ϵ_{ij} are the Lagrange multipliers. Further, we can exploit the fact that the $\mathcal{L} \in \mathbb{R}$ and the properties of the scalar product to verify that $\epsilon_{ji} = \epsilon_{ij}^*$.
The expectation value E_0 of the single Slater determinant can be determined by the Slater-Condon rules introduced in Sec. 2.5

$$E_0[\{\chi_i\}] = \sum_{i=1}^N \langle i|\hat{h}|i\rangle + \frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N \Big(\langle ij|ij\rangle - \langle ij|ji\rangle\Big). \tag{3.2}$$

For the sake of compactness, we carry on with the shorthand notation of the spin orbitals $\chi_i \equiv i$. The minimization of E_0 is now obtained by minimizing the Lagrangian $\mathcal{L}$

$$\delta\mathcal{L} = \delta E_0 - \sum_{ij}^N \epsilon_{ji}\,\delta\langle i|j\rangle = 0. \tag{3.3}$$

Since the variation acts like a differential operator, we investigate each term in Eq. 3.3 separately. Let us examine the constraint term first:

$$\begin{aligned}\sum_{ij}\epsilon_{ji}\,\delta\langle i|j\rangle &= \sum_{ij}\epsilon_{ji}\langle\delta i|j\rangle + \sum_{ij}\epsilon_{ji}\langle i|\delta j\rangle = \sum_{ij}\epsilon_{ji}\langle\delta i|j\rangle + \sum_{ij}\epsilon_{ij}\langle j|\delta i\rangle \\ &= \sum_{ij}\epsilon_{ji}\langle\delta i|j\rangle + \sum_{ij}\epsilon_{ji}^*\langle\delta i|j\rangle^* = \sum_{ij}\epsilon_{ji}\langle\delta i|j\rangle + c.c.\end{aligned}$$

We use the fact that $\epsilon_{ji} = \epsilon_{ij}^*$ and that i and j are just dummy variables that can be interchanged. In the last line *c.c.* denotes the complex conjugate of the first term. Next, we look at the variation of the energy stemming from the one-electron integrals:

$$\delta \sum_i \langle i|\hat{h}|i\rangle = \sum_i \langle \delta i|\hat{h}|i\rangle + \sum_i \langle i|\hat{h}|\delta i\rangle = \sum_i \langle \delta i|\hat{h}|i\rangle + \sum_i \langle \delta i|\hat{h}|i\rangle^* = \sum_i \langle \delta i|\hat{h}|i\rangle + c.c.$$

Finally, we apply the variational ansatz on the energy contribution from the two-electron integrals:

$$\begin{aligned}
&\delta \frac{1}{2} \sum_{ij} \Big(\langle ij|ij\rangle - \langle ij|ji\rangle \Big) \\
&= \frac{1}{2} \sum_{ij} \Big(\langle \delta ij|ij\rangle + \langle i\delta j|ij\rangle + \langle ij|\delta ij\rangle + \langle ij|i\delta j\rangle \Big) - \frac{1}{2} \sum_{ij} \Big(\langle \delta ij|ji\rangle + \langle i\delta j|ji\rangle + \langle ij|\delta ji\rangle + \langle ij|j\delta i\rangle \Big) \\
&= \frac{1}{2} \sum_{ij} \Big(\langle \delta ij|ij\rangle + \langle \delta ij|ij\rangle - \langle \delta ij|ji\rangle - \langle \delta ij|ji\rangle + c.c. \Big) = \sum_{ij} \Big(\langle \delta ij|ij\rangle - \langle \delta ij|ji\rangle + c.c. \Big),
\end{aligned}$$

where we exploit the fact that $\langle \delta ij|ij\rangle^* = \langle ij|\delta ij\rangle$ and $\langle i\delta j|ij\rangle = \langle \delta ji|ji\rangle \overset{i\leftrightarrow j}{=} \langle \delta ij|ij\rangle$.
Combing these three terms yields the first variation in $\mathcal{L}$:

$$\delta\mathcal{L} = \sum_i^N \langle \delta i|\hat{h}|i\rangle + \sum_{ij}^N \langle \delta ij|ij\rangle - \langle \delta ij|ji\rangle - \sum_{ij}^N \epsilon_{ji} \langle \delta i|j\rangle + c.c. = 0. \tag{3.4}$$

Further, we can rewrite Eq. 3.4 in terms of integrals:

$$\begin{aligned}
\delta\mathcal{L} = \sum_i^N \int d\vec{x}_1 \, \delta\chi_i^*(1) \Bigg[&\hat{h}(1)\chi_i(1) + \sum_j^N \int d\vec{x}_2 \, \frac{||\chi_j(2)||^2}{\vec{r}_{12}} \chi_i(1) \\
&- \sum_j^N \int d\vec{x}_2 \, \frac{\chi_j^*(2)\chi_i(2)}{\vec{r}_{12}} \chi_j(1) - \sum_j^N \epsilon_{ji}\chi_j(1) \Bigg] + c.c. = 0,
\end{aligned} \tag{3.5}$$

where the spin orbitals $\chi_i(1)$ and $\chi_i(2)$ refer to two arbitrary electrons of the system. Since electrons are indistinguishable, we could use every labeling, but for convenience, we stick with electron 1 and electron 2. Also, as mentioned earlier, in the derivation of the Hartree-Fock method we stick to the general $3D$ Coulomb potential $1/\vec{r}_{12} = 1/|\vec{r}_1 - \vec{r}_2|$ instead of the $1D$ two-electron potential of our model.
The first two-electron integral represents the averaged Coulomb potential at $\vec{x}_1$ arising from an electron in χ_j. By summing over all orbitals, we obtain the total averaged potential acting on the electron in χ_i. The Coulomb operator acting on the spin orbital χ_i reads:

$$\hat{J}_j(1)\chi_i(1) = \left[\int dx_2 \frac{||\chi_j(2)||^2}{\vec{r}_{12}} \right] \chi_i(1). \tag{3.6}$$

Furthermore, due to the antisymmetric nature of the single Slater determinant, we get a second two-integral term, which has no classical interpretation. However, we can imagine this additional contribution as a correction to the overestimation of the repulsion between electrons which are naively assumed to move independently, but in reality are correlated with each other. Similarly, we define this so-called exchange operator

$$\hat{K}_j(1)\chi_i(1) = \left[\int dx_2 \frac{\chi_j^*(2)\chi_i(2)}{\vec{r}_{12}}\right]\chi_j(1). \tag{3.7}$$

In contrast to the Coulomb term, we find a non-local potential in the exchange term. If we want to express the expectation values of these two operators with respect to the spin orbital χ_i, we get:

$$\langle\chi_i(1)|\hat{J}_j(1)|\chi_i(1)\rangle = \langle ij|ij\rangle = J_{ij},$$
$$\langle\chi_i(1)|\hat{K}_j(1)|\chi_i(1)\rangle = \langle ij|ji\rangle = K_{ij}.$$

Now we come back to Eq. 3.5. Here, $\delta\chi_i^*(1)$ is arbitrary and therefore the quantity in the square brackets is zero for all i. With the definition of the Coulomb and exchange operator we obtain:

$$\underbrace{\left[\hat{h}(1) + \sum_j^N \hat{J}_j(1) - \hat{K}_j(1)\right]}_{\hat{f}(1)}\chi_i(1) = \sum_j^N \epsilon_{ji}\chi_j(1). \tag{3.8}$$

Eploiting the definition of the Fock operator $\hat{f}(1)$ from Eq. 3.8 yields:

$$\hat{f}(1)\chi_i(1) = \sum_j^N \epsilon_{ji}\chi_j(1). \tag{3.9}$$

Finally, we need to consider that spin orbitals can be mixed amongst each other without changing the expectation value $E^{HF} = \langle\Psi^{HF}|\hat{H}|\Psi^{HF}\rangle$. For a single determinant, any expectation value is invariant under unitary transformation of the spin orbitals. Thus, there exists no unique set of spin orbitals and subsequently no significant physical meaning can be given to a certain set of them. Hence, more localized spin orbitals are equally valid as delocalized ones. However, the invariance of a single determinant under unitary transformation can be used to simplify Eq. 3.9 to an eigenvalue equation for a particular set of spin orbitals. For this, we consider the following unitary transformation

$$\tilde{\chi} = U\chi,$$

where $U^\dagger = U^{-1}$ holds. Since the trace of the Fock operator $\hat{f}$ is independent of the choice of the basis set, $\hat{f}$ transforms like:

$$\tilde{\hat{f}} = U\hat{f}U^{\dagger}.$$

Now Eq. 3.9 may be written as

$$\tilde{\hat{f}}\tilde{\chi} = U^{\dagger}\epsilon U\tilde{\chi},$$

where ϵ is a Hermitian matrix that reasons to choose U in such a way that it diagonalizes the matrix of the Lagrange multipliers. By doing so, we get the canonical Hartree-Fock equations

$$\hat{f}\,|\chi_i\rangle = \epsilon_i\,|\chi_i\rangle\,. \tag{3.10}$$

Here, we drop the tildes and remark that the ϵ_i are the eigenvalues of the Fock operator $\hat{f}$. To find a solution to the canonical Hartree-Fock equation we need to introduce a basis set and consequently solve a set of matrix equations.

3.2 Koopman and Brillouin Theorem

Before we go into the details of the actual solution of the Hartree-Fock equations, we want to discuss some aspects of the eigenvalue equation itself. As we have seen, for an N-electron system with the single Slater determinant $|\chi_1\chi_2\dots\chi_i\chi_j\dots\chi_N\rangle$ we get the following equation:

$$\hat{f}\,|\chi_i\rangle = \epsilon_i\,|\chi_i\rangle\,,$$

for the N occupied spin orbitals $\{\chi_i\}$. Once the set $\{\chi_i\}$ is known, the Fock operator $\hat{f}$ is a well-defined Hermitian operator with an infinite number of eigenfunctions

$$\hat{f}\,|\chi_p\rangle = \epsilon_p\,|\chi_p\rangle\,. \tag{3.11}$$

We want to point out again that the subscript i refers to occupied spin orbitals while p indicates generic spin orbitals. The eigenvalues ϵ_p are the orbital energies, whereas the first N-lowest belong to the occupied orbitals. The remaining number of spin orbitals are attributed to the unoccupied or virtual orbitals, labeled with $(a, b, \dots)$. We use the shorthand notation and project the state $\langle q|$ onto Eq. 3.11 to get:

$$\begin{aligned}
&\langle q|\hat{f}|p\rangle = \epsilon_p\,\langle q|p\rangle\,,\\
&\left(\langle q|\hat{h}|p\rangle + \sum_{r}^{N}\langle qr||pr\rangle\right)\delta_{pq} = \epsilon_p\delta_{pq}.
\end{aligned}$$

Only for $p = q$ we get a non-vanishing solution and thus the orbital energies to the corresponding orbitals. Further, the occupied spin orbital $|i\rangle$ with eigenvalue ϵ_i can be interpreted as the energy of an electron in that orbital. On the other side, an unoccupied spin orbital $|a\rangle$ with the orbital energy ϵ_a represents the

energy needed to add one additional electron to that state.
This brings us to the Koopman theorem, where we only present the main outcomes. The significance of the orbital energies can be pointed out by looking at two cases: Adding and removing an electron with respect to an N-electron state declared as $|^N\Psi^{HF}\rangle$. These two processes can be viewed in terms of second quantization introduced in Sec. 2.6

$$|^{N+1}\Psi^a\rangle = a_a^\dagger\,|^N\Psi^{HF}\rangle\,,$$
$$|^{N-1}\Psi_i\rangle = a_i\,|^N\Psi^{HF}\rangle\,,$$

where in the first line one electron is added to the virtual spin orbital $|a\rangle$ and in the second line an electron is removed from spin orbital $|i\rangle$. Now, by comparing the energies of these determinants with a different number of electrons, the ionization potential can be defined. The result is stated in Koopman's theorem, namely that the ionization potential for removing an electron from $|i\rangle$ is just the negative orbital energy

$$IP = {}^{N-1}E_i - {}^{N}E_0 = -\epsilon_i.$$

Similarly, adding an electron to $|a\rangle$ gives the electron affinity

$$EA = {}^{N}E_0 - {}^{N+1}E^a = -\epsilon_a.$$

Koopman's theorem yields an approximation since the relaxation of the orbitals is not considered (frozen orbital approximation). Further, the approach with the single determinant as well as neglecting the electron-electron correlation effects obtained by post-Hartree-Fock methods leads to inaccuracies.

The second important theorem resulting from the Hartree-Fock eigenvalue equation is the so-called Brillouin theorem. It states that singly excited determinants $|\Psi_i^a\rangle$ do not interact with the reference Hartree-Fock determinant $|\Psi^{HF}\rangle$. Again, we don't provide the full derivation, but with the Slater-Condon rules introduced in Sec. 2.5, it follows that

$$\langle\Psi^{HF}|\hat{H}|\Psi_i^a\rangle = \langle i|\hat{h}|a\rangle + \sum_p^N \langle ip||ap\rangle = f_{ia} = \epsilon_a\delta_{ia}.$$

An intuitive explanation is given considering the matrix element which mixes singly excited determinants with $|\Psi^{HF}\rangle$. This matrix element must be equal to an off-diagonal element of the Fock matrix. However, solving the Hartree-Fock eigenvalue equations implies that the off-diagonal f_{pq} are zero. Therefore solving the Hartree-Fock equation is equivalent to ensuring that the reference determinant $|\Psi^{HF}\rangle$ does not mix with any of the singly excited ones. For the total derivation, we recommend the book of Szabo and Ostlund (s. Ref. [24]).

3.3 Restricted Closed-Shell Hartree-Fock Theory

In order to solve the Hartree-Fock equation, we need to specify what kind of orbitals we are dealing with. Namely, it is differentiated between unrestricted and restricted as well as open- and closed-shell spin orbitals. Thereby it is distinguished between the type of spatial functions for the two spin functions α (spin up) and β (spin down) as well as the occupation of the spin orbitals. In this work, our main interest lies in

(restricted) closed-shell spin orbitals where all electrons are paired such that all spatial orbitals are doubly occupied. This further implies an even number of electrons. We write an arbitrarily restricted spin orbital according to Eq. 2.8 as follows:

$$\chi_p(\vec{x}) = \begin{cases} \Phi(\vec{r})\alpha(\bar{\omega}) \\ \Phi(\vec{r})\beta(\bar{\omega}) \end{cases},$$

together with the restricted ground state function

$$|\Psi^{HF}\rangle = |\chi_1\chi_2\dots\chi_N\rangle = |\Phi_1\bar{\Phi}_1\dots\Phi_{N/2}\bar{\Phi}_{N/2}\rangle\,,$$

where for the last equality we used the fact that two spatial orbitals, belonging to the two different spin functions, are equal. We can exploit this further in order to derive the Hartree-Fock equation as a spatial eigenvalue equation. We project the spin up function on the following Hartree-Fock equation:

$$\hat{f}(\vec{x}_1)\Phi_p(\vec{r}_1)\alpha(\bar{\omega}_1) = \epsilon_p\Phi_p(\vec{r}_1)\alpha(\bar{\omega}_1).$$

Multiplying by $\alpha^*(\bar{\omega}_1)$ gives:

$$\left[\int d\bar{\omega}_1\,\alpha^*(\bar{\omega}_1)\hat{f}(\vec{x}_1)\alpha(\bar{\omega}_1)\right]\Phi_p(\vec{r}_1) = \epsilon_p\Phi_p(\vec{r}_1),$$

where we use $\int d\bar{\omega}\ \alpha^*(\bar{\omega})\alpha(\bar{\omega}) = 1$ to obtain the right hand side. Inserting the definition of the Fock operator (s. Eq. 3.8) yields

$$\begin{aligned}
\hat{f}(\vec{r}_1)\Phi_p(\vec{r}_1) &= \hat{h}(\vec{r}_1)\Phi_p(\vec{r}_1) \\
&+ \sum_i^{N/2}\int d\bar{\omega}_1\int d\bar{\omega}_2\int d\vec{r}_2\,\alpha^*(\bar{\omega}_1)\Phi_i^*(\vec{r}_2)\alpha^*(\bar{\omega}_2)\vec{r}_{12}^{-1}\Phi_i(\vec{r}_2)\alpha(\bar{\omega}_2)\alpha(\bar{\omega}_1)\Phi_p(\vec{r}_1) \\
&+ \sum_i^{N/2}\int d\bar{\omega}_1\int d\bar{\omega}_2\int d\vec{r}_2\,\alpha^*(\bar{\omega}_1)\Phi_i^*(\vec{r}_2)\beta^*(\bar{\omega}_2)\vec{r}_{12}^{-1}\Phi_i(\vec{r}_2)\beta(\bar{\omega}_2)\alpha(\bar{\omega}_1)\Phi_p(\vec{r}_1) \\
&- \sum_i^{N/2}\int d\bar{\omega}_1\int d\bar{\omega}_2\int d\vec{r}_2\,\alpha^*(\bar{\omega}_1)\Phi_i^*(\vec{r}_2)\alpha^*(\bar{\omega}_2)\vec{r}_{12}^{-1}\Phi_i(\vec{r}_1)\alpha(\bar{\omega}_1)\alpha(\bar{\omega}_2)\Phi_p(\vec{r}_2) \\
&- \sum_i^{N/2}\int d\bar{\omega}_1\int d\bar{\omega}_2\int d\vec{r}_2\,\alpha^*(\bar{\omega}_1)\Phi_i^*(\vec{r}_2)\beta^*(\bar{\omega}_2)\vec{r}_{12}^{-1}\Phi_i(\vec{r}_1)\beta(\bar{\omega}_1)\alpha(\bar{\omega}_2)\Phi_p(\vec{r}_2) \\
&= \epsilon_p\Phi_p(\vec{r}_1),
\end{aligned} \tag{3.12}$$

where the closed-shell Fock operator reads $\hat{f}(\vec{r}_1) = \int d\bar{\omega}_1\,\alpha^*(\bar{\omega}_1)\hat{f}(\vec{x}_1)\alpha(\bar{\omega}_1)$. In addition, the four sums in Eq. 3.12 arise due to the splitting of the sum over all N electrons in the Coulomb and exchange term to sums over the occupied spin orbitals with the α and β spin functions, respectively. In passing we mention that the same results are obtained by projection of β on the Hartree-Fock equation. Carrying out the integrations

over the spin variables results in a vanishing last term because of the spin orthogonality relations (s. Eq. 2.7). The closed-shell Fock operator reads

$$\hat{f}(\vec{r}_1) = \hat{h}(\vec{r}_1) + \sum_i^{N/2} \int d\vec{r}_2 \, \Phi_i^*(\vec{r}_2)(2 - \hat{P}_{12})\vec{r}_{12}^{-1}\Phi_i(\vec{r}_2),$$

or more compactly

$$\hat{f}(1) = \hat{h}(1) + \sum_i^{N/2} 2\hat{J}_i(1) - \hat{K}_i(1). \tag{3.13}$$

Here, the Coulomb and exchange operators are defined like in Eqs. 3.6 and 3.7 but with spatial orbitals. Comparing the general Fock operator (s. Eq. 3.8) with the derived closed-shell Fock operator (s. Eq. 3.13), we see two similar expressions but the latter has a factor 2 in front of the Coulomb operator and a sum running over $N/2$ electrons. For a closed-shell determinant we are now able to write the ground state energy

$$\begin{aligned} E^{HF} = \langle\Psi_0|\hat{H}|\Psi_0\rangle &= 2\sum_i^{N/2}(i|\hat{h}|i) + \sum_{ij}^{N/2} 2(ij|ij) - (ij|ji) \\ &= 2\sum_i^{N/2} h_{ii} + \sum_{ij}^{N/2} 2J_{ij} - K_{ij}, \end{aligned} \tag{3.14}$$

where the factor 2 in front of the first term stems from the fact that the core Hamiltonian is the same for both spin functions in the same orbital. Furthermore, the orbital energies in the close-shell formalism are given by

$$\begin{aligned} \epsilon_p = f_{pp} &= (p|\hat{h}|p) + \sum_j^{N/2} 2(jp|jp) - (jp|pj) \\ &= h_{pp} + \sum_j^{N/2} 2J_{jp} - K_{jp}. \end{aligned} \tag{3.15}$$

The sum of the orbital energies ϵ_p, however, does not give the ground state energy E^{HF}. This is due to the fact that the electron-electron interaction energy is counted twice. Taking that into account, E^{HF} can also be written in terms of the occupied orbital energies

$$E^{HF} = 2\epsilon_i - \sum_{ij}^{N/2} 2J_{ij} - K_{ij}. \tag{3.16}$$

This can be easily verified by comparing Eq. 3.14 and Eq. 3.15. We want to stress that in contrast to the orbital energies, the ground state energy is calculated only over the occupied orbitals.

At this point, we want to make a connection to our model system. Up to now, the derivation of the Hartree-Fock equation and the corresponding energies is done with the general electronic Hamiltonian usually used for the calculation of molecules. The equations do not change for our model, i.e. we just need to plug in our core Hamiltonian and two-electron potential (s. Eq. 2.6). However, the derived Hartree-Fock equation represents an integro-differential equation. The numerical solution of such a problem is difficult. Therefore, it is common to introduce a basis set with already-known functions in order to convert the problem to a set of algebraic equations which can be solved by standard matrix techniques. The general approach is to expand the spatial orbitals in a linear expansion

$$\Phi_p(\vec{r}) = \sum_{\mu}^{K} C_{\mu p}\, \phi_\mu(\vec{r}) \qquad p = 1, 2, \ldots, K,$$

where $\{\phi_\mu(\vec{r})\}$ is a set of already known basis functions and $C_{\mu p}$ are the corresponding coefficients. Note, that the Greek subscripts denote the representation in the respective basis. With this, the Hartree-Fock problem reduces to finding the optimal coefficients $C_{\mu p}$. One of the standard basis sets for ϕ_μ is the Gaussian basis set. Returning to the comparison with our problem, we want to emphasize that we don't need to choose a basis set that provides the most efficient and accurate functions possible as is the case for realistic systems. As we show later, our approach is less educated but nevertheless valid allowing to obtain accurate results.
Still, we want to summarize the main outcome when solving the Hartree-Fock equations for realistic systems. The integration of the integro-differential Hartree-Fock equation over the basis functions leads to the Roothan equation

$$FC = SC\epsilon,$$

which is a matrix equation with the Fock matrix $F_{\mu\nu} = \int d\vec{r}_1 \phi^*_\mu(1)\hat{f}(1)\phi_\nu(1)$, the C matrix containing the basis coefficients and the overlap matrix $S_{\mu\nu} = \int d\vec{r}_1 \phi^*_\mu(1)\phi_\nu(1)$. Here, the columns of C represent the molecular orbitals and ϵ the corresponding energies. In practice, the Roothan equation is transformed into an eigenvalue equation. This is done by diagonalizing the overlap matrix S with a unitary matrix U according to $X = Us^{-1/2}$, where s is a diagonal matrix containing the square roots of the eigenvalues of S. The final result is:

$$\tilde{F}\tilde{C} = \tilde{C}\epsilon.$$

This representation consequently also allows us to write the ground state energy in terms of the Fock matrix. Without proof (s. Ref. [24]), this gives

$$E^{HF} = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}(H^{core}_{\mu\nu} + F_{\mu\nu}),$$

where $H^{core}_{\mu\nu} = \int d\vec{r}_1 \phi^*_\mu(1)\hat{h}(1)\phi_\nu(1)$ is the core Hamiltonian matrix and $P_{\mu\nu} = 2\sum_i^{N/2} C_{\mu i}C^*_{\nu i}$ is known as the density matrix.

3.4 Self-Consistent Field Procedure

Now we are able to describe the actual steps for the implementation of the solver for the restricted closed-shell Hartree-Fock equation. Since the Hartree-Fock equation is a non-linear equation, the solution needs to be found iteratively. Therefore, the Hartree-Fock method is often also called the self-consistent field (SCF) method.
The main steps of the SCF procedure are:

1. Specify system-specific parameters (number of electrons, nuclear coordinates, atomic numbers, basis set).
2. Calculate all necessary integrals to get: H^{core}, S, J and K.
3. Diagonalize the overlap matrix according to $s = U^{\dagger}SU$ and calculate the transformation matrix $X = Us^{-1/2}$.
4. Take an initial guess of the coefficients C and thus the density matrix P.
5. Set up the Fock matrix F and transform it $\tilde{F} = X^{\dagger}FX$.
6. Diagonalize $\tilde{F}$ to get $\tilde{C}$ and the orbital energies ϵ.
7. Do the back transformation $C = X\tilde{C}$ and calculate the new density matrix P.
8. Check for convergence of, for example, the density matrix. If convergence is not obtained, return to step 5.
9. Calculate desired quantities, such as E^{HF}.

After specifying the initial parameters, the one- and two-electron integrals arising in $\hat{H}^{core}$ and $\hat{S}$ as well as $\hat{J}$ and $\hat{K}$ need to be calculated. If we assume e.g. a Gaussian basis set, as mentioned earlier, we need to consider that the Gaussian functions are normalized, but not orthogonal. Therefore, the overlap matrix must be transformed using a symmetric orthogonalization method or a canonical orthogonalization. Before going into the SCF iteration, an initial guess for the coefficient matrix C is needed. More trivial choices are just random numbers or even the zero matrix. However, there also exist more advanced choices for the initial coefficients [29]. After this, in the SCF loop the Fock matrix is set up and transformed according to the introduced orthogonal transformation. The diagonalization procedure of the Fock matrix yields the transformed expansion coefficients and orbital energies. Then, the expansion coefficients are back-transformed in order to update the density matrix, which is used to calculate the updated Fock matrix. The SCF algorithm terminates when convergence is reached, and there are various options available for selecting a convergence criterion. The change in energy of the lowest orbital, e.g. , turns out to be computationally very efficient. However, more evolved criteria consider the change of the orbital wave functions or the density matrix.

3.5 Implementation

The above listing provides the fundamentals for realistic molecule calculations. In our Hartree-Fock procedure, we have a fixed number of electrons, namely two. At first, we set up the core Hamiltonian (s. Eq. 2.6). Numerical calculations require the introduction of a finite grid, which we choose to have a linear spacing between the grid points

$$x_i = -x_{min} + i\Delta x, \qquad i = 0, 1, 2, \ldots, N.$$

Since we are working in the position space, the potential term is just a diagonal matrix. On the other hand, the kinetic term is not, so we use the central finite difference method where the second derivative is approximated by:

$$\left.\frac{d^2}{dx^2}\Phi(x)\right|_{x=i} = \frac{\Phi(x_{i+1}) - 2\Phi(x_i) + \Phi(x_{i-1})}{\Delta x^2} + \mathcal{O}(\Delta x)^2,$$

which can be represented as a tridiagonal matrix. Thus, the core Hamiltonian matrix is calculated by a simple matrix addition. The eigenfunctions of $\hat{H}^{core}$, defined via the time-independent Schrödinger equation $\langle x_i|\hat{H}^{core}|x_j\rangle \langle x_j|\Phi_p\rangle = E_p \langle x_i|\Phi_p\rangle$, are obtained by diagonalization of the core Hamiltonian matrix with already existing numerical packages. Special attention needs to be paid to the dimensionality of $|\Phi_p\rangle$. Although the eigenvalue solver gives orthonormal eigenfunctions, the physical dimension is not considered. Therefore, the eigenfunctions need to be additionally normalized:

$$\langle x_j|\Phi\rangle_{norm} = \frac{\langle x_j|\Phi\rangle}{\sqrt{\Delta x}},$$

In this way, the eigenfunctions represent also physically valid wave functions, i.e. expectation values such as $\langle\Phi_p|\hat{H}|\Phi_p\rangle = E$ are of dimension energy. However, in the rest of the thesis we assume that this additional normalization is implicitly in $|\Phi_p\rangle$.
Now, the steps which do not need to be repeated in the SCF algorithm (s. grey boxes in Fig. 1) are completed. The obtained wave functions are used to construct the reduced density matrix. The density operator reads:

$$\hat{\rho} = \sum_{pq} |\Phi_p\rangle \langle\Phi_q| ,$$

with the generic spatial orbitals $|\Phi_p\rangle$. Introducing the reduced density matrix by taking the partial trace of $\hat{\rho}$ allows to look at the subsystem containing a sum only over the occupied spatial orbitals. Since in the restricted closed-shell formalism each spatial orbital contains two electrons with opposite spin, the sum goes to $N/2$. The reduced density in matrix form is

$$\rho_{ij}^{red} = \sum_{ij}^{N/2} |\Phi_i\rangle \langle\Phi_j| .$$

The eigenfunctions $|\Phi_i\rangle$ are already orthonormal and we don't need to introduce a basis set. Consequently, in contrast to the general procedure (s. Sec. 3.4), no overlap matrix needs to be set up and no unitary transformation is needed either. In our case, we can immediately calculate the Coulomb and exchange matrices in order to construct the Fock matrix. We calculate the Coulomb and exchange matrices according to Eqs. 3.6 and 3.7 using the reduced density matrix:

$$J_{ii} = \sum_j \Delta x \, \frac{\rho_{ii}^{red}}{|x_i - x_j| + \sigma},$$
$$K_{ij} = \sum_j \Delta x \, \frac{\rho_{ij}^{red}}{|x_i - x_j| + \sigma}.$$

While the Coulomb operator is a diagonal matrix, the non-local exchange operator has also off-diagonal elements. This is expressed via $\rho_{ij}^{red} = \sum_{ij}^{N/2} |\Phi_i\rangle \langle\Phi_j|$. The Fock matrix results in $F_{ij} = H_{ij}^{core} + 2J_{ii} - K_{ij}$. For the case $i = j$, the Coulomb and exchange operator are the same, and therefore they cancel each other. This case is especially interesting for our model since we only have two electrons. As discussed in section 3.1, the exchange operator only acts on electrons with the same spin. If both electrons are in the lowest energy state $|\Phi^{HF}\rangle$, this could also be viewed as an interaction of the electron with itself. Nevertheless, numerically we get the same contribution. Similar to the general case, we diagonalize the Fock matrix to obtain the orbital energies as well as the orbital wave functions. In the last step of one SCF iteration (s. yellow boxes in Fig. 1) we calculate the Hartree-Fock ground state energy and afterward for convergence is checked (s. blue box in Fig. 1). As mentioned in the previous section, there are several ways to do so. While the comparison of the orbital energies ϵ_p between two iterations is the computationally most efficient way, we choose to look at the Hartree-Fock ground state energy E^{HF}. Hence, we additionally calculate E^{HF} for each iteration according to Eq. 3.16 which is computationally more expensive, but due to the size of our system irrelevant. Convergence is obtained if the condition

$$|{}^{(N)}E^{HF} - {}^{(N-1)}E^{HF}| < \epsilon, \tag{3.17}$$

between iteration N and $N-1$ is fulfilled. Often not only the change in energy ΔE^{HF} but also the change of the reduced density matrix $\Delta\rho^{red}$ is considered. Since comparing two matrixes is more complex than comparing two scalars like the energy, the threshold should be adapted accordingly. We choose the maximum norm to compare the two matrices

$$||{}^{(N)}\rho_{ij}^{red} - {}^{(N-1)}\rho_{ij}^{red}||_{max} < \epsilon. \tag{3.18}$$

In our implementation we choose an accuracy of $\epsilon = 10^{-6}$ for both criteria, which is especially for the comparison of the density matrices a rather strict condition. Finally, after the SCF algorithm is converged, different expectation values can be calculated (s. green box in Fig. 1).

To summarize, we derive the canonical Hartree-Fock equations by applying the variational theorem on the respective Lagrangian functional. The exact wave function is approximated by a single Slater determinant which in turn we use to minimize the ground state energy. By introducing a basis, we are able to convert the Hartree-Fock equation into a set of algebraic equations, solvable in terms of matrix techniques. Finally, we specify the types of orbitals introducing the restricted closed-shell formalism.

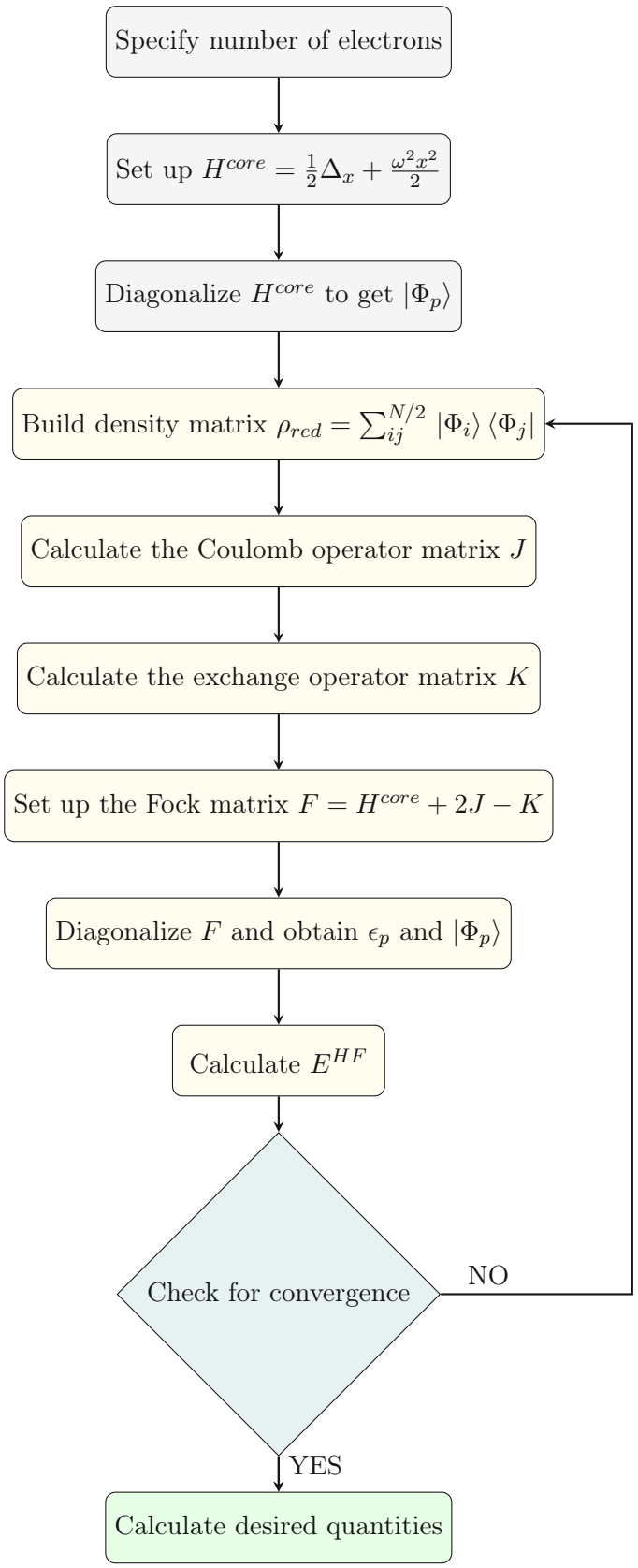

Figure 1: Flowchart of the SCF algorithm.

4 Post-Hartree-Fock Methods

We introduced the Hartree-Fock method based on the idea of approximating the electronic wave function by a single Slater determinant. Thereby the electron correlation is treated in a mean-field manner which yields inaccurate energies. Nevertheless, Hartree-Fock wave functions partly provide qualitative correct information about trends of important quantities such as the energy with respect to certain system parameters. Nowadays, post-Hartree-Fock methods are a standard and more accurate tool in computational chemistry to describe a given system. The terminology post-Hartree-Fock covers a number of different methods sharing the fact for being based on the results of the SCF algorithm. Møller Plesset, coupled cluster and equation of motion coupled cluster theory are among the best-known representatives. All methods have in common that, in contrast to Hartree-Fock, they no longer treat the electron-electron effects as an averaged field but introduce actual electron correlations, which represents a more accurate way to describe the repulsion between electrons. Thus, post-Hartree-Fock methods typically give more accurate results. However, the gained accuracy comes with a significantly increased computational cost.

This chapter is divided into three parts. In the first part, we introduce Møller-Plesset perturbation theory, which is justified by the fact that it is to some extent contained in the theory of coupled clusters. At the same time, the Coulomb integrals, which are constructed using the orbital wave functions from the Hartree-Fock method, are incorporated. In the second part, coupled cluster theory is discussed, where we pay special attention to the formalism of singles and doubles. Hereby, we use the formalism of the second quantization, introduced in chapter 2, to derive the set of coupled cluster equations for the implementation. In the last part, we extend the coupled cluster formalism in order to be able to describe also excited states. We briefly review this so-called equation of motion coupled cluster theory to go into detail about the singles and doubles formalism. Also here, the equations which can be readily implemented into code are provided.

4.1 Møller-Plesset Perturbation Theory

The first natural step beyond Hartree-Fock is to employ perturbation theory in order to account for electron correlations. In general, many-body perturbation theory is known as Rayleigh-Schrödinger perturbation theory. The idea is to split the exact Hamiltonian $\hat{H}$ in an unperturbed $\hat{H}_0$ and perturbation $\hat{V}$ part. The concept of perturbation theory can be applied if the latter contribution $\hat{V}$ is small compared to $\hat{H}_0$. In the case of $\hat{H}_0$ being the Hartree-Fock Hamiltonian, the perturbation theory is called Møller-Plesset perturbation theory (MPPT).

4.1.1 Theoretical Basics

We start with the general formulation of the Rayleigh-Schrödinger perturbation theory by constructing the exact Hamiltonian operator as follows:

$$\hat{H} = \hat{H}_0 + \hat{V}.$$

The goal is to solve the time-independent Schrödinger equation for this Hamiltonian

$$\hat{H} \left|\Phi_i\right\rangle = \epsilon_i \left|\Phi_i\right\rangle, \tag{4.1}$$

which, however, becomes an unsolvable task for systems with more than two electrons. Therefore, one exploits the fact that the eigenfunctions and eigenvalues of the unperturbed Hamiltonian $\hat{H}_0$ are known:

$$\hat{H}_0 \left|\phi_i^{(0)}\right\rangle = E_i^{(0)} \left|\phi_i^{(0)}\right\rangle ,$$

where the superscript (n) refers to the n-th order of perturbation, i.e. (0) denotes 0-th order perturbation or the unperturbed quantity, respectively. Assuming $\hat{V}$ is small, we expect $|\phi_i^{(0)}\rangle$ to be close to $|\Phi_i\rangle$ as well as $E_i^{(0)}$ to ϵ_i. This small deviation is parameterized as follows:

$$\hat{H} = \hat{H}_0 + \lambda \hat{V},$$

where $0 \leq \lambda \leq 1$ is the order parameter which will later be set equal to one. Now we can expand the exact eigenfunctions and eigenvalues in a Taylor expansion:

$$\begin{aligned} \epsilon_i &= E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots , \\ |\Phi_i\rangle &= |\phi_i^{(0)}\rangle + \lambda \left|\phi_i^{(1)}\right\rangle + \lambda^2 \left|\phi_i^{(2)}\right\rangle + \dots . \end{aligned} \tag{4.2}$$

We take the wave functions of $\hat{H}_0$ to be normalized, i.e. $\langle\phi_i^{(0)}|\phi_i^{(0)}\rangle = 1$ and choose the so-called intermediate normalization for the exact wave functions $\langle\phi_i^{(0)}|\Phi_i\rangle = 1$ assuming that they are not orthogonal. By comparing coefficients, we obtain:

$$\langle\phi_i^{(0)}|\phi_i^{(n)}\rangle = 0 \qquad n = 1, 2, 3, \dots .$$

We insert the Taylor expanded expressions for the energy and the wave function (s. Eq. 4.2) into the exact Schrödinger equation from Eq. 4.1 and are able to equate coefficients of λ^n to find a recursive system of equations. For λ^0 we get as expected just the unperturbed eigenvalue equation. The remaining expression for the energy in the n-th order results in:

$$E_i^{(n)} = \langle\phi_i^{(0)}|\hat{V}|\phi_i^{(n)}\rangle \qquad n \geq 1.$$

Finally, we want to determine the equations for the wave functions $|\phi_i^{(n)}\rangle$ in order to get an explicit energy expression depending on known quantities only. By doing so, we see that we have to deal with integro-differential equations. One ansatz is to expand the wave functions $|\phi_i^{(n)}\rangle$ with the complete set of the eigenfunctions of $\hat{H}_0$ [24]

$$\left|\phi_i^{(n)}\right\rangle = \sum_j c_{ij}^{(n)} \phi_j^{(0)},$$

where the coefficients are defined as $c_{ij}^{(n)} = \langle\phi_j^{(0)}|\phi_i^{(n)}\rangle$. The first few contributions to the energy are given by

$$
\begin{aligned}
E_i^{(1)} &= \langle \phi_i^{(0)} | \hat{V} | \phi_i^{(0)} \rangle \,, \\
E_i^{(2)} &= \sum_{j \neq i} \frac{\langle \phi_i^{(0)} | \hat{V} | \phi_j^{(0)} \rangle \, \langle \phi_j^{(0)} | \hat{V} | \phi_i^{(0)} \rangle}{E_i^{(0)} - E_j^{(0)}} = \sum_{j \neq i} \frac{\left| \langle \phi_i^{(0)} | \hat{V} | \phi_j^{(0)} \rangle \right|^2}{E_i^{(0)} - E_j^{(0)}}, \\
E_i^{(3)} &= \sum_{j,k \neq i} \frac{\langle \phi_i^{(0)} | \hat{V} | \phi_j^{(0)} \rangle \, \langle \phi_j^{(0)} | \hat{V} | \phi_k^{(0)} \rangle \, \langle \phi_k^{(0)} | \hat{V} | \phi_i^{(0)} \rangle}{(E_i^{(0)} - E_j^{(0)})(E_i^{(0)} - E_k^{(0)})} - E_i^{(1)} \sum_{j \neq i} \frac{\left| \langle \phi_i^{(0)} | \hat{V} | \phi_j^{(0)} \rangle \right|^2}{(E_i^{(0)} - E_j^{(0)})^2}.
\end{aligned}
\tag{4.3}
$$

Now we want to apply the general Rayleigh-Schrödinger perturbation theory to the Hartree-Fock model yielding the MPPT. Starting point is the Hartree-Fock Hamiltonian from Eq. 3.13:

$$\hat{H}_0 = \sum_i \hat{f}(i) = \sum_i \hat{h}(i) + v^{HF}(i).$$

It is evident that the Hartree-Fock Hamiltonian, being an approximation to the exact electronic Hamiltonian (s. Eq. 2.4), represents the unperturbed term $\hat{H}_0$ with the known solutions obtained from the SCF method. At the same time, the perturbation $\hat{V}$ reads

$$\hat{V} = \hat{H} - \hat{H}_0 = \sum_{i<j} \frac{1}{\vec{r}_{ij}} - \sum_i v^{HF}(i),$$

where $\sum_{i<j} \vec{r}_{ij}^{-1}$ contains the full two-electron interaction and $v^{HF}(i) = \sum_j \hat{J}_j - \hat{K}_j$ is the familiar Hartree-Fock potential for the i-th electron. Here, we only study corrections to the ground state energy. The zeroth-order energy correction is, again, just the expectation value of the unperturbed, i.e. the Hartree-Fock Hamiltonian with the Hartree-Fock determinant $|\Psi^{HF}\rangle$

$$E_0^{(0)} = \langle \Psi^{HF} | \hat{H}_0 | \Psi^{HF} \rangle = \sum_i \epsilon_i = \sum_i \langle i | \hat{h} | i \rangle + \sum_{ij} \langle ij || ij \rangle \,.$$

Consequently, the first-order correction is:

$$E_0^{(1)} = \langle \Psi^{HF} | \hat{V} | \Psi^{HF} \rangle = \sum_{i<j} \langle \Psi^{HF} | \frac{1}{\vec{r}_{ij}} | \Psi^{HF} \rangle - \sum_i \langle \Psi^{HF} | v^{HF}(i) | \Psi^{HF} \rangle \,.$$

To get a feasible expression for $E_0^{(1)}$, we again resort to the Slater-Condon rules (s. Eq. 2.20)

$$E_0^{(1)} = \frac{1}{2} \sum_{ij} \langle ij || ij \rangle - \sum_{ij} \langle ij || ij \rangle = -\frac{1}{2} \sum_{ij} \langle ij || ij \rangle \,.$$

At this point, we shall look at the energy obtained with first-order perturbation theory. Here, the expansion is stopped after the second term of the perturbation series which results in

$$E_0 \approx E_0^{(0)} + E_0^{(1)} = \sum_i \langle i|\hat{h}|i\rangle + \frac{1}{2}\sum_{ij}\langle ij||ij\rangle = E_0^{HF}.$$

As a result, the first two energy contributions in the MPPT give exactly the Hartree-Fock energy defined in Eq. 3.2. This means at the same time, that in order to improve the accuracy of our electronic structure calculations beyond Hartree-Fock, we need to include higher-order corrections of the MPPT. These higher-order terms are calculated using an expansion in excited states of the zeroth-order Hamiltonian. Due to this, and because $|\phi_i^{(0)}\rangle = |\Psi^{HF}\rangle$, most of the expectations values arising in the energy corrections of MPPT are zero. On one hand expectation values with triple and higher excited determinants can be neglected due to the two-electron nature of the many-body Hamiltonian, and on the other hand, expectation values involving the Hartree-Fock ground state determinant and single excited vanish due to the Brillouin theorem (s. Sec. 3.2). Therefore, the remaining quantities to calculate are the doubly excited determinants $|\Psi_{ij}^{ab}\rangle$. We further exploit the fact that double excitations $|\Psi_{ij}^{ab}\rangle$ are eigenstates of the Hartree-Fock Hamiltonian

$$\hat{H}^{HF}\,|\Psi_{ij}^{ab}\rangle = (E_0^{(0)} - \epsilon_i - \epsilon_j + \epsilon_a + \epsilon_b)\,|\Psi_{ij}^{ab}\rangle\,.$$

According to Eqs. 4.3 the second-order perturbation energy results in:

$$E_0^{(2)} = \sum_{i<j}^{N}\sum_{a<b}\frac{|\langle ij||ab\rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}, \tag{4.4}$$

where both sums are restricted in order to avoid double counting. As introduced in chapter 2, the indices $(i, j, \dots)$ denote the occupied spin orbitals that's why in the first sum it is summed over N electrons. The second sum with the virtual spin orbitals $(a, b, \dots)$ is unrestricted.

The energy correction to the Hartree-Fock ground state energy is the so-called correlation energy. In the formalism of MPPT we introduced the second order correction $E_0^{(2)}$, also referred to as MP2 correlation energy. However, it has become common practice that the MP2 energy denotes the total energy, i.e. $E^{MP2} = E^{HF} + E_0^{(2)} = E^{HF} + E_{cor}^{MP2}$, introducing E_{cor}^{MP2} as the correlation energy itself.

4.1.2 Implementation

Before going into the details of the coding, we shall investigate one crucial ingredient of the post-Hartree-Fock methods and even more of this thesis. Until now, we have not dealt with the calculation and implementation of two-electron integrals. They were formally introduced in chapter 2 and then also in the Hartree-Fock method in chapter 3 at the definition of the Coulomb and exchange operator. In the implementation of the SCF code, we escaped from the direct evaluation of these integrals by solving matrix equations. Anyway, now we have to analyze the implementation of the two-electron integrals since we need them as an important building block for the present and following sections.

Two-Electron Integrals

We begin with the general definition of the two-electron integrals according to Eq. 2.16

$$(pq|rs) = \int dx_1 \int dx_2 \; \Phi_p^*(x_1)\Phi_q^*(x_2)\frac{1}{|x_1 - x_2| + \sigma}\Phi_r(x_1)\Phi_s(x_2), \tag{4.5}$$

where from now on, we work with the two-electron potential from our model (s. Eq. 2.6) as well as with the generic indices $(p, q, \dots)$. Further, we want to mention that in the context of electronic structure theory, this type of two-electron integrals are also referred to as Coulomb integrals.
Note, that from now on the integration is over the one-dimensional x-axis, not be confused with the four-dimensional spin orbital coordinate $\vec{x} = \{\vec{r}, \bar{\omega}\}$. In addition, we immediately work with spatial orbitals. The reason for this is that we are concerned with the restricted closed-shell Hartree-Fock method, where the spatial wave functions for up- and down-spin electrons in the same shell are equal. Therefore, the main task is to evaluate these spatial two-electron integrals and then account for both electrons. For the integration itself, we make use of the trapezoidal integration rule in one dimension which states

$$\int_{x_0}^{x_f} f(x)\, dx \approx \sum_{k=1}^{N} \frac{f(x_{k-1}) + f(x_k)}{2}\Delta x_k,$$

where $f(x)$ is an arbitrary function defined on $[x_0, x_f]$ and Δx_k represents the spacing between the points $f(x_{k-1})$ and $f(x_k)$. In our case we choose an equidistant grid ($\Delta x_k = \Delta x$) and the above expression simplifies to:

$$\int_a^b f(x)\, dx \approx \frac{\Delta x}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{N-1}) + f(x_N)\right].$$

It is straightforward to apply this rule in two dimensions

$$\begin{aligned}\int_{x_0}^{x_f}\int_{y_0}^{y_f} f(x,y)\, dx\, dy \approx \frac{\Delta x \Delta y}{4}\Bigg[&f(x_0,y_0) + f(x_0,y_f) + f(x_f,y_0) + f(x_f,y_f)\\ &+ 2\cdot\sum_{i=2}^{N_x} f(x_0,y_i) + f(x_f,y_i) + 2\cdot\sum_{j=2}^{N_x} f(x_j,y_0) + f(x_j,y_f)\\ &+ 4\cdot\sum_{i=2}^{N_x}\sum_{j=2}^{N_y} f(x_j,x_i)\Bigg].\end{aligned}$$

We can further simplify this expression by assuming an equal number of grid points and equal spacing

$$\begin{aligned}\int_{x_0}^{x_f}\int_{y_0}^{y_f} f(x,y)\, dx\, dy \approx \frac{\Delta^2}{4}\Bigg[&f(x_0,y_0) + f(x_0,y_f) + f(x_f,y_0) + f(x_f,y_f)\\ &+ 2\cdot\sum_{i=2}^{N} f(x_0,y_i) + f(x_f,y_i) + 2\cdot\sum_{j=2}^{N} f(x_j,y_0) + f(x_j,y_f)\\ &+ 4\cdot\sum_{i,j=2}^{N} f(x_j,x_i)\Bigg],\end{aligned} \tag{4.6}$$

with $\Delta x = \Delta y = \Delta$ and $N_x = N_y = N$.
At this point we are able to evaluate the two-electron integral for a given set of parameters $\{p, q, r, s\}$ by identifying

$$f(x_1, x_2) \mathrel{\widehat{=}} \Phi_p^*(x_1)\Phi_q^*(x_2)\frac{1}{|x_1 - x_2| + \sigma}\Phi_r(x_1)\Phi_s(x_2).$$

In the following, we define the four-dimensional Coulomb tensor V_{rs}^{pq}, which contains the two-electron integrals for all possible combinations of four arbitrary generic indices:

$$V_{rs}^{pq} \equiv \langle pq|rs\rangle \qquad \forall\, p, q, r, s \in N_{SO}, \tag{4.7}$$

where N_{SO} is the number of spin orbitals.
In fact, we need all possible integrals from Eq. 4.5 to be able to evaluate expressions like $E_0^{(2)}$ in MPPT (s. Eq. 4.4) with a sum over several indices efficiently. To also consider the spin dependency, we calculate the spatial two-electron integral twice, according to the two possible spin states. We further must consider that only p and r as well as q and s orbitals with the same spin contribute. The calculation of V_{rs}^{pq} is computationally expensive which is why only a small number of spin orbitals is investigated. The introduction of the four-dimensional tensor V_{rs}^{pq} makes it easier for us to deal with the coupled cluster algorithm, which requires such types of objects as input. To be precise, the post-Hartree-Fock methods require the antisymmetrized four-dimensional tensor:

$$V^{-pq}_{rs} = \langle pq|rs\rangle - \langle pq|sr\rangle = \langle pq||rs\rangle \qquad \forall\, p, q, r, s \in N_{SO}, \tag{4.8}$$

where the minus in the superscript indicates the antisymmetrization. While the antisymmetrized integrals $\langle pq||rs\rangle$ are used for the derivation of the coupled cluster equations (s. Sec. 4.2.2), the analysis of the two-electron, respectively Coulomb integrals $\langle pq|rs\rangle$ itself is presented in chapter 5.

MP2 Energy

Now we are at a point where we may use the antisymmetrized four-dimensional tensor to evaluate the second-order energy expression of the MPPT. In our code, we use a slightly different formula for the MP2 energy compared to the derived Eq. 4.4:

$$E_0^{(2)} = \frac{1}{4}\sum_{i,j=1}^{N}\sum_{a,b=N+1}^{N_{SO}}\frac{\langle ij||ab\rangle\,\langle ab||ij\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}, \tag{4.9}$$

where for the implementation the second sum runs from $N+1$ to the remaining number of unoccupied spin orbitals. Eq. 4.9 follows directly from the fact that we do not constrain the sums (see Eq. 4.4). Accordingly, we need to compensate for this with a factor of 1/4. Further, we are able to express the MP2 energy in terms of V^{-pq}_{rs}:

$$E_0^{(2)} = \frac{1}{4} V^{-ij}_{\ ab}\, V^{-ab}_{\ ij}\, (\mathcal{E}^{ij}_{ab})^{-1}, \tag{4.10}$$

where we define $(\mathcal{E}^{ij}_{ab})^{-1}$ as the orbital energy denominator from Eq. 4.9. In addition, this notation allows us to readily evaluate the MP2 expression by means of tensor contraction. We elaborate more on tensor contraction in chapter 5.

In the end, the MP2 code consists of four loops representing the sum over a set of indices $\{p, q, r, s\}$, where for all values the antisymmetrized tensor $V^{-pq}_{\ rs}$ as well as the orbital energies ϵ_i, are evaluated according to Eq. 4.9. We want to mention that in our case, the orbital energies are given by the diagonal elements of the Fock matrix, i.e. $\epsilon_i = f_{ii}$. A more elegant way to evaluate Eq. 4.9 is the mentioned tensor contraction, where the *NumPy* package *einsum* may be used.

4.2 Coupled Cluster Theory

Similar to MPPT, also coupled cluster (CC) theory is introduced as an extension to the Hartree-Fock method in a perturbative regime. The concept of coupled cluster theory relies on an exponential ansatz with the cluster operator $\hat{T}$ for the many-electron wave function. However, since in the exponential ansatz the numerical cost increases dramatically with each order, the cluster operator must be truncated. For us, the contributions up to the second order in $\hat{T}$, the so-called coupled cluster singles and doubles (CCSD), are of particular interest.
In the upcoming sections, we first introduce the coupled cluster method conceptually. Building on this, we derive the coupled cluster equations for the special case of CCSD. Our derivations follow closely those of Ref. [30].

4.2.1 Theoretical Basics

We consider the time-independent non-relativistic Schrödinger equation

$$\hat{H} \left|\Psi\right\rangle = E \left|\Psi\right\rangle .$$

The coupled cluster wave function $\left|\Psi^{CC}\right\rangle$ defined as:

$$\left|\Psi^{CC}\right\rangle \equiv e^{\hat{T}} \left|\Psi^{HF}\right\rangle , \tag{4.11}$$

shall approximate the exact solution $\left|\Psi\right\rangle$. We again take the reference state $\left|\Psi^{HF}\right\rangle$ as the Hartree-Fock ground state wave function. The cluster operator can formally be written as

$$\hat{T} \equiv \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots . \tag{4.12}$$

By specifying the order of excitation, $\hat{T}$ can also be expressed via second quantization

$$\hat{T}_n = \left(\frac{1}{n!}\right)^2 \sum_{ij\ldots ab\ldots} t_{ij\ldots}^{ab\ldots} \hat{a}_a^\dagger \hat{a}_b^\dagger \ldots \hat{a}_j \hat{a}_i, \tag{4.13}$$

where $t_{ij\ldots}^{ab\ldots}$ are the so-called cluster amplitudes and n refers to the number of excitations. By the structure of the cluster operator in Eq. 4.13 we see that when acting on a reference wave function by the creation and annihilation operators excited states arise, which can be expressed by excited Slater determinants. For illustration we consider the action of $\hat{T}_1$ and $\hat{T}_2$ on the reference state:

$$\begin{aligned}
\hat{T}_1 \left|\Psi^{HF}\right\rangle &= \sum_{ia} t_i^a \hat{a}_a^\dagger \hat{a}_i \left|\Psi^{HF}\right\rangle = \sum_{ia} t_i^a \left|\Psi_i^a\right\rangle, \\
\hat{T}_2 \left|\Psi^{HF}\right\rangle &= \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i \left|\Psi^{HF}\right\rangle = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \left|\Psi_{ij}^{ab}\right\rangle.
\end{aligned}$$

To truncate the exponential coupled cluster ansatz, we investigate the exponential operator itself:

$$e^{\hat{T}} = \sum_{k=0}^{n} \frac{1}{k!} \hat{T}^k = 1 + \hat{T} + \frac{1}{2}\hat{T}^2 + \ldots, \tag{4.14}$$

with the truncation up to the n-th order. Exploiting the definition of the cluster operator from Eq. 4.12, we find

$$\begin{aligned}
e^{\hat{T}} &= 1 + (\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots) + \frac{1}{2}(\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots)^2 + \ldots \\
&= 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2}\hat{T}_1^2) + (\hat{T}_3 + \frac{1}{3!}\hat{T}_1 + \hat{T}_2\hat{T}_1) + \ldots .
\end{aligned}$$

In the last line the sequence of the cluster operators is sorted by the order of excitation. The exponential ansatz of the CCSD theory reads

$$e^{\hat{T}} \approx e^{\hat{T}_1 + \hat{T}_2} = 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2}\hat{T}_1^2) + \ldots . \tag{4.15}$$

Returning to the general coupled cluster theory, the main task remains to determine the energy and the cluster amplitudes. Projecting $\langle\Psi^{HF}|$ on the approximated Schrödinger equation constructed with the coupled cluster wave functions yields

$$\langle\Psi^{HF}|\hat{H}e^{\hat{T}}|\Psi^{HF}\rangle = E\,\langle\Psi^{HF}|e^{\hat{T}}|\Psi^{HF}\rangle = E, \tag{4.16}$$

where again intermediate normalization $\langle \Psi^{HF} | \Psi^{CC} \rangle = 1$ is assumed. Similarly, we obtain the amplitude equations by left projecting the excited determinants

$$\langle \Psi_{ij\ldots}^{ab\ldots} | \hat{H} e^{\hat{T}} | \Psi^{HF} \rangle = E \, \langle \Psi_{ij\ldots}^{ab\ldots} | e^{\hat{T}} | \Psi^{HF} \rangle \, . \tag{4.17}$$

We refer to Eqs. 4.16 and 4.17 as the coupled cluster equations, where the former is also called the CC energy equation and the latter the CC amplitude equations.
We further want to highlight, that the structure of $\hat{H}$ implies a natural truncation rather than the artificial one like in Eq. 4.15. Assuming the Hartree-Fock single Slater determinant reference wave function, the Slater-Condon rules (s. Eq. 2.5) state that matrix elements of the Hamiltonian between determinants that differ by more than two spin orbitals are zero. As a result, the energy and amplitude equations simplify in a natural manner, which is only due to the form of the Hamiltonian and depends neither on $\hat{T}$ nor on the number of electrons.

To obtain practical equations for the numerical implementation of Eq. 4.16 and Eq. 4.17 additional manipulations are required. To do this, all terms must be expressed in the form of one-electron and two-electron integrals derived from the Hamiltonian and cluster amplitudes.

4.2.2 Coupled Cluster Singles and Doubles

The formal starting point of the CCSD formalism is given by the truncation from Eq. 4.15 assuming $\hat{T} \approx \hat{T}_1 + \hat{T}_2$. Before we address the consequences of this truncation and the resulting CCSD equations, we return to the results of Sec. 2.6, where the formalism of the second quantization was introduced. We continue with the normal-ordered electronic Hamiltonian in the particle-hole formalism from Eq. 2.25. We refer to the first expression as the normal-ordered two-electron potential $\hat{V}_N$ and identify the second and third term as the normal-ordered Fock operator $\hat{f}_N$ as well as the last two terms as the Hartree-Fock energy. With this, we may write the electronic Hamiltonian as follows:

$$\begin{aligned} \hat{H} &= \hat{f}_N + \hat{V}_N + \langle \Psi^{HF} | \hat{H} | \Psi^{HF} \rangle \, , \\ \hat{H}_N \equiv \hat{f}_N + \hat{V}_N &= \hat{H} - \langle \Psi^{HF} | \hat{H} | \Psi^{HF} \rangle \, , \end{aligned} \tag{4.18}$$

where the subscript N stands for the normal-ordering.
We simultaneously introduce the normal-ordered Hamiltonian $\hat{H}_N$, which can be understood as a correlation operator since the Hartree-Fock energy is subtracted. Furthermore, we introduce the similarity-transformed Hamiltonian $\bar{H}$ via projection of $e^{-\hat{T}}$ on the left of the Schrödinger equation (s. Eq. 4.16)

$$\langle \Psi^{HF} | e^{-\hat{T}} \hat{H}_N e^{\hat{T}} | \Psi^{HF} \rangle = \langle \Psi^{HF} | \bar{H} | \Psi^{HF} \rangle = \underbrace{\langle \Psi^{HF} | e^{-\hat{T}} e^{\hat{T}} | \Psi^{HF} \rangle}_{=1} E. \tag{4.19}$$

Note that in contrast to $\hat{H}_N$ the similarity-transformed Hamiltonian $\bar{H}$ is not Hermitian. We emphasize that now the amplitude equations are decoupled from the energy equation, which is very convenient. Exploiting the Baker-Campbell-Hausdorff formula yields:

$$
\begin{aligned}
\bar{H} = \hat{H}_N &+ [\hat{H}_N, \hat{T}_1] + [\hat{H}_N, \hat{T}_2] \\
&+ \frac{1}{2}\Big([[\hat{H}_N, \hat{T}_1], \hat{T}_1] + [[\hat{H}_N, \hat{T}_2], \hat{T}_2] + [[\hat{H}_N, \hat{T}_1], \hat{T}_2] + [[\hat{H}_N, \hat{T}_2], \hat{T}_1]\Big) \\
&+ \frac{1}{3!}\Big([[[\hat{H}_N, \hat{T}_1], \hat{T}_1], \hat{T}_1] + [[[\hat{H}_N, \hat{T}_1], \hat{T}_1], \hat{T}_2] + [[[\hat{H}_N, \hat{T}_1], \hat{T}_2], \hat{T}_1] + [[[\hat{H}_N, \hat{T}_2], \hat{T}_1], \hat{T}_1] \\
&+ [[[\hat{H}_N, \hat{T}_1], \hat{T}_2], \hat{T}_2] + [[[\hat{H}_N, \hat{T}_2], \hat{T}_1], \hat{T}_2] + [[[\hat{H}_N, \hat{T}_2], \hat{T}_2], \hat{T}_1] + [[[\hat{H}_N, \hat{T}_2], \hat{T}_2], \hat{T}_2]\Big) \\
&+ \dots,
\end{aligned}
\tag{4.20}
$$

where the expansion naturally terminates at quadruply nested commutators.
We now are left evaluating the expectation values from the coupled cluster equations according to Eq. 4.16 and Eq. 4.17 with the similarity-transformed Hamiltonian $\bar{H}$. In doing so, we look at a couple of examples, however, due to the tedious formalism we do not derive all contributions stemming from Eq. 4.20. For a more detailed documentation we refer to Ref. [30]. In passing, we want to mention that for the sake of compactness in the remaining chapters the summation over the virtual and occupied orbitals indices $(a, b, c, \dots)$ and $(i, j, k, \dots)$, respectively, is not specified. However, our shorthand notation implies

$$
\sum_{i,j,k,\dots} \widehat{=} \sum_{i,j,k,\dots=1}^{N} \quad ; \qquad \sum_{a,b,c,\dots} \widehat{=} \sum_{a,b,c,\dots=N+1}^{N_{SO}} \quad ,
$$

where N is the number of electrons and N_{SO} the number of spin orbitals.

We begin our derivations by considering the commutator $[\hat{f}_N, \hat{T}_1]$ arising from Eq. 4.20. In second quantization, the product of the normal-ordered Fock operator $\hat{f}_N$ and the cluster operator $\hat{T}_1$ yields:

$$
\begin{aligned}
\hat{f}_N \hat{T}_1 &= \sum_{pqia} f_{pq} t_i^a \{\hat{a}_p^\dagger \hat{a}_q\}\{\hat{a}_a^\dagger \hat{a}_i\} \\
&= \sum_{pqia} f_{pq} t_i^a \Big(\{\hat{a}_p^\dagger \hat{a}_q \hat{a}_a^\dagger \hat{a}_i\} + \{\hat{a}_p^\dagger \hat{a}_q \hat{a}_a^\dagger \hat{a}_i\} + \{\hat{a}_p^\dagger \hat{a}_q \hat{a}_a^\dagger \hat{a}_i\} + \{\hat{a}_p^\dagger \hat{a}_q \hat{a}_a^\dagger \hat{a}_i\}\Big) \\
&= \sum_{pqia} f_{pq} t_i^a \Big(\{\hat{a}_p^\dagger \hat{a}_q \hat{a}_a^\dagger \hat{a}_i\} + \delta_{pi}\{\hat{a}_q \hat{a}_a^\dagger\} + \delta_{qa}\{\hat{a}_p^\dagger \hat{a}_i\} + \delta_{pi}\delta_{qa}\Big).
\end{aligned}
$$

Here, f_{pq} denotes the Fock matrix element. We further use normal-ordering as well as the generalized Wick theorem, where the contractions between the two strings of operators are considered. Consequently, the full commutator is given by:

$$
\begin{aligned}
[\hat{f}_N, \hat{T}_1] &= \hat{f}_N \hat{T}_1 - \hat{T}_1 \hat{f}_N \\
&= \sum_{pqia} f_{pq} t_i^a \Big(\{\hat{a}_p^\dagger \hat{a}_q \hat{a}_a^\dagger \hat{a}_i\} + \delta_{pi}\{\hat{a}_q \hat{a}_a^\dagger\} + \delta_{qa}\{\hat{a}_p^\dagger \hat{a}_i\} + \delta_{pi}\delta_{qa} - \{\hat{a}_a^\dagger \hat{a}_i \hat{a}_p^\dagger \hat{a}_q\}\Big) \\
&= \sum_{qia} f_{iq} t_i^a \{\hat{a}_q \hat{a}_a^\dagger\} + \sum_{pia} f_{pa} t_i^a \{\hat{a}_p^\dagger \hat{a}_i\} + \sum_{ia} f_{ia},
\end{aligned}
\tag{4.21}
$$

where the contribution $\{a_a^\dagger a_i a_p^\dagger a_q\}$ of $\hat{T}_1\hat{f}_N$ cancels out. Wick's theorem for the first doubly nested commutator in Eq. 4.20 reads

$$\begin{aligned}
\frac{1}{2}\hat{f}_N\hat{T}_1^2 = \frac{1}{2}\sum_{pqiajb} f_{pq}t_i^a t_j^b \Big(&\{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} \\
&+ \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} \\
&+ \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\}\Big) \\
= \frac{1}{2}\sum_{pqiajb} t_i^a t_j^b \Bigg(&\sum_{pq} f_{pq}\{\hat{a}_p^\dagger\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + 2\sum_{q} f_{iq}\{\hat{a}_q\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\} \\
&+ 2\sum_{p} f_{pb}\{\hat{a}_p^\dagger\hat{a}_a^\dagger\hat{a}_i\hat{a}_j\} + 2f_{ib}\{\hat{a}_a^\dagger\hat{a}_i\} + 2f_{ja}\{\hat{a}_i\hat{a}_b^\dagger\}\Bigg).
\end{aligned}$$

For the last equality, we use the fact that several terms can be summarized by renaming the indices. The main difference for the two remaining expressions of the commutator is that the sequence of the creation and annihilation operators is exchanged according to the ordering of $\hat{f}_N$ and $\hat{T}_1$. For brevity, we just give the final results:

$$\begin{aligned}
\hat{T}_1\hat{f}_N\hat{T}_1 &= \sum_{aibj} t_i^a t_j^b \left(\sum_{pq} f_{pq}\{\hat{a}_a^\dagger\hat{a}_i\hat{a}_p^\dagger\hat{a}_q\hat{a}_b^\dagger\hat{a}_j\} + \sum_{q} f_{jq}\{\hat{a}_a^\dagger\hat{a}_i\hat{a}_q\hat{a}_b^\dagger\} + \sum_{p} f_{pb}\{\hat{a}_a^\dagger\hat{a}_i\hat{a}_p^\dagger\hat{a}_j\} + f_{ib}\{\hat{a}_a^\dagger\hat{a}_i\}\right), \\
\frac{1}{2}\hat{T}_1^2\hat{f}_N &= \frac{1}{2}\sum_{aibjpq} f_{pq}t_i^a t_j^b\{\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\hat{a}_p^\dagger\hat{a}_q\},
\end{aligned}$$

which leads to

$$\frac{1}{2}[[\hat{f}_N,\hat{T}_1],\hat{T}_1] = \frac{1}{2}\hat{f}_N\hat{T}_1^2 - \hat{T}_1\hat{f}_N\hat{T}_1 + \frac{1}{2}\hat{T}_1^2\hat{f}_N = \sum_{aibj} f_{ja}t_i^a t_j^b\{\hat{a}_i\hat{a}_b^\dagger\}.$$

In this case, only one term of the whole commutator survives. However, without proof, we are able to make an important generalization. The application of Wick's theorem to the commutators of Eq. 4.20 entails that only expressions in which the Hamiltonian $\hat{H}_N$ has at least one contraction with each cluster operator $\hat{T}_N$ on its right-hand side are non-zero [30]. In other words, the Hamiltonian must share at least one index with every cluster operator in the final result. Consequently, we are able to simplify Eq. 4.20 such that only the so-called connected terms are displayed

$$\begin{aligned}\bar{H} = \Big(& \hat{H}_N + \hat{H}_N\hat{T}_1 + \hat{H}_N\hat{T}_2 + \frac{1}{2}\hat{H}_N\hat{T}_1^2 + \frac{1}{2}\hat{H}_N\hat{T}_2^2 + \hat{H}_N\hat{T}_1\hat{T}_2 \\ & + \frac{1}{6}\hat{H}_N\hat{T}_1^3 + \frac{1}{2}\hat{H}_N\hat{T}_1^2\hat{T}_2 + \frac{1}{2}\hat{H}_N\hat{T}_1\hat{T}_2^2 + \frac{1}{6}\hat{H}_N\hat{T}_2^3 + \frac{1}{24}\hat{H}_N\hat{T}_1^4 \\ & + \frac{1}{6}\hat{H}_N\hat{T}_1^3\hat{T}_2 + \frac{1}{6}\hat{H}_N\hat{T}_1^2\hat{T}_2^2 + \frac{1}{6}\hat{H}_N\hat{T}_1\hat{T}_2^3 + \frac{1}{24}\hat{H}_N\hat{T}_2^4 \Big)_c ,\end{aligned} \tag{4.22}$$

where the subscript c stands for the connected terms. Eq. 4.22 is also referred to as the connected cluster form of the similarity-transformed Hamiltonian [31], which we will use for the rest of this thesis. Further, the natural truncation of the Baker-Campbell-Hausdorff formula is verified since the Hamiltonian contains at most four creation and annihilation operators which in turn means that only four cluster operators can be connected simultaneously.
Now we are at a point where we shall derive the CCSD equation for the energy. The energy results from the expectation value of the similarity-transformed Hamiltonian

$$E_{cor}^{CCSD} = \langle \Psi^{HF} | \bar{H} | \Psi^{HF} \rangle - E^{HF}.$$

In passing we mention that by construction of the normal-ordered Hamiltonian, the energy produced by the CCSD method represents a correlation energy. By convention, however, the total CCSD energy is defined as $E^{CCSD} = E^{HF} + E_{cor}^{CCSD}$ giving the ground state energy.
The leading term is just the expectation value of the normal-ordered Hamiltonian, for which by definition $\langle \Psi^{HF} | \hat{H}_N | \Psi^{HF} \rangle = 0$ holds. The remaining expressions may be evaluated by considering only the fully contracted terms. We begin with the first non-vanishing term in Eq. 4.22

$$(\hat{H}_N\hat{T}_1)_c = (\hat{f}_N\hat{T}_1)_c + (\hat{V}_N\hat{T}_1)_c,$$

where we have already evaluated the first expression in Eq. 4.21. However, with only one fully contracted term this contribution gives

$$\langle \Psi^{HF} | (\hat{f}_N\hat{T}_1)_c | \Psi^{HF} \rangle = \sum_{ia} f_{ia} t_i^a.$$

Further, there are no fully contracted terms resulting from the expectation value with $(\hat{V}_N\hat{T}_1)_c$. Similarly, we find no fully contracted terms for $(\hat{f}_N\hat{T}_2)_c$ from the third expression in Eq. 4.22. The next non-zero term is the two-electron component

$$\begin{aligned}
\langle \Psi^{HF}|(\hat{V}_N\hat{T}_2)_c|\Psi^{HF}\rangle &= \frac{1}{16}\sum_{pqrs}\sum_{aibj}\langle pq||rs\rangle\, t_{ij}^{ab}\,\langle \Psi^{HF}|\{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\}\{\hat{a}_a^\dagger\hat{a}_b^\dagger\hat{a}_j\hat{a}_i\}|\Psi^{HF}\rangle \\
&= \frac{1}{16}\sum_{pqrs}\sum_{aibj}\langle pq||rs\rangle\, t_{ij}^{ab}\Bigg(\{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_b^\dagger\hat{a}_j\hat{a}_i\} + \{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_b^\dagger\hat{a}_j\hat{a}_i\} \\
&+ \{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_b^\dagger\hat{a}_j\hat{a}_i\} + \{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_b^\dagger\hat{a}_j\hat{a}_i\}\Bigg) \\
&= \frac{1}{16}\sum_{pqrs}\sum_{aibj}\langle pq||rs\rangle\, t_{ij}^{ab}\,(\delta_{pi}\delta_{qj}\delta_{ra}\delta_{sb} + \delta_{pj}\delta_{qi}\delta_{rb}\delta_{sa} - \delta_{pj}\delta_{qi}\delta_{ra}\delta_{sb} - \delta_{pi}\delta_{qj}\delta_{rb}\delta_{sa}) \\
&= \frac{1}{4}\sum_{aibj}\langle pq||rs\rangle\, t_{ij}^{ab}.
\end{aligned}$$

While $\frac{1}{2}(\hat{f}_N\hat{T}_1^2)_c$ again vanishes, $\frac{1}{2}(\hat{V}_N\hat{T}_1^2)_c$ produces four equivalent full contractions

$$\begin{aligned}
\frac{1}{2}\langle \Psi^{HF}|(\hat{V}_N\hat{T}_1^2)_c|\Psi^{HF}\rangle &= \frac{1}{8}\sum_{pqrs}\sum_{aibj}\langle pq||rs\rangle\, t_i^a t_j^b\,\langle \Psi^{HF}|\{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\}\{\hat{a}_a^\dagger\hat{a}_i\}\{\hat{a}_b^\dagger\hat{a}_j\}|\Psi^{HF}\rangle \\
&= \frac{1}{8}\sum_{pqrs}\sum_{aibj}\langle pq||rs\rangle\, t_i^a t_j^b\Bigg(\{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} \\
&+ \{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\} + \{\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_s\hat{a}_r\hat{a}_a^\dagger\hat{a}_i\hat{a}_b^\dagger\hat{a}_j\}\Bigg) \\
&= \frac{1}{8}\sum_{pqrs}\sum_{aibj}\langle pq||rs\rangle\, t_i^a t_j^b\,(-\delta_{pj}\delta_{qi}\delta_{ra}\delta_{sb} + \delta_{pj}\delta_{qi}\delta_{rb}\delta_{sa} + \delta_{pi}\delta_{qj}\delta_{ra}\delta_{sb} - \delta_{pi}\delta_{qj}\delta_{rb}\delta_{sa}) \\
&= \frac{1}{2}\sum_{aibj}\langle ij||ab\rangle\, t_i^a t_j^b.
\end{aligned}$$

All other terms vanish since in all of them we find more creation and annihilation operators than in the Hamiltonian. Therefore, no fully contracted products can be formed. This may also be interpreted in terms of Slater-Condon rules which state that contributions with excitations of third order and higher are zero. Finally, we arrive at the expression of the CCSD energy equation:

$$E_{cor}^{CCSD} = \sum_{ia} f_{ia}t_i^a + \frac{1}{4}\sum_{aibj}\langle pq||rs\rangle\, t_{ij}^{ab} + \frac{1}{2}\sum_{aibj}\langle ij||ab\rangle\, t_i^a t_j^b. \tag{4.23}$$

The first term vanishes if the Brillouin theory applies. The second term confirms the obtained expressions from the Møller-Plesset perturbation theory (s. Eq. 4.10). Thus, the last term in Eq. 4.23 may be viewed as the next higher order in the perturbation scheme. However, the derived CCSD energy equation is

exact, also for higher-order theories. This is because higher excitation cluster operators cannot produce fully contracted terms with the Hamiltonian. The contribution of the higher excitations is only reflected indirectly via the amplitude equations.

Now we shall investigate how to evaluate the matrix elements arising in the cluster amplitude equations (s. Eq. 4.17). The restriction to singly and doubly excited determinants leaves us with two distinct equations

$$\begin{aligned} \langle \Psi_i^a | \bar{H} | \Psi^{HF} \rangle &= 0, \\ \langle \Psi_{ij}^{ab} | \bar{H} | \Psi^{HF} \rangle &= 0. \end{aligned} \tag{4.24}$$

The first line in Eq. 4.24 is also referred to as the $\hat{T}_1$ amplitude equation while the second line is denoted the $\hat{T}_2$ amplitude equation. To be able to evaluate the fully contracted terms from Wick's theorem, the excited determinants need to be expressed in terms of the reference determinant $|\Psi^{HF}\rangle$:

$$\begin{aligned} \langle \Psi_i^a | &= \langle \Psi^{HF} | \, \hat{a}_i^\dagger \hat{a}_a, \\ \langle \Psi_{ij}^{ab} | &= \langle \Psi^{HF} | \, \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a. \end{aligned} \tag{4.25}$$

In the following, we again only examine some expectation values. It is reasonable to start with the evaluation of the bare normal ordered electronic Hamiltonian $\hat{H}_N$. For the $\hat{T}_1$ amplitude equation it holds:

$$\langle \Psi_i^a | \hat{f}_N + \hat{V}_N | \Psi^{HF} \rangle = \sum_{pq} f_{pq} \, \langle \Psi^{HF} | \{ \hat{a}_i^\dagger \hat{a}_a \} \{ \hat{a}_p^\dagger \hat{a}_q \} | \Psi^{HF} \rangle + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \, \langle \Psi^{HF} | \{ \hat{a}_i^\dagger \hat{a}_a \} \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \} | \Psi^{HF} \rangle \, .$$

Obviously, the two-electron component does not lead to fully contracted terms. The one with the Fock matrix, however, yields the following contribution:

$$\begin{aligned} \langle \Psi_i^a | \hat{f}_N | \Psi^{HF} \rangle &= \sum_{pq} f_{pq} \, \langle \Psi^{HF} | \{ \hat{a}_i^\dagger \hat{a}_a \} \{ \hat{a}_p^\dagger \hat{a}_q \} | \Psi^{HF} \rangle \\ &= \sum_{pq} f_{pq} \{ \hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q \} = \sum_{pq} f_{pq} \delta_{iq} \delta_{ap} \\ &= f_{ai}. \end{aligned}$$

The contributions of $\hat{H}_N$ regarding the $\hat{T}_2$ amplitude equation is somewhat similar except in this case the one-electron component vanishes while the two-electron component is fully contracted.

$$
\begin{aligned}
\langle \Psi_{ij}^{ab} | \hat{V}_N | \Psi^{HF} \rangle &= \frac{1}{4} \sum_{pqrs} \langle pq||rs \rangle \, \langle \Psi^{HF} | \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a\} \{\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} | \Psi^{HF} \rangle \, . \\
&= \frac{1}{4} \sum_{pqrs} \langle pq||rs \rangle \Bigg(\{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} \\
&+ \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} \Bigg) \\
&= \frac{1}{4} \sum_{pqrs} \langle pq||rs \rangle \, (\delta_{pa}\delta_{qb}\delta_{ri}\delta_{sj} - \delta_{pb}\delta_{qa}\delta_{ri}\delta_{sj} - \delta_{pa}\delta_{qb}\delta_{rj}\delta_{si} + \delta_{pb}\delta_{qa}\delta_{rj}\delta_{si}) \\
&= \langle ab||ij \rangle \, .
\end{aligned}
$$

As our last example we investigate the contributions which are linear in $\hat{T}_1$ (second term in Eq. 4.22). For both, the one-electron and two-electron operators, full contractions are possible. For the former we get:

$$
\begin{aligned}
\langle \Psi_i^a | (\hat{f}_N \hat{T}_1)_c | \Psi^{HF} \rangle &= \sum_{pq} \sum_{jb} f_{pq} \, t_j^b \, \langle \Psi^{HF} | \{\hat{a}_i^\dagger \hat{a}_a\} (\{\hat{a}_p^\dagger \hat{a}_q\} \{\hat{a}_b^\dagger \hat{a}_j\})_c | \Psi^{HF} \rangle \, . \\
&= \sum_{pq} \sum_{jb} f_{pq} t_j^b \Big(\{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q \hat{a}_b^\dagger \hat{a}_j\} + \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q \hat{a}_b^\dagger \hat{a}_j\} \Big) \\
&= \sum_{pq} \sum_{jb} f_{pq} t_j^b \, (\delta_{ij}\delta_{ap}\delta_{qb} - \delta_{iq}\delta_{ab}\delta_{pj}) \\
&= \sum_{b} f_{ab} t_i^b - \sum_{j} f_{ji} t_j^a ,
\end{aligned}
$$

while the latter contribution simplifies to:

$$
\begin{aligned}
\langle \Psi_i^a | (\hat{V}_N \hat{T}_1)_c | \Psi_0 \rangle &= \frac{1}{4} \sum_{pqrs} \sum_{jb} \langle pq||rs \rangle \, t_j^b \, \langle \Psi_0 | \{\hat{a}_i^\dagger \hat{a}_a\} (\{\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} \{\hat{a}_b^\dagger \hat{a}_j\})_c | \Psi_0 \rangle \, . \\
&= \frac{1}{4} \sum_{pqrs} \sum_{jb} \langle pq||rs \rangle \, t_j^b \Bigg(\{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_b^\dagger \hat{a}_j\} + \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_b^\dagger \hat{a}_j\} \\
&+ \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_b^\dagger \hat{a}_j\} + \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_b^\dagger \hat{a}_j\} \Bigg) \\
&= \frac{1}{4} \sum_{pqrs} \sum_{jb} \langle pq||rs \rangle \, t_j^b \, (-\delta_{pa}\delta_{qj}\delta_{rb}\delta_{si} + \delta_{pj}\delta_{qa}\delta_{rb}\delta_{si} + \delta_{pa}\delta_{qj}\delta_{ri}\delta_{sb} - \delta_{pj}\delta_{qa}\delta_{ri}\delta_{sb}) \\
&= \sum_{bj} \langle ja||bi \rangle \, t_j^b .
\end{aligned}
$$

To complete the linear term in $\hat{T}_1$ from Eq. 4.22, the contributions to the $\hat{T}_2$ amplitude equations are still to be found. The one-electron contribution

$$\langle \Psi_{ij}^{ab} | (\hat{f}_N \hat{T}_1)_c | \Psi^{HF} \rangle = \sum_{pq} \sum_{kc} f_{pq}\, t_k^c \, \langle \Psi^{HF} | \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a\} (\{\hat{a}_p^\dagger \hat{a}_q\}\{\hat{a}_c^\dagger \hat{a}_k\})_c | \Psi^{HF} \rangle = 0, \tag{4.26}$$

vanishes due to the missing connections of the cluster operator and the Hamiltonian elements. The two-electron term reads:

$$\begin{aligned}
\langle \Psi_{ij}^{ab} | (\hat{V}_N \hat{T}_1)_c | \Psi^{HF} \rangle &= \frac{1}{4} \sum_{pqrs} \sum_{kc} \langle pq || rs \rangle \; t_k^c \, \langle \Psi^{HF} | \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a\} (\{\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\}\{\hat{a}_c^\dagger \hat{a}_k\})_c | \Psi^{HF} \rangle \, . \\
&= \frac{1}{4} \sum_{pqrs} \sum_{kc} \langle pq || rs \rangle \; t_k^c \Bigg(\{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \\
&+ \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \\
&+ \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \\
&+ \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \\
&+ \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \\
&+ \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \Bigg) \\
&= \sum_{pq} \sum_{kc} f_{pq}\, t_k^c \, (\delta_{pa}\delta_{qb}\delta_{rc}\delta_{sj}\delta_{ik} - \delta_{pb}\delta_{qa}\delta_{rc}\delta_{sj}\delta_{ik} - \delta_{pa}\delta_{qb}\delta_{rc}\delta_{si}\delta_{jk} + \delta_{pb}\delta_{qa}\delta_{rc}\delta_{si}\delta_{jk} \\
&- \delta_{pa}\delta_{qb}\delta_{rj}\delta_{sc}\delta_{ik} + \delta_{pb}\delta_{qa}\delta_{rj}\delta_{sc}\delta_{ik} + \delta_{pa}\delta_{qb}\delta_{ri}\delta_{sc}\delta_{jk} - \delta_{pb}\delta_{qa}\delta_{ri}\delta_{sc}\delta_{jk} \\
&- \delta_{pa}\delta_{qk}\delta_{ri}\delta_{sj}\delta_{bc} + \delta_{pb}\delta_{qk}\delta_{ri}\delta_{sj}\delta_{ac} - \delta_{pb}\delta_{qk}\delta_{rj}\delta_{si}\delta_{ac} + \delta_{pa}\delta_{qk}\delta_{rj}\delta_{si}\delta_{bc} \\
&+ \delta_{pk}\delta_{qa}\delta_{ri}\delta_{sj}\delta_{bc} - \delta_{pk}\delta_{qb}\delta_{ri}\delta_{sj}\delta_{ac} - \delta_{pk}\delta_{qa}\delta_{rj}\delta_{si}\delta_{bc} + \delta_{pk}\delta_{qb}\delta_{rj}\delta_{si}\delta_{ac}) \\
&= \sum_c (\langle ab || cj \rangle \, t_i^c - \langle ab || ci \rangle \, t_j^c) + \sum_k (\langle ij || bk \rangle \, t_k^a - \langle ij || ak \rangle \, t_k^b).
\end{aligned}$$

The already rather tedious derivation of the second term in Eq. 4.22 shows the complexity of the coupled cluster amplitude equations. Once all contributions are computed using Wick's theorem we get the $\hat{T}_1$ and $\hat{T}_2$ amplitude equations [30] (s. Appendix A).

In summary, the coupled cluster singles and doubles equations are derived by introducing the similarity-transformed Hamiltonian $\bar{H}$ allowing us to define three distinct CCSD equations. The first one describes the CCSD energy (s. Eq. 4.23) while Eq. A.1 and Eq. A.2 give the $\hat{T}_1$ and $\hat{T}_2$ amplitude equations, respectively. With this, we are now able to perform the CCSD method only by evaluating the Fock matrix elements

f_{pq} as well as the antisymmetrized two-electron integrals $\langle pq||rs\rangle$. In the following, we focus on the actual implementation of the CCSD algorithm.

4.2.3 Implementation

While the implementation of the derived CCSD energy equation (s. Eq. 4.23) can be realized directly, the CCSD amplitude equations (s. Eqs. A.1 and A.2) are not readily solvable at first glance. However, rearranging the $\hat{T}_1$ and $\hat{T}_2$ equations leads to expressions that can be solved iteratively. Eq. A.1 may also be written as:

$$0 = f_{ai} + f_{aa}t_i^a - f_{ii}t_i^a + \sum_c (1-\delta_{ca}) f_{ac} t_c^i - \sum_k (1-\delta_{ik}) f_{ki} t_k^a + \dots \ .$$

We extract the diagonal elements of the Fock matrix and define

$$D_i^a \equiv f_{ii} - f_{aa}, \tag{4.27}$$

in order to get the reformulated $\hat{T}_1$ amplitude equation:

$$D_i^a t_i^a = f_{ai} + \sum_c (1-\delta_{ca}) f_{ac} t_c^i - \sum_k (1-\delta_{ik}) f_{ki} t_k^a + \dots \ .$$

By defining

$$D_{ij}^{ab} \equiv f_{ii} + f_{jj} - f_{aa} - f_{bb}, \tag{4.28}$$

we similarly get the reformulated $\hat{T}_2$ amplitude equation:

$$D_{ij}^{ab} t_{ij}^{ab} = \langle ab||ij\rangle + \hat{P}_{ab} \sum_c (1-\delta_{bc}) f_{bc} t_{ij}^{ac} - \hat{P}_{ij} \sum_k (1-\delta_{kj}) f_{kj} t_{ik}^{ab} + \dots \ .$$

In addition, we are able to reduce the computational cost by factorization of multiple terms in the CCSD amplitude equations. The basic idea is to identify factors that require computationally expensive operations. These can then be decomposed into a product of two or more terms by defining so-called intermediate quantities. However, there is no unique set of such intermediates. The most efficient factorization of the amplitude equations depends on both the structure of the code and the model under study. The latter entails that for closed-shell Hartree-Fock methods different intermediates are used than, e.g. for open-shell theories. We illustrate this factorization with one term in the $\hat{T}_2$ amplitude equation (s. last term in the third line of Eq. A.2):

$$\frac{1}{4} \sum_{klcd} \langle kl||cd\rangle \, t_{ij}^{cd} t_{kl}^{ab} = \frac{1}{2} \sum_{kl} t_{kl}^{ab} \frac{1}{2} \sum_{cd} \langle kl||cd\rangle \, t_{ij}^{cd} = -\hat{P}_{ab} \frac{1}{2} \sum_{kl} t_{kl}^{ab} X_{ij}^{kl},$$

where we introduce the X intermediate as:

$$X_{ij}^{kl} = \frac{1}{2}\sum_{cd} \langle kl||cd\rangle\, t_{ij}^{cd}.$$

The original term scales with $\mathcal{O}(N_o^4 N_v^4)$ with N_o and N_v being the number of the occupied and virtual spin orbitals, respectively. The factorized term, in contrast, scales only with $\mathcal{O}(N_o^4 N_v^2)$. This is due to the fact that we are able to first construct X_{ij}^{kl} and then perform the contraction with the amplitudes t_{kl}^{ab}. Thus, we can conclude that the coupled cluster amplitude equations may be rewritten in terms of only linear components in $\hat{T}$ by factorizing every non-linear term. Consequently, the most expensive computations in our CCSD algorithm scale at worst with $\mathcal{O}(N_{SO}^6)$ where N_{SO} is the number of spin orbitals.

In this work, we use the factorization scheme deduced in Refs. [32, 33]. Thereby, the one and two-particle intermediates $\mathcal{F}_{pq}$ and $\mathcal{W}_{pqrs}$ are introduced. In the following, we provide the final equations of these intermediates as well as the final CCSD equations for the implementation. For more details we refer to aforementioned Refs. [32, 33]:

$$\begin{aligned}
\mathcal{F}_{ae} &= (1-\delta_{ae})f_{ae} - \frac{1}{2}\sum_{m} f_{me}t_m^a + \sum_{mf} t_m^f \langle ma||fe\rangle - \frac{1}{2}\sum_{mnf} \tilde{\tau}_{mn}^{af} \langle mn|ef\rangle\,, \\
\mathcal{F}_{mi} &= (1-\delta_{mi})f_{mi} + \frac{1}{2}\sum_{e} t_i^e f_{me} + \sum_{en} t_n^e \langle mn||ie\rangle + \frac{1}{2}\sum_{nef} \tilde{\tau}_{in}^{ef} \langle mn|ef\rangle\,, \\
\mathcal{F}_{me} &= f_{me} + \sum_{nf} t_n^f \langle mn||ef\rangle\,, \\
\mathcal{W}_{mnij} &= \langle mn||ij\rangle + \hat{P}_{ij}\sum_{e} t_j^e \langle mn||ie\rangle + \frac{1}{4}\sum_{ef} \tau_{ij}^{ef} \langle mn||ef\rangle\,, \\
\mathcal{W}_{abef} &= \langle ab||ef\rangle - \hat{P}_{ab}\sum_{m} t_m^b \langle am||ef\rangle + \frac{1}{4}\sum_{ef} \tau_{mn}^{ab} \langle mn||ef\rangle\,, \\
\mathcal{W}_{mbej} &= \langle mb||ej\rangle + \sum_{f} t_j^f \langle mb||ef\rangle - \sum_{n} t_n^b \langle mn||ej\rangle - \sum_{nf}(\frac{1}{2}t_{jn}^{fb} + f_j^f t_n^b) \langle mn||ef\rangle\,.
\end{aligned} \tag{4.29}$$

Here, the effective two-particle excitation operators are introduced as:

$$\begin{aligned}
\tilde{\tau}_{ij}^{ab} &= t_{ij}^{ab} + \frac{1}{2}(t_i^a t_j^b - t_i^b t_j^a), \\
\tau_{ij}^{ab} &= t_{ij}^{ab} + (t_i^a t_j^b - t_i^b t_j^a).
\end{aligned}$$

In the end, we are able to provide the three CCSD equations which can readily be implemented:

$$\begin{aligned}
t_i^a D_i^a = f_{ia} &+ \sum_{e} t_i^e \mathcal{F}_{ae} - \sum_{m} t_m^a \mathcal{F}_{mi} + \sum_{me} t_{im}^{ae} \mathcal{F}_{me} - \sum_{nf} t_n^f \langle na||if\rangle \\
&- \frac{1}{2}\sum_{mef} t_{im}^{ef} \langle ma||ef\rangle - \frac{1}{2}\sum_{men} t_{mn}^{ae} \langle nm||ei\rangle\,.
\end{aligned} \tag{4.30}$$

$$
\begin{aligned}
t_{ij}^{ab} D_{ij}^{ab} = \langle ij||ab\rangle &+ \hat{P}_{ab} \sum_{e} t_{ij}^{ae} (\mathcal{F}_{be} - \frac{1}{2} \sum_{m} t_{m}^{b} \mathcal{F}_{me}) - \hat{P}_{ij} \sum_{m} t_{im}^{ab} (\mathcal{F}_{mj} - \frac{1}{2} \sum_{e} t_{j}^{e} \mathcal{F}_{me}) \\
&+ \frac{1}{2} \sum_{mn} \tau_{mn}^{ab} \mathcal{W}_{mnij} + \frac{1}{2} \sum_{ef} \tau_{ij}^{ef} \mathcal{W}_{abef} + \hat{P}_{ij} \hat{P}_{ab} \sum_{me} (t_{im}^{ae} \mathcal{W}_{mbej} - t_{i}^{e} t_{m}^{a} \langle mb||ej\rangle) \\
&+ \hat{P}_{ij} \sum_{e} t_{i}^{e} \langle ab||ej\rangle - \hat{P}_{ab} \sum_{m} t_{m}^{a} \langle mb||ij\rangle \,.
\end{aligned} \tag{4.31}
$$

We want to emphasize that the CCSD energy equation does not need to be expressed via intermediate quantities and is therefore already provided by Eq. 4.23.

The derived CCSD equation system is a coupled non-linear problem, which as mentioned earlier may be solved iteratively. Thus, we need to self-consistently update the $\hat{T}_1$ and $\hat{T}_2$ amplitudes till convergence is reached. For the necessary initial guess, all amplitudes on the right-hand side of Eq. 4.30 are set to zero. This allows us to calculate the $\hat{T}_1$ amplitude directly using the elements of the Fock matrix

$$t_i^a = \frac{f_{ia}}{D_i^a}.$$

Similarly, in the first iterative step the $\hat{T}_2$ amplitude reads:

$$t_{ij}^{ab} = \frac{\langle ij||ab\rangle}{D_{ij}^{ab}}.$$

The initial guess for the $\hat{T}_2$ amplitude equations is the same as the first-order wave function from the Møller-Plesset perturbation theory. The $\hat{T}_1$ amplitude guess vanishes because the Fock matrix elements are just the orbital energies.

The algorithm itself, as in MPPT, requires the Fock matrix f_{pq} and the antisymmetrized two-electron integrals given by the Coulomb tensor V_{rs}^{-pq} (s. Eq. 4.8). Afterwards, the intermediates $\mathcal{F}$ and $\mathcal{W}$ as well as the $\hat{T}$ amplitudes, starting from the initial estimate, are iteratively updated. In addition, the CCSD energy is calculated in each iteration. Convergence is reached when the energy no longer changes between two iterations up to a certain tolerance. Similarly to our Hartree-Fock algorithm, the CCSD algorithm breaks when the energy difference falls below 10^{-6}.
Our algorithm relies on the repository *pyqchem*, which converts the equations described above from Ref. [32] into code language which turned out to be useful when dealing with the tedious index notation of the Eqs. 4.29, 4.30 and 4.31.

4.3 Equation of Motion Coupled Cluster Theory

The Equation of Motion coupled cluster formalism (EOM-CC) is a powerful tool for the calculation of excitation energies. In particular, very accurate and robust descriptions of excited states can be obtained, with the only drawback that for large molecules the computational cost becomes significant. The EOM-CC ansatz relies on the diagonalization of the similarity-transformed Hamiltonian $\bar{H}$ in a chosen subspace of the Fock space. Thereby, it is distinguished between excited electron, electron-attached and ionized states. Since we focus exclusively on the excited electron formalism (EE EOM-CC) in the remainder of

this chapter, we continue to use the term EOM-CC for simplicity. For this case, the number of electrons remains unchanged.

4.3.1 Theoretical Basics

Starting point for the EOM-CC formalism is again the time-independent Schrödinger equation

$$\hat{H}_N \left|\Psi\right\rangle = E \left|\Psi\right\rangle ,$$

where we use the normal-ordered Hamiltonian $\hat{H}_N$ introduced in Eq. 4.18. The EOM-CC ansatz for the exact wave function $|\Psi\rangle$ reads [16]:

$$\left|\Psi\right\rangle = \hat{R} \left|\Psi^{CC}\right\rangle = \hat{R} e^{\hat{T}} \left|\Psi^{HF}\right\rangle ,$$

where $\hat{R}$ is a linear excitation operator with the same properties as the cluster operator $\hat{T}$. The excitation operator may be written as $\hat{R} = \hat{R}_0 + \hat{R}_1 + \hat{R}_2 + \dots$ and can additionally be expressed in terms of creation and annihilation operators

$$\hat{R}_n = \left(\frac{1}{n!}\right)^2 \sum_{ij\dots ab\dots} r_{ij\dots}^{ab\dots} \hat{a}_a^\dagger \hat{a}_b^\dagger \dots \hat{a}_j \hat{a}_i ,$$

where $r_{ij\dots}^{ab\dots}$ are the excitation amplitudes. Consequently, the Schrödinger equation reads:

$$\hat{H}_N \hat{R} e^{\hat{T}} \left|\Psi^{HF}\right\rangle = E \hat{R} e^{\hat{T}} \left|\Psi^{HF}\right\rangle .$$

Since $\hat{R}$ and $\hat{T}$ commute, we may express the Schrödinger equation in terms of the similarity-transformed Hamiltonian $\bar{H}$:

$$e^{-\hat{T}} \hat{H}_N e^{\hat{T}} \hat{R} \left|\Psi^{HF}\right\rangle = \bar{H} \hat{R} \left|\Psi^{HF}\right\rangle = E \hat{R} \left|\Psi^{HF}\right\rangle . \tag{4.32}$$

$\bar{H}$ is not Hermitian and therefore we must distinguish between left and right eigenvectors. Similarly, we may define the de-excitation operator $\hat{L}$:

$$\hat{L} = \hat{L}_0 + \hat{L}_1 + \hat{L}_2 + \dots ,$$
$$\hat{L}_n = \left(\frac{1}{n!}\right)^2 \sum_{ij\dots ab\dots} l_{ij\dots}^{ab\dots} \hat{a}_i^\dagger \hat{a}_j^\dagger \dots \hat{a}_b \hat{a}_a .$$

From the properties of the non-Hermitian eigenvalue problem, it follows that the two sets of solutions are biorthogonal

$$\langle \Psi^{HF} \hat{L}_i e^{-\hat{T}} | e^{\hat{T}} \hat{R}_j \Psi^{HF} \rangle = c\delta_{ij} .$$

Choosing the constant c to be unity, leads to the EOM-CC normalization condition. Therefore, we can express the EOM-CC energy as follows:

$$E = \langle \Psi^{HF} \hat{L} | \bar{H} | \hat{R} \Psi^{HF} \rangle \,.$$

However, if one is only interested in the excitation energies it is sufficient to solve the right eigenvalue problem with the $\hat{R}$ operator. We may further subtract the ground state CC equation $\hat{R}\bar{H}\,|\Psi^{HF}\rangle = \hat{R}E^{CC}\,|\Psi^{HF}\rangle$ from Eq. 4.32:

$$\begin{aligned} (\bar{H}\hat{R} - \hat{R}\bar{H})\,|\Psi^{HF}\rangle &= (E\hat{R} - \hat{R}E^{CC})\,|\Psi^{HF}\rangle\,, \\ [\bar{H}, \hat{R}]\,|\Psi^{HF}\rangle &= \Delta E \hat{R}\,|\Psi^{HF}\rangle\,, \end{aligned} \tag{4.33}$$

where $\Delta E = E - E^{CC}$ are the excitation energies which may again be seen as correlation energy since the ground state energy of the coupled cluster theory is deducted. Accordingly, the full energy spectrum is given by

$$E_i = E^{CC} + \Delta E_i = E^{HF} + E^{CC}_{cor} + \Delta E_i. \tag{4.34}$$

Unlike Eq. 4.33, we introduce the subscript i to explicitly emphasize the distinct energy levels. Moreover, the form of Eq. 4.33 explains the name equation of motion because its structure is similar to the time-dependent Heisenberg picture differential equation for the time evolution of an operator. We point out that the commutator $[\bar{H}, \hat{R}]$ can be simplified, like in the CCSD theory, such that only contributions from connected terms need to be considered. Finally, we are left to solve:

$$(\bar{H}\hat{R})_c\,|\Psi^{HF}\rangle = \Delta E \hat{R}\,|\Psi^{HF}\rangle\,. \tag{4.35}$$

4.3.2 Equation of Motion Coupled Cluster Singles and Doubles

Applying the general theory of the EOM-CC method to the restriction to singles and doubles (EOM-CCSD) implies, again, the truncation of the cluster as well as the excitation operator

$$\begin{aligned} \hat{T} &\approx \hat{T}_1 + \hat{T}_2, \\ \hat{R} &\approx \hat{R}_1 + \hat{R}_2. \end{aligned}$$

In contrast to the theoretical introduction, we don't consider the $\hat{R}_0$ contribution, which is simply the already calculated CCSD ground state solution. This may be viewed as a projection onto the singles and doubles excitation manifold of the Hilbert space. Nevertheless, we need to solve the eigenvalue problem derived in Eq. 4.35. This corresponds to the solution of the following matrix equations [16, 34, 35]:

$$\underbrace{\begin{bmatrix} \bar{H}_{SS} & \bar{H}_{SD} \\ \bar{H}_{DS} & \bar{H}_{DD} \end{bmatrix}}_{\bar{H}^{EOM-CCSD}} \begin{bmatrix} \hat{R}_1 \\ \hat{R}_2 \end{bmatrix} = \Delta E \begin{bmatrix} \hat{R}_1 \\ \hat{R}_2 \end{bmatrix}, \tag{4.36}$$

where $\bar{H}_{SS} = \langle \Psi_i^a | \bar{H} | \Psi_k^c \rangle$ corresponds to the singles-singles block of the Hamiltonian matrix, $\bar{H}_{SD}$ to the singles-doubles block and so on. For convenience, we also introduce the matrix $\bar{H}^{EOM-CCSD} = \langle \mu | \bar{H} | \nu \rangle$ with $\mu, \nu = \{\Psi_r^p, \Psi_{rs}^{pq}\}$ containing all four submatrices.

Thus, the determination of the EOM-CCSD equations is a very similar task to the one already discussed in CCSD theory. Nevertheless, we want to give some exemplary calculations in order to convey the basic idea behind it. We start by looking at the singles-singles block of the Hamiltonian matrix in Eq. 4.36. Considering the form of the connected similarity-transformed Hamiltonian $\bar{H}$ (s. Eq. 4.22), the lowest order contributions read

$$\langle \Psi_i^a | \hat{H}_N + (\hat{H}_N \hat{T}_1)_c + (\hat{H}_N \hat{T}_2)_c + \ldots | \Psi_k^c \rangle \,.$$

With the definition of the normal-order Hamiltonian $\hat{H}_N$, we may evaluate:

$$\begin{aligned} \langle \Psi_i^a | \hat{f}_N | \Psi_k^c \rangle &= \sum_{pq} f_{pq} \, \langle \Psi^{HF} | \{\hat{a}_i^\dagger \hat{a}_a\} \{\hat{a}_p^\dagger \hat{a}_q\} \{\hat{a}_c^\dagger \hat{a}_k\} | \Psi^{HF} \rangle \\ &= \sum_{pq} f_{pq} \left(\{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q \hat{a}_c^\dagger \hat{a}_k\} \right) \\ &= \sum_{pq} f_{pq} (\delta_{ik} \delta_{ap} \delta_{qc} - \delta_{iq} \delta_{ac} \delta_{pk}) \\ &= f_{ac} \delta_{ik} - f_{ik} \delta_{ac}, \end{aligned}$$

as well as

$$\begin{aligned} \langle \Psi_i^a | \hat{V}_N | \Psi_k^c \rangle &= \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \, \langle \Psi^{HF} | \{\hat{a}_i^\dagger \hat{a}_a\} \{\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\} \{\hat{a}_c^\dagger \hat{a}_k\} | \Psi^{HF} \rangle \\ &= \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \left(\{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \right. \\ &\left. + \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} + \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_k\} \right) \\ &= \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \, (-\delta_{pa} \delta_{qk} \delta_{rc} \delta_{si} + \delta_{pk} \delta_{qa} \delta_{rc} \delta_{si} + \delta_{pa} \delta_{qk} \delta_{ri} \delta_{sc} - \delta_{pk} \delta_{qa} \delta_{ri} \delta_{sc}) \\ &= \langle ak || ic \rangle \,. \end{aligned}$$

As our last example, we evaluate the $\hat{f}_N\hat{T}_1$ contribution. The only non-vanishing terms are:

$$\begin{aligned}
\langle\Psi_i^a|(\hat{f}_N\hat{T}_1)_c|\Psi_k^c\rangle &= \sum_{pq}\sum_{bj} f_{pq}t_j^b \langle\Psi^{HF}|\{\hat{a}_i^\dagger\hat{a}_a\}(\{\hat{a}_p^\dagger\hat{a}_q\}\{\hat{a}_b^\dagger\hat{a}_j\})_c\{\hat{a}_c^\dagger\hat{a}_k\}|\Psi^{HF}\rangle \\
&= \sum_{pq}\sum_{bj} f_{pq}t_j^b \left(\overbrace{\{\hat{a}_i^\dagger\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_b^\dagger\hat{a}_j\hat{a}_c^\dagger\hat{a}_k\}} + \overbrace{\{\hat{a}_i^\dagger\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_b^\dagger\hat{a}_j\hat{a}_c^\dagger\hat{a}_k\}} \right) \\
&= \sum_{pq}\sum_{bj} f_{pq}t_j^b(-\delta_{ij}\delta_{ac}\delta_{pk}\delta_{qb} - \delta_{ik}\delta_{ab}\delta_{pj}\delta_{qc}) \\
&= -f_{kb}t_i^b\delta_{ac} - f_{jc}t_j^a\delta_{ik}.
\end{aligned}$$

For the remaining blocks, we only evaluate the first expectation value of $\bar{H}$ due to the raising complexity of the equations. For the $\bar{H}_{SD}$ part, the contribution of the normal-ordered Fock operator reads:

$$\begin{aligned}
\langle\Psi_i^a|\hat{f}_N|\Psi_{kl}^{cd}\rangle &= \sum_{pq} f_{pq} \langle\Psi^{HF}|\{\hat{a}_i^\dagger\hat{a}_a\}\{\hat{a}_p^\dagger\hat{a}_q\}\{\hat{a}_c^\dagger\hat{a}_d^\dagger\hat{a}_l\hat{a}_k\}|\Psi^{HF}\rangle \\
&= \sum_{pq} f_{pq} \left(\overbrace{\{\hat{a}_i^\dagger\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_c^\dagger\hat{a}_d^\dagger\hat{a}_l\hat{a}_k\}} + \overbrace{\{\hat{a}_i^\dagger\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_c^\dagger\hat{a}_k^\dagger\hat{a}_l\hat{a}_k\}} \right) \\
&= \sum_{pq} f_{pq}\delta_{ik}\delta_{ad}\delta_{pl}\delta_{qc} = f_{lc}\delta_{ik}\delta_{ad}.
\end{aligned}$$

The contribution of $\hat{f}_N$ to the doubles-singles block $\bar{H}_{DS}$ vanishes similarly as in Eq. 4.26. The lowest order term for the doubles-doubles block results in:

$$\begin{aligned}
\langle\Psi_{ij}^{ab}|\hat{f}_N|\Psi_{kl}^{cd}\rangle &= \sum_{pq} f_{pq} \langle\Psi^{HF}|\{\hat{a}_i^\dagger\hat{a}_j^\dagger\hat{a}_b\hat{a}_a\}\{\hat{a}_p^\dagger\hat{a}_q\}\{\hat{a}_c^\dagger\hat{a}_d^\dagger\hat{a}_l\hat{a}_k\}|\Psi^{HF}\rangle \\
&= \sum_{pq} f_{pq} \left(\overbrace{\{\hat{a}_i^\dagger\hat{a}_j^\dagger\hat{a}_b\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_c^\dagger\hat{a}_d^\dagger\hat{a}_l\hat{a}_k\}} + \overbrace{\{\hat{a}_i^\dagger\hat{a}_j^\dagger\hat{a}_b\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_c^\dagger\hat{a}_d^\dagger\hat{a}_l\hat{a}_k\}} \right. \\
&\left. + \overbrace{\{\hat{a}_i^\dagger\hat{a}_j^\dagger\hat{a}_b\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_c^\dagger\hat{a}_d^\dagger\hat{a}_l\hat{a}_k\}} + \overbrace{\{\hat{a}_i^\dagger\hat{a}_j^\dagger\hat{a}_b\hat{a}_a\hat{a}_p^\dagger\hat{a}_q\hat{a}_c^\dagger\hat{a}_d^\dagger\hat{a}_l\hat{a}_k\}} \right) \\
&= \sum_{pq} f_{pq} \left(\delta_{jk}\delta_{il}\delta_{ad}\delta_{bp}\delta_{qc} - \delta_{jk}\delta_{il}\delta_{bd}\delta_{ap}\delta_{qc} + \delta_{jl}\delta_{ad}\delta_{bc}\delta_{iq}\delta_{pk} - \delta_{il}\delta_{ad}\delta_{bc}\delta_{jq}\delta_{pk}\right) \\
&= f_{bc}\delta_{jk}\delta_{il}\delta_{ad} - f_{ac}\delta_{jk}\delta_{il}\delta_{bd} + f_{ki}\delta_{jl}\delta_{ad}\delta_{bc} - f_{kj}\delta_{il}\delta_{ad}\delta_{bc}.
\end{aligned}$$

It is obvious that also in the case of the EOM-CCSD equations, Wick's theorem can be applied. In this case, the algebraic effort increases, since not only expectation values of operators between an excited state and the ground state must be calculated, but expectation values of operators between two excited states.

The increased amount of work is due to the fact that for each additional excitation, a pair of creation and annihilation operators has to be added and then to be contracted as well.
Nevertheless, the matrix eigenvalue equation ansatz (s. Eq. 4.36) turns out to be a very powerful technique, as it simplifies the task of evaluating the aforementioned expectation values. This means, that no iterative procedure such as for the CCSD amplitude equations is needed. In the next section, we discuss the implementation of the EOM-CCSD method, where we also provide the full set of the EOM-CCSD equations.

4.3.3 Implementation

For our model, we obtain the excitation energies by performing an exact diagonalization of the EOM-CCSD Hamiltonian matrix. We want to mention that, due to the high computational expenses, for systems with a large basis set and many electrons Eq. 4.36 is solved by means of the Davidson algorithm or variants of it. Before we are able to diagonalize the full Hamiltonian in the singles and doubles subspace, we need to construct the respective Hamiltonian matrix $\bar{H}^{EOM-CCSD}$. For this, all contributions to each block in Eq. 4.36, presented in Appendix B, need to be taken into account. The final dimension of the matrix $\bar{H}^{EOM-CCSD}$ may be viewed as follows. The individual blocks H_{SS}, H_{SD}, H_{DS} and H_{DD} have the dimensions $N_S \times N_S$, $N_S \times N_D$, $N_D \times N_S$ and $N_D \times N_D$. Then $\bar{H}^{EOM-CCSD}$ is obtained by vertically stacking the first and last two submatrices, respectively. The resulting block matrix has accordingly the dimension of $(N_S + N_D) \times (N_S + N_D)$.

Since the EOM-CC method does not require any self-consistent algorithm, there is no core part regarding the implementation. However, it is important to note that the amplitudes t_r^p and t_{rs}^{pq} for single and double excitations, respectively, are obtained by the CCSD algorithm. Moreover, our algorithm computes Eqs. B.1-B.4 to build the final EOM-CCSD matrix. After that, the eigenvalue problem is solved by means of a standard diagonalization package. Finally, the eigenvalues corresponding to the excitation energies are obtained. It is worthwhile noting that the eigenvalues, due to the nature of the non-Hermitian matrix, can also be complex. Complex eigenvalues have the interpretation of resonance or metastable states. The imaginary part corresponds thereby to the inverse of the decay time, i.e. the decay rate. The real part represents the actual energy. Thus, if the obtained eigenvalue has an imaginary part, it is an indication that the corresponding state will eventually decay or transition to another state.
Our implemented code is again oriented to the repository *pyqchem*, which turned out to be useful due to the error-prone formalism of the derived EOM-CCSD equations.

In summary, the investigated post-Hartree-Fock methods give more accurate results compared to the Hartree-Fock method itself. Thereby, MPPT represents an intuitive extension by means of Rayleigh-Schrödinger perturbation theory. We gain further accuracy by introducing the CC formalism. By evaluating the matrix elements concerning the similarity-transformed Hamiltonian with the Hartree-Fock reference state, we derive the CC energy and amplitude equations. These are solved iteratively, whereby the first iteration gives exactly the MPPT energy. Furthermore, we investigate not only ground state energies but the entire energy spectrum. The EOM-CC theory provides a powerful tool for the calculation of the aforementioned excited states energies. Thereby, by investigating only electronic excitations, the ansatz with the excitation operator $\hat{R}$ leads to a standard eigenvalue problem. In the subspace of the single and double excitations this yields a matrix equation, which in our case is solved by means of exact diagonalization. Similarly, to the CCSD method, expectation values between excited states with the similarity-transformed Hamiltonian are evaluated in order to construct the EOM-CCSD Hamiltonian itself.

5 Low-Rank Approximation

Besides the implementation of the Hartree-Fock, CCSD and EOM-CCSD theories for our model Hamiltonian, our major focus lies on the low-rank approximation of the two-electron integrals. One of the main limiting factors of the coupled cluster methods is that they are computationally rather expensive. CCSD and EOM-CCSD e.g. scale with $\mathcal{O}(N_{SO}^6)$ with N_{SO} being the number of molecular spin orbitals. Thereby, the most expensive calculations represent the operations involving the two-electron integrals. In the following, we present two techniques allowing us to reduce the overall computational cost of these coupled cluster methods to $\mathcal{O}(N_{SO}^5)$ [36], namely the singular value decomposition (SVD) and the canonical polyadic decomposition (CPD).

5.1 Motivation

We define an arbitrary tensor in the sense of multidimensional arrays

$$\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k},$$

where k denotes the order of the tensor and $n_1 \times n_2 \times \cdots \times n_k$, with $n_{i=1,2,\ldots,k}$ integers, its size. As can readily be seen, the number of entries of a tensor increases exponentially with its rank for constant $n_i = n$. Therefore, high-dimensional objects lead to huge memory requirements. Alone a rank 15 tensor with $n = 5$ entries per dimension requires more than 2 gigabyte storage space. It is therefore inevitable to reduce the rank of the tensor to be able to perform numerical calculations. Besides the storage, we also need to consider the computation time. In general, the computational cost of an arbitrary tensor operation is proportional to the number of floating point operations (FLOPs). The number of FLOPs depends on the size and the rank of the involved tensors, as well as on the number of involved contractions. For example, two tensors A_x^y and B_y^z are contracted along the particular dimension y eliminating the size of that dimension from the resulting tensor

$$\sum_y A_x^y B_y^z = A_x^y B_y^z = C_x^z,$$

where it is summed over all indices appearing more than once. In the second equality, the Einstein sum convention is applied. The contraction of this example represents a simple matrix multiplication which scales naively with $\mathcal{O}(N_x N_y N_z)$. Further factors for increased computation times may be of numerical nature such as the structure of the algorithm.

The goal of low-rank approximation techniques is to express a tensor of k-th order as a sum of tensors m-th order, where $m < k$ holds. This approximation allows not only to reduce the number of FLOPs required for the specific tensor contractions but also saves storage. However, one must consider the tradeoff of information loss, i.e., accuracy.

5.2 Singular Value Decomposition

We proceed from the definition of the Coulomb integral tensor V_{rs}^{pq} in Eq. 4.7, being a tensor of fourth order with dimension $N_{SO} \times N_{SO} \times N_{SO} \times N_{SO}$. In the following, we investigate how to decompose the Coulomb tensor in terms of the singular value decomposition.
The SVD as a generalization of the eigendecomposition allows the factorization of non-square matrices.

Since a matrix is a tensor of rank two, we compress our Coulomb tensor by swapping the indices q and r and defining the compound indices $P \equiv \{p, r\}$ and $Q \equiv \{q, s\}$. Consequently, the SVD reads:

$$V_Q^P = U_{N_\sigma}^P \Sigma_{N_\sigma}^{N_\sigma} W_Q^{N_\sigma}, \tag{5.1}$$

N_σ represents the dimension of the matrix Σ and thus denotes the number of singular values of the system. Now, the Coulomb tensor is represented via two third-order tensors $U_{N_\sigma}^P$ and $W_Q^{N_\sigma}$, and a second-order tensor $\Sigma_{N_\sigma}^{N_\sigma}$. We may exploit the fact that the singular values σ contained in Σ are in descending order and decay super-exponentially. This allows us to truncate the SVD according to:

$$\frac{\sigma_i}{\sigma_0} > \epsilon, \tag{5.2}$$

where σ_0 is the first, and therefore largest singular value. ϵ represents the threshold that determines the number of remaining singular values $\tilde{N}_\sigma$ after the truncation. In addition, ϵ may also be viewed as the accuracy describing how much information of the original tensor is retained. Since the singular values decay super-exponentially, most information is contained in the first few singular values and vectors, respectively. It is noteworthy that the SVD does not change the size of the involved dimensions but results in a reduction of storage since the original tensor is represented with fewer values which in turn requires fewer bits. Also, the computational cost scales as $\mathcal{O}(N_P N_{\tilde{N}_\sigma} N_Q)$. Therefore, the number of truncated singular values $\tilde{N}_\sigma$ affects the computation time, which is a beneficial property. Thereafter, the Coulomb tensor V_{qs}^{pr} is reconstructed with the truncated quantities:

$$\bar{V}_Q^P = U_{\tilde{N}_\sigma}^P \Sigma_{\tilde{N}_\sigma}^{\tilde{N}_\sigma} W_Q^{\tilde{N}_\sigma}. \tag{5.3}$$

We call $\bar{V}_Q^P$ the SVD truncated Coulomb tensor and in the next section we analyze a further factorization of it.

5.3 Canonical Polyadic Decomposition

Despite the fact that the SVD leads to a reduction of the rank of the original Coulomb tensor, it is reasonable to decompose the obtained tensors further. We assume the Coulomb tensor V_{qs}^{pr} to be already SVD decomposed and continue therefore with the truncated Coulomb tensor $\bar{V}_{qs}^{pr}$. First, we introduce the so-called optimized auxiliary field Coulomb vertex [36]

$$\Gamma_{\tilde{N}_\sigma}^{pr} \equiv U_{\tilde{N}_\sigma}^{pr} \left(\Sigma_{\tilde{N}_\sigma}^{\tilde{N}_\sigma}\right)^{1/2}. \tag{5.4}$$

This ansatz allows us to formally write the truncated Coulomb tensor as follows:

$$\bar{V}_{qs}^{pr} \approx \Gamma_{\tilde{N}_\sigma}^{pr} \Gamma_{qs}^{\tilde{N}_\sigma}. \tag{5.5}$$

The exchanged lower and upper indices indicate the conjugate transpose of the Coulomb vertex. This representation follows immediately from the relation $U = W^T$ from the singular value decomposition. Similarly to the Coulomb tensor, we may apply also a tensor rank decomposition for the optimized auxiliary field Coulomb vertex in Eq. 5.5. Thereby, $\Gamma^{pr}_{\tilde{N}_\sigma}$ is represented as a product of three tensors of second order

$$\Gamma^{pr}_{\tilde{N}_\sigma} \approx \Pi^p_R \Pi^r_R \Lambda^R_{\tilde{N}_\sigma}. \tag{5.6}$$

Here, $\Lambda^R_{\tilde{N}_\sigma}$ is called the Coulomb factor and Π^r_R are referred to as the factor orbitals [36]. Furthermore, we define N_R as the number of vertex indices R. N_R influences the dimension of the Coulomb factors and the factor orbitals and may therefore be interpreted as a measure of accuracy. The factorization in Eq. 5.6 is known as the canonical polyadic decomposition [37].

In the end, the CPD yields the desired approximate factorization of the Coulomb integrals. The final form for the approximated Coulomb tensor reads:

$$\bar{V}^{pr}_{qs} \approx \Gamma^{pr}_{\tilde{N}_\sigma} \Gamma^{\tilde{N}_\sigma}_{qs} \approx \Pi^p_R \Pi^r_R \Lambda^R_{\tilde{N}_\sigma} \Lambda^{\tilde{N}_\sigma}_S \Pi^S_q \Pi^S_s. \tag{5.7}$$

In summary, we are able to factorize the Coulomb integrals in such a way that the original fourth-order tensor is represented by a sequence of matrices, whereas the scaling of the computational cost with the system size does not exceed $\mathcal{O}(N^4)$.

5.4 Implementation

The SVD and CPD decomposed Coulomb integrals derived in Eq. 5.7 are now to be implemented numerically. Thereby, we want to point out that a tensor rank decomposition such as CPD has no unique solution. On the contrary, there are infinitely many Coulomb factors and factor orbitals satisfying Eq. 5.6. Thus, we are left with an optimization problem that is high dimensional, at least for realistic systems. The Coulomb factors and factor orbitals may be fit in the following way:

$$(\Lambda, \Pi) = \underset{\Lambda,\Pi}{argmin} ||\Pi^p_R \Pi^r_R \Lambda^R_{\tilde{N}_\sigma} - \Gamma^{pr}_{\tilde{N}_\sigma}||^2_F, \tag{5.8}$$

where $||\ ||_F$ denotes the Frobenius norm. There exist several ways to tackle this optimization problem in Eq. 5.8. While local algorithms tend to require up to thousands of steps [38], global algorithms have turned out to be more suitable for such problems. In this work we want to follow the ideas of Ref. [36] and use a regularized alternating least squares (RALS) algorithm which represents an extension to the standard alternating least squares (ALS) [39] algorithm defined via Eq. 5.8.
For simplicity, the following derivations are done with arbitrary tensors without Greek notation. We define the third order tensor T_{ijk} which shall be approximated by the matrices A_{iR}, B_{jR} and C_{kR}. In contrast to Eq. 5.8 where the factor orbitals appear quadratically, these three matrices are assumed to be distinct. The ALS algorithm results in three minimization problems for the three matrices

$$
\begin{aligned}
A^{(n+1)} &= \underset{A}{argmin}||A_{iR}B_{jR}^{(n)}C_{kR}^{(n)} - T_{ijk}||_F^2, \\
B^{(n+1)} &= \underset{B}{argmin}||A_{iR}^{(n+1)}B_{jR}C_{kR}^{(n)} - T_{ijk}||_F^2, \\
C^{(n+1)} &= \underset{C}{argmin}||A_{iR}^{(n+1)}B_{jR}^{(n+1)}C_{kR} - T_{ijk}||_F^2.
\end{aligned} \tag{5.9}
$$

Here n stands for the number of iterations. To obtain programmable equations, we can write the explicit form for each matrix for the $(n+1)$-th iteration. Here we take advantage of the fact that the argument minimum appearing in each equation is equivalent to finding the roots. For the matrix A we may therefore write:

$$
0 = \frac{\partial}{\partial A_{iR}}||A_{iR}B_{jR}^{(n)}C_{kR}^{(n)} - T_{ijk}||_F^2 = -2B_{jR}^{(n)}C_{kR}^{(n)}T_{ijk} + 2A_{iR}\left(B_{jR}^{(n)}C_{kR}^{(n)}\right)^2. \tag{5.10}
$$

With this, we are able to obtain the final equations for the implementation

$$
\begin{aligned}
A_{iR}^{(n+1)} &= T_{ijk}B^{jS}C^{kS}\left(G_{SR}^{+}\right)_A, \\
\left(G_{RS}\right)_A &= B^{jS}C^{kS}B_{jR}C_{kR}.
\end{aligned} \tag{5.11}
$$

For brevity, we are omitting the iteration specification. G_{SR}^{+} is the Moore-Penrose pseudo inverse [40, 41] of the Gramian matrix G_{RS}. For the matrices B and C the derivation is similar. In the final, we get:

$$
\begin{aligned}
B_{jR}^{(n+1)} &= T_{ijk}A^{iS}C^{kS}\left(G_{SR}^{+}\right)_B, \\
\left(G_{RS}\right)_B &= A^{iS}C^{kS}A_{iR}C_{kR}, \\
C_{kR}^{(n+1)} &= T_{ijk}A^{iS}B^{jS}\left(G_{SR}^{+}\right)_C, \\
\left(G_{RS}\right)_C &= A^{iS}B^{jS}A_{iR}B_{jR}.
\end{aligned} \tag{5.12}
$$

Returning to our optimization problem, the computationally most expensive steps in Eqs. 5.11 and 5.12 are the calculation of G^{+} scaling with $\mathcal{O}(N_R^3)$ and the contraction of T_{ijk} with another matrix scaling as $\mathcal{O}(N_{SO}^2\tilde{N}_\sigma N_R)$.

Albeit the ALS algorithm guarantees the convergence of the product of the three second-order tensors A, B and C to the third-order tensor T, we additionally consider the RALS algorithm [42]. The reason is that there may be several similar minima for A in different regions. Then, updating B and C may change the order of minima of A and hence the choice of the optimal A between iterations which significantly slows down the convergence. Consequently, these fluctuations also occur when B and C are optimized. Such behavior of slow convergence of the ALS algorithm is referred to as swamping [43]. To get a faster convergence to a region for which each factor is globally minimized, an additional regularization is introduced. The extension to the RALS algorithm is justified by the fact that considerably faster convergence has been observed for real systems, such as for water adsorption on the hexagonal boron nitride monolayer [36]. For this ansatz, the minimization problem reads

$$
\begin{aligned}
A^{(n+1)} &= \underset{A}{argmin} ||A_{iR} B^{(n)}_{jR} C^{(n)}_{kR} - T_{ijk}||^2_F + \lambda^{(n)}_A ||A_{iR} - A^{(n)}_{iR}||^2_F, \\
B^{(n+1)} &= \underset{B}{argmin} ||A^{(n+1)}_{iR} B_{jR} C^{(n)}_{kR} - T_{ijk}||^2_F + \lambda^{(n)}_B ||B_{jR} - B^{(n)}_{jR}||^2_F, \\
C^{(n+1)} &= \underset{C}{argmin} ||A^{(n+1)}_{iR} B^{(n+1)}_{jR} C_{kR} - T_{ijk}||^2_F + \lambda^{(n)}_C ||C_{kR} - C^{(n)}_{kR}||^2_F,
\end{aligned}
\tag{5.13}
$$

where λ denotes the regularization parameter.
The explicit solution is found by solving the minimization problem from Eq. 5.13 with the same approach as in Eq. 5.10 for the ALS method. The final equations are

$$
\begin{aligned}
A^{(n+1)}_{iR} &= \left[T_{ijk} B^{jS} C^{kS} + \lambda^{(n)}_A A^{(n)}_{iS} \right] \left(G^{+}_{SR} \right)_A , \\
\left(G_{RS} \right)_A &= B^{jS} C^{kS} B_{jR} C_{kR} + \lambda^{(n)}_A \delta_{RS}, \\
B^{(n+1)}_{jR} &= \left[T_{ijk} A^{iS} C^{kS} + \lambda^{(n)}_B B^{(n)}_{jS} \right] \left(G^{+}_{SR} \right)_B , \\
\left(G_{RS} \right)_B &= A^{iS} C^{kS} A_{iR} C_{kR} + \lambda^{(n)}_B \delta_{RS}, \\
C^{(n+1)}_{kR} &= \left[T_{ijk} A^{iS} B^{jS} + \lambda^{(n)}_C C^{(n)}_{kS} \right] \left(G^{+}_{SR} \right)_C , \\
\left(G_{RS} \right)_C &= A^{iS} B^{jS} A_{iR} B_{jR} + \lambda^{(n)}_C \delta_{RS}.
\end{aligned}
\tag{5.14}
$$

One central point in the RALS method is to determine the value of the regularization parameter λ. On one hand, low values might not influence the behavior of the optimization algorithm enough and lead to possible swamping like in the ALS method. On the other hand, too large values slow down the convergence significantly. In the following, we make two assumptions as in Ref. [36], namely, the fit quality of the entire problem, as well as the local change of the fit quality for each individual factor, do not change much between two iterations. This allows us to define the so-called relative step size for each factor

$$
s^{(n)}_X \equiv \frac{||X^{(n)} - X^{(n-1)}||_F}{||X^{(n)}||_F}. \tag{5.15}
$$

For the sake of brevity, we define the relative step size by an arbitrary tensor X. Additionally, the relative step size in each iteration is set to be approximately as large as a defined swamping threshold s_0. This estimation leads to the so-called estimated regularization parameter

$$
\hat{\lambda}^{(n)}_X \equiv \lambda^{(n-1)}_X \frac{s^{2}_X{}^{(n)}}{s_0}, \tag{5.16}
$$

allowing to change the regularization between iterations according to the deviation of the relative step size to the swamping threshold. Finally, the regularization parameter λ is given by a mixing of the estimated regularization parameters and λ itself from the previous iteration

$$\lambda_X^{(n)} \equiv \alpha \hat{\lambda}_X^{(n)} + (1-\alpha)\lambda_X^{(n-1)}. \tag{5.17}$$

Here the parameter α describes the weight of the estimated regularization parameter $\hat{\lambda}_X^{(n)}$ and the regularization parameter $\lambda_X^{(n-1)}$ from the previous iteration for the new $\lambda_X^{(n)}$.

In our code, we implement two main functions. The first one computes the singular value decomposition of a given tensor of rank four and returns the truncated object. The second function executes the canonical polyadic decomposition, where firstly the Coulomb tensor, given from the input, is again SVD truncated and decomposed. The returned U and Σ tensors are used to build the optimized Coulomb vertex Γ (s. Eq. 5.4). Then, a tensor rank decomposition in terms of CPD of the optimized Coulomb vertex according to Eq. 5.6 is performed. The latter optimization problem is consequently solved by means of the ALS or RALS algorithm. To get an insight into the convergence behavior of our system both algorithms are implemented. We formulate the convergence criterion of the underlying optimization problem as

$$\frac{||ABC - T||_F}{||T||_F}. \tag{5.18}$$

This is the Frobenius norm of the difference between the matrices A, B and C and the tensor T normalized with respect to the Frobenius norm of T. For brevity, we call Eq. 5.18 the normalized Frobenius norm. Since ALS and RALS are iterative algorithms, we need to make initial guesses for the first iteration. While the three matrices $A^{(0)}$, $B^{(0)}$ and $C^{(0)}$ are initialized as random matrices, the swamping threshold s_0 and the initial guess for the regularization parameter $\lambda^{(0)}$ are set to 1. The mixing parameter α is set to 0.8. In our convergence analysis of the ALS and RALS algorithm, we modify also the values of s_0 and $\lambda^{(0)}$ as well as α to get insights of the system behavior.
After performing the low-rank approximation with the necessary input data, the coupled cluster energies are analyzed. Our main focus lies in the analysis of the accuracy of the obtained energies using the low-rank approximation with respect to the order of truncation and decomposition. Thereby, we explore the extent to which utilizing low-rank approximated Coulomb integrals still give accurate results without losing too much computational efficiency.

In summary, we present two low-rank factorization techniques. The SVD provides efficient treatment of tensors by reducing storage requirements as well as computation time. This is accomplished with the parameter N_σ giving the order of truncation. The second method discussed is the CPD, where we go a step beyond by applying an additional tensor rank decomposition. In this case, the number of vertex indices N_R is the parameter determining the trade-off between computational efficiency and accuracy.

6 Results and Discussion

In this chapter we present the results of the introduced methods for electronic structure calculations. We analyze the Hartree-Fock algorithm and investigate its convergence with various system parameters. The improvements in terms of electron correlations using the MP2 and CCSD methods are investigated. In the final part, we give the main results of the calculations with the EOM-CCSD algorithm to obtain the excited energy spectrum and analyze the impact of low-rank approximation techniques. Besides, we discuss convergence criteria regarding our implemented tensor-rank decomposition algorithms.
All computations are performed with our model Hamiltonian according to Eq. 2.6. Further, if not specified, all results are displayed in atomic units. We introduce the short forms HF and HO for Hartree-Fock and harmonic oscillator, respectively.

6.1 Hartree-Fock Theory

We obtain the solutions of the Hartree-Fock equations using the introduced SCF algorithm (s. Sec. 3.5). The first natural step is to investigate the change of the Hartree-Fock energy including two-electron interactions compared to the non-interacting harmonic oscillator problem. The solution of the latter is well-known and may be represented in the occupation number representation

$$E_n^{HO} = \sum_{i=1}^{2} \epsilon_n^{HO}(i) = \omega\left(n + \frac{1}{2}\right) + \omega\left(n + \frac{1}{2}\right) = \omega\,(2n+1)\,.$$

Since Hartree-Fock theory solves for the ground state energy of the system including two-electron interactions, most interest goes into the comparisons with $E_0^{HO} = w$. The comparison with the harmonic oscillator turns out to be very useful, mainly to estimate the influence of the averaged electron-electron interaction described by the Coulomb and exchange term. In addition, the ground state energy of the harmonic oscillator is used as a reference when it comes to changing the system parameters and analyzing the convergence behavior of the algorithm.

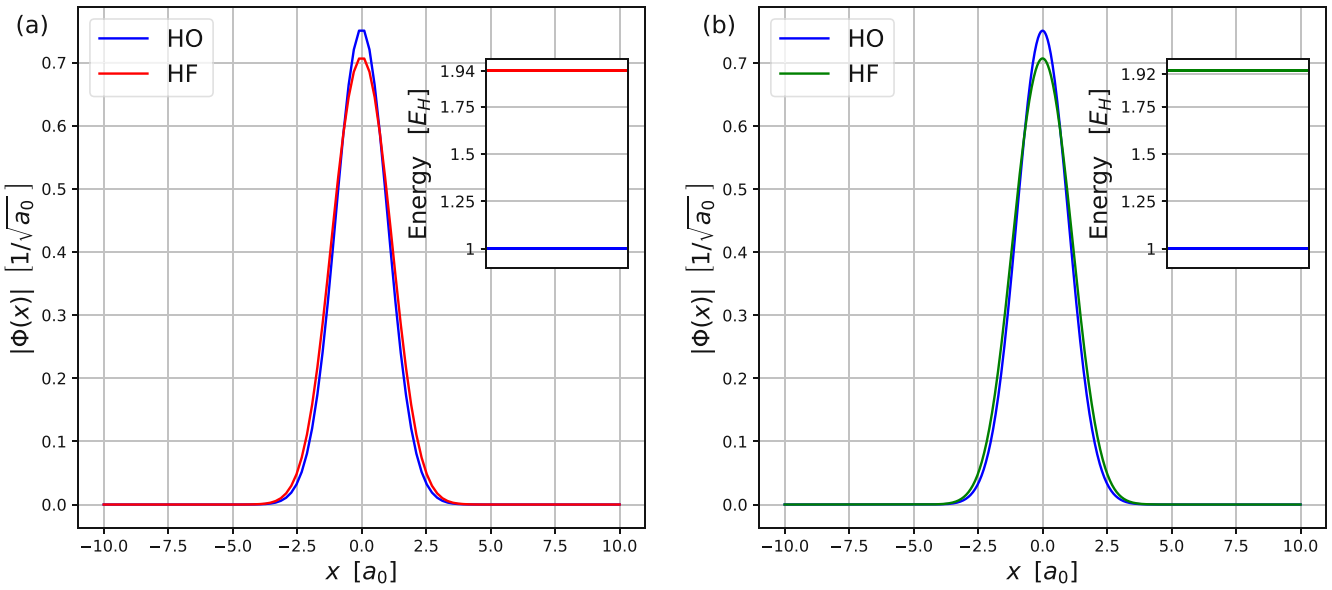

Figure 2: Comparison between the HO and the HF ground state wave function and ground state energy. The calculations are performed with $\omega = 1\tau_0^{-1}$ and $\sigma = 0.5$ and with (a) $N_{grid} = 100$ and (b) $N_{grid} = 5000$.

Fig. 2 shows the differences in energies and wave functions between the non-interacting (HO) and interacting (HF) calculations. The first thing to recognize is that the ground state wave functions have a Gaussian-type shape. In both plots the amplitude of the HF ground state wave function is smaller compared to the HO one. In return, we find a slightly greater variance for the interacting model. Furthermore, we would like to highlight two important points. The HF wave function does not change significantly with an increasing number of grid points, however, the ground state energy E^{HF} does. This follows from comparing the Hartree-Fock energies obtained from 100 grid points (red) and 5000 grid points (green). The HF energy is expected to converge with larger number of grid points N_{grid}. As expected, the ground state energy of the harmonic oscillator (blue) is less dependent on the grid size.

After the introductory comparison of the non-interacting and the interacting problem, we go into more detail about the performed HF calculations. Thereby we want to take a closer look at the dependence of E^{HF} on the number of grid points and investigate the influence of the two system parameters ω and σ.

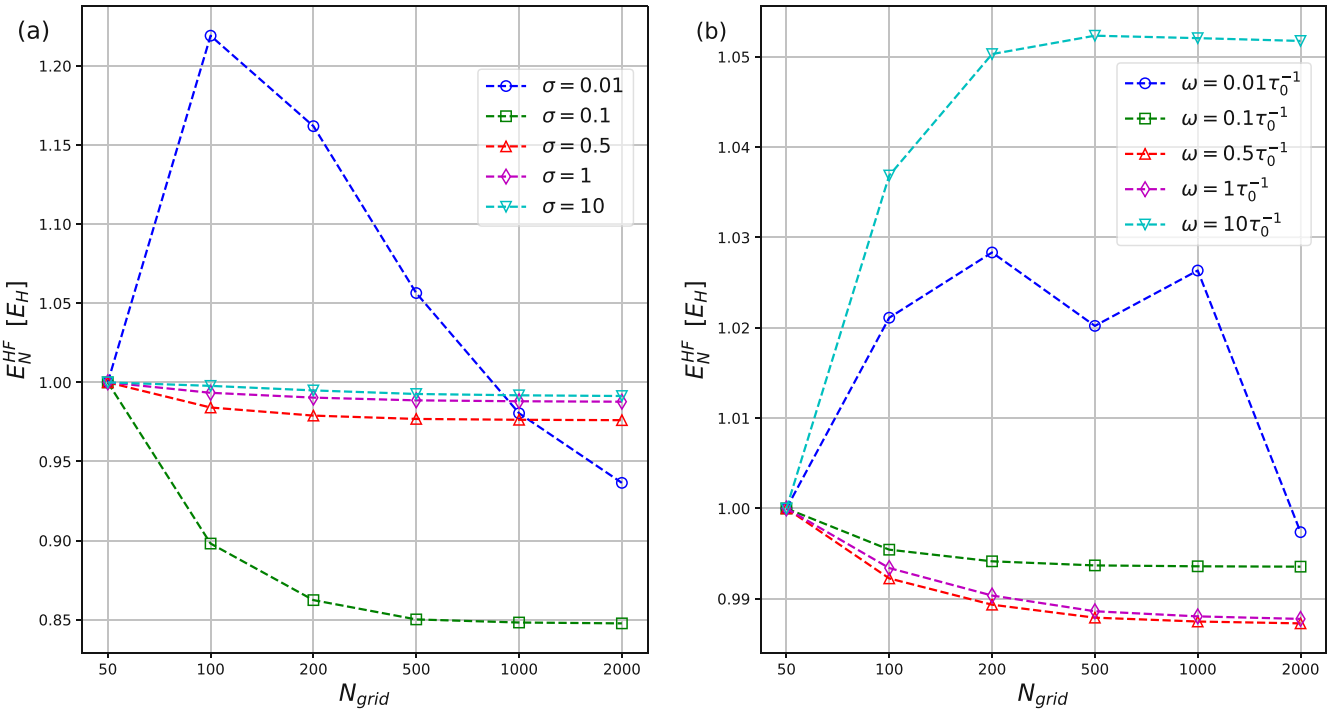

Figure 3: Dependence of the normalized Hartree-Fock energy on the number of grid points N_{grid}. (a) The energy is calculated with several σ values for $\omega = 1\tau_0^{-1}$. (b) Different ω values are used to compute the energy with $\sigma = 1$. The energies are normalized with respect to $E^{HF}(N_{grid} = 50)$ (s. Eq. 6.1). For small ω values we adapt the grid to $x \in [-100a_0, 100a_0]$ in order to obtain correct results.

In Fig. 3 we introduced the following normalization:

$$E^{HF}(N_{grid}) \to E_N^{HF}(N_{grid}) = \frac{E^{HF}(N_{grid})}{E^{HF}(50)}, \tag{6.1}$$

with the reasoning that we are at this point only interested in the qualitative behavior of our model with respect to the system parameters. This gives not only a better understanding of our model but also allows

us to identify the range of parameters most suitable for the upcoming coupled cluster calculations.
We first discuss Fig. 3a showing $E^{HF}(N_{grid})$ for different σ values. For small values like $\sigma = 0.01$, $E^{HF}(N_{grid})$ converges significantly slower compared to larger values of σ. Even when increasing N_{grid} from 1000 to 2000 no convergence is observed. This suggests that for $\sigma = 0.01$ more grid points would be required to obtain relatively accurate results because in general convergence of the energy with the number of grid points is expected. However, such a large number of grid points significantly increases the numerical cost due to the most demanding operations in our SCF algorithm, the computations of the Coulomb and exchange term, which scale with $\mathcal{O}(N_{grid}^2)$. We conclude that very small values for σ result in fluctuations for small N_{grid} and slow convergence.
For the remaining σ plots, we observe a faster convergence. However, for $\sigma = 0.1$ still a relatively slow convergence is shown which can be attributed to the more amplified Coulomb potential. Thus, small σ values still lead to slow convergence. Looking at the form of the Coulomb potential in our Hamiltonian, we recall the $1/\sigma$ dependence. Thus, we may also argue that for small σ the Coulomb term is much larger than the kinetic and harmonic potential one which is the reason for the strong deviations of the energy compared to larger σ. On the other hand, for large σ values only small deviations are noticeable. Here we argue similarly, namely the influence of the interaction term is vanishingly small. In summary, Fig. 3a suggests reasonable results for σ values in the range of $[0.1, 1]$ as well as a number of grid points equal or greater than 500.
Fig. 3b shows a similar behavior of the energy as a function of the number of grid points for $\omega \in [0.01\tau_0^{-1}, 1\tau_0^{-1}]$. We start with the interpretation for the smallest value $\omega = 0.01\tau_0^{-1}$. Again, an increase in the energy with N_{grid} is recognizable. In this case, however, we find no convergence with the number of grid points. We explain this behavior by investigating the Hamiltonian (s. Eq. 2.6). For small ω, the harmonic potential term is strongly suppressed. This implies that the functional of the energy is dominated by the $1/x$ Coulomb-like term. Numerically, small numbers of grid points, are not able to fully resolve this behavior, similarly to the $\sigma = 0.01$ plot of the left graph. In this case, however, this is noticeable up to $N_{grid} = 1000$ due to the x^2-dependence of the harmonic potential and the wider grid $[-100a_0, 100a_0]$ used. Only for $N_{grid} \geq 2000$ the suppression of the harmonic potential is captured correctly and consequently convergence to lower energy values is expected. For $\omega = 0.1\tau_0^{-1}$ smaller changes in the energy are seen due to the consequences of the wider grid. Lastly, $\omega = 10\tau_0^{-1}$ results in a big weight of the harmonic oscillator potential as well as a strong electron correlation due to the narrowed harmonic potential. Thus, an increase with N_{grid} is expected up to a certain number of grid points. In our case it is 500, where we find enough supporting points to fully represent the potential and to achieve convergence.

In conclusion, we find that the computed HF energies are difficult to converge for small and large values for σ and ω, respectively. Up to this point we were only concerned about the convergence of the energy as a function of the number of grid points. We further want to investigate the convergence by means of our SCF algorithm. We analyze the two numerical convergence conditions, namely the convergence of the energy and the reduced density operator (s. Eq. 3.17 and Eq. 3.18). We find both convergence criteria are fulfilled after a few steps for $\sigma, \omega \in [0.1, 1]$. Thus, we are left investigating more extreme values. We present one particular example with $\omega = 1\tau_0^{-1}$ and $\sigma = 0.01$ while using 500 grid points (s. Fig 3).

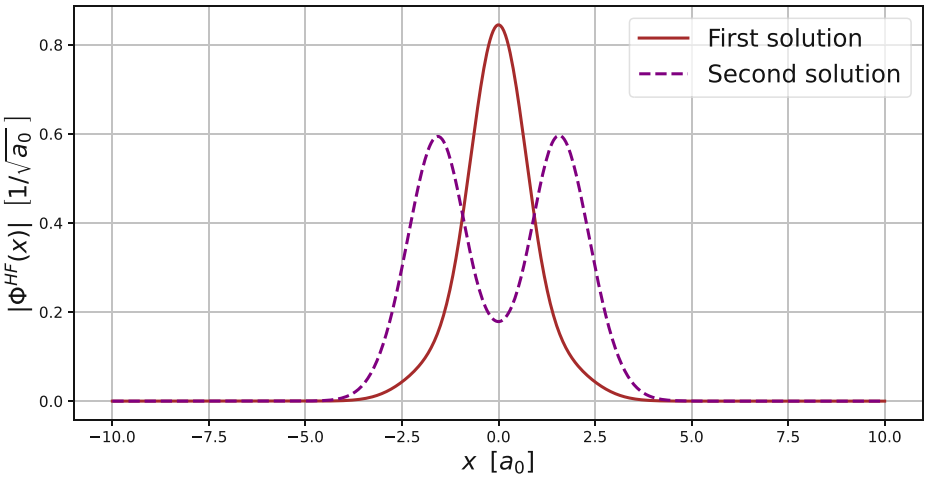

Figure 4: Degeneracy of the HF ground state for the parameters $\omega = 1\tau_0^{-1}$ and $\sigma = 0.01$. The two possible ground state wave functions using 500 grid points are shown.

The numerical results show that for this example convergence is not reached. However, it is important to point out, that while E^{HF} is converged with sufficient accuracy, the reduced density matrix, which as mentioned in Sec. 3.5, fails to converge using the employed SCF algorithm. For this example, Eq. 3.18 converges to a constant value greater than our convergence criterion $\epsilon < 10^{-6}$. A constant change of the reduced density matrix can be interpreted in a way that the algorithm jumps back and forth between two solutions, which have non-vanishing off-diagonal Fock matrix elements. These convergence difficulties together with the slow convergence of the energy with N_{grid} have led us to conclude that this parameter range will not be suitable for the purpose of the present study. Fig. 4 shows the two alternating states.
In consequence, it is reasonable to work within the range of values $[0.1, 1]$ to obtain fast convergence, not only with respect to the number of grid points N_{grid} (s. Fig 3), but also with respect to the convergence condition from our SCF algorithm itself. Through empirical investigations, it turns out that the value set $\sigma, \omega = \{0.5, 1\tau_0^{-1}\}$ as well as $N_{grid} = 500$ are good choices. We will justify the specific choice of σ and ω in the next chapter on coupled cluster theory in more detail.

We conclude our analysis of the Hartree-Fock theory by investigating the influence of the two-electron interaction term. We define the parameter $0 \leq g \leq 1$ describing the interaction strength of the system. The modified Hamiltonian consequently reads:

$$\hat{H} = \sum_{i=1}^{2} \left[-\frac{1}{2}\frac{\partial^2}{\partial x_i^2} + \frac{1}{2}\omega^2 x_i^2 \right] + \frac{g}{|x_1 - x_2| + \sigma}.$$

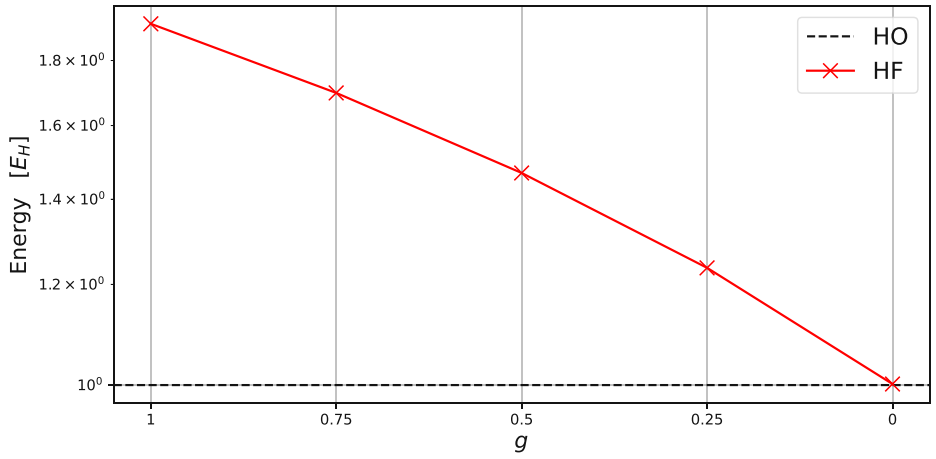

Figure 5: HF energy as a function of the interaction strength g. For comparison, we additionally plot the HO ground state energy. We choose $\omega = 1\tau_0^{-1}$, $\sigma = 0.5$ and 500 grid points.

As the interaction strength is reduced linearly, the Hartree-Fock energy shows a polynomial decrease (s. Fig. 5). This implies that with decreasing g, the interaction strength and thus the repulsion between the electrons becomes weaker leading to a reduction in the overall energy of the system. The non-linear behavior of the energy can be attributed to the Coulomb-like potential term. The Fock matrix is set up with the Coulomb and exchange matrices which are depending on g. In each SCF iteration, the wave function is updated by diagonalization of the Fock matrix. This wave function is then used to compute the ground state energy as well as the new reduced density matrix which in turn flows back into the calculation of the J and K. Consequently, no linear behavior can be expected for $0 \leq g \leq 1$.

6.2 Post-Hartree-Fock Theories

All post-Hartree-Fock methods require two precomputed quantities, the orbital energies ϵ_p and the antisymmetrized Coulomb tensor V^{-pq}_{rs}. The orbital energies are obtained from the SCF algorithm and the antisymmetrized Coulomb tensor is computed using the introduced trapezoidal integration method. In passing, we mention that the calculation of V^{pq}_{rs} by integration techniques can partly be verified such that:

$$\langle \Phi^{HF} | 2\hat{J} - \hat{K} | \Phi^{HF} \rangle = \int dx_1 \int dx_2 \; \Phi_0^*(x_1)\Phi_0^*(x_2) \frac{1}{|x_1 - x_2| + \sigma} \Phi_0(x_1)\Phi_0(x_2).$$

For the specific case with the set of generic indices $\{p = 1, q = 1, r = 1, s = 1\}$, the two-electron integral gives the same result as the expectation value $\langle 2\hat{J} - \hat{K} \rangle$ representing the interaction term in the Hartree-Fock equations. Although the integral can only be verified with this combination of spin orbital indices, it is an indication of the validity of our integration method.

Next, we analyze the energy dependence of our system with different parameters similar to the discussion on the Hartree-Fock theory. Thereby we especially focus on the influence of ω and σ while taking a fixed number of 500 grid points which has proven to be a favorable balance between accuracy and computational cost (s. Fig. 3).

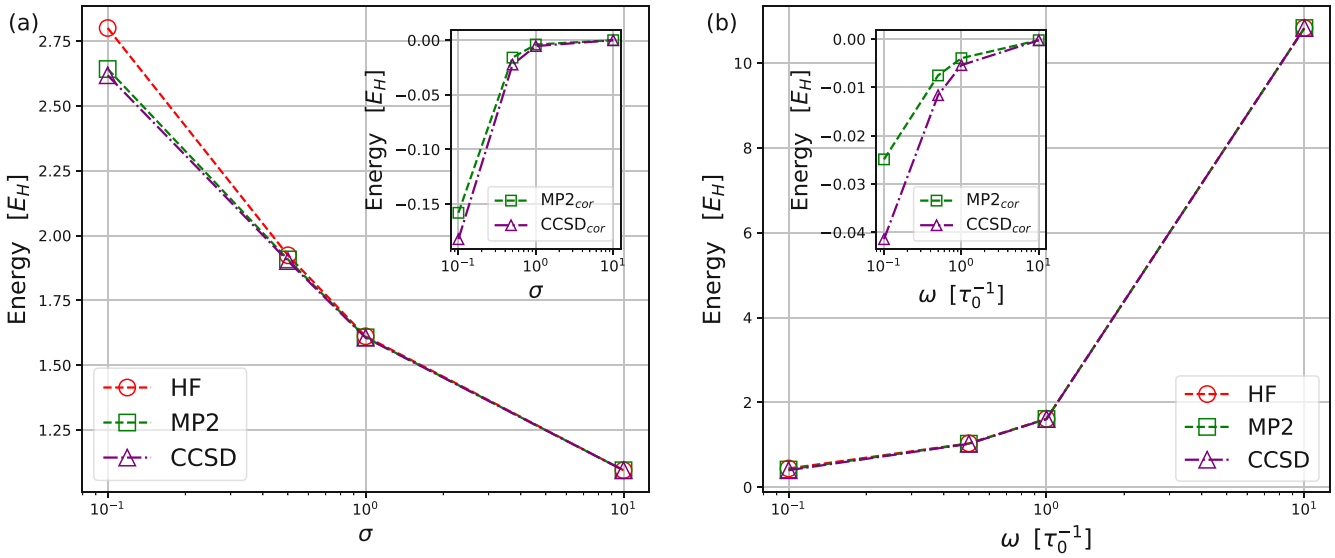

Figure 6: Comparison of the HF, MP2 and CCSD energies with variation of (a) σ and (b) ω. In addition, we plot the MP2 and CCSD correlation energies, respectively. The calculations are performed with 500 grid points and 4 spin orbitals.

First of all, the comparison of the two graphs in Fig. 6 shows that our system is much more sensitive to changes in σ than for ω, i.e. the energies of the Hartree-Fock and post-Hartree-Fock methods differ most for very small σ values (s. Fig. 6a). This means at the same time that the singularity behavior of the Coulomb-like potential is favored. In other words, small σ enhance the electron-electron interaction term in our Hamiltonian such that the kinetic and harmonic potential parts become less important. The difference between the Hartree-Fock and post-Hartree-Fock energies becomes vanishingly small for larger σ, implying a smaller and smaller electron-electron correlation. This is also in agreement with the MP2 and CCSD correlation energies shown in the subplot of Fig. 6a where the correlation energies converge to zero as σ increases. The difference between MP2 and CCSD is manifested best for small σ, where CCSD leads to lower energies. This is consistent with the order of perturbation applied in MP2 compared to CCSD.
Fig. 6b shows the expected behavior of the energies with increasing ω. Bigger values imply a greater contribution of the harmonic potential to the energy. Furthermore, in the subplot it is shown that the correlation energy is again converging to zero as ω increases and the Coulomb potential is suppressed. In contrast to the correlation energies of Fig. 6a, the values of ω do not influence the total energy that much. We conclude that a reasonable choice of σ and ω has to be determined by the behavior of the algorithm. Thereby, we are more flexible in designating a value for ω. The value for σ, however, has to be chosen more carefully. The second criterion for the choice of these values is that our system, although it is just an artificial model, should represent the same behavior as a real physical system. Commonly, the correlation energies are in the percentage range of the total energies [44]. For our system we find that the value pair $\{\sigma, \omega\} = \{0.5, 1\tau_0^{-1}\}$ is a valid candidate. With this not only a smooth convergence of our algorithm is guaranteed but also the behavior of real physical models can be simulated.

After the determination of the system-specific parameters we investigate the model-independent parameters N_{grid} and N_{SO}, representing the number of grid points and the number of spin orbitals, respectively.

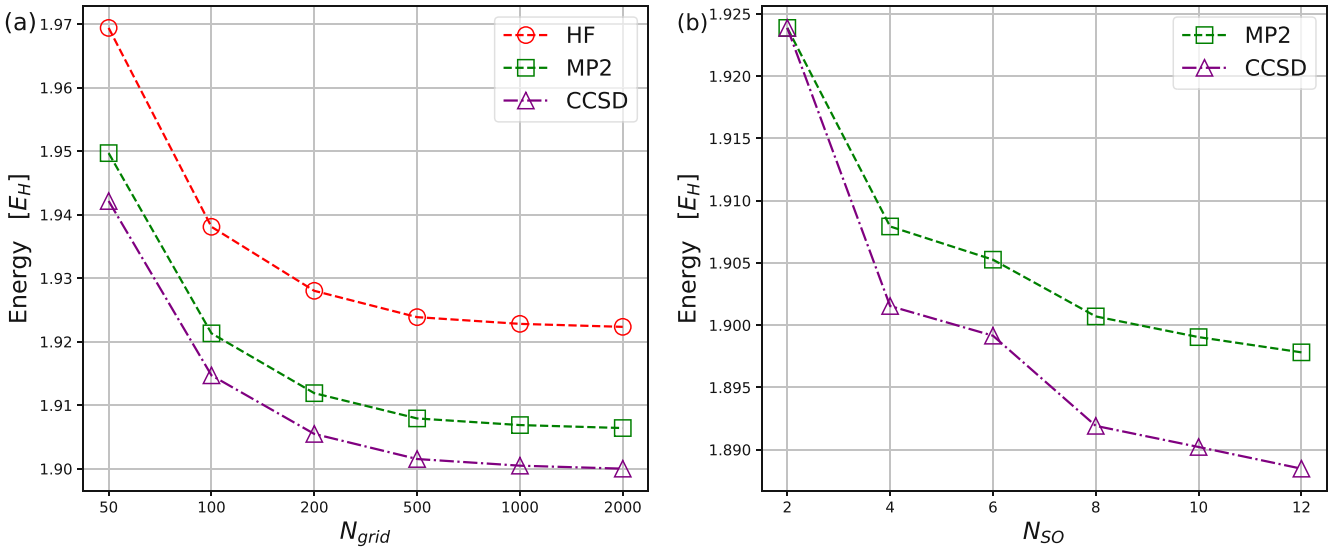

Figure 7: (a) Energy dependence on the number of grid points N_{grid} using 4 spin orbitals for the HF, MP2 and CCSD methods. (b) Energy dependence of the MP2 and CCSD methods on the number of spin oribtals N_{SO} for 500 grid points. In both cases, we use $\omega = 1\tau_0^{-1}$ and $\sigma = 0.5$.

Fig. 7a shows the energy dependence of the Hartree-Fock and both post-Hartree-Fock methods on the number of grid points. As N_{grid} increases all three energies converge to a certain value. This behavior is expected since the energy should only depend on N_{grid} up to a rather low number of mesh points and then be independent. Further, we see that the improvements in accuracy gained from the post-Hartree-Fock methods result in a reduction of energy. It is expected that the CCSD corrections provide more accurate results because the MP2 correction is just the first iteration of the CCSD algorithm. Furthermore, CCSD is exact for two-electron systems.
Fig. 7b gives the energy dependence on the number of spin orbitals. Since Hartree-Fock assumes the lowest energy configuration where the two electrons occupy the lowest spatial orbital, the first value corresponds exactly to E^{HF}. Thus, also with the MP2 and CCSD methods no improvements can be achieved when considering only 2 spin orbitals. The more spin orbitals are involved, the better the accuracy. A larger number of spin orbitals means at the same time that more excitations are considered and with this, we are able to capture the real behavior of our system better and better. In passing we mention that formally CCSD is exact for systems with two electrons since higher excitations do not contribute. Although a higher N_{SO} provides more accurate outcomes, we stick for the rest of this chapter to $N_{SO} = 4$. This allows a lucid and concise discussion of the EOM-CCSD results. Further, if not already precomputed, the computational cost is kept within limits since the EOM-CCSD algorithm scales with $\mathcal{O}(N_{SO}^6)$.

The last part of this section is dedicated to the EOM-CCSD calculations. Recall that our algorithm relies on the direct diagonalization of the EOM-CCSD matrix. Depending on the considered number of spin orbitals a different number of energies is obtained. As in the discussion of Fig. 7b, we restrict ourselves to four spin orbitals, also for the reasoning that only the lowest few excited states are of interest. The EOM-CCSD matrix has the dimension of $(N_S + N_D) \times (N_S + N_D)$ with $N_S = 4$ and $N_D = 16$. To justify this, we look at the matrix elements of the singles-singles block of the EOM-CCSD matrix $\langle \Psi_i^a | \bar{H} | \Psi_k^c \rangle$. Thereby, the orbitals i and a implicitly account for the two possible spin configurations. Thus, there are a total of 2^2 possible single excitations. The same considerations apply to the doubly excited matrix elements $\langle \Psi_{ij}^{ab} | \bar{H} | \Psi_{kl}^{cd} \rangle$, where we find the 2^4 possible ways of two electrons to be excited according to the possible

combinations of the indices i, j and a, b. Consequently, the EOM-CCSD matrix with dimension 20×20 yields twenty energies by its diagonalization.

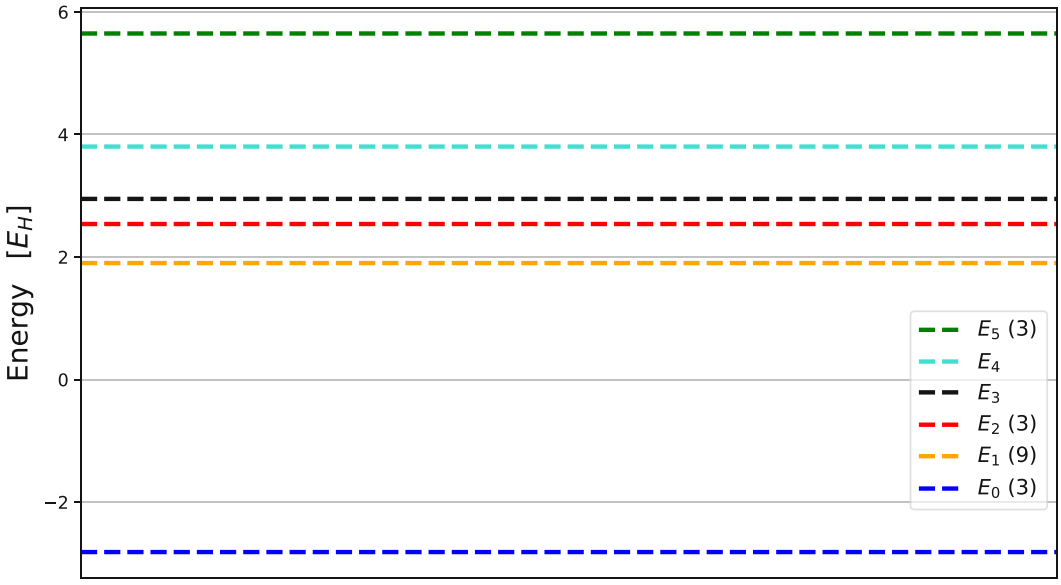

Figure 8: Energy spectrum obtained by the EOM-CCSD algorithm for 4 spin orbitals. We are using $\omega = 1\tau_0^{-1}$, $\sigma = 0.5$ and $N_{grid} = 500$. The numbers in the brackets denote the degree of degeneracy.

In Fig. 8 we plot the energies resulting from the EOM-CCSD algorithm using four spin orbitals. The corresponding values are shown in Tab. 1. Of the total of twenty energies, only six are distinctive. These degeneracies stem from the possible spin configurations of a system with two electrons

$$\begin{aligned} S = 1: &\qquad M_S = -1, 0, 1\,, \\ S = 0: &\qquad M_S = 0, \end{aligned}$$

where S denotes the total spin quantum number and M_S the magnetic quantum number along a selected axis. States attributed with $S = 1$ are the so-called triplet states and those with $S = 0$ are the singlet states. In Fig. 8, e.g. , the ground state E_0 is a triplet state with $S = 1$.

Table 1: Obtained energy values for the excited energy spectrum by the EOM-CCSD algorithm, rounded to two decimal places.

i	E_i $[E_H]$
0	-2.81
1	1.90
2	2.54
3	2.95
4	3.80
5	5.65

At this point, however, we need to further distinguish between physically possible and forbidden states. The eigenvalues obtained by our EOM-CCSD code are built from the fact that each index associated with an electron introduces a factor of 2 in terms of the possible spin orientations. But, we did not take into account that for electrons the Pauli exclusion principle and the demand of antisymmetry must hold too. That being said, the EOM-CCSD code also produces non-fermionic states, where e.g. two particles with the same spin occupy the same spatial orbital. Despite this non-physicality, we are eager to look also at the eigenvalues of the non-fermionic eigenstates to learn more about the low-rank factorization accuracy in the upcoming section. The physically relevant states, however, are shown in Fig. 9.

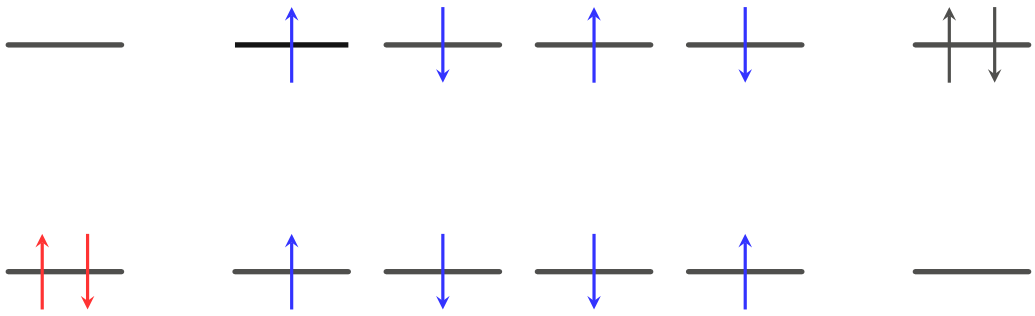

Figure 9: Six allowed electron configurations for our model considering two spatial orbitals: Hartree-Fock ground state singlet (red), triplet and singlet state arising from single excitations (blue) and singlet resulting for the double excitation (black).

Fig. 9 shows on one hand the Hartree-Fock electron configuration describing the lowest spatial orbital occupied by two spins with opposite orientations. On the other hand, four possible scenarios for an electron to be excited in the orbital of higher energy arise. In Fig. 9, the first two configurations seen in blue describe the triplet states for $M_S = 1$ and $M_S = -1$. The other two describe the triplet state with $M_S = 0$ but also the singlet state with the same magnetic quantum number. Both form a superposition of these two configurations, where the triplet state is symmetric, and the singlet state is antisymmetric. The last configuration represents the only possibility for a doubly excited state with a total spin of zero.
To identify the allowed states from our EOM-CCSD outcomes, we do a further analysis of the eigenvectors R and their components R_1 and R_2 from the EOM-CCSD equation (s. Eq. 4.35). The four possible single excited states are found by demanding $|R_1|^2 \approx 1$, i.e. states with almost zero R_2 contribution. The doubly excited state is found by applying the Pauli principle and the antisymmetry condition to every R_2 vector. We find that the lowest excited energy is the triplet with $E = 2.54E_H$. The singlet arising from the single excited state has an energy of $E = 2.95E_H$ and the singlet from the double excitation occupies the highest state of these five with $E = 3.80E_H$. The CCSD ground state energy $E = 1.90E_H$ remains unchanged since our EOM-CCSD method only considers the subspace of single and double excitations. It is worth noting that with a very small σ, the contribution of the electron interaction can be amplified to such an extent that the triplet state becomes energetically more favorable compared to the ground state energy of CCSD. This could also be the case with the inclusion of more spin orbitals, where the size of the basis set is expanded allowing for a more accurate description of the electronic wave function.

Table 2: Non-fermionic ground state energy E_0 of the EOM-CCSD method for different numbers of spin orbitals N_{SO}.

N_{SO}	E_0 $[E_H]$
4	-2.81
6	-4.78
8	-6.69
10	-8.61
12	-10.54

Tab. 2 shows the change of the non-fermionic ground state energy E_0 with the inclusion of more and more spin orbitals. We conclude that E_0 decreases linearly by approximately the same magnitude as the CCSD ground state energy for every additional two spin orbitals. This observation suggests that the additional pair of spin orbitals captures a significant amount of the correlation energy. However, for an increasing number of spin orbitals the ground state energy is expected to converge. The proper convergence behavior may be explained in terms of the Kato cusp condition [45] which states that the electronic wave function must satisfy a set of cusp conditions that prescribe the derivative discontinuity at the coalescence point of two electrons. At zero interparticle separation, the Coulomb interaction diverges, necessitating the kinetic energy also to diverge such that the sum of both is finite. The latter divergence is displayed as a cusp in the electronic wave function at the point of collision. For a three-dimensional system, the Kato cusp condition for two electrons reads:

$$\left.\frac{\partial \hat{\Psi}}{\partial \vec{r}_{12}}\right|_{\vec{r}_{12}=0} = \frac{1}{2}\Psi(\vec{r}_{12}=0),$$

where $\hat{\Psi}$ is Ψ averaged over a small sphere about the singularity. In our case, we attained the electronic wave function with single-particle orbitals giving very accurate energies. However, such finite linear combinations are not capable of representing points of discontinuity in the wave function. In conclusion, smooth basis functions slow down the rapid convergence due to the misrepresentation of the cusp. Due to the high computational demand, it is not easy to predict the convergence of the system under study. We perform our EOM-CCSD calculations exclusively for the N_{SO} in Tab. 2. It is estimated that with around 100 spin orbitals convergence might be reached. However, further investigations are required to confirm the exact number of spin orbitals.

6.3 Low-Rank Approximation

In this section, the focus is on solving the CCSD and especially the EOM-CCSD algorithm by LRA methods. We put special focus on the thresholds where the description with the LRA-fed quantities leads to unacceptable deviations. This gives us an estimate of how coarse our approximations can be made to still obtain accurate results. Low-rank approximations approximate the Coulomb tensor V_{rs}^{pq} by tensors of lower rank. In the following, we primarily focus on the physical results rather than going into the details of constructing the approximated Coulomb tensors. However, we investigate the convergence of the RALS and ALS algorithms. This is essential since particularly the RALS algorithm relies on system-specific parameters which influence its convergence which in turn impacts the computation time and the accuracy of the results.

As a first step, we consider the CCSD algorithm, which is solved with the approximated Coulomb tensor $\tilde{V}_{rs}^{pq}$, obtained by the singular value decomposition.

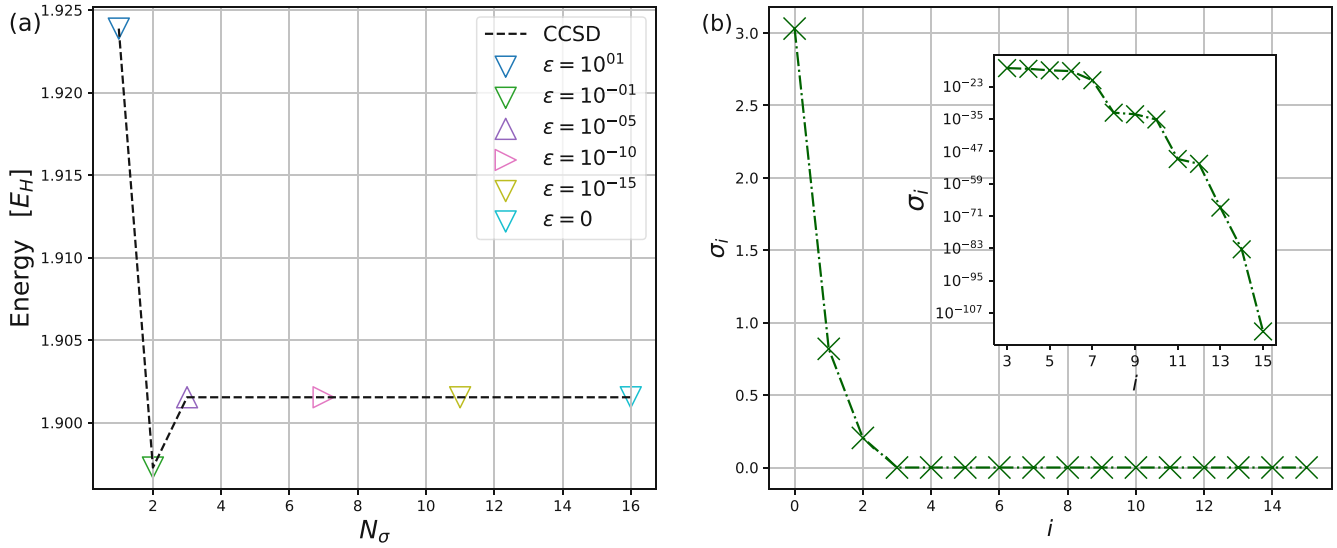

Figure 10: (a) Computation of the CCSD energy with the SVD truncated Coulomb tensor for several numbers of singular values N_σ or equivalently for several accuracies ϵ. (b) Singular values of the Coulomb tensor V_Q^P (s. Eq. 5.1). We use $\omega = 1\tau_0^{-1}$, $\sigma = 0.5$, $N_{grid} = 500$ and 4 spin orbitals.

Fig. 10a shows the ground state energy resulting from our CCSD algorithm computed with the approximated Coulomb tensor. Thereby, different accuracies ϵ are chosen which determine how many singular values are kept (s. Eq. 5.2). The number of total singular values available is determined by the size of the SVD decomposed quantities according to Eq. 5.1 yielding $N_\sigma = pr = qs$. Thus, choosing four spin orbitals implies that the maximum number of singular values is $4 \cdot 4 = 16$. Fig. 10a indicates that the truncation of the original Coulomb tensor up to 3 singular values ($\tilde{N}_\sigma = 3$ in Eq. 5.3), corresponding to an accuracy of $\epsilon = 10^{-5} - 10^{-15}$, gives reasonably accurate results. This suggests that we can truncate the Coulomb tensor by more than 3/4 without losing any relevant information. With this, both computation time and storage are reduced.
Fig. 10b gives the connection between the accuracies ϵ shown in Fig. 10a and the singular values σ_i. It is clear that high accuracies are achieved with a small number of singular values since they decay super-exponentially. The subplot shows the range $\sigma_i \in [3, 15]$ where the logarithmic y-axis confirms the explained behavior. This also underlines why three singular values are kept for a relatively wide range of ϵ, allowing one to freely choose the lowest accuracy, here $\epsilon \approx 10^{-5}$, without losing information. Finally, we want to mention that the choice of the perfect N_σ is not only system specific but also depends on the considered number of spin orbitals and desired accuracy.

We apply the same considerations to the EOM-CCSD method.

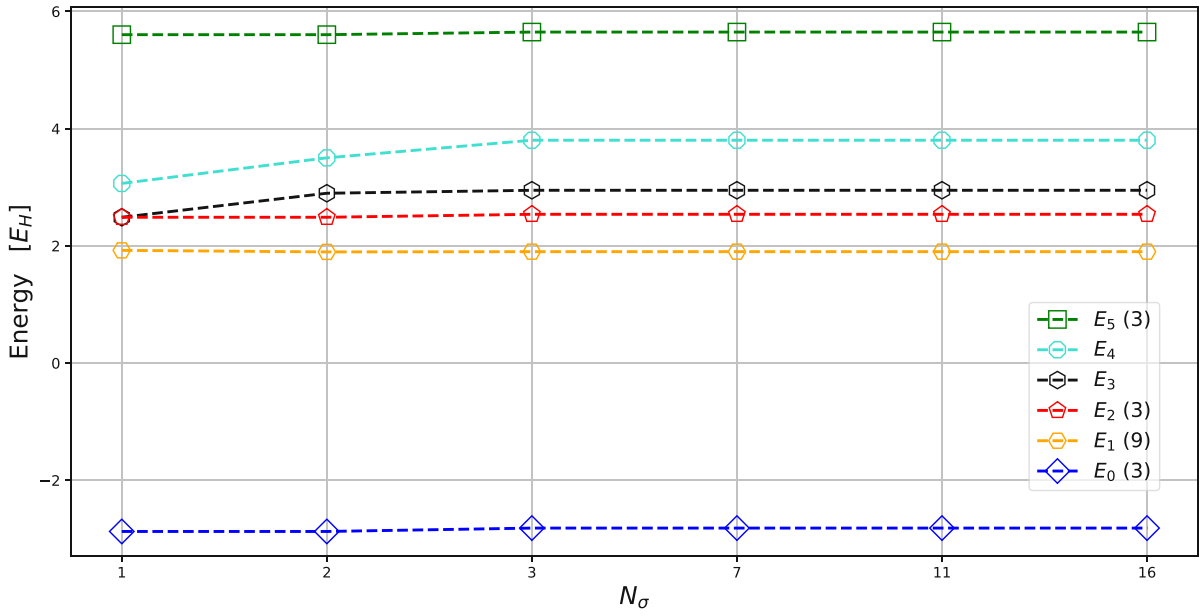

Figure 11: Energy spectrum obtained by the EOM-CCSD algorithm for 4 spin orbitals with the SVD truncated Coulomb tensor as a function of N_σ. We are using $\omega = 1\tau_0^{-1}$, $\sigma = 0.5$ and $N_{grid} = 500$.

Fig. 11 shows the already familiar energy spectrum from Fig. 8. To be precise for $N_\sigma = 16$, exactly the same plot is shown. Now, with decreasing number of singular values, the energies are slightly shifted. While E_0, E_1 and E_5 remain relatively constant, the energies levels E_2-E_4 vary, especially for $N_\sigma \leq 2$. The latter observation is consistent with the behavior observed in the CCSD method. In addition, with only one singular value retained, there seem to be more degeneracies as E_2 and E_3 merge.

In the next part of this section, we focus on the convergence analysis of the ALS and RALS algorithms, respectively. As mentioned earlier, this is a rather numerically motivated analysis, but it is essential for the CPD especially when it comes to realistic systems with large basis sets.

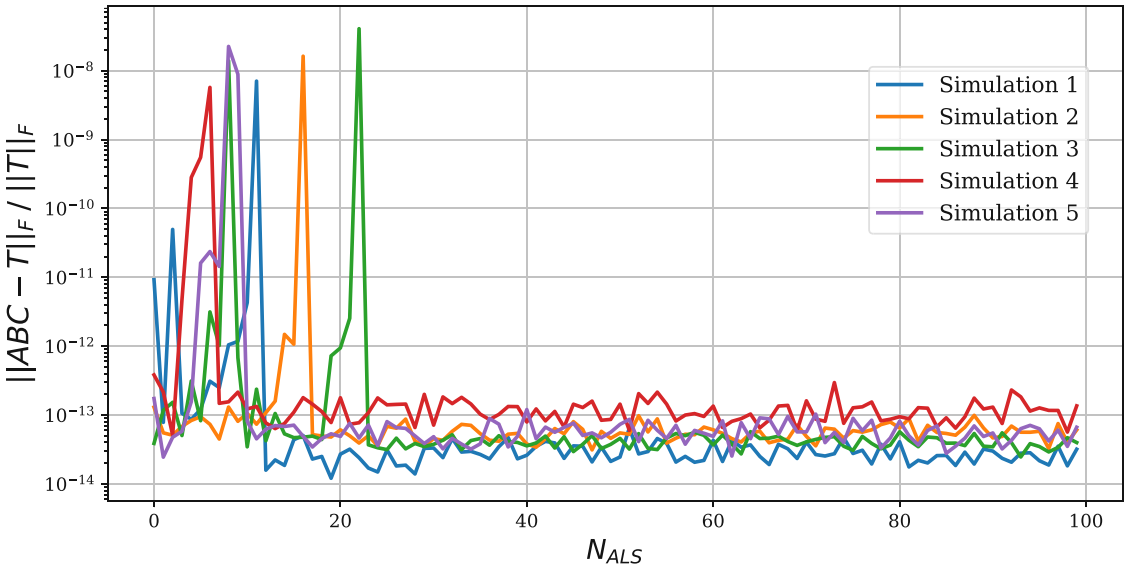

Figure 12: Convergence analysis of the ALS algorithm. The normalized Frobenius norm with respect to the number of ALS iterations is shown. We perform five independent simulations with the Coulomb vertex $\Gamma^{pr}_{N_\sigma}$ as input.

Fig. 12 shows the convergence criterion defined by Eq. 5.18 with respect to the number of self-consistent ALS iterations for five individual simulations. It is shown for all simulations that the normalized Frobenius norm fluctuates strongly in the first 20 iteration steps. Afterwards, all simulations show a common convergence which, however, oscillates around a certain value. The initial guess for the matrices denoted by A, B and C is random. This suggests that certain guesses yield more accurate results than others. We want to mention that empirical studies showed that, due to the nature of the random initial matrices, also less accurate simulations with $||ABC - T||_F \,/\, ||T||_F \approx 10^{-11}$ are possible. Also, some simulations showed no fluctuations in the first 20 iteration steps resulting in instant convergence. In summary, for large enough values of N_{ALS} convergence can be achieved independent of the initial guesses for the matrices. Further, depending on the initial guess, one may immediately find a minimum and obtain instant convergence. However, such behavior is only expected for relatively small systems. In general, the convergence of the ALS algorithm usually requires several hundreds of steps.

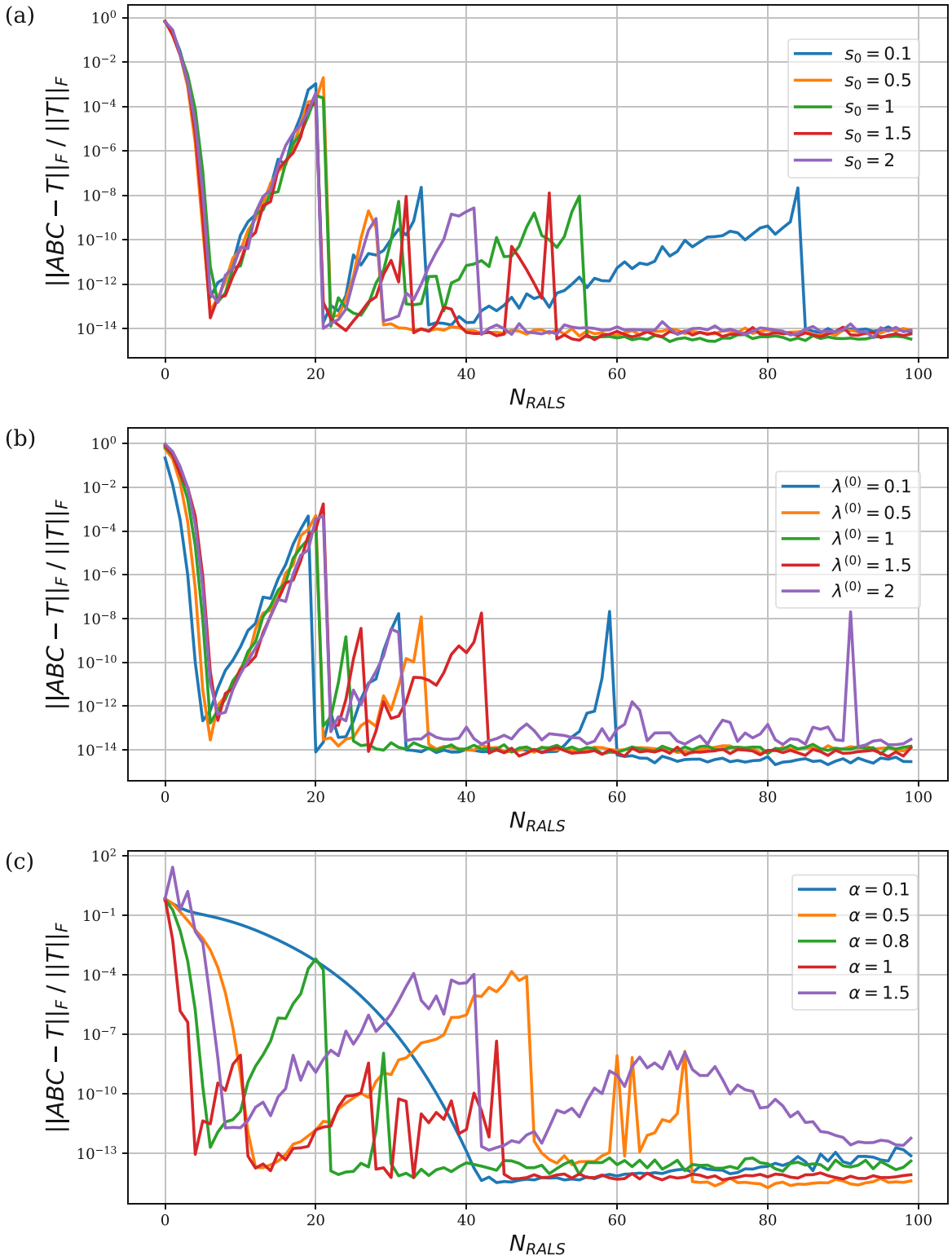

Figure 13: Convergence analysis of the RALS algorithm. The normalized Frobenius norm with respect to the number of RALS iterations is shown. We specifically investigate the convergence behavior by varying (a) the swamping threshold s_0, (b) the regularization parameter $\lambda^{(0)}$ and (c) the mixing factor α. As input the Coulomb vertex $\Gamma^{pr}_{N_\sigma}$ is used.

We now turn to the discussion of the implemented RALS algorithm. Fig. 13 shows the convergence behavior for different choices of the three tunable parameters in our RALS algorithm. Thereby, one parameter is varied while the others are kept fixed. The values of the fixed parameters are chosen to be the default ones introduced in Sec. 5.4. The variation of s_0 and λ in Fig. 13a and Fig. 13b, respectively, looks quite similar. The swamping threshold s_0 influences the estimated regularization parameter $\hat{\lambda}$ (s. Eq. 5.16) in such a way that values greater than one result in a bigger contribution of $\lambda^{(n-1)}$ to $\lambda^{(n)}$ and vice versa for values smaller than one. The influence of $\lambda^{(0)}$-values greater than one may be viewed as a bigger contribution from the old matrix, e.g. $A^{(n)}$ (s. Eq. 5.14). Otherwise, small $\lambda^{(0)}$-values lead to a smaller contribution of $A^{(n)}$ from the previous iteration. For both cases, the quantities of the previous iteration are more weighted in the present iteration for values greater than one and less weighted for values smaller than one. Furthermore, both figures have in common that in the first 7-8 steps a monotonic convergence is observed. After that, however, the normalized Frobenius norm increases again to find convergence after the mentioned 20 steps. This behavior can be understood as the algorithm trying to find a local minimum but overshooting it and thus having to find a new one, which in turn causes the norm to increase.
Fig. 13c on the other hand depicts the dependence of the convergence on the variation of α. Again, local minima are found after a few iterations, but the algorithm seems to have more trouble finding and staying at a minimum, unlike for the other two parameters. For $\alpha = 1$, the contribution of $\lambda^{(n-1)}$ vanishes and for values of the mixing parameter greater than one, $\lambda^{(n)}$ is even diminished by the amount of $\lambda^{(n-1)}$. Very small values like $\alpha = 0.1$ lead to slow convergence. This suggests that the contribution of $\lambda^{(n-1)}$ to $\lambda^{(n)}$ should be small compared to the one of $\hat{\lambda}^{(n)}$ (s. Eq. 5.17).
In summary, the default values of the three investigated parameters seem to be also suitable for our system. Especially for the reason that the change of one parameter, in the end, influences the other two, which makes their determination even more extensive. Further, we want to point out, that also for the RALS algorithm, the random nature of the initialized matrices A, B and C may lead to different convergence behaviors for different simulations. Nevertheless, the obtained accuracy is again sufficient for our purposes to compute of the energy spectrum.

After the confirmation that the default values of the RALS parameters are also suitable for our model, we want to go further and extend the discussed LRA obtained with the SVD by the CPD. Thereby, the SVD truncated Coulomb tensor $\tilde{V}_{rs}^{pq}$ is further decomposed. At this point, the accuracy ϵ of the SVD can be chosen arbitrarily, however, we continue with the found truncation to 3 singular values since it is computationally efficient while yielding accurate results.

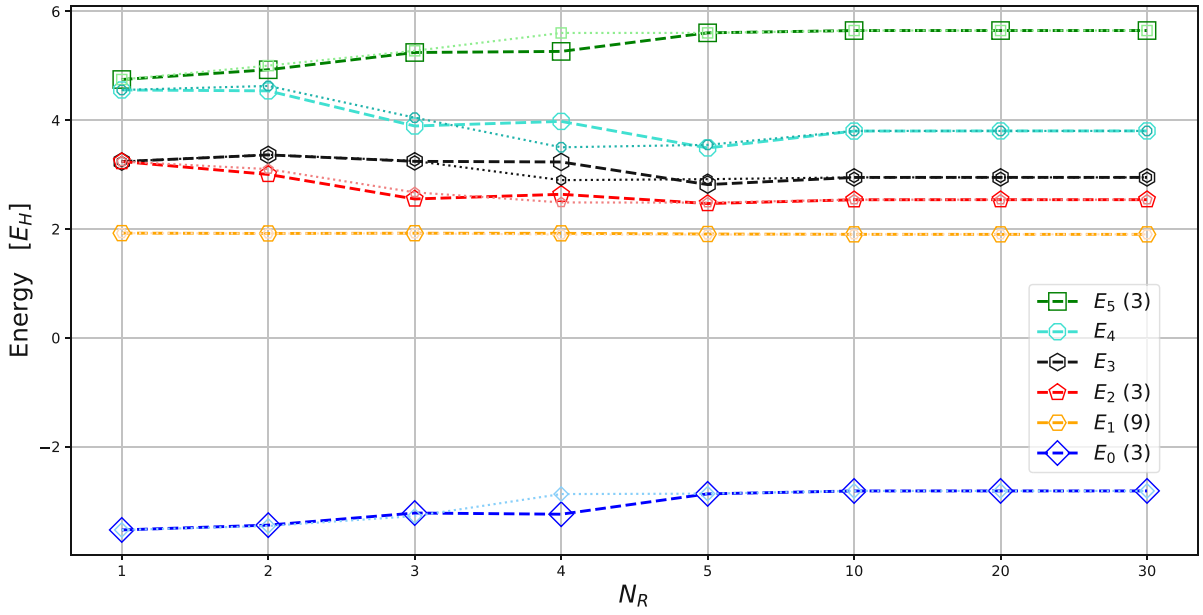

Figure 14: Energy spectrum obtained by the EOM-CCSD algorithm for 4 spin orbitals with the CPD truncated Coulomb tensor as a function of the number of vertex indices N_R. As the starting point for the CPD algorithm, we choose $N_\sigma = 3$ corresponding to an SVD accuracy of $\epsilon = 10^{-5}$. The RALS algorithm is executed with a total of 5 iterations. In addition, we plot another set of the five energy levels (each in a light color and with dotted lines) but with 100 iteration steps. The parameters $\omega = 1\tau_0^{-1}$, $\sigma = 0.5$ and $N_{grid} = 500$ are set.

In Fig. 14 the EOM-CCSD energy spectrum as a function of the number of vertex indices N_R is displayed. N_R is linked to the dimension of the Coulomb factor and the factor orbitals (s. Eq. 5.6) which are determined by the RALS algorithm. At this point, however, the choice of N_R is arbitrary. Fig. 14 shows, starting from the SVD approximated energies with $N_\sigma = 3$, that for $N_R \geq 10$ no visible change in the energy is present. Similarly, to the energy spectrum obtained with the SVD, we can conclude that with $N_R = 10$ reasonable accurate results are still obtained. For a lower number of vertex indices, all energy levels, except E_1, change.

It is important to note that in Fig. 14 our RALS algorithm stops after five iterations. We justify this choice by the fact that for sufficiently large N_R, in our case 10, the normalized Frobenius norm introduced in Eq. 5.18 converges very fast. Therefore, for these cases, on the one hand, we do not need a convergence condition that breaks the RALS algorithm, since convergence is guaranteed anyway. On the other hand, if there are only a few iterations, the speed of convergence is more apparent. For the case of smaller N_R, we found empirically that even for 100 iterations the accuracy as for larger N_R is not achieved. In this case the Frobenius norm is on the order of 10^{-1}. The latter finding therefore justifies our choice of a few iterations. Still, Fig. 14 also shows the results obtained with 100 iterations in the respective light color and dotted lines. For $N_R \geq 10$ there is no visible difference between both calculations. For smaller numbers of vertex indices slight deviations are recognizable, but sufficient accuracy is not guaranteed. Accordingly, due to the random nature of the initialized matrices and since convergence is not reached, every CPD simulation results in slightly different energies for small N_R. We want to emphasize here that for real systems the RALS algorithm with the convergence condition and not with a fixed number of iterations is recommended due to the probable slower convergence.

In conclusion, with $N_\sigma = 3$ and $N_R = 10$ accurate results compared to the reference results are obtained. More generally speaking, we found a low-rank approximation of the Coulomb tensor using a combination of SVD and CPD obtained by a RALS algorithm. As a result, while the overall computational cost of an iterative EOM-CCSD calculation could be substantially reduced by employing such a LRA, the same results up to a certain accuracy from the reference EOM-CCSD algorithm can be acquired.
Furthermore, while initially taking the choice of N_R arbitrarily, it turns out that $N_R \approx 5 \cdot N_\sigma$ represents a reasonable estimation. This is in agreement with Ref. [36], where N_R was found to be an order of magnitude lower than the number of real space grid points of the original factors of the Coulomb integrals.
More generally, we have found that the interplay between N_σ and N_R can be used to systematically reduce the computational cost while obtaining arbitrary accurate results.

7 Conclusion and Outlook

This work presents the application of low-rank approximation (LRA) techniques to improve correlated wave function-based theories, focusing primarily on the equation of motion coupled cluster singles and doubles (EOM-CCSD) theory. We examine, as a proof of principle, a system with two interacting electrons confined in a one-dimensional harmonic potential.

To achieve our main objective, we initially set up a closed-shell Hartree-Fock algorithm that provides the Hartree-Fock ground state energy and orbital energies by treating the electron-electron interaction in a mean-field approximation. In addition, we investigate the convergence behavior of the algorithm with respect to various system parameters, allowing us to obtain a balance between accuracy and computational cost. The Hartree-Fock wave functions are employed to compute the two-electron (Coulomb) integrals, based on which the Coulomb tensor is constructed. The Hartree-Fock orbital energies and the antisymmetrized Coulomb tensor serve as the starting point for the post-Hartree-Fock methods. By applying the second-order Møller Plesset perturbation (MP2) and coupled cluster singles and doubles (CCSD) theories, we achieve improvements in the ground state energy by incorporating a portion of the electronic correlation energy. Similarly to the analysis of the Hartree-Fock method, also for the CCSD method a convergence analysis is carried out enabling us to vary our system parameters such that the behavior of real molecular systems can be simulated. Through the EOM-CCSD algorithm, not only the ground state energies but also the excited state energies are obtained. The code requires the same input as the other two post-Hartree-Fock methods, along with the additional singles and doubles amplitudes calculated by the CCSD algorithm itself. The results of the EOM-CCSD algorithm reveal the presence of singlet and triplet states; they are however not only restricted to the fermionic case. The further requirement of antisymmetry and the Pauli principle yields the physically allowed solutions for the two interacting electrons.
In the final part of this thesis, we analyze the impact of employing the LRA to the Coulomb integrals on the outcomes of the aforementioned algorithms. Initially, we focus on the singular value decomposition (SVD) technique. Our findings demonstrate that, for our model, a substantial truncation of the Coulomb tensor is feasible while still maintaining accurate results within a given threshold. Subsequently, we go a step further and utilize the results of the SVD-decomposed and truncated Coulomb tensor to perform a canonical polyadic decomposition (CPD). This tensor factorization is obtained using a regularized alternating least squares algorithm (RALS) where the Coulomb tensor is approximated by six matrices, allowing us to potentially further reduce computational costs of EOM-CCSD calculations without sacrificing accuracy. We investigate the resulting energy spectrum for both the SVD and CPD methods and observe that the truncation of the Coulomb tensor in SVD, as well as its decomposition in CPD, influences the accuracy. This flexibility enables us to strike a balance between numerical efficiency and physical accuracy, depending on the specific problem at hand. However, the computation time and the memory footprint pose significant bottlenecks for most molecular systems making it essential to select the aforementioned interplay to achieve optimal computational efficiency.

In summary, we have successfully achieved the primary objective of this thesis, which is the application of LRA techniques to the Coulomb integrals in EOM-CCSD theory. The discussion and implementation of the methods leading to the development of the EOM-CCSD theory are crucial and constitute a significant portion of our work. We are able to confirm that the two tensor decomposition methods introduced herein have the potential to reduce the computational cost while maintaining the desired level of accuracy. The trade-off between computational efficiency and accuracy can thereby systematically be varied by the two parameters N_σ and N_R. This permits an efficient computation using the EOM-CCSD algorithm, facilitating the investigation of excited energy spectra.

Potential future work could include the extension of the presented equation of motion algorithm to the perturbative triplets regime, EOM-CCSD(T), which is recognized as the current benchmark (gold standard) in computational chemistry. Moreover, many opportunities arise when it comes to the reduction of computational expenses by means of the LRA of the Coulomb integrals. This includes exploring alternative tensor decomposition techniques, such as the Cholesky decomposition, as well as improving and extending the presented approaches. One possible enhancement to the presented RALS algorithm, as proposed in Ref. [36], could involve the additional treatment of doubly occurring factor orbitals. Another avenue worth considering is the accelerated RALS technique [46] as one of several potential extensions.
A further very interesting topic arises regarding the *ab inito* prediction of electronically and vibrationally excited state properties for real materials employing time-dependent EOM-CCSD theory.

Appendices

A CCSD Amplitude Equations

$\hat{T}_1$ amplitude equation:

$$\begin{aligned}
0 = {} & f_{ai} + \sum_{c} f_{ac} t_c^i - \sum_{k} f_{ki} t_k^a + \sum_{kc} \langle ka||ci\rangle \, t_k^c + \sum_{kc} f_{kc} t_{ik}^{ac} + \frac{1}{2} \sum_{kcd} \langle ka||cd\rangle \, t_{ki}^{cd} \\
& - \frac{1}{2} \sum_{klc} \langle kl||ci\rangle \, t_{kl}^{ca} - \sum_{kc} f_{kc} t_i^c t_k^a - \sum_{klc} \langle kl||ci\rangle \, t_k^c t_l^a - \sum_{kcd} \langle ka||cd\rangle \, t_k^c t_i^d \\
& - \sum_{klcd} \langle kl||cd\rangle \, t_k^c t_i^d t_l^a + \sum_{klcd} \langle kl||cd\rangle \, t_k^c t_{li}^{da} - \frac{1}{2} \sum_{klcd} \langle kl||cd\rangle \, t_{ki}^{cd} t_l^a - \frac{1}{2} \sum_{klcd} \langle kl||cd\rangle \, t_{kl}^{ca} t_i^d.
\end{aligned} \tag{A.1}$$

$\hat{T}_2$ amplitude equation:

$$\begin{aligned}
0 = {} & \langle ab||ij\rangle + \sum_{c} (f_{bc} t_{ij}^{ac} - f_{ac} t_{ij}^{bc}) - \sum_{k} (f_{kj} t_{ik}^{ab} - f_{ki} t_{jk}^{ab}) + \frac{1}{2} \sum_{kl} \langle kl||ij\rangle \, t_{kl}^{ab} \\
& + \frac{1}{2} \sum_{cd} \langle ab||cd\rangle \, t_{ij}^{cd} + \hat{P}_{ij} \hat{P}_{ab} \sum_{kc} \langle kb||cj\rangle \, t_{ik}^{ac} + \hat{P}_{ij} \sum_{c} \langle ab||cj\rangle \, t_i^c \\
& - \hat{P}_{ab} \sum_{k} \langle kb||ij\rangle \, t_k^a + \frac{1}{2} \hat{P}_{ij} \hat{P}_{ab} \sum_{klcd} \langle kl||cd\rangle \, t_{ik}^{ac} t_{lj}^{db} + \frac{1}{4} \sum_{klcd} \langle kl||cd\rangle \, t_{ij}^{cd} t_{kl}^{ab} \\
& - \frac{1}{2} \hat{P}_{ab} \sum_{klcd} \langle kl||cd\rangle \, t_{ij}^{ac} t_{kl}^{bd} - \frac{1}{2} \hat{P}_{ij} \sum_{klcd} \langle kl||cd\rangle \, t_{ik}^{ab} t_{jl}^{cd} + \frac{1}{2} \hat{P}_{ab} \sum_{kl} \langle kl||ij\rangle \, t_k^a t_l^b \\
& + \frac{1}{2} \hat{P}_{ij} \sum_{cd} \langle ab||cd\rangle \, t_i^c t_j^d - \hat{P}_{ij} \hat{P}_{ab} \sum_{kc} \langle kb||ic\rangle \, t_k^a t^b c_j + \hat{P}_{ab} \sum_{kc} f_{kc} t_k^a t_{ij}^{bc} \\
& + \hat{P}_{ij} \sum_{kc} f_{kc} t_i^c t_{jk}^{bc} - \hat{P}_{ij} \sum_{klc} \langle kl||ci\rangle \, t_k^c t_{lj}^{ab} + \hat{P}_{ab} \sum_{kcd} \langle ka||cd\rangle \, t_k^c t_{ij}^{db} \\
& + \hat{P}_{ij} \hat{P}_{ab} \sum_{kcd} \langle ak||dc\rangle \, t_i^d t_{jk}^{bc} + \hat{P}_{ij} \hat{P}_{ab} \sum_{klc} \langle kl||ic\rangle \, t_l^a t_{jk}^{bc} + \frac{1}{2} \hat{P}_{ij} \sum_{klc} \langle kl||cj\rangle \, t_i^c t_{kl}^{ab} \\
& - \frac{1}{2} \hat{P}_{ab} \sum_{kcd} \langle kb||cd\rangle \, t_k^a t_{ij}^{cd} - \frac{1}{2} \hat{P}_{ij} \hat{P}_{ab} \sum_{kcd} \langle kb||cd\rangle \, t_i^c t_k^a t_j^d + \frac{1}{2} \hat{P}_{ij} \hat{P}_{ab} \sum_{klc} \langle kl||cj\rangle \, t_i^c t_k^a t_l^b \\
& - \hat{P}_{ij} \sum_{klcd} \langle kl||cd\rangle \, t_k^c t_i^d t_{lj}^{ab} - \hat{P}_{ab} \sum_{klcd} \langle kl||cd\rangle \, t_k^c t_l^a t_{ij}^{db} + \frac{1}{4} \hat{P}_{ij} \sum_{klcd} \langle kl||cd\rangle \, t_i^c t_j^d t_{kl}^{ab} \\
& + \frac{1}{4} \hat{P}_{ab} \sum_{klcd} \langle kl||cd\rangle \, t_k^a t_l^b t_{ij}^{cd} + \hat{P}_{ij} \hat{P}_{ab} \sum_{klcd} \langle kl||cd\rangle \, t_i^c t_l^b t_{kj}^{ad} + \frac{1}{4} \hat{P}_{ij} \hat{P}_{ab} \sum_{klcd} \langle kl||cd\rangle \, t_i^c t_k^a t_j^d t_l^b.
\end{aligned} \tag{A.2}$$

Here, $\hat{P}_{pq}$ denotes the permutation operator which action is defined as:

$$\hat{P}_{pq} f_{pq} = f_{pq} - f_{qp}.$$

B Contributions to the EOM-CCSD Matrix

Singles-singles block:

$$\begin{aligned}
\bar{H}_{SS} &= \langle \Psi_i^a | \bar{H} | \Psi_k^c \rangle \\
&= f_{ac}\delta_{ik} - f_{ik}\delta_{ac} + \langle ak||ic\rangle + \sum_e \left(\langle ak||ec\rangle \, t_i^e - f_{ke} t_i^e \delta_{ac} \right) + \sum_m \Big(- \langle mk||ic\rangle \, t_m^a - f_{mc} t_m^a \delta_{ik} \Big) \\
&+ \sum_{me} \Big(\langle am||ce\rangle \, t_m^e \delta_{ik} - \langle km||ie\rangle \, t_m^e \delta_{ac} + \langle km||ce\rangle \, t_{mi}^{ea} - \langle mk||ec\rangle \, t_i^e t_m^a \Big) + \sum_{men} \\
&- \frac{1}{2} \langle mn||ce\rangle \, t_{mn}^{ae} \delta_{ik} + \sum_{mef} \Big(- \frac{1}{2} \langle km||ef\rangle \, t_{im}^{ef} \delta_{ac} - \langle km||ef\rangle \, t_i^e t_m^f \delta_{ac} \Big) \\
&- \sum_{mnf} \langle mn||cf\rangle \, t_m^a t_n^f \delta_{ik}.
\end{aligned} \tag{B.1}$$

Singles-doubles block:

$$\begin{aligned}
\bar{H}_{SD} = \langle \Psi_i^a | \bar{H} | \Psi_{kl}^{cd} \rangle = f_{ld}\delta_{ik}\delta_{ac} + \frac{1}{2} \Big(\langle al||cd\rangle \, \delta_{ik} - \langle kl||id\rangle \, \delta_{ac} \Big) \\
- \sum_e \frac{1}{2} \langle kl||ed\rangle \, t_i^e \delta_{ac} - \sum_m \frac{1}{2} \langle ml||cd\rangle \, t_m^a \delta_{ik} + \sum_{me} \langle mk||ec\rangle \, t_m^e \delta_{il}\delta_{ad}.
\end{aligned} \tag{B.2}$$

Doubles-singles block:

$$\begin{aligned}
\bar{H}_{DS} &= \langle \Psi_{ij}^{ab} | \bar{H} | \Psi_k^c \rangle \\
&= \langle ab||cj\rangle \, \delta_{ik} - \langle ab||ci\rangle \, \delta_{jk} + \langle ka||ij\rangle \, \delta_{bc} - \langle kb||ij\rangle \, \delta_{ac} \\
&+ \sum_e \Big(\langle ab||ce\rangle \, t_j^e \delta_{ik} - \langle ab||ce\rangle \, t_i^e \delta_{jk} + \langle ka||ej\rangle \, t_i^e \delta_{bc} - \langle ka||ei\rangle \, t_j^e \delta_{bc} - \langle kb||ej\rangle \, t_i^e \delta_{ac} \\
&+ \langle kb||ei\rangle \, t_j^e \delta_{ac} + f_{ke} t_{ij}^{ea} \delta_{bc} - f_{ke} t_{ij}^{eb} \delta_{ac} + \langle ka||ce\rangle \, t_{ij}^{eb} - \langle kb||ce\rangle \, t_{ij}^{ea} \Big) \\
&+ \sum_{ef} \Big(\frac{1}{2} \langle ka||ef\rangle \, t_{ij}^{ef} \delta_{bc} - \frac{1}{2} \langle kb||ef\rangle \, t_{ij}^{ef} \delta_{ac} - \langle kb||ef\rangle \, t_i^e t_j^f \delta_{ac} + \langle ka||ef\rangle \, t_i^e t_j^f \delta_{bc} \Big) \\
&+ \sum_m \Big(\langle km||ij\rangle \, t_m^b \delta_{ac} - \langle km||ij\rangle \, t_m^a \delta_{bc} + \langle mb||ci\rangle \, t_m^a \delta_{jk} - \langle ma||ci\rangle \, t_m^b \delta_{jk} - \langle mb||cj\rangle \, t_m^a \delta_{ik} \\
&+ \langle ma||cj\rangle \, t_m^b \delta_{ik} + f_{mc} t_{ni}^{ab} \delta_{jk} - f_{mc} t_{nj}^{ab} \delta_{ik} + \langle km||cj\rangle \, t_{mi}^{ab} - \langle km||ci\rangle \, t_{mj}^{ab} \Big) \\
&+ \sum_{mn} \Big(\frac{1}{2} \langle mn||cj\rangle \, t_{mn}^{ab} \delta_{ik} - \frac{1}{2} \langle mn||ci\rangle \, t_{mn}^{ab} \delta_{jk} - \langle mn||cj\rangle \, t_m^a t_n^b \delta_{ik} + \langle mn||ci\rangle \, t_m^a t_n^b \delta_{jk} \Big) \\
&+ \sum_{me} \Big(\langle am||ce\rangle \, t_{mj}^{eb} \delta_{ik} - \langle bm||ce\rangle \, t_{mj}^{ea} \delta_{ik} - \langle am||ce\rangle \, t_{mi}^{eb} \delta_{jk} + \langle bm||ce\rangle \, t_{mi}^{ea} \delta_{jk} \\
&+ \langle km||ie\rangle \, t_{mj}^{ea} \delta_{bc} - \langle km||je\rangle \, t_{mi}^{ea} \delta_{bc} - \langle km||ie\rangle \, t_{mj}^{eb} \delta_{ac} + \langle km||je\rangle \, t_{mi}^{eb} \delta_{ac} \\
&+ \langle mb||ce\rangle \, t_m^a t_i^e \delta_{jk} - \langle ma||ce\rangle \, t_m^b t_i^e \delta_{jk} - \langle mb||ce\rangle \, t_m^a t_j^e \delta_{ik} + \langle ma||ce\rangle \, t_m^b t_j^e \delta_{ik} \\
&+ \langle km||ej\rangle \, t_i^e t_m^b \delta_{ac} - \langle km||ei\rangle \, t_j^e t_m^b \delta_{ac} - \langle km||ej\rangle \, t_i^e t_m^a \delta_{bc} + \langle km||ei\rangle \, t_j^e t_m^a \delta_{bc} \\
&+ \langle km||ce\rangle \, t_j^e t_{mi}^{ab} - \langle km||ce\rangle \, t_i^e t_{mj}^{ab} + \langle km||ce\rangle \, t_m^b t_{ij}^{ea} - \langle km||ce\rangle \, t_m^a t_{ij}^{eb} \Big)
\end{aligned} \tag{B.3}$$

$$
\begin{aligned}
&+\sum_{men}\Big(\frac{1}{2}\langle mn||ce\rangle\, t_j^e t_{mn}^{ab}\delta_{ik}-\frac{1}{2}\langle mn||ce\rangle\, t_i^e t_{mn}^{ab}\delta_{jk}+\langle mn||ce\rangle\, t_m^a t_{ni}^{eb}\delta_{jk}-\langle mn||ce\rangle\, t_m^b t_{ni}^{ea}\delta_{jk}\\
&-\langle mn||ce\rangle\, t_m^a t_{nj}^{eb}\delta_{ik}+\langle mn||ce\rangle\, t_m^b t_{nj}^{ea}\delta_{ik}+\langle mn||ec\rangle\, t_m^e t_{ni}^{ab}\delta_{jk}-\langle mn||ec\rangle\, t_m^e t_{nj}^{ab}\delta_{ik}\\
&-\langle mn||ec\rangle\, t_i^e t_m^a t_n^b\delta_{jk}+\langle mn||ec\rangle\, t_j^e t_m^a t_n^b\delta_{ik}\Big)\\
&+\sum_{mef}\Big(\frac{1}{2}\langle km||ef\rangle\, t_m^b t_{ij}^{ef}\delta_{ac}-\frac{1}{2}\langle km||ef\rangle\, t_m^a t_{ij}^{ef}\delta_{bc}+\langle km||ef\rangle\, t_i^e t_{mj}^{fa}\delta_{bc}-\langle km||ef\rangle\, t_j^e t_{mi}^{fa}\delta_{bc}\\
&-\langle km||ef\rangle\, t_i^e t_{mj}^{fb}\delta_{ac}+\langle km||ef\rangle\, t_j^e t_{mi}^{fb}\delta_{ac}+\langle mk||ef\rangle\, t_m^e t_{ij}^{fa}\delta_{bc}-\langle mk||ef\rangle\, t_m^e t_{ij}^{fb}\delta_{ac}\\
&-\langle mk||ef\rangle\, t_i^e t_m^a t_j^f\delta_{bc}+\langle mk||ef\rangle\, t_i^e t_m^b t_j^f\delta_{ac}\Big).
\end{aligned}
$$

Doubles-doubles block:

$$
\begin{aligned}
\bar{H}_{DD} &= \langle\Psi_{ij}^{ab}|\bar{H}|\Psi_{kl}^{cd}\rangle\\
&= f_{bc}\delta_{jk}\delta_{il}\delta_{ad}-f_{ac}\delta_{jk}\delta_{il}\delta_{bd}+f_{ki}\delta_{jl}\delta_{ad}\delta_{bc}-f_{kj}\delta_{il}\delta_{ad}\delta_{bc}\\
&+\frac{1}{2}\langle ab||cd\rangle\,\delta_{ik}\delta_{jl}-\frac{1}{2}\langle kl||ij\rangle\,\delta_{ac}\delta_{bd}\\
&+\langle kb||cj\rangle\,\delta_{il}\delta_{ad}-\langle kb||ci\rangle\,\delta_{jl}\delta_{ad}-\langle ka||cj\rangle\,\delta_{il}\delta_{bd}+\langle ka||ci\rangle\,\delta_{jl}\delta_{bd}\\
&+\sum_{e}\Big(f_{ke}t_j^e\delta_{ac}\delta_{il}\delta_{bd}-f_{ke}t_i^e\delta_{ac}\delta_{jl}\delta_{bd}+\langle ak||ec\rangle\, t_i^e\delta_{jl}\delta_{bd}\\
&-\langle ak||ec\rangle\, t_j^e\delta_{il}\delta_{bd}-\langle bk||ec\rangle\, t_i^e\delta_{jl}\delta_{ad}+\langle bk||ec\rangle\, t_j^e\delta_{il}\delta_{ad}\\
&+\frac{1}{2}\langle kl||ej\rangle\, t_i^e\delta_{bd}\delta_{ac}-\frac{1}{2}\langle kl||ei\rangle\, t_j^e\delta_{bd}\delta_{ac}+\frac{1}{2}\langle kl||ce\rangle\, t_{ij}^{ea}\delta_{bd}-\frac{1}{2}\langle kl||ce\rangle\, t_{ij}^{eb}\delta_{ad}\Big)\\
&+\sum_{ef}\Big(\frac{1}{4}\langle kl||ef\rangle\, t_{ij}^{ef}\delta_{ac}\delta_{bd}+\frac{1}{2}\langle kl||ef\rangle\, t_i^e t_j^f\delta_{ac}\delta_{bd}\Big)\\
&+\sum_{m}\Big(f_{mc}t_m^b\delta_{ik}\delta_{jl}\delta_{ad}-f_{mc}t_m^a\delta_{ik}\delta_{jl}\delta_{bd}+\langle mk||jc\rangle\, t_m^a\delta_{il}\delta_{bd}-\langle mk||ic\rangle\, t_m^a\delta_{jl}\delta_{bd}\\
&-\langle mk||jc\rangle\, t_m^b\delta_{il}\delta_{ad}+\langle mk||ic\rangle\, t_m^b\delta_{jl}\delta_{ad}+\frac{1}{2}\langle ma||cd\rangle\, t_m^b\delta_{jl}\delta_{ik}-\frac{1}{2}\langle mb||cd\rangle\, t_m^a\delta_{jl}\delta_{ik}\\
&+\frac{1}{2}\langle km||cd\rangle\, t_{mi}^{ab}\delta_{jl}-\frac{1}{2}\langle km||cd\rangle\, t_{mj}^{ab}\delta_{il}\Big)\\
&+\sum_{mn}\Big(\frac{1}{4}\langle mn||cd\rangle\, t_{mn}^{ab}\delta_{ik}\delta_{jl}+\frac{1}{2}\langle mn||cd\rangle\, t_m^a t_n^b\delta_{ik}\delta_{jl}\Big)\\
&+\sum_{me}\Big(\langle ma||ec\rangle\, t_m^e\delta_{ik}\delta_{jl}\delta_{bd}-\langle mb||ec\rangle\, t_m^e\delta_{ik}\delta_{jl}\delta_{ad}+\langle mk||ej\rangle\, t_m^e\delta_{ac}\delta_{il}\delta_{bd}\\
&-\langle mk||ei\rangle\, t_m^e\delta_{ac}\delta_{jl}\delta_{bd}+\langle km||ce\rangle\, t_{mi}^{ea}\delta_{jl}\delta_{bd}-\langle km||ce\rangle\, t_{mj}^{ea}\delta_{il}\delta_{bd}-\langle km||ce\rangle\, t_{mi}^{eb}\delta_{jl}\delta_{ad}\\
&+\langle km||ce\rangle\, t_{mj}^{eb}\delta_{il}\delta_{ad}+\langle mk||ec\rangle\, t_i^e t_m^b\delta_{lj}\delta_{ad}-\langle mk||ec\rangle\, t_i^e t_m^a\delta_{lj}\delta_{bd}-\langle mk||ec\rangle\, t_j^e t_m^b\delta_{li}\delta_{ad}\\
&+\langle mk||ec\rangle\, t_j^e t_m^a\delta_{li}\delta_{bd}\Big)
\end{aligned}
\tag{B.4}
$$

$$+\sum_{men}\Big(\frac{1}{2}\langle mn||ec\rangle\, t^{be}_{nm}\delta_{ik}\delta_{jl}\delta_{ad}-\frac{1}{2}\langle mn||ec\rangle\, t^{ae}_{nm}\delta_{ik}\delta_{jl}\delta_{bd}+\langle mn||ec\rangle\, t^{e}_{m}t^{b}_{n}\delta_{ik}\delta_{jl}\delta_{ad}$$
$$-\langle mn||ec\rangle\, t^{e}_{m}t^{a}_{n}\delta_{ik}\delta_{jl}\delta_{bd}\Big)+\sum_{mef}\Big(\frac{1}{2}\langle mk||ef\rangle\, t^{fe}_{jm}\delta_{ac}\delta_{il}\delta_{bd}-\frac{1}{2}\langle mk||ef\rangle\, t^{fe}_{im}\delta_{ac}\delta_{jl}\delta_{bd}$$
$$+\langle mk||ef\rangle\, t^{f}_{j}t^{e}_{m}\delta_{ik}\delta_{ac}\delta_{il}\delta_{bd}-\langle mk||ef\rangle\, t^{f}_{i}t^{e}_{m}\delta_{jk}\delta_{ac}\delta_{il}\delta_{bd}\Big).$$

List of Figures

List of Tables

References

[1] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Physical Review*, vol. 136, no. 3B, pp. B864–B871, 1964.

[2] R. G. Parr, "Density Functional Theory of Atoms and Molecules," in *Horizons of Quantum Chemistry* (K. Fukui, ed.), Académie Internationale Des Sciences Moléculaires Quantiques / International Academy of Quantum Molecular Science, pp. 5–15, Dordrecht: Springer, 1980.

[3] J. P. Perdew, "Density-functional approximation for the correlation energy of the inhomogeneous electron gas," *Physical review. B, Condensed matter*, vol. 33, no. 12, pp. 8822–8824, 1986.

[4] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Challenges for density functional theory," *Chemical reviews*, vol. 112, no. 1, pp. 289–320, 2012.

[5] T. Gruber and A. Grüneis, "Ab initio calculations of carbon and boron nitride allotropes and their structural phase transitions using periodic coupled cluster theory," *Physical Review B*, vol. 98, no. 13, 2018.

[6] H. Stoll and K. Doll, "Approaching the bulk limit with finite cluster calculations using local increments: the case of LiH," *The Journal of Chemical Physics*, vol. 136, no. 7, p. 074106, 2012.

[7] A. Grüneis, "Efficient Explicitly Correlated Many-Electron Perturbation Theory for Solids: Application to the Schottky Defect in MgO," *Physical review letters*, vol. 115, no. 6, p. 066402, 2015.

[8] A. D. Boese and J. Sauer, "Accurate adsorption energies of small molecules on oxide surfaces: CO-MgO(001)," *Physical chemistry chemical physics : PCCP*, vol. 15, no. 39, pp. 16481–16493, 2013.

[9] F. Hummel, T. Gruber, and A. Grüneis, "A many-electron perturbation theory study of the hexagonal boron nitride bilayer system*," *The European Physical Journal B*, vol. 89, no. 11, 2016.

[10] K. Rościszewski, B. Paulus, P. Fulde, and H. Stoll, "Ab initio calculation of ground-state properties of rare-gas crystals," *Physical review. B, Condensed matter*, vol. 60, no. 11, pp. 7905–7910, 1999.

[11] S. J. Nolan, M. J. Gillan, D. Alfè, N. L. Allan, and F. R. Manby, "Calculation of properties of crystalline lithium hydride using correlated wave function theory," *Physical Review B*, vol. 80, no. 16, 2009.

[12] F. Coester, "Bound states of a many-particle system," *Nuclear Physics*, vol. 7, pp. 421–424, 1958.

[13] F. Coester and H. Kümmel, "Short-range correlations in nuclear wave functions," *Nuclear Physics*, vol. 17, pp. 477–485, 1960.

[14] J. Čížek, "On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell–Type Expansion Using Quantum–Field Theoretical Methods," *The Journal of Chemical Physics*, vol. 45, no. 11, pp. 4256–4266, 1966.

[15] J. Paldus, J. Čížek, and I. Shavitt, "Correlation Problems in Atomic and Molecular Systems. IV. Extended Coupled-Pair Many-Electron Theory and Its Application to the B H3 Molecule," *Physical Review A*, vol. 5, no. 1, pp. 50–67, 1972.

[16] J. F. Stanton and R. J. Bartlett, "The equation of motion coupled–cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties," *The Journal of Chemical Physics*, vol. 98, no. 9, pp. 7029–7039, 1993.

[17] R. J. Bartlett and M. Musiał, "Coupled-cluster theory in quantum chemistry," *Reviews of Modern Physics*, vol. 79, no. 1, pp. 291–352, 2007.

[18] J. McClain, Q. Sun, G. K.-L. Chan, and T. C. Berkelbach, "Gaussian-Based Coupled-Cluster Theory for the Ground-State and Band Structure of Solids," *Journal of chemical theory and computation*, vol. 13, no. 3, pp. 1209–1218, 2017.

[19] N. Shenvi, H. van Aggelen, Y. Yang, W. Yang, C. Schwerdtfeger, and D. Mazziotti, "The tensor hypercontracted parametric reduced density matrix algorithm: coupled-cluster accuracy with O(r(4)) scaling," *The Journal of Chemical Physics*, vol. 139, no. 5, p. 054110, 2013.

[20] E. G. Hohenstein, R. M. Parrish, and T. J. Martínez, "Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Møller-Plesset perturbation theory," *The Journal of Chemical Physics*, vol. 137, no. 4, p. 044103, 2012.

[21] R. M. Parrish, E. G. Hohenstein, T. J. Martínez, and C. D. Sherrill, "Tensor hypercontraction. II. Least-squares renormalization," *The Journal of Chemical Physics*, vol. 137, no. 22, p. 224106, 2012.

[22] E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, and T. J. Martínez, "Communication: Tensor hypercontraction. III. Least-squares tensor hypercontraction for the determination of correlated wavefunctions," *The Journal of Chemical Physics*, vol. 137, no. 22, p. 221101, 2012.

[23] U. Benedikt, K.-H. Böhm, and A. A. Auer, "Tensor decomposition in post-Hartree-Fock methods. II. CCD implementation," *The Journal of Chemical Physics*, vol. 139, no. 22, p. 224101, 2013.

[24] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. New York: Dover Publications, 1996.

[25] J. C. Slater, "The Theory of Complex Spectra," *Physical Review*, vol. 34, no. 10, pp. 1293–1322, 1929.

[26] E. U. Condon, "The Theory of Complex Spectra," *Physical Review*, vol. 36, no. 7, pp. 1121–1133, 1930.

[27] F. A. Berazin, *The Method of Second Quantization*. Oxford: Elsevier Science, 1966.

[28] V. Fock, "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems," *Zeitschrift für Physik*, vol. 61, no. 1-2, pp. 126–148, 1930.

[29] S. Lehtola, "Assessment of Initial Guesses for Self-Consistent Field Calculations. Superposition of Atomic Potentials: Simple yet Efficient," *Journal of chemical theory and computation*, vol. 15, no. 3, pp. 1593–1604, 2019.

[30] T. D. Crawford and H. F. Schaefer, *An Introduction to Coupled Cluster Theory for Computational Chemists*. Reviews in Computational Chemistry, 2000.

[31] J. Čížek, "On the Use of the Cluster Expansion and the Technique of Diagrams in Calculations of Correlation Effects in Atoms and Molecules," in *Advances in Chemical Physics* (R. LeFebvre and C. Moser, eds.), Advances in Chemical Physics, pp. 35–89, Hoboken, NJ, USA: John Wiley & Sons, Inc, 1969.

[32] J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, "A direct product decomposition approach for symmetry exploitation in many–body methods. I. Energy calculations," *The Journal of Chemical Physics*, vol. 94, no. 6, pp. 4334–4345, 1991.

[33] J. Gauss, J. F. Stanton, and R. J. Bartlett, "Coupled–cluster open–shell analytic gradients: Implementation of the direct product decomposition approach in energy gradient calculations," *The Journal of Chemical Physics*, vol. 95, no. 4, pp. 2623–2638, 1991.

[34] J. Geertsen, M. Rittby, and R. J. Bartlett, "The equation-of-motion coupled-cluster method: Excitation energies of Be and CO," *Chemical Physics Letters*, vol. 164, no. 1, pp. 57–62, 1989.

[35] J. J. Goings, M. Caricato, M. J. Frisch, and X. Li, "Assessment of low-scaling approximations to the equation of motion coupled-cluster singles and doubles equations," *The Journal of Chemical Physics*, vol. 141, no. 16, p. 164116, 2014.

[36] F. Hummel, T. Tsatsoulis, and A. Grüneis, "Low rank factorization of the Coulomb integrals for periodic coupled cluster theory," *The Journal of Chemical Physics*, vol. 146, no. 12, p. 124105, 2017.

[37] F. L. Hitchcock, "The Expression of a Tensor or a Polyadic as a Sum of Products," *Journal of Mathematics and Physics*, vol. 6, no. 1-4, pp. 164–189, 1927.

[38] G. Venter, "Review of Optimization Techniques," in *Encyclopedia of Aerospace Engineering* (R. Blockley and W. Shyy, eds.), Chichester, UK: John Wiley & Sons, Ltd, 2010.

[39] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[40] E. H. Moore, "The fourteenth western meeting of the American Mathematical Society," *Bulletin of the American Mathematical Society*, vol. 26, no. 9, pp. 385–397, 1920.

[41] R. Penrose, "A generalized inverse for matrices," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 3, pp. 406–413, 1955.

[42] A. Cichocki and R. Zdunek, "Regularized alternating least squares algorithms for non-negative matrix/tensor factorization," *Advances in neural networks - ISNN 2007, pp. 793–802, Springer*, 2007.

[43] S. Kindermann and C. Navasca, "Analysis and Approximation of the Canonical Polyadic Tensor Decomposition," *arXiv preprint*, 2011.

[44] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic–Structure Theory*. Wiley, 2000.

[45] T. Kato, "On the eigenfunctions of many-particle systems in quantum mechanics," *Communications on Pure and Applied Mathematics*, vol. 10, no. 2, pp. 151–177, 1957.

[46] X. Wang, C. Navasca, and S. Kindermann, "On accelerating the regularized alternating least-squares algorithm for tensors," *ETNA - Electronic Transactions on Numerical Analysis*, vol. 48, pp. 1–14, 2018.