Implementation and Application of the Core Polarization Potential Ansatz in Quantum Chemical Systems

INAUGURAL–DISSERTATION
zur Erlangung des Doktorgrades
der Mathematisch–Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
SASCHA BUBECK
aus Esslingen am Neckar

Köln 2021

Berichterstatter:	Prof. Dr. Michael Dolg
	PD Dr. Michael Hanrath
Tag der mündlichen Prüfung:	07. Dezember 2021

Abstract

Modern quantum chemical simulations are computationally demanding and one approach to reduce these demands is the pseudopotential ansatz. This ansatz, however, approximates the interaction of the valence and core electrons in a way that neglects the polarizability of the atomic core. While the core polarization potential (CPP) ansatz addresses this neglected polarization, exhaustive studies of its influence in molecular simulations are still scarce. Here, we present a customized implementation of the CPP ansatz to establish an entry point for such studies in the Quantum Objects Library, a program package of the Institute for Theoretical Chemistry (University of Cologne). We successfully tested the implementation in atomic and molecular systems, both on the Hartree-Fock and the electron-correlation level. Additionally, we investigated the influence of CPPs in two systems - Hg_2 and HgF_4 - employing scalar-relativistic small-core pseudopotentials for mercury. For this, we developed a new approach for generating CPPs relying purely on *ab initio* data. The influence of the generated CPPs on the molecular properties of the investigated systems was small and requires a more detailed investigation.

Contents

1 Introduction

Simulating chemical systems is one of the major goals in quantum chemistry. Over the last decades, the advances in computer technology and in quantum chemical approaches opened new possibilities for simulating increasingly larger systems. Computational demands, however, restrict the size of such systems such that approximations are required to reduce these demands. Especially the simulation of systems containing heavy atoms is computationally demanding due to the great number of electrons present in these systems. Additionally, relativistic effects, which are pronounced in such systems, have to be taken into account, thereby further increasing the computational demand.

One approach, which both incorporates relativistic effects and reduces the computational demands for heavy atoms, is the use of effective core potentials (ECPs) [1]. The most widely applied form of the ECP ansatz are pseudopotentials (PPs). In this ansatz one fundamental consideration is the so-called frozen-core approximation, in which the inner electrons are excluded from the quantum chemical optimizations of the wave function. However, such treatment results in a missing description of the polarizability of the core and thus in a missing core-valence and core-core correlation [1]. To correct for this missing correlation, the core polarization potential (CPP) ansatz can be employed [2, 3]. This ansatz was employed in simulations of various systems over the last decades. Such simulations containing main group [4–7] and transition metal [8, 9] elements and f-elements were reported [10, 11]. For employing the CPP ansatz in modern quantum chemistry program packages, Schwerdtfeger and Silberbach proposed a way to solve the multi-center integrals in this ansatz using Cartesian Gaussian functions [12]. Their work represents the foundation for the only currently commercially available implementation in the quantum chemistry program package MOLPRO [13] by Nicklaß [14].

The goal of this work consists of two parts aiming to extend the current knowledge of the CPP ansatz and its application. First, a custom implementation of CPP ansatz will be incorporated into the quantum chemistry program package of the University of Cologne, namely the Quantum Objects Library [15]. This implementation will be tested in multiple test systems to ensure the correct behaviour both on the Hartree-Fock and electron-correlation level. Second, the influence of the CPP ansatz

in simulations of Hg_2 and HgF_4 will be investigated. For this, we will generate CPPs purely relying on *ab initio* data and employ them to calculate spectroscopic constants of the investigated systems. The results will be compared to experimental and theoretical studies to evaluate the influence of the CPPs.

2 Theory

2.1 Hartree-Fock (HF)

The Hartree-Fock (HF) method is one of the most fundamental wave function-based methods for simulating many-body systems by approximately solving the Schrödinger equation [16, 17]. The method uses two major approximations, the Born-Oppenheimer approximation and an approximated electron-electron interaction [18–20]. The goal of the Hartree-Fock method is to obtain approximate solutions of the Schrödinger equation in order to get insights into the properties of different chemical systems. Moreover it provides a starting point for higher level treatments including electron correlation. Since we focus exclusively on time-independent properties in this work, we only have to solve the time-independent form of the Schrödinger equation

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

in which $\hat{\mathcal{H}}$ is the Hamiltonian, $|\Psi\rangle$ the many-body wave function and E the energy eigenvalue.

2.1.1 Born-Oppenheimer approximation (adiabatic approximation)

The Born-Oppenheimer approximation [21] allows for treating the electrons of a system separately from the nuclei. For this, the electronic Hamiltonian $\hat{\mathcal{H}}_{\text{el}}$ in atomic units (a.u.) is used

$$\hat{\mathcal{H}}_{\text{el}} = -\frac{1}{2}\sum_{i=1}^{n}\Delta_i - \sum_{i=1}^{n}\sum_{A=1}^{N}\frac{Z_A}{r_{iA}} + \sum_{i<j}^{n}\frac{1}{r_{ij}} \ . \tag{2.2}$$

In this Hamiltonian the kinetic energy of the nuclei is neglected and the Coulomb repulsion between the nuclei is considered constant. This constant value is added to the energy eigenvalue after applying the Hamiltonian, i.e. after solving the electronic Schrödinger equation. In equation 2.2, the first term represents the kinetic energy of the electrons, the second term the Coulomb attraction between the electrons and nuclei and the last term the electron-electron interaction [22].

2.1.2 Slater determinant

For describing systems with multiple electrons, the wave function is approximated as a product of one-electron spin orbitals $|\chi\rangle$, termed the Hartree product. Due to the Pauli exclusion principle [23, 24], the wave function has to be antisymmetric with respect to the exchange of electrons. Since a single Hartree product cannot fulfill this requirement, an antisymmetrized linear combination of all possible Hartree products, i.e. a Slater determinant, is used:

$$|\Phi_{\text{Slater}}\rangle = \frac{1}{\sqrt{n!}} \sum_{k=1}^{n!} (-1)^{p_k} \hat{P}_k \prod_{i=1}^{n} |\chi_i(i)\rangle \ , \tag{2.3}$$

in which $\hat{P}_k$ is the permutation operator and $n!$ the number of all possible permutations of n electrons in n spin orbitals.

2.1.3 Hartree-Fock equation

In the closed-shell Hartree-Fock method, the approximated wave function is described by a single Slater determinant. The variational principle is used for finding the minimal energy and thus the optimal wave function by approximately solving the Schrödinger equation using the electronic Hamiltonian [22]. This principle states that the variation of an approximated wave function, in this case $|\Phi_{\text{Slater}}\rangle$, always leads to an energy which is greater than or equal to the exact ground-state energy E_0:

$$\min \left\{ \frac{\langle \Phi_{\text{Slater}} | \hat{\mathcal{H}}_{\text{el}} | \Phi_{\text{Slater}} \rangle}{\langle \Phi_{\text{Slater}} | \Phi_{\text{Slater}} \rangle} : \chi_i \right\} \geq E_0 \ . \tag{2.4}$$

Varying the spin orbitals χ_i, which are part of the wave function Φ_{Slater} (eq. 2.3), gives the optimized energy [22]. The variation principle and the expression of the single Slater determinant complete the description of the HF method; the energy can now be calculated by this method. For the evaluation of the integrals $\langle \Phi_{\text{Slater}} | \hat{\mathcal{H}}_{\text{el}} | \Phi_{\text{Slater}} \rangle$ and $\langle \Phi_{\text{Slater}} | \Phi_{\text{Slater}} \rangle$, the Slater-Condon rules [25–27] are commonly employed.

For the application of the Slater-Condon rules on the variation principle, the spin orbitals are required to be orthonormal while varying them. This orthonormality requirement can be fulfilled using the Lagrange method, which introduces the so-called Lagrange multiplier [22]. By varying the spin orbitals by an infinitesimal amount and setting the first variation of the Langrangian to be zero, the following

equations can be derived:

$$\left[\hat{h}(1) + \sum_{j=1}^{n}\left(\hat{\mathcal{J}}_j(1) - \hat{\mathcal{K}}_j(1)\right)\right]|\chi_i(1)\rangle = \epsilon_i\,|\chi_i(1)\rangle \ , \qquad i = 1, 2, \ldots, n \ . \tag{2.5}$$

Here, $\hat{\mathcal{J}}_j(1)$ is the Coulomb operator and $\hat{\mathcal{K}}_j(1)$ the exchange operator. The expression inside the brackets in equation 2.5 is the so-called Fock operator $\hat{f}$ and equation 2.5 can be expressed as

$$\hat{f}\,|\chi_i\rangle = \epsilon_i\,|\chi_i\rangle \ , \tag{2.6}$$

which is the canonical HF equation. This canonical HF equation is a pseudo-eigenvalue problem since the Fock operator, which operates on the one-particle space, is functionally dependent on the spin orbitals. This functional dependence renders the equation nonlinear such that is has to be solved in an iterative manner. Furthermore, the HF equation (or the HF equations, for simulating a many-electron system; eq. 2.6) can be solved exactly only for atoms [22] (and in some special cases for highly symmetric molecules [28–30]). To solve the HF equations for arbitrary molecules, a set of basis functions is introduced yielding the Roothaan-Hall equations [31, 32].

2.1.4 Roothaan-Hall equations: restricted Hartree-Fock (RHF)

To derive the Roothaan-Hall equations, the spin orbitals $|\chi\rangle$ are restricted, i.e. α and β spins are constrained to the same spatial function $\psi(r)$

$$|\chi(x)\rangle = \begin{cases} |\psi(r)\alpha(\omega)\rangle \\ |\psi(r)\beta(\omega)\rangle \end{cases} \ . \tag{2.7}$$

This restriction of the spin orbitals results in the spatial HF equation

$$\hat{f}\,|\psi_i\rangle = \epsilon_i\,|\psi_i\rangle \ , \tag{2.8}$$

with

$$\hat{f}(1) = \hat{h}(1) + \sum_{i}^{n/2}\left(2\hat{\mathcal{J}}_i(1) - \hat{\mathcal{K}}_i(1)\right) . \tag{2.9}$$

To solve equation (2.8), numerical solutions for molecules can be obtained by introducing a set of K basis functions and performing a linear expansion of the

form

$$|\psi_i\rangle = \sum_{\mu=1}^{K} c_{\mu i} |\phi_\mu\rangle \ , \qquad i = 1, 2, \ldots, K \ , \tag{2.10}$$

which leads to the HF equation

$$\sum_{\nu} c_{\nu i} \langle \psi_\mu | \hat{f} | \psi_\nu \rangle = \epsilon_i \sum_{\nu} c_{\nu i} \langle \psi_\mu | \psi_\nu \rangle \ . \tag{2.11}$$

or short, by using matrix notation for the entire set of basis functions,

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon} \ , \tag{2.12}$$

in which $\mathbf{F}$ is the Fock matrix and $\mathbf{S}$ the overlap matrix. $\mathbf{C}$ is a $K \times K$ matrix of the expansion coefficients $c_{\mu i}$ and $\boldsymbol{\epsilon}$ the diagonal matrix of the orbital energies ϵ_i.

These Roothaan equations can only be solved iteratively due to the functional dependency of the Fock operator. To obtain a usual matrix eigenvalue problem, the coefficient matrix $\mathbf{C}$ can be transformed such, that the overlap matrix $\mathbf{S}$ becomes the unit matrix, resulting in the following equation:

$$\tilde{\mathbf{F}}\tilde{\mathbf{C}} = \tilde{\mathbf{C}}\boldsymbol{\epsilon} \ . \tag{2.13}$$

This eigenvalue problem can be solved for the eigenvectors $\tilde{\mathbf{C}}$ and the eigenvalues $\boldsymbol{\epsilon}$ using the self-consistent field (SCF) method [22]. After the SCF procedure converges the total energy can be expressed as

$$E_{\mathrm{HF,tot}} = E_{\mathrm{HF}} + \sum_{A<B} \frac{Z_A Z_B}{R_{AB}} \ , \tag{2.14}$$

in which E_{HF} is the energy that can be calculated from the coefficients obtained in the final SCF iteration. The second term in equation 2.14 is the constant term of the nuclei repulsion, which is added after the SCF procedure as a consequence of the Born-Oppenheimer approximation (section 2.1.1). The choice of the basis set is of great importance, because the more complete the basis set, the more precise the result of the calculation. Hypothetically, choosing an infinite, i.e. complete, basis set, allows for obtaining the exact HF solution. At this so-called HF limit, the best possible HF energy E_{HF} is reached. This energy, however, differs from the exact energy E_{exact}, which is obtained when solving the Schrödinger equation exactly. This deviation is usually about 0.5 % of the state energy, which is not accurate enough

for describing chemical phenomena [33]. The "missing" energy, which is the reason for the high deviation, is often referred to as the correlation energy E_{corr}

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}} \ , \tag{2.15}$$

which is less than or equals to zero since the HF energy represents the upper boundary to the exact solution [22].

2.2 Electron correlation methods

The major goal of correlated methods is to improve the HF approximation by retrieving correlation energy. This can be achieved by using more than one Slater determinant for describing the wave function [22, 34].

2.2.1 Coupled cluster (CC)

The coupled cluster (CC) theory is a widely used wave function-based correlation method due to its size consistency, which is lacking in other wave function-based methods. In this ansatz, the exact ground state wave function is described as follows:

$$|\Psi_{\text{CC}}\rangle = e^{\hat{T}} |\Phi_0\rangle \ , \tag{2.16}$$

in which $\hat{T}$ is the so-called cluster operator and is generally defined as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \ldots + \hat{T}_n \tag{2.17}$$

$$\hat{T}_\nu = \frac{1}{(\nu!)^2} \sum_{\substack{a_1,\ldots,a_n \\ i_1,\ldots,i_n}} t_{i_1,\ldots,i_n}^{a_1,\ldots,a_n} \hat{a}_{a_1}^\dagger \cdots \hat{a}_{a_n}^\dagger \hat{a}_{i_n} \cdots \hat{a}_{i_1} \ . \tag{2.18}$$

$e^{\hat{T}}$ in equation 2.16 is the so-called exponential operator, and $\hat{a}_{a_n}^\dagger$ and $\hat{a}_{i_n}$ in equation 2.18 are the creation and annihilation operators, respectively, known from the theory of second quantization, while $t_{i_1,\ldots,i_n}^{a_1,\ldots,a_n}$ are the so-called amplitudes [35, 36].

The Schrödinger equation is approximated using the CC ansatz by projecting it onto the reference wave function resulting in the following expression:

$$E_{\text{CC}} = \langle \Phi_0 | \hat{\mathcal{H}} e^{\hat{T}} | \Phi_0 \rangle \ , \tag{2.19}$$

which can be expanded into the following form using spin orbitals, which were

obtained by a HF calculation:

$$E_{\mathrm{CC}} = E_{\mathrm{HF}} + \sum_{\substack{i<j \\ a<b}} (t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a)(\langle \chi_i \chi_j | \chi_a \chi_b \rangle - \langle \chi_i \chi_j | \chi_b \chi_a \rangle) \,. \tag{2.20}$$

This equation indicates that the correlation energy in the CC ansatz is entirely determined by the singles and doubles amplitudes and the two-electron molecular orbital (MO) integrals [36]. To obtain these amplitudes, a similarity transformation of the Hamiltonian is performed by employing $e^{-\hat{T}}$. By projecting onto substituted determinants, the following amplitude equations for different substitutions are obtained:

$$\begin{aligned} \langle \Phi_I^A | e^{-\hat{T}} \mathcal{H} e^{\hat{T}} | \Phi_0 \rangle &= 0 \,, \\ \langle \Phi_{IJ}^{AB} | e^{-\hat{T}} \mathcal{H} e^{\hat{T}} | \Phi_0 \rangle &= 0 \,, \\ \langle \Phi_{IJK}^{ABC} | e^{-\hat{T}} \mathcal{H} e^{\hat{T}} | \Phi_0 \rangle &= 0 \,, \\ &\vdots \quad . \end{aligned}$$

In general, optimizing the parameters is a nonlinear optimization problem of the following form:

$$\vec{\Omega}\left(\vec{t}\right) = \vec{0} \,, \tag{2.21}$$

which can be solved using different approaches [37, 38].

The full expansion of n substitutions, however, is computationally too demanding to be feasible. Usually, the expansion is truncated after the singles and doubles substitutions, resulting in the coupled cluster singles and doubles (CCSD) method. For improving the CCSD method, the triples substitutions can be perturbatively added, resulting in the CCSD(T) method, which is known as the "gold standard" of quantum chemistry [39]. This method is able to account for more than 99 % of the correlation energy [40]. Besides the high computational demands, the CCSD(T) method is a single reference method, rendering it not accurate for simulations requiring a multi-reference treatment. For multi-reference cases, one of the multitude of multi-reference CC (MRCC) approaches can be used [40–42].

2.3 Relativistic effects

2.3.1 The Dirac equation

Relativistic effects significantly influence the chemical properties of heavy elements and are thus important for understanding the nature of these elements [43–46]. The most well-known example for the contribution of relativistic effects is gold with its characteristic color, as well as its nobility, often unusual short bond lengths and possible trivalency [45, 47, 48]. Furthermore, properties like high oxidation states, high electron affinities, spin-spin coupling and contraction of bond lengths of heavy elements can only be explained by relativistic effects [48].

For describing relativistic effects in the most accurate way, the Dirac (D) one-particle Hamiltonian $\hat{h}_\mathrm{D}$ with

$$\hat{h}_\mathrm{D}(i) = c\hat{\vec{\alpha}}_i \cdot \hat{\vec{p}}_i + (\hat{\beta}_i - \mathbf{I}_4)c^2 + \sum_\lambda \hat{V}_\lambda(r_{i\lambda}) \; , \tag{2.22}$$

is employed in the following Hamiltonian to solve the Dirac equation within the Born-Oppenheimer approximation:

$$\hat{\mathcal{H}}_\mathrm{el} = \sum_i^n \hat{h}_\mathrm{D}(i) + \sum_{i<j}^n \hat{g}(i,j) + \sum_{\lambda<\mu}^N \frac{Z_\lambda Z_\mu}{r_{\lambda\mu}} \; . \tag{2.23}$$

In equation 2.22 the variable c represents the speed of light, $\hat{\vec{\alpha}}_i$ a vector of 4×4 Dirac matrices and $\hat{\beta}_i$ a single 4×4 Dirac matrix in the "standard representation" [49], $\mathbf{I}_4$ a 4×4 identity matrix and $\hat{\vec{p}}_i$ the momentum operator

$$\hat{\vec{p}}_i = -i\hat{\vec{\nabla}}_i \quad \text{with} \quad \hat{\vec{\nabla}}_i = \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i} \right) . \tag{2.24}$$

The last term in 2.22 represents the electrostatic potential generated by the λ-th nucleus as follows:

$$\hat{V}_\lambda(r_{i\lambda}) = -\frac{Z_\lambda}{r_{i\lambda}} \; . \tag{2.25}$$

The Dirac-Hamiltonian does not only describe the electron but also its counterpart, the positron. Thus, the relativistic wave function consists of four components, which are described by a four-component vector, the so-called four-spinor. The upper bispinor is large for electronic states, while the lower bispinor is large for positronic states. Due to the focus of quantum chemistry on electrons, the upper and lower

spinors are often referred to as large and small components, respectively [1].

The simplest way of choosing the two-particle interaction term $g(i,j)$ is the following:

$$\hat{g}_{\mathrm{C}}(i,j) = \frac{1}{r_{ij}}\,, \tag{2.26}$$

which is the non-relativistic Coulomb (C) interaction, leading to the so-called Dirac-Coulomb (DC) Hamiltonian. This form of the two-particle interaction term is sufficient, because the contribution of the one-particle Dirac-Hamiltonian to the relativistic effects is more dominant than the relativistic corrections to the Coulomb electron-electron repulsion [50]. For a more accurate description of the two-particle interaction, the frequency-independent Breit (B) interaction can be added, yielding the following term:

$$\hat{g}_{\mathrm{CB}}(i,j) = \frac{1}{r_{ij}} - \frac{1}{2r_{ij}}\left[\hat{\vec{\alpha}}_i \cdot \hat{\vec{\alpha}}_j + \frac{(\hat{\vec{\alpha}}_i \cdot \hat{\vec{r}}_{ij})(\hat{\vec{\alpha}}_j \cdot \hat{\vec{r}}_{ij})}{r_{ij}^2}\right]\,, \tag{2.27}$$

which accounts for the magnetic interaction of two electrons and the retardation of this interaction due to the finite velocity of light. Contributions of higher-order corrections, like the vacuum polarization, are often negligible, because their contribution to chemical properties is not significant [1]. Importantly, the Dirac-Hamiltonian is not bound from below, which may lead to a variational collapse during the energy-minimization. Furthermore, a mixing of electronic and positronic states can occur, which is known as Brown-Ravenhall continuum dissolution or Brown-Ravenhall disease [1, 51].

Both the DC- or DCB-Hamiltonian can be applied in methods, in which their non-relativistic counterparts are used (e.g in the Hartree-Fock equation resulting in the Dirac-Hartree-Fock (DHF)-equation). While solving this equation is also done self-consistently, the biggest differences compared to the non-relativistic SCF is due to the substitution of the one-component, the wave function, by the four-component vector, the spinor. These spinors should obey the kinetic balance conditions. Hence, the basis sets become significantly larger and the DHF method more computationally demanding than its non-relativistic counterpart [52]. Furthermore, the spinors are not optimized to result in a minimized, but in a stationary energy [49].

2.4 Effective core potentials

The general idea of effective core potential (ECP) methods is to describe the valence electrons independent of the atomic nucleus and the core electrons in order to reduce computational demands. For this, the atomic nucleus and the core electrons are combined to one potential, which interacts with the valence electrons. Hence, for describing the valence electrons explicitly, a valence-only (VO) model Hamiltonian can be written as follows:

$$\hat{\mathcal{H}}_{\mathrm{v}} = \sum_{i}^{n_{\mathrm{v}}} \hat{h}_{\mathrm{v}}(i) + \sum_{i<j}^{n_{\mathrm{v}}} \hat{g}_{\mathrm{v}}(i,j) + \hat{V}_{\mathrm{cpp}} + V_{\mathrm{cc}} \ , \tag{2.28}$$

in which c and v are abbreviations for core and valence electrons, respectively. n_v is the number of valence electrons, $\hat{V}_{\mathrm{cpp}}$ is the core polarization potential (CPP), which accounts for core-core and core-valence correlations, while V_{cc} represents the repulsion of the cores and nuclei in the system. The valence-only model Hamiltonian of the following form:

$$\hat{h}_{\mathrm{v}}(i) = -\frac{1}{2}\hat{\vec{\nabla}}_i^2 + \hat{V}_{\mathrm{cv}}(i) \quad \text{and} \quad \hat{g}_{\mathrm{v}}(i,j) = \frac{1}{r_{ij}} \ , \tag{2.29}$$

can be used for non-relativistic, scalar-relativistic and quasi-relativistic ECPs [1]. It is usually assumed that all relativistic effects can be considered by the parameterization of the effective core potential $\hat{V}_{\mathrm{cv}}$. This assumption holds true, only if the non-relativistic form of the kinetic energy operator and the Coulomb interaction are sufficient for describing the system of valence electrons. Additionally, the ECP $\hat{V}_{\mathrm{cv}}$ accounts for all interactions of the valence electrons with the nucleus and the removed core electrons and ensures the core valence-orthogonality constraints [1].

For creating computationally practical VO schemes, approximations have to be made. One of these approximations is to assume that the core and valence electrons can be separated, which is only feasible on an independent-particle level, i.e. on the Hartree-Fock level. This approximation, however, is not exact on a correlated level of theory and thus the core-valence correlation is neglected in correlation methods. This correlation can be taken into account by CPPs, which are widely used, especially for large-core ECPs [1].

Another approximation made for creating VO schemes is the frozen-core approximation, in which the core orbitals are frozen for a special state on an independent-particle level. This means that the atomic cores are assumed to be inert leading to the frozen-

core (FC) approximation. Choosing the size of the core is a crucial part in this approximation. Employing a small core results in more accurate but more computationally demanding results, while large core simulations are more efficient but less precise. To eliminate the core electrons and core orbitals form the calculation, the contributions of the valence electrons in the all-electron FC Hamiltonian are replaced by non-local HF potentials. Finally, the atomic cores are assumed to interact only by their Coulomb repulsion:

$$V_{\mathrm{cc}} = \sum_{\lambda<\nu}^{N} \frac{Q_\lambda Q_\nu}{r_{\lambda\nu}} \, . \tag{2.30}$$

If this assumption is violated by the cores overlapping or penetrating each other, V_{cc} has to be corrected to account for deviations from the point charge approximation [1].

Since the development of the ECP theory, various schemes for generating ECPs have been proposed [1, 53–56]. However, only two of these schemes, namely the pseudopotential (PP) and model potential (MP) approaches, are widely used in the literature. In both schemes, the VO Hamiltonian is fitted to all-electron (AE) calculations. The MP approach uses valence orbitals with an identical nodal structure to that of the AE valence orbitals. In contrast, the PP methods use a simpler nodal structure by employing a pseudovalence orbital transformation [55], which renders PP less computationally demanding. While the PP method is less accurate than the MP method, the accuracy of the PPs is sufficient enough for an accurate description of molecular systems. The reduced computational demands often render the PPs superior to the MPs and are thus the most widely used ECPs in the literature [1]. Hence, only PPs will be discussed in the following.

2.4.1 Pseudopotentials

For generating PPs, which are both accurate and reduce the computational demands, the semilocal PP ansatz is employed [57, 58]. For this, the semilocal PP for a core λ of the following form is used:

$$\Delta\hat{V}_{\mathrm{cv}}^{\lambda}(i) \cong \Delta\hat{V}_{\mathrm{PP}}^{\lambda}(i) = V_L^{\lambda}(r_{i\lambda}) + \sum_{l=0}^{L-1}[V_l^{\lambda}(r_{i\lambda}) - V_L^{\lambda}(r_{i\lambda})]\hat{P}_l^{\lambda}(i) \, , \tag{2.31}$$

in which $\hat{P}_l^{\lambda}(i)$ is the angular momentum projection operator based on the spherical harmonics $|lm,\lambda\rangle$ with respect to the core angular momentum l. The semilocal PP consists of a sum of local potentials $V_l^{\lambda}(r_{i\lambda})$ acting on the angular momentum symmetries and a common local potential $V_L^{\lambda}(r_{i\lambda})$ [1]. In practice, the L of the

local potential $V_L^\lambda(r_{i\lambda})$ is chosen such that $(L-1)$ represents the largest angular momentum used by the core orbitals [58].

Furthermore, the spin-orbit (SO) interaction is included resulting in a modified analytic form of the PPs [59]. For this, the PP needs to be dependent on the angular momentum j of the orbital leading to an lj-dependent semilocal PP. This PP can be written as follows:

$$\Delta\hat{V}_{\mathrm{PP}}^\lambda(i) = \Delta\hat{V}_{\mathrm{PP,av}}^\lambda(i) + \Delta\hat{V}_{\mathrm{PP,so}}^\lambda(i) \ , \tag{2.32}$$

in which $\hat{V}_{\mathrm{PP,av}}^\lambda(i)$ is a scalar-relativistic PP and $\Delta\hat{V}_{\mathrm{PP,so}}^\lambda$ is the SO PP of the following form:

$$\Delta\hat{V}_{\mathrm{PP,so}}^\lambda(i) = \sum_{l=1}^{L-1} \frac{\Delta V_l^\lambda(r_{i\lambda})}{2l+1}[l\hat{P}_{l,l+1/2}^\lambda(i) - (l+1)\hat{P}_{l,l-1/2}^\lambda(i)] \ , \tag{2.33}$$

in which ΔV_l^λ is the difference between the corresponding relativistic PPs and $\hat{P}_{lj}^\lambda(i)$ being the lj-dependent projection operator as follows:

$$\hat{P}_{lj}^\lambda(i) = \sum_{m=-j}^{m=j} |ljm,\lambda\rangle \langle ljm,\lambda| \ . \tag{2.34}$$

For efficient use in combination with Gaussian basis sets, PPs are constructed by linear combinations of radial Gaussian functions multiplied by the electron core distance. The PPs are transformed into the following form:

$$V_m^\lambda = \sum_k A_{km}^\lambda r_{i\lambda}^{n_{km}} e^{-a_{km}^2 r_i^2 \lambda} \ . \tag{2.35}$$

These pseudovalence orbitals and the corresponding PPs, however, are not unique [60]. Two approaches to make them unique are the shape-consistent and the energy-consistent pseudopotentials [1]. In this work, we only focus on only energy-consistent pseudopotentials.

Energy-consistent pseudopotentials

The energy-consistent *ab initio* PP approach is based on the semi-empirical PP approach, in which the free PP parameters are fitted to experimental data [54]. Due to various difficulties using experimental data for parametrizing the PP, modern PPs are fitted to *ab initio* data, which is justified by the findings of Topp and Hopfield [61]. In their work they state that if the eigenvalues of the ground and excited states agree for any angular momentum, the logarithmic derivatives of the wave function will also be correct. Hence, energy-consistent PPs are additionally shape-consistent, i.e. the

orbital energies and the shape of the valence orbitals in the spatial valence region agree with the all-electron valence orbitals. Energy-consistent pseudopotentials were used throughout this work, mainly the Stuttgart-Cologne energy-consistent PPs [62].

The parametrization of energy-consistent *ab initio* PPs is commonly performed on two types of relativistic theory. One approach employs the scalar-relativistic one-component Wood-Boring Hamiltonian for generating the AE reference energies of multiple electron configurations [63, 64]. The second approach uses the relativistic four-component Dirac-Coulomb Hamiltonian, with or without the inclusion of the Breit term [65]. In general, the AE reference and the PP adjustment calculations are performed on a finite grid within a finite-difference method, which avoids basis set incompleteness. In both approaches the functional of the following form:

$$S = \sum_I w_I (E_I^{\mathrm{PP}} - E_I^{\mathrm{AE}} + \Delta E_{\mathrm{shift}})^2 \,, \tag{2.36}$$

is minimized. The summation runs over all configurations, LS states and/or J levels [1]. E_I^{AE} and E_I^{PP} represent the total valence energies, i.e. the energy of the Ith reference energy subtracted by the core electron energy, for the AE and the PP calculations, respectively. w_I are weighting factors, which represent the non-relativistic configuration independent from the number of LS states or J levels. $\Delta E_{\mathrm{shift}}$ can be used for relaxing the PP and AE total energies and thus allows for a better fit [1].

2.5 Core polarization

As mentioned above, ECPs are able to reduce computational demands, while incorporating relativistic effects into simulations. However, due to the core-valence separation and the frozen core approximation, the polarization of the core is neglected. For nearly 100 years [66], the core polarization was studied to understand the influence of the core polarization in quantum chemical calculations and how to incorporate this phenomenon with little computational effort.

2.5.1 Polarization of the atomic core

Born and Heisenberg were the first to describe the interaction energy of the valence electron of a system with its core [66]. Using a classical approach, they derived the following equation for the interaction energy:

$$E_{\mathrm{pol}} = -\frac{\alpha}{2r^4} \,, \tag{2.37}$$

in which α is the dipole polarizability of the core. Here, a singularity of the expression occurs in the vicinity of the core. Besides this singularity, they were able to describe the interaction of one valence electron with a noble gas core (Na, Mg^+, Al^{2+}, etc.) in ionic crystals [66], indicating the importance of the core polarizability in such systems.

2.5.2 Limitations of effective core potentials

Since the introduction of the effective core potential approach [57], it has become the method of choice for reducing computational demands in non-relativistic simulations, while maintaining accuracy [67–69]. However, due to the core-valence separation and the frozen core approximation [70] it became obvious that for certain chemical properties an accurate description of the core polarization is needed [6].

The work of Bauschlicher *et al.* [71] demonstrated how in frozen core AE simulations the missing contributions of the core polarization can be retrieved. For this, they chose a MCSCF/MRCI method to reproduce FCI simulations of Be, C, CH^+ for various excitations energies and spectroscopic constants. By comparing both approaches, they investigated the missing correlation. For this, they separated the missing correlation energy into two parts, the core-valence and core-core correlation. The former describes the correlation between valence electrons with core electrons of all centers, while the latter describes the correlation between core electrons of all centers with each other. The authors demonstrated that the inclusion of single substitutions, starting from the core electrons in the valence space, could contribute for the majority of the core-valence correlation, if the core-core correlation is neglected. The core-core correlation itself, however, could not be described correctly by this method.

While the contributions of the core polarization can be incorporated into AE simulations, as shown by Bauschlicher *et al.*, this is not the case for ECP simulations, because no substitutions out of the core space can be performed. Hence, other approaches are needed.

2.5.3 Core polarization in pseudopotential simulations

Core polarization in single-electron systems

The first work to explicitly incorporate core-polarization into PP simulation was performed for the potential energy curves for Li_2^+ and Li_2 [72]. Here, an effective potential of the form:

$$V_{\text{eff}}(r) = -\frac{1}{r} - \frac{\alpha_{\text{D}}}{2\left(r^2+d^2\right)^2} - \frac{\alpha_{\text{Q}}}{2\left(r^2+d^2\right)^3} + V_{\text{SR}}(r)\ , \tag{2.38}$$

was chosen, in which the second and third term contribute to the dipole α_{D} and quadrupole α_{Q} polarizability of the core. The V_{SR} term represents the potential of the effective core potential ansatz, which was fitted via *ab initio* calculations. Additionally, the cutoff parameter d was introduced to account for the divergence of the potential near the nucleus (see section 2.5.1). With this approach it was possible to reproduce the dissociation energies and bond distances for the two molecular systems, when compared to other *ab initio* simulations [72].

To improve upon the ansatz by Bardsley *et al.* (equation 2.38), Fuentealba *et al.* [4, 5, 73] used one of the proposed cutoff functions by Müller *et al.* [2, 3] of the following form:

$$\omega(r,\gamma) = \left(1 - e^{-\left(\frac{r}{\gamma}\right)^2}\right)^2\ , \tag{2.39}$$

in which, γ is the cutoff factor, which was fitted to experimental atomic ionization energies for Li, Na and K. The corresponding polarization potential V_{pol} was chosen

$$V_{\text{pol}}(r,R) = -\omega(r,\gamma)\frac{\alpha_{\text{D}}}{2r^4} + \omega(r,\gamma)\frac{\alpha_{\text{D}} r\cdot R}{r^3R^3} - \frac{\alpha_{\text{D}}}{2R^4}\ , \tag{2.40}$$

in which r and R are the distances between the core and the valence electron and the core and another nucleus/core, respectively.
This ansatz incorporates three different contributions to the core polarization: the induced dipole by the single electron, the single nucleus and the combination of the single electron and the nucleus (Fig. 2.1). Employing this ansatz results in an accurate description of the dimer and hydride ions of the elements Li, Na and K. This made the description of core polarization in simulations of single electron-systems using the framework of effective core potentials possible.

This work by Fuentealba *et al.* represents the first published application of the

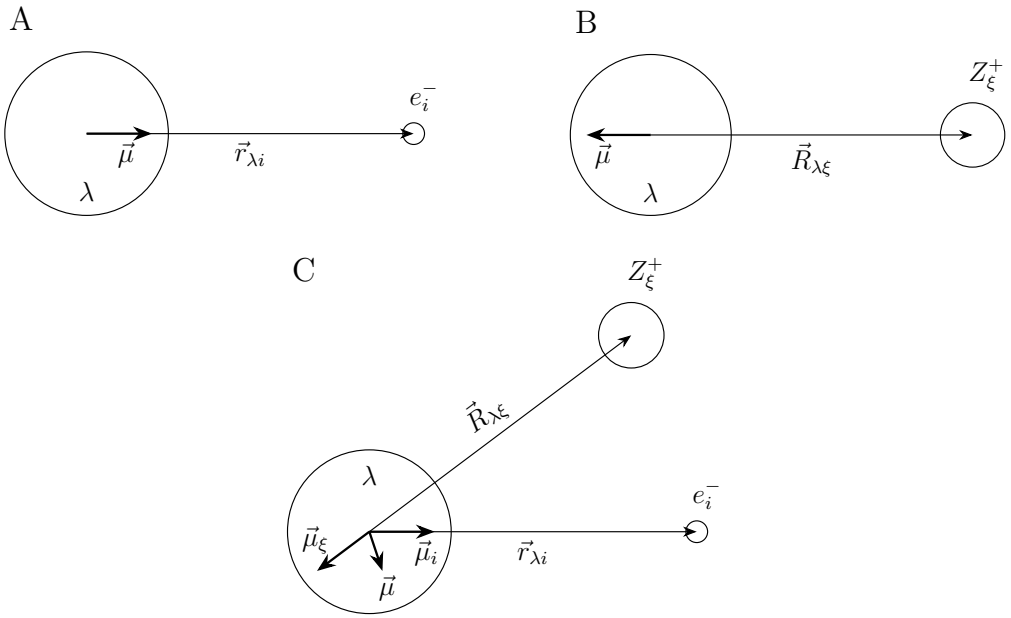

Fig. 2.1: Graphical representation of the induced dipole $\vec{\mu}$ at a core λ by a single electron (A), a single nucleus/core (B), and an electron and a nucleus/core (C).

core polarization potential (CPP) ansatz by Müller *et al.* [2, 3].

Multi-electron ansatz

The work of Fuentealba *et al.* [6, 73] pointed out, that for multiple-electron systems it is important to incorporate terms of two electrons and two nuclei inducing a dipole moment at a given core (Fig. 2.2).

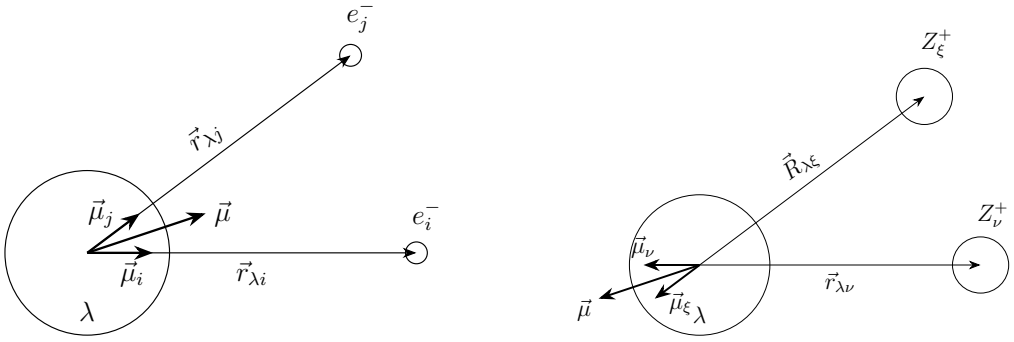

Fig. 2.2: Graphical representation of the induced dipole $\vec{\mu}$ at a core λ by two electrons (left) and two nuclei/cores (right).

This ansatz represents the fundamental work towards the incorporation of core polarization beyond alkali and earth alkali elements for all main group elements [6, 7, 74], transition elements (see e.g. [8, 9]) and f-elements (see e.g. [75]) using only

the dipole polarizability of the atomic core.

l-dependent cutoff functions

One way to improve upon the CPP ansatz was proposed by Foucrault *et al.* [76] by using a cutoff function, which uses a projection operator on the spherical harmonics, dependent on the angular momentum l of the following form:

$$\omega\left(r_\lambda, \gamma_\lambda\right) = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l} F_l(r_\lambda, \gamma_\lambda^l) \left|lm\lambda\right\rangle \left\langle lm\lambda\right| \ , \tag{2.41}$$

with

$$F_l\left(r_\lambda, \gamma_\lambda^l\right) = \begin{cases} 0 & \text{for} \quad r < \gamma_\lambda^l \\ 1 & \text{for} \quad r \geq \gamma_\lambda^l \end{cases} \quad . \tag{2.42}$$

With this ansatz highly, accurate simulations of alkali compounds can be performed. For example, the approach was recently employed to simulate alkali-noble gas dimers [77, 78] or alkali hydrides [79].

Core polarization beyond the dipole polarizability

An alternative to the above described approaches by Fuentealba/Müller and Foucrault *et al.*, which introduced a potential operator to account for the core polarization, is a second order perturbation theory ansatz [80]. One advantage of this ansatz is that no cutoff function has to be introduced, because no $\frac{1}{r^4}$ operator is added to the Hamiltonian [81, 82]. Dalgarno *et al.* incorporated not only contributions of the dipole moment, but also higher multipole and adiabatic corrections [80]. Using a similar approach, it was possible to accurately describe core-valence correlation in alkali dimers and determine spectroscopic properties in better agreement with experimental values, when compared with the approach by Bardsley *et al.* [72] (eq. 2.38).

2.5.4 Generalization of the CPP ansatz

While the fundamental work of Fuentealba/Müller allowed for incorporating core polarization via the dipole polarizability, the CPP ansatz was not applicable to general molecular systems and other multipole polarizabilities. The work of Schwerdtfeger and Silberach [12, 83, 84] and Nicklaß [14] generalizes the CPP ansatz and allows to perform simulations beyond single- and two-electron systems. Furthermore, in this framework not only the dipole polarizability α_D, but also the quadrupole polarizability

α_{Q} and the adiabatic correction to the dipole polarizability β_{D} are considered. Here, the electric field of a single electron $\vec{f}_\lambda$ is described as:

$$\vec{f}_\lambda = \sum_i \frac{\vec{r}_{\lambda i}}{r_{\lambda i}^3}\omega\left(r_{\lambda i}, \gamma_\lambda\right) - \sum_{\mu\neq\lambda} \frac{\vec{R}_{\lambda\mu}}{R_{\lambda\mu}^3} Q_\mu \ , \tag{2.43}$$

in which cutoff functions of the Müller-type are employed [2, 3]. For incorporating the multipole and adiabatic corrections, the following potentials are added to the Hamiltonian:

$$V_{\mathrm{di}} = -\frac{1}{2}\sum_\lambda \alpha_{\mathrm{D}}^\lambda \vec{f}_\lambda^{\,2} \ , \tag{2.44}$$

$$V_{\mathrm{quad}} = -\frac{1}{12}\sum_\lambda \alpha_{\mathrm{Q}}^\lambda \sum_{A\in\{X,Y,Z\}} \left(\frac{\partial \vec{f}_\lambda}{\partial A_\lambda}\right)^2 \ , \tag{2.45}$$

$$V_{\mathrm{adia}} = \sum_i \sum_\lambda \frac{3\beta_{\mathrm{D}}^\lambda}{r_{\lambda i}^6} \ . \tag{2.46}$$

By substituting $\vec{f}_\lambda$ with the expression in equation 2.43 for the core-core interaction Ω, the single-electron V' and the two-electron V'' term can be obtained:

$$\begin{aligned}
\Omega = -\frac{1}{2}\sum_\lambda \Bigg\{ & \alpha_{\mathrm{D}}^\lambda \sum_{\mu\neq\lambda} \frac{Q_\mu^2}{R_{\lambda\mu}^4}\omega^4(R_{\lambda\mu},\gamma_\lambda) \\
& + \alpha_{\mathrm{Q}}^\lambda \sum_{\mu\neq\lambda} \frac{Q_\mu^2}{R_{\lambda\mu}^6}\omega^6(R_{\lambda\mu},\gamma_\lambda) \\
& + 2\alpha_{\mathrm{D}}^\lambda \sum_{\substack{\mu<\nu \\ \nu\neq\lambda}} \frac{\vec{R}_{\lambda\mu}\cdot\vec{R}_{\lambda\nu}}{R_{\lambda\mu}^3 R_{\lambda\nu}^3} Q_\mu Q_\nu \omega^2(R_{\lambda\mu},\gamma_\lambda)\omega^2(R_{\lambda\nu},\gamma_\lambda) \\
& + \alpha_{\mathrm{Q}}^\lambda \sum_{\substack{\mu<\nu \\ \nu\neq\lambda}} \frac{Q_\mu Q_\nu}{R_{\lambda\mu}^3 R_{\lambda\nu}^3} \left[\frac{3(\vec{R}_{\lambda\mu}\cdot\vec{R}_{\lambda\nu})^2}{R_{\lambda\mu}^2 R_{\lambda\nu}^2} - 1\right] \omega^3(R_{\lambda\mu},\gamma_\lambda)\omega^3(R_{\lambda\nu},\gamma_\lambda) \Bigg\}
\end{aligned} \tag{2.47}$$

$$
\begin{aligned}
V' = -\frac{1}{2}\sum_i\sum_\lambda \Bigg\{ & \alpha_{\mathrm{D}}^\lambda \left[\frac{1}{r_{\lambda i}^4}\omega^4(r_{\lambda i},\gamma_\lambda) - 2\left(\frac{x_{\lambda i}}{r_{\lambda i}^3}B_X^\lambda + \frac{y_{\lambda i}}{r_{\lambda i}^3}B_Y^\lambda + \frac{z_{\lambda i}}{r_{\lambda i}^3}B_Z^\lambda \right)\omega^2(r_{\lambda i},\gamma_\lambda) \right] \\
& + \left(\alpha_{\mathrm{Q}}^\lambda - 6\beta_{\mathrm{D}}^\lambda\right)\frac{1}{r_{\lambda i}^6}\omega^6(r_{\lambda i},\gamma_\lambda) \\
& + \alpha_{\mathrm{Q}}^\lambda \left[\frac{1}{r_{\lambda i}^3}D^\lambda - 3\left(\frac{x_{\lambda i}^2}{r_{\lambda i}^5}C_{XX}^\lambda + \frac{y_{\lambda i}^2}{r_{\lambda i}^5}C_{YY}^\lambda + \frac{z_{\lambda i}^2}{r_{\lambda i}^5}C_{ZZ}^\lambda \right) \right. \\
& \left. - 6\left(\frac{x_{\lambda i}y_{\lambda i}}{r_{\lambda i}^5}C_{XY}^\lambda + \frac{x_{\lambda i}z_{\lambda i}}{r_{\lambda i}^5}C_{XZ}^\lambda + \frac{y_{\lambda i}z_{\lambda i}}{r_{\lambda i}^5}C_{YZ}^\lambda \right) \right] \omega^3(r_{\lambda i},\gamma_\lambda) \Bigg\}
\end{aligned} \tag{2.48}
$$

$$
\begin{aligned}
V'' = -\sum_{i<j}\sum_\lambda \Bigg\{ & \alpha_{\mathrm{D}}^\lambda \left[\frac{x_{\lambda i}}{r_{\lambda i}^3}\cdot\frac{x_{\lambda j}}{r_{\lambda j}^3} + \frac{y_{\lambda i}}{r_{\lambda i}^3}\cdot\frac{y_{\lambda j}}{r_{\lambda j}^3} + \frac{z_{\lambda i}}{r_{\lambda i}^3}\cdot\frac{z_{\lambda j}}{r_{\lambda j}^3} \right]\omega^2(r_{\lambda i},\gamma_\lambda)\omega^2(r_{\lambda j},\gamma_\lambda) \\
& + \frac{1}{2}\alpha_{\mathrm{Q}}^\lambda \Bigg[-\frac{1}{r_{\lambda i}^3}\cdot\frac{1}{r_{\lambda j}^3} + 3\left(\frac{x_{\lambda i}^2}{r_{\lambda i}^5}\cdot\frac{x_{\lambda j}^2}{r_{\lambda j}^5} + \frac{y_{\lambda i}^2}{r_{\lambda i}^5}\cdot\frac{y_{\lambda j}^2}{r_{\lambda j}^5} + \frac{z_{\lambda i}^2}{r_{\lambda i}^5}\cdot\frac{z_{\lambda j}^2}{r_{\lambda j}^5} \right) \\
& + 6\left(\frac{x_{\lambda i}y_{\lambda i}}{r_{\lambda i}^5}\cdot\frac{x_{\lambda j}y_{\lambda i}}{r_{\lambda j}^5} + \frac{x_{\lambda i}z_{\lambda i}}{r_{\lambda i}^5}\cdot\frac{x_{\lambda j}z_{\lambda i}}{r_{\lambda j}^5} + \frac{y_{\lambda i}z_{\lambda i}}{r_{\lambda i}^5}\cdot\frac{y_{\lambda j}z_{\lambda i}}{r_{\lambda j}^5} \right) \Bigg] \\
& \cdot \omega^3(r_{\lambda i},\gamma_\lambda)\omega^2(r_{\lambda j},\gamma_\lambda) \Bigg\}
\end{aligned} \tag{2.49}
$$

Here, B_A, C_{AB} and D ($A, B \in \{X, Y, Z\}$) are geometric constants:

$$
B_A^\lambda = \sum_{\mu\neq\lambda} \frac{Q_\mu A_{\lambda\mu}}{R_{\lambda\mu}^3}\omega^2(R_{\lambda\mu},\gamma_\lambda) \tag{2.50}
$$

$$
C_{AB}^\lambda = \sum_{\mu\neq\lambda} \frac{Q_\mu A_{\lambda\mu}B_{\lambda\mu}}{R_{\lambda\mu}^5}\omega^3(R_{\lambda\mu},\gamma_\lambda) \tag{2.51}
$$

$$
D^\lambda = \sum_{\mu\neq\lambda} \frac{Q_\mu}{R_{\lambda\mu}^3}\omega^3(R_{\lambda\mu},\gamma_\lambda) \tag{2.52}
$$

Ω, V' and V'' represent the general analytic form of the CPP ansatz [2, 3], for which a matrix-representation was derived by Schwerdtfeger and Silberbach for the use in modern quantum chemical program packages [12, 83, 84]. In their work, they showed, that employing Müller-type cutoff functions allows for solving the matrix elements

$$
\langle\phi|x^{m_1}y^{m_2}z^{m_3}|\vec{r}|^{-k}|\phi\rangle \quad (m_i \in \mathbb{N}_0, k \in \mathbb{Z}), \tag{2.53}
$$

which are in general not solvable for arbitrary k greater than two. Their work is of fundamental importance, because solving matrix elements for V' (eq. 2.48) and V'' (eq. 2.49) involves values of k which are greater than two. This ansatz of solving the CPP matrix elements was used for a general implementation of the CPP ansatz in the program package MOLPRO [13] (and MELD [85–88]) by Nicklaß [14]. The present work is based on the implementation of Nicklaß following the ansatz and assumptions made by Schwerdtfeger and Silberbach to solve the matrix-representation of the CPP integrals and the numerical approximations for solving the mathematical functions within.

Results and Discussion

3 Implementation of the CPP ansatz

3.1 Quantum Objects Library (QOL)

The Quantum Objects Library (QOL) [15] is a software library, which was designed and implemented by M. Hanrath in 2006. Since then, the QOL provides a means for implementing various *ab initio* methods and tools written in the programming language C++. For this work, the two most important parts of the QOL are (i) the implementation [89, 90] and optimization [91, 92] of the closed-shell Hartree-Fock program, namely the SCF, and (ii) the evaluation of the ARGOS [93–95] pseudopotential integrals via a wrapper [96]. Additionally, the SCF can be combined with the spin-orbital coupled cluster program of M. Hanrath, which is a separate program, used to perform correlation calculations with arbitrary substitution levels. For enabling the use of CPPs in the QOL we followed the implementation by Nicklaß [14]. For this, the SCF program was extended with the evaluation of additional integrals needed for the CPPs.

3.2 Modifying the QOL SCF

In general, the SCF code of the QOL solves the general Roothaan-Hall eigenvalue problem (eq. 2.8) by employing a modified [97–99] Obara-Saika scheme [100, 101] for the evaluation of the SCF integrals [102]. In addition to the SCF integrals, pseudopotential integrals are evaluated using the pseudopotential routine of the ARGOS code. These pseudopotential integrals are represented by a matrix $\mathbf{H}^{\mathrm{PP}}$, which is added to the core Hamiltonian matrix $\mathbf{H}^{\mathrm{core}}$ during the SCF procedure (Fig. 3.1).

In this work, we follow the design principle of the pseudopotential routine. First, we evaluate the one- and two-electron CPP integrals and represent them as one- and two-electron matrices, namely $\mathbf{V}'$ and $\mathbf{V}''$, respectively. After that, the matrices are combined with the respective integral matrices during the SCF procedure (Fig. 3.1). Thus, the incorporation of the CPP integrals is straightforward and there is no need to modify the complex and efficient integral code. Additionally, the generated matrices can be compared to the matrices in the MOLPRO implementation, analogous to the one-electron matrices without CPP contributions as previously described [91].

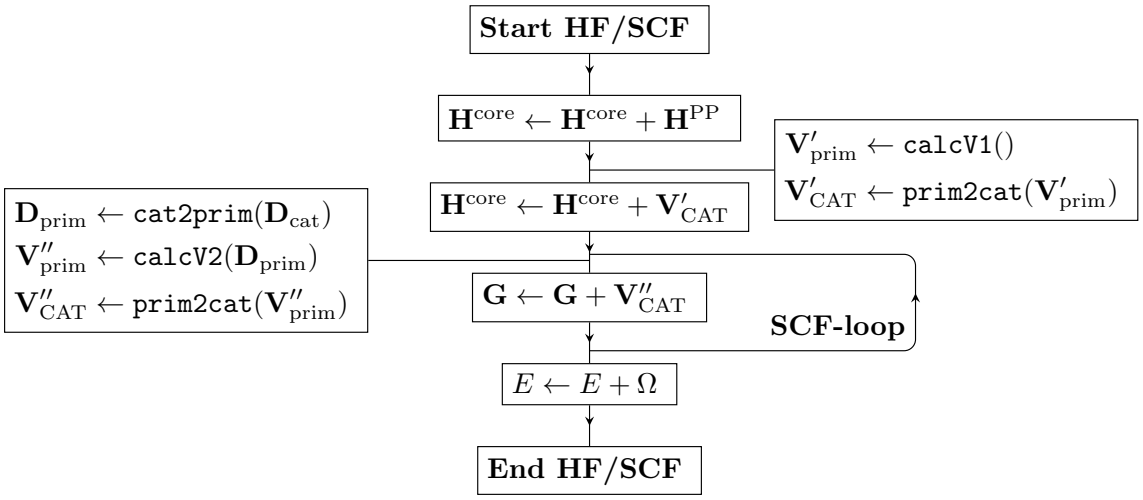

Fig. 3.1: Schematic of the modified SCF for incorporating the CPP contributions $\mathbf{V}'$, $\mathbf{V}''$ and Ω.

Importantly, the CPP matrices have to be represented in the same basis before they are combined with the corresponding SCF integral matrices. Note, that the QOL uses different basis representations for integrals and matrices - the primitive basis (prim), in which primitive Cartesian Gaussians (PCGs) [103] are used, and the contracted angular transformed (CAT) basis described by J. Held [91]. For the evaluation and thus representation of the CPP integrals, we use PCGs in accordance with the original work of Nicklaß [14]. To combine the CPP integrals with the SCF integrals, we therefore have to transform them into the CAT basis (Fig. 3.1). Importantly, the density matrix $\mathbf{D}$ is transformed into the primitive basis before it is used for the density contraction for the two-electron CPP matrix $\mathbf{V}''$. All the transformations (`prim2cat` and `cat2prim`) are performed with the `BasisTransformation`-Class written by J. Held [91]. To incorporate the nucleus repulsion energy contribution Ω into the simulation, we added it after the converged SCF procedure.

3.3 The two-electron matrix $\mathbf{V}''$

Setting up the $\mathbf{V}''$ matrix is computationally the most demanding component of the CPP implementation. The $\mathbf{V}''$ matrix is the result of a density contraction, which is analogous to the density contraction of the SCF integrals. The latter results in the two-electron matrix $\mathbf{G}$ by combining the density matrix $\mathbf{D}$ and the electron repulsion integrals (ERIs) $\langle ik|jl\rangle$ and $\langle ik|lj\rangle$ via

$$G_{ij} = \sum_{kl} D_{kl} \left[\langle ik|jl\rangle - \frac{1}{2} \langle ik|lj\rangle \right] . \tag{3.1}$$

Analogously, setting up the two-electron CPP matrix $\mathbf{V}''$ involves a density contraction, but instead of the ERIs we use the CPP two-electron integrals $\langle ik|\hat{v}''|jl\rangle$ and $\langle ik|\hat{v}''|lj\rangle$ via

$$V''_{ij} = \sum_{kl} D_{kl} \left[\langle ik|\hat{v}''|jl\rangle - \frac{1}{2} \langle ik|\hat{v}''|lj\rangle \right] . \tag{3.2}$$

Constructing both matrices according to the same equation allows for combining them correctly via matrix-matrix addition (Fig. 3.1).

Listing 3.1: Pseudocode for a straightforward implementation of the density contraction.

```
1 init V''
2 loop over ijkl
3     ⟨ik|v̂''|jl⟩ ← calcIntegral()
4     ⟨ik|v̂''|lj⟩ ← calcIntegral()
5     V''_ij      ← V''_ij + D_kl · ⟨ik|v̂''|jl⟩
6     V''_ij      ← V''_ij − ½ D_kl · ⟨ik|v̂''|lj⟩
7 end
8 return V''
```

A straightforward way of implementing the density contraction according to equation 3.2 is to use a loop over all four summation indices of the two-electron integrals $\langle ik|v''|jl\rangle$ and $\langle ik|v''|lj\rangle$, adding them with the corresponding prefactors (1 and $-\frac{1}{2}$) to the $\mathbf{V}''$ matrix (Listing 3.1). Here, `calcIntegral` represents a function for calculating the corresponding two-electron CPP integral. In this implementation, the integrals are calculated "on the fly", i.e. no integrals are saved and reused, but instead evaluated when needed. With the loop over the four indices (i,j,k and l), the runtime scaling of the implementation is $\mathcal{O}(N^4)$, where N is the number of basis functions. For reducing this runtime scaling, the two-electron integrals can be separated into multiple products of two one-electron integrals [14]. With this, the two-electron integrals can be expressed as

$$\langle ik|\hat{v}''|jl\rangle = \sum_{\lambda} \sum_{t} c_{\lambda t} \cdot \langle i|\hat{v}^{\lambda t}|j\rangle \cdot \langle k|\hat{v}^{\lambda t}|l\rangle \ , \tag{3.3}$$

in which t describes different types of one-electron integrals in the CPP ansatz, λ

the center in the molecular system and $c_{\lambda t}$ a center- and integral-specific prefactor. In this work, the integrals of the type $\langle i|\hat{v}^{\lambda t}|j\rangle$ are not calculated "on the fly", but stored in a matrix for easy access. Hence, we introduce the notation

$$\langle i|\hat{v}^{\lambda t}|j\rangle = v_{ij}^{\lambda t} \; . \tag{3.4}$$

Finally, equation 3.2 can be rewritten as:

$$V_{ij}'' = \sum_{\lambda t} c_{\lambda t} V_{ij}''^{\lambda t} \tag{3.5}$$

$$V_{ij}''^{\lambda t} = \sum_{kl} D_{kl} \cdot v_{ij}^{\lambda t} \cdot v_{kl}^{\lambda t} - \tfrac{1}{2} \sum_{kl} D_{kl} \cdot v_{il}^{\lambda t} \cdot v_{kj}^{\lambda t} \; . \tag{3.6}$$

With this, we express the density contraction using the one-electron integrals without changing the scaling. To finally reduce the runtime scaling, we have to rearrange the above equation. For this, we first rearrange the first term on the right-hand side of equation 3.6. Here, the indices of $v_{ij}^{\lambda t}$ are independent of the summation indices k and l and thus independent of the summation.

$$V_{ij}''^{\lambda t} = v_{ij}^{\lambda t} \sum_{kl} D_{kl} \cdot v_{kl}^{\lambda t} - \tfrac{1}{2} \sum_{k} \sum_{l} D_{kl} \cdot v_{il}^{\lambda t} \cdot v_{kj}^{\lambda t} \; . \tag{3.7}$$

Introducing the intermediate scalar $X^{\lambda t}$ for the summation with the indices k and l, we can rewrite equation 3.7 as:

$$V_{ij}''^{\lambda t} = v_{ij}^{\lambda t} \cdot X^{\lambda t} - \frac{1}{2} \sum_{k} \sum_{l} D_{kl} \cdot v_{il}^{\lambda t} \cdot v_{kj}^{\lambda t} \tag{3.8}$$

This reduces the scaling of the first term from $\mathcal{O}(N^4)$ to $\mathcal{O}(N^2)$, because only two indices are used during the density contraction. The separation of the integrals and the density matrix, however, is not possible for the term on the right-hand side, because only one index of the corresponding integral matrix $\mathbf{v}$ is connected to one of the two summation indices (l and k). However, instead of contracting both indices at once, it is possible to contract each index on its own. Hence, we first contract index k

$$V_{ij}''^{\lambda t} = v_{ij}^{\lambda t} \cdot X^{\lambda t} - \tfrac{1}{2} \sum_{l} v_{il}^{\lambda t} \cdot \sum_{k} D_{kl} \cdot v_{kj}^{\lambda t} \; , \tag{3.9}$$

which results in an intermediate matrix $\mathbf{Y}^{\lambda t}$

$$V''^{\lambda t}_{ij} = v^{\lambda t}_{ij} \cdot X^{\lambda t} - \tfrac{1}{2} \sum_{l} v^{\lambda t}_{il} \cdot Y^{\lambda t}_{lj} \ . \tag{3.10}$$

Then, we contract $\mathbf{Y}^{\lambda t}$ with the second summation index l, resulting in another intermediate matrix $\mathbf{Z}^{\lambda t}$. Thus, the two-electron matrix $\mathbf{V}''$ can be expressed using only two indices

$$V''^{\lambda t}_{ij} = v^{\lambda t}_{ij} \cdot X^{\lambda t} - \tfrac{1}{2} Z^{\lambda t}_{ij} \ . \tag{3.11}$$

This formulation is the foundation of the current implementation for the construction of $\mathbf{V}''$ (Listing 3.2).

Listing 3.2: Pseudocode for the final implementation of the density contraction.

```
1  init V'', X^λt, Y^λt, Z^λt, v^λt
2  loop over λ,t
3      v^λt ← calcIntegralMatrix(λ,t)
4      loop over kl
5          X^λt ← X^λt + D_kl · v^λt_kl
6      end
7      Z^λt = 0
8      loop over jl
9          Y = 0
10         loop over k
11             Y^λt ← Y^λt + D_kl · v^λt_kj
12         end
13         loop over i
14             Z^λt_ij ← Z^λt_ij + Y^λt · v^λt_il
15         end
16     end
17     loop over ij
18         V''_ij ← V''_ij + c_λt · v^λt_ij · X^λt
19         V''_ij ← V''_ij − c_λt · ½ Z^λt_ij
20     end
21 end
22 return V''
```

First, all required one-electron integrals for a given integral type t and a core λ are

calculated and stored in the corresponding matrix $\mathbf{v}^{\lambda t}$ (line 3). Then, $X^{\lambda t}$ (left term of equation 3.11) (lines 4-6) and matrix $\mathbf{Z}^{\lambda t}$ (right term of equation 3.11) (lines 7-16) are constructed. For setting up $\mathbf{Z}^{\lambda t}$, two looping branches with three indices each are needed. In the first branch (j, l, k), a scalar $Y^{\lambda t}$ is initialized and subsequently used in the second branch (j, l, i) to calculate the intermediate matrix $\mathbf{Z}^{\lambda t}$. Finally, both $X^{\lambda t}$ and $\mathbf{Z}^{\lambda t}$ are used to output $\mathbf{V}''$ (eq. 3.11) (lines 17-22).

With regards to computational demands, the bottleneck of this algorithm is setting up the $\mathbf{Z}^{\lambda t}$ matrix. In contrast to the runtime scaling of the generation of $X^{\lambda t}$, which is $\mathcal{O}(N^2)$, generating the $\mathbf{Z}^{\lambda t}$ matrix with $2\mathcal{O}(N^3)$, because two summations with three indices each are used. This is an improvement by a factor of N when compared to the straightforward implementation (Listing 3.1).

3.4 Evaluating the CPP integrals

For generating the $\mathbf{V}'$ and $\mathbf{V}''$ matrices, CPP integrals need to be evaluated. For this, we follow the original implementation by Nicklaß [14] in MOLPRO [13]. Parts of the original code were taken as inspiration for the custom implementation in this work. In the next sections we provide a compact overview over all equations and relations needed to understand the evaluation and thus the implementation in detail. For further details, consider the literature:

- Derivation of the equations [12]
- Corrections [14] of and errata [83, 84] to the original work (derivation of the equations)
- Implementation [14] and usage in MOLPRO [13, 14]

3.4.1 The one- and two-electron CPP integrals

The one- and two-electron integrals for the CPP ansatz are defined as follows:

$$\begin{aligned}\langle i|v'|j\rangle = -\tfrac{1}{2}\sum_{\lambda}\Bigg\{&\alpha_{\mathrm{D}}^{\lambda}\left[v_{ij}^{\lambda 0004} - 2\left(v_{ij}^{\lambda 1003}B_X^{\lambda} + v_{ij}^{\lambda 0103}B_Y^{\lambda} + v_{ij}^{\lambda 0013}B_Z^{\lambda}\right)\right]\\ &+\alpha_{\mathrm{Q}}^{\lambda}\Bigg[v_{ij}^{\lambda 0006} + v_{ij}^{\lambda 0003}D^{\lambda}\\ &-3\left(v_{ij}^{\lambda 2005}C_{XX}^{\lambda} + v_{ij}^{\lambda 0205}C_{YY}^{\lambda} + v_{ij}^{\lambda 0025}C_{ZZ}^{\lambda}\right)\\ &-6\left(v_{ij}^{\lambda 1105}C_{XY}^{\lambda} + v_{ij}^{\lambda 1015}C_{XZ}^{\lambda} + v_{ij}^{\lambda 0115}C_{YZ}^{\lambda}\right)\\ &-6\beta_{\mathrm{D}}^{\lambda}v_{ij}^{\lambda 0006}\Bigg]\Bigg\}, \end{aligned} \tag{3.12}$$

$$\begin{aligned}\langle ij|v''|kl\rangle = -\sum_{\lambda}\Bigg\{&\alpha_{\mathrm{D}}^{\lambda}\left[v_{ik}^{\lambda 1003}v_{jl}^{\lambda 1003} + v_{ik}^{\lambda 0103}v_{jl}^{\lambda 0103} + v_{ik}^{\lambda 0013}v_{jl}^{\lambda 0013}\right]\\ &+\tfrac{1}{2}\alpha_{\mathrm{Q}}^{\lambda}\Bigg[-v_{ik}^{\lambda 0003}v_{jl}^{\lambda 0003}\\ &+3\left(v_{ik}^{\lambda 2005}v_{jl}^{\lambda 2005} + v_{ik}^{\lambda 0205}v_{jl}^{\lambda 0205} + v_{ik}^{\lambda 0025}v_{jl}^{\lambda 0025}\right)\\ &+6\left(v_{ik}^{\lambda 1105}v_{jl}^{\lambda 1105} + v_{ik}^{\lambda 1015}v_{jl}^{\lambda 1015} + v_{ik}^{\lambda 0115}v_{jl}^{\lambda 0115}\right)\Bigg]\Bigg\}. \end{aligned} \tag{3.13}$$

Here, the one- and two-electron integrals are defined with respect to two and four PCGs, respectively. These basis functions are indexed by i, j, k and l. λ represents the index for all cores/centers in the molecular system. B_A^{λ}, C_{AB}^{λ} and D^{λ} are constants, which are described in the equations 2.50-2.52. $\alpha_{\mathrm{D}}^{\lambda}$, $\alpha_{\mathrm{Q}}^{\lambda}$ and $\beta_{\mathrm{D}}^{\lambda}$ represent the dipole, quadrupole polarizability and the adiabatic correction to the dipole polarizability, respectively [12, 80].

$v_{ij}^{\lambda t}$ represents the integral for a given type t ($t = m_1m_2m_3k$), core λ and two PCGs, indexed by i and j

$$\begin{aligned} v_{ij}^{\lambda m_1m_2m_3k} &= \langle i|\hat{v}^{\lambda t}|j\rangle \\ &= \left\langle i\left|\frac{(x-x_\lambda)^{m_1}(y-y_\lambda)^{m_2}(z-z_\lambda)^{m_3}}{\left|\vec{r}-\vec{R}_\lambda\right|^{k}}\cdot\omega^{k-(m_1+m_2+m_3)}\left(\left|\vec{r}-\vec{R}_\lambda\right|,\gamma_\lambda\right)\right|j\right\rangle . \end{aligned} \tag{3.14}$$

Here, the cutoff function ω ensures the correct numeric behaviour at the core region [12]. Of the multiple types of cutoff functions described in the literature, we

use one function of the Müller-type [2]:

$$\omega(r, \gamma_\lambda) = 1 - e^{-\gamma_\lambda r^2} \,, \tag{3.15}$$

in which γ_λ is the so-called cutoff parameter. This is the only function of the Müller-type that allows for employing the quadrupole polarizability α_{Q} without any further precautions [14].

3.4.2 Evaluation of the one- and two-electron integrals

The mathematical existence of the solution of the $v_{ij}^{\lambda t}$ integral is crucially linked to the combinations of m_1, m_2, m_3 and k [12]. These values are used to define some new constants, which are later used for the evaluation of the $v_{ij}^{\lambda t}$ integrals. First, k_1 and k_2 can be extracted from the original k using the following relation:

$$(k_1, k_2) = \begin{cases} \left(\frac{k}{2}, 0\right), & \text{for even } k \\ \left(\frac{k-1}{2}, 1\right), & \text{for odd } k \end{cases} \,. \tag{3.16}$$

Another constant derived from the $v_{ij}^{\lambda t}$ integral indices is q, which is later used throughout the integral evaluation

$$q = k - (m_1 + m_2 + m_3) \in \mathbb{N} \,. \tag{3.17}$$

Importantly, the expression of q depends on the chosen cutoff function (eq. 3.15) [14]. Furthermore, while we discuss the evaluation of the $v_{ij}^{\lambda t}$ integrals in the next section, there is one integral, which cannot be evaluated in this manner. This integral $v_{ij}^{\lambda 0003}$ is expressed and thus evaluated using three other integrals

$$v_{ij}^{\lambda 0003} = v_{ij}^{\lambda 2005} + v_{ij}^{\lambda 0205} + v_{ij}^{\lambda 0025} \,. \tag{3.18}$$

General evaluation of $v_{ij}^{\lambda t}$

With the above constants at hand, the general $v_{ij}^{\lambda t}$ integrals can be evaluated via

$$\begin{aligned} v_{ij}^{\lambda m_1 m_2 m_3 k} =& \frac{\sqrt{\pi^3}}{\Gamma(k/2)} e^{-\beta_i r_i^2 - \beta_j r_j^2} \\ & \cdot \sum_{\eta_{1i}=0}^{n_{1i}} \sum_{\eta_{2i}=0}^{n_{2i}} \sum_{\eta_{3i}=0}^{n_{3i}} \sum_{\eta_{1j}=0}^{n_{1j}} \sum_{\eta_{2j}=0}^{n_{2j}} \sum_{\eta_{3j}=0}^{n_{3j}} \binom{n_{1i}}{\eta_{1i}} \binom{n_{2i}}{\eta_{2i}} \binom{n_{3i}}{\eta_{3i}} \binom{n_{1j}}{\eta_{1j}} \binom{n_{2j}}{\eta_{2j}} \binom{n_{3j}}{\eta_{3j}} \\ & \cdot (-1)^{\sum_{o=1}^{3} \sum_{h=1}^{2} (n_{oh} - \eta_{oh})} x_i^{n_{1i}-\eta_{1i}} y_i^{n_{2i}-\eta_{2i}} z_i^{n_{3i}-\eta_{3i}} x_j^{n_{1j}-\eta_{1j}} y_j^{n_{2j}-\eta_{2j}} z_j^{n_{3j}-\eta_{3j}} \\ & \cdot \sum_{\zeta_1=0}^{l_1^\eta} \sum_{\zeta_2=0}^{l_2^\eta} \sum_{\zeta_3=0}^{l_3^\eta} \prod_{c=1}^{3} \binom{l_c^\eta}{\zeta_c} \frac{(\zeta_c - 1)!!}{2^{\zeta_c/2}} b_c^{l_c^\eta - \zeta_c} \sum_{\mu=0}^{q} \binom{q}{\mu} (-1)^\mu \mathtt{u}_{k_2}^{s_{l^\eta}^\eta}(k_1, a_\mu, |\vec{b}|) \, . \end{aligned} \tag{3.19}$$

Here, β are the exponents of the basis functions and $\vec{b}$ is a vector, which is later used to determine how to evaluate the integral. The vector and its norm $|\vec{b}|$ are defined as:

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \beta_i \vec{r_i} + \beta_j \vec{r_j} \quad \text{and} \quad |\vec{b}| = \sqrt{b_1^2 + b_2^2 + b_3^2} \, . \tag{3.20}$$

Additionally, more constants are introduced by the above equation:

$$a_\mu = \beta_i + \beta_j + \mu \gamma_\lambda \, , \tag{3.21}$$

$$s_l^\eta = l_1^\eta + l_2^\eta + l_3^\eta - \frac{\zeta_1 + \zeta_2 + \zeta_3}{2} \, , \tag{3.22}$$

with

$$l_c^\eta = \eta_{c1} + \eta_{c2} + m_c \in \mathbb{N}_0 \, . \tag{3.23}$$

$\zeta_1, \zeta_2, \zeta_3$ are constrained to be even integers by incrementing the index by two. Γ is the gamma function of the following form:

$$\Gamma(z) = \int_0^x x^{z-1} e^{-x} dx \, . \tag{3.24}$$

$\mathtt{u}_{k_2}^{s_{l^\eta}^\eta}$ (eq. 3.19) is a function to evaluate the rest of the integral expression depending on the value of k_2 being odd or even and, if odd, whether the subtraction $s - k_1$ is a positive or negative integer.

3.4.3 Evaluation of u for even k

With k being even ($k_1 = \frac{k}{2}, k_2 = 0$), the u function can be calculated using the d function [12] derived by Schwerdtfeger and Silberbach as follows:

$$\mathtt{u}_0^s(k_1, a, b) = b^{-2s+2k_1-3} \sum_{\nu=0}^{k_1-1} \binom{k_1-1}{\nu} (-1)^\nu \left(\frac{a}{b^2}\right)^\nu \mathtt{d}\left(s - k_1 + \nu + \frac{3}{2}, \frac{b^2}{a}\right) . \quad (3.25)$$

The evaluation of the d function differs depending on the first argument ($p + \frac{1}{2}$) and second argument (x). For positive p and x, the d function is evaluated using

$$\mathtt{d}\left(p + \frac{1}{2}, x\right) = \frac{(2p-1)!!}{2^{p-1}} (-1)^p e^x \left(\mathtt{daw}(\sqrt{x}) - \sum_{i=0}^{p-1} (-1)^i x^{i+\frac{1}{2}} \frac{2^i}{(2i+1)!!}\right) . \quad (3.26)$$

Here, daw is the Dawson function

$$\mathtt{daw}(x) = e^{-x^2} \int_0^x e^{t^2} dt . \quad (3.27)$$

For x values smaller than 0.01 d is evaluates via

$$\mathtt{d}(p + \tfrac{1}{2}, x) = \sum_{n=0}^{\infty} \frac{x^{p+n+\frac{1}{2}}}{n!(p + n + \frac{1}{2})} , \quad (3.28)$$

which is truncated after $n = 6$ in the implementation. For negative p values, the d equation is evaluated using $p' = -p$ in

$$\mathtt{d}\left(\frac{1}{2} - p', x\right) = e^x \left(2C_{p'p'}\mathtt{daw}(\sqrt{x}) - \sum_{i=1}^{p'} C_{ip'} x^{i-p'-\frac{1}{2}}\right) , \quad (3.29)$$

with

$$C_{ip} = \prod_{\nu=1}^{i} \left(p + \frac{1}{2} - \nu\right)^{-1} = \frac{2^i (2p - 2i - 1)!!}{(2p-1)!!} . \quad (3.30)$$

3.4.4 Evaluation of u for odd k

For odd k ($k_1 = \frac{k-1}{2}, k_2 = 1$), the u function is evaluated via

$$\mathtt{u}_1^s(k_1, a, b) = a^{k_1-s-1} e^{b^2/a} \sum_{\nu=0}^{s-k_1} \binom{s-k_1}{\nu} (-1)^\nu \left(\frac{a}{b^2}\right)^{k_1+\nu+\frac{1}{2}} \gamma\left(k_1 + \nu + \frac{1}{2}, \frac{b^2}{a}\right) , \quad (3.31)$$

for $s - k_1 \geq 0$, with γ being the incomplete gamma function

$$\gamma\left(p + \frac{1}{2}, x\right) = \frac{(2p-1)!!}{2^{p-1}} \left(\mathtt{erf}(\sqrt{\pi}) - e^{-x} \sum_{\mu=0}^{p-1} x^{\mu+\frac{1}{2}} \frac{2^\mu}{(2\mu+1)!!} \right) , \tag{3.32}$$

with the error function

$$\mathtt{erf}(x) = \int_0^x e^{-t^2} dt \, . \tag{3.33}$$

For $s - k_1 < 0$, the $\mathtt{u}$ function is evaluated by

$$\mathtt{u}_1^s(k_1, a, b) = 2a^{k_1-s-1} e^{b^2/a} \sum_{\nu=0}^{k_1} \binom{k_1}{\nu} (-1)^\nu \mathtt{H}\left(k_1 - s - \nu, \frac{b^2}{a}\right) , \tag{3.34}$$

which employs another custom function $\mathtt{H}$, which is evaluated recursively

$$\begin{aligned} \mathtt{H}(p, x) &= \frac{2(p+x)-3}{2(p-1)} \cdot \mathtt{H}(p-1, x) - \frac{x}{p-1} \cdot \mathtt{H}(p-2, x) \, , \\ \mathtt{H}(0, x) &= a^{-\frac{1}{2}} \mathtt{erf}(\sqrt{x}) \, , \\ \mathtt{H}(1, x) &= 2e^{-x} \cdot \mathtt{m}(\sqrt{x}) \, , \end{aligned} \tag{3.35}$$

with the Dawson-error-hybrid function

$$\mathtt{m}(x) = \int_0^x e^{t^2} \cdot \mathtt{erf}(t) dt \, , \tag{3.36}$$

for positive arguments of x. For negative values this relation is applied:

$$\mathtt{H}(-p, a) = \frac{1}{2} \sum_{\nu=0}^{p} \binom{p}{\nu} (-1)^\nu a^{-\nu-\frac{1}{2}} \cdot \gamma\left(\nu + \frac{1}{2}, a\right) \, . \tag{3.37}$$

Simplified evaluation of the integral

The above described cases and custom functions are used to evaluate the general form of the integral. Besides this general form (eq. 3.19), there is a simplified form of the integral. This simplified integral can be evaluated if $\vec{b}$ is equal to $\vec{0}$. This is the case when both basis functions are located at the current core or when the relation

$$-\beta_i \vec{r}_i = \beta_j \vec{r}_j \tag{3.38}$$

holds true. In the latter case (eq. 3.38) the general integral transforms to

$$\begin{aligned} v_{ij}^{\lambda m_1 m_2 m_3 k} = &\sum_{\eta_{1i}=0}^{n_{1i}} \sum_{\eta_{2i}=0}^{n_{2i}} \sum_{\eta_{3i}=0}^{n_{3i}} \sum_{\eta_{1j}=0}^{n_{1j}} \sum_{\eta_{2j}=0}^{n_{2j}} \sum_{\eta_{3j}=0}^{n_{3j}} \binom{n_{1i}}{\eta_{1i}} \binom{n_{2i}}{\eta_{2i}} \binom{n_{3i}}{\eta_{3i}} \binom{n_{1j}}{\eta_{1j}} \binom{n_{2j}}{\eta_{2j}} \binom{n_{3j}}{\eta_{3j}} \\ &\cdot (-1)^{\sum_{o=1}^{3} \sum_{h=1}^{2} (n_{oh} - \eta_{oh})} x_i^{n_{1i}-\eta_{1i}} y_i^{n_{2i}-\eta_{2i}} z_i^{n_{3i}-\eta_{3i}} x_j^{n_{1j}-\eta_{1j}} y_j^{n_{2j}-\eta_{2j}} z_j^{n_{3j}-\eta_{3j}} \\ &\cdot e^{-\beta_i \vec{r}_i^{\,2} - \beta_j \vec{r}_j^{\,2}} L_{ij}^{kq\lambda}(\eta_{1i} + \eta_{1j} + m_1, \eta_{2i} + \eta_{2j} + m_2, \eta_{3i} + \eta_{3j} + m_3) \ , \end{aligned} \tag{3.39}$$

while in the former case ($\vec{r}_i = \vec{r}_j = \vec{0}$) the integral simplifies even more to

$$v_{ij}^{\lambda m_1 m_2 m_3 k} = L_{ij}^{kq\lambda}(n_{1i} + n_{1j} + m_1, n_{2i} + n_{2j} + m_2, n_{3i} + n_{3j} + m_3) \ , \tag{3.40}$$

with

$$L_{ij}^{kq\lambda}(p_1, p_2, p_3) = \int_{\mathbb{R}^3} x^{p_1} y^{p_2} z^{p_3} \frac{1}{r^k} \sum_{\mu=0}^{q} \binom{q}{\mu} (-1)^{\mu} e^{-[\beta_i + \beta_j + \mu\gamma_\lambda] r^2} d\tau \ . \tag{3.41}$$

All p_i are constrained to be even, otherwise the whole integral would become zero and no explicit evaluation is needed [14]. The evaluation of this integral is split into two cases. For even k the integral can be evaluated as

$$\begin{aligned} L_{ij}^{kq\lambda}(p_1, p_2, p_3) = &\pi^{\frac{3}{2}} \frac{(p_1 - 1)!!(p_2 - 1)!!(p_3 - 1)!!}{2^s} \prod_{c=0}^{\frac{k}{2}-1} \frac{1}{s - c + \frac{1}{2}} \\ &\cdot \sum_{\mu=0}^{q} \binom{q}{\mu} (-1)^{\mu} (\beta_i + \beta_j + \mu\gamma)^{\frac{k}{2} - s - \frac{3}{2}} \ , \end{aligned} \tag{3.42}$$

while for odd k the integral is evaluated as

$$\begin{aligned} L_{ij}^{kq\lambda}(p_1, p_2, p_3) = &2\pi (p_1 - 1)!!(p_2 - 1)!!(p_3 - 1)!! \frac{(s - \frac{k-1}{2})!}{(2s+1)!!} \\ &\cdot \sum_{\mu=0}^{q} \binom{q}{\mu} (-1)^{\mu} (\beta_i + \beta_j + \mu\gamma_\lambda)^{\frac{k-1}{2} - s - 1} \ , \end{aligned} \tag{3.43}$$

with

$$s = \frac{p_1 + p_2 + p_3}{2} \ . \tag{3.44}$$

For this last form of the integral, there is the additional condition that if $s - j < 0$ holds true, the integral evaluates to zero [14].

3.5 Implementation and design decisions

3.5.1 `Cpp` and `Integral` class

The implementation in this work is based on the object-oriented programming paradigm, which expresses the functionalities of the program via so-called classes. These classes can then be used to build the final program. To design the program structure in an intuitive way, we implemented a "top-level" class, the `Cpp` class (Fig. 3.2), which performs all crucial steps during the SCF procedure (section 3.2). This class is initialized by a molecule, a basis representation, CPP parameters (γ_λ, $\alpha_\mathrm{D}^\lambda$, $\alpha_\mathrm{Q}^\lambda$ and β_D^λ) and molecule-dependent constants (B_A^λ, C_{AB}^λ and D^λ), which are calculated before the SCF procedure. The class is then used for calculating the CPP contributions Ω, $\mathbf{V}'$ and $\mathbf{V}''$ via `calcNucNuc`, `calcOneEle` and `calcTwoEle`, respectively (Fig. 3.2).

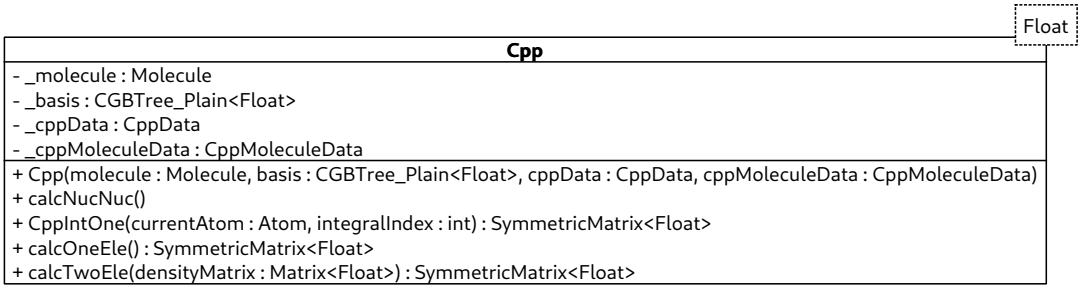

Fig. 3.2: Schematic representation of the `Cpp` class, which performs all crucial operations for the contributions of the CPP ansatz (Ω, $\mathbf{V}'$ and $\mathbf{V}''$) to the SCF code. The class is templated with the template argument `Float`, which represents a generic type of floating point number.

Both functions `calcOneEle` and `calcTwoEle` employ the `Integral` class, which represents the integral evaluation (Fig. 3.3) described in section 3.4. This class is a so-called base class, which allows for deriving two other classes, namely `CPPIntegralGeneral` and `CPPIntegralSimple`. These two classes evaluate the general (eq. 3.19) and the simplified form (eq. 3.40) of the $v_{ij}^{\lambda t}$ integral via their `evaluate` member-function. Of note, most classes in this work are template-classes allowing for evaluating the integrals with different floating point precision, if needed.

3.5.2 Decision making during the integral evaluation

With `CppIntegralGeneral` and `CppIntegralSimple` at hand, we then implemented the integral evaluation to set up the one- and two-electron matrices ($\mathbf{V}'$ and $\mathbf{V}''$). The

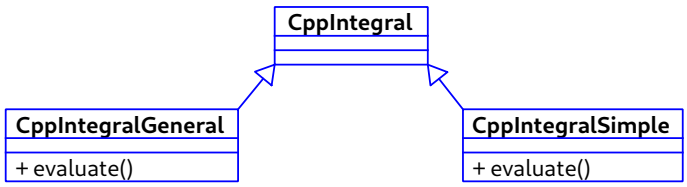

Fig. 3.3: Simplified schematic representation of the `CppIntegral` base class and the derived classes `CppIntegralGeneral` and `CppIntegralSimple`.

only missing part is a way to decide which integral evaluation is needed according to the conditions described in section 3.4. For this, the class `CppIntegralEvaluator` was implemented (Listing 3.3).

Listing 3.3: Pseudocode to evaluate the $v_{ij}^{\lambda t}$ using either the the simple or the general form of the $v_{ij}^{\lambda t}$ integral during the initialization of the `CppIntegralEvaluator` class.

```
1 if (r_i^2 + r_j^2 == 0) && (b^2 == 0)
2   value = CppIntegralSimple.evaluate();
3 else if (b^2 == 0)
4   tmp = 0
5   loop over n_1i, n_2i, n_3i, n_1j, n_2j, n_3j
6     tmp = tmp + ... · CppIntegralSimple.evaluate();
7   value = e^(−β_i·r_i^2 − β_j·r_j^2) · tmp
8 else
9   value = CppIntegralGeneral.evaluate();
```

Here, we first check two cases, in which the $v_{ij}^{\lambda t}$ integral can be evaluated via the simplified form (line 1 and 3). If this is not the case, we refer to the general form of the integral evaluation (line 9).

3.5.3 Setting up the $\mathbf{v}^{\lambda t}$ and $\mathbf{V}'$ matrix

With the above equations and the decision making process via `CppIntegralEvaluator`, the $v_{ij}^{\lambda t}$ can be evaluated. For the construction of $\mathbf{V}'$ and $\mathbf{V}''$, we use the matrix $\mathbf{v}^{\lambda t}$ to store all $v_{ij}^{\lambda t}$ integrals for a given core λ and a given type of integral $t = m_1 m_2 m_3 k$. Constructing this matrix is straightforward and is achieved by iterating over all n basis functions, resulting in a $n \times n$ matrix.

This matrix is then used for constructing $\mathbf{V}'$ by summing over all $\mathbf{v}^{\lambda t}$ and a prefactor $c_{\lambda t}$ according to equation 3.12. Hence, when only contributions of the core polarization potentials using the dipole polarizeability $\alpha_{\mathrm{D}}^{\lambda}$ are considered the sum would look like this

$$\mathbf{V}' = -\tfrac{1}{2} \cdot \sum_{\lambda} \sum_{t \in \mathbb{D}_1} \alpha_{\mathrm{D}}^{\lambda} c_{\lambda t} \mathbf{v}^{\lambda t} \quad , \quad \mathbb{D}_1 = \{0004, 1003, 0103, 0013\} \ . \tag{3.45}$$

Additional contributions of the quadrupole polarizability $\alpha_{\mathrm{Q}}^{\lambda}$ and the adiabatic correction to the dipole polarizability $\beta_{\mathrm{D}}^{\lambda}$ extend the sum to

$$\mathbf{V}' = -\tfrac{1}{2} \cdot \sum_{\lambda} \left(\sum_{t \in \mathbb{D}_1} \alpha_{\mathrm{D}}^{\lambda} c_{\lambda t} \mathbf{v}^{\lambda t} + \sum_{t \in \mathbb{D}_2} \alpha_{\mathrm{Q}}^{\lambda} c_{\lambda t} \mathbf{v}^{\lambda t} - 6 \beta_{\mathrm{D}}^{\lambda} \mathbf{v}^{\lambda 0006} \right) \ , \tag{3.46}$$

with

$$\mathbb{D}_2 = \{0006, 0003, 2005, 0205, 0025, 1105, 1015, 0115\} \ .$$

Here, $c_{\lambda t}$ are the prefactors in equation 3.12, e.g. 1 and $-6C_{XY}^{\lambda}$ for $v_{ij}^{\lambda 0004}$ and $v_{ij}^{\lambda 1105}$, respectively. This concludes the construction of $\mathbf{V}'$ and $\mathbf{V}''$. The next step is to implement a way of passing CPP parameters to the program.

3.5.4 Input for CPP parameters

For running simulations using CPPs, the parameters of the CPPs ($\gamma_\lambda, \alpha_{\mathrm{D}}^{\lambda}, \alpha_{\mathrm{Q}}^{\lambda}, \beta_{\mathrm{D}}^{\lambda}$) are read from an input file. Additionally we specify the charge of the valence electrons q_λ of a core λ. Here, the `SimpSimpleInputParser`, written by J. Held, is extended by the `-data` flag, which receives the file path of the input file as an argument. Then, the parameters are read from a file starting with the `CppData` identifier. All values are read line-by-line until the `CppDataEnd` identifier is encountered (line 2-3 in Listing 3.4). In each line the `CppData` identifier, and the atomic label are read as string (`std::string`) and all following values ($\alpha_{\mathrm{D}}^{\lambda}$, $\alpha_{\mathrm{Q}}^{\lambda}$, $\beta_{\mathrm{D}}^{\lambda}$, γ_λ and q_λ) are read as floating point numbers (`double`) separated by whitespace characters. Everything before and after these lines is ignored (line 1). All processing is done by the `CppData` class, which is then passed to the `CppMolecularData` and `Cpp` class.

Listing 3.4: Example of a CPP parameter input file.

```
1   Atom  alpha_D alpha_Q beta_D  gamma QCharge
2 CppData Na 0.9947 0.0000 0.0000 0.6200 1.0000
3 CppData Br 6.6660 2.2222 1.0101 1.0000 1.0000 CppDataEnd
```

3.5.5 Technical details

The QOL was compiled with the GCC C++ Compiler (version 10.2) [104]. Additionally, three external libraries were used: BOOST (version 1.75.0) [105], GNU

Scientific Library (gsl [106]) for special mathematical functions (table 3.1) and the INTEL *Math Kernel Library* (version 11.3) [107] for calling LAPACK/BLAS routines throughout the SCF procedure.

Tab. 3.1: Mathematical functions from external libraries used in this work.

Mathematical function	equation	Library	Library function
error function `erf`	3.33	BOOST	`erf`
binomial coefficient		BOOST	`binomial_coefficients<T>`
gamma function Γ	3.24	BOOST	`tgamma_lower`
dawson function `daw`	3.27	gsl	`gsl_sf_dawson`

All other function, e.g. the incomplete gamma function γ (eq. 3.32) or the Dawson-error hybrid function `m` (eq. 3.36), were implemented according to Nicklaß [14].

3.6 Verifying the SCF CPP implementation

3.6.1 Computational details

To verify the implementation of the CPP ansatz in the SCF code of the QOL, we performed simulations of atomic and dimer systems using energy-consistent small-core pseudopotentials of the Stuttgart-Cologne group [62]. These simulations were compared to MOLPRO (version 2020.2) [13] simulations with same inputs. Hereby, only closed-shell systems are considered. The QOL was compiled on a PC (Intel(R) Core(R) i3-8100; 16 GB RAM; name: pc11) as described in section 3.5.5. The simulations were performed on a cluster node (Intel(R) Xeon(R) CPU E7-4809 v3; 500 GB RAM; name: dc01). All CPP parameters were chosen arbitrarily (Appendix A.1). The absolute energy difference between the SCF QOL and MOLPRO closed-shell Hartree-Fock (rhf) ground state energy was chosen as criterion for the successful implementation. In both SCF programs a energy convergence threshold of 10^{-7} a.u. was chosen and the energies were compared with a precision of 12 floating point digests. All deviations smaller than 10^{-12} a.u. will be displayed as $0.0000 \cdot 10^{-00}$ a.u.

3.6.2 Atomic systems

Atomic systems represent the least complex system to verify the implementation of the CPP ansatz into the QOL. Since there is only one core center in theses simulations, only the `CppIntegralSimple` is called during the simulations. Here, different PPs with different basis set sizes were compared (Tab. 3.2). The energy differences were lower than the energy convergence threshold of 10^{-7} a.u., indicating that the CPP code behaves the same in QOL as in MOLPRO. Thus, the correct behavior of `CppIntegralSimple` confirms the correct implementation. Since the CPP parameters were chosen arbitrarily, it is not possible to assess the contribution of the CPP to the energy in a physically meaningful way. To get an impression of this influence, we would like to mention, that the contribution of the CPP amounts to 4 to 12 % of the total energy (Appendix A.2). Such a significant contribution allows for a safe comparison of both implementations.

Tab. 3.2: Absolute energy differences between the MOLPRO and QOL implementation of the CPP ansatz for different PPs in atomic systems. All energy differences are given in atomic units.

Atom	PP/Basis	Contraction	Number of Iterations		$\lvert\Delta E\rvert$ / a.u.
			QOL	MOLPRO	
Mg	10SDF [108]	4s4p	6	6	$0.0000 \cdot 10^{-00}$
Yb	60MWB [64, 109]	(7s6p5d)/[5s4p3d]	8	8	$2.3704 \cdot 10^{-11}$
Hg	60MHF[a]	(8s7p6d)/[6s5p3d][a]	7	8	$1.5632 \cdot 10^{-11}$
	60MWB [110, 111]	(8s7p6d2f1g)/[6s5p3d2f1g]	7	6	$9.7913 \cdot 10^{-11}$

[a] unpublished; see [62]

3.6.3 Molecular systems

For the evaluation of the CPP integrals in simulations of molecular systems, not only `CppIntegralSimple`, but also `CppIntegralGeneral` is employed. Hence, these systems allow to verify the correct implementation of both classes. As in the atomic test systems, no significant deviations in the energy are observed when the QOL energy values are compared to MOLPRO (Tab. 3.3). Thus, both classes were implemented correctly for multiple PPs and different basis sets. The only difference can be observed in the convergence of both SCF procedures. In most molecular simulations the SCF of the QOL converges faster. This, however, was not observed in the atomic simulations and thus this not a significant observation. Again, in these simulations the contribution of the CPPs ranges between 0.1 and 5 % of the total

energy (Appendix A.2), allowing for a fair comparison of both implementations.

Tab. 3.3: Absolute energy differences between the MOLPRO and QOL implementation of the CPP ansatz for different PPs in molecular systems. All energy differences are given in atomic units.

Dimer	PP/Basis	Contraction	Number of Iterations		$\lvert\Delta E\rvert$ / a.u.
			QOL	MOLPRO	
Na_2	10SDF [5]	(4s4p)/[2s2p][a]	6	7	$2.1656 \cdot 10^{-12}$
K_2	10MWB [112]	(7s6p)/[5s4p][a]	9	9	$9.3265 \cdot 10^{-10}$
Sc_2	10MHF [113]	(8s7p6d)/[2s1p1d][a]	9	11	$3.4099 \cdot 10^{-10}$
Hg_2	60MHF[a]	(8s7p6d)/[6s5p3d][a]	9	11	$1.2074 \cdot 10^{-09}$
	60MWB [110, 111]	(8s7p6d2f1g)/[6s5p3d2f1g]	11	12	$8.4702 \cdot 10^{-10}$

[a] unpublished; see [62]

3.7 Verifying the implementation of CPPs in CC simulations

Having confirmed the correct incorporation of the CPP integrals into the SCF procedure, we next performed CC simulations. To perform these CC simulations in combination with CPPs, we extended the already existing interface between the SCF method and the CC code, the so-called MO transformation. The CC program was provided by M. Hanrath.

3.7.1 MO transformation

Molecular orbitals

A converged SCF simulation yields the molecular orbital (MO) coefficients C_{il}. These coefficients relate the atomic orbitals (AOs) to the MOs by

$$|\mathrm{MO}_i\rangle = \sum_a^n C_{ia} \, |\mathrm{AO}_a\rangle \ . \tag{3.47}$$

These MOs represent the optimized spin orbitals of the Hartree-Fock method. Using these MOs in a correlation method is advantageous, because they describe the best possible solution for a single determinate description of the wave function. Additionally, while performing electron correlation methods with AOs is possible, MOs are simpler to implement and converge typically faster in these procedures [114].

SCF integral MO transformation

The QOL SCF code and the CC code produce separate binaries and the former does not pass the MO coefficients to the latter. For being able to perform a CC simulation, we thus have to (i) transform the AOs to MOs using the MO coefficients and (ii) pass the MOs to the CC program. Because the SCF and the CC program are independent of each other, this MO transformation is implemented in an intermediary program. For the MO transformation, the MO coefficients C_{il} are used to transform the one- and two-electron integrals from the AO-basis to the MO-basis via

$$\langle i|\hat{h}|j\rangle = \sum_{a}^{n} C_{ia} \sum_{b}^{n} C_{jb} \langle a|\hat{h}|b\rangle \tag{3.48}$$

$$\langle ij|kl\rangle = \sum_{a}^{n} C_{ia} \sum_{b}^{n} C_{jb} \sum_{c}^{n} C_{kc} \sum_{d}^{n} C_{ld} \langle ab|cd\rangle \tag{3.49}$$

where $a, b, \ldots$ and $i, j, \ldots$ are the indices of the AO and MO basis, respectively. These transformed integrals are then passed to the CC method via an integral file.

Transforming the integrals, especially the two-electron integrals, scales with $\mathcal{O}(n^8)$ when implementing this transformation via equation 3.49. The current implementation of the MO transformation in the QOL, scales only with $\mathcal{O}(n^5)$, which is achieved by matrix-matrix multiplications and subsequent shuffling. The latest changes in the implementation of the MO transformation were performed by M. Hülsen [115], J. Ciupka [89] and J. Held [91]. Please consider their work for a more detailed explanation. For the efficient transformation of the two-electron integrals $\langle ab|cd\rangle$, they are described by a four tensor $\mathbf{G}_{abcd}$, which is indexed by only two indices, of which one is a super index of three indices, e.g. $\mathbf{G}_{abc,d}$. With this, a MO transformation can be performed by a single matrix multiplication

$$\sum_{d} c_{ld} \cdot g_{abc,d} = g_{abc,l} \; ; \quad \mathbf{C}_{ld}\mathbf{G}^{T}_{abc,d} = \mathbf{G}_{abc,l} \; . \tag{3.50}$$

After this step, the indices are shuffled so that the next AO index can contracted to a MO index, e.g.

$$\mathbf{G}_{abc,i} \xrightarrow{\text{shuffle}} \mathbf{G}_{iab,c} \; . \tag{3.51}$$

Repeating these steps three times results in a fully contracted tensor $\mathbf{G}_{ijkl}$, which is then further transformed and saved to a file. The resulting file, containing the transformed integrals, is part of the input for the CC program.

CPP MO transformation

To make the CPP integrals available for the CC simulation, they have to be incorporated in the MO transformation. For this, the already implemented MO transformation of J. Ciupka and J. Held was modified by adding the CPP integrals onto the AO integrals. The first implementation added the CPP integrals onto the AO integrals as they are being calculated

$$\langle a|\hat{h}|b\rangle \leftarrow \langle a|\hat{h}|b\rangle + \langle a|v'|b\rangle \ , \tag{3.52}$$

$$\langle ab|cd\rangle \leftarrow \langle ab|cd\rangle + \langle ab|v''|cd\rangle \ . \tag{3.53}$$

The modified integrals $\langle a|\hat{h}|b\rangle$ and $\langle ab|cd\rangle$ were then transformed with two and four matrix multiplication, respectively.

Here, however, it is possible to split the two-electron CPP integrals into products of one-electron CPP integrals, as described in section 3.3. This allows for an implementation, in which the SCF integrals and the CPP integrals are calculated and transformed separately, because there is no need to incorporate the CPP integrals before the MO transformation of the SCF integrals. Hence, the one-electron CPP integrals ($v_{ij}^{\lambda t}$) are transformed by two matrix-matrix multiplications and combined according to equation 3.3. For this no shuffling of indices is needed. This way of transforming the integrals represents the current implementation, which can be performed with the `addTwoEleCpp` member function of the `Cpp` class. In this implementation the two-electron integrals (electron repulsion and CPP) are stored linear in the memory. Thus, combining the two integrals, after the transformation, is done from the first (0) to the last ($n^4 - 1$) index of these integrals.

3.7.2 Verifying the implementation (MO transformation)

To verify the implementation of the CPP MO transformation, simulations in atomic and dimer closed-shell systems were performed using the CCSD method. The ground state energy was compared to energies obtained using MOLPRO. This allowed for verifying the CPP integral routine on a correlated level.

3.7.3 Computational details

All simulations were performed analogous to the SCF test simulations (section 3.6.1). The only difference to the previous SCF simulations, was that we chose an energy

convergence threshold of 10^{-9} a.u. to avoid an error propagation to the coupled cluster simulation. Here, we chose the CCSD method because both programs are capable of performing such simulations. Here, the convergence threshold of 10^{-7} a.u. for the correlation energy was chosen and all electrons were explicitly correlated (MOLPRO `core` flag). All other parameters were chosen according to the default values of MOLPRO. All deviations smaller than 10^{-12} a.u. will be displayed as $0.0000 \cdot 10^{-00}$ a.u.

3.7.4 Atomic and molecular systems

Tab. 3.4: Comparison of the correlation energies of CCSD performed with MOLPRO and the combination of QOL MO transformation and CC program in atomic systems. The absolute energy differences $|\Delta E_{\text{corr}}|$ are given in atomic units.

Atom	PP/Basis	Contraction	$\lvert\Delta E_{\text{corr}}\rvert$ / a.u.
Mg	10SDF [108]	4s4p	$0.0000 \cdot 10^{-00}$
Yb	60MWB [64, 109]	(7s6p5d)/[5s4p3d]	$7.9111 \cdot 10^{-10}$
Hg	60MHF[a]	(8s7p6d)/[6s5p3d][a]	$8.4099 \cdot 10^{-09}$
	60MWB [110, 111]	(8s7p6d2f1g)/[6s5p3d2f1g]	$1.9628 \cdot 10^{-08}$

[a] unpublished; see [62]

We observe that both the atomic (Tab. 3.4) and molecular (Tab. 3.5) test simulations deviate from the MOLPRO energies by approximately 10^{-8} to 0 (smaller than 10^{-12}) a.u. Thus, both CCSD implementations using CPPs behave the same with respect to the convergence threshold of 10^{-7} a.u.

Tab. 3.5: Comparison of the correlation energies of CCSD performed with MOLPRO and the combination of QOL MO transformation and CC program in molecular systems. The absolute energy differences $|\Delta E_{\text{corr}}|$ are given in atomic units.

Dimer	PP/Basis	Contraction	$\lvert\Delta E_{\text{corr}}\rvert$ / a.u.
Na_2	10SDF [5]	(4s4p)/[2s2p][a]	$6.5875 \cdot 10^{-09}$
K_2	10MWB [112]	(7s6p)/[5s4p][a]	$5.4548 \cdot 10^{-08}$
Sc_2	10MHF [113]	(8s7p6d)/[2s1p1d][a]	$2.2732 \cdot 10^{-09}$
Hg_2	60MHF[a]	(8s7p6d)/[6s5p3d][a]	$4.9281 \cdot 10^{-09}$
	60MWB [110, 111]	(8s7p6d2f1g)/[6s5p3d2f1g]	$5.9430 \cdot 10^{-08}$

[a] unpublished; see [62]

Again, the influences of the CPPs are large enough for both simulations to be compared (Appendix A.2). In contrast to the SCF simulations, however, the contribution

is even larger. In the atomic magnesium simulation for example, the correlation energy from the simulation with a CPP is about 20 % smaller than that from the simulation without CPP. For both simulations, the biggest gain in correlation energy can be observed with the 60MWB PP for mercury (about 100 %) and ytterbium (266 %). These large changes in correlation energy allow for a fair comparison of both programs to verify the successful implementation in the QOL.

4 CPPs in molecular simulations

To investigate the influence of CPPs in molecular systems, we focused on two test systems, Hg_2 and HgF_4. The simulations were centered around the energy-constant small-core PP of mercury, 60MDF [116]. Although, the two test systems have been previously described both in experimental and theoretical investigations, theoretical investigations employing CPPs are still lacking. The two test systems are both closed-shell systems, which, in principle, allows for performing these simulations using the QOL. While the CC program does allow for arbitrary substitutions levels, it does not allow to perform CCSD(T) simulations. To reproduce values in the literature, however CCSD(T) was required and we thus used the MOLPRO program package to perform al simulations instead, which also offers a significantly faster implementation of the CPP ansatz.

4.1 Generating the CPP

A prerequisite for the simulations are CPP parameters for the respective pseudopotential. Since these parameters have not been previously reported, we first generated them for the mercury PP 60MDF [116]. The generation of CPPs in the literature involves two steps: first, the dipole polarizability α_D is obtained by an *ab initio* simulation and second, the cutoff factor γ is fitted, so that the simulations accurately describe ionization potentials (IPs) [10, 11]. We modified this procedure by purely relying on *ab initio* methods and data. The quadrupole polarizability α_Q and the adiabatic correction to the dipole polarizability β_D were not considered and thus not generated.

4.1.1 Dipole polarizability

To obtain the dipole polarizability α_D, we performed a DHF simulation for Hg^{20+}, which corresponds to the ECP60MDF core, by using the program provided by D. Kolb [117]. We obtained a value of 0.10662239 a.u. for α_D. This values was used throughout all following simulations, both for the fitting of the CPPs and the molecular simulations.

4.1.2 Cut-off parameter

In previous works, the cutoff parameter γ was fitted to ionization potentials of the corresponding element [10, 11]. Here, we propose a different approach. We first perform AE-DKH2 (second-order Douglas-Kroll-Hess Hamiltonian) simulations for the Hg^{20+} and Hg^{19+} ground states. AE-DKH2 simulations were provided by M. Dolg employing the ANO-RCC basis of Roos *et al.* [118, 119]. Both states were simulated using the multi-reference SCF (MCSCF) method to obtain the energy difference of both states with respect to the static core-valence correlation. To obtain the combination of static and dynamic core-valence correlation, the CCSD(T) method was employed. The energy difference between both states represents an ionization potential, which was chosen as the fitting criterion for PP simulations employing the 60MDF PP. In this simulation the one electron system (PP Hg^{19+}) was fitted to accurately describe the AE IP. The resulting CPPs are referred to as stat and statDyn for the fit onto the static core-valence correlation energy and the combination of static and dynamic core-valence correlation energy, respectively. Additionally, one fit was performed for the difference between stat and statDyn to approximate a core-valence correlation solely relying on dynamic correlation. This CPP is called dyn (Tab. 4.1).

Tab. 4.1: Obtained γ values for ECP60MDF using an α_D of 0.10662239 a.u.

stat	dyn	statDyn
1.4403	1.9580	2.5792

4.1.3 Testing the generated CPPs

For testing the generated CPPs, we performed a CCSD(T) simulation of the first four ionization potentials. These values were compared to ECP60MDF simulations without CPPs and experimental values (Tab. 4.2). The spin-orbit (SO) contribution was estimated by M. Dolg for all IPs and subtracted form the experimental values (exp.-SO). This allowed for the direct comparison of the ECP60MDF simulations with the experimental values. We observed that the values of the first three IPs were higher than when comparing the simulations with and without CPPs. While stat CPP describes all IPs well, the dyn and statDyn CPPs significantly deviate from the first two IPs in the experimental literature. In general, the third and fourth IPs were not effected by the different cutoff parameter, varying by only, 0.03 and 0.01 eV,

Tab. 4.2: Comparison of the calculated first four ionization potentials, given in electronvolts (eV), for the mercury atom between simulations with and without CPPs and experimental values. All simulations were performed using the ECP 60MDF and different cutoff parameters. α_D is 0.10662239 a.u. The experimental values were corrected by the SO contribution by M. Dolg.

	exp.-SO	no CPP	cutoff parameter name/value			
			stat	dyn	statDyn	expFit
			1.4403	1.9580	2.5792	7.0850
IP1 / eV	10.44 [120, 121]	10.35	10.48	10.55	10.66	10.42
IP2 / eV	18.76 [122]	18.65	18.88	18.99	19.16	18.76
IP3 / eV	35.27 ± 0.05 [123]	35.00	35.20	35.19	35.17	35.01
IP4 / eV	48.67 ± 0.15 [124]	48.51	48.73	48.74	48.74	48.52

respectively.

To test these simulations, we also fitted a CPP with respect to the first two experimental IPs (expFit). Here, we employed a least-square fit, which resulted in a γ value of 7.0850. The fitted CPP accurately describes the first two IPs, which was expected since these two were used for the fitting procedure (Tab. 4.1). While the third and fourth IPs were less accurately described, they were still within the experimental variance.

To further test the generated CPPs and their influence in molecular simulations, we simulated spectroscopic constants of the two systems Hg_2 and HgF_4.

4.2 Hg_2

The mercury dimer Hg_2 represents the least complex molecular system containing mercury, for which experimental and theoretical data have been reported in the literature. However, none of the theoretical approaches in the literature have employed the CPP ansatz as derived by Schwerdtfeger and Silberbach.

4.2.1 Computational details

For investigating the influence of the CPPs, we used the four generated CPPs in CCSD(T) simulations. Additionally, the counterpoise (CP) correction [125] was used to account for the basis set superposition error (BSSE) [126, 127]. To probe for the influence, three spectroscopic constants of the dimer were determined: the bond

distance r_e, the dissociation energy D_e and the harmonic frequency ω_e. All values were obtained by single point simulations for bond distances in a range between 6.5 and 8.0 Å with a step size of 0.1 Å. Seven of the obtained data points were fitted by a polynomial of order five using the program `mniv` provided by M. Dolg. To understand the influence of the CP correction, the CPPs and the contraction of the basis set, we performed four sets of simulations for each CPP (non, stat, dyn, statDyn and expFit). These combinations vary in the contraction of the basis and the use of the CP correction. The two possible contractions are the contracted basis sets ((12s12p9d3f2g)/[6s6p4d3f2g]) and the uncontracted basis set which was augmented by additional basis functions (15s14p11d5f3g). A detailed description of the basis sets can be found elsewhere [116].

The goal of these simulations was to reproduce literature values and further extend them by employing the CPP method. Note, that the MOLPRO implementation [13] of the CPP ansatz is limited to angular momenta in the basis set, which are smaller than five. Hence, no basis set extrapolation (BSE) was performed.

4.2.2 Spectroscopic constants

First we compared the simulations with and without the CP correction. This correction has the biggest influence on the spectroscopic constants. For example, when comparing the simulations with and without the CP correction using the contracted basis set, the r_e, D_e and ω_e differ by about 0.13 Å, 150 cm^{-1} and 3.5 cm^{-1}, respectively (Tab. 4.3). This also holds true for the uncontracted basis set.

The second biggest influence with a similar order of magnitude is the use of different basis sets. Here, the deviation between both basis sets is about 0.1 Å for the bond distance r_e. Changes in D_e and w_e differed depending on the use of the CP correction. With the CP correction the difference in D_e is about 120 cm^{-1}, while without the CP correction it is only about 20 cm^{-1}. For w_e the changes are about 2.1 cm^{-1} and 1.1 cm^{-1} for with and without CP correction, respectively. While we could not reproduce the spectroscopic constants published by Figgen *et al.* [116], the qualitative behaviour regarding the CP correction can be observed.

Employing the different CPPs yields different values when compared to the simulations without CPPs. Of note, these changes were not very pronounced in simulations without the CP correction. For example, the differences between stat and no CPP for the contracted basis set without CP correction were only 0.0037 Å, 1.36 cm^{-1} and 0.06 cm^{-1}, whereas for simulations with the CP correction they accounted for 0.025 Å,

Tab. 4.3: Spectroscopic constants of Hg_2 obtained by using different cutoff functions according to section 4.1. The values are compared to both experimental and theoretical literature values. Additionally, the influence of the CP correction and the type of basis set is depicted.

CPP	correction	r_e / Å	D_e / cm^{-1}	ω_e / cm^{-1}
		(12s12p9d3f2g)/[6s6p4d3f2g]		
no		3.7738	404.93	19.57
stat		3.7701	406.29	19.63
dyn		3.7702	405.35	19.58
statDyn		3.7701	404.05	19.63
expFit		3.7629	403.00	19.76
no	CP	3.9093	265.98	16.41
stat	CP	3.9343	254.63	16.03
dyn	CP	3.9358	253.84	15.94
statDyn	CP	3.9366	253.16	16.04
expFit	CP	3.9358	250.35	16.30
		15s14p11d5f3g		
no		3.6744	426.99	20.68
stat		3.6713	427.77	20.90
dyn		3.6714	427.01	20.86
statDyn		3.6710	426.47	20.90
expFit		3.6653	426.09	20.86
no	CP	3.7928	344.43	18.56
stat	CP	3.8178	329.67	18.15
dyn	CP	3.8196	328.52	18.12
statDyn	CP	3.8208	327.53	18.15
expFit	CP	3.8223	324.51	18.19
Literature (theoretical)				
Figgen *et al.*[a]		3.760	372.00	19.00
	CP	3.830	327.00	17.90
CBS/CCSD(T)+SO+ΔT[b]		3.679	392.00	20.40
BSS+CCSD(T)[c]	CP	3.744	403.28	-
Literature (experimental)				
Koperski *et al.*		3.69 ± 0.01 [128]	380 ± 15 [129]	19.6 ± 0.5 [129]
Koperski *et al.*		3.605 ± 0.009 [130]		
Greif and Hensel			407[d]	

[a] Reference [116]
[b] Reference [131]
[c] Reference [132]
[d] PhD thesis: [133]; mentioned in Ref. [131].

11.35 cm^{-1} and 0.38 cm^{-1}. Importantly, the observed changes did not significantly depend on the used CPP. This holds especially true for the three *ab initio* CPPs (stat, dyn, statDyn). Here, the maximal deviation was 0.003 Å (uncontracted basis set, CP), 2 cm^{-1} (uncontracted basis set, CP) and 0.1 cm^{-1} (contracted basis set, CP). In most cases, the values obtained by the expFit CPP are not different from the other values. An exception was the bond distance, for which the deviation was more pronounced for simulations without the CP correction (0.008 and 0.005 Å for the uncontracted and contracted basis, respectively).

For the discussion of the obtained values with theoretical and experimental literature values, we will only refer to the results obtained with the uncontracted basis set and CP correction. This is due to the fact that the uncontracted basis set uses more basis functions for describing the orbitals and should thus result in the more accurate simulations. Additionally, only CP-corrected values will be discussed since all theoretical literature values correct for the BSSE. For the comparison with theoretical literature values, we chose the following two studies:

- Complete basis set (CBS) CCSD(T) simulations by Pahl *et al.*, which were corrected by spin-orbit (SO) contributions and a triples correction [131].
- AE CCSD(T) simulations by Borschevsky *et al.*, which employ infinite-order two-component relativistic Hamiltonian obtained by a Barysz-Sadlej-Snijders (BSS) transformation. Counterpoise corrections were performed [132].

All our obtained values deviate significantly from the simulations performed in the literature. The bond distance is at least 0.05 Å larger then the literature values, while the dissociation energy and harmonic frequency are at least 48 cm^{-1} and 1.84 cm^{-1} smaller. Furthermore, we observe that the use of CPPs increases these deviations, rather than improve the accuracy of the performed simulations.

This is even more pronounced when comparing the obtained values to experimental values. This, however, is due to the fact that the theoretical literature values overestimate the bond distance and underestimate the dissociation energy and harmonic frequency. Overall, our simulations are not accurate enough to describe the experimental values. This could be due to a missing SO correction and the approximation of the BSSE via the CP correction. To investigate this further, future work should perform a basis set extrapolation and use more accurate CC methods, like CCSDT and CCSDTQ, in the PP framework. If the results from these improved simulations show better agreement with the experimental values, CPPs could then be added to further improve the results.

4.3 HgF_4

In addition to Hg_2, we investigated a second test system, mercury (IV) fluoride (HgF_4), which is not only of theoretical but also experimental interest. From a theoretical chemist's point of view it is interesting, because it is stabilized by "relativistic effects", which is not the case for the lighter homologous molecules ZnF_4 and CdF_4 [134, 135]. For an experimentalist, it is interesting due to its weak stability and therefore difficult synthesize. This is also the reason why the experimental confirmation [135] of this molecule took place 13 years after its theoretical prediction [134]. The experimental confirmation was achieved by matrix isolation in a solid argon phase with IR spectroscopy at 4 K and represents the first (successful) measurement of a (+IV) mercury compound. In recent years, the existence of another mercury fluoride (HgF_6) was predicted to be at least kinetically stable [136].

4.3.1 Computational details

We simulated the HgF_4 system by varying the Hg-F bond distance in a range from 1.830 to 1.910 Å in 0.005 Å increments using the MOLPRO program package (version 2020.2) [13]. Analogous to Hg_2 (section 4.2), the single point calculations were fitted to a polynomial of order five to obtain the bond distance r_e, the harmonic frequency ω_e, the force constant k and the dissociation energy D_e. To reduce computational demands and incorporate scalar-relativistic effects, we employed the 60MDF pseudopotential and the corresponding basis set for Hg [116]. For F, we employed the aug-cc-pVQZ [137] basis set. Additionally, we used the CP correction to address the BSSE. The simulations were performed with and without the expFit CPP. Furthermore, we performed an AE-DKH3 (third-order Douglas-Kroll-Hess Hamiltonian) CCSD(T) simulation, which also employed the CP correction. For the AE simulations the aug-cc-pVTZ-DK3 [138] and aug-cc-pVQZ [137] basis sets were used for Hg and F, respectively.

4.3.2 Spectroscopic constants

The CCSD(T) simulation without the expFit yielded a bond distance and harmonic frequency, which are in agreement with other theoretical studies in the literature [135, 136, 139] (Tab. 4.4). Here, the deviation of r_e ranges between 0.0007 and 0.0057 Å, while the deviation of ω_e ranges between 5.09 and 7.79 cm^{-1}. Employing the expFit CPP decreases the bond distance by 0.0029 Å and increases ω_e by 3.40 cm^{-1}. These

obtained values also agree with other theoretical studies [135, 136, 139]. Note, that all theoretical studies in the literature try to reproduce the values published by Wang *et al.* [135]. To establish a new reference values we performed a different simulation. Performing an AE-DKH3 CCSD(T) simulation results in spectroscopic constants, which are in the range of the literature values, but differ from our CCSD(T) simulations without the expFit CPP (Tab. 4.4). To understand this deviation in more detail, we also compared k and D_e. We assume that the values obtained by the AE-DKH3 simulation are the most accurate values, because the simulation was performed on the highest level of theory. Hence, if the CPP is able to contribute for missing core-valence correlation, including the expFit CPP should result in values closer to these values. Employing the CPP resulted in a decreased deviation of ω_e, k and D_e from 4.95 cm^{-1}, 0.0144 a.u. and 0.0014 a.u. to 1.55 cm^{-1}, 0.0045 a.u. and 0.0009 a.u. While this is an improvement, the deviation for r_e increased from 0.0015 to 0.0044 Å. Thus, an increased accuracy is not achieved for all values. Of note, there are no experimental values for r_e, k and D_e published in the literature. Thus, the results of our simulations should be interpreted with caution.

Tab. 4.4: Comparison of the spectroscopic constants of HgF_4 by PP simulations with and without expFit CPP with an AE simulation and other theoretical and experimental studies.

	r_e / Å	ω_e / cm^{-1}	k / a.u.	D_e / a.u.
CCSD(T)[a]+CP	1.8897	693.79	1.0037	0.2661
CCSD(T)[a]+CP+expFit	1.8868	697.19	1.0136	0.2656
AE-DKH3 CCSD(T)+CP	1.8912	697.69	1.0151	0.2647
CCSD(T)[b] [135]	1.885	686.00		
PBE0[a]+SO [139]	1.884	688.70		
CCSD(T)[a]+SO [139]	1.889	687.60		
X1C/X2C-PBE0[a]+SO [136]	1.885/1.883	687.00/-		
exp. (predicted) [135]		$\sim$ 690.00		
exp. (neon-matrix) [135]		703.00		
exp. (argon-matrix) [135]		682.00		

[a] ECP60MDF [116]
[b] ECP60MHF [140]

4.4 Conclusion of the CPP simulations

The two test systems Hg_2 and HgF_4 allow for a preliminary investigation of how CPPs affect the simulations of these systems in the context of the spectroscopic constants obtained. We observed that these values change when employing CPP, but that these changes are relatively small. Furthermore, these changes can decrease the accuracy of the simulations as seen in Hg_2 or increase and decrease the accuracy dependent on the spectroscopic constant as seen in HgF_4. In general, the choice of the CPP does not alter the result when comparing simulations with and without CPPs. This is plausible, because the CPPs only differ in the cutoff parameter, which ensures the numeric stability of the CPP ansatz. We can envision multiple explanations for the unexpectedly small changes of employing CPPs. First, due to the small core of the pseudopotential, the improvement was expected to be smaller than of a large core would have been chosen [1]. A good example for the effective use of CPPs in large-core pseudopotentials is the work of Weigand *et al.* for the di-, tri- and tetravalent actinides and lanthanides [10, 11]. There, performing large-core pseudopotentials in combination with custom CPPs yielded simulated properties of equal accuracy as the small-core pseudopotential counterpart. Second, we suspect that the closed-shell character and the symmetry of the investigated systems minimize the contribution of the CPP ansatz. The reason behind this may be that the induced dipole moments in these systems cancel each other out. Finally, our results should be interpreted with caution, because either our underlying simulations are not of satisfying accuracy (Hg_2) or there are not enough experimental values (HgF_4) available to obtain a reasonable and predictive conclusion.

Summary and Outlook

CPP implementation

We successfully implemented the CPP ansatz in both the integral evaluation of the SCF program and the interface between the SCF and the CC program, namely the MO transformation. In both cases, the two-electron CPP integrals were expressed using the one-electron integrals $v_{ij}^{\lambda t}$, which decreased the runtime scaling. Combining the SCF integrals and the CPP integrals was achieved through straightforward matrix-matrix additions during the SCF procedure. To verify our implementation, we then tested it in multiple atomic and molecular test simulations both on the Hartree-Fock and CC level. We observed no significant deviations from similar simulations performed with MOLPRO, confirming the successful implementation of the CPP ansatz.

The herein described implementation extends the features of the QOL and provides a framework for future investigations of the influence of the CPP in molecular systems in greater detail. The current implementation of the ARGOS [95] integral evaluation for the pseudopotential ansatz in the QOL is limited to the use of basis set functions of angular momenta smaller than five. Overcoming this limitation by implementing an updated version of the integral evaluation would enable basis set extrapolations. Furthermore, a thorough analysis of numerical errors of the CPP implementation with respect to larger basis sets ($l > 4$) should be performed to ensure a correct behavior on both the SCF and CC level, because this has not been done in the original implementation of Nicklaß (MOLPRO). While a basis set extrapolation is not possible with neither MOLRPO nor QOL yet, the implementation of the QOL allows for separately employing CPPs in the SCF and QOL program. Thus, incorporating core-valence and valence-valence correlation on only the correlated level would allow for reducing the active space in the CC program. In the case of atomic mercury this could be envisioned as follows: First we perform an SCF simulation using the 60MDF PP (20 valence electrons) without CPP. Second, we perform a CC simulation, in which the 5s and 5p electrons are frozen, and employ a custom CPP generated for a 68 valence electron core to describe the polarization of the frozen orbitals. Hence, CC simulations could be performed with drastically reduced computational demands. If this method is successful, it could be transferred to molecular systems, which, however, would require a frozen-core SCF implementation in the QOL. Another advantage of the QOL over MOLRPO is that CC simulations with arbitrary substitutions could be performed. This would allow for a more detailed investigation of the importance of the CPP while increasing the many-particle basis.

Furthermore, the implementation presented herein opens up new opportunities to extend and modify the CPP implementation in the QOL. This lifts the restrictions imposed by the fact that MOLRPO is a commercial product and no source code is published nor accessible. Furthermore, the angular momentum-dependent CPP ansatz by Foucrault *et al.* [76] could be implemented for an even more detailed investigation of CPPs in general. All this would allow for a thorough investigation of the influence and importance of CPPs in molecular systems.

Simulations of Hg_2 and HgF_4

To get a first impression on the influence of the CPPs in molecular systems, we simulated the potential curves of Hg_2 and HgF_4. Here, we extracted the bond distance, harmonic frequency, force constant and the dissociation energy and compared them to theoretical and experimental values. All simulations were performed using the 60MDF small-core energy-consistent PP. For this, we generated new CPPs and investigated their influence on the spectroscopic constants. We observed that the influence of the CPPs was small and did not always result in improved values. These small changes can be explained by the use of the small-core PP, for which the core polarization and core-valence correlation was expected to be small and thus difficult to correct for. The lack of improvement in all values was not expected. This conclusion, however, requires further validation, because the performed simulations were not accurate enough or there are not enough experimental values published.

To continue the investigation of the influence on CPPs in molecular systems, simulations of higher accuracy should be performed. This includes the calculation of spin-orbit contributions and a better description of the BSSE. To account for the BSSE and the basis set incompleteness, a basis set extrapolation is needed. For this, a CPP implementation without limitations on the basis set would be required. To widen the scope of application, additional molecular properties, like transition energies and vibrational frequencies, could be investigated to observe the influence of the CPPs. Furthermore, the fitting of the CPPs should be extended by also optimizing the used basis set or by introducing an angular momentum-dependent cutoff function using the approach by Foucrault *et al.* [76]. With this, the use of CPPs may improve the simulations. At last, the investigation should be extended to other molecular systems to allow for a broad understanding of the influence of the CPPs. Of note, CPPs are commonly used in large-core PPs and thus, the investigation

of the CPPs should be extended for larger cores, too. Here, accurate CPPs could enhance simulations of large molecular systems, by efficiently describing the core polarization and core-valence correlation.

Appendix

A.1 CPP parameters (section 3.6 and 3.7)

Tab. A.1: CPP parameters for the test simulations with both the QOL and MOLPRO CPP implementation. The core charge q is only needed for the simulations with QOL implementation and is linked to the corresponding PP core size.

Element	α_D / a.u.	α_Q / a.u.	β_D / a.u.	γ	q / e^-
Na	0.9947	0.0000	0.0000	0.6200	1.0000
K	1.0000	0.0000	0.0000	1.0000	9.0000
Sc	1.0000	0.0000	0.0000	0.1000	11.000
Mg	1.2345	0.0000	0.0000	1.5000	2.0000
Yb	2.9876	0.0000	0.0000	5.0000	2.0000
Hg	1.0000	0.0000	0.0000	20.000	20.000

A.2 Absolute energies (section 3.6 and 3.7)

Tab. A.2: Comparison of ground state energies obtained by SCF simulations performed with and without CPPs given in a.u. Additionally, the relative influence of the CPPs are given in %.

System	PP	E_{SCF} / a.u.	$E_{\mathrm{SCF+CPP}}$ / a.u.	$\frac{E_{\mathrm{SCF+CPP}}}{E_{\mathrm{SCF}}} - 1$ / %
Mg	10SDF	$-7.8559 \cdot 10^{-01}$	$-8.3152 \cdot 10^{-01}$	5.8
Yb	60MWB	$-3.2541 \cdot 10^{+01}$	$-3.6772 \cdot 10^{+01}$	13.0
Hg	60MHF	$-1.4808 \cdot 10^{+02}$	$-1.5414 \cdot 10^{+02}$	4.1
	60MWB	$-1.5255 \cdot 10^{+02}$	$-1.6005 \cdot 10^{+02}$	4.9
Na_2	10SDF	$-3.5699 \cdot 10^{-01}$	$-3.7137 \cdot 10^{-01}$	4.0
K_2	10MWB	$-5.6044 \cdot 10^{+01}$	$-5.6531 \cdot 10^{+01}$	0.9
Sc_2	10MHF	$-9.2093 \cdot 10^{+01}$	$-9.2099 \cdot 10^{+01}$	0.01
Hg_2	60MHF	$-2.9613 \cdot 10^{+02}$	$-3.0826 \cdot 10^{+02}$	4.1
	60MWB	$-3.0508 \cdot 10^{+02}$	$-3.2008 \cdot 10^{+02}$	4.9

Tab. A.3: Comparison of CCSD correlation energies obtained by simulations performed with and without CPPs given in a.u. Additionally, the relative influence of the CPPs are given in %.

System	PP	E_{corr} / a.u.	$E_{\mathrm{corr+CPP}}$ / a.u.	$\frac{E_{\mathrm{corr+CPP}}}{E_{\mathrm{corr}}} - 1$ / %
Mg	10SDF	$-3.3696 \cdot 10^{-02}$	$-2.6432 \cdot 10^{-02}$	-21.6
Yb	60MWB	$-1.7199 \cdot 10^{-01}$	$-6.3028 \cdot 10^{-01}$	266.5
Hg	60MHF	$-1.2633 \cdot 10^{-01}$	$-1.8154 \cdot 10^{-01}$	43.7
	60MWB	$-5.6632 \cdot 10^{-01}$	$-1.1393 \cdot 10^{+00}$	101.2
Na_2	10SDF	$-2.1431 \cdot 10^{-02}$	$-1.9723 \cdot 10^{-02}$	-8.0
K_2	10MWB	$-1.0180 \cdot 10^{-01}$	$-9.9853 \cdot 10^{-02}$	-1.9
Sc_2	10MHF	$-2.8661 \cdot 10^{-01}$	$-2.8668 \cdot 10^{-01}$	0.02
Hg_2	60MHF	$-2.6441 \cdot 10^{-01}$	$-3.7820 \cdot 10^{-01}$	43.0
	60MWB	$-1.1447 \cdot 10^{+00}$	$-2.3089 \cdot 10^{+00}$	101.7

A.3 Reproducibility

All simulation data presented in this work are stored on the server of the Institute of Theoretical Chemistry. These data are located in the home directory (either `home` or `home_noBackup`) of S. Bubeck in a separate directory named `ThesisData`.

References

1. Dolg, M. & Cao, X. Relativistic Pseudopotentials: Their Development and Scope of Applications. *Chem. Rev.* **112,** 403–480 (2012).
2. Müller, W., Flesch, J. & Meyer, W. Treatment of intershell correlation effects in *ab initio* calculations by use of core polarization potentials. Method and application to alkali and alkaline earth atoms. *J. Chem. Phys.* **80,** 3297–3310 (1984).
3. Müller, W. & Meyer, W. Ground-state properties of alkali dimers and their cations (including the elements Li, Na, and K) from *ab initio* calculations with effective core polarization potentials. *J. Chem. Phys.* **80,** 3311 (1984).
4. Fuentealba, P. On the reliability of semiempirical pseudopotentials: dipole polarisability of the alkali atoms. *J. Phys. B: At. Mol. Phys.* **15,** L555–L558 (1982).
5. Fuentealba, P., Preuss, H., Stoll, H. & Szentpály, L. V. A proper account of core-polarization with pseudopotentials: single valence-electron alkali compounds. *Chem. Phys. Lett.* **89,** 418–422 (1982).
6. G. Igel-Mann, U. W., Fuentealba, P. & Stoll, H. Ground-state properties of alkali dimers XY (X, Y=Li to Cs). *J. Chem. Phys.* **84,** 5007–5012 (1986).
7. Igel-Mann, G., Stoll, H. & Preuss, H. Pseudopotentials for main group elements (IIIa through VIIa). *Mol. Phys.* **65,** 1321–1328 (1988).
8. Schwerdtfeger, P., Fischer, T., Dolg, M., Igel-Mann, G., Nicklass, A., Stoll, H. & Haaland, A. The accuracy of the pseudopotential approximation. I. An analysis of the spectroscopic constants for the electronic ground states of InCl and $InCl_3$ using various three valence electron pseudopotentials for indium. *J. Chem. Phys.* **102,** 2050–2062 (1995).
9. Lee, Y. & Needs, R. J. Core-polarization potentials for Si and Ti. *Phys. Rev. B* **67** (2003).
10. Moritz, A., Cao, X. & Dolg, M. Quasirelativistic energy-consistent 5f-in-core pseudopotentials for divalent and tetravalent actinide elements. *Theor. Chem. Acc.* **118,** 845–854 (2007).
11. Weigand, A., Cao, X., Yang, J. & Dolg, M. Quasirelativistic f-in-core pseudopotentials and core-polarization potentials for trivalent actinides and lanthanides: molecular test for trifluorides. *Theor. Chem. Acc.* **126,** 117–127 (2009).
12. Schwerdtfeger, P. & Silberbach, H. Multicenter integrals over long-range operators using Cartesian Gaussian functions. *Phys. Rev. A* **37,** 2834–2842 (1988).
13. Werner, H.-J., Knowles, P. J., Manby, F. R., Black, J. A., Doll, K., Heßelmann, A., Kats, D., Köhn, A., Korona, T., Kreplin, D. A., Ma, Q., Miller, T. F., Mitrushchenkov, A., Peterson, K. A., Polyak, I., Rauhut, G. & Sibaev, M. The Molpro quantum chemistry package. *J. Chem. Phys.* **152,** 144107 (2020).
14. Nicklaß, A. *Pseudopotentiale für die Edelgase* MA thesis (Universität Stuttgart, 1990).
15. Hanrath, M. Quantum Objects Library (*QOL*) (2006-2012).
16. Schrödinger, E. Quantisierung als Eigenwertproblem. *Ann. Phys.* **384,** 361–376 (1926).
17. Schrödinger, E. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.* **28,** 1049–1070 (1926).

18. Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z. Phys.* **61,** 126–148 (1930).
19. Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Math. Proc. Camb. Phil. Soc.* **24,** 89 (1928).
20. Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion. *Math. Proc. Camb. Phil. Soc.* **24,** 111 (1928).
21. Born, M. & Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **389,** 457–484 (1927).
22. Szabo, A. & Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Eletronic Structure Theory* (Mineola: Dover Publications Inc, 1989).
23. Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Z. Phys.* **31,** 765–783 (1925).
24. Pauli, W. The Connection Between Spin and Statistics. *Phys. Rev.* **58,** 716–722 (1940).
25. Slater, J. C. The Theory of Complex Spectra. *Phys. Rev.* **34,** 1293–1322 (1929).
26. Condon, E. U. The Theory of Complex Spectra. *Phys. Rev.* **36,** 1121–1133 (1930).
27. Löwdin, P.-O. Quantum Theory of Many-Particle Systems. I. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction. *Phys. Rev.* **97,** 1474–1489 (1955).
28. Laaksonen, L., Pyykkö, P. & Sundholm, D. Two-dimensional fully numerical solutions of molecular Schrödinger equations. I. One-electron molecules. *Int. J. Quantum Chem.* **23,** 309–317 (1983).
29. Laaksonen, L., Pyykkö, P. & Sundholm, D. Two-Dimensional fully numerical solutions of molecular Schrödinger equations. II. Solution of the Poisson equation and results for singlet states of H_2 and HeH^+. *Int. J. Quantum Chem.* **23,** 319–323 (1983).
30. Kobus, J., Laaksonen, L. & Sundholm, D. A numerical Hartree-Fock program for diatomic molecules. *Comput. Phys. Commun.* **98,** 346–358 (1996).
31. Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **23,** 69–89 (1951).
32. Hall, G. G. The Molecular Orbital Theory of Chemical Valency. VIII. A Method of Calculating Ionization Potentials. *Proc. Roy. Soc. A* **205,** 541–552 (1951).
33. Levine, I. N. *Quantum Chemistry* 5th ed. (Prentice Hall: Upper Saddle River, NJ, 1999).
34. Atkins, P. W. & Friedman, R. S. *Molecular quantum mechanics* 4th ed. (Oxford University Press, 2005).
35. Crawford, T. D. & Schaefer, H. F. in *Reviews in Computational Chemistry* 33–136 (John Wiley & Sons, Inc., 2007).
36. Jensen, F. *Introduction to Computational Chemistry* 2nd ed. (John Wiley & Sons, Ltd, 2007).
37. Surján, P. R. & Szabados, Á. in *Recent Progress in Coupled Cluster Methods: Theory and Applications* (eds Cársky, P., Paldus, J. & Pittner, J.) 513–534 (Springer Netherlands, Dordrecht, 2010).

38. Kutzelnigg, W. Error analysis and improvements of coupled-cluster theory. *Theor. Chim. Acta* **80,** 349–386 (1991).
39. Ramabhadran, R. O. & Raghavachari, K. Extrapolation to the Gold-Standard in Quantum Chemistry: Computationally Efficient and Accurate CCSD(T) Energies for Large Molecules Using an Automated Thermochemical Hierarchy. *J. Chem. Theory Comput.* **9,** 3986–3994 (2013).
40. Bartlett, R. J. & Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **79,** 291–352 (1 2007).
41. Hanrath, M. Multi-reference coupled-cluster study of the ionic-neutral curve crossing LiF. *Mol. Phys.* **106,** 1949–1957 (2008).
42. Lyakh, D. I., Musiał, M., Lotrich, V. F. & Bartlett, R. J. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chem. Rev.* **112,** 182–243 (2012).
43. Pyykkö, P. in *Advances in Quantum Chemistry Volume 11* 353–409 (Elsevier BV, 1978).
44. Pitzer, K. S. Relativistic effects on chemical properties. *Acc. Chem. Res.* **12,** 271–276 (1979).
45. Pyykkö, P. & Desclaux, J. P. Relativity and the periodic system of elements. *Acc. Chem. Res.* **12,** 276–281 (1979).
46. Pyykkö, P. Relativistic effects in structural chemistry. *Chem. Rev.* **88,** 563–594 (1988).
47. Autschbach, J., Siekierski, S., Seth, M., Schwerdtfeger, P. & Schwarz, W. H. E. Dependence of relativistic effects on electronic configuration in the neutral atoms of d- and f-block elements. *J. Comput. Chem.* **23,** 804–813 (2002).
48. Pyykkö, P. Relativistic Effects in Chemistry: More Common Than You Thought. *Annu. Rev. Phys. Chem.* **63,** 45–64 (2012).
49. Reiher, M. & Wolf, A. *Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science* (Wiley-VCH, 2014).
50. Visser, O., Visscher, L., Aerts, P. J. C. & Nieuwpoort, W. C. Relativistic all-electron molecular Hartree-Fock-Dirac-(Breit) calculations on CH_4, SiH_4, GeH_4, SnH_4, PbH_4. *Theor. Chim. Acta* **81,** 405–416 (1992).
51. Brown, G. E. & Ravenhall, D. G. On the Interaction of Two Electrons. *Proc. Roy. Soc. A* **208,** 552–559 (1951).
52. Dyall, K., Grant, I., Johnson, C., Parpia, F. & Plummer, E. GRASP: A general-purpose relativistic atomic structure program. *Comput. Phys. Commun.* **55,** 425–456 (1989).
53. Dolg, M. On the accuracy of valence correlation energies in pseudopotential calculations. *J. Chem. Phys.* **104,** 4061 (1996).
54. Dolg, M. *Effective Core Potentials* in *Modern Methods and Algorithms of Quantum Chemistry* (ed Grotendorst, J.) **3** (John von Neumann Institute for Computing, 2000), 507–540.
55. Dolg, M. in *Theoretical and Computational Chemistry* 793–862 (Elsevier BV, 2002).
56. Stoll, H., Metz, B. & Dolg, M. Relativistic energy-consistent pseudopotentials - Recent developments. *J. Comput. Chem.* **23,** 767–778 (2002).
57. Schwarz, W. H. E. Das Kombinierte Näherungsverfahren. *Theor. Chim. Acta* **11,** 307–324 (1968).
58. Kahn, L. R. & III, W. A. G. *Ab Initio* Effective Potentials for Use in Molecular Calculations. *J. Chem. Phys.* **56,** 2685–2701 (1972).

59. Ermler, W. C., Lee, Y. S., Christiansen, P. A. & Pitzer, K. S. *Ab initio* effective core potentials including relativistic effects. A procedure for the inclusion of spin-orbit coupling in molecular wavefunctions. *Chem. Phys. Lett.* **81,** 70–74 (1981).
60. Cohen, M. H. & Heine, V. Cancellation of Kinetic and Potential Energy in Atoms, Molecules, and Solids. *Phys. Rev.* **122,** 1821–1826 (1961).
61. Topp, W. C. & Hopfield, J. J. Chemically Motivated Pseudopotential for Sodium. *Phys. Rev. B* **7,** 1295–1303 (1973).
62. `http://www.tc.uni-koeln.de/PP/index.en.html`.
63. Dolg, M., Stoll, H. & Preuss, H. Energy-adjusted *ab initio* pseudopotentials for the rare earth elements. *J. Chem. Phys.* **90,** 1730 (1989).
64. Dolg, M., Stoll, H., Savin, A. & Preuss, H. Energy-adjusted pseudopotentials for the rare earth elements. *Theor. Chim. Acta* **75,** 173–194 (1989).
65. Dolg, M., Stoll, H., Preuss, H. & Pitzer, R. M. Relativistic and correlation effects for element 105 (hahnium, Ha): a comparative study of M and MO (M = Nb, Ta, Ha) using energy-adjusted *ab initio* pseudopotentials. *J. Phys. Chem.* **97,** 5852–5859 (1993).
66. Born, M. & Heisenberg, W. Über den Einfluß der Deformierbarkeit der Ionen auf optische und chemische Konstanten. I. *Z. Phys.* **23,** 388–410 (1924).
67. Chang, T. C., Habitz, P., Pittel, B. & Schwarz, W. H. E. Accuracy and limitations of the pseudopotential method. *Theor. Chim. Acta* **34,** 263–275 (1974).
68. Chang, T. C., Habitz, P. & Schwarz, W. H. E. Accuracy and limitations of the pseudopotential method. *Theor. Chim. Acta* **44,** 61–76 (1977).
69. Bardsley, J. in *Case Studies in Atomic Physics 4* 299–368 (Elsevier, 1974).
70. Edward S. Sachs Juergen Hinze, N. H. S. Frozen core approximation, a pseudopotential method tested on six states of NaH. *J. Chem. Phys.* **62,** 3393 (1975).
71. Bauschlicher, C. W., Langhoff, S. R. & Taylor, P. R. Core–core and core–valence correlation. *J. Chem. Phys.* **88,** 2540–2546 (1988).
72. Bardsley, J. Pseudopotential calculations of alkali interactions. *Chem. Phys. Lett.* **7,** 517–520 (1970).
73. Fuentealba, P. *Die Rolle der Rumpfpolarisierung in Pseudopotentialverfahren* PhD thesis (Universität Stuttgart, 1982).
74. Igel-Mann, G. *Semiempirsche Pseudopotentiale: Untersuchung an Hauptgruppenelementen und Nebengruppenelementen mit abgeschlossener d-Schale* PhD thesis (Universität Stuttgart, 1987).
75. Weigand, A., Cao, X., Hangele, T. & Dolg, M. Relativistic Small-Core Pseudopotentials for Actinium, Thorium, and Protactinium. *J. Phys. Chem. A* **118,** 2519–2530 (2014).
76. Foucrault, M., Millie, P. & Daudey, J. P. Nonperturbative method for core-valence correlation in pseudopotential calculations: Application to the Rb_2 and Cs_2 molecules. *J. Chem. Phys.* **96,** 1257–1264 (1992).
77. Mtiri, S. & Oujia, B. Spectroscopic, vibrational and structural properties analysis of $CaXe_n$ (n = 1-4) clusters. *Comput. Theor. Chem.* **1151,** 58–71 (2019).
78. Ayed, M. B. H., Mtiri, S. & Ghalla, H. Adiabatic investigation of the electronic and dipolar properties of the LiNe system in ground state and numerous excited states. *J. Quant. Spectrosc. Radiat. Transfer* **274,** 107864 (2021).

79. Souissi, H., Mejrissi, L., Habli, H., Alsahhaf, M., Oujia, B. & Gadéa, F. X. *Ab initio* diabatic and adiabatic calculations for francium hydride FrH. *New J. Chem.* **44,** 5572–5587 (2020).
80. Bottcher, C. & Dalgarno, A. A constructive model potential method for atomic interactions. *Proc. Roy. Soc. London* **340,** 187–198 (1974).
81. Jeung, G. H., Malrieu, J. P. & Daudey, J. P. Inclusion of core-valence correlation effects in pseudopotential calculations. I. Alkali atoms and diatoms. *J. Chem. Phys.* **77,** 3571–3577 (1982).
82. Jeung, G. H., Malrieu, J. P. & Daudey, J. P. Inclusion of core-valence correlation effects in pseudopotential calculations. II. K_2 and KH lowest Σ^+ potential curves from valence-correlated wavefunctions. *J. Chem. Phys.* **16,** 699–714 (1983).
83. Schwerdtfeger, P. & Silberbach, H. Erratum: Multicenter integrals over long-range operators using Cartesian Gaussian functions. *Phys. Rev. A* **42,** 665–665 (1990).
84. Schwerdtfeger, P. & Silberbach, H. Erratum: Multicenter integrals over long-range operators using Cartesian Gaussian functions [Phys. Rev. A 37 , 2834 (1988)]. *Phys. Rev. A* **103** (2021).
85. Davidson, E. R. The Right Answer for the Right Reason: My Personal Goal for Quantum Chemistry. *Annu. Rev. Phys. Chem.* **70,** 1–20 (2019).
86. Elbert, S. T. & Davidson, E. R. Evaluation of electron repulsion integrals over gaussian lobe basis functions. *J. Comput. Phys.* **16,** 391–395 (1974).
87. McMurchie, L. E. & Davidson, E. R. One- and two-electron integrals over cartesian gaussian functions. *J. Comput. Phys.* **26,** 218–231 (1978).
88. McMurchie, L. E. & Davidson, E. R. Calculation of integrals over *ab initio* pseudopotentials. *J. Comput. Phys.* **44,** 289–301 (1981).
89. Ciupka, J. *Localization Scheme for Relativistic Spinors and Optimized Electron Repulsion Integral Evaluation and Transformation* PhD thesis (University of Cologne, 2013).
90. Ciupka, J., Hanrath, M. & Dolg, M. Localization scheme for relativistic spinors. *J. Chem. Phys.* **135,** 244101 (2011).
91. Held, J. *Closed-Shell Hartree-Fock: an Efficient Implementation Based on the Contraction of Integrals in the Primitive Basis* PhD thesis (University of Cologne, 2017).
92. Held, J., Hanrath, M. & Dolg, M. An Efficient Hartree–Fock Implementation Based on the Contraction of Integrals in the Primitive Basis. *J. Chem. Theory Comput.* **14,** 6197–6210 (2018).
93. Pitzer, R. M. Contribution of atomic orbital integrals to symmetry orbital integrals. *J. Chem. Phys.* **58,** 3111 (1973).
94. Harrison, R. J. & Kendall, R. A. A parallel version of ARGOS: A distributed memory model for shared memory UNIX computers. *Theor. Chim. Acta* **79,** 337–347 (1991).
95. Pitzer, R. M. & Winter, N. W. Spin-orbit (core) and core potential integrals. *Int. J. Quantum Chem.* **40,** 773–780 (1991).
96. Wiebke, J. *Kramers-restricted self-consistent 2-spinor fields for heavy-element chemistry* PhD thesis (University of Cologne, 2010).
97. Head-Gordon, M. & Pople, J. A. A method for two-electron Gaussian integral and integral derivative evaluation using recurrence relations. *J. Chem. Phys.* **89,** 5777–5786 (1988).

98. Hamilton, T. P. & Schaefer, H. F. New variations in two-electron integral evaluation in the context of direct SCF procedures. *Chem. Phys.* **150,** 163–171 (1991).
99. Lindh, R., Ryu, U. & Liu, B. The reduced multiplication scheme of the Rys quadrature and new recurrence relations for auxiliary function based two-electron integral evaluation. *J. Chem. Phys.* **95,** 5889–5897 (1991).
100. Obara, S. & Saika, A. Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *J. Chem. Phys.* **84,** 3963–3974 (1986).
101. Obara, S. & Saika, A. General recurrence formulas for molecular integrals over Cartesian Gaussian functions. *J. Chem. Phys.* **89,** 1540–1559 (1988).
102. Helgaker, T., Jorgensen, P. & Olsen, J. *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd, Chichester New York, 2000).
103. Boys, S. F. Electronic wave functions - I. A general method of calculation for the stationary states of any molecular system. *Proc. Roy. Soc. London* **200,** 542–554 (1950).
104. *GNU Compiler Collection (10.2)* `https://gcc.gnu.org`.
105. *BOOST C++ Libraries (1.75.0)* `https://www.boost.org`.
106. *GNU Scientific Library* `https://www.gnu.org/software/gsl/`.
107. Intel. *Math Kernel Library (11.3)* `https://software.intel.com/oneapi/onemkl`.
108. Fuentealba, P., von Szentpaly, L., Preuss, H. & Stoll, H. Pseudopotential calculations for alkaline-earth atoms. *J. Phys. B: At. Mol. Phys.* **18,** 1287–1296 (1985).
109. Dolg, M., Stoll, H. & Preuss, H. A combination of quasirelativistic pseudopotential and ligand field calculations for lanthanoid compounds. *Theor. Chim. Acta* **85,** 441–450 (1993).
110. Andrae, D., Häußermann, U., Dolg, M., Stoll, H. & Preuß, H. Energy-adjusted *ab initio* pseudopotentials for the second and third row transition elements. *Theor. Chim. Acta* **77,** 123–141 (1990).
111. Martin, J. M. L. & Sundermann, A. Correlation consistent valence basis sets for use with the Stuttgart-Dresden-Bonn relativistic effective core potentials: The atoms Ga-Kr and In-Xe. *J. Chem. Phys.* **114,** 3408–3420 (2001).
112. Leininger, T., Nicklass, A., Küchle, W., Stoll, H., Dolg, M. & Bergner, A. The accuracy of the pseudopotential approximation: non-frozen-core effects for spectroscopic constants of alkali fluorides XF (X = K, Rb, Cs). *Chem. Phys. Lett.* **255,** 274–280 (1996).
113. Dolg, M., Wedig, U., Stoll, H. & Preuss, H. Energy-adjusted *ab initio* pseudopotentials for the first row transition elements. *J. Chem. Phys.* **86,** 866 (1987).
114. Hanrath, M. personal communication. 2019.
115. Hülsen, M. *Optimization of Integral Transformations and Quantum Chemical Calculations on Organometallic Templates* PhD thesis (University of Cologne, 2012).
116. Figgen, D., Rauhut, G., Dolg, M. & Stoll, H. Energy-consistent pseudopotentials for group 11 and 12 atoms: adjustment to multi-configuration Dirac–Hartree–Fock data. *Chem. Phys.* **311,** 227–244 (2005).
117. Kolb, D., Johnson, W. R. & Shorer, P. Electric and magnetic susceptibilities and shielding factors for closed-shell atoms and ions of high nuclear charge. *Phys. Rev. A* **26,** 19–31 (1982).

118. Roos, B. O., Lindh, R., Malmqvist, P.-Å., Veryazov, V. & Widmark, P.-O. New Relativistic ANO Basis Sets for Transition Metal Atoms. *J. Phys. Chem. A* **109,** 6575–6579 (2005).
119. `https://www.molcas.org/ANO/ANO-RCC`.
120. Zia, M. A., Suleman, B. & Baig, M. A. Two-photon laser optogalvanic spectroscopy of the Rydberg states of mercury by RF discharge. *J. Phys. B: At., Mol. Opt. Phys.* **36,** 4631–4639 (2003).
121. Zia, M. A. & Baig, M. A. Two-step laser optogalvanic spectroscopy of the odd-parity Rydberg states of atomic mercury. *Eur. Phys. J. D* **28,** 323–330 (2004).
122. Sansonetti, C. J. & Reader, J. Spectrum and Energy Levels of Singly-Ionized Mercury (Hg II). *Phys. Scr.* **63,** 219–242 (2001).
123. Rashid, A. & Tauheed, A. Revised analysis of the third spectrum of mercury: Hg III. *J. Quant. Spectrosc. Radiat. Transfer* **233,** 119–133 (2019).
124. Huttula, M., Huttula, S.-M., Lablanquie, P., Palaudoux, J., Andric, L., Eland, J. H. D. & Penent, F. Spectroscopy of triply and quadruply ionized states of mercury. *Phys. Rev. A* **83** (2011).
125. Boys, S. & Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **19,** 553–566 (1970).
126. Davidson, E. R. & Feller, D. Basis set selection for molecular calculations. *Chem. Rev.* **86,** 681–696 (1986).
127. Richard, R. M., Bakr, B. W. & Sherrill, C. D. Understanding the Many-Body Basis Set Superposition Error: Beyond Boys and Bernardi. *J. Chem. Theory Comput.* **14,** 2386–2400 (2018).
128. Koperski, J., Atkinson, J. & Krause, L. The $G0_u^+(6^1P_1) - X0_g^+$ Excitation and Fluorescence Spectra of Hg_2 Excited in a Supersonic Jet. *J. Mol. Spectrosc.* **184,** 300–308 (1997).
129. Koperski, J., Atkinson, J. & Krause, L. The $0_u^+(6^3P_1)\leftarrow X0^+g$ spectrum of Hg_2 excited in a supersonic jet. *Chem. Phys. Lett.* **219,** 161–168 (1994).
130. Koperski, J., Qu, X., Meng, H., Kenefick, R. & Fry, E. Rotational analysis of the (57,0) band of the $D1_u \leftarrow X0_g^+$ triplet-singlet transition in Hg_2 produced in a free-jet expansion beam. *Chem. Phys.* **348,** 103–112 (2008).
131. Pahl, E., Figgen, D., Thierfelder, C., Peterson, K. A., Calvo, F. & Schwerdtfeger, P. A highly accurate potential energy curve for the mercury dimer. *J. Chem. Phys.* **132,** 114301 (2010).
132. Borschevsky, A., Pershina, V., Eliav, E. & Kaldor, U. Relativistic coupled cluster study of diatomic compounds of Hg, Cn, and Fl. *J. Chem. Phys.* **141,** 084301 (2014).
133. Greif-Wüstenbecker, J. N. PhD thesis (Philipps University, 2000).
134. Kaupp, M., Dolg, M., Stoll, H. & von Schnering, H. G. Oxidation State +IV in Group 12 Chemistry. *Ab Initio* Study of Zinc(IV), Cadmium(IV), and Mercury(IV) Fluorides. *Inorg. Chem.* **33,** 2122–2131 (1994).
135. Wang, X., Andrews, L., Riedel, S. & Kaupp, M. Mercury Is a Transition Metal: The First Experimental Evidence for HgF_4. *Angew. Chem. Int. Ed.* **46,** 8371–8375 (2007).

136. Gao, C., Hu, S.-X., Han, H., Guo, G., Suo, B. & Zou, W. Exploring the electronic structure and stability of HgF_6: Exact 2-Component (X2C) relativistic DFT and NEVPT2 studies. *Comput. Theor. Chem.* **1160,** 14–18 (2019).
137. Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90,** 1007–1023 (1989).
138. Figgen, D., Peterson, K. A., Dolg, M. & Stoll, H. Energy-consistent pseudopotentials and correlation consistent basis sets for the 5d elements Hf–Pt. *J. Chem. Phys.* **130,** 164108 (2009).
139. Kim, J., Ihee, H. & Lee, Y. S. Spin-orbit density functional and *ab initio* study of HgX_n (X=F, Cl, Br, and I: n=1, 2, and 4). *J. Chem. Phys.* **133,** 144309 (2010).
140. Häussermann, U., Dolg, M., Stoll, H., Preuss, H., Schwerdtfeger, P. & Pitzer, R. Accuracy of energy-adjusted quasirelativistic *ab initio* pseudopotentials. *Mol. Phys.* **78,** 1211–1224 (1993).

Abbreviations and acronyms

AE	all-electron
AO	atomic orbital
ANO-RCC	all natural orbitals-relativistic and (semi-)core correlation
B	Breit
BLAS	Basic Linear Algebra Subprograms
BSE	basis set extrapolation
BSS	Barysz-Sadlej-Snijders
BSSE	basis set superposition error
C	Coulomb
CAT	contracted angular transformed
CBS	complete basis set
(MR)CC	(multi-reference) coupled cluster
CCSD	coupled cluster singles and doubles
CCSD(T)	coupled cluster singles, doubles and perturbative triples
CCSDT	coupled cluster singles, doubles and triples
CCSDTQ	coupled cluster singles, doubles, triples and quadruples
(MR)CI	(multi-reference) configuration interaction
CP	counterpoise
CPP	core polarization potential
D	Dirac
DHF	Dirac-Hartree-Fock
DKH2	second-order Douglas-Kroll-Hess
DKH3	third-order Douglas-Kroll-Hess
ECP	effective core potential
ERI	electron repulsion integral
FC	frozen core
GCC	GNU Compiler Collection
gsl	GNU Scientific Library
(R)HF	(restricted) Hartree-Fock
IP	ionization potential
LAPACK	Linear Algebra PACKage
MCSCF	multi-configuration self-consistent field
MO	molecular orbital
MP	model potential
PCG	primitive Cartesian Gaussian
PP	pseudopotential
QOL	Quantum Objects Library
RAM	random access memory
SCF	self-consistent field
SO	spin-orbit
VO	valence-only

Danksagung

Vor ungefähr sechs Jahren kam ich zum Masterstudiengang und zur Promotion nach Köln und konnte mich nicht nur akademisch sondern vor allem auch persönlich weiterentwickeln. Dies ist natürlich euch geschuldet und ich möchte mich bei euch bedanken.

Danke, Herr Dolg, für die hervorragende Betreuung, das Vertrauen und die stets aufbauenden Worte. Außerdem möchte ich mich für die Erfahrungen in der gemeinsamen Lehre bedanken. Für mich waren Sie ein ausgezeichneter Chef und Doktorvater.

Zu einem tollen Doktorvater gehört natürlich noch ein cooler Doktoronkel. Dieser warst du, Michael Hanrath, und dafür bin ich dir sehr dankbar. Danke für deine ansteckende Begeisterung für alles rund um den Computer, die mich während meiner Doktorarbeit zusätzlich motiviert hat.

Herr Wickleder, Ihnen möchte ich für die hilfreichen Anregungen danken, dir mir eine neue Perspektive auf mein Projekt eröffnet haben, nämlich die des experimentellen Chemikers.

Danke, Xiaoyan Cao-Dolg, für die anregenden Kommentaren während des Gruppenseminars, der leckeren Hotpots während der Weihnachtsfeiern und der immer herzlichen und freundlichen Art.

Nach den TC-Familienoberhäuptern, möchte ich nun meinen Kollegen, Mitstreitern und Freunden Nils, Niko, Ilyas und Simon für die schönen Jahre danken. Seien es Hackathons, Konferenzen, (Austausch-)Seminare, Kaffeepausen, Betriebsfeiern oder private Treffen - das gemeinsame Lernen, Lachen, Diskutieren, sich Austauschen oder Feiern waren eine unglaubliche Bereicherung und Unterstützung. Ich bin froh, dass wir zur selben Zeit am selben Ort aufeinander getroffen sind.

Außerdem danke ich euch, Oli, Anna, Max, Dana, Jan, Daniel, Sarah und Norah, dass ihr als Ehemaligen mit mir in der TC tätig wart. Vor allem dir, Oli, bin ich dankbar, dass du mich in meinen ersten Praktika betreut und in die TC intrigiert hast. Anna, dir will ich ganz besonders danken, da du mir mit deinem Wissen rund um PPs und CPPs sehr geholfen und stets mein Projekt mit Interesse verfolgt hast.

Der letzte Dank in Bezug auf die TC gebührt euch, Frau Börsch-Pulm, Martin und Ansgar. An euch konnte man sich jederzeit mit nicht-wissenschaftlichen Problem (Internetauftritt, Anträge und kaputte Netzteile) wenden.

Ich war stets gerne ein Teil des Instituts für Theoretische Chemie und freue mich, in Zukunft an besonderen Veranstaltungen wie beispielweise der Weihnachtsfeier und dem Wandertag als Ehemaliger teilzunehmen.

Danke, Flore, Yannic, Dominik, Bernhard, Christian, RochelsL, Thom Thom, Karin, Svenja, Esra, Lukas, Mattes und Tim, dass ihr mich so herzlich in euren Reihen aufgenommen habt. Dank euch habe ich mich seit meinem ersten Tag in Köln aufgehoben gefühlt. Ohne euch wär das Studium wesentlich einsamer gewesen.

Danke, Annika Bande, für den ersten praktischen Einblick in die theoretische Chemie und für die Ratschläge auch weit nach der Bachelorarbeit.

Danke, Ralf Kellner, für die Vorbildfunktion, die du eingenommen hast und mir damit den universitären Weg aufgezeigt hast.

Danke, Daniel R., Daniel F., Patrick und Fabian, für die jahrelange Freundschaft, die trotz unterschiedlichsten persönlichen Entwicklungen und räumlichen Distanzen immer noch besteht.

Danke, Mama, Papa und Colin, für die Unterstützung und den Rückhalt über all die Jahre.

Danke, Mihaela, meiner besten Freundin, Mitstreiterin und Partnerin, für die Unterstützung, den Zuspruch und das gemeinsame Bewältigen von Tiefen während des Studiums und der Promotionszeit. Обичам те.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen:

_____

(Ort, Datum)

_____ _____

(Name) (Unterschrift)