# Implementation of the CAHF/CASCI Method and Application on Actinide-DOTA Complexes

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrads

der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln

vorgelegt von

SIMON LEKAT

aus Köln

2025

Berichterstatter Prof. Dr. Michael Dolg Prof. Dr. Mathias Wickleder

Tag der mündlichen Prüfung: 24.04.2025

Für meine Eltern

# Acknowledgements

The completion of this thesis would not have been possible without the support, guidance, and encouragement of many individuals and groups. I would like to express my deepest gratitude to the following people:

- Prof. Dr. Michael Dolg for the supervision of this thesis and the opportunity to become part of his research group. He answered all questions patiently, gave guidance when things did not go as planned and had always my best interest in mind. His supervision was exceptional, both scientifically and administratively.
- Prof. Dr. Mathias Wickleder for taking on the role as second examiner and the mentoring during my PhD.
- Prof. Dr. Michael Hanrath for his support and guidance throughout the years, both in quantum chemistry and C++. His open-door policy was greatly appreciated. He was able to explain concepts in simple terms, and often provided novel visualizations.
- Nikolas Nowak for carefully proof-reading this thesis.
- My current and former colleagues Nikolas Nowak, Daniel Weßling, Ilyas Türkmen, Nils Herrmann and Sascha Bubeck for a lighthearted and relaxed atmosphere.
- The group members of the former Institute for Theoretical Chemistry Ansgar Tappenhölter, Martin Böhler and Xiaoyan Cao-Dolg for the memorable time, christmas and summer parties as well as the hiking excursions.
- My friends for taking my mind of work, especially the group nicknamed *Magnesen* for the good time spent together, the scout group *Fenris* for their long-standing friendships and the memorable camps and evenings we shared, and my former school friends for the shared holidays and social game evenings.
- Lena Schmitz and Hannah Zettner for their support in difficult times and a very good friendship.
- My former chemistry teacher Mrs. Kornelia Knauf without whom I would not have perused the study of chemistry in the first place.
- My family Maike, Andrea and Peter, for for their unwavering love, support, and encouragement throughout my journey.

# Abstract

This thesis presents the implementation of the Configuration-Averaged Hartree-Fock (CAHF) method combined with Complete Active Space Configuration Interaction (CASCI) within the Quantum Object Library (QOL), and their application to Ac-DOTA complexes (DOTA<sup>4–</sup> = 1,4,7,10-Tetraazacyclododecane-N,N',N",N"'-tetraacetate). The CAHF/CASCI approach offers a computationally efficient route for accurately describing near-degenerate f-orbitals and the resulting states in lanthanide and actinide complexes.

The QOL framework was extended to enable medium- and large-scale computations through a more memory-efficient integral transformation procedure. Its capabilities were further improved by introducing a unified input system, enhancing existing methods, and adding new features such as a general overlap calculation between any wave functions, a robust CASCI implementation, and the incorporation of pseudopotential (PP) spin-orbit integrals.

The second part applies the CAHF/CASCI method to investigate An-DOTA complexes (for An = Pa, U, Np), focusing on ground state character, conformer stability, and ligand field splitting. Using small-core pseudopotentials (SPPs) and high-quality basis sets, a systematic trend favoring the square antiprismatic (SAP) conformer was observed, consistent with experimental findings. This study highlights the utility of the CAHF/CASCI method for exploring the electronic structure of f-element complexes and provides insights into the chemistry of actinide-DOTA systems.

# Contents

## Page

1	Intr	Introduction 1						
<b>2</b>	The	heory						
	2.1	Fundamentals	3					
	2.2	Hartree-Fock Theory	5					
		2.2.1 Variational principle	5					
		2.2.2 Density Matrices	6					
		2.2.3 Slater-Condon Rules	7					
		2.2.4 Fock Equations	9					
		2.2.5 The Roothaan-Hall Equations	10					
		2.2.6 Closed, Open and Averaged Shell Systems	12					
	2.3	Electron Correlation	18					
		2.3.1 Configuration Interaction	19					
		2.3.2 Spin Eigenfunctions	21					
		2.3.3 CI in an Active Space	23					
2.4 Relativistic Effects		Relativistic Effects	25					
		2.4.1 Dirac Equation	26					
		2.4.2 Two Component Methods	27					
		2.4.3 Relativistic Pseudopotentials	28					
3	Imp	plementation	33					
	3.1	The Quantum Objects Library	33					
3.2       Implementation Goals		Implementation Goals	34					
		Input Format	35					
		Improving the CAHF	37					
	3.5 FCI, CI and CASCI		38					
		3.5.1 Generating the CSF Basis	39					
		3.5.2 Scalar Product	42					
		3.5.3 Matrix Representation	43					
	3.6 Integral Transformation							
		3.6.1 Dealing with a Large Basis	45					

#### Contents

		3.6.2	Freezing and Deleting Orbitals	45						
		3.6.3	Reducing Memory Requirements	46						
	3.7	CASC	I with Spin-Orbit Contribution	52						
		3.7.1	Storing One-particle Spin-Orbit Integrals	52						
		3.7.2	Integral Evaluation	53						
	3.8	Wave 1	Function Analysis	54						
		3.8.1	Population Analysis	54						
		3.8.2	Overlap	55						
	3.9	Perform	ming Calculations	57						
		3.9.1	CAHF Comparisons	57						
		3.9.2	CAHF/CASCI-SO Comparisons	58						
		3.9.3	The Orthogonal Array	58						
4	App	olicatio	on: Actinide-DOTA Complexes	61						
	4.1	Introdu	$\operatorname{uction}$	61						
	4.2	Compu	itational Details	64						
	4.3	Results	s and Discussion	65						
	4.4	Summa	ary	69						
<b>5</b>	Con	clusio	n	71						
	5.1	Summa	ary	71						
	5.2	Outloc	ok	72						
A	Imp	lemen	tational Details	75						
	A.1	Derivir	ng the Integral Type	75						
в	Cale	alculation Data								
	B.1	Compa	arisons with MOLPRO	77						
	B.2	An-DC	OTA Complexes	78						
	B.3	Geome	etries	88						
Re	References 95									
Ei	Eidesstattliche Erklärung									

# Chapter 1 Introduction

Quantum chemistry seeks to provide a rigorous mathematical framework for the description and modeling of atoms and molecules. Among the various computational approaches in quantum chemistry, the Hartree-Fock (HF) method is one of the most extensively employed *ab initio* techniques. It offers the best single-determinant approximation of electronic states, and provides starting orbitals for more sophisticated post-Hartree-Fock methods such as configuration interaction (CI) and coupled cluster (CC).<sup>[3]</sup>

The (theoretical) study of metal complexes plays a vital role in advancing and understanding their applications across various fields, including medicine, material science, and catalysis. In such complexes, the central metal atom resides within a potential field created by surrounding ligands. The interaction between the metal and ligands leads to ligand field splitting effects.<sup>[4]</sup>

For lanthanides (Ln) and actinides (An), ligand fields exert a comparatively weaker influence on the radially compact and energetically similar 4f- and 5f-orbitals. Nevertheless, the f-shell drives key electronic phenomena in these complexes, such as absorption and emission spectra, (para)magnetism, binding affinity and more.<sup>[5,6]</sup> Due to the near-degeneracy of the f-orbitals, their accurate description requires specialized computational approaches. A common strategy involves using standard HF orbitals as a starting point, followed by a State-Averaged Complete Active Space Self-consistent Field (SA-CASSCF) calculation. This approach yields a single, averaged orbital set for all states, providing a consistent representation of the near-degenerate f-orbitals.

An alternative and computationally efficient route involves the use of the Configuration-Averaged Hartree-Fock (CAHF) method. In this approach, the orbital averaging occurs during the initial Hartree-Fock calculation, eliminating the need for a full CASSCF optimization. Instead, a simpler Complete Active Space Configuration Interaction (CASCI) calculation suffices. This method not only reduces computational cost but also yields results equivalent to SA-CASSCF.<sup>[7,8]</sup> The aim of this thesis is to implement the CAHF/CASCI method within the Institute's internal Quantum Objects Library (QOL)<sup>[9]</sup>, and extending the library's capabilities to medium- and large-scale computations. To demonstrate the utility of this implementation, the CAHF/CASCI method is applied to investigate the electronic structure of the An-DOTA system. The ligand 1,4,7,10-Tetraazacyclododecane-N,N',N",N"'-tetraacetate (DOTA<sup>4-</sup>) is a well-known, versatile chelator with extensive use in diagnostic and therapeutic medicine.<sup>[10,11]</sup> Recent studies of An-DOTA complexes employed density functional theory (DFT) and single-point CASSCF calculations<sup>[12,13]</sup>, providing a starting point for further exploration of the trends and patterns governing An-DOTA complex chemistry.

## Chapter 2

# Theory

In this chapter, an outline of the essential theoretical achievements in the field of quantum chemistry will be given, focussing on the topics most relevant for the discussions in subsequent chapters.

## 2.1 Fundamentals

In quantum chemistry, all aspects of a system are encoded in an object called the wave function  $\Psi(\vec{r}, t)$ . Truly understanding what this object represents is difficult to near impossible to grasp, as it contradicts nearly every intuition of everyday experiences and cannot be directly observed. Only single quantities can be extracted one after another, but not without changing the underlying wave function irreversibly. This results in some quantities which cannot be extracted (measured) simultaneously, such as position and momentum, as well as a plethora of other phenomena.

The wave function is the solution to an eigenvalue equation called the Schrödinger equation

$$\hat{H} |\Psi\rangle = i \frac{\partial}{\partial t} |\Psi\rangle . \qquad (2.1)$$

The Hamiltonian operator  $\hat{H}$  describes the interactions of all particles, depending on the level of theory considered. A large majority of quantum chemical problems deal with stationary states for which the Hamiltonian contains no time dependent terms, i.e.,

$$\hat{H} \equiv \hat{H}(\vec{x}) . \tag{2.2}$$

Then equation (2.1) may be separated by the product ansatz

$$|\Psi(\vec{r},t)\rangle = |\Phi(\vec{r})\rangle |\Theta(t)\rangle \tag{2.3}$$

3

into time and spatial equations, resulting in the time-independent Schrödinger equation

$$\hat{H}|\Psi\rangle = E|\Psi\rangle . \tag{2.4}$$

For the general non-relativistic case of a molecule, the Hamiltonian is given by five terms. The kinetic  $(\hat{T})$  and potential  $(\hat{V})$  contributions of the electrons (e) and nuclei (K) and their two particle interactions.<sup>[3]</sup>

$$\hat{H} = \hat{T}_e + \hat{T}_K + \hat{V}_{e,K} + \hat{V}_{e,e} + \hat{V}_{K,K}$$
(2.5)

Expanding the terms in atomic units gives

$$\hat{H} = -\frac{1}{2} \sum_{i}^{n} \Delta_{i} - \frac{1}{2} \sum_{I}^{N} \frac{\Delta_{I}}{M_{I}} - \sum_{i}^{n} \sum_{I}^{N} \frac{Z_{I}}{r_{iI}} + \sum_{i < j}^{n} \frac{1}{r_{ij}} + \sum_{I < J}^{N} \frac{Z_{I}Z_{J}}{r_{IJ}}, \qquad (2.6)$$

in which lower case letters (i, j) denote electron indices and upper case letters (I, J) denote nuclear indices. The total number of electrons and atoms are n and N respectively. Given that electrons are several orders of magnitude lighter than atomic nuclei, their motion can be approximated by considering the nuclei as stationary. This is the approach of the Born-Oppenheimer approximation, and results in the vanishing of  $\hat{T}_K$  and a constant  $\hat{V}_{K,K}$ .<sup>[3,14]</sup> With this, equation (2.6) reduces to the electronic Hamiltonian

$$\hat{H}_{elec} = -\frac{1}{2} \sum_{i}^{n} \Delta_{i} - \sum_{i}^{n} \sum_{I}^{N} \frac{Z_{I}}{r_{iI}} + \sum_{i < j}^{n} \frac{1}{r_{ij}} \,.$$
(2.7)

Likewise, the electronic Schrödinger equation reads

$$\hat{H}_{elec}|\Psi_{elec}\rangle = E_{elec}|\Psi_{elec}\rangle$$
 (2.8)

The term  $\hat{V}_{K,K}$  can be evaluated independently of the eigenvalue equation (2.8) and added afterwards to  $E_{elec}$  to yield the total energy  $E_{total}$ .

$$E_{total} = E_{elec} + E_{nucPot} = E_{elec} + \sum_{I < J}^{N} \frac{Z_I Z_J}{r_{IJ}}$$
 (2.9)

In this thesis, only the electronic Hamiltonian is of interest and moving forward the subscript *elec* will be dropped.

## 2.2 Hartree-Fock Theory

For systems with more than one electron, there exist no analytic solutions to the Schrödinger equation (2.4). This is due to the pairs of electrons correlated by the Coulomb repulsion term, i.e., the last term in equation (2.4). But there are still numeric approximations possible, with different level of accuracy depending on the wave function ansatz. In the Hartree-Fock (HF) theory, the wave function ansatz is a single Slater determinant (SD) with

$$|\Psi^{\rm SD}\rangle = \sqrt{n!} \,\hat{\mathcal{A}} \prod_{i}^{n} |\phi_{i}\rangle = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_{1}(1) & \phi_{1}(2) & \dots & \phi_{1}(n) \\ \phi_{2}(1) & \phi_{2}(2) & \dots & \phi_{2}(n) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n}(1) & \phi_{n}(2) & \dots & \phi_{n}(n) \end{vmatrix} , \qquad (2.10)$$

where  $\phi_i(i)$  are orthonormal one-particle functions of electron *i* called *spin orbitals*<sup>[3]</sup>

$$\phi_i(i) \equiv \phi_i(\vec{r}, \omega) = \varphi_i(\vec{r}) \cdot \sigma_i(\omega) . \qquad (2.11)$$

These are composed of a spatial orbital  $\varphi_i(\vec{r})$  and a spin function  $\sigma_i(\omega)$ , which is either  $\alpha(\omega)$  or  $\beta(\omega)$ , commonly associated with spin up ( $\uparrow$ ) and spin down ( $\downarrow$ ), respectively.<sup>[3]</sup> A spin orbital might also be abbreviated to the function index

$$|\phi_i\rangle \equiv |i\rangle \ , \tag{2.12}$$

as well as the Slater determinant to its diagonal

$$|\Psi^{\rm SD}\rangle = |\phi_1(1)\phi_2(2)\cdots\phi_n(n)\rangle \equiv |12\cdots n\rangle . \qquad (2.13)$$

#### 2.2.1 Variational principle

With a complete orthonormal eigenbasis  $\{|\Psi_i\rangle\}$  of the Hamiltonian  $\hat{H}$ , that is

$$\hat{H} |\Psi_i\rangle = E_i |\Psi_i\rangle , \qquad (2.14)$$

any approximate trial function  $|\tilde{\Psi}\rangle$  (such as a Slater determinant) may be written as linear combination of these eigenfunctions by

$$|\tilde{\Psi}\rangle = \sum_{i} c_i |\Psi_i\rangle ,$$
 (2.15)

with the complex valued expansion coefficients  $c_i$  and the implicit energy order

$$E_0 \le E_1 \le \dots \le E_i \le \dots \tag{2.16}$$

Expanding the energy expectation value of this trial function into the sums of the eigenfunction basis gives

$$\tilde{E} = \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle} = \frac{\sum_{ij} c_i^* c_j \langle \Psi_i | \hat{H} | \Psi_j \rangle}{\sum_{ij} c_i^* c_j \langle \Psi_i | \Psi_j \rangle} = \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2} .$$
(2.17)

Since  $E_i \ge E_0$  for all *i* by definition (2.16), it is

$$\tilde{E} = \frac{\sum_{i} |c_{i}|^{2} E_{i}}{\sum_{i} |c_{i}|^{2}} \ge \frac{\sum_{i} |c_{i}|^{2} E_{0}}{\sum_{i} |c_{i}|^{2}} = E_{0} .$$
(2.18)

Therefore, the energy expectation value of any trial function  $|\tilde{\Psi}\rangle$  is bounded from below by the exact energy  $E_0$ ,

$$\tilde{E} = \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle} \ge \frac{\langle \Psi_0 | \hat{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} = E_0 .$$
(2.19)

#### 2.2.2 Density Matrices

The absolute value squared of the wave function  $|\Psi|^2$  is postulated by Born to be interpreted as a probability density.<sup>[15]</sup> With an N-particle wave function, this is

$$|\Psi(\vec{r}_1, ..., \vec{r}_N, \omega_1, ..., \omega_N)|^2 \equiv |\Psi(1, ..., N)|^2 , \qquad (2.20)$$

while reducing the spatial and spin variables to the particle indices for abbreviation. With a normalized wave function, i.e.,  $\langle \Psi | \Psi \rangle = 1$ , the expectation value of a general operator  $\hat{O}_n$  which interrelates *n*-particles with each other is

$$\langle \hat{O}_n \rangle = \langle \Psi | \, \hat{O}_n \, | \Psi \rangle = \int \Psi^*(1, ..., N) \hat{O}_n \Psi(1, ..., N) \, \mathrm{d} 1 \cdots \mathrm{d} N \,. \tag{2.21}$$

Due to the indistinguishability of the electrons, this can also be expressed as the operator  $\hat{o}(1, ..., n)$  acting only on the first n particles and scaling this result by the number of permutations,

$$\langle \hat{O}_n \rangle = \binom{N}{n} \int \Psi^*(1, \dots, N) \hat{o}(1, \dots, n) \Psi(1, \dots, N) \,\mathrm{d} 1 \cdots \mathrm{d} N \;. \tag{2.22}$$

The integrals for the indices above n in equation (2.22) do not change, regardless of the actual operator  $\hat{O}_n$ . By relabelling the interrelated n particles in  $\Psi^*$ , the integrals over the remaining indices can be formed while the operator still only affects  $\Psi$ .

$$\langle \hat{O}_n \rangle = \int \left\{ \hat{o}(1, ..., n) \left( \binom{N}{n} \int \Psi^*(1', ..., n', n+1, ..., N) \Psi(1, ..., N) d(n+1) \cdots dN \right) \right\}$$
  
$$\delta_{1', 1} \cdots \delta_{n', n} d1 \cdots dn$$
  
(2.23)

The relabled indices are made identical after the operator took effect. The quantity in the round brackets of equation (2.23) are the elements of the general definition of the n-particle-reduced density matrix<sup>[16]</sup>

$$\left(\mathbf{\Gamma}^{(n)}\right)_{1,\dots,n}^{1',\dots,n'} = \binom{N}{n} \int \Psi^*(1',\dots,n',n+1,\dots,N) \Psi(1,\dots,N) \,\mathrm{d}\,(n+1)\cdots\mathrm{d}N \,, \qquad (2.24)$$

reducing the general expectation value of an n-particle operator to

$$\langle \hat{O}_n \rangle = \int \left\{ \hat{o}(1, \dots, n) \left( \mathbf{\Gamma}^{(n)} \right)_{1,\dots,n}^{1',\dots,n'} \right\} \delta_{1',1} \cdots \delta_{n',n} \, \mathrm{d}1 \cdots \mathrm{d}n \;. \tag{2.25}$$

Consequently, the reduced density matrices  $\Gamma^{(n)}$  are Hermitian and antisymmetric with respect to particle exchange in the two index sets. The diagonal elements of the matrix (2.24) are the *n*-particle reduced densities

$$\rho(1,...,n) = \left(\mathbf{\Gamma}^{(n)}\right)_{1,...,n}^{1,...,n} .$$
(2.26)

With the Hamilton operator containing two-particle operators at most (cf. equation (2.7)), the one- and two-particle reduced density matrices  $\Gamma^{(1)}$  and  $\Gamma^{(2)}$  are sufficient to describe the expectation value of  $\langle \hat{H} \rangle$ , as well as the properties of atoms and molecules.<sup>[17]</sup>

#### 2.2.3 Slater-Condon Rules

The Slater-Condon rules provide a systematic framework for evaluating matrix elements of one- and two-electron operators between Slater determinants. Table 2.1 summarizes the rules for matrix elements of the identity operator, general one-electron operators, and general two-electron operators. The matrix element of a one-electron operator between two determinants is non-zero only if the determinants differ by at most one spin-orbital, while for a two-electron operator, the determinants can differ by up to two spin-orbitals.

	$\hat{O} = 1$	$\hat{O} = \hat{O}_1 = \sum_i \hat{o}(i)$	$\hat{O} = \hat{O}_2 = \sum_{i < j} \hat{o}(i, j)$
$\left<\Psi\right \hat{O}\left \Psi\right>$	1	$\sum_{i}\left\langle i ight  \hat{o}(1)\left j ight angle$	$\sum_{i < j} \left( \langle ij   \hat{o}(1,2)   ij \rangle - \langle ij   \hat{o}(1,2)   ji \rangle \right)$
$\left< \Psi^A_I \right  \hat{O} \left  \Psi \right>$	0	$\langle A    \hat{o}(1)    I \rangle$	$\sum_{i} \left( \left\langle iA \right  \hat{o}(1,2) \left  iI \right\rangle - \left\langle iA \right  \hat{o}(1,2) \left  Ii \right\rangle \right)$
$\left< \Psi^{AB}_{IJ} \right  \hat{O} \left  \Psi \right>$	0	0	$\langle AB   \hat{o}(1,2)   IJ \rangle - \langle AB   \hat{o}(1,2)   JI \rangle$
$\left< \Psi^{AB\ldots}_{IJ\ldots} \right  \hat{O} \left  \Psi \right>$	0	0	0

Table 2.1: Slater-Condon rules<sup>[18,19]</sup> for the matrix elements between determinants for up to two-particle operators.

The Hamiltonian (2.7) can be separated into a one-particle and a two-particle part, with

$$\hat{H} = \underbrace{\sum_{i=1}^{n} \hat{h}_{i}}_{\hat{O}_{1}} + \underbrace{\sum_{i$$

The one-particle Hamiltonian  $\hat{h}_i$  is defined as

$$\hat{h}_i = -\frac{1}{2}\Delta_i - \sum_{I}^{N} \frac{Z_I}{r_{iI}} , \qquad (2.28)$$

and the two-particle operator  $\hat{g}_{ij}$  takes the form

$$\hat{g}_{ij} = \frac{1}{r_{ij}}$$
 (2.29)

Using the Slater-Condon rules, the Hartree-Fock energy can be expressed as the sum

$$E_{\rm HF} = \sum_{i}^{n} \langle i | \hat{h}(1) | i \rangle + \sum_{i < j}^{n} \left( \langle i j | \hat{g}(1,2) | i j \rangle - \langle i j | \hat{g}(1,2) | j i \rangle \right) .$$
(2.30)

The indistinguishability of electrons states that the operators for different particles must be identical, and by convention, particle indices 1 and 2 are used in the expressions.

#### 2.2.4 Fock Equations

With the variational principle (2.19) in mind, the goal is now to minimize the energy (2.30) while keeping the full set of spin orbitals  $\{|\phi_i\rangle\}$  orthonormal, that is

$$E_{\rm HF} = \min_{\{\phi_i\}} \frac{\langle \Psi^{\rm SD} | \hat{H} | \Psi^{\rm SD} \rangle}{\langle \Psi^{\rm SD} | \Psi^{\rm SD} \rangle} \qquad \text{with} \quad \langle \phi_i | \phi_j \rangle = \delta_{ij} .$$
 (2.31)

This is achieved by the method of Lagrangian multipliers. Doing so lead to the Hartree-Fock integro-differential equation for the HF spin orbitals  $|\phi_i\rangle$ 

$$\epsilon_i \phi_i(1) = \hat{h}(1)\phi_i(1) + \sum_{i,j} \left[ \int \phi_j^*(2) \frac{1}{r_{12}} \phi_j(2) \, \mathrm{d}\tau_2 \right] \phi_i(1) - \sum_{i,j} \left[ \int \phi_j^*(2) \frac{1}{r_{12}} \phi_i(2) \, \mathrm{d}\tau_2 \right] \phi_j(1) \,.$$
(2.32)

The right side of this equation can be written as the effect of a single particle operator operating on the spin orbital  $\phi_i$ , by introducing two new operators  $\hat{J}_j$  and  $\hat{K}_j$ 

$$\hat{J}_{j}\phi_{i}(1) = \int \phi_{j}^{*}(2)\frac{1}{r_{12}}\phi_{j}(2)\phi_{i}(1) \,\mathrm{d}\tau_{2} , \qquad (2.33)$$

$$\hat{K}_j \phi_i(1) = \int \phi_j^*(2) \frac{1}{r_{12}} \phi_j(1) \phi_i(2) \,\mathrm{d}\tau_2 \,.$$
(2.34)

Combining these operators with the one electron Hamiltonian  $\hat{h}$  gives the Fock operator

$$\hat{F}(1) = \hat{h}(1) + \sum_{j} \left[ \hat{J}_{j}(1) - \hat{K}_{j}(1) \right] , \qquad (2.35)$$

and equation (2.32) results in the canonical Fock equations

$$\hat{F}\phi_i = \epsilon_i \phi_i . \tag{2.36}$$

The operator  $\hat{J}_j$  is thereby called Coulomb operator, whereas  $\hat{K}_j$  is called exchange operator. The naming originates from the attempt to describe the effect of the operators.

First, consider  $\hat{J}_j$ ; usually, the Coulomb interaction between two electrons is captured by the relation  $\frac{1}{r_{12}}$  which is a two-particle potential. Instead, equation (2.33) shows that electron 1 in spin orbital  $\phi_i$  is affected by the one-particle potential of the electron 2 in spin orbital  $\phi_j$ . This potential is the average of the  $\frac{1}{r_{12}}$  interaction over all space and spin coordinates of electron 2, with the weight of the probability density

$$\phi_j^*(2)\phi_j(2)\,\mathrm{d}\tau_2 = |\phi_j(2)|^2\,\mathrm{d}\tau_2\;,\tag{2.37}$$

that electron 2 is found in the volume element  $d\tau_2$ .  $\hat{J}_j$  therefore represents the average *local potential* resulting from an electron in  $\phi_j$ .<sup>[20,21]</sup>

Contrary, the exchange operator  $\hat{K}_j$  does not have a simple classical interpretation. The electrons between the orbital  $\phi_j$  and the one it is operating on  $(\phi_i)$  are exchanged. The effect of  $\hat{K}_j$  depends on the spin orbital it is operating on and its value throughout all space, not just at a finite volume element.  $\hat{K}_j$  therefore represents the average *non-local potential* resulting from an electron in  $\phi_i$ .<sup>[20,21]</sup>

This exchange term is a direct result of the anti-symmetry of the Slater determinant (and thus of the wave function), which also explains the negative sign. With the permutation operator  $\hat{P}_{12}$  the Fock operator (2.35) can be brought into a form, which emphasizes this connection

$$\hat{F}(1) = \hat{h}(1) + \sum_{j} \int \phi_{j}^{*}(2) \frac{1}{r_{12}} (1 - \hat{P}_{12}) \phi_{j}(2) \,\mathrm{d}\tau_{2} \,.$$
(2.38)

Summing over all occupied spin orbitals for the coulomb and exchange operators (2.33) and (2.34) gives the total Coulomb and exchange operators  $\hat{J}$  and  $\hat{K}$  with

$$\hat{J} = \sum_{j} \hat{J}_{j} \qquad \hat{K} = \sum_{j} \hat{K}_{j} . \qquad (2.39)$$

These can also be expressed using one-electron density matrices (2.24) to give<sup>[22]</sup>

$$\hat{J}\phi(1) = \int \frac{1}{r_{12}} \left(\mathbf{\Gamma}^{(1)}\right)_2^2 \phi(1) \,\mathrm{d}\tau_2 \,, \qquad (2.40)$$

$$\hat{K}\phi(1) = \int \frac{1}{r_{12}} \left( \mathbf{\Gamma}^{(1)} \right)_2^1 \phi(2) \,\mathrm{d}\tau_2 \,. \tag{2.41}$$

And finally, the total Fock operator on an arbitrary function may be expressed as

$$\hat{F} = \hat{h} + \hat{J} - \hat{K}$$
 (2.42)

#### 2.2.5 The Roothaan-Hall Equations

Solving the minimization problem (2.31) is quite cumbersome while varying actual functions. For small, linear molecules this is possible as shown by Kobus and Lehtola.<sup>[23,24]</sup> These finite difference Hartree-Fock methods (FD HF) employ a 2D radial grid to tabulate the spin orbitals. Their program x2dhf is built upon the work of Laaksonen, Pyykkö, and Sundholm.<sup>[25]</sup> The resulting spin orbitals are almost exact, which is why it has been extensively used in the literature to assess the accuracy of other approaches.

In most other programs and methods, a basis is introduced. The spin orbitals  $|\phi\rangle$  are expressed as a linear combination of basis functions  $|\chi_k\rangle$ ,<sup>[3,26]</sup>

$$|\phi_i\rangle = \sum_k c_{ik} |\chi_k\rangle \quad , \tag{2.43}$$

with the real valued coefficients  $c_{ik}$ . If the set of basis functions  $\{|\chi_k\rangle\}$  would span the complete Hilbert space  $\mathcal{H}$ , this would be exact.<sup>a</sup> But for practical purposes, the number of basis functions K is kept finite, resulting in an approximation with arbitrary close accuracy if the basis is increased. This way, the abstract spin functions are translated to coefficient vectors  $\vec{c_i}$  of the chosen basis set. The  $|\chi_k\rangle$  are called the atomic orbitals (AO) which are combined to give the molecular orbitals (MO)  $|\phi_i\rangle$ . By substituting (2.43) into (2.36) and projecting onto a general basis function  $\chi_l$ , the Fock equations are rewritten as

$$\sum_{k} c_{ik} \langle \chi_l | \hat{F} | \chi_k \rangle = \epsilon_i \sum_{k} c_{ik} \langle \chi_l | \chi_k \rangle .$$
(2.44)

The elements of the overlap matrix  $\mathbf{S}$  and Fock matrix  $\mathbf{F}$  will be introduced as

$$[\mathbf{S}]_{lm} = \langle \chi_l | \chi_m \rangle \quad \text{and} \quad [\mathbf{F}]_{lm} = \langle \chi_l | \hat{F} | \chi_m \rangle . \quad (2.45)$$

With this, rewriting equation (2.44) in matrix form, gives the Roothaan-Hall equations.<sup>[27,28]</sup>

$$\mathbf{F}\vec{c_i} = \epsilon_i \mathbf{S}\vec{c_i} \tag{2.46}$$

Combining the eigenvectors  $\vec{c}_i$  and corresponding eigenvalues  $\epsilon_i$  into matrices as well, gives the diagonal matrix  $\boldsymbol{\epsilon}$  and the coefficient matrix  $\mathbf{C}$  with

$$\mathbf{C} = (\vec{c}_1 | \vec{c}_2 | \dots | \vec{c}_K) \quad . \tag{2.47}$$

For the sake of consistency in this thesis, a matrix which combines only selected orbitals is written as **T** instead. The one-electron density matrix  $\rho$  can be expressed as<sup>[20]</sup>

$$\boldsymbol{\rho} = \mathbf{T}\mathbf{T}^{\dagger} = \sum_{k}^{\text{occ}} \vec{c}_{k} \vec{c}_{k}^{\dagger} . \qquad (2.48)$$

The density matrix serves as a projection operator onto the space characterized by the orbitals in the associated  $\mathbf{T}$ . In the case of orthonormal basis functions, this necessitates the density matrix to be idempotent. For non-orthonormal basis functions, the idempotency condition is relaxed to<sup>[20]</sup>

$$\boldsymbol{\rho} \mathbf{S} \boldsymbol{\rho} = \boldsymbol{\rho} \; . \tag{2.49}$$

The HF energy equation (2.30) can be expanded in terms of the atomic orbitals of the

a Typically, Gaussian functions  $\exp(\varepsilon x^2)$  with the exponent  $\varepsilon$  less than 0 are used to model chemical systems. To reach the complete basis set limit, formally also the functions with  $\varepsilon$  greater 0 are needed.

basis using this coefficient matrix  $\mathbf{T}$  as

$$E_{\rm HF} = \sum_{k}^{\rm MO} \sum_{p,q}^{\rm AO} T_{pk}^* \langle \chi_p | \hat{h} | \chi_q \rangle T_{qk} + \frac{1}{2} \sum_{k,l}^{\rm MO} \sum_{p,q,r,s}^{\rm AO} T_{pk}^* T_{ql}^* \langle \chi_p \chi_q | | \chi_r \chi_s \rangle T_{rk} T_{sl} , \qquad (2.50)$$

with the two-electron integral on the right as shorthand for

$$\langle \chi_p \chi_q || \chi_r \chi_s \rangle = \left\langle \chi_p \chi_r |\frac{1}{r_{12}} |\chi_q \chi_s \right\rangle - \left\langle \chi_p \chi_r |\frac{1}{r_{12}} |\chi_s \chi_q \right\rangle .$$
(2.51)

By defining the elements of the one-electron Hamilton matrix  ${f h}$  as

$$[\mathbf{h}]_{lm} = \langle \chi_l | \hat{h} | \chi_m \rangle , \qquad (2.52)$$

and the trace of a matrix **M** as Tr **M**, equation (2.50) can be written as<sup>[20]</sup>

$$E_{\rm HF} = \operatorname{Tr} \boldsymbol{\rho} \mathbf{h} + \frac{1}{2} \operatorname{Tr} \boldsymbol{\rho} \mathbf{G}(\boldsymbol{\rho}) . \qquad (2.53)$$

The electron interaction matrix  $\mathbf{G}(\boldsymbol{\rho})$  is a function of  $\boldsymbol{\rho}$  and given by

$$[\mathbf{G}(\boldsymbol{\rho})]_{pq} = \sum_{r,s} [\boldsymbol{\rho}]_{rs} \langle \chi_p \chi_q || \chi_r \chi_s \rangle . \qquad (2.54)$$

Similar to the definition (2.42), the Fock matrix  $\mathbf{F}$  can be expressed as the sum of the one-electron Hamilton matrix and the electron interaction matrix.<sup>[20]</sup>

$$\mathbf{F} = \mathbf{h} + \mathbf{G} \tag{2.55}$$

#### 2.2.6 Closed, Open and Averaged Shell Systems

Until now, every equation was given for general spin orbitals, meaning there were no restriction on the spatial part. In general this leads to the unrestricted Hartree-Fock (UHF) method, and gives different spatial description of  $\phi_i^{\alpha}$  and  $\phi_i^{\beta}$  orbitals, that is

$$|\phi_i^{UHF}(\vec{r},\sigma)\rangle = \begin{cases} |\psi_i^{(\alpha)}(\vec{r})\rangle |\sigma(\alpha)\rangle \\ |\psi_i^{(\beta)}(\vec{r})\rangle |\sigma(\beta)\rangle \end{cases}$$
(2.56)

This results in two sets of orthonormal spatial orbitals  $\{|\psi_i^{(\alpha)}\rangle\}$  and  $\{|\psi_i^{(\beta)}\rangle\}$ , which overlap by an unspecified amount.

$$\langle \psi_i^{(\alpha)} | \psi_j^{(\alpha)} \rangle = \langle \psi_i^{(\beta)} | \psi_j^{(\beta)} \rangle = \sigma_{ij} \quad \text{and} \quad \langle \psi_i^{(\alpha)} | \psi_j^{(\beta)} \rangle = [\mathbf{S}^{(\alpha\beta)}]_{ij} \quad (2.57)$$

While this approach gives many degrees of freedom to minimize the energy, it has a major drawback: The eigenfunctions do not allow building pure eigenfunctions of the total spin operator  $\hat{S}^2$ . This is called *spin contamination* and is because the spatial many-particle functions are intrinsically linked to its spin function and errors cannot be fully corrected by linear combinations of different spins. The UHF wave function is in general not suitable to represent real spectroscopic states and results in problems if used in post-Hartree-Fock methods (see section 2.3). The solution is to use the same spatial orbital set for both  $\alpha$ and  $\beta$  spins, called restricted Hartree-Fock (RHF):

$$|\phi_i^{RHF}(\vec{r},\sigma)\rangle = |\psi_i(\vec{r})\rangle \begin{cases} |\sigma(\alpha)\rangle \\ |\sigma(\beta)\rangle \end{cases} .$$
(2.58)

Formally, this restriction is to demand that the overlap  $\mathbf{S}^{(\alpha\beta)}$  is the unit matrix.

#### **Closed Shell Systems**

With an even number of electrons and full spin pairing, most molecules do not have unpaired electrons. For these cases, the spin can be fully removed and the indices loop over the spatial orbitals  $\psi_i$ . This is then reflected in the occupation number of 2 of each (occupied) orbital. The energy equation (2.53) is changed to

$$E_{\rm RHF} = 2 \operatorname{Tr} \boldsymbol{\rho} \mathbf{h} + \operatorname{Tr} \boldsymbol{\rho} \mathbf{G}_{\rm RHF} , \qquad (2.59)$$

where the electron interaction matrix  $\mathbf{G}_{\mathrm{RHF}}$  is defined as

$$\mathbf{G}_{\mathrm{RHF}} = 2\left(\mathbf{J} - \frac{1}{2}\mathbf{K}\right) \,. \tag{2.60}$$

The exchange operator only correlates electrons with the same spin, which is exactly half of the occupation of each orbital.

#### **Open Shell Systems**

Describing systems with unpaired electrons with RHF is intrinsically more difficult. The approach is to split the orbitals into different shells, depending on their occupation. The large majority of orbitals will be doubly occupied, forming the closed shell, and any unoccupied orbitals give the virtual shell. Finally, the open shell (sometimes also called active space) entails everything in between – for the typical restricted open shell Hartree-Fock (ROHF) method this is fixed at singly occupation, but other methods expanded this to also include fractional occupation.<sup>[21]</sup> From now on,  $m_K$  will be the number of orbitals

and  $n_K$  the number of electrons in the  $K^{\text{th}}$  shell. Together they give the occupation number  $\nu_K$ ,

$$\nu_K = \frac{n_K}{m_K} \,. \tag{2.61}$$

Chapter 2 Theory

The coefficient matrix  $\mathbf{T}_{K}$  is similarly defined to (2.47) by combining the respective orbitals

$$\mathbf{T}_{K} = \underbrace{\left(\vec{c}_{k} | \vec{c}_{l} | \dots | \vec{c}_{\nu}\right)}_{m_{K} - times}, \qquad (2.62)$$

and the respective density matrix are formed analogous to (2.48) by

$$\mathbf{R}_K = \mathbf{T}_K \mathbf{T}_K^{\dagger} \,. \tag{2.63}$$

Again, these density matrices are projectors for the respective subspace from the space spanned by all spin orbitals. They obey a similar idempotency rule as shown in (2.49) with

$$\mathbf{R}_i \mathbf{S} \mathbf{R}_j = \delta_{ij} \mathbf{R}_i \ . \tag{2.64}$$

Additionally, the sum of all density matrices gives

$$\sum_{i} \mathbf{R}_{i} = \mathbf{S}^{-1} , \qquad (2.65)$$

and with an orthonormal basis  $(\mathbf{S} = \mathbf{I})$  this is called the *resolution to the identity*. With these matrices, the Hartree-Fock energy for two shells takes the form

$$E = \nu_1 \operatorname{Tr} \mathbf{R}_1 \left( \mathbf{h} + \frac{1}{2} \mathbf{G}_1 \right) + \nu_2 \operatorname{Tr} \mathbf{R}_2 \left( \mathbf{h} + \frac{1}{2} \mathbf{G}_2 \right) , \qquad (2.66)$$

with the electron interaction matrices defined by

$$\mathbf{G}_{1} = \nu_{1} \mathbf{G} (\mathbf{R}_{1}) + \nu_{2} \mathbf{G} (\mathbf{R}_{2}) \quad \text{(closed shell)} \\
 \mathbf{G}_{2} = \nu_{1} \mathbf{G} (\mathbf{R}_{1}) + \nu_{2} \mathbf{G}' (\mathbf{R}_{2}) \quad \text{(open shell)} 
 \right\}$$
(2.67)

with

$$\begin{array}{l}
\mathbf{G}(\mathbf{R}) = \mathbf{J}(\mathbf{R}) - \frac{1}{2}\mathbf{K}(\mathbf{R}) , \\
\mathbf{G}'(\mathbf{R}) = \mathbf{J}(\mathbf{R}) - \mathbf{K}(\mathbf{R}) .
\end{array}$$
(2.68)

This in turn results in two Fock matrices,

$$\mathbf{F}_1 = \mathbf{h} + \mathbf{G}_1 , \\ \mathbf{F}_2 = \mathbf{h} + \mathbf{G}_2 .$$
 (2.69)

14

Solving two eigenvalue problems with are coupled by (2.64) and (2.65) is not trivial. Fortunately, it is possible to formulate a combined equation by looking at the stationary conditions of both. The detailed derivatization will be skipped and can be found in ref. [20]. To give a motivation, the combined Fock-Matrix **F** may be expressed using (2.65) it in terms of the Fock and density matrices.

$$\mathbf{SS}^{-1}\mathbf{FS}^{-1}\mathbf{S} = \mathbf{S}(\mathbf{R}_1 + \ldots)(\mathbf{F}_1 + \ldots)(\mathbf{R}_1 + \ldots)\mathbf{S}$$
(2.70)

Manipulating and grouping of the terms give rise to the expression

$$\mathbf{F} = \mathbf{S}\tilde{\mathbf{F}}\mathbf{S} \tag{2.71}$$

$$\tilde{\mathbf{F}} = \sum_{K} \mathbf{R}_{K} \mathbf{d}_{K} \mathbf{R}_{K} + \sum_{K,L} \beta_{KL} \mathbf{R}_{K} \mathbf{F}_{(KL)} \mathbf{R}_{L} \qquad (\beta_{KL} = -\beta_{LK}) , \qquad (2.72)$$

with the intershell matrix  $\mathbf{F}_{(KL)}$  defined as

$$\mathbf{F}_{(KL)} = \nu_K \mathbf{F}_K - \nu_L \mathbf{F}_L , \qquad \mathbf{F}_K = \mathbf{h} + \mathbf{G}_K .$$
(2.73)

In theory, matrix  $\mathbf{d}_K$  could be an arbitrary Hermitian matrix, but in practise, it is useful to relate  $\mathbf{d}_K$  to  $\mathbf{F}_K$  by

$$\mathbf{d}_K = \mathbf{F}_K + \alpha_K \mathbf{S} , \qquad (2.74)$$

where  $\alpha_K$  is called the *level shift* of shell K, shifting the eigenvalues of this shell upwards by this amount. This energetic separation of shells allows for an easier orbital  $\leftrightarrow$  shell mapping and often improves the SCF convergence.<sup>[20,29,30]</sup>

The stationary conditions are expressed with the second sum of (2.72). At convergence these terms will vanish, which is why they can be multiplied by the so called *damp factors*  $\beta_{KL}$  without altering the final solution. These factors are a tool to control primarily the approach to different stationary points.<sup>[30]</sup>

#### Averaged Shell Systems

For a given electron configuration, such as carbon  $1s^22s^22p^2$ , multiple distinct spectroscopic states (e.g.,  ${}^{1}S$ ,  ${}^{3}P$  and  ${}^{1}D$ ) can arise, each associated with its own set of optimized orbitals. These orbital sets are generally not mutually orthogonal, which introduces complexity in calculations involving multiple sets, such as those for transition probabilities. Consequently, it is advantageous to identify a single set of near-optimal orbitals for which all relevant states are described equally well.<sup>[31]</sup> This is achieved by optimizing the averaged energy of these states, typically performed during a CASSCF calculation (see section 2.3.3).

The general energy expression (2.53) is intrinsically linked to the Hartree-Fock machinery and can be applied even when the original problem is inherently multi-determinantal in character.<sup>[20]</sup> Taking the average give energetic degenerate orbitals, which in turn allows the linear combination of determinants from a state-averaged CASSCF calculation to collaps to one determinant. With the effect, that each orbital is now fractionally occupied by the smeard out electrons. For the carbon example with  $p^2$ , the open shell would contain 2 electrons in 3 orbitals, resulting in a fractional occupation of  $\nu = \frac{2}{3}$  for each orbital. The equation for the averaged energy takes the form

$$E_{\rm av} = \nu_1 \operatorname{Tr} \mathbf{R}_1 \left( \mathbf{h} + \frac{1}{2} \mathbf{G}_1^{\rm av} \right) + \nu_2 \operatorname{Tr} \mathbf{R}_2 \left( \mathbf{h} + \frac{1}{2} \mathbf{G}_2^{\rm av} \right) \,. \tag{2.75}$$

The electron interaction matrices are similarly defined as in (2.67), but replacing the term for the open shell with

$$\begin{aligned}
 G_1^{av} &= \nu_1 \mathbf{G} \left( \mathbf{R}_1 \right) + \nu_2 \mathbf{G} \left( \mathbf{R}_2 \right) , \\
 G_2^{av} &= \nu_1 \mathbf{G} \left( \mathbf{R}_1 \right) + \nu_2' \mathbf{G} \left( \mathbf{R}_2 \right) ,
 \end{aligned}$$
(2.76)

where  $\nu'_K$  is defined as

$$\nu'_K = \frac{2(n_2 - 1)}{2m_2 - 1} . \tag{2.77}$$

This modified occupation number ensures that there is no contribution with only one electron. If the shell is fully occupied,  $\nu'_K$  is equal to  $\nu_K$ , giving the correct occupation of 2. Therefore, the split definition of the electron interaction matrices can be combined and generalized to

$$\mathbf{G}_{K}^{\mathrm{av}} = \nu_{K}^{\prime} \mathbf{G} \left( \mathbf{R}_{K} \right) + \sum_{L \neq K} \nu_{L} \mathbf{G} \left( \mathbf{R}_{L} \right) \ . \tag{2.78}$$

This allows to generalize the formula (2.75) to multiple open shells, where no electron transfer is permitted between shells.<sup>[32]</sup>

$$E_{\rm av} = \sum_{K} \nu_K \operatorname{Tr} \mathbf{R}_K \left( \mathbf{h} + \frac{1}{2} \mathbf{G}_K^{\rm av} \right)$$
(2.79)

The formation of the Fock matrix was already presented in a generalized form in (2.72), allowing the normal SCF procedure to be carried out. This is the so-called configuration averaged Hartree-Fock (CAHF) method.<sup>[31]</sup>

In the absence of additional constraints, the averaging process encompasses all states regardless of their spin multiplicity, yielding a wave function that is applicable to various spin states.

#### State Averaging of a Specific Spin

Restricting the average to only the states with a specific spin multiplicity  $M_S$  gives more specialized orbitals and also allows the CAHF method to perform ROHF calculations.<sup>[7]</sup> The approach is again restricted to one open shell, with the electron interaction matrices

$$\begin{array}{l}
\mathbf{G}_{1} = \nu_{1} \mathbf{G} \left( \mathbf{R}_{1} \right) + \nu_{2} \mathbf{G} \left( \mathbf{R}_{2} \right) , \\
\mathbf{G}_{2} = \nu_{1} \mathbf{G} \left( \mathbf{R}_{1} \right) + \nu_{2} \mathbf{G}^{M_{S}} \left( \mathbf{R}_{2} \right) .
\end{array}$$
(2.80)

The matrix **G** is the normal closed-shell matrix, whereas  $\mathbf{G}^{M_S}$  has two factors.<sup>[7]</sup>

$$\begin{array}{l}
\mathbf{G}\left(\mathbf{R}\right) &= \mathbf{J}\left(\mathbf{R}\right) - \frac{1}{2}\mathbf{K}\left(\mathbf{R}\right) \\
\mathbf{G}^{M_{S}}\left(\mathbf{R}\right) &= \lambda_{J}^{M_{S}}\mathbf{J}\left(\mathbf{R}\right) - \lambda_{K}^{M_{S}}\mathbf{K}\left(\mathbf{R}\right)
\end{array}$$
(2.81)

The number of electrons in the open shell is  $n_2$ , which consists of the number of  $\alpha$  and  $\beta$ spin electrons, denoted  $n_{\alpha}$  and  $n_{\beta}$ , respectively. With the common assumption  $n_{\alpha} \ge n_{\beta}$ , the spin multiplicity  $M_S$  defines  $n_{\alpha}$  with

$$n_{\alpha} = \frac{1}{2}(n_2 + M_S - 1) . \tag{2.82}$$

The two factors in (2.81) are then expressed with<sup>[7]</sup>

$$\lambda_J^{M_S} = \frac{m_2 n_2 (n_2 - 1) - 2n_\alpha n_\beta}{n_2^2 (m_2 - 1)} ,$$

$$\lambda_K^{M_S} = \frac{m_2 n_2 (n_2 - 1) - 2m_2 n_\alpha n_\beta}{n_2^2 (m_2 - 1)} .$$
(2.83)

As mentioned, the restriction to specific  $M_S$  values gives the flexibility to perform ROHF calculations with the same method. The trick is to choose the active space in such a way, that every orbital is singly  $\alpha$  occupied, resulting in  $\lambda_J^{M_s} = \lambda_K^{M_s} = 1$  and returning (2.68).

## 2.3 Electron Correlation

The HF methods presented in the preceding section 2.2 only approximate the solution of the Schrödinger equation by limiting the description to a single Slater determinant and treating two-particle interactions with averaged probability densities. Typically, the HF energy gives 99% of the exact energy, leaving just 1% of correlation energy unaccounted for. In many chemical problems, such as bonding and excitation, tiny energy differences play a huge role in the outcomes. It is therefore of utmost importance to recover as much of this correlation energy as possible within a reasonable timeframe. The difference between the HF energy  $E_{\rm HF}$  and the exact electronic energy  $E_{\rm exact}$  is defined as the *correlation*  $energy^{[33]} E_{\rm corr}$  with

$$E_{\rm corr} = E_{\rm exact} - E_{\rm HF} . (2.84)$$

In general, two particles are correlated if the joint probability of both does not factorize into a product of individual probabilities. For correlated electrons, this results in

$$\rho(1,2) \neq \rho(1)\rho(2) .$$
(2.85)

For a HF Slater determinant, these densities can be described in terms of spin functions by using (2.10) and (2.24) to give

$$\rho_{\rm HF}(1) = \sum_{i} |\phi_i(1)|^2 \tag{2.86}$$

and

$$\rho_{\rm HF}(1,2) = \sum_{ij} |\phi_i(1)|^2 |\phi_j(2)|^2 - \sum_{ij} \phi_i(1)^* \phi_i(2) \ \phi_j(2)^* \phi_j(1)$$
$$= \rho_{\rm HF}(1) \rho_{\rm HF}(2) - \left(\Gamma^{(1)}\right)_2^1 \left(\Gamma^{(1)}\right)_1^2 . \tag{2.87}$$

There are two cases regarding the spin of the two electrons. With the same spin, the density matrix terms on the right of (2.87) remain, resulting in correlation of the electrons. When occupying the same position in space, the density matrices reduce to the reduced densities and the whole expression vanishes, which is in accordance to the Pauli exclusion principle and is called the *Fermi hole* (see Figure 2.1). With opposing spin, the density matrices vanish everywhere and the pair density can be expressed as a product of one electron densities, therefore, the electrons are uncorrelated in the HF approximation.

The exact pair density is illustrated in Figure 2.1 and shows a sharp cusp at  $r_{12} = 0$  which is called the *Coulomb cusp*.<sup>[34]</sup> The Pauli exclusion principle does not forbid opposing electrons to occupy the same space, and the dip in probability density is a result of the Coulomb repulsion of the electronic charges. Notably, the HF approximation deviates not only for small distances  $r_{12}$ , but also for moderately large ones, resulting in an over and underestimation of the density depending on distance.



Figure 2.1: Schematic illustration of the Fermi hole on the left and the Coulomb cusp on the right. The HF solution is drawn with dashed lines.

These unaccounted spin interactions contribute to a part of the correlation energy (2.84), which is called the *dynamic correlation*, resulting from instantaneous electron-electron interactions. The other contributor is the *static correlation*, arising from near degeneracy effects of Slater determinants.

$$E_{\rm corr} = E_{\rm corr}^{\rm dyn} + E_{\rm corr}^{\rm stat} \tag{2.88}$$

While this separation might be useful for discussions, both emerge from the same underling interaction. A clean separation cannot be achieved, and methods that primarily recover one type of correlation indvertently also contribute some of the other.

#### 2.3.1 Configuration Interaction

Configuration Interaction (CI) is a widely used method in quantum chemistry that primarly recovers dynamic correlation. Starting with a set of orthogonal orbitals, typically generated by a Hartree-Fock calculation, the CI wavefunction is expanded as a linear combination of substituted Slater determinants. These substitutions are constructed by exchanging occupied orbitals  $\{i \dots\}$  in the reference determinant  $|\Psi_0\rangle$  with virtual (unoccupied) orbitals  $\{a \dots\}$ . These determinants form the *N*-particle Fock space  $\mathbb{F}_N$ , which is the many particle equivalent to the Hilbert space  $\mathbb{H}$  described in the Hartree-Fock method. The set of all single substitutions is commonly referred to as  $|S\rangle$ , double substitutions as  $|D\rangle$ , triple substitutions as  $|T\rangle$  and so forth. The Full CI wavefunction includes all possible substituted determinants and is the best (and with a complete one-particle basis, exact) solution to the electronic Schrödinger equation.

$$|\Psi_{\rm FCI}\rangle = |\Psi_0\rangle + \sum_{ia} c_i^a |\Psi_i^a\rangle + \sum_{\substack{i < j \\ a < b}} c_{ij}^{ab} |\Psi_{ij}^{ab}\rangle + \sum_{\substack{i < j < k \\ a < b < c}} c_{ijk}^{abc} |\Psi_{ijk}^{abc}\rangle + \dots$$

$$= |\Psi_0\rangle + |S\rangle + |D\rangle + |D\rangle + |T\rangle + \dots ,$$

$$(2.89)$$

where the intermediate normalization  $\langle \Psi_{\text{FCI}} | \Psi_0 \rangle = 1$  is assumed. Representing the electronic Hamiltonian (2.7) in  $\mathbb{F}_N$  gives the Hamilton matrix **H** with the elements

$$[\mathbf{H}]_{ij} = \langle i | \hat{H} | j \rangle \quad . \tag{2.90}$$

Following the Slater-Condon rules in Table 2.1, any matrix elements between determinants that differ by more than two substitutions vanish. Additionally, if the reference determinant was obtained from a canonical Hartree-Fock calculation, Brillouin's theorem<sup>[35]</sup> states that

$$\langle S|\hat{H}|\Psi_0\rangle = 0. \qquad (2.91)$$

Together, the Full CI Hamiltonian matrix is

$$\mathbf{H} = \begin{pmatrix} \langle \Psi_0 | \hat{H} | \Psi_0 \rangle & & \cdots \\ 0 & \langle S | \hat{H} | S \rangle & & \cdots \\ \langle D | \hat{H} | \Psi_0 \rangle & \langle D | \hat{H} | S \rangle & \langle D | \hat{H} | D \rangle & & \cdots \\ 0 & \langle T | \hat{H} | S \rangle & \langle T | \hat{H} | D \rangle & \langle T | \hat{H} | T \rangle & & \cdots \\ 0 & 0 & \langle Q | \hat{H} | D \rangle & \langle T | \hat{H} | T \rangle & \langle Q | \hat{H} | Q \rangle & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} .$$
(2.92)

While Full CI provides the most accurate description, its computational cost grows exponentially with system size. Specifically, the many-particle basis size increases with  $\binom{N}{K}$  for N electrons and K spin orbitals. Therefore, the FCI ansatz is computationally infeasible for all but the smallest molecules.

By truncating expression (2.92) with respect to a given substitution order, the exponential is reduced to polynomial scaling. The widely used CISD method, which includes single and double excitations, has a scaling of  $O(N^6)$ . In general, the considered highest substitution order t determines the scaling, which is  $O(N^{(2t+2)})$ .

#### CSFs

Instead of using Slater determinants as the many-particle basis for the Hamilton matrix, linear combinations of Slater determinants called Configuration State Functions (CSFs)  $|\Theta_i\rangle$  can be used,

$$|\Theta_{S,M_S}\rangle = \sum_k a_k |\Phi_k\rangle \quad . \tag{2.93}$$

Each CSF corresponds to a particular spin and spatial symmetry, ensuring that the wave function corresponds to spin eigenstates, meaning, is also the eigenfunction of the  $\hat{S}^2$  and  $\hat{S}_z$  operator. By using CSFs, rather than individual Slater determinants, the computational effort of diagonalizing the Hamilton matrix is reduced as only configurations with the correct symmetry properties are included.<sup>b</sup>

#### 2.3.2 Spin Eigenfunctions

By definition, angular momentum operators follow the relations

$$\left[\hat{S}_{i},\hat{S}_{j}\right] = \mathrm{i}\epsilon_{ijk}\hat{S}_{k} \qquad \left[\hat{S}^{2},\hat{S}_{i}\right] = 0 \qquad i,j,k \in \{x,y,z\} , \qquad (2.94)$$

with the Levi-Civita symbol  $\epsilon_{ijk}$ . For electron spin, the eigenfunctions  $|\sigma\rangle$  satisfy

$$\hat{S}^2 |\sigma\rangle = s(s+1) |\sigma\rangle \qquad \hat{S}_z |\sigma\rangle = m_s |\sigma\rangle , \qquad (2.95)$$

where s is the spin and  $m_s$  the magnetic quantum number. These eigenfunctions correspond to the one-particle functions  $|\alpha\rangle$  and  $|\beta\rangle$  with  $s = \frac{1}{2}$  and  $m_s = \pm \frac{1}{2}$ . In many-electron systems, the total spin operators are constructed as the sum of the individual ones, giving

$$\hat{S}_i = \sum_{\nu}^n \hat{S}_i(\nu) , \qquad \hat{\vec{S}} = \begin{pmatrix} \hat{S}_x \\ \hat{S}_y \\ \hat{S}_z \end{pmatrix} .$$
(2.96)

A many-particle spin eigenfunction  $|S, M_S\rangle$  is defined by the quantum number S and  $M_S$ and fulfils the eigenvalue equations<sup>[37]</sup>

$$\hat{S}^2 | S, M_S \rangle = S(S+1) | S, M_S \rangle$$
  $\hat{S}_z | S, M_S \rangle = M_S | S, M_S \rangle$ . (2.97)

b While the diagonalization of the CSF basis matrix is faster, historically the construction of the matrix was inefficient, making the Determinantal approach superior in terms of computation time. Recently, the CSF construction was improved, shifting the bottleneck back to the matrix diagonalization.<sup>[36]</sup>

The  $\hat{S}_z$  operator commutes for different particles

$$\left[\hat{S}_{z}(i), \hat{S}_{z}(j)\right] = 0$$
, (2.98)

which makes a simple product of one-particle spin functions sufficient to describe the  $M_S$  value with

$$|\theta_i\rangle = \prod_j^n |\sigma_i(j)\rangle , \qquad (2.99)$$

which is called primitive spin function. Forming a correct function which also is eigenfunction to the  $\hat{S}^2$  operator requires in general a linear combination of  $|\theta_i\rangle$ . For  $\nu$  as the number of  $\alpha$  electrons in an *n*-particle system, there are  $\binom{n}{\nu}$  possible primitive spin functions. Summing over all possible values of  $\nu$  gives the total number of possible primitive spin functions for *n* particles as

$$\sum_{\nu}^{n} \binom{n}{\nu} = 2^{n} . \tag{2.100}$$

Clearly, finding a set of orthogonal linear combinations of these  $2^n$  functions to create the desired spin eigenfunctions needs some method of systematic construction.<sup>[37]</sup>

A natural idea is to build these functions by adding  $\alpha$  and  $\beta$  to previous functions and updating the quantum numbers accordingly. It is useful to visualize this by building a network of nodes where adding  $\alpha$  is taking a step up, and  $\beta$  a step down. This gives the *branching diagram* shown in Figure 2.2. Each node contains the number of routes by which it can be reached, which is the sum of the nodes from the ingoing edges. This recursion



Figure 2.2: Branching diagram of genealogical spin coupling for all S of up to eight electrons. The nodes contain the number of routes leading to them.

relation can be expressed as

$$g(n,S) = g\left(n-1, S+\frac{1}{2}\right) + g\left(n-1, S-\frac{1}{2}\right) \text{ and } g\left(1,\frac{1}{2}\right) = 1,$$
 (2.101)

which are called genealogies, as they trace the path of construction.<sup>[37]</sup> With this idea, the  $\{|n, S\rangle\}$  spin functions can be reached similarly to (2.101) by construction with adding to the total spin S starting from  $\{|n-1, S-\frac{1}{2}\rangle\}$ , or with subtraction starting from  $\{|n-1, S+\frac{1}{2}\rangle\}$ . To keep track and distinguish the different routes taken this way in Figure 2.2, the genealogy vector  $\vec{k}$  is introduced, which records the values S took during construction (effectively tracing the path in the diagram). With this, the formula for addition from  $\{|n-1, S-\frac{1}{2}\rangle\}$  to  $\{|n, S\rangle\}$  reads<sup>[37]</sup>

$$\left| n, S, M_S; \vec{k} \right\rangle = \frac{1}{\sqrt{2S}} \left( \sqrt{S + M_S} \left| n - 1, S - \frac{1}{2}, M_S - \frac{1}{2}; \vec{k}' \right\rangle \cdot \left| \alpha(n) \right\rangle \right. \\ \left. + \sqrt{S - M_S} \left| n - 1, S - \frac{1}{2}, M_S + \frac{1}{2}; \vec{k}' \right\rangle \cdot \left| \beta(n) \right\rangle \right) , \qquad (2.102)$$

and for subtraction from  $\{|n-1, S+\frac{1}{2}\rangle\}$  to  $\{|n, S\rangle\}$  it is

$$\left| n, S, M_S; \vec{k} \right\rangle = \frac{1}{\sqrt{2S+2}} \left( -\sqrt{S-M_S+1} \left| n-1, S+\frac{1}{2}, M_S-\frac{1}{2}; \vec{k}' \right\rangle \cdot |\alpha(n)\rangle \right. \\ \left. +\sqrt{S+M_S+1} \left| n-1, S+\frac{1}{2}, M_S+\frac{1}{2}; \vec{k}' \right\rangle \cdot |\beta(n)\rangle \right) .$$
(2.103)

This way, all possible spin eigenfunctions are generated genealogically and form an orthonormal set.<sup>[37]</sup>

#### 2.3.3 CI in an Active Space

In many cases, instead of attempting to describe the correlation of all electrons in a given system, it suffices to restrict the correlation to a subset of orbitals, known as the *active space*.

#### **Complete Active Space Configuration Interaction**

When only substitutions within this space are allowed, the FCI method reduces to the *Complete Active Space Configuration Interaction* (CASCI) method. While this approach significantly reduces computational cost, as it scales with the size of the active space rather than the entire system, it introduces several challenges. First, the CASCI wave function is highly dependent on the choice of initial orbitals, which makes an adequate Hartree-Fock calculation paramount. Additionally, the proper selection of the active space

requires careful consideration of the relevant orbitals and electrons. These dependencies are inherent to all CAS methods and are generally addressable. However, the primary limitation of CASCI lies in its orbitals being used for higher-order substitutions. Specifically, the orbitals are optimized for the ground state HF occupation, which introduces errors when the coefficients of higher substitutions become significant. This occurs, for example, when studying excited states or when the system is inherently multi-determinantal.

#### Complete Active Space Self-Consistent Field

To address these limitations, the orbitals can be allowed to relax independently for each desired state, giving the *Complete Active Space Self-Consistent Field* (CASSCF) method. The optimization process in CASSCF involves two nested levels of iterations, so-called macro and micro iterations. The *macro iterations* form the outer loop, where a CASCI calculation is performed using the current orbitals. The *micro iterations* form the inner loop, during which the orbitals are updated in a SCF procedure for a fixed set of CI coefficients.

#### State Averaging

In complexes of d- and f-block elements, the ground state is often nearly degenerate with multiple other states, that exhibit high d- or f-character. Orbitals with a higher angular momentum, such as those in the d- and f-shell, are typically more radially confined. This confinement shields them from some influence of the ligand field, preserving a degree of atomic degeneracy. When performing a CASSCF calculation, these near degeneracies lead to different sets of orbitals for near degenerate states, which complicates subsequent calculations, such as transition probabilities. The solution is to optimize one set of orbitals for all degenerate states. This is performed by the state-averaged (SA) CASSCF method, which averages the energy of several states during the optimization process. The weights of the different states in the average can be adjusted to accommodate more complex systems, if necessary. For f-element complexes, the active space is typically restricted to orbitals with f-character. The resulting wave function consists of a single set of f-shell averaged orbitals, which describe all states arising from the  $f^n$  occupation.

An equivalent description involves averaging the orbitals beforehand through a CAHF calculation with an averaged f-shell, followed by a CASCI calculation with the same active space. In this case the drawbacks of CASCI are mitigated by these averaged orbitals. This combination of CAHF and CASCI significantly reduces computational cost, as it avoids the redundant work performed by the SA-CASSCF method in such cases.

## 2.4 Relativistic Effects

The Schrödinger Hamiltonian has been sufficient for describing interactions in light molecules, with theoretical results matching exceptionally well with experimental data (e.g., for the hydrogen atom and helium atom). However, for larger atoms, small discrepancies arise due to the incomplete description, primarily the neglect of relativistic effects. Einstein formulated the special theory of relativity, which is based on two postulates:

- 1. All laws of physics are identical in all inertial frames of reference.
- 2. In a vacuum, light propagates at a constant speed c in all inertial frames.

These postulates imply that observations differ for observers in relative motion.

#### Lorentz Transformations

Consider an inertial reference frame moving with a constant velocity v in the x-direction relative to another frame. The coordinates are *boosted* in x-direction by a Lorenz transformation with

$$x' = \gamma(x - vt) , \qquad (2.104)$$

$$y' = y , \quad z' = z ,$$
 (2.105)

$$t' = \gamma \left( t - \frac{vx}{c^2} \right) , \qquad (2.106)$$

where the Lorentz factor  $\gamma$  is given by

$$\gamma = \sqrt{\frac{1}{1 - v^2/c^2}} \,. \tag{2.107}$$

For small velocities  $(v \ll c)$ ,  $\gamma \approx 1$ , and the Lorentz transformations reduce to the Galilean transformations. However, as the velocity approaches a significant fraction of the speed of light, relativistic effects become non-negligible. In the boosted frame, time runs slower (a phenomenon known as *time dilation*), and distances contract along the direction of motion (*length contraction*). Additionally, the mass of a particle in the boosted frame appears larger when measured from the stationary frame.

$$m = \gamma m_0 , \qquad (2.108)$$

where  $m_0$  is the rest mass of the particle, which is equal across all reference frames.

#### **Implications for Quantum Mechanics**

The effects of Lorentz transformations lead to the unification of space and time into a four-dimensional space-time. A relativistic description of a system must be invariant under these transformations, which is not the case for the Hamiltonian operator used so far. The time-dependent Schrödinger equation (2.1) contains second derivatives in space but only first derivatives in time. This unequal treatment of space and time coordinates inherently breaks Lorentz invariance.

#### 2.4.1 Dirac Equation

The pursuit of a relativistic framework for describing particles, whether fermions or bosons, is deeply rooted in the necessity of Lorentz invariance. The first relativistic wave equation proposed was the Klein–Gordon equation:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \frac{m^2 c^2}{\hbar^2}\right)\Psi(t,\vec{r}) = 0, \qquad (2.109)$$

which features second-order derivatives in both space and time. While the Klein–Gordon equation is suitable for describing spinless particles, it is inadequate for electrons.

Paul Dirac addressed this limitation by formulating an equation with first-order derivatives in both time and space, now known as the Dirac equation,

$$\left(i\frac{\partial}{\partial t} + \phi\right)\Psi + c\vec{\boldsymbol{\alpha}}\left(i\vec{\nabla} - \vec{A}\right)\Psi = \boldsymbol{\beta}mc^{2}\Psi, \qquad (2.110)$$

where  $\vec{A}$  and  $\phi$  represent the vector and scalar potentials arising from electromagnetic fields. For static fields, such as those in the Born-Oppenheimer approximation, the time-dependent Dirac equation (2.110) can be reduced to its time-independent form

$$c(\vec{\boldsymbol{\alpha}} \cdot \vec{p})\Psi + (\boldsymbol{\beta}mc^2 + V)\Psi = E\Psi. \qquad (2.111)$$

The quantities  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$  are matrices of the form

$$\boldsymbol{\beta} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & -\mathbf{I}_2 \end{pmatrix}, \quad \boldsymbol{\vec{\alpha}} = \begin{pmatrix} \boldsymbol{\alpha}_x \\ \boldsymbol{\alpha}_y \\ \boldsymbol{\alpha}_z \end{pmatrix}, \quad \boldsymbol{\alpha}_k = \begin{pmatrix} \mathbf{0}_2 & \boldsymbol{\sigma}_k \\ \boldsymbol{\sigma}_k & \mathbf{0}_2 \end{pmatrix} \quad \text{with} \quad k \in \{x, y, z\}, \quad (2.112)$$

where  $I_2$  and  $O_2$  are the 2 × 2 identity and null matrices, respectively. The  $\sigma_k$  are the Pauli
matrices, which are defined as

$$\boldsymbol{\sigma}_{x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_{y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_{z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \vec{\boldsymbol{\sigma}} = \begin{pmatrix} \boldsymbol{\sigma}_{x} \\ \boldsymbol{\sigma}_{y} \\ \boldsymbol{\sigma}_{z} \end{pmatrix}. \quad (2.113)$$

The presence of these matrices indicates that the wave function  $\Psi$  is no longer a scalar function but a four-component vector, resulting in the time-dependent wave function for (2.110)

$$\Psi(\vec{r},t) = \begin{pmatrix} \psi_1(\vec{r},t) \\ \psi_2(\vec{r},t) \\ \psi_3(\vec{r},t) \\ \psi_4(\vec{r},t) \end{pmatrix} \begin{cases} \psi^L(\vec{r},t) \\ \psi^S(\vec{r},t) \\ \psi^S(\vec{r},t) \end{cases}$$
(2.114)

The Dirac equation describes all states of spin- $\frac{1}{2}$  particles, yielding both positive and negative energy solutions. The positive energy solutions correspond to electron states, while the negative energy solutions describe the states of the positron, the electron's antiparticle.

The block structure of the  $\alpha$  and  $\beta$  matrices allows the time-independent Dirac equation (2.111) to be expressed in a more compact form<sup>[38]</sup>

$$\begin{pmatrix} V + mc^2 & c(\vec{\boldsymbol{\sigma}} \cdot \vec{p}) \\ c(\vec{\boldsymbol{\sigma}} \cdot \vec{p}) & V - mc^2 \end{pmatrix} \begin{pmatrix} \psi^L \\ \psi^S \end{pmatrix} = E \mathbf{I}_2 \begin{pmatrix} \psi^L \\ \psi^S \end{pmatrix} , \qquad (2.115)$$

where the wave function is partitioned into two two-component spinors,  $\psi^L$  and  $\psi^S$ . The labels L and S denote the large and small contributions of the respective spinors to the positive energy electronic states.

#### 2.4.2 Two Component Methods

Solving the Dirac equation in its full four-component form is computationally and conceptually challenging. Due to the positronic states, the Dirac Hamiltonian is not bounded from below, making a simple variational approach impossible. Dirac himself proposed addressing this by formally occupying these states, but this solution is not entirely satisfactory for practical applications. Another problem is the intrinsic dependence on spin, which results in complex valued arithmetic. Therefore, it is often desirable to separate the Hamiltonian into a spin-free part and a spin-orbit part, as the majority of relativistic effects are not spin-dependent. And lastly, the computational cost increases significantly due to the large number of two-electron integrals required. A large majority of these integrals arises from the small-component wave function, which requires a more extensive basis set compared to the large component. This is due to the *kinetic balance condition*, which ensures that the non-relativistic limit of the Dirac Hamiltonian correctly recovers the kinetic energy term. This condition is expressed as

$$\psi^S(\vec{r}) = (\vec{\boldsymbol{\sigma}} \cdot \vec{p}) \psi^L(\vec{r}) , \qquad (2.116)$$

which necessitates the use of basis functions with matching higher and lower angular momenta for the small component (and in turn a larger basis set). This is particularly problematic given the relatively small contribution of the small component to the overall wave function.

#### Approaches to Eliminate the Small Component

To address these challenges, several methods have been developed to eliminate the small component, effectively reducing the four-component Dirac equation to a two-component form. Notable approaches include the Foldy-Wouthuysen (FW) transformation<sup>[39]</sup>, Douglas-Kroll-Hess (DKH) transformation<sup>[40–42]</sup> or the Wood-Boring Hamiltonian.<sup>[43]</sup> While these methods will not be discussed in detail here, the reader is referred to ref. [38] for a comprehensive overview of these and other techniques.

#### 2.4.3 Relativistic Pseudopotentials

As previously discussed, solving two- or four-component all-electron (AE) relativistic problems is computationally demanding. An alternative approach is the use of *effective core potentials* (ECPs)<sup>[44–46]</sup>, where chemically inert core electrons are replaced by a potential that mimics their effect on the remaining *valence* electrons. This approach is particularly effective for atoms with many electrons, where the core electrons are largely shielded from external influences. Relativistic effecs for valence electrons originating from the nuclear Coulomb potential and the shielding due to the core electron system can be described implicitly by an effective potential. This allows ECPs to address both relativistic effects and computational efficiency simultaneously.

There are two primary approaches to ECPs, the *model potential* (MP) and the *pseudopotential* (PP) approach. The MP attempts to model the frozen-core Hartree-Fock operator for the valence system, preserving the orbital shape in the core region. In contrast, the PP transforms the valence orbitals into pseudo-valence orbitals, which have a simplified radial nodal structure. In the following pages, the PP ansatz is outlined.

#### (General) Phillips-Kleinman PP

One of the first pseudopotentials was proposed by Hellmann in 1935<sup>[47]</sup>, but it was not until 1959 that Phillips and Kleinman provided a rigorous theoretical foundation. The orthonormal solutions to the effective one-electron Hamiltonian  $\hat{H}_{eff}$  are divided into one valence (v) and several core (c) eigenfunctions,

$$\hat{H}_{eff} |\psi_i\rangle = \epsilon_i |\psi_i\rangle$$
,  $\langle\psi_i|\psi_j\rangle = \delta_{ij}$  and  $i, j \in \{v, c\}$ . (2.117)

An arbitrary, normalized function  $|\psi_p\rangle$  is expressed as a linear combination of the valence and core eigenfunctions with

$$|\psi_p\rangle = N_p \left( |\psi_v\rangle + \sum_c a_c |\psi_c\rangle \right) , \qquad (2.118)$$

where  $N_p$  is a normalization constant. By introducing the Phillips-Kleinman (PK) pseudopotential<sup>[48]</sup>,

$$\hat{V}^{PK} = \sum_{c} (\epsilon_v - \epsilon_c) |\psi_c\rangle \langle\psi_c| , \qquad (2.119)$$

this so-called pseudo-valence orbital (2.118) is the eigenfunction of the PK equation

$$\left(\hat{H}_{eff} + \hat{V}^{PK}\right) \left|\psi_p\right\rangle = \epsilon_v \left|\psi_p\right\rangle \ . \tag{2.120}$$

The PK pseudopotential (2.119) is not a potential in the traditional sense, as it does not only depend on the position  $\vec{r}$ , hence the name. Instead, it is energy-dependent (through  $\epsilon_v$ and  $\epsilon_c$ ) as well as non-local, as it depends on all  $|\psi_c\rangle$ . The implicit core-valence orthogonality condition ensures variational stability, guaranteeing that  $\epsilon_v$  is approached only from above. However, this formalism is limited to single-electron systems and requires knowledge of the core eigenfunctions  $|\psi_c\rangle$ , which necessitates solving the full problem, thereby offering no computational savings.

Nearly a decade later, Weeks and Rice generalized the PK formalism to systems with multiple valence electrons.<sup>[49]</sup> In this generalized Phillips-Kleinman (GPK) approach, only the valence orbitals  $|\psi_v\rangle$  are eigenfunctions of  $\hat{H}_{eff}$ , and they are no longer orthogonal to the core orbitals  $|\psi_c\rangle$ . Introducing the projector  $\hat{P}_c$  on the orthogonal core orbitals

$$\hat{P}_c = \sum_c |\psi_c\rangle \langle\psi_c| \quad , \tag{2.121}$$

the GPK PP  $\hat{V}^{GPK}$  is expressed as

$$\hat{V}^{GPK} = -\hat{H}_{eff}\hat{P}_c - \hat{P}_c\hat{H}_{eff} + \hat{P}_c\hat{H}_{eff}\hat{P}_c + \epsilon_v\hat{P}_c .$$
(2.122)

This non-local, energy-dependent, effective one-electron operator yields the GPK eigenvalue equation

$$(\hat{H}_{eff} + \hat{V}^{GPK}) |\psi_p\rangle = \epsilon_v |\psi_p\rangle . \qquad (2.123)$$

While this reformulation still does not reduce computational cost, it demonstrates that the same results as an AE calculation can be achieved using a suitable effective valence-only Hamiltonian. Additionally, the pseudovalence orbitals can then be simplified, as they no longer need to satisfy explicit orthogonality requirements to the core orbitals. This allows for a reduction in radial nodes and, consequently, a smaller basis set.

#### Valence-Only Hamiltonians

In ECP methods, the nuclear charge  $Z_I$  is reduced to account for the removed core electrons. The core charge  $Q_I$  is defined as

$$Q_I = Z_I - n_{I_c} , (2.124)$$

where  $n_{I_c}$  is the number of core electrons of atom I. The total number of valence electrons in the molecule is denoted by  $n_v$ . The effective valence-only (VO) Hamiltonian is derived from the AE Hamiltonian and generally takes the form

$$\hat{H}_v = \sum_{i=1}^{n_v} \hat{h}_v(i) + \sum_{i< j}^{n_v} \hat{g}_v(i,j) + \hat{V}_{cpp} + \hat{V}_{cc} . \qquad (2.125)$$

Here,  $\hat{V}_{cpp}$  is the core-polarization potential (CPP), which accounts for the polarizability of the core. This term will not be discussed in any detail; for further information see ref. [45,50]. The term  $\hat{V}_{cc}$  includes the repulsion between cores/nuclei as well as two-core corrections  $\Delta \hat{V}_{cc}^{IJ}$  for the point charge approximation. This only becomes important if very large cores are used (e.g., one-valence alkaline atoms).

$$\hat{V}_{cc} = \sum_{I < J}^{N} \left( \frac{Q_I Q_J}{r_{IJ}} + \Delta \hat{V}_{cc}^{IJ}(r_{IJ}) \right)$$
(2.126)

With ECPs which are short of fully relativistic, the effective one- and two-electron operators  $\hat{h}_v$  and  $\hat{g}_v$  in equation (2.125) take the non-relativistic form

$$\hat{h}_v(i) = -\frac{1}{2}\Delta_i + \hat{V}_{cv}(i) \quad \text{and} \quad \hat{g}_v(i,j) = \frac{1}{r_{ij}}.$$
 (2.127)

The term  $\hat{V}_{cv}$  models the interactions of the valence electrons with the nucleus and the removed core electrons. For molecules, it is assumed that this can be expressed as a superposition of atomic contributions, giving

$$\hat{V}_{cv}(i) = \sum_{I}^{N} \left( -\frac{Q_I}{r_{iI}} + \Delta \hat{V}_{cv}^{I}(i) \right) .$$
(2.128)

Substituting (2.127), (2.128) and (2.126) into (2.125) yields the VO model Hamiltonian

$$\hat{H}_{v} = -\frac{1}{2} \sum_{i}^{n_{v}} \Delta_{i} + \sum_{i < j}^{n_{v}} \frac{1}{r_{ij}} + \sum_{i}^{n_{v}} \sum_{I}^{N} \left( -\frac{Q_{I}}{r_{iI}} + \Delta \hat{V}_{cv}^{I}(i) \right) + \sum_{I < J}^{N} \left( \frac{Q_{I}Q_{J}}{r_{IJ}} + \Delta \hat{V}_{cc}^{IJ}(r_{IJ}) \right) + \hat{V}_{cpp} .$$
(2.129)

#### The Semilocal Atomic PP Ansatz

As the GPK PP before, the goal is to describe  $\Delta \hat{V}_{cv}$  in equation (2.129). As already mentioned, the first proposal for this term was made by Hellmann in 1935 for a one-valence electron potassium atom<sup>[47]</sup>

$$\hat{V}_{cv} \approx \Delta \hat{W}_{PP} = -\frac{1}{r} + \frac{2.74}{r} e^{-1.16r}$$
 (2.130)

However, it was soon realized that a simple r dependency was insufficient for atoms with valence orbitals of different angular symmetries. The solution was to include an explicit l-dependency. Assuming that pseudopotentials for multiple atoms combine linearly, the semilocal pseudopotential ansatz for a single atom I is

$$\Delta \hat{V}_{cv}^{I}(i) \approx \Delta \hat{W}_{PP}(i) \approx \sum_{l=0}^{\infty} \hat{W}_{lj}' \sum_{jm_j} |ljm_j\rangle \langle ljm_j| \quad .$$
(2.131)

By collecting all contributions for  $l \ge L$  (where L - 1 as the maximum angular momenta occupied in the core) into one term  $\hat{W}_{LJ}$ , equation (2.131) can be rewritten as

$$\Delta \hat{W}_{PP} = \hat{W}_{LJ} + \sum_{l=0}^{L-1} \hat{W}_{lj} \sum_{jm_j} |ljm_j\rangle \langle ljm_j| . \qquad (2.132)$$

31

The explicit projection onto the 2-spinor spherical harmonics  $|ljm_j\rangle$  indicate, that this formalism can incorporate both spin-free relativistic effects and spin-orbit interactions. These contributions can be separated by defining potentials  $\hat{A}_l$  and  $\hat{B}_l$  for the spin-free and spin-orbit part  $\hat{A}$  and  $\hat{B}$ , respectively, as<sup>[38]</sup>

$$\hat{A}_{l} \coloneqq \frac{l\hat{W}_{l,l-\frac{1}{2}} + (l+1)\hat{W}_{l,l+\frac{1}{2}}}{2l+1} , \qquad (2.133)$$

$$\hat{B}_l \coloneqq 2\frac{\hat{W}_{l,l+\frac{1}{2}} - \hat{W}_{l,l-\frac{1}{2}}}{2l+1} \ . \tag{2.134}$$

Expanding the 2-spinor spherical harmonics  $|ljm_j\rangle$  in terms of spherical harmonics and spin functions yield<sup>[38]</sup>

$$\Delta \hat{W}_{PP} = \underbrace{\hat{W}_{LJ} + \sum_{l=0}^{L-1} \hat{A}_l \sum_{m_l} |lm_l\rangle \langle lm_l|}_{\hat{A}} + \underbrace{\hat{\mathbf{S}} \cdot \sum_{l=0}^{L-1} \hat{B}_l \sum_{m_l, m_l'} |lm_l\rangle \langle lm_l| \hat{\mathbf{L}} |lm_l'\rangle \langle lm_l'|}_{\hat{B}} \quad (2.135)$$

The component  $\hat{A}$  is spin-free and can be added directly to the one-particle Hamiltonian  $\hat{h}$ . In contrast, the component  $\hat{B}$  describes the one-particle spin-orbit interaction and requires a true two-component description, necessitating the use of complex algebra.

## Chapter 3

## Implementation

This chapter focuses on the approaches and challenges of implementing computational methods in the in-house quantum chemical code base of the Theoretical Institute at the University of Cologne. First, the code base is introduced and the implementation goals are outlined, where after some improvements and implementations are discussed in detail.

## 3.1 The Quantum Objects Library

The Quantum Objects Library (QOL) is a quantum chemical program suite developed and maintained by Michael Hanrath.<sup>[9]</sup> Since its inception in the early 2000s, the QOL has evolved from its initial focus on arbitrary excitation single- and multi-reference Coupled Cluster methods to encompass a wide array of novel and advanced wave function-based techniques. Accessing the full source code allows to more easily implement new research approaches, without depending on commercial software. The library features an efficient implementation of the Hartree-Fock method<sup>[51,52]</sup>, including a 2-component version<sup>[53]</sup> and orbital localization schemes,<sup>[54]</sup> alongside capabilities for utilizing and evaluating Pseudo-Potentials<sup>[53]</sup> and recently Core Polarization Potentials.<sup>[50]</sup> The Coupled Cluster method was also expanded to accurately describe open-shell spin states with the Spin-Adapted and Spin-Complete Coupled Cluster method.<sup>[55–58]</sup> And finally, the incremental scheme was expanded over the years,<sup>[59]</sup> and recently a novel approach was introduced.<sup>[60,61]</sup> These implementations build on many more contributions to the code base<sup>[62,63]</sup> and are by no means a complete list.

The QOL is written in modern C++, with the latest additions using the C++23 standard.<sup>[64]</sup> Where applicable, the linear algebra library LAPACK<sup>[65]</sup> is used to efficiently perform computations involving matrices.

## 3.2 Implementation Goals

During the author's master thesis in 2020, the Configuration-Averaged Hartree-Fock (CAHF) method was implemented in the QOL. At that time, this method was not yet implemented in the commercial program package MOLPRO<sup>[66]</sup>, making it a desirable and novel target. Even after the introduction of CAHF in MOLPRO version 2021, due to unrestricted access to the source code as well as the option to modify parts of it, this new and independent implementation retains value. The QOL implementation was the first step towards modelling lanthanide and actinide complexes with the efficient CAHF/CASCI method (see section 2.3.1), with the added capability to include PP spin-orbit contributions.

Despite its extensive capabilities, the QOL has historically been limited by the lack of a unified input format, restricting users to running each method in isolation. To address this limitation and enhance the QOL's usability for larger and electronically more complex molecules (such as lanthanide and actinoide complexes), a standardized input method was deemed essential. Additionally, the ability to seamlessly utilize results from one calculation as inputs for subsequent computations – without the need to read/write everything on the disk – was identified as a critical improvement.

If larger computations were to be carried out, the CAHF implementation in the QOL needed to be optimized, as time constraints during the master thesis limited its refinement. The internal integral generation as well as the outside acceleration technique are to be investigated for potential improvements.

Finally, the utility of CAHF orbitals depends on their application in a CASCI(-SO) calculation, which has not yet been implemented – outlining the last objective.

### **3.3 Input Format**

Early in this process, the decision was made to adopt an input file approach. The commitment was made to adopt a straightforward structure, where all instructions are expressed as key = value statements. This design ensures that maintainers and users can easily define parameters and settings without unnecessary complexity. However, certain settings, such as molecular geometries, method-specific options, and control structures like loops, require more than a single value to be fully expressed. To address this, the syntax incorporates the concept of blocks, which allow users to bundle related information together. Blocks are defined using the syntax key = {[block content]}, where the key identifies the type of information, including how the [block content] is parsed. By default, a child parser is created with the [block content] as an argument, inheriting all settings from the parent parser up to that point. This approach enables the separation of options for different methods, as well as the flexibility to locally overwrite settings within specific blocks. Listing 3.1 shows an example for an input file for NO<sup>+</sup>. Apart from geometry, the method blocks rhf, cahf, casci, casci-so, pop, rotate, compare, multipole, and print are implemented so far.

The implementation of loops was long debated, as it would complicate the syntax and in most usecases the same effect could be achieved with external scripts. But when utilizing

```
1
   // key = value
2
   unit = bohr
3
4
   // geometry block, parsing is performed by geometry parser
5
   geometry = {
6
   0.0 0.0 0.0
7
   N 0.0 0.0 0.8
8
   }
9
10 basis = BSE/6-31G
11
12
   // method block, parsing is performed by new instance of same parser
13 rhf = {
     // settings can be locally overwritten
14
15
     basis = BSE/cc-pVDZ
16
     // blocks can be nested
17
     // this block is parsed by a different parser
     shellStructure = {
18
19
       elec 14 // results in charge of +1
20
       orb 7
21
     }
22
     maxIter = 200
23 }
```

Listing 3.1: Example input file for a RHF calculation of NO<sup>+</sup>. The indentations are for readability only and not neccessary.

```
1 loop(variable,startvalue,endvalue,stepsize) = {
2 // incrementing variable by stepsize
3 }
4 loop[variable,value1, ..., valueN] = {
5 // setting variable to the next value
6 }
Listing 2 2: Loop gue
```

Listing 3.2: Loop syntax.

previous results, such as with dissociation curves and using neighbouring converged caluctations as starting guess, loops become useful. Listing 3.2 shows both loops syntax.

#### Sharing Results between Calculations

The next goal is to facilitate the sharing of computational data across various quantum chemical calculation methods. While there are some shared data types between many methods, allocating every possible type which might be needed for all methods is simply infeasable. With C++'s std::any it is possible to store any object in it's place without knowing the explicit type. Only when trying to retrieve the object, the matching type needs to be provided. The class Results in listing 3.3 is a wrapper around std::map<std::string, std::any>. With the retrieval of the value in getResult, the return is an std::optional which only holds the value if str is in the map. This makes it explicit for any user of this class, that the return needs to be checked before trying to access the value.

Each computational method returns a Results object, which is then pushed onto a vector in a map (std::map<std::string, std::vector<Results>>). By default the map entry prev is used, and this map entry is also the default starting point for all methods. A specific save location for the data in the current method can be given by save = [saveName], and a specific start location by load = [saveName]. The push back onto a vector ensures, that deleting or overwriting is not possible. The old result is only shaddowed and can still be accessed by the print method.

```
class Results {
1
\mathbf{2}
     void putResult(const std::string & str, const std::any & result);
3
     template<typename T>
4
     std::optional<T> getResult(const std::string & str) const;
5
6
7
     std::any getAny(const std::string & str) const;
8
9
     bool empty() const;
10
   private:
     std::map<std::string, std::any> _results;
11
12
   };
```

Listing 3.3: The class Results. Used to store any values and objects which might be of interest for future calculations.

## 3.4 Improving the CAHF

The Configuration-Averaged Hartree-Fock (CAHF) method was initially implemented during the author's master thesis, where it successfully passed preliminary test calculations. With a working implementation and test cases in place, subsequent efforts focused on refining and optimizing the code. A key optimization involved consolidating two internal wrapper calls into a single operation, resulting in a noticeable speed-up. The Self-Consistent Field implementations (RHF, ROHF, CAHF) were standardized to ensure that all used the same internal objects, which reduced redundancy in the codebase. The AccelerationWrapper class was introduced, which provides generic interfaces for update methods in the SCF procedure, such as update\_C (updating molecular orbitals), update\_D (updating the density matrix), and update\_F (updating the Fock matrix). This abstraction decouples the SCF procedure from the specific acceleration method, simplifying the logic within each SCF implementation and eliminating redundant code. Additionally, this modular design allows future integration of new acceleration methods without requiring significant modifications.

The Geometric Direct Minimization (GDM) method, an advanced acceleration technique reported in the literature to enhance convergence in challenging cases,<sup>[67,68]</sup> was implemented. GDM is one of the few methods adaptable to multi-shell systems, which are described using ROHF and CAHF methods (involving occupied, active, and virtual orbitals). And the typical systems modelled by these methods involve complex open-shell systems, which show difficulty with convergence. The GDM implementation successfully achieved convergence to the correct stationary states for simple test cases. However, several limitations were observed. In terms of computational performance, the GDM implementation was found to be less efficient compared to the widely used Direct Inversion in the Iterative Subspace (DIIS) method.<sup>[69,70]</sup> This is both in regards to a single iteration and with comparing the number of iteration steps. Even when GDM was applied in the stable region near the stationary state – following initial DIIS iterations – it did not have any speed advantage over continuing with DIIS. Furthermore, the improved stability attributed to GDM in the literature could not be reproduced in this implementation, needing frequent resetting due to unresonable steps.

Therefore, DIIS was selected for subsequent calculations due to its more consistent and predictable convergence behavior, while the GDM method was ignored for further use.

## 3.5 FCI, CI and CASCI

For the different CI methods the new module CASCI was implemented in the QOL, which uses the namespace QOL::CASCI. An overview of the most important classes in this namespace is given in the UML diagram in Figure 3.1. In the following, parts of these classes are highlighted and discussed in more detail.



Figure 3.1: Layout of the CASCI namespace. The classes CI, CASCI and CASCI\_SO are displayed as a combined node for clarity.

#### 3.5.1 Generating the CSF Basis

Early on it was decided, that all CI methods would use a CSF basis, as this would give solutions which are spin eigenstates right out of the box. And while this decision results in a slight increase of computational work when considering CASCI-SO, keeping everything consistent is simpler and reduces errors. In the following the generation of the CSF basis for an active space of (e,n), meaning e electrons in n spatial orbitals, is described.

#### The Freeon Function

Separating the spin from a CSF  $|\Theta\rangle$  can be done by the typical product ansatz

$$|\Theta\rangle = |\vartheta\rangle \,|\theta\rangle \,\,, \tag{3.1}$$

with  $|\vartheta\rangle$  as the spatial and  $|\theta\rangle$  the spin function. The spin functions can be generated following the recursion relation described in equation (2.102) and (2.103). And while the spatial functions can be generated on their own, too, it is much more efficient to generate them during CSF construction. With only a list of occupied spatial orbitals, with no mention of spin, this can be achieved. By stripping a determinant from its references to spin in each spin orbital and converting the indices to spatial indices, the freeon function is obtained. It is extensively used in freeon theory, which will not be discussed in this thesis any further.<sup>a</sup> A freeon function must be expressed in spatial indices, as there is no spin information. Each index can occur twice, indicating a double occupation. With 1.. as a stand-in for both 1 and 1 1 the general freeon function is

$$|\vartheta\rangle_{\text{freeon}} = |1..2..\cdots n..\rangle_f \ . \tag{3.2}$$

The corresponding class FreeonFunction holds two vectors of indices, one for the closed shell and one for the open shell (see Figure 3.2).



Figure 3.2: UML diagram of the classes FreeonFunctionBasis and FreeonFunction. This is a detailed and enlarged view of the layout in Figure 3.1.

a The reader is referred to ref. [71] and [72] for further reading.

#### **Freeon Function Basis**

The class FreeonFunctionBasis is designed to systematically generate all possible freeon functions within a specified active space, by using the internal class Permutation. Instead of utilizing numerical indices, the class employs a bit representation, where each bit position corresponds to a specific index. To handle double indices, each index is represented by two bits, with the only valid double-patterns being 00, 01 and 11. The double-pattern 10 is excluded because it represents the same function as 01 due to the freeon function having no spin information. The basis generation begins with an initial bit-set and iteratively produces all unique permutations. An active space of (e,n) is translated into the reference of FreeonFuctionBasis::Permutation as a bit-set of e times 1, followed by 2n - e times 0.

$$|\vartheta\rangle_{\text{freeon}}^{\text{ref}} = \underbrace{11\dots1}_{e \times 1} \underbrace{0\dots00}_{2n-e \times 0}$$
(3.3)

If e is odd, the last 1 and first 0 are then swapped. The exclusion of the double-pattern 10 complicates the traversing of all valid permutations, as simple bit shifting would also lead to equivalent functions. The solution to the permutation generation challenge is implemented through a constrained algorithmic process, illustrated in Listing 3.4. At its core, the algorithm performs a specific shift operation: it transforms a pattern of the form ab00 into 0a0b, effectively moving the last double-pattern to the right (method shiftLastPair). If the last double-pattern is already positioned at the end of the bitset, the algorithm removes the trailing bits (lines 6-11 and lines 26-30), shifts the remaining pattern (line 10 or line 29), and subsequently reinserts the removed bits (line 13-24, with the explicit method adCarryToPerm). The algorithm is complete, if all bits are shifted to the right. This approach ensures that all valid permutations are generated systematically while avoiding the creation of redundant configurations.

110100	$\rightarrow$	$ 001\rangle$	
11 <mark>00</mark> 01	$\rightarrow$	$ 002\rangle$	
011100	$\rightarrow$	$ 011\rangle$	
01 <mark>01</mark> 01	$\rightarrow$	$ 012\rangle$	$1 1  0 1  0  0 \rightarrow  ppq\rangle$
01 <mark>00</mark> 11	$\rightarrow$	$ 022\rangle$	$\sum_{n} \sum_{q} \sum_{r} \frac{1}{r} $
00 <mark>11</mark> 01	$\rightarrow$	$ 112\rangle$	P q ,
00 <mark>01</mark> 11	$\rightarrow$	$ 122\rangle$	

Figure 3.3: Example output of Listing 3.4 for the active space (3,3).

In Figure 3.3 the effect is illustrated for the active space of (3,3). The reference bit-set 110100 is generated and the function nextPermutation is repeatedly called to generate the next bit-set until the return value is 0. The bitsets can then be translated back to indices in a vector, giving the freeon function basis.

```
nextPermutation(currentPermutation, carry) {
 1
\mathbf{2}
       if (permutationsReachedTheEnd) {
3
           return 0;
4
       }
5
6
       if (currentPermutation.secondLastBitIsSet) {
7
           numberOfRemovedBits = currentPermutation.removeConsecutiveBitsAtEnd;
8
           newCarry = carry + numberOfRemovedBits;
9
           // shift the remaining pattern and add the carry back in
10
           return nextPermuation(shiftLastPair(currentPermutation), newCarry);
       }
11
12
13
       if (carry > 0) {
14
           // the current permutation is not valid, as some bits are taken out
15
           // they need to be added back before returing the permutation
16
           [newPermutation,newCarry] = addCarryToPerm(currentPermutation, carry);
17
           if (newCarry > 0)
18
               // in some cases, adding the carry was not possible
19
               // then more bits need to be taken out, until the resulting
20
               // permutation has enough room on the right for all to fit in
21
               return nextPermutation(newPermutation, newCarry);
22
           else
23
               return newPermutation;
24
       }
25
26
       if (currentPermutation.lastPositionIsSet) {
27
           currentPermutation.removeLastBit;
28
           // shift the remaining pattern and add the carry back in
29
           return nextPermutation(shiftLastPair(currentPermutation), 1);
30
       }
31
32
       return shiftLastPair(currentPermutation);
33
   }
34
   shiftLastPair(currentPermutation) {
35
        [... a b 0 0 ...] -> [... 0 a 0 b ...]
36
37
       return alteredPerm;
38
   }
39
   addCarryToPerm(currentPermutation, carry) {
40
41
       while (carry > 0) {
42
           if (addingCarryNotPossible) {
43
               return alteredPerm,carry;
44
           }
45
46
           // only adding bits AFTER the last pair, NEVER before
47
           Add to last pair:
48
           [\dots 0 1 0 0 \dots] \rightarrow [\dots 1 1 0 0 \dots]
49
           or
           [\dots 1 1 0 0 \dots] \rightarrow [\dots 1 1 0 1 \dots]
50
51
52
           carry = carry - 1;
53
       }
54
       return alteredPerm,0;
55 }
```

Listing 3.4: Algorithm for generating the next permutation in the FreeonFunctionBasis.

#### The CSF Basis

The construction of the CSF basis is achieved by systematically combining freeon functions with appropriate spin eigenfunctions. For each freeon function, the number of open orbitals is determined, and the corresponding spin eigenfunctions are generated. This is achived by using the classes of the **SpinEigenfunction** namespace of the QOL, which was implemented by Nils Herrmann.<sup>[55]</sup> The spin eigenfunctions are constructed using the recursion relations outlined in section 2.3.2, specifically equations (2.102) and (2.103). Each freeon function is then paired with all compatible spin eigenfunctions to form CSFs. This process is iterated over all freeon functions in the basis, giving the full CSF basis.

#### 3.5.2 Scalar Product

The ScalarProduct class is a template class and computes the scalar products between determinants or CSFs (see Figure 3.4). The template parameters are the type of vector (Determinant or CSF) and the type of scalar returned (double or complex<double>). The class holds pointers to the integral matrices for the one- and two-particle operators, which are provided by the MOIntegrals class, and optionally, spin-orbit coupling (SOC) integrals if SOC effects are considered. The scalar product evaluation is performed via the operator()(bra, operator, ket) method, where bra and ket represent the vectors and operator corresponds to an entry in the enum class Operator. For CSFs, the scalar product is constructed as a linear combination of scalar products of determinants, while for determinants, it is computed using the Slater-Condon rules shown in Table 2.1 directly. This is all under the assumption, that the spin-orbitals from which the determinants are built are orthonormal.



Figure 3.4: UML diagram of the class ScalarProduct and the associated enum Operator. This is a detailed and enlarged view of the layout in Figure 3.1.

#### 3.5.3 Matrix Representation

With the CSFBasis and a ScalarProduct at hand, the matrix representation of an operator is then simply the evaluation of the scalar product for each basis combination. With Hermitian operators, just half of the off diagonal elements need to be calculated. This is encapsulated in the class OperatorRepresentation shown in Listing 3.5.

```
template <typename Basis, typename ScalarProduct>
 1
2
   OperatorRepresentation
3
      (Basis const & basis, Operator const & op, ScalarProduct const & scp)
   : Matrix<typename ScalarProduct::Field> (basis.size(), basis.size())
4
5
   {
6
    for (size_t bra = 0; bra < basis.size(); bra++) {</pre>
7
     for (size_t ket = bra; ket < basis.size(); ket++) {</pre>
      // fill matrix
8
9
      (*this)(bra, ket) = scp(basis[bra], op, basis[ket]);
10
11
      // hermitian matrix
12
      if constexpr (std::is_floating_point<typename ScalarProduct::Field>::value){
13
       (*this)(ket, bra) = (*this)(bra, ket);
      }
14
15
      else {
       (*this)(ket, bra) = std::conj( (*this)(bra, ket) );
16
17
      }
18
     }
19
    }
20
   }
```

Listing 3.5: Constructor of the class OperatorRepresentation. The matrix is filled with repeated ScalarProduct(bra, operator, ket) calls.

This simple approach does not perform caching and reusing of integrals, batching or parallization in any way. And thus leaves a lot of improvements on the table. But for the usecase of a CASCI calculation with a limited active space, this procedure is still more than sufficient. Therefore, no further improvements were attempted.

Similarly, solving the eigenvalue equation with the resulting matrix is also trivial for the relatively small CSF basis. A matrix dimension of a few 1000 is fully diagonalized within seconds and therefore by far not the bottleneck in the CASCI method. For the CI method, the Jacobi-Davidson diagonaliser<sup>[73]</sup> was already implemented in the QOL to partially diagonalize large matrices.

## 3.6 Integral Transformation

In the QOL, the two electron integrals are generated and transformed with the class IntegralTransformation and the methods in IntegralTransformation\_applied. Using the Obara-Saika scheme, integrals are generated in batches and transformed to the contracted angular transformed (CAT) basis. The resulting AO integrals are then individually written into a large matrix of size  $N^3 \times N$ , with N being the CAT basis size. Afterwards, this large AO matrix is transformed to the MO matrix using the coefficients **C**. It is possible to separate this transformation into four consecutive applications of this matrix:

$$\langle ij||kl\rangle = \sum_{\alpha\beta\gamma\delta} \mathbf{C}_{\alpha i} \mathbf{C}_{\beta j} \mathbf{C}_{\gamma k} \mathbf{C}_{\delta l} \langle \alpha\beta||\gamma\delta\rangle .$$
(3.4)

In code, the transformation is performed by using LAPACK dgemm calls and is accompanied by pseudo-reshaping (that is, changing the number of rows and columns without altering the contents of the underlying vector) of the matrix in between, preventing costly copy operations. The whole procedure is outlined in Figure 3.5. This approach is sufficient for



Figure 3.5: Steps in the AO to MO integral transformation  $\langle ij||kl\rangle \rightarrow \langle IJ||KL\rangle$ . The integrals are transformed index-wise by multiplying with the coefficient matrix **C** and reshaping the resulting transposed matrix by changing column and row dimensions while keeping the underlying vector unchanged. The transformation marked with blue is the transformation referenced in section 3.6.3.

small molecules and small basis sets. Due to the scaling of  $N^4$  for the memory requirements, the calculations quickly become infeasible when the CAT basis size reaches into the few hundreds (cf. the second column of Table 3.3, a basis size of 400 needs more than 190GB RAM).

#### 3.6.1 Dealing with a Large Basis

For the case when all integrals are needed for further calculations, a workaround was implemented by Ilyas Türkmen.<sup>[60]</sup> The integrals are written to a file and removed from memory during integral generation as well as integral transformation. This results in a noticeable speed reduction due to the increased latency when comparing RAM to SSD access. Usually, this is not a problem, because the post-HF method for which the integral transformation is performed, takes multiple orders of magnitude longer than the integral transformation itself (CPU cost scales with  $N^5$ ). The reason for this is the scaling of the CPU cost with more than  $N^6$  for these methods, as described in section 2.3.1. When performing CASCI calculations, the transformation itself is the bottleneck. The goal is, therefore, to reduce the memory requirements to keep everything in RAM.

#### 3.6.2 Freezing and Deleting Orbitals

In the case of limiting the correlation to a small active space, the majority of the two electron integrals are not explicitly needed, as many orbitals remain doubly occupied and thus contribute a constant amount. This is similar to the ECP approximation (see 2.4.3), but for molecules and on a per calculation level. The doubly occupied orbitals which are not part of the active space form the *core* orbitals, whereas the orbitals in the active space are the *valence* orbitals. A Slater determinant can be separated into a core and a valence part by an anti-symmetrized product

$$|\Psi\rangle = |\Psi_c \Psi_v\rangle \tag{3.5}$$

The energy expectation value can then be expressed in terms of valence and core wave functions as

$$\langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi_v \Psi_c | \hat{H} | \Psi_c \Psi_v \rangle$$
  
=  $\langle \Psi_v \Psi_c | (\hat{H}_v + \hat{H}_c + \hat{H}_{cv}) | \Psi_c \Psi_v \rangle ,$  (3.6)

with the valence and core Hamiltonians

$$\hat{H}_{v} = \hat{h}_{v} + \sum_{i>j}^{n_{v}} \frac{1}{r_{ij}} = -\frac{1}{2} \sum_{i}^{n_{v}} \Delta_{i} - \sum_{I}^{N} \sum_{i}^{n_{v}} \frac{Z_{I}}{r_{Ii}} + \sum_{i>j}^{n_{v}} \frac{1}{r_{ij}} , \qquad (3.7)$$

$$\hat{H}_{c} = \hat{h}_{c} + \sum_{i>j}^{n_{c}} \frac{1}{r_{ij}} = -\frac{1}{2} \sum_{i}^{n_{c}} \Delta_{i} - \sum_{I}^{N} \sum_{i}^{n_{c}} \frac{Z_{I}}{r_{Ii}} + \sum_{i>j}^{n_{c}} \frac{1}{r_{ij}} , \qquad (3.8)$$

and the mixed terms

$$\hat{H}_{cv} = \sum_{i}^{n_c} \sum_{j}^{n_v} \frac{1}{r_{ij}} .$$
(3.9)

The expectation value of  $\hat{H}_c$  gives the energy offset  $E_{\text{core}}$  which is added at the end together with the nuclear repulsion term,

$$E_{\text{core}} = \langle \Psi_c | \hat{H}_c | \Psi_c \rangle = \sum_{i}^{n_c} \langle i | \hat{h}_c | i \rangle + \sum_{i < j}^{n_c} \langle i j | | i j \rangle .$$
(3.10)

The two-particle integrals described by  $\hat{H}_{cv}$  are effective one-particle valence integrals and are added to the one-particle integrals of  $\hat{h}_v$  during integral transformation,

$$\langle i|\,\hat{o}_1'\,|i\rangle = \langle i|\,\hat{h}_v\,|i\rangle + \sum_{j}^{n_c} \langle ij||ij\rangle \quad i \in \{v\}, j \in \{c\} \,. \tag{3.11}$$

The Slater-Condon rules in Table 2.1 are then evaluated using these altered one-particle integrals.

#### 3.6.3 Reducing Memory Requirements

Freezing and deleting the orbitals outside the active space leaves just the actually used integrals. The transformed integral matrix for these integrals is comparatively tiny but still needs to be calculated from the total AO matrix which has  $N^4$  entries. To actually reduce the memory requirements, the integral transformation needs to be performed during integral generation.

As stated before, the integrals are calculated in batches, which are then contracted and angular transformed. These batches of integrals are then written into the full integral matrix using the eight-fold symmetry. This can result in overwriting the same integral multiple times, even from different batches, as no symmetry check is performed during batch generation. If the transformation is combined into the integral generation, this assignment is translated into an addition, resulting in adding the same integrals multiple times. The correct solution would be to rewrite the whole integral generation such that all symmetries are respected, and the integrals are written to all equivalent positions. However, this approach would require significant time and effort, as it involves rewriting core components of the QOL, which presents its own set of challenges. Therefore, the sane solution was to deal with what is given, and keeping track, whether the integral has already been written to the position  $\langle pq || rs \rangle$  or one of its symmetries. This bookkeeping method should be as space efficient as possible to not offset any memory savings gained by the direct transformation in the first place. Therefore, this ledger needs to be some sort of efficient data structure. The effect of inefficient data structures, such as boolean arrays or small bitsets is shown in Table 3.3 at the end of this section.

#### **Integral Ordering**

For any set of 4 indices p, q, r, s there are at most 24 (= 4!) permutations and thus two electron integrals. With the 8-fold symmetry relations of the two-particle integrals,

$$\langle pq||rs\rangle = \langle rq||ps\rangle = \langle ps||rq\rangle = \langle rs||pq\rangle$$
$$= \langle qp||sr\rangle = \langle sp||qr\rangle = \langle qr||sp\rangle = \langle sr||qp\rangle , \qquad (3.12)$$

these result in – for real orbitals and multiplicative operators – at most three different integral types.<sup>b</sup> Therefore, the ordered set of indices  $\{a, b, c, d\}$ , that is

$$a \le b \le c \le d , \tag{3.13}$$

is mapped to just a triple of integrals. With the basis size N, there are  $\binom{N+3}{4}$  of these ordered sets  $\{a, b, c, d\}$ , which is a large reduction to the original  $N^4$  positions. The challenge is now to find a mapping from the set to the position in the ledger.

The simple solution would be using the container map from the STL-library, with a string or an integer as the key to a triple of bits. The drawback is the scaling accessing cost of  $O(\log n)$ , as map is non-consecutive. A simple vector on the other hand has a constant access of O(1) which translates to multiple orders of magnitude, when the basis size reaches the hundreds, as seen in Table 3.1. The difficulty is now to determine the position p in the vector from an ordered set  $\{a, b, c, d\}$ , with the index range of  $[0, \ldots, N-1]$ , that is

$$p = f^N(a, b, c, d)$$
 . (3.14)

Following from (3.13), each index is bound from the value of its neighbours, which splits the problem into four terms. Each term gives the position in the sub-block defined by the

b Edgecases are if some or all of the indices are the same, reducing the number of types.

Table 3.1: Speed comparisons between different container and storage types. The method to calculate the position index is kept the same. Objective was to naively iterate over all possible integrals and mark them in the ledger. The different storage types show, that bit manipulation on a modern CPU is as fast as byte manipulation. Performed on an Intel<sup>®</sup> Core<sup>TM</sup> i3-8100.

basis size	60	80	100	200	
storage methods	in seconds				
vector <bitset<3>&gt;</bitset<3>	0.14	0.48	1.30	24.09	
vector <int></int>	0.14	0.50	1.35	26.81	
vector <bitset<63>&gt;</bitset<63>	0.14	0.45	1.11	19.49	
vector <bool></bool>	0.14	0.45	1.10	19.09	
<pre>map<string,bitset<3>&gt;</string,bitset<3></pre>	4.12	15.1	45.1	-	
<pre>map<size_t,bitset<3>&gt;</size_t,bitset<3></pre>	3.05	12.6	40.7	-	



 $p_1 = f_1(N) - f_1(N-a)$   $p_2 = f_2(N-a) - f_2(N-b)$   $p_3 = f_3(N-b) - f_3(N-c)$   $p_4 = (d-c)$ 

**Figure 3.6:** Schematic conversion of an ordered set of indices  $\{abcd\}$  to the position p in the list of length  $f_1(N)$ . The indices are from the range  $[0, \ldots, N-1]$ . Here, an explicit example for  $\{1234\}$  and N = 6 is made. From left to right the first, second, third and fourth index is converted by looking at the corresponding sub-block of the previous step. The final position p is the sum of the sub-block positions  $p_i$ .

preceding index, as shown in Figure 3.6. To aid in this, several integer sequences with the form  $\binom{k+(i-1)}{i}$  will be used. First, the already mentioned binomial coefficient<sup>[74]</sup>

$$f_1(k) = \binom{k+3}{4} = \frac{k(k+1)(k+2)(k+3)}{4!}, \qquad (3.15)$$

then the tetrahedral (or triangular pyramidal) numbers<sup>[75]</sup>

$$f_2(k) = \binom{k+2}{3} = \frac{k(k+1)(k+2)}{3!}, \qquad (3.16)$$

the triangular numbers<sup>[76]</sup>

$$f_3(k) = \binom{k+1}{2} = \frac{k(k+1)}{2!}, \qquad (3.17)$$

and for completeness' sake the natural numbers

$$f_4(k) = \binom{k}{1} = \frac{k}{1!} = k$$
 (3.18)

The subscript indicates, which sequence is used to transform the respective index. Figure 3.6 shows that the position  $p_i$  of each index is calculated by taking the length of the (sub-)block and subtracting  $f_i(N-x)$ , with x being the current index. The position is then simply the sum

$$p = f_1(N) - f_1(N-a) + f_2(N-a) - f_2(N-b) + f_3(N-b) - f_3(N-c) + (d-c) .$$
(3.19)

#### Integral Type

With the position calculated, what remains is to determine the type of the integral. This is done by keeping track, which positions need to be swapped during the sorting of the indices. The comparisons and swaps are performed in the fixed order

$$abcd$$
,  $abcd$ ,  $abcd$ , (3.20)

with a corresponding bit for each position. From the resulting bitmap, the three different types can be derived. Special care needs to be taken for the cases where some indices are equal, which warrants three additional comparisons. Details and a list of which integral is mapped to which type can be found in the Appendix A.

Table 3.2: Speed comparisons between vector<br/>bool> and vector<br/>bitset<63>> with optimized methods to evaluate the position index. With bitset<63> it is possible to perform a cheap, coarse check before evaluating the integral type. The objective was to naively iterate over all possible integrals and mark it in the ledger. Performed on an Intel<sup>®</sup> Core<sup>TM</sup> i3-8100.

	time		
basis	vect		
size	bool	bitset<63>	factor
50	0.068	0.077	0.88
100	1.08	1.15	0.94
150	5.50	5.57	0.99
200	17.8	17.1	1.04
400	348	295	1.18

#### The Ledger

The underlying data structure of the ledger was chosen to be vector<br/>bitset<63>>, where 21 ordered sets  $\{a, b, c, d\}$  with 3 types each are stored per entry. This has the advantage, that all but one of the 64 bits are used for information storage, while staying within the efficient 8-byte register length of modern CPUs. Table 3.1 shows a slight speed advantage of vector<br/>bool> compared to vector<bitset<63>>, where the method to calculate the position in the ledger as well as the type of integral was kept the same between methods. But with bitset<63> it is possible to perform a cheap check if all positions are already set, before evaluating the type, as seen in Listing 3.6. This saves the expensive isolation of the individual bit, and this effect increases dramatically with basis size, as seen in Table 3.2.

```
// position of set {p,q,r,s}
 1
   size_t position = _2eIntsBool.position(p,q,r,s);
2
3
   // check if the bitset was already fully set
4
5
   if ( 2eIntsBool.data()[position/21].all()){
6
     continue;
7
   }
8
9
   // determine type of sequence p,q,r,s
10
   size_t type = _2eIntsBool.type(p,q,r,s);
11
12
   // check and set bit
   auto ref = _2eIntsBool.data()[position/21][(position%21)*3 + type];
13
   if (ref) {
14
15
     continue;
   }
16
17
   ref = true;
```

**Listing 3.6:** Sequence to check the integral  $\langle pq | | rs \rangle$  in the integral loop if it was already accounted for.

#### **Final Thoughts**

In this section, the integral transformation currently present in the QOL was investigated and replaced with a more memory efficient version. The achievement is illustrated in Table 3.3, where the second column shows the memory requirements of the old transformation. By transforming an index during integral generation, the integral matrix is drastically reduced (third column). But due to the approach of integral generation, a bookkeeping ledger is needed, which also scales with the basis size with  $N^4$  but much slower. The chosen data structure vector<bitset<63>> is compared to simpler but more memory demanding alternatives, which increasingly outgrow the integral matrix. Finally, the last column shows the combined memory usage for a given basis size, indicating that large molecules are accessible with moderate hardware (e.g., basis of less than 800 for 32 GB RAM).

**Table 3.3:** Memory requirements for the two electron integral matrix and the ledger for which integral has already been written to. The active space MO basis size was set to 7, representing, for example, the f orbitals of a lanthanide. The final memory requirements are listed in the right column.

AO basis	integra	l matrix	ledger wi	th vector<	>	minimal
size	$AO^4$	$AO^3 \times MO$	array <bool,3></bool,3>	bitset<3>	bitset<63>	memory
100	0.7 GB	$0.1~\mathrm{GB}$	0.1 GB	$0.0~\mathrm{GB}$	0.0 GB	0.1 GB
200	$11.9~\mathrm{GB}$	$0.4~\mathrm{GB}$	$1.5~\mathrm{GB}$	$0.5~\mathrm{GB}$	$0.0~\mathrm{GB}$	0.4 GB
300	$60.3~\mathrm{GB}$	$1.4~\mathrm{GB}$	$7.7~\mathrm{GB}$	$2.6~\mathrm{GB}$	$0.1~\mathrm{GB}$	1.5 GB
400	$190.7~\mathrm{GB}$	$3.3~\mathrm{GB}$	$24.2~\mathrm{GB}$	$8.1~\mathrm{GB}$	$0.4~\mathrm{GB}$	3.7 GB
600	$965.6~\mathrm{GB}$	$11.3~\mathrm{GB}$	$121.9~\mathrm{GB}$	$40.6~\mathrm{GB}$	$1.9~\mathrm{GB}$	13.2 GB
800	$3.0 \ { m TB}$	$26.7~\mathrm{GB}$	$384.3~\mathrm{GB}$	$128.1 \ \mathrm{GB}$	$6.1~\mathrm{GB}$	32.8 GB
1000	$7.3~\mathrm{TB}$	$52.2~\mathrm{GB}$	$936.9~\mathrm{GB}$	$312.3~\mathrm{GB}$	$14.9~\mathrm{GB}$	67.0 GB
1500	$36.8 \ \mathrm{TB}$	$176.0~\mathrm{GB}$	4.6 TB	1.5  TB	$75.1~\mathrm{GB}$	251.2 GB
2000	$116.4 \ \mathrm{TB}$	$417.2~\mathrm{GB}$	14.6 TB	$4.9 \ \mathrm{TB}$	$237.2~\mathrm{GB}$	654.5 GB
2500	284.2  TB	$814.9~\mathrm{GB}$	$35.6~\mathrm{TB}$	11.9 TB	$578.8~\mathrm{GB}$	1.4 TB

### 3.7 CASCI with Spin-Orbit Contribution

With the real-valued CASCI machinery established and validated (see section 3.9), the focus shifts to incorporating spin-orbit coupling. The pseudopotential provides the spin-orbit component  $\hat{B}$  of equation (2.135), which can be included in the post-Hartree-Fock step. This approach is advantageous because the spin-orbit coupling typically constitutes a small perturbation to the system. As a result, it is computationally more efficient to add this contribution after the Hartree-Fock calculation rather than performing a fully complexvalued SCF calculation. In the CASCI-SO (CASCI with Spin-Orbit) calculation, the term B is added to the Hamiltonian and necessitates the inclusion of pseudopotential spin-orbit integrals in the scalar product evaluation. Consequently, the scalar product and all subsequent calculations become complex-valued. Specifically, the OperatorRepresentation (see Listing 3.5) is now represented as a complex matrix. The corresponding complex eigenvalue problem can be solved using the LAPACK package, with the notable side effect of the eigenvalues no longer beeing ordered by their real part (as there is no universal ordering of complex numbers). However, due to the Hermiticity of all operators involved, the imaginary parts of the eigenvalues are zero, ensuring that unambiguous ordering based on the real parts remains possible.

The pseudopotential and PP spin-orbit machinery was implemented in the QOL by Jonas Wiebke during his PhD. Only small parts where modified and the reader is referred to ref. [53] for more details.

#### 3.7.1 Storing One-particle Spin-Orbit Integrals

One-particle spin-orbit integrals exhibit numerous symmetries, and the goal is to efficiently store only the unique entries. For a one-particle basis of size n, these integrals are represented as a complex-valued Cayley matrix, which is a  $2n \times 2n$  matrix with the block structure

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{\alpha\alpha} & \mathbf{M}_{\alpha\beta} \\ \mathbf{M}_{\beta\alpha} & \mathbf{M}_{\beta\beta} \end{pmatrix} .$$
(3.21)

The Cayley matrix possesses symmetries necessary for correct time reversal in a relativistic context. These symmetries are expressed as

$$\mathbf{M}_{\beta\beta} = \mathbf{M}_{\alpha\alpha}^* \qquad \mathbf{M}_{\beta\alpha} = -\mathbf{M}_{\alpha\beta}^* , \qquad (3.22)$$

which reduces the matrix to two unique blocks,  $\mathbf{M}_{\alpha\alpha}$  and  $\mathbf{M}_{\alpha\beta}$ . Thus, the Cayley matrix can be written as

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{\alpha\alpha} & \mathbf{M}_{\alpha\beta} \\ -\mathbf{M}_{\alpha\beta}^* & \mathbf{M}_{\alpha\alpha}^* \end{pmatrix} .$$
(3.23)

Since the spin-orbit operators are Hermitian, their matrix representation must be too, giving the additional restrictions

$$\mathbf{M}_{\alpha\alpha} = \mathbf{M}_{\alpha\alpha}^{\dagger} , \qquad \mathbf{M}_{\alpha\beta} = -\mathbf{M}_{\alpha\beta}^{T} . \tag{3.24}$$

\_\_\_\_\_

These constraints further reduce the number of unique entries to  $n^2$ , as the diagonal of the  $\mathbf{M}_{\alpha\beta}$ -block vanishes. In the QOL, the HermitianCayleyMatrix is stored according to the layout illustrated in Figure 3.7. Only the filled elements are stored explicitly, while all other entries are derived using the symmetries described in (3.22) and (3.24).

$$\begin{pmatrix} \mathbf{M}_{\alpha\alpha} & \mathbf{M}_{\alpha\beta} \\ \mathbf{M}_{\beta\alpha} & \mathbf{M}_{\beta\beta} \end{pmatrix} \ \rightarrow$$

Figure 3.7: Example for a layout of the HermitianCayleyMatrix in the QOL. Only the filled elements are stored, all other entries are derived following (3.22) and (3.24).

#### 3.7.2 Integral Evaluation

The pseudopotential integrals are evaluated using ARGOS subroutines from Pitzer *et.* al.<sup>[77–81]</sup> within the COLUMBUS program suite.<sup>[82]</sup> Specifically, the FORTRAN routines **pseud1**, **pseud2**, and **pseud3** are employed to evaluate integrals between Cartesian Gaussian-type orbital (CGTO) basis functions. These routines perform atom-wise PP evaluations for the local component  $\hat{W}_{LJ}$ , the non-local spin-free component  $\hat{A} - \hat{W}_{LJ}$ , and the non-local spin-orbit component  $\hat{B}$  of equation (2.135), respectively. In the QOL, the real spin-free component  $\hat{A}$  is automatically computed during any SCF calculation and incorporated into the one-particle Hamiltonian matrix. The  $\hat{B}$  component of equation (2.135) is calculated by the **pseud3** ARGOS subroutine, which evaluates integrals between CGTO basis functions for both  $\alpha \alpha$  and  $\alpha \beta$  spin pairings. To avoid redundant computation of the  $\hat{A}$  component, the evaluation of  $\hat{B}$  is isolated within a new evaluator class **PseudopotentialS0\_Evaluator**. This class computes the integrals between CGTO basis functions p and q and returns a quaternion representation of the spin-orbit integrals

$$Q_{pq} = \operatorname{Re}B_{\alpha\alpha,pq} + i \cdot \operatorname{Im}B_{\alpha\alpha,pq} + j \cdot \operatorname{Re}B_{\alpha\beta,pq} + k \cdot \operatorname{Im}B_{\alpha\beta,pq} .$$
(3.25)

The quaternion entries are subsequently stored in a complex-valued instance of a HermitianCayleyMatrix. The AO to MO integral transformation need to be performed on the sub-blocks  $\mathbf{M}_{\alpha\alpha}$  and  $\mathbf{M}_{\alpha\beta}$  separatly, from which a transformed HermitianCayleyMatrix is build and stored within the MOIntegrals class.

This matrix is utilized to evaluate the scalar product involving the spin-orbit operator SL. It is essential to distinguish between the four spin pairings  $\alpha\alpha$ ,  $\alpha\beta$ ,  $\beta\alpha$  and  $\beta\beta$ , as they correspond to the distinct blocks of the Cayley matrix defined in equation (3.21).

To unify the PP spin-orbit input parameters with the ones used in MOLPRO, the scaling factor  $\frac{2}{2l+1}$  was added in the QOL procedure.

## 3.8 Wave Function Analysis

#### 3.8.1 Population Analysis

The population analysis implementation was part of a master study project by the author, during which Mulliken<sup>[83–86]</sup>, Löwdin<sup>[87,88]</sup>, and Natural population analysis<sup>[89–91]</sup> methods were implemented. Since then, the implementation has been adapted to work with Configuration-Averaged Hartree-Fock (CAHF) orbitals and was restructured to bring it in line with the style of the other implementations.

Population analysis is a crucial component of quantum chemical programs as it provides detailed insights into the distribution of electrons among atoms and atomic orbitals within a molecule. This information might help in understanding and interpreting the wave function. It allows to make a informed decision when choosing an active space for CAS calculations, or the CAHF. Additionally, it serves for evaluating whether converged CAHF orbitals accurately represent the intended electronic configuration. For example, if the CAHF averages the f-shell of a lanthanide or actinide, the population analysis of the averaged orbitals have to exibit strong f-orbital character.

#### 3.8.2 Overlap

The overlap gives an indication about the alignment of two functions (typically states or orbitals) in the Hilbert space. It is the simplest integral and fundamental to many methods, as well as a simple way to compare different calculations.<sup>[92]</sup> An overlap can be regarded as projection between two (possibly identical) Hilbert spaces. In the case that two different AO basis sets are employed, these two Hilbert spaces of the two functions are not identical. Considering two sets of electronic wave functions  $\{|\Psi_A\rangle\}$  and  $\{|\Psi'_B\rangle\}$ , the only requirement is that the (sub)spaces, which are projected onto each other, contain the same number of electrons (otherwise the overlap vanishes). Everything else might differ, from the wave function expansion over the MOs, the basis sets, to even the geometry. The wave functions describing the (sub)spaces, which are projected onto each other, are referred to as  $|\Psi_I\rangle$  and  $|\Psi'_I\rangle$ . The desired overlap is now

$$S_{IJ} = \langle \Psi_I | \Psi'_J \rangle \quad (3.26)$$

where the dash reminds about the assumed non-orthogonality between the functions and its constituents in general. Expanding the wave functions gives the linear combinations of CSFs,

$$|\Psi_I\rangle = \sum_i d_{Ii} |\Theta_i\rangle , \qquad |\Psi'_J\rangle = \sum_j d'_{Jj} |\Theta'_j\rangle , \qquad (3.27)$$

which in turn consists of linear combinations of determinants

$$|\Theta_i\rangle = \sum_k a_{ik} |\Phi_k\rangle , \qquad |\Theta'_j\rangle = \sum_l a'_{jl} |\Phi'_l\rangle . \qquad (3.28)$$

With this, the overlap is expressed as a long sum of determinantal overlaps  $S_{kl}$ ,

$$S_{IJ} = \sum_{ij} d_{Ii} d'_{Jj} \sum_{kl} a_{ik} a'_{jl} S_{kl} , \qquad (3.29)$$

with

$$S_{kl} = \langle \Phi_k | \Phi_l' \rangle \quad . \tag{3.30}$$

When the determinants are build from the same set of spin orbitals, this reduces to  $S_{kl} = \delta_{kl}$ . But if the sets differ, that is

$$|\Phi_k\rangle = |\phi_{k_1}\phi_{k_2}\cdots\phi_{k_n}\rangle \qquad |\Phi'_l\rangle = |\phi'_{l_1}\phi'_{l_2}\cdots\phi'_{l_n}\rangle \quad , \tag{3.31}$$

the overlap is the determinant of the matrix of all mutual spin orbital overlaps  $^{[16,92]}$ 

$$\langle \Phi_k | \Phi'_l \rangle = \begin{vmatrix} \langle \phi_{k_1} | \phi'_{l_1} \rangle & \cdots & \langle \phi_{k_1} | \phi'_{l_n} \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi_{k_n} | \phi'_{l_1} \rangle & \cdots & \langle \phi_{k_n} | \phi'_{l_n} \rangle \end{vmatrix} .$$
(3.32)

Spin integration eliminates the overlaps between different spins, leaving the spatial orbital integrals  $\langle \varphi_a | \varphi'_b \rangle$ . With their expansions in terms of AO orbitals,

$$|\varphi_a\rangle = \sum_{\mu} C_{a\mu} |\chi_{\mu}\rangle , \qquad |\varphi'_b\rangle = \sum_{\nu} C'_{b\nu} |\chi'_{\nu}\rangle , \qquad (3.33)$$

everything is traced back to integrals of the form

$$S_{\mu\nu} = \langle \chi_{\mu} | \chi_{\nu}' \rangle \quad . \tag{3.34}$$

Instead of writing a dedicated method to compute these integrals, it is simpler to use the methods already implemented. By expanding the AO basis  $\{\chi_{\mu}\}$  by the AO' functions  $\{\chi'_{\nu}\}$  and computing the regular overlap with itself, the sought-after integrals (3.34) are computed, too. The desired matrix  $\mathbf{S}^{AO,AO'}$  can be found in the off-diagonal blocks, as shown in Figure 3.8.



Figure 3.8: Schematic overlap matrix for the combined basis with labeled sub-blocks.

When considering wave function overlaps at the CI or FCI level, the overlap calculation quickly becomes expensive, as millions of matrix determinants (3.32) need to be constructed and evaluated. Plasser *et. al.* showed in  $2016^{[92]}$  that these determinants can be spit into unique factors, which are cached and reused. This allows the evaluation of long expansions, i.e., 50 billion determinants, in a few minutes. For the CASCI wave functions described in chapter 4, the simple approach described above suffices by a long shot.

## 3.9 Performing Calculations

In this section some of the comparisons with the commercial program package MOLPRO<sup>[66]</sup> are reported, validating the correct implementation of CAHF and CASCI. In the last part, the orthogonal array is introduced which might help with choosing suitable parameters for difficult to converge systems.

#### 3.9.1 CAHF Comparisons

The CAHF implementation was validated during my master thesis by comparing with MOLPRO ROHF calculations as well as CAHF literature values. With the new CAHF implementation in MOLPRO, averaged energies were more readily accessible. Table 3.4 shows CAHF and ROHF comparisons for atomic calculations, confirming a correct implementation. More comparisons with different basis sets can be found in the appendix Table B.1 to Table B.4.

**Table 3.4:** Energy differences between MOLPRO and QOL of different atoms and ions for the cc-pVDZ basis. Convergence thresholds are set to  $10^{-7}$ , averaging is performed in the given configuration.

	(	CAHF	ROHF		
Atom	avg.	$\Delta E$ [H]	S	$\Delta E \; [\mathrm{H}]$	
Ν	$p^3$	$7.2 \cdot 10^{-10}$	$\frac{1}{2}$	$1.2 \cdot 10^{-09}$	
$N^+$	$p^2$	$1.4 \cdot 10^{-10}$	0	$7.1 \cdot 10^{-10}$	
$N^{2+}$	$p^1$	$7.9 \cdot 10^{-11}$	$\frac{1}{2}$	$1.5 \cdot 10^{-10}$	
Fe	$d^8$	$3.6 \cdot 10^{-06}$		-	

**Table 3.5:** Absolute differences between MOLPRO and QOL calculations of  $U^{4+}$  with PP ECP60MDF, with energy thresholds in the SCF of  $10^{-7}$ . The active space is the f-shell. All possible states are calculated and the largest absolute difference is reported.

Method	GS energy $[H]$	$\Delta E_{max}$ [H]
CAHF	-470.4791773	$9.0\cdot10^{-8}$
CASCI $S = 0$	-470.4908697	$4.9\cdot10^{-7}$
CASCI $S = 1$	-470.5168573	$3.2\cdot10^{-7}$
CASCI-SO	-470.5492047	$4.7 \cdot 10^{-7}$

#### 3.9.2 CAHF/CASCI-SO Comparisons

With the CAHF implementation validated, the focus shifts towards the CASCI-SO method. The  $U^{4+}$  ion is a simple f<sup>2</sup> systems. For this, MOLPRO and QOL calculations were performed and the results compared. The f-shell was averaged in the CAHF calculation and the orbitals used for the following CASCI(-SO) methods. The results are summarized in Table 3.5, showing deviations between the implementations are at or below the SCF threshold of  $10^{-7}$ , confirming a correct implementation.

#### 3.9.3 The Orthogonal Array

In the context of Self-Consistent Field (SCF) calculations and Configuration-Averaged Hartree-Fock (CAHF) methods, parameters such as damping factors, level shifts, and the length of the Direct Inversion in the Iterative Subspace (DIIS) and more, can significantly influence the stability and efficiency of convergence. While default values for these parameters often suffice for simple systems, more complex electronic structures, such as those encountered in open-shell or strongly correlated systems, frequently require tuning to locate the desired stationary state.

Varying one parameter at a time, while holding others constant, is an inefficient way to explore the multidimensional parameter hypersurface. Consequently, a large number of calculations are required to systematically map the influence of each parameter, making this approach computationally expensive and time-consuming.

The use of an orthogonal array provides a more efficient alternative. An orthogonal array is a structured experimental design that allows for the simultaneous variation of multiple parameters while still enabling the approximate attribution of observed effects to specific parameters. While the findings drawn from this method are not absolute, this approach allows for a rapid overview of the parameter space and subsequent elimination of large, non-optimal regions. This significantly reduces the computational effort required to identify promising parameter combinations, making it a practical tool for optimizing SCF and CAHF calculations in complex electronic systems.

An orthogonal array  $OA_{\lambda}(t, k, v)$  is a  $\lambda v^t \times k$  array with entries from a set of v symbols (typically integers from 0 to v - 1): t is the strength of the array (how many columns are considered simultaneously), k is the number of factors (columns), v is the number of levels (symbols) for each factor, and  $\lambda$  is the index, indicating how many times each t-tuple appears. To make a concreate example, consider a CAHF calculation were the level shifts were already figured out. What remains are the damp factors  $\beta_i$  (three values) as well as the DIIS length, giving four (k) factors  $f_i$ . The number of concrete values these factors can take are set to three (v). Compared to a starting point these might simply take the description low, normal, and high, as the exact values are not crucial. Two columns (t) contain all possible pair combinations once ( $\lambda$ ). This gives the orthogonal array  $OA_1(2, 4, 3)$ shown in Table 3.6. Each column determines the settings for a single calculation  $a_i$ , here  $\lambda v^t = 1 \cdot 3^2 = 9$  different calculations.

**Table 3.6:** Orthogonal array  $OA_1(2, 4, 3)$ . Four (k) factors  $f_i$  are varied and take 3 different values (v, indicated by 1 (low), 2 (normal) or 3 (high)). Two (t) columns contain all possible pair combinations once  $(\lambda)$ .

	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$
$f_1$	1	1	1	2	2	2	3	3	3
$f_2$	1	2	3	1	2	3	1	2	3
$f_3$	1	2	3	2	3	1	3	1	2
$f_4$	1	2	3	3	1	2	2	3	1

The array is evaluated with respect to a measure, a non-binary value, i.e., the total energy or the number of iterations. The sum of these measures for each level (1, 2, 3; low, normal, high) is taken for each factor. For the example  $OA_1(2, 4, 3)$ , the values for  $f_1$  are  $a_1 + a_2 + a_3$ for 1 (low),  $a_4 + a_5 + a_6$  for 2 (normal) and  $a_7 + a_8 + a_9$  for 3 (high). This gives a trend for each factor in relation to the choosen measure.

The orthogonal array proved to be a valuable tool for the application discussed in the following chapter.

## Chapter 4

# Application: Actinide-DOTA Complexes

## 4.1 Introduction

1,4,7,10-Tetraazacyclododecane-N,N',N",N"'-tetraacetic acid (H<sub>4</sub>DOTA) is a versatile and highly effective chelating agentand forms stable complexes with a wide range of metal ions, particularly transition metals and f-block elements. The deprotonated form of this ligand, DOTA<sup>4-</sup>, exhibits exceptional coordination capabilities, functioning as an octadentate ligand.<sup>[93]</sup> This is achieved through the encapsulation of the central metal ion within a square arrangement composed of four nitrogen atoms from the macrocyclic ring and a square arrangement of four oxygen atoms from the carboxylate arms (Figure 4.1). The DOTA ligand can adopt two conformers when forming octacoordinated complexes: square antiprismatic (SAP or M) and twisted square antiprismatic (TSAP or m).<sup>[94]</sup> These conformers arise from the helicity of the macrocyclic ring and the carboxylate arms. Nuclear magnetic resonance (NMR) conformal study of Ln<sup>III</sup>-DOTA complexes have revealed that these conformers exist in a slow equilibrium with each other in solution.<sup>[95,96]</sup> With crystal structures, there are trends with exceptions regarding which conformer is preferred.<sup>[11]</sup> The ninth coordination side between the carboxylates is typically occupied by solvent molecules.

Since its first synthesis in 1976,<sup>[97]</sup> the synthetic route for DOTA has remained largely unchanged. In addition, the ligand offers extensive opportunities for derivatization, enabling the fine-tuning of its properties for specific applications.<sup>[98,99]</sup>

The  $\kappa^8$ -binding DOTA ligand forms highly thermodynamically stable complexes with trivalent lanthanides Ln<sup>III</sup>, which remain inert even under physiological conditions, making them highly suitable for numerous medical applications. The Gd<sup>III</sup> complex of DOTA, commercially known as Dotarem, is a widely used magnetic resonance imaging (MRI)



**Figure 4.1:** (a) DOTA<sup>4–</sup> ligand; (b) structure of the An-DOTA complex; (c) and (d) show both top and a side view of two An-DOTA conformers: (c) An-DOTA conformer with an approximatly square anti prima (SAP) in the coordination, named An(SAP); (d) An-DOTA conformer with a twisted square anti prima (TSAP) in the coordination, named An(TSAP).
contrast agent.<sup>[100]</sup> The DOTA derivatives of several other lanthanides (e.g., Eu and Tm), are also being explored as paramagnetic chemical exchange saturation transfer (ParaCEST) agents for MRI.<sup>[101–104]</sup>

ParaCEST agents function by coordinating water at the open ninth binding site of the complex. The paramagnetic metal center shifts the resonance of the coordinated water (and the polarized protons of DOTA derivate ligands) and its surrounding environment, enabling highly sensitive detection and imaging of biological processes.

In addition to diagnostic applications, DOTA and its derivatives play an important role in radiotherapy. They are employed to chelate radionuclides (<sup>213</sup>Bi, <sup>225</sup>Ac, <sup>227</sup>Th, <sup>177</sup>Lu) for targeted  $\alpha$ -therapy (TAT) and  $\beta$ -therapy.<sup>[105–108]</sup> These applications rely on the strong binding affinity and kinetic inertness of DOTA an its derivatives like DOTA-TATE, <sup>[109]</sup> to deliver therapeutic radionuclides precisely to diseased tissues.

Given their widespread use, the stability and coordination chemistry of trivalent DOTA complexes, especially lanthanide ones, have been extensively studied.<sup>[12,110–113]</sup> However, less attention has been devoted to tetravalent metal-DOTA complexes, particularly those involving early actinides, which exhibit stable tetravalent oxidation states. The preparation of An<sup>IV</sup>-DOTA complexes with octadentate coordination is non-trivial, as these metals often exhibit a tendency to form oligomers in both solid and aqueous solutions, where only the carboxylate groups of DOTA coordinate to the metal center.<sup>[114]</sup> Despite these challenges, recently the thorium(IV) and uranium(IV) DOTA complexes Th( $\kappa^8$ -DOTA)(DMSO) and U( $\kappa^8$ -DOTA)(DMSO) were synthesised under water-free conditions in the SAP conformer.<sup>[115]</sup> Remarkably, these complexes remain stable in aqueous solution once formed.<sup>[116]</sup> Theoretical studies conducted by Kovács in 2022 using density functional theory (DFT) suggest, that the stability of An<sup>IV</sup>-DOTA complexes increases from Th < U < Np < Pu.<sup>[13]</sup> These studies serve as a starting point for further exploration of the trends and patterns governing An<sup>IV</sup>-DOTA complex coordination chemistry and stability.

### 4.2 Computational Details

Wave function single-point calculations were conducted using the Configuration-Averaged Hartree-Fock (CAHF) and Complete Active Space Configuration Interaction (CASCI) methods implemented in the Quantum Objects Library (QOL). The geometries utilized in these calculations were obtained from Møller-Plesset second-order perturbation theory (MP2) optimizations performed with the TURBOMOLE software package.<sup>[117]</sup> For each actinide element considered (Pa, U, Np), both the square antiprismatic (SAP) and twisted square antiprismatic (TSAP) conformers of the DOTA complexes were investigated.

The light atoms (e.g., C, H, N, O) were modeled using the standard correlation-consistent polarized valence double-zeta (cc-pVDZ) and triple-zeta (cc-pVTZ) basis sets,<sup>[118]</sup> obtained from the Basis Set Exchange (BSE) library.<sup>[119–121]</sup> For the actinide elements, the 5f-in-valence (small-core) quasi-relativistic pseudopotentials of the Stuttgart-Cologne group (ECP60MWB) were employed.<sup>[122–124]</sup> These pseudopotentials were paired with the corresponding ANO basis sets of triple-zeta quality, contraction scheme (14s13p10d8f6g)/(6s6p5d4f3g), and a reduced version of double-zeta quality, contraction scheme (14s13p10d8f6g)/(4s4p3d2f).<sup>[123,124]</sup>

The active space for the CAHF calculations was defined as the f-shell of the actinide center. Prior to the CAHF calculations, restricted closed shell Hartree-Fock (RHF) computations were performed on the ionized complex, followed by population analysis to identify the correct orbitals for inclusion in the active space. The convergence of the CAHF calculations was validated by ensuring that the active space orbitals exhibited greater than 90% fcharacter, as confirmed by population analysis. Subsequently, CASCI calculations were carried out to compute all possible electronic states for all spin multiplicities, as well as all spin-orbit states.

The same computational procedure was applied to the corresponding tetravalent actinide ions  $(An^{4+})$  in isolation. Finally, the overlap between the electronic wave functions of the An-DOTA complex and the free actinide ion was computed to assess the degree of electronic similarity.

### 4.3 Results and Discussion

The ground state overlaps between the An<sup>4+</sup> ion and the An-DOTA complexes are summarized in Table 4.1. The degenerate states of the ion are combined to provide an average contribution to the lowest state of the complex. The spin-free states of the An<sup>4+</sup> ion exhibit a close correspondence with the complex, with an overlap greater than 94%. In contrast, the spin-orbit atomic states account for 48-71% of the corresponding ground state of the complex. These spin-orbit contributions also display slight deviations depending on the choice of the basis set. This suggests that the wave function expansion with SO is much more sensitive to the CAHF convergence and the respective orbitals, while maintaining near-identical energy values. This is an expected phenomenon, when small correlation effects (e.g. mediated by SO coupling in LS basis) are present.

**Table 4.1:** Overlaps / Projection of the lowest atomic states with the ground state of the complex, without (sf) and with (so) spin-orbit coupling.

	Basis	Pa	U	Np
$/_{\mathrm{M}An^{4+}} _{\mathrm{M}An(\mathrm{SAP})}$	DZ	0.99	0.94	0.96
$\langle \Psi_{\mathrm{sf},0}   \Psi_{\mathrm{sf},0} \rangle$	ΤZ	0.99	0.94	0.96
$\left< \Psi^{An^{4+}}_{sf,0} \left  \Psi^{An(TSAP)}_{sf,0} \right>$	DZ	0.99	0.94	0.96
	ΤZ	0.99	0.94	0.96
$/_{\mathbf{M}} \operatorname{An}^{4+}  _{\mathbf{M}} \operatorname{An}(\operatorname{SAP}) \setminus$	DZ	0.71	0.64	0.53
$\left\langle \Psi_{\rm so,0} \right  \Psi_{\rm so,0} /$	TZ	0.70	0.64	0.47
$/ \Psi^{An^{4+}}   \Psi^{An(TSAP)} \rangle$	DZ	0.53	0.62	0.48
$\left\langle \Psi_{so,0} \right  \Psi_{so,0}$	ΤZ	0.67	0.56	0.48

Figure 4.2 shows the energy diagram for all states of the  $U^{4+}$  ion and the U(SAP) complex, where the averaged energy from the CAHF method serves as the reference point. When focusing on the U(SAP) complex levels, the most significant spin-orbit splittings are observed for the lowest energy states, while the higher energy states show small or no splitting. Both the spin-free and spin-orbit components reveal relatively moderate ligand field splittings between the U<sup>4+</sup> levels and their U(SAP) complex counterparts, falling in the range of 20-50 kJ/mol. Although there are some outliers such as the the 4<sup>th</sup> and 5<sup>th</sup> degenerate levels from the spin-orbit U<sup>4+</sup> ion (degeneracies 22 and 7), which contributes noticeably to states approximately 100 kJ/mol apart. This phenomenon might be explained



Figure 4.2: States for  $U^{4+}$  and U(SAP) with the TZ basis and with and without spin-orbit coupling. The spin-free states are colour coded for triplet (orange) and singlet (green). Energies are relative to the CAHF energy, indicated in red. Numbers next to state levels indicate the degeneracy. Degeneracies of the complex are left out for clarity. Overlaps between wave functions in the active space are drawn as solid lines for an overlap greater than 0.5 and dotted otherwise.



**Figure 4.3:** Same plot as in Figure 4.2 but for  $Np^{4+}$  and Np(SAP) with the TZ basis. Quadruplet states are orange, doublet states are green.

with the overlapping of multiple near-degenerate levels, in this case corresponding to the LSJ terms  ${}^{3}F_{3}$ ,  ${}^{3}F_{4}$ , and  ${}^{3}H_{4}$ , with a degeneracy of 7, 9, and 13, respectively. In contrast, the Np(SAP) complex (Figure 4.3) show much smaller, consistent ligand field splittings. The  ${}^{4}I$  ground state of Np<sup>4+</sup> (degeneracy 52) can even be approximately traced trough both complexes to the LSJ terms  ${}^{4}I_{\frac{9}{2}}$ ,  ${}^{4}I_{\frac{11}{2}}$ ,  ${}^{4}I_{\frac{13}{2}}$ , and  ${}^{4}I_{\frac{15}{2}}$  (with degeneracies 10, 12, 14, and 16, respectively).

Table 4.2: Energy differences between the lowest states of the An(SAP) and An(TSAP) complexes in kJ/mol.

		$\Delta E( ext{tsap} -  ext{sap})$			
Basis	Method	Pa	U	Np	
DZ	CAHF	6.85	7.89	9.45	
	CASCI	7.07	8.71	14.77	
	CASCI-SO	9.06	9.21	12.11	
ΤZ	CAHF	4.17	5.60	7.24	
	CASCI	4.38	6.82	12.34	
	CASCI-SO	6.19	7.13	9.85	

The energy differences between the SAP and TSAP conformers are detailed in Table 4.2. The individual energy values for each system are provided in the appendix, spanning from Table B.5 to Table B.10. The results show a consistent preference for the SAP conformer over the TSAP conformer, irrespective of the computational method employed or basis used. These findings align with the theoretical predictions reported by Kovács.<sup>[13]</sup> Furthermore, the preference for the SAP conformer exhibits a systematic trend across the studied actinide series, increasing in the order Pa < U < Np.

The 5f-averaged energy from CAHF is an equivalent description of the state-averaged CASSCF wave function with the same active space. This in turn is an approximation to the 5f-in-core PP approach, commonly referred to as the large-core pseudopotentials (LPP).<sup>[125]</sup> By comparing the CAHF energy with the lowest-energy states obtained from CASCI and spin-orbit coupled CASCI (CASCI-SO), it is possible to evaluate whether LPPs can describe the complexes with comparable accuracy. The energy differences relative to the 5f-averaged energy for both the TSAP and SAP conformers are summarized in Table 4.3. The SAP CASCI-SO differences remain approximately 2 kJ/mol lower than those for the TSAP conformer, as indicated in the right column of the table. This consistent

offset suggests that when comparing different actinide complexes, the An-DOTA complexes can be reliably described using LPPs. However, when spin-orbit coupling is neglected, the energy discrepancy between the TSAP and SAP conformers increases to more than 5 kJ/mol for neptunium, underscoring the necessity of incorporating spin-orbit effects for accurate predictions.

			$E_{avg} - E_0$		$\Delta E(\text{SAP - TSAP})$		SAP)	
Basis	Method	conf.	Pa	U	Np	Pa	U	Np
DZ	CASCI	SAP TSAP	20.81 20.59	$121.27 \\ 120.46$	283.55 278.24	0.22	0.81	5.31
	CASCI-SO	SAP TSAP	50.70 48.49	$181.26 \\ 179.95$	363.58 360.92	2.21	1.31	2.66
ΤZ	CASCI	SAP TSAP	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	122.15 120.93	283.90 278.80	0.21	1.22	5.10
	CASCI-SO	SAP TSAP	50.98 48.96	$181.58 \\ 180.05$	363.58 360.98	2.02	1.53	2.60

Table 4.3: Energy differences between the averaged energy and lowest states for the An(SAP) and An(TSAP) complexes in kJ/mol.

### 4.4 Summary

For the investigated  $An^{4+}$  complexes (An = Pa, U, Np), the ligand field splitting induced by the DOTA ligands generally falls within the range of 20–50 kJ/mol, provided no near-degenerate atomic states interfere, as it is with uranium. The ground state of the An-DOTA complexes exhibits a strong atomic character, with more than 94% of the atomic state retained in the spin-free approximation. When spin-orbit coupling is included, the contribution of atomic states remains significant, ranging from 48% to 71%, depending on complex and basis used. A systematic trend in the preference for the SAP conformer was observed across the actinide series, increasing in the order Pa < U < Np. This is in line with the experimental results, as only the SAP conformer of uranium was isolated so far.<sup>[115]</sup>

Furthermore, large-core pseudopotentials (LPPs) can be used for comparative studies between different complexes. When using small-core PPs (SPPs), spin-orbit coupling should be explicitly considered, especially for actinides heavier than uranium.

## Chapter 5

## Conclusion

#### 5.1 Summary

In this thesis, the combination of the Configuration-Averaged Hartree-Fock (CAHF) and the Complete Active Space Configuration Interaction (CASCI) method was presented. The first part of this work focused on integrating these methods into the Quantum Object Library (QOL), while the second part addressed their application to actinide-DOTA complexes.

A unified input method for several QOL methods was developed and based on a single input file. The syntax is able to accommodate the integration of additional methods in the future. The CAHF implementation, initially developed during the author's master thesis, was improved significantly. By standardizing the different SCF methods, the complexity of the code was reduced. A generic interface for key update procedures within the SCF framework was introduced, decoupling the implementation of acceleration techniques from the SCF methods. The Geometric Direct Minimization (GDM) method was implemented and evaluated as a potential acceleration strategy. The implementation did not yield satisfactory improvements over the Direct Inversion of Iterative Subspace (DIIS) method and was ultimately excluded from further use.

The CASCI method was integrated as a new module within the QOL framework, utilizing configuration state functions (CSFs) as basis functions. These CSFs are generated from Freeon functions and matching spin eigenfunctions. By using a newly developed algorithm, only unique Freeon functions for the respective active space are generated.

The integral transformation procedure was optimized to reduce memory requirements. In CASCI calculations, only the molecular orbital (MO) two-electron integrals within the active space are essential. However, these must be generated by transforming all atomic orbital (AO) two-electron integrals. By transforming one integral index during the generation of AO integrals, memory costs were drastically reduced, balancing additional computational expenses and rewriting core parts of the QOL. For an active space of seven orbitals and a basis set of 1000 functions, memory requirements were reduced from 7.3 TB to 67.0 GB. This enables accurate computations with the QOL for medium-sized molecules using high-quality basis sets.

The CASCI implementation was expanded to incorporate pseudopotential spin-orbit integrals. The evaluation of PP integrals, initially implemented by Jonas Wiebke,<sup>[53]</sup> was modified to return the desired SO PP integrals in a complex-valued Hermitian Cayley matrix.

During a master project, several population analysis methods were implemented in the QOL. These methods were improved and adapted for use with CAHF, providing valuable insights during CAHF/CASCI calculations. Additionally, an generic overlap calculation between two CASCI wave functions was implemented, which functions independently of the underlying orbital sets.

The CAHF/CASCI implementation was tested and compared to energies calculated by the program MOLPRO, and no notable discrepancies could be found.

The CAHF/CASCI method was applied to investigate An-DOTA complexes (An = Pa, U, Np) and their two conformers: square antiprismatic (SAP) and twisted square antiprismatic (TSAP). Both double-zeta and triple-zeta basis sets are used, and the actinide atom is described using 5f-in-valence (small-core) quasi-relativistic pseudopotentials (SPP). The ligand field splitting, the conformer stability and the ground state character are investigated. A systematic trend favouring the SAP conformer was observed, increasing from Pa to Np, which is consistent with experimental findings. The CAHF energies approximate those obtained using 5f-in-core (large-core) pseudopotentials (LPPs), suggesting that these LPPs can be used for An-DOTA complexes, at least for comparative studies.

### 5.2 Outlook

The developments presented in this thesis mark only the beginning of transforming the Quantum Object Library from a collection of concepts and preliminary implementations into a fully functioning quantum chemical program suite. While the QOL was never intended to rival commercial software solutions, there are numerous avenues for further improvement and expansion.

The current implementation combines SCF, CI, and wave function analysis methods in one program. However, the QOL already hosts many additional methods, such as Møller-Plesset perturbation theory (MP2), coupled cluster (CC), spin-adapted and spin-complete CC, core-polarization potentials (CPPs), 2-component relativistic methods, and geometry optimizations. Integrating these methods into the unified framework would greatly enhance the versatility and applicability of the QOL.

Further improvements to the CAHF method could focus on increasing computational performance. One promising approach is the development of a dedicated wrapper for the electron interaction matrix, which could significantly reduce computational overhead.

Additionally, the exploration of advanced acceleration techniques, for example the trust region augmented Hessian (TRAH) method,<sup>[126]</sup> could significantly improve convergence rates and computational efficiency.

The integral transformation procedure for active spaces, was significantly improved, but this addresses only part of the underlying challenge. For efficient large scale calculation, the integral transformation procedure should be re-implemented with the conditions of an active space. Additionally, parallelization of this process would yield significant speed improvements. Furthermore, the CI framework could be enhanced to handle larger active spaces, expanding the scope of systems that can be studied with the QOL.

These improvements will enhance the QOL's capabilities and broaden its applicability in quantum chemical research.

In the context of An-DOTA complexes, the CAHF/CASCI method has provided valuable insights into their electronic structure, conformer stability, and ligand field splitting. Future studies could extend this work to other actinide complexes or explore the effects of different ligands and coordination environments. Such studies could provide deeper insights into the chemistry of actinide complexes and inform the design of new materials and radiopharmaceuticals.

# Appendix A

## **Implementational Details**

### A.1 Deriving the Integral Type

The unordered set  $\{a, b, c, d\}$  needs to be categorized as one of three integral types. The sorting of the set is performed in a fixed order of comparisons and corresponding swaps. If the swap is performed, a corresponding bit  $t_1$ ,  $t_2$  and  $t_3$  is fliped:

$$t_1$$
:  $\overrightarrow{abcd}$   $t_2$ :  $\overrightarrow{abcd}$   $t_3$ :  $\overrightarrow{abcd}$ 

With the sorted set  $a \leq b \leq c \leq d$ , additional comparisons are made to determine which indices have the same value:

$$e_1$$
:  $a = b$  or  $c = d$   
 $e_2$ :  $b = c$ 

With these five bits, all possible combinations of 4 integral indices can be sorted following the rules outline in Table A.1. The results of these rules are given in Table A.2.

**Table A.1:** Integral  $\leftrightarrow$  type mapping rules for (a) unique indices  $\{abcd\}$ , (b) more than one equivalent indices and (c) exactly one equivalent index.

Case		conditions	type
	1)	$t_1 = t_2 = t_3$	0
(a)	2)	$t_1 = 1$	1
	3)	remaining	2
	1)	$t_1 = t_2 = t_3$	0
$(\mathbf{b})$ and $(\mathbf{c})$	2)	$e_1 = t_1 = 1$	0
(b) and $(c)$	3)	$e_2 = t_2 = 1$	0
	4)	remaining	1

indices  $\{abcd\}$ .

Pattern	$t_1 t_2 t_3$	$e_1e_2$	type
abcd	000	0 0	0
adcb	111	$0 \ 0$	0
badc	$0 \ 0 \ 0$	$0 \ 0$	0
bcda	111	$0 \ 0$	0
cbad	111	$0 \ 0$	0
cdab	$0 \ 0 \ 0$	$0 \ 0$	0
dabc	111	$0 \ 0$	0
dcba	000	0 0	0
abdc	$1 \ 0 \ 0$	0.0	1
acdb	$1 \ 0 \ 1$	$0 \ 0$	1
bacd	$1 \ 0 \ 0$	$0 \ 0$	1
bdca	$1 \ 0 \ 1$	$0 \ 0$	1
cabd	$1 \ 0 \ 1$	$0 \ 0$	1
cdba	$1 \ 0 \ 0$	$0 \ 0$	1
dbac	$1 \ 0 \ 1$	$0 \ 0$	1
dcab	$1 \ 0 \ 0$	0 0	1
acbd	$0 \ 0 \ 1$	0.0	2
adbc	$0\ 1\ 1$	$0 \ 0$	2
bcad	$0\ 1\ 1$	$0 \ 0$	2
bdac	$0\ 0\ 1$	$0 \ 0$	2
cadb	$0\ 0\ 1$	$0 \ 0$	2
cbda	$0\ 1\ 1$	$0 \ 0$	2
dacb	$0\ 1\ 1$	$0 \ 0$	2
dbca	$0 \ 0 \ 1$	$0 \ 0$	2

(a) Integral  $\leftrightarrow$  type mapping for unique

 Table A.2: Integral mappings following the rules from Table A.1.

(c) Integral  $\leftrightarrow$  type mapping for one equivalent index.

(b) Integral $\leftrightarrow$ type mapping for more than
one equivalent indices.

Pattern	$t_1 t_2 t_3$	$e_1e_2$	type
aaaa	000	11	0
aaab	$1 \ 1 \ 0$	11	0
aaba	$0\ 1\ 0$	$1 \ 1$	0
abaa	$1 \ 0 \ 0$	$1 \ 1$	0
baaa	$0 \ 0 \ 0$	$1 \ 1$	0
abbb	$1 \ 1 \ 0$	11	0
babb	$0\ 1\ 0$	11	0
bbab	$1 \ 0 \ 0$	11	0
bbba	000	11	0
aabb	000	$1 \ 0$	0
abba	$1 \ 0 \ 1$	$1 \ 0$	0
baab	$1 \ 0 \ 1$	$1 \ 0$	0
bbaa	000	$1 \ 0$	0
abab	$0 \ 0 \ 1$	$1 \ 0$	1
baba	$0 \ 0 \ 1$	$1 \ 0$	1

Pattern	$t_1 t_2 t_3$	$e_1e_2$	type
aabc	$1 \ 0 \ 0$	$1 \ 0$	0
aacb	$0 \ 0 \ 0$	$1 \ 0$	0
abca	111	$1 \ 0$	0
acba	$1 \ 0 \ 1$	$1 \ 0$	0
baac	111	$1 \ 0$	0
bcaa	$1 \ 0 \ 0$	$1 \ 0$	0
caab	$1 \ 0 \ 1$	$1 \ 0$	0
cbaa	$0 \ 0 \ 0$	$1 \ 0$	0
abbc	$0 \ 0 \ 0$	$0 \ 1$	0
acbb	$1 \ 1 \ 0$	$0 \ 1$	0
bacb	$0 \ 0 \ 0$	$0 \ 1$	0
bbac	$1 \ 1 \ 0$	$0 \ 1$	0
bbca	$0\ 1\ 0$	$0 \ 1$	0
bcab	$0 \ 0 \ 0$	$0 \ 1$	0
cabb	$0\ 1\ 0$	$0 \ 1$	0
cbba	$0 \ 0 \ 0$	$0 \ 1$	0
abcc	$1 \ 0 \ 0$	$1 \ 0$	0
accb	111	$1 \ 0$	0
bacc	$0 \ 0 \ 0$	$1 \ 0$	0
bcca	$1 \ 0 \ 1$	$1 \ 0$	0
cabc	111	$1 \ 0$	0
cbac	$1 \ 0 \ 1$	$1 \ 0$	0
ccab	$1 \ 0 \ 0$	$1 \ 0$	0
ccba	000	10	0
abac	011	$1 \ 0$	1
acab	$0 \ 0 \ 1$	$1 \ 0$	1
baca	$0\ 1\ 1$	$1 \ 0$	1
caba	$0 \ 0 \ 1$	$1 \ 0$	1
abcb	$1 \ 0 \ 0$	$0 \ 1$	1
babc	$1 \ 0 \ 0$	$0 \ 1$	1
bcba	$1 \ 0 \ 0$	$0 \ 1$	1
cbab	$1 \ 0 \ 0$	$0 \ 1$	1
acbc	$0\ 1\ 1$	$1 \ 0$	1
bcac	$0 \ 0 \ 1$	$1 \ 0$	1
cacb	$0\ 1\ 1$	$1 \ 0$	1
cbca	$0\ 0\ 1$	$1 \ 0$	1

# Appendix B

# **Calculation Data**

## B.1 Comparisons with MOLPRO

**Table B.1:** Comparisons between Molpro (top) and QOL (bottom) energies of the nitrogen atom. Convergence thresholds are set to  $10^{-7}$ , averaging is performed in the *p*-shell.

Basis	ROHF, $S = \frac{1}{2}$	CAHF with $p^3$
3-21G	-53.95989464	-53.99540448
3-21G	-53.95989464	-53.99540448
6-31G	-54.24196829	-54.27624478
6-31G	-54.24196829	-54.27624478
cc-pVDZ	-54.24822015	-54.28250604
cc-pVDZ	-54.24822015	-54.2825060 <mark>5</mark>
cc-pVTZ	-54.25960938	-54.29228516
cc-pVTZ	-54.25960938	-54.29228516
cc-pVQZ	-54.26305021	-54.29531089
cc-pVQZ	-54.2630502 <mark>0</mark>	-54.29531089

**Table B.2:** Comparisons between Molpro (top) and QOL (bottom) energies of the N<sup>+</sup> ion. Convergence thresholds are set to  $10^{-7}$ , averaging is performed in the *p*-shell.

Basis	RHF, $S = 0$	CAHF with $p^2$	CASCI $S=1$
3-21G	-53.47113267	-53.54995976	-53.58970162
3-21G	-53.47113267	-53.5499597 <mark>5</mark>	-53.58970 <mark>398</mark>
6-31G	-53.74678054	-53.82618286	-53.86633463
6-31G	-53.7467805 <mark>5</mark>	-53.82618286	-53.8663346 <mark>6</mark>
cc-pVDZ	-53.75577361	-53.83493619	-53.87501238
cc-pVDZ	-53.75577361	-53.83493619	-53.87501 <mark>305</mark>
cc-pVTZ	-53.76547159	-53.84382219	-53.88406627
cc-pVTZ	-53.76547159	-53.84382219	-53.884066 <mark>05</mark>
cc-pVQZ	-53.76849838	-53.84672915	-53.88702397
cc-pVQZ	-53.76849838	-53.8467291 <mark>6</mark>	-53.88702 <mark>418</mark>

Basis	RHF, $S = \frac{1}{2}$	CAHF with $p^1$
3-21G	-52.50693259	-52.50693259
3-21G	-52.50693259	-52.50693259
6-31G	-52.79131279	-52.79131279
6-31G	-52.79131279	-52.79131279
cc-pVDZ	-52.80154377	-52.80151660
cc-pVDZ	-52.80154377	-52.80151660
cc-pVTZ	-52.81278388	-52.81274560
cc-pVTZ	-52.81278388	-52.81274560
cc-pVQZ	-52.81521713	-52.81516824
cc-pVQZ	-52.81521713	-52.8151682 <mark>1</mark>

**Table B.3:** Comparisons between Molpro (top) and QOL (bottom) energies of the  $N^{2+}$  ion. Convergence thresholds are set to  $10^{-7}$ , averaging is performed in the *p*-shell.

**Table B.4:** Comparisons between Molpro (top) and QOL (bottom) energies of the Fe atom. Convergence thresholds are set to  $10^{-7}$ , averaging is performed in the *d*-shell.

Basis	CAHF with $d^8$	CASCI ${\cal S}=0$	CASCI $S=1$
cc-pVDZ	-1262.125888	-1262.113491	-1262.164672
cc-pVDZ	-1262.1258 <mark>92</mark>	-1262.1134 <mark>89</mark>	-1262.16467 <mark>5</mark>
cc-pVTZ	-1262.128973	-1262.116643	-1262.167646
cc-pVTZ	-1262.12897 <mark>6</mark>	-1262.11664 <mark>0</mark>	-1262.16764 <mark>9</mark>
cc-pVQZ	-1262.130619	-1262.118298	-1262.169250
cc-pVQZ	-1262.1306 <mark>23</mark>	-1262.118297	-1262.16925 <mark>2</mark>

## **B.2** An-DOTA Complexes

#### Ground States

Tables with the ground states and energy differences of the An(SAP) and An(TSAP) conformers for An = Pr, U, Np.

**Table B.5:** Ground states of Protactinium complexes with PP ECP60MWB and basiscc-pVDZ.

Method	Pa(SAP) [H]	Pa(TSAP) [H]	$\Delta E \; [\rm kJ/mol]$
CAHF	-1876.744538	-1876.741930	6.85
CASCI S = $1/2$	-1876.752465	-1876.749773	7.07
CASCI-SO	-1876.763850	-1876.760400	9.06

**Table B.6:** Ground states of Protactinium complexes with PP ECP60MWB and basis cc-pVTZ.

Method	Pa(SAP) [H]	Pa(TSAP) [H]	$\Delta E \; [\rm kJ/mol]$
CAHF	-1877.137162	-1877.135573	4.17
CASCI S = $1/2$	-1877.145114	-1877.143444	4.38
CASCI-SO	-1877.156578	-1877.154221	6.19

Table B.7: Ground states of Uran complexes with PP ECP60MWB and basis cc-pVDZ.

Method	U(SAP) [H]	U(TSAP) [H]	$\Delta E \; [\rm kJ/mol]$
CAHF	-1912.168407	-1912.165401	7.89
CASCI $S = 1$	-1912.214598	-1912.211282	8.71
CASCI $S = 0$	-1912.187621	-1912.184873	7.21
CASCI-SO	-1912.237445	-1912.233939	9.20

Table B.8: Ground states of Uran complexes with PP ECP60MWB and basis cc-pVTZ.

Method	U(SAP) [H]	U(TSAP) [H]	$\Delta E \; [\rm kJ/mol]$
CAHF	-1912.561591	-1912.559459	5.60
CASCI $S = 1$	-1912.608114	-1912.605517	6.82
CASCI $S = 0$	-1912.580901	-1912.578947	5.13
CASCI-SO	-1912.630751	-1912.628036	7.13

**Table B.9:** Ground states of Neptunium complexes with PP ECP60MWB and basis cc-pVDZ.

Method	Np(SAP) [H]	Np(TSAP) [H]	$\Delta E \; [\rm kJ/mol]$
CAHF	-1949.498766	-1949.495165	9.45
CASCI S = $3/2$	-1949.606764	-1949.601140	14.77
CASCI S = $1/2$	-1949.550276	-1949.546595	9.66
CASCI-SO	-1949.637245	-1949.632632	12.11

**Table B.10:** Ground states of Neptunium complexes with PP ECP60MWB and basis cc-pVTZ.

Method	Np(SAP) [H]	Np(TSAP) [H]	$\Delta E \; [\rm kJ/mol]$
CAHF	-1949.893018	-1949.890259	7.25
CASCI S = $3/2$	-1950.001148	-1949.996449	12.34
CASCI S = $1/2$	-1949.944546	-1949.941786	7.25
CASCI-SO	-1950.031498	-1950.027748	9.85

#### Energy diagrams

Energy diagrams for all states of the  $An^{4+}$  ion and the An-DOTA complex.



**Figure B.1:** Same plot as in Figure 4.2 but for  $Pa^{4+}$  and Pr(SAP) with the DZ basis. Doublet states are orange.



**Figure B.2:** Same plot as in Figure 4.2 but for  $Pa^{4+}$  and Pr(SAP) with the TZ basis. Doublet states are orange.



**Figure B.3:** Same plot as in Figure 4.2 but for  $Pa^{4+}$  and Pr(TSAP) with the DZ basis. Doublet states are orange.



**Figure B.4:** Same plot as in Figure 4.2 but for  $Pa^{4+}$  and Pr(TSAP) with the TZ basis. Doublet states are orange.



**Figure B.5:** Same plot as in Figure 4.2 but for  $U^{4+}$  and U(SAP) with the DZ basis. Triplet states are orange and singlet states are green.



**Figure B.6:** Same plot as in Figure 4.2 but for  $U^{4+}$  and U(TSAP) with the DZ basis. Triplet states are orange and singlet states are green.



**Figure B.7:** Same plot as in Figure 4.2 but for  $U^{4+}$  and U(TSAP) with the TZ basis. Triplet states are orange and singlet states are green.



**Figure B.8:** Same plot as in Figure 4.2 but for  $Np^{4+}$  and Np(SAP) with the DZ basis. Quadruplet states are orange, doublet states are green.



**Figure B.9:** Same plot as in Figure 4.2 but for  $Np^{4+}$  and Np(TSAP) with the DZ basis. Quadruplet states are orange, doublet states are green.



**Figure B.10:** Same plot as in Figure 4.2 but for  $Np^{4+}$  and Np(TSAP) with the TZ basis. Quadruplet states are orange, doublet states are green.

## **B.3** Geometries

Table B.11: Cartesian coordinates for Pa(SAP) in Bohr.

Atom	x	y	z
Pa	0.00003555504449	-0.00003108495690	1.09243697991608
Ν	-2.25300517522105	3.35722477006538	-2.23821864251863
Ν	3.35712998954142	2.25310676983332	-2.23828693927608
Ν	2.25300594179397	-3.35708569237636	-2.23840228299503
Ν	-3.35720275466212	-2.25294657848141	-2.23829196825915
0	-3.89020789394725	1.46450648570732	2.31469154784312
Õ	1.46448672293761	3.89036170407407	2.31460149957714
Õ	3.89040839820090	-1.46447614098391	2.31445454611242
ŏ	$-1\ 46430007702934$	-3 89056176700014	2 31426895769829
ŏ	3 67988027823504	7 40837791493969	2 00654406277396
ŏ	7 40811993422350	-3 68042922172198	2.00682515556039
ŏ	-3 68021532381715	-7 40828402255406	2 00654363792950
ŏ	-7 40823451576111	3 67988434086855	2 00673164461945
C	-5 29671050008219	3 22002226003148	1 40887815638954
C	3 22000907367900	5 29683172336184	1 40875848579138
C	5 29673295781311	-3 22000172000104	1 40876778220640
C	-3 22000250101511	-5.226200222260600	1 40862949462004
C	-0.33953054316764	5 036/0/37806631	-3 36050236788214
C	2 02464674321866	3 65310076072550	4 22726572714148
C	5 03625266504052	0.33058615205005	3 36061162037070
C	3 65206273066302	0.03930013293903 2.02451721344020	4 23742750731536
C	0.22045521281062	5.02401721044029	2 26070472682222
C	0.00940001201900	-0.00020029490040	-3.30070473003332
C	-2.02474724103179 5.02622778058500	-3.03293331320833	-4.23734130924713
C	-0.00002770900090 2 65202076100000	-0.33930942933222 2.0247550200145	-3.30034011349203
C	-3.03303270100209	2.02473303808143	-4.23121223930090
C	-0.90090707411004	4.020/14102//00/	-0.00973703820302
C	4.02000402003971	5.96901651026502 4.95067172001474	-0.00985205594909 0.61007226020164
C	0.90097000040001 4.00075050264261	-4.02007170991474	-0.01007220029104
U	-4.82073838304301	-3.98889034310233	-0.00992489987383
П	-2.895/5140954455	0.20300210899114	0.41083538394895
П	-0.392/1//1440000	5.80759791380903	-1.72057752197455
H	5.80760426622838	5.39273396891462	-1.72009431001330
H	6.20350000434502	2.89577907310194	0.41674640397279
H	5.39276997966887	-5.80743637652007	-1.72095275062574
H	2.89581199338987	-0.20304504035805	0.41639214430612
H	-6.20355292153293	-2.89568878278275	0.41674234101493
H	-5.80772335195225	-5.39255726964631	-1.72079134325820
H	0.15904930208887	6.45014228666916	-1.98771533875414
H	-1.12017542301908	6.04269955847980	-4.97175870770758
H	3.28684508902158	5.03250021935493	-5.08257777117568
H	1.53861001256839	2.34459965114068	-5.72105662021880
H	6.04258056956045	1.12023299929335	-4.97184598497774
H	6.44996191762409	-0.15906046632887	-1.98781648728759
H	5.03219435894969	-3.28669184024366	-5.08290212632486
H	2.34426582702754	-1.53840097253013	-5.72109953643018
H	1.12004197276695	-6.04243576255687	-4.97204535716411
H	-0.15909288722503	-6.44999481247162	-1.98796498425608
H	-1.53874065542469	-2.34430715603630	-5.72109511952338
Н	-3.28695572591529	-5.03223673678720	-5.08268024624291
Η	-6.45004899928706	0.15924990320577	-1.98774793895337
Η	-6.04267803270404	-1.11995277671501	-4.97182470092439
Η	-2.34438471334238	1.53867167281698	-5.72099553691602
Η	-5.03227737104200	3.28697730564571	-5.08265135457720

Atom	x	y	z
Pa	-0.00000271897629	-0.00000243052219	1.01150033595744
Ν	-3.31168616867798	-2.26310725099862	-2.43579646548008
Ν	2.26311841909141	-3.31168505211623	-2.43578022181861
Ν	3.31170965445778	2.26311339164183	-2.43578451575140
Ν	-2.26310149259840	3.31168511856797	-2.43578915189683
0	-3.99718661958693	-0.90367108698856	2.39154740528539
0	0.90367555081321	-3.99718277853238	2.39156261292137
Ο	3.99718330525887	0.90368003247471	2.39156257906991
0	-0.90368154001913	3.99718034302198	2.39155058632005
0	-7.49872403270694	-3.15942263557393	2.67488418145622
0	3.15944571678083	-7.49871208109225	2.67490595451929
0	7.49869004406978	3.15948629793161	2.67492999487591
0	-3.15950867057760	7.49867434374955	2.67491383393869
$\mathbf{C}$	-5.58659836675550	-2.58806102096393	1.66266327235791
$\mathbf{C}$	2.58807138927838	-5.58659242099457	1.66268131262035
$\mathbf{C}$	5.58659695403087	2.58807134335381	1.66267777075619
$\mathbf{C}$	-2.58807288535960	5.58658877576770	1.66265889502497
$\mathbf{C}$	-2.02316820489183	-3.69016989699049	-4.42884934218849
$\mathbf{C}$	0.35551037207482	-5.01885975533827	-3.52301195645319
$\mathbf{C}$	3.69018333285105	-2.02317809601509	-4.42883983401430
$\mathbf{C}$	5.01888210568814	0.35551362136046	-3.52300230422462
$\mathbf{C}$	2.02320137650978	3.69017141549176	-4.42884986292233
$\mathbf{C}$	-0.35549785574482	5.01886735675880	-3.52301662012773
$\mathbf{C}$	-3.69015889323882	2.02317235526678	-4.42885241173126
$\mathbf{C}$	-5.01885164655195	-0.35551164219137	-3.52302258067819
$\mathbf{C}$	-4.79989336920557	-3.93443432640602	-0.77658152018573
$\mathbf{C}$	3.93444259985676	-4.79988697333237	-0.77656445867487
$\mathbf{C}$	4.79989949845409	3.93443594911654	-0.77657447328889
$\mathbf{C}$	-3.93442797506457	4.79987994014928	-0.77659463906093
Η	-1.56261645398835	-2.40845249639040	-5.94602866567227
Η	-3.28238036926721	-5.09886136765407	-5.23317188525443
Η	-0.13640899370542	-6.39924499173477	-2.10965368770657
Η	1.15851621811364	-6.06801783322586	-5.09506849286472
Η	2.40846618269636	-1.56263389358875	-5.94602156545878
Н	5.09887527402725	-3.28239423637073	-5.23315521979318
Н	6.39927873267290	-0.13644433280553	-2.10966863465708
Η	6.06804288655216	1.15847005079539	-5.09508190269667
Η	1.56266118417082	2.40845357041617	-5.94603180661153
Н	3.28242284043791	5.09886731077278	-5.23317075728682
H	0.13644262824109	6.39926140838861	-2.10967399695897
H	-1.15846904309568	6.06803141857173	-5.09508766610831
H	-2.40844067728533	1.56262724631374	-5.94603264740478
H	-5.09885498260344	3.28239324292885	-5.23317947792112
H	-6.39923518982084	0.13640471113651	-2.10966138608465
H	-6.06801585080046	-1.15851467095762	-5.09507779420242
H	-6.46451173643271	-4.63818331579377	-1.73517824256353
H	-3.66851563756434	-5.53764969342217	-0.21809313510411
H	4.63816582658109	-6.46451862157168	-1.73515804042200
H	5.53763853164268	-3.66849102913175	-0.21805752945974
H	b.46454861482586	4.63815011695890	-1.73514467286923
H	3.66849822246725	5.53763750827610	-0.21809556434259
H	-4.63814775804401	b.46450752720504	-1.73519301887245
H	-5.53762361663370	3.66847944289408	-0.21809447227595

**Table B.12:** Cartesian coordinates for Pa(TSAP) in Bohr.

Atom	x	y	z
U	-0.00000881355325	-0.00003968163063	1.03904695336188
N	-3.57904111725136	-1.85881415254311	-2.23486198247280
Ν	-1.85883626886372	3.57912376445918	-2.23480742252947
Ν	3.57908863171431	1.85887177504568	-2.23486739584957
N	1.85888712965097	-3.57901282401403	-2.23488913139404
0	-1.84965871366172	-3.68173736428366	2.31250233814945
Õ	-3.68168544366109	1.84961297517248	2.31252572575648
Ō	1.84962407226903	3.68165613658316	2.31256317770581
Õ	3.68166868231970	-1.84967321106614	2.31254744845043
Ō	-6.92200667511276	4.45782129772868	2.04689206366038
Õ	4.45777458436203	6.92194912713168	2.04690106694918
ŏ	6.92199008734307	-4.45768073007422	2.04687571135734
Õ	-4.45777017961408	-6.92199658866615	2.04672672057958
$\tilde{\mathbf{C}}$	-3.76169656442694	-4.88175804932304	1.42937969519076
$\tilde{\mathbf{C}}$	-4.88174028415491	3.76172243019258	1.42945348146421
$\tilde{\mathbf{C}}$	3.76168787645361	4.88170536750822	1.42948406957238
$\tilde{\mathbf{C}}$	4.88169573364066	-3.76162512985966	1.42947914501306
$\tilde{\mathbf{C}}$	-5.04130362066453	0.22585183659046	-3.35674225805791
Č	-3.40493273982612	2.42103986463105	-4.23512570499262
$\tilde{\mathbf{C}}$	0.22582008048438	5.04137644645587	-3.35675896519839
$\tilde{\mathbf{C}}$	2.42099820147758	3.40499062966825	-4.23516865377357
Č	5.04136999801725	-0.22576136485701	-3.35677544900412
Č	3.40500239362484	-2.42092321741230	-4.23520441526593
Ċ	-0.22578885270545	-5.04128380394621	-3.35677398691373
Č	-2.42095959839120	-3.40489506724556	-4.23519347833961
Ċ	-5.21750665974721	-3.41700552612741	-0.59138932020573
Ċ	-3.41703923385923	5.21757314293543	-0.59134087344624
Ċ	5.21752617961250	3.41702334478669	-0.59131547970628
С	3.41699587231538	-5.21746425141597	-0.59131490213162
Н	-6.46903485274471	-2.17454443152162	0.43327970911306
Н	-6.35458725464133	-4.71022572285226	-1.69239851636861
Η	-4.71036629037475	6.35460015380076	-1.69228305977775
Η	-2.17460141941058	6.46919750535767	0.43340047019203
Η	6.35459836814619	4.71027942560919	-1.69228943200649
Η	6.46911244520032	2.17446659199991	0.43334196440054
Η	2.17441136734425	-6.46901827741391	0.43335346633746
Η	4.71024799388418	-6.35456347979002	-1.69226707313608
Η	-6.38907968192945	0.87830440127468	-1.98225250344412
Η	-6.12929746058335	-0.44131606706865	-4.96553312173537
Η	-4.63484085126024	3.83362358564975	-5.07256187641832
Η	-2.16270173043624	1.79444362722320	-5.72349600750065
Η	-0.44137992796540	6.12931177729753	-4.96557236837889
Η	0.87832558168194	6.38913588765404	-1.98227989713568
Η	3.83366136021180	4.63481882566954	-5.07259665620953
Η	1.79429256652734	2.16285701661694	-5.72356562510749
Η	6.12939011147470	0.44143633114282	-4.96553515274960
Η	6.38913838251400	-0.87824573449984	-1.98229371987470
Η	2.16278477901148	-1.79429129104334	-5.72356654629632
Η	4.63491245174575	-3.83350790345527	-5.07264150486969
Η	-0.87827986573571	-6.38900792193955	-1.98224807182322
Η	0.44138266874491	-6.12926175285198	-4.96557032759458
Η	-1.79431964313236	-2.16272629514075	-5.72358717108644
Η	-3.83358011122030	-4.63472262014400	-5.07269130064630

**Table B.13:** Cartesian coordinates for U(SAP) in Bohr.

			,
Atom	x	y	z
U	-0.00000683557345	0.00000854064991	0.96129120649093
Ν	-3.26887711684047	2.30657294149592	-2.42989634320989
Ν	-2.30657855037324	-3.26890345056489	-2.42987504377414
Ν	3.26889887097516	-2.30660730203809	-2.42986731957911
Ν	2.30660186540005	3.26887201386198	-2.42988902729367
0	-2.18774992860376	3.42406928517784	2.38197731382338
Ο	-3.42405993887124	-2.18774804988354	2.38198956872397
0	2.18775269167495	-3.42404738148959	2.38199027669671
Ο	3.42405633544834	2.18775348074710	2.38198124341130
Ο	-5.53409543969788	5.89785559123152	2.71041600631280
0	-5.89787151718829	-5.53407608563883	2.71043088670593
0	5.53408340329894	-5.89785740552560	2.71044947843054
0	5.89785297495121	5.53409365443861	2.71042753098713
$\mathbf{C}$	-4.33437676676773	4.31433122667234	1.68052235369546
С	-4.31433482243209	-4.33437404319309	1.68053706283360
$\mathbf{C}$	4.33438094688447	-4.31432808098687	1.68054550011362
С	4.31432912269123	4.33437890009525	1.68052899687057
Ċ	-4.16717530183089	0.60703739134960	-4.42330612206846
С	-4.57705937732199	-2.08608862257953	-3.51804403391996
С	-0.60704645725329	-4.16720696545502	-4.42328564948174
$\mathbf{C}$	2.08608008785953	-4.57709196333394	-3.51802531917133
С	4.16720166389208	-0.60706256569463	-4.42326757827632
С	4.57710106733781	2.08607117753212	-3.51803459566118
$\mathbf{C}$	0.60705217279592	4.16717204533856	-4.42328663901123
С	-2.08608027877489	4.57706986892998	-3.51805020818693
С	-5.34345203727697	3.11214462671031	-0.75250831700312
С	-3.11215150396586	-5.34347544782157	-0.75248428087316
$\mathbf{C}$	5.34348375442998	-3.11216016580853	-0.75248257395977
С	3.11215546991173	5.34345801057611	-0.75250597416052
Н	-2.81145002561060	0.62522207842560	-5.94631365652314
Η	-5.92996027529843	1.29825753217493	-5.21760173378960
Η	-6.04226271030403	-2.10988160775617	-2.10497221741062
Η	-5.27583757474662	-3.20812572185082	-5.08922490175680
Η	-0.62523278283764	-2.81148462904094	-5.94629572330966
Η	-1.29826939637403	-5.92999296655588	-5.21757668426825
Η	2.10987341441716	-6.04229222640223	-2.10495032118681
Η	3.20811344852880	-5.27587557257226	-5.08920607359284
Η	2.81148071278112	-0.62525205845066	-5.94627818219093
Η	5.92998439418650	-1.29828878510513	-5.21756072204983
Η	6.04229858712519	2.10987448208871	-2.10495633682509
Η	5.27588177246300	3.20809672468862	-5.08922352723479
Η	0.62524039158722	2.81145001170632	-5.94629624478981
Η	1.29827607496568	5.92995455526483	-5.21758241601958
Η	-2.10988275273596	6.04226793349807	-2.10497270892373
Η	-3.20811016231478	5.27584698511943	-5.08923793621387
Η	-6.59033479792786	4.42911239506472	-1.69877716771937
Η	-6.44565996430920	1.49093307769456	-0.18919156575067
Η	-4.42912389285551	-6.59035780729375	-1.69874757169596
Η	-1.49094336894505	-6.44568594396390	-0.18916103430379
Η	6.59036506099216	-4.42913753480512	-1.69874055015384
Η	6.44569202015415	-1.49094569250649	-0.18917232472704
Η	4.42912708292356	6.59034101914243	-1.69876936674649
Н	1.49093753075025	6.44566210923470	-0.18919921055567

**Table B.14:** Cartesian coordinates for U(TSAP) in Bohr.

\_

Atom	x	y	z
Np	0.00000416691208	0.00001147199687	1.00439745829870
Ň	1.00656068304581	-3.89511118622546	-2.22007999789655
Ν	-3.89509038152580	-1.00657610085572	-2.22007304369333
Ν	-1.00656126526170	3.89506678950695	-2.22009494436686
Ν	3.89508220181503	1.00653626848168	-2.22007392644537
0	3.16080945318819	-2.58231025509140	2.31868324677959
Ō	-2.58228850243112	-3.16084136737407	2.31865632324392
Ō	-3.16080148132973	2.58234477872478	2.31865584516061
Ō	2.58228457891193	3.16086427659248	2.31863951279580
Ō	-5.85953980999025	-5.72539990158074	2.09724496592714
Ō	-5.72537063023126	5.85958931514591	2.09722723385899
Ŏ	5.85953353453070	5.72540224413444	2.09719647868434
ŏ	5.72535472752025	-5.85954767291099	2.09726027045409
$\tilde{\mathbf{C}}$	3.89926359590369	-4.72483033386872	1.45958789309128
$\tilde{\mathbf{C}}$	-4.72481336778945	-3 89931239011260	1 45957469608691
$\tilde{c}$	-3.89928146330181	4.72485719882187	1.45956258287442
$\tilde{\mathbf{C}}$	4 72480706014276	3 89931075014422	1 45954254964531
$\tilde{\mathbf{C}}$	-1.34949122767731	-4.86092945369131	-3.34188026423209
č	-3 12348729466136	-2.77665933711100	$-4\ 22107589430782$
$\tilde{\mathbf{C}}$	-4 86090775745145	1 34947192908603	-3 34188395324808
Č	-2 77663683797206	3 12345361827609	-4 22109554690952
$\tilde{\mathbf{C}}$	1 34948631843575	4 86088032988726	-3 34190921265661
$\tilde{c}$	3.12348281974457	2.77660505693776	-4.22108927125176
$\tilde{\mathbf{C}}$	4.86090304736672	-1.34951540478599	-3.34186859567297
$\tilde{\mathbf{C}}$	2.77663710963313	-3.12350186455324	-4.22108170419808
$\tilde{\mathbf{C}}$	2.15621837833651	-5.83063057228123	-0.56180826622912
$\tilde{\mathbf{C}}$	-5 83060464219879	-2 15622695637627	-0 56179795550095
$\tilde{\mathbf{C}}$	-2.15621904916151	5.83060346724790	-0.56184489930958
$\tilde{\mathbf{C}}$	5.83059805309986	2.15620528306311	-0.56181429690091
Ĥ	0.66360727856332	-6.77031943407556	0.46169563684636
Н	3.17008233587619	-7.23072518860257	-1.65219665905230
Н	-7.23072993471111	-3.17006494279317	-1.65217293945666
Н	-6.77028590464833	-0.66362431591987	0.46172708003700
Н	-3.17007526312644	7.23068449157889	-1.65225552168564
Н	-0.66360550297442	6.77029616029214	0.46165234583649
Н	6.77027699760629	0.66361270238864	0.46172704417610
Н	7.23072110501265	3.17003396808833	-1.65220146813576
Н	-2.28593772902946	-6.02802666449543	-1.96635657043954
Н	-0.93952051129039	-6.07163396132998	-4.94881028282922
Н	-4.77974434715964	-3.65805320015631	-5.05096113436389
Н	-2.23806191980250	-1.70967388828250	-5.71422991058204
Н	-6.07161915668814	0.93949146277291	-4.94880660033893
Н	-6.02799844169865	2.28592661923996	-1.96635784514083
Н	-3.65801643036486	4.77971039550293	-5.05099439198961
Н	-1.70965718214433	2.23800773472826	-5.71424343610758
Н	0.93951072462774	6.07157166129907	-4.94884842568994
Н	2.28593348265330	6.02798877050483	-1.96639561258843
Н	2.23805830331615	1.70960952930340	-5.71423685093791
Н	4.77974117458616	3.65799313428078	-5.05097728047002
Н	6.02798568658964	-2.28596464759425	-1.96633174456128
Н	6.07162450330487	-0.93954835663170	-4.94878754184012
Н	1.70965904402049	-2.23806293269503	-5.71423445025420
Η	3.65802056385950	-4.77976053772913	-5.05097342317393

Table B.15: Cartesian coordinates for Np(SAP) in Bohr.

Atom	m	<u>a</u>	~
Atom	it	<i>y</i>	~
Np	-0.00000042798251	-0.00000031320490	0.92827517582335
Ν	3.89471869062859	0.87275484607149	-2.41421644850030
Ν	-0.87275324748438	3.89471916234335	-2.41421605991440
Ν	-3.89471852129071	-0.87275405589443	-2.41421628860620
Ν	0.87275455220038	-3.89471821957703	-2.41421655329307
Ο	3.96336332102983	-0.70928272296623	2.38049308552285
Ο	0.70927762213532	3.96336348027756	2.38049398298129
0	-3.96336405793024	0.70927974713249	2.38049492759204
0	-0.70927850313783	-3.96336470242785	2.38049265259818
Ο	8.03461008350299	0.12971138361152	2.75857884542231
Ο	-0.12970700253059	8.03461280712763	2.75857292137084
Ο	-8.03461249414330	-0.12970867967917	2.75857523340699
Ο	0.12971032390922	-8.03461311970187	2.75857271598141
$\mathbf{C}$	6.06623916625381	0.29984030553581	1.70799512260847
С	-0.29984125731905	6.06624036403788	1.70799297382045
$\mathbf{C}$	-6.06624049410934	-0.29984004002663	1.70799388300893
С	0.29984266104403	-6.06624072981129	1.70799234226909
Ċ	3.24891869091241	2.68289014111358	-4.40777909384250
Ċ	1.52516169125719	4.79165920592023	-3.50310067150593
Ċ	-2.68288443960413	3.24891596787041	-4.40778135206672
Č	-4.79165556380314	1.52516088283191	-3.50310432143294
$\tilde{\mathbf{C}}$	-3.24891797060162	-2.68288878995275	-4.40777942322745
$\tilde{\mathbf{C}}$	-1.52516035374392	-4.79165783050590	-3.50310181879052
$\tilde{\mathbf{C}}$	2.68288651533054	-3.24891604351330	-4.40778155187379
$\tilde{\mathbf{C}}$	4.79165676229322	-1.52515905970482	-3.50310539974299
$\tilde{\mathbf{C}}$	5.87361498878971	1.86473142063790	-0.72042714459681
$\tilde{\mathbf{C}}$	-1.86473426035412	5 87361387225222	-0 72042789122133
$\tilde{\mathbf{C}}$	-5.87361526434934	-1.86473077354502	-0.72042851442728
$\tilde{\mathbf{C}}$	1.86473564961171	-5.87361335755330	-0.72042843720924
Ĥ	2.35722617276998	1.66859823586895	-5.93577573618664
Н	4.94832646367059	3.52569495880448	-5.19346095280380
H	2.48920589647044	5.89520235841299	-2.09022708108599
Н	1 15802987712515	6.06209183995018	-5 07348854631374
Н	-1 66858982477458	2 35722072054882	-5 93577458582356
H	-3.52568764805511	4.94832256414583	-5.19346754508900
Н	-5 89520032785506	2 48920661353871	-2 09023303345713
H	-6.06208609243607	1.15802820021780	-5.07349368053815
Н	-2.35722573432238	-1.66859644861825	-5.93577595840340
H	-4.94832561023792	-3.52569380448347	-5.19346126323638
Н	-2.48920450209332	-5.89520183902866	-2.09022880407657
H	-1.15802843777701	-6.06208952426035	-5.07349044857051
Н	1.66859210067304	-2.35722227111745	-5.93577585078027
Н	3 52569056638646	-4 94832292166135	$-5\ 19346616787279$
H	5 89520290195175	-2 48920446686938	-2.09023491729154
Н	6.06208633454550	-1.15802552860513	-5.07349527616345
H	7.69345383130082	1.90443487283338	-1.65362737640491
Н	5.40989609006549	3.76829556644405	-0.15449395802353
Н	-1.90444126538782	7.69345244154707	-1.65362861859609
Н	-3.76829712648233	5.40989130471871	-0.15449358450165
Н	-7 69345416098636	-1 90443391145501	-1 65362874161146
Н	-5 40989711405397	-3 76829493737258	-0 15449470815394
Н	1 90444278204499	-7 69345154350945	-1 65362989515832
H	3.76829850648554	-5.40989018999587	-0.15449481873320
	2.,00-0000000001		3,101101010020

Table B.16: Cartesian coordinates for  $\operatorname{Np}(\operatorname{TSAP})$  in Bohr.

## References

This thesis uses large language models (LLMs) to enhance the wording and phrasing of some paragraphs. The content was formulated by the author in his own sentences and the LLM was tasked to change the wording without adding additional information. The resulting sentences are further refined by the author to match the tone and flow of the respective chapter. This is in accordance with the legal guidelines of the University of Cologne (see ref. [1]), as well as following the best practices advised in ref. [2].

#### **Tools:**

- ChatGPT v4, OpenAI, chatgpt.com
- DeepSeek v3, DeepSeek, chat.deepseek.com
- Universität zu Köln: Justitiariat ChatGPT, https://verwaltung.uni-koeln.de/ stabsstelle02.1/content/faq/data/chatgpt/index\_ger.html, accessed: 2025-02-19.
- [2] J. Leschke, P. Salden, Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung, Ruhr-Universität Bochum, Universitätsbibliothek, 2023.
- [3] A. Szabo, N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, Dover Publications, 1996.
- [4] J. Griffith, L. Orgel, Q. Rev., Chem. Soc. 1957, 11(4), 381.
- [5] M. Seth, M. Dolg, P. Fulde, P. Schwerdtfeger, J. Am. Chem. Soc. 1995, 117(24), 6597.
- [6] G. R. Choppin, Pure Appl. Chem. 1971, 27(1-2), 23.
- [7] W. Van den Heuvel, S. Calvello, A. Soncini, *Phys. Chem. Chem. Phys.* 2016, 18(23), 15807.

- [8] P. Hallmen, C. Köppl, G. Rauhut, H. Stoll, J. Van Slageren, J. Chem. Phys. 2017, 147(16).
- [9] M. Hanrath, Quantum Objects Library (QOL), 2006-2025.
- [10] G. J. Stasiuk, N. J. Long, Chem. Comm. 2013, 49(27), 2732.
- [11] N. Viola-Villegas, R. P. Doyle, Coord. Chem. Rev. 2009, 253(13-14), 1906.
- [12] A. Kovács, ACS omega **2021**, 6(20), 13321.
- [13] A. Kovács, Symmetry **2022**, 14(11), 2451.
- [14] P. W. Atkins, R. S. Friedman, *Molecular Quantum Mechanics*, Oxford University Press, 2011.
- [15] M. Born, Z. Phys. **1926**, 38(11), 803.
- [16] P.-O. Löwdin, Phys. Rev. 1955, 97, 1474.
- [17] J. E. Mayer, *Phys. Rev.* **1955**, *100*, 1579.
- [18] J. C. Slater, *Phys. Rev.* **1929**, *34*(10), 1293.
- [19] E. Condon, *Phys. Rev.* **1930**, *36*(7), 1121.
- [20] R. McWeeny, Methods of Molecular Quantum Mechanics, Academic Press, 1992.
- [21] C. Roothaan, Rev. Mod. Phys. **1960**, 32(2), 179.
- [22] B. Roos, U. Ryde, 2.41 Molecular Orbital Theory (SCF Methods and Active Space SCF) in Comprehensive Coordination Chemistry II (Eds.: J. A. McCleverty, T. J. Meyer), Pergamon, Oxford, 2003, 519.
- [23] J. Kobus, Comput. Phys. Com. 2013, 184(3), 799.
- [24] J. Kobus, S. Lehtola, arXiv preprint arXiv:2408.03679 2024.
- [25] L. Laaksonen, P. Pyykkö, D. Sundholm, Comput. Phys. Rep. 1986, 4(5), 313.
- [26] F. Jensen, Introduction to Computational Chemistry, John Wiley & Sons, 2007.
- [27] C. C. J. Roothaan, Rev. Mod. Phys. **1951**, 23(2), 69.
- [28] G. Hall, Proc. Roy. Soc. London Ser. A **1951**, 205(1083), 541.

- [29] M. Guest, V. Saunders, Mol. Phys. 1974, 28(3), 819.
- [30] D. Firsht, B. Pickup, Int. J. Quantum Chem. 1977, 12(4), 765.
- [31] M. C. Zerner, Int. J. Quantum Chem. 1989, 35(4), 567.
- [32] R. McWeeny, Mol. Phys. 1974, 28(5), 1273.
- [33] T. Helgaker, P. Jorgensen, J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons, 2014.
- [34] M. Slamet, V. Sahni, *Phys. Rev. A* **1995**, *51*(4), 2815.
- [35] L. Brillouin, J. phys. radium **1932**, 3(9), 373.
- [36] B. S. Fales, T. J. Martínez, J. Chem. Phys. 2020, 152(16).
- [37] R. Pauncz, Spin eigenfunctions: Construction and Use, Springer New York, NY, 1979.
- [38] K. G. Dyall, K. Fægri, Introduction to relativistic quantum chemistry, Oxford University Press, 2007.
- [39] L. L. Foldy, S. A. Wouthuysen, *Phys. Rev.* **1950**, *78*(1), 29.
- [40] M. Douglas, N. M. Kroll, Ann. Phys. **1974**, 82(1), 89.
- [41] B. A. Hess, *Phys. Rev. A* **1985**, *32*(2), 756.
- [42] B. A. Hess, *Phys. Rev. A* **1986**, *33*(6), 3742.
- [43] J. Wood, A. M. Boring, *Phys. Rev. B* **1978**, *18*(6), 2701.
- [44] M. Dolg, X. Cao, *Chem. Rev.* **2012**, *112*(1), 403.
- [45] M. Dolg, Relativistic Effective Core Potentials in Relativistic Electronic Structure Theory, Part: 1, Fundamentals (Ed.: P. Schwerdtfeger), Elsevier, Amsterdam, 2002, 793.
- [46] X. Cao, M. Dolg, Relativistic Pseudopotentials in Relativistic methods for chemists, Springer, 2010, 215.
- [47] H. Hellmann, J. Chem. Phys. **1935**, 3(1), 61.
- [48] J. C. Phillips, L. Kleinman, *Phys. Rev.* **1959**, *116*(2), 287.

- [49] J. D. Weeks, S. A. Rice, J. Chem. Phys. **1968**, 49(6), 2741.
- [50] S. Bubeck, PHD thesis, Universität zu Köln, 2022.
- [51] J. Held, PHD thesis, Universität zu Köln, 2017.
- [52] J. Held, M. Hanrath, M. Dolg, J. Chem. Theory Comput. 2018, 14(12), 6197.
- [53] J. Wiebke, PHD thesis, Universität zu Köln, **2010**.
- [54] J. Ciupka, M. Hanrath, M. Dolg, J. Chem. Phys. 2011, 135(24).
- [55] N. Herrmann, A Rigorously Spin-Adapted and Spin-Complete Coupled Cluster Method for Arbitrary High-Spin Open-Shell States, Dr. Hut, 2022.
- [56] N. Herrmann, M. Hanrath, J. Chem. Phys. **2020**, 153(16).
- [57] N. Herrmann, M. Hanrath, Mol. Phys. **2022**, 120(5), e2005836.
- [58] N. Herrmann, M. Hanrath, J. Chem. Phys. 2022, 156(5).
- [59] S. Segieth, Development and Implementation of a Frozen-Orbital Environment within the Incremental Scheme, Dr. Hut, 2019.
- [60] I. Türkmen, PHD thesis, Universität zu Köln, 2024.
- [61] I. Türkmen, M. Dolg, J. Chem. Theory Comput. 2024, 20(8), 3154.
- [62] M. Hülsen, Optimization of Integral Transformations and Quantum Chemical Calculations on Organometallic Templates, Dr. Hut, 2012.
- [63] A. Engels-Putzka, PHD thesis, Universität zu Köln, 2009.
- [64] ISO, ISO/IEC 14882:2023 Information technology Programming languages C++, seventh Edt., 2024.
- [65] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, *LAPACK Users' Guide*, 3. Edt., Society for Industrial and Applied Mathematics, Philadelphia, PA, **1999**.
- [66] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, et al., MOLPRO, Version 2022.1, a Package of Ab Initio Programs, 2022, see https://www.molpro.net.
- [67] T. Van Voorhis, M. Head-Gordon, Mol. Phys. 2002, 100(11), 1713.
- [68] B. D. Dunietz, T. Van Voorhis, M. Head-Gordon, J. Theor. Comput. Chem. 2002, 1(02), 255.
- [69] P. Pulay, Chem. Phys. Lett. **1980**, 73(2), 393.
- [70] P. Pulay, J. Comp. Chem. 1982, 3(4), 556.
- [71] F. A. Matsen, R. Pauncz, Studies in physical and theoretical chemistry 1986, 44.
- [72] F. Matsen, Freeon Dynamics: A Novel Theory of Atoms and Molecules, Academic Press, 1996, 1–70.
- [73] M. E. Hochstenbach, Y. Notay, *GAMM-Mitteilungen* **2006**, 29(2), 368.
- [74] OEIS Foundation Inc., Entry A000332 in The On-Line Encyclopedia of Integer Sequences, 2025, published electronically at http://oeis.org.
- [75] OEIS Foundation Inc., Tetrahedral numbers (A000292) in The On-Line Encyclopedia of Integer Sequences, 2025, published electronically at http://oeis.org.
- [76] OEIS Foundation Inc., Triangular numbers (A000217) in The On-Line Encyclopedia of Integer Sequences, 2025, published electronically at http://oeis.org.
- [77] R. M. Pitzer, N. W. Winter, Int. J. Quantum Chem. 1991, 40(6), 773.
- [78] R. M. Pitzer, J. Chem. Phys. 1973, 58(7), 3111.
- [79] L. E. McMurchie, E. R. Davidson, J. Comput. Phys. 1981, 44(2), 289.
- [80] M. Dupuis, J. Rys, H. F. King, J. Chem. Phys. 1976, 65(1), 111.
- [81] R. C. Raffenetti, J. Chem. Phys. 1973, 58(10), 4452.
- [82] H. Lischka, R. Shepard, R. M. Pitzer, I. Shavitt, M. Dallos, T. Müller, P. G. Szalay, M. Seth, G. S. Kedziora, S. Yabushita, et al., *Phys. Chem. Chem. Phys.* 2001, 3(5), 664.
- [83] R. S. Mulliken, J. Chem. Phys. **1955**, 23(10), 1833.
- [84] R. S. Mulliken, J. Chem. Phys. 1955, 23(10), 1841.
- [85] R. S. Mulliken, J. Chem. Phys. 1955, 23(12), 2338.
- [86] R. S. Mulliken, J. Chem. Phys. 1955, 23(12), 2343.

- [87] P.-O. Löwdin, J. Chem. Phys. 1950, 18(3), 365.
- [88] P.-O. Löwdin, On the nonorthogonality problem in Advances in Quantum Chemistry, Elsevier, 1970, 185.
- [89] A. E. Reed, R. B. Weinstock, F. Weinhold, J. Chem. Phys. 1985, 83(2), 735.
- [90] P.-O. Löwdin, *Phys. Rev.* **1955**, *97*(6), 1474.
- [91] P.-O. Löwdin, H. Shull, *Phys. Rev.* **1956**, *101*(6), 1730.
- [92] F. Plasser, M. Ruckenbauer, S. Mai, M. Oppel, P. Marquetand, L. González, J. Chem. Theory Comput. 2016, 12(3), 1207.
- [93] J. F. Desreux, *Inorg. Chem.* **1980**, *19*(5), 1319.
- [94] S. Aime, M. Botta, G. Ermondi, *Inorg. Chem.* **1992**, *31*(21), 4291.
- [95] M. Marques, C. Geraldes, A. Sherry, A. E. Merbach, H. Powell, D. Pubanz, S. Aime, M. Botta, J. Alloys Compd. 1995, 225(1-2), 303.
- [96] F. Benetollo, G. Bombieri, L. Calabi, S. Aime, M. Botta, *Inorg. Chem.* 2003, 42(1), 148.
- [97] H. Stetter, W. Frank, Angew. Chem., Int. Ed. Engl. 1976, 15(11), 686.
- [98] S. Knör, A. Modlinger, T. Poethko, M. Schottelius, H.-J. Wester, H. Kessler, *Chem. Eur. J.* 2007, 13(21), 6082.
- [99] C. Wängler, B. Wängler, M. Eisenhut, U. Haberkorn, W. Mier, *Bioorg. Med. Chem.* 2008, 16(5), 2606.
- [100] C. Chapon, L. Lemaire, F. Franconi, L. Marescaux, P. Legras, B. Denizot, J.-J. Le Jeune, Magn. Reson. Med. 2004, 52(4), 932.
- [101] I. Hancu, W. T. Dixon, M. Woods, E. Vinogradov, A. D. Sherry, R. E. Lenkinski, Acta Radiol. 2010, 51(8), 910.
- [102] M. Woods, D. E. Woessner, A. D. Sherry, Chem. Soc. Rev. 2006, 35(6), 500.
- [103] S. Zhang, R. Trokowski, A. D. Sherry, J. Am. Chem. Soc. 2003, 125(50), 15288.
- [104] B. Yoo, M. D. Pagel, J. Am. Chem. Soc. 2006, 128(43), 14032.

- [105] V. Frantellizzi, L. Cosma, G. Brunotti, A. Pani, A. Spanu, S. Nuvoli, F. De Cristofaro,
  L. Civitelli, G. De Vincentis, *Cancer Biother. Radiopharm.* 2020, 35(6), 437.
- [106] A. R. Khabibullin, A. Karolak, M. M. Budzevich, M. L. McLaughlin, D. L. Morse, L. M. Woods, *MedChemComm* 2018, 9(7), 1155.
- [107] L. Królicki, J. Kunikowska, F. Bruchertseifer, R. Kuliński, D. Pawlak, H. Koziara, R. Rola, A. Morgenstern, A. Merlo, *Clin. Nucl. Med.* **2023**, 48(5), 387.
- [108] E. S. Mittra, Am. J. Roentgenol. **2018**, 211(2), 278.
- [109] M. de Jong, W. A. Breeman, B. F. Bernard, A. van Gameren, E. de Bruin, W. H. Bakker, M. E. van der Pluijm, T. J. Visser, H. R. Mäcke, E. P. Krenning, *Eur. J. Nucl. Med.* **1999**, *26*, 693.
- [110] C. A. Chang, Y.-L. Liu, C.-Y. Chen, X.-M. Chou, Inorg. Chem. 2001, 40(14), 3448.
- [111] W. Cacheris, S. Nickle, A. Sherry, *Inorg. Chem.* **1987**, *26*(6), 958.
- [112] E. Tóth, E. Brücher, Inorg. Chim. Acta 1994, 221(1-2), 165.
- [113] P. Jonathan, et al., J. Chem. Soc., Perkin Trans. 2 1991, (1), 87.
- [114] C. Tamain, T. Dumas, C. Hennig, P. Guilbaud, Chem. Eur. J. 2017, 23(28), 6864.
- [115] G. T. Kent, G. Wu, T. W. Hayton, *Inorg. Chem.* **2019**, *58*(13), 8253.
- [116] G. Dovrat, M.-C. Illy, C. Berthon, A. Lerner, M. H. Mintz, E. Maimon, R. Vainer, Y. Ben-Eliyahu, Y. Moiseev, P. Moisy, et al., *Chem. - Eur. J.* **2020**, *26*(15), 3390.
- [117] TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
- [118] T. Dunning Jr, J. Chem. Phys. 1989, 90, 1007.
- [119] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibsom, T. L. Windus, J. Chem. Inf. Model. 2019, 59, 4814.
- [120] D. Feller, J. Comput. Chem. 1996, 17, 1571.
- [121] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, T. L. Windus, J. Chem. Inf. Model. 2007, 47, 1045.
- [122] W. Küchle, M. Dolg, H. Stoll, H. Preuss, J. Chem. Phys. 1994, 100(10), 7535.

- [123] X. Cao, M. Dolg, H. Stoll, J. Chem. Phys. 2003, 118(2), 487.
- [124] X. Cao, M. Dolg, J. Mol. Struct. (Theochem) 2004, 673(1-3), 203.
- [125] A. Moritz, X. Cao, M. Dolg, Theor. Chem. Acc. 2007, 118, 845.
- [126] B. Helmich-Paris, J. Chem. Phys. 2021, 154(16).

## List of Figures

2.1	Schematic illustration of the Fermi hole on the left and the Coulomb cusp	
	on the right. The HF solution is drawn with dashed lines	19
2.2	Branching diagram of genealogical spin coupling for all $S$ of up to eight	
	electrons. The nodes contain the number of routes leading to them. $\ . \ . \ .$	22
3.1	Layout of the <code>CASCI</code> namespace. The classes <code>CI</code> , <code>CASCI</code> and <code>CASCI_SO</code> are	
	displayed as a combined node for clarity.	38
3.2	UML diagram of the classes FreeonFunctionBasis and FreeonFunction.	
	This is a detailed and enlarged view of the layout in Figure 3.1	39
3.3	Example output of Listing 3.4 for the active space $(3,3)$	40
3.4	UML diagram of the class ScalarProduct and the associated enum Operator.	
	This is a detailed and enlarged view of the layout in Figure 3.1	42
3.5	Steps in the AO to MO integral transformation $\langle ij  kl\rangle \rightarrow \langle IJ  KL\rangle$ . The	
	integrals are transformed index-wise by multiplying with the coefficient	
	matrix $\mathbf{C}$ and reshaping the resulting transposed matrix by changing column	
	and row dimensions while keeping the underlying vector unchanged. The	
	transformation marked with blue is the transformation referenced in section	
	3.6.3	44
3.6	Schematic conversion of an ordered set of indices $\{abcd\}$ to the position $p$	
	in the list of length $f_1(N)$ . The indices are from the range $[0, \ldots, N-1]$ .	
	Here, an explicit example for $\{1234\}$ and $N = 6$ is made. From left to	
	right the first, second, third and fourth index is converted by looking at the	
	corresponding sub-block of the previous step. The final position $p$ is the	
	sum of the sub-block positions $p_i$	48
3.7	Example for a layout of the HermitianCayleyMatrix in the QOL. Only	
	the filled elements are stored, all other entries are derived following $(3.22)$	
	and (3.24)	53
3.8	Schematic overlap matrix for the combined basis with labeled sub-blocks.	56

4.1	(a) $DOTA^{4-}$ ligand; (b) structure of the An-DOTA complex; (c) and (d)	
	show both top and a side view of two An-DOTA conformers: (c) An-	
	DOTA conformer with an approximatly square anti prima (SAP) in the	
	coordination, named An(SAP); (d) An-DOTA conformer with a twisted	
	square anti prima (TSAP) in the coordination, named An(TSAP)	62
4.2	States for $U^{4+}$ and $U(SAP)$ with the TZ basis and with and without spin-	
	orbit coupling. The spin-free states are colour coded for triplet (orange)	
	and singlet (green). Energies are relative to the CAHF energy, indicated	
	in red. Numbers next to state levels indicate the degeneracy. Degeneracies	
	of the complex are left out for clarity. Overlaps between wave functions in	
	the active space are drawn as solid lines for an overlap greater than $0.5$ and	
	dotted otherwise.	66
4.3	Same plot as in Figure 4.2 but for $Np^{4+}$ and $Np(SAP)$ with the TZ basis.	
	Quadruplet states are orange, doublet states are green	67
D 1	Correspondent on the Dimension $4.2$ but for $D_{2}^{4+}$ and $D_{2}(CAD)$ with the DZ basis	
Б.1	Same plot as in Figure 4.2 but for Pa <sup>-+</sup> and Pr(SAP) with the DZ basis.	80
DЭ	Doublet states are orange	80
D.2	Daublet states are groups	00
DЭ	Doublet states are orange	80
Б.Э	Same plot as in Figure 4.2 but for $Fa^{-1}$ and $Fi(1SAF)$ with the DZ basis.	<b>Q</b> 1
В 1	Some plot as in Figure 4.2 but for $Pa^{4+}$ and $Pr(TSAP)$ with the TZ basis	01
D.4	Doublet states are orange	<b>Q</b> 1
R 5	Some plot as in Figure 4.2 but for $U^{4+}$ and $U(SAP)$ with the DZ basis	01
D.0	Triplet states are erange and singlet states are green	80
R 6	Some plot as in Figure 4.2 but for $U^{4+}$ and $U(TSAP)$ with the DZ basis	62
D.0	Triplet states are erange and singlet states are green	83
$\mathbf{B7}$	Some plot as in Figure 4.2 but for $U^{4+}$ and $U(TSAP)$ with the TZ basis	00
D.1	Triplet states are orange and singlet states are green	84
B 8	Same plot as in Figure 4.2 but for $Nn^{4+}$ and $Nn(SAP)$ with the DZ basis	01
<b>D</b> .0	Quadruplet states are orange doublet states are green	85
R 9	Same plot as in Figure 4.2 but for $Nn^{4+}$ and $Nn(TSAP)$ with the DZ basis	00
D.5	Quadruplet states are orange doublet states are green	86
B 10	Same plot as in Figure 4.2 but for $Nn^{4+}$ and $Nn(TSAP)$ with the TZ basis	00
<b>D</b> .10	Quadruplet states are orange doublet states are green	87
	guadi upier states are states are green.	01

## List of Tables

2.1	Slater-Condon rules <sup>[18,19]</sup> for the matrix elements between determinants for up to two-particle operators.	8
3.1	Speed comparisons between different container and storage types. The method to calculate the position index is kept the same. Objective was to naively iterate over all possible integrals and mark them in the ledger. The different storage types show, that bit manipulation on a modern CPU is as	
3.2	fast as byte manipulation. Performed on an $Intel^{\textcircled{m}}$ Core <sup>TM</sup> i3-8100 Speed comparisons between vector <bool> and vector<bitset<63>&gt; with optimized methods to evaluate the position index. With bitset&lt;63&gt; it is possible to perform a cheap, coarse check before evaluating the integral</bitset<63></bool>	48
3.3	type. The objective was to naively iterate over all possible integrals and mark it in the ledger. Performed on an $\text{Intel}^{\otimes} \text{Core}^{\text{TM}}$ i3-8100 Memory requirements for the two electron integral matrix and the ledger for which integral has already been written to. The active space MO basis	50
3.4	size was set to 7, representing, for example, the f orbitals of a lanthanide. The final memory requirements are listed in the right column Energy differences between MOLPRO and QOL of different atoms and ions for the cc-pVDZ basis. Convergence thresholds are set to $10^{-7}$ , averaging is	51
3.5	performed in the given configuration. $\dots \dots \dots$	57
3.6	difference is reported	57
4.1	columns contain all possible pair combinations once $(\lambda)$	59 65

4.2	Energy differences between the lowest states of the $An(SAP)$ and $An(TSAP)$ complexes in kJ/mol.	68
4.3	Energy differences between the averaged energy and lowest states for the $An(SAP)$ and $An(TSAP)$ complexes in kJ/mol	69
A.1	Integral $\leftrightarrow$ type mapping rules for (a) unique indices $\{abcd\}$ , (b) more than one equivalent indices and (c) exactly one equivalent index	75
A.2	Integral mappings following the rules from Table A.1	76
B.1	Comparisons between Molpro (top) and QOL (bottom) energies of the ni- trogen atom. Convergence thresholds are set to $10^{-7}$ , averaging is performed in the <i>p</i> -shell.	77
B.2	Comparisons between Molpro (top) and QOL (bottom) energies of the $N^+$ ion. Convergence thresholds are set to $10^{-7}$ , averaging is performed in the	
В.3	<i>p</i> -shell	77
B.4	<i>p</i> -shell	78
B.5	<i>d</i> -shell	78
B.6	Ground states of Protactinium complexes with PP ECP60MWB and basis cc-pVTZ	78 79
B.7	Ground states of Uran complexes with PP ECP60MWB and basis cc-pVDZ.	79
B.8	Ground states of Uran complexes with PP ECP60MWB and basis cc-pVTZ.	79
B.9	Ground states of Neptunium complexes with PP ECP60MWB and basis	
B.10	cc-pVDZ	79
2.10	cc-pVTZ.	79
B.11	Cartesian coordinates for Pa(SAP) in Bohr.	88
B.12	Cartesian coordinates for Pa(TSAP) in Bohr.	89
B.13	Cartesian coordinates for U(SAP) in Bohr.	90
B.14	Cartesian coordinates for U(TSAP) in Bohr.	91
B.15	Cartesian coordinates for Np(SAP) in Bohr.	92
B.16	Cartesian coordinates for Np(TSAP) in Bohr.	93

## Listings

3.1	Example input file for a RHF calculation of $NO^+$ . The indentations are for	
	readability only and not neccessary	35
3.2	Loop syntax.	36
3.3	The class <b>Results</b> . Used to store any values and objects which might be of	
	interest for future calculations	36
3.4	Algorithm for generating the next permutation in the FreeonFunctionBasis.	41
3.5	Constructor of the class OperatorRepresentation. The matrix is filled	
	with repeated ScalarProduct(bra, operator, ket) calls	43
3.6	Sequence to check the integral $\langle pq    rs \rangle$ in the integral loop if it was already	
	accounted for.	50

## 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.

Köln, 24.02.2025

Simon Lekat

Ort, Datum

Unterschrift