



Department of Chemistry

# Explicitly Correlated Orbital Optimized Contracted Pair Correlation Methods

By the Faculty of Mathematics and Natural science  
of the Carl von Ossietzky University Oldenburg  
to obtain the degree and title of a  
Doctor of Natural Sciences  
accepted dissertation of

Christian Lasar

June, 2018

First Referee : Prof. Dr. T. Klüner

Second Referee : Prof. Dr. M. Holthaus

Date of Disputation : 27.04.2018

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Theory</b>	<b>7</b>
2.1	Second Quantization . . . . .	7
2.1.1	Excited Determinants . . . . .	12
2.1.2	Orbital Transformations . . . . .	13
2.1.3	Spin Integration . . . . .	16
2.2	Quasi-Newton Minimization . . . . .	24
2.3	Local Correlation Methods . . . . .	27
<b>3</b>	<b>Contracted Pair Correlation Methods</b>	<b>30</b>
3.1	Ansatz for the Wave Function . . . . .	30
3.2	Integral Evaluation . . . . .	31
3.3	Energy Optimization . . . . .	32
3.4	Improved Scaling . . . . .	35
3.5	Results . . . . .	36
3.6	Summary and Outlook . . . . .	40
<b>4</b>	<b>Orbital Optimized Contracted Pair Correlation Methods</b>	<b>42</b>
4.1	Ansatz for the Wave Function . . . . .	42
4.2	Energy Optimization . . . . .	43
4.3	Results . . . . .	46
4.4	Summary and Outlook . . . . .	50
<b>5</b>	<b>Explicitly Correlated Orbital Optimized Contracted Pair Correlation Methods</b>	<b>51</b>
5.1	Ansatz for the Wave Function . . . . .	51
5.2	Modified Working Equations . . . . .	54
5.3	Explicitly Correlated Two-Electron Integrals . . . . .	57
5.4	Results . . . . .	59
5.5	Summary and Outlook . . . . .	66

<b>6</b>	<b>Summary</b>	<b>67</b>
<b>7</b>	<b>German Summary</b>	<b>69</b>
<b>8</b>	<b>Appendix</b>	<b>71</b>
8.1	Generalised Transformation Rules . . . . .	71
8.2	Additional Contributions for Explicit Correlation . . . . .	72
<b>9</b>	<b>List of Abbreviations</b>	<b>73</b>
<b>10</b>	<b>Bibliography</b>	<b>75</b>
<b>11</b>	<b>Acknowledgments</b>	<b>79</b>
<b>12</b>	<b>Declaration</b>	<b>80</b>

# Chapter 1

## Introduction

Theoretical chemistry summarizes the field of developing and using software to simulate chemical properties and reactions. These properties can for example be reaction energies or barriers and all kinds of spectroscopic data. Simulating these properties can save a lot of measurements and expensive chemicals, when searching for new chemical compounds. Since large compute servers are becoming cheaper and more powerful over the years, these kind of simulations will become even more important in future research and industrial applications.[1] The methods available for simulations can roughly be divided into wave function and density based methods. Density based methods try to predict the electronic density of a molecule by applying a functional of the density. These functionals often contain empirical parameters which are fitted to certain sets of molecules and properties. They work very well for these molecules and properties but might get inaccurate for extrapolations outside the set. Wave function based methods usually avoid empirical parameters and are therefore generally applicable. However, they often need much longer simulation times and demand significantly larger computational resources.

Pair correlation approaches belong to the wave function based methods and represent an efficient and accurate way to describe chemical properties and predict chemical reactions by simulations.[2, 3, 4] However, their disadvantage is that a huge number of variables has to be optimized to achieve this accuracy. The following manuscript introduces contracted pair correlation methods. Contraction strongly reduces the number of variables which have to be stored and manipulated, while maintaining a large fraction of the accuracy of uncontracted pair correlation methods (see chapter 3). The pair correlation wave functions under consideration can be reduced to configuration interaction singles doubles.[2] We will leave out the single excitations at

this point to reduce the complexity.

$$|\Psi_{(\mathbf{A})}\rangle = |\Phi\rangle + \sum_{ijab} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle \quad (1.1)$$

Here,  $|\Phi\rangle$  denotes the Hartree-Fock determinant and  $|\Phi_{ij}^{ab}\rangle$  are doubly excited determinants. This ansatz introduces a huge number of variables, namely the coefficients  $A_{ij}^{ab}$  which need to be optimized and stored. In our new ansatz, we introduce a contraction scheme which drastically reduces the number of coefficients to be stored. The conventional coefficients  $A_{ij}^{ab}$  are approximated by a linear combination of integrals  $A_{ij(\mathbf{B})}^{ab}$ .

$$A_{ij}^{ab} \approx A_{ij(\mathbf{B})}^{ab} = \sum_{klg} B_{ij}^{klg} R_{kl}^{abg} \quad (1.2)$$

The large number of Slater determinants  $\Phi_{ij}^{ab}$  are contracted by a weighted sum over the virtual space  $a, b$ . The contraction coefficients are given by the integral expressions  $R_{kl}^{abg}$ . Only the remaining coefficients  $B_{ij}^{klg}$  need to be optimized to minimize the energy. They depend on the occupied space  $i, j, k, l$  which is in general much smaller than the virtual space.

$$\begin{aligned} |\Psi_{(\mathbf{B})}\rangle &= |\Phi\rangle + \sum_{ijklg} B_{ij}^{klg} |\Xi_{ij}^{klg}\rangle \\ |\Xi_{ij}^{klg}\rangle &= \sum_{ab} R_{kl}^{abg} |\Phi_{ij}^{ab}\rangle \end{aligned} \quad (1.3)$$

Here,  $R_{kl}^{abg}$  are defined as integrals over the operators  $\hat{f}_{12}^g$ , which are Gaussian type two electron functions with different exponents and angular momenta.

$$\begin{aligned} R_{kl}^{abg} &= \langle ab | \hat{f}_{12}^g | kl \rangle \\ \hat{f}_{12}^g &= x_{12}^{i_g} y_{12}^{j_g} z_{12}^{k_g} \exp(-\alpha_g r_{12}^2) \end{aligned} \quad (1.4)$$

The formal scaling of memory requirements can be drastically reduced using this way of contracting the wave function. While uncontracted pair correlation methods formally require  $\frac{1}{2}n_o(n_o + 1)n_v^2$  variables to be stored, we only need to store  $\frac{1}{2}n_o(n_o + 1)n_o^2n_g$  variables ( $n_o$  number of occupied orbitals,  $n_v$  number of virtual orbitals,  $n_g$  number of geminals). Note that the storage requirement of the contracted approach is independent of the size of the virtual space and thereby independent of the choice of the basis set (see figure 1.1). In combination with algorithms which use the short range behavior of correlation effects (see chapter 2.3), this approach will lead to high accuracy methods applicable to large molecular systems.

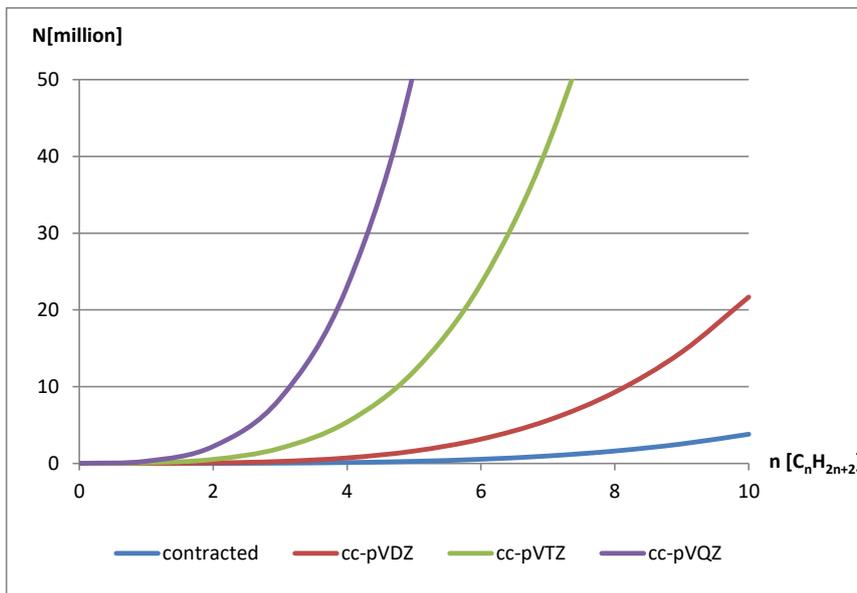


Figure 1.1: Formal number  $N$  of coefficients ( $A_{ij}^{ab}$  or  $B_{ij}^{klg}$ ) which need to be stored for the  $C_nH_{2n+2}$  molecule using different basis sets [5] compared with the contracted ansatz with two s- and p-type Gaussian geminals.[6]

To further improve on the accuracy of molecular properties one can extend pair correlation methods by orbital optimization.[7, 8, 9, 10] This was e.g. shown by Bozkaya *et. al.* for a test set of molecules and reactions using the orbital optimized coupled electron pair approximation OCEPA(0).[7]. With orbital optimized contracted pair correlation methods one can additionally describe single bond breaking with high accuracy. This process is very important in many chemical reactions. Our novel contraction scheme can be simply extended to this kind of wave functions.

$$\begin{aligned}
 |\Psi_{(\mathbf{B}, \boldsymbol{\kappa})}\rangle &= |\Phi\rangle + \sum_{ijklg} B_{ij}^{klg} |\Xi_{ij}^{klg}(\boldsymbol{\kappa})\rangle \\
 |\Xi_{ij}^{klg}\rangle &= \sum_{ab} R_{kl(\boldsymbol{\kappa})}^{abg} |\Phi_{ij}^{ab}(\boldsymbol{\kappa})\rangle
 \end{aligned} \tag{1.5}$$

Here, we introduced the orbital rotation parameter matrix  $\boldsymbol{\kappa}$ . The dependency of the wave function on these parameters can be elegantly formulated in second quantization, as shown in section 2.1.2. Details on the energy optimization and the working equations can be found in chapter 4.

A very general problem of wave function based methods is the slow con-

vergence of chemical properties (e.g. reaction energies) with respect to the size of the one-particle basis set. To avoid usage of large and therefore computationally expensive basis sets, explicitly correlated methods were introduced into electronic structure theory.[11, 12, 13, 14] We also extend our approach to explicitly correlated orbital optimized contracted pair correlation methods (see chapter 5). The contraction approach shown above is the basis for explicitly correlated wave functions. While in current explicitly correlated theory the linear combination of Gaussian geminals is kept fixed[15, 16, 12, 13, 14], we allow its optimization by the dependence of  $B_{ij}^{klg}$  on the geminal index  $g$  and therefore increase the variational freedom of the wave function. The ansatz for the wave function is modified in the following way:

$$\begin{aligned}
|\Xi_{ij}^{klg}\rangle &= \sum_{ab} R_{kl}^{abg} |\Phi_{ij}^{ab}\rangle + \frac{1}{2} \left( \sum_{PQ} R_{kl}^{PQg} |\Phi_{ij}^{PQ}\rangle \right. \\
&\quad \left. - \sum_{pq} R_{kl}^{pqg} |\Phi_{ij}^{pq}\rangle \right)
\end{aligned} \tag{1.6}$$

The indices  $p, q$  correspond to the current basis set while the capital  $P, Q$  introduce a complete one particle basis. This complete one particle basis introduces new integrals over Gaussian geminals into the working equations. The explicit calculation of these new integrals drastically improves on the basis set convergence. Note that if the basis set becomes complete  $p \rightarrow P$ , the last part of the wave function vanishes and we return to the conventional ansatz of contracted pair correlation methods.

All these features of our new ansatz for the wave function will be discussed in detail in the remainder of this manuscript. However, we would first like to present a short introduction of the basic theory for non-expert readers in the following chapter.

# Chapter 2

## Theory

In the following chapters we would like to discuss the mathematical details which are necessary to solve the equations given in the main part. For a more detailed overview the reader is referred to the literature.[17] We present the usefulness of second quantization to derive the working equations and how to optimize the energy using Newton and quasi-Newton methods. The common index conventions are adopted for this purpose. The indices  $ijkl$  label the occupied space, while  $abcd$  labels the virtual space. The labels  $PQRS$  are used for the complete one particle basis and  $pqrs$  label the basis used for actual computations, consisting of the occupied and the virtual space.

### 2.1 Second Quantization

A necessary condition for a wave function of fermions is the Pauli-principle. It states, in mathematical sense, that the wave function has to be anti-symmetric with respect to the permutation of two electronic coordinates. A simple ansatz for this condition is to describe the wave function as a determinant or a linear combination of determinants constructed from the molecular spin-orbitals  $\phi_i(i)$ . The number of orbitals and electrons is abbreviated by  $n = n_{ele}$ .

$$|\Phi\rangle = ||\dots\phi_i\dots\rangle = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \dots & \phi_1(i) & \dots & \phi_1(n) \\ \vdots & & \vdots & & \vdots \\ \phi_i(1) & \dots & \phi_i(i) & \dots & \phi_i(n) \\ \vdots & & \vdots & & \vdots \\ \phi_n(1) & \dots & \phi_n(i) & \dots & \phi_n(n) \end{vmatrix} \quad (2.1)$$

The permutation of two electrons then corresponds to the interchange of two rows or columns in the matrix. A property of determinants is that this results in a change of sign. Second quantization represents a very elegant way to simplify the solution of equations containing integrals over determinants. Since this is the main task when deriving new wave function based electronic structure methods, it is broadly applied in this field of theoretical chemistry and will be summarized here. In contrast to physics where the reference state often is the vacuum state  $|\rangle$ , we normally use the Hartree-Fock determinant  $|\Phi\rangle$  as a reference.

$$|\Phi\rangle \rightarrow \prod_{i=1}^{n_{ele}} a_i^\dagger |\rangle \quad (2.2)$$

The operator  $a_i^\dagger$  is called creation operator. Its action upon the vacuum state will create the state  $i$  in the vacuum state  $|\rangle$  which belongs to the orbital  $\phi_i$  in first quantization. The product of all creation operators of the occupied space will then create all occupied states from the vacuum state. The resulting expression is the second quantization representation of the Hartree-Fock determinant  $|\Phi\rangle$  and is called a state vector. The creation operators have special properties to ensure the anti-symmetry relations to be preserved. The permutation of two electrons can be translated to the interchange of two creation operators.

$$\begin{aligned} |\dots\phi_j\dots\phi_i\dots\rangle &= -|\dots\phi_i\dots\phi_j\dots\rangle \\ \dots a_j^\dagger\dots a_i^\dagger\dots|\rangle &= -\dots a_i^\dagger\dots a_j^\dagger\dots|\rangle \end{aligned} \quad (2.3)$$

This holds due to the anti-commutator property:

$$a_P^\dagger a_Q^\dagger = -a_Q^\dagger a_P^\dagger \quad ; \quad i.e. \quad [a_P^\dagger, a_Q^\dagger]_+ = 0 \quad (2.4)$$

of the creation operators and ensures that the Pauli-principle will always be fulfilled. For the definition and solution of integral expressions we will need some additional operators and properties. The adjoint of the creation operator  $a_P^\dagger$  is called annihilation operator  $a_P$ . If it acts upon a state vector containing the one-particle state  $P$ , this state will be annihilated. In the first quantization picture, the corresponding row and column would be removed from the determinant. If  $a_P$  acts upon a state vector which does not contain the one-particle state  $P$ , the state can not be removed and the result will be zero. Similarly, if  $a_P^\dagger$  acts upon a state vector which already contains the one-particle state  $P$ , the result will be zero. Translated to first quantization this corresponds to a determinant where two rows or columns are equal violating

the Pauli-principle.

$$\begin{aligned} a_P |\Phi\rangle &= 0 & P \notin \{\text{occupied}\} \\ a_P^\dagger |\Phi\rangle &= 0 & P \in \{\text{occupied}\} \end{aligned} \quad (2.5)$$

The same holds for the adjoint of the expressions above, namely:

$$\begin{aligned} 0 &= 0^\dagger = (a_P |\Phi\rangle)^\dagger = \langle \Phi | a_P^\dagger & P \notin \{\text{occupied}\} \\ 0 &= 0^\dagger = (a_P^\dagger |\Phi\rangle)^\dagger = \langle \Phi | a_P & P \in \{\text{occupied}\} \end{aligned} \quad (2.6)$$

We will additionally need the rule resulting from creation and annihilation operators acting upon each other, namely:

$$a_P^\dagger a_Q = \delta_{PQ} - a_Q a_P^\dagger \quad ; \quad i.e. \quad [a_P^\dagger, a_Q]_+ = \delta_{PQ} \quad (2.7)$$

Note that this rule only holds if the one-particle basis is orthonormal. With these three simple operator properties (cf. equations 2.4, 2.5 and 2.7) we can now derive all equations needed in wave function based electronic structure theory. Before we start evaluating integrals over determinants, we have to define the representation of an operator in second quantization.

$$\begin{aligned} \hat{o}_1 &\rightarrow \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle a_P^\dagger a_Q \\ \hat{o}_{12} &\rightarrow \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle a_P^\dagger a_Q^\dagger a_S a_R \end{aligned} \quad (2.8)$$

Note that the indices  $PQRS$  span the complete one-particle basis. A similar expression in first quantization can be defined by using the projection operator of the complete one-particle basis. In any Hilbert space the sum over the dyadic product of all orthonormal basis vectors reduces to a unity operator defined within this Hilbert space. In Dirac notation with the orthonormal one-particle basis this results in:

$$\begin{aligned} \hat{1}_1 &= \sum_P |\phi_P\rangle \langle \phi_P| \\ \hat{o}_1 &= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle |\phi_P\rangle \langle \phi_Q| \\ \hat{o}_{12} &= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle |\phi_P \phi_Q\rangle \langle \phi_R \phi_S| \end{aligned} \quad (2.9)$$

Now, we need to integrate these expressions using a specific wave function  $\Psi$ .

$$\begin{aligned}
\langle \Psi | \hat{o}_1 | \Psi \rangle &= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle \langle \Psi | | \phi_P \rangle \langle \phi_Q | | \Psi \rangle \\
&= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle \langle \Psi | a_P^\dagger a_Q | \Psi \rangle \\
&= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle \gamma_{PQ} \\
\langle \Psi | \hat{o}_{12} | \Psi \rangle &= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle \langle \Psi | | \phi_P \phi_Q \rangle \langle \phi_R \phi_S | | \Psi \rangle \\
&= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle \langle \Psi | a_P^\dagger a_Q^\dagger a_S a_R | \Psi \rangle \\
&= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle \Gamma_{PQRS} \tag{2.10}
\end{aligned}$$

Here, an exact identity holds and  $|\Psi\rangle$  represents a function in first quantization and a state vector in second quantization. The transition between first and second quantization is therefore given by the first order ( $\gamma_{PQ}$ ) and second order ( $\Gamma_{PQRS}$ ) reduced density matrices. To illustrate how the simplification of those density matrices in second quantization works, we will use the Hartree-Fock determinant as a wave function  $\Psi = \Phi$ .

$$\gamma_{PQ}^{HF} = \langle \Phi | a_P^\dagger a_Q | \Phi \rangle \tag{2.11}$$

We may first use equation 2.5 and see that only orbitals from the occupied space can be annihilated to the right hand side, since only they are included in the state vector  $|\Phi\rangle$ . The same holds for the adjoint expression  $\langle \Phi | a_P^\dagger$ :

$$\gamma_{PQ}^{HF} = \delta_{Pi} \delta_{Qj} \langle \Phi | a_i^\dagger a_j | \Phi \rangle \tag{2.12}$$

We may then apply equation 2.7 to move the creation operator to the right hand side:

$$\begin{aligned}
\gamma_{PQ}^{HF} &= \delta_{Pi} \delta_{Qj} \langle \Phi | \delta_{ij} - a_j a_i^\dagger | \Phi \rangle \\
&= \delta_{Pi} \delta_{Qj} (\langle \Phi | \Phi \rangle \delta_{ij} - \langle \Phi | a_j a_i^\dagger | \Phi \rangle) \tag{2.13}
\end{aligned}$$

Since the state  $i$  is already existent in the state vector  $|\Phi\rangle$  the last term is zero (see eqn. 2.5). Using the fact that the state vector is normalized  $\langle \Phi | \Phi \rangle = 1$ ,

we arrive at the final expression.

$$\begin{aligned}
\gamma_{PQ}^{HF} &= \delta_{Pi}\delta_{Qj}\delta_{ij} \\
\langle \Phi | \hat{o}_1 | \Phi \rangle &= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle \delta_{Pi}\delta_{Qj}\delta_{ij} \\
&= \sum_{ij} \langle \phi_i | \hat{o}_1 | \phi_j \rangle \delta_{ij} \\
&= \sum_i \langle \phi_i | \hat{o}_1 | \phi_i \rangle
\end{aligned} \tag{2.14}$$

This corresponds to the expectation value of a one-electron operator with the Hartree-Fock wave function in first quantization. One can derive the expectation value of the two electron operator applying the same simple rules.

$$\Gamma_{PQRS}^{HF} = \langle \Phi | a_P^\dagger a_Q^\dagger a_S a_R | \Phi \rangle \tag{2.15}$$

Note that for the first step, equation 2.4 will additionally be used:

$$\begin{aligned}
\Gamma_{PQRS}^{HF} &= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} \langle \Phi | a_i^\dagger a_j^\dagger a_l a_k | \Phi \rangle \\
&= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\langle \Phi | a_i^\dagger a_k | \Phi \rangle \delta_{jl} - \langle \Phi | a_i^\dagger a_l a_j^\dagger a_k | \Phi \rangle) \\
&= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\langle \Phi | \Phi \rangle \delta_{jl}\delta_{ik} - \langle \Phi | a_k a_i^\dagger | \Phi \rangle \delta_{jl} - \langle \Phi | a_i^\dagger a_l a_j^\dagger a_k | \Phi \rangle) \\
&= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\delta_{jl}\delta_{ik} - \langle \Phi | a_i^\dagger a_l a_j^\dagger a_k | \Phi \rangle) \\
&= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\delta_{jl}\delta_{ik} - \langle \Phi | a_i^\dagger a_l | \Phi \rangle \delta_{jk} + \langle \Phi | a_i^\dagger a_l a_k a_j^\dagger | \Phi \rangle) \\
&= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\delta_{jl}\delta_{ik} - \langle \Phi | \Phi \rangle \delta_{jk}\delta_{il} + \langle \Phi | a_l a_i^\dagger | \Phi \rangle \delta_{jk}) \\
&= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\delta_{jl}\delta_{ik} - \delta_{jk}\delta_{il})
\end{aligned} \tag{2.16}$$

We may use this simplified expression for the density matrix to calculate the expectation value of a two electron operator with the Hartree-Fock wave function:

$$\begin{aligned}
\Gamma_{PQRS}^{HF} &= \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\delta_{jl}\delta_{ik} - \delta_{jk}\delta_{il}) \\
\langle \Phi | \hat{o}_{12} | \Phi \rangle &= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle \delta_{Pi}\delta_{Qj}\delta_{Rk}\delta_{Sl} (\delta_{jl}\delta_{ik} - \delta_{jk}\delta_{il}) \\
&= \sum_{ijkl} \langle \phi_i \phi_j | \hat{o}_{12} | \phi_k \phi_l \rangle (\delta_{jl}\delta_{ik} - \delta_{jk}\delta_{il}) \\
&= \sum_{ij} \langle \phi_i \phi_j | \hat{o}_{12} | \phi_i \phi_j \rangle - \langle \phi_i \phi_j | \hat{o}_{12} | \phi_j \phi_i \rangle
\end{aligned} \tag{2.17}$$

This procedure also works for any linear combination of determinants in the reduced densities. Even for this general ansatz, the order of simplifications

can be used as defined above, and the aim is to get all creation operators to the right hand side and all annihilation operators to the left hand side. Eventually, only overlap elements and zero contributions remain, and the integral expressions can be simplified.

### 2.1.1 Excited Determinants

This work extends the applicability of pair correlation methods. The most primitive ansatz for a pair correlation wave function is named configuration interaction singles doubles (CISD):

$$|\Psi^{CISD}\rangle = |\Phi\rangle + \sum_{ia} A_i^a |\Phi_i^a\rangle + \sum_{ijab} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle \quad (2.18)$$

The determinants  $\Phi_i^a$  and  $\Phi_{ij}^{ab}$  are called singly and doubly excited determinants. This name results from the idea that an occupied orbital  $\phi_i$  with low energy is replaced by a virtual orbital  $\phi_a$  with higher energy in these determinants. Note that this does not necessarily mean that these determinants actually describe an excited state, since a faithful description of excited states will again be a linear combination of determinants. In other words: One should not over interpret the meaning of virtual orbitals. The second quantization representation of excited determinants is given by:

$$\begin{aligned} |\Phi_i^a\rangle &= a_a^\dagger a_i |\Phi\rangle \\ |\Phi_{ij}^{ab}\rangle &= a_b^\dagger a_j a_a^\dagger a_i |\Phi\rangle = a_a^\dagger a_b^\dagger a_j a_i |\Phi\rangle \end{aligned} \quad (2.19)$$

The simple interpretation of the equations above is that one first annihilates corresponding occupied orbitals in the Hartree-Fock determinant and afterwards replaces them by virtual orbitals. As mentioned above, this does not change the rules used for the derivation of equations in second quantization. The expressions just become more lengthy. The reduced density matrices for a configuration interaction doubles wave function are for example given by:

$$\begin{aligned} \gamma_{PQ} &= \langle \Psi^{CID} | a_P^\dagger a_Q | \Psi^{CID} \rangle \\ &= \langle \Phi | a_P^\dagger a_Q | \Phi \rangle + 2 \sum_{ijab} A_{ij}^{ab} \langle \Phi | a_P^\dagger a_Q a_a^\dagger a_b^\dagger a_j a_i | \Phi \rangle \\ &\quad + \sum_{ijab} \sum_{klcd} A_{ij}^{ab} A_{kl}^{cd} \langle \Phi | a_i^\dagger a_j^\dagger a_b a_a a_P^\dagger a_Q a_c^\dagger a_d^\dagger a_l a_k | \Phi \rangle \\ \Gamma_{PQRS} &= \langle \Psi | a_P^\dagger a_Q^\dagger a_S a_R | \Psi \rangle \\ &= \langle \Phi | a_P^\dagger a_Q^\dagger a_S a_R | \Phi \rangle + 2 \sum_{ijab} A_{ij}^{ab} \langle \Phi | a_P^\dagger a_Q^\dagger a_S a_R a_a^\dagger a_b^\dagger a_j a_i | \Phi \rangle \\ &\quad + \sum_{ijab} \sum_{klcd} A_{ij}^{ab} A_{kl}^{cd} \langle \Phi | a_i^\dagger a_j^\dagger a_b a_a a_P^\dagger a_Q^\dagger a_S a_R a_c^\dagger a_d^\dagger a_l a_k | \Phi \rangle \end{aligned} \quad (2.20)$$

We left out the single excitations, since they are approximately described by orbital optimization as shown below.

## 2.1.2 Orbital Transformations

To describe the dependency of the wave function on orbital transformations, we may again start in first quantization. To maintain orthonormality between the orbitals all transformations have to be unitary. This can be ensured by describing the unitary matrix  $\mathbf{U}$  with the matrix exponential  $\exp(-\boldsymbol{\kappa})$  if the matrix  $\boldsymbol{\kappa}$  is anti-symmetric. The dependency of an orbital on  $\boldsymbol{\kappa}$  can be described as [18]:

$$|\phi_{i(\boldsymbol{\kappa})}\rangle = \sum_P |\phi_P\rangle U_{Pi(\boldsymbol{\kappa})} = \sum_P |\phi_P\rangle [\exp(-\boldsymbol{\kappa})]_{Pi} \quad (2.21)$$

We may describe a unitary transformed creation operator by:

$$|\phi_{i(\boldsymbol{\kappa})}\rangle = \sum_P a_P^\dagger | \rangle U_{Pi(\boldsymbol{\kappa})} \rightarrow a_{i(\boldsymbol{\kappa})}^\dagger = \sum_P a_P^\dagger U_{Pi(\boldsymbol{\kappa})} \quad (2.22)$$

An orbital product or a Slater determinant can be rewritten as:

$$\begin{aligned} ||\phi_{1(\boldsymbol{\kappa})} \dots \phi_{i(\boldsymbol{\kappa})} \dots \phi_{n(\boldsymbol{\kappa})} | \rangle &= \left( \sum_P U_{P1(\boldsymbol{\kappa})} \right) \dots \left( \sum_Q U_{Qi(\boldsymbol{\kappa})} \right) \\ &\dots \left( \sum_R U_{Rn(\boldsymbol{\kappa})} \right) ||\phi_P \dots \phi_Q \dots \phi_R | \rangle \end{aligned} \quad (2.23)$$

Consequently, the corresponding state vector is defined as:

$$\begin{aligned} a_{1(\boldsymbol{\kappa})}^\dagger \dots a_{i(\boldsymbol{\kappa})}^\dagger \dots a_{n(\boldsymbol{\kappa})}^\dagger | \rangle &= \left( \sum_P U_{P1(\boldsymbol{\kappa})} \right) \dots \left( \sum_Q U_{Qi(\boldsymbol{\kappa})} \right) \\ &\dots \left( \sum_R U_{Rn(\boldsymbol{\kappa})} \right) a_P^\dagger \dots a_Q^\dagger \dots a_R^\dagger | \rangle \end{aligned} \quad (2.24)$$

This can also be generalized to a linear combination of determinants. In second quantization, the expression above can be further simplified by a redefinition of the creation operators.

$$\begin{aligned} a_{P(\boldsymbol{\kappa})}^\dagger &= \exp(-\hat{\kappa}) a_P^\dagger \exp(\hat{\kappa}) \\ \hat{\kappa} &= \sum_{PQ} \kappa_{PQ} a_P^\dagger a_Q \end{aligned} \quad (2.25)$$

Here,  $\kappa_{PQ}$  are the matrix elements of the anti-symmetric matrix  $\boldsymbol{\kappa}$ . The derivation of the right hand side of equation 2.25 can be proven by expanding the exponential functions using the Baker-Campbell-Hausdorff (BCH) formula.[19]

$$a_{P(\boldsymbol{\kappa})}^\dagger = a_P^\dagger + [a_P^\dagger, \hat{\kappa}] + \frac{1}{2!} [[a_P^\dagger, \hat{\kappa}], \hat{\kappa}] + \dots \quad (2.26)$$

The commutators can be further simplified:

$$\begin{aligned} [a_P^\dagger, \hat{\kappa}] &= \sum_{QR} \kappa_{QR} [a_P^\dagger, a_Q^\dagger a_R] \\ &= \sum_{QR} \kappa_{QR} (a_P^\dagger a_Q^\dagger a_R - a_Q^\dagger a_R a_P^\dagger) \\ &= \sum_{QR} \kappa_{QR} (-a_Q^\dagger \delta_{PR}) \\ &= - \sum_Q a_Q^\dagger [\boldsymbol{\kappa}]_{QP} \end{aligned} \quad (2.27)$$

Here, we reintroduced the anti-symmetric matrix  $\boldsymbol{\kappa}$ . The nested commutators can be simplified using the equation above which can be applied from the outside to the inside in a similar form. Note that care must be taken concerning the labeling of summation indices when applying this rule multiple times.

$$\begin{aligned} [[a_P^\dagger, \hat{\kappa}], \hat{\kappa}], \hat{\kappa}] &= - \sum_Q [a_Q^\dagger, \hat{\kappa}] [\boldsymbol{\kappa}]_{QP} \\ &= \sum_{QR} a_R^\dagger [\boldsymbol{\kappa}]_{RQ} [\boldsymbol{\kappa}]_{QP} \\ &= \sum_Q a_Q^\dagger [\boldsymbol{\kappa}^2]_{QP} \end{aligned} \quad (2.28)$$

The n-folded nested commutator is then given by:

$$[\dots [[a_P^\dagger, \hat{\kappa}], \hat{\kappa}], \hat{\kappa}], \dots] = (-1)^n \sum_Q a_Q^\dagger [\boldsymbol{\kappa}^n]_{QP} \quad (2.29)$$

We may now use these definitions for the BCH expansion above:

$$\begin{aligned} a_{P(\boldsymbol{\kappa})}^\dagger &= \sum_Q a_Q^\dagger (\delta_{QP} - [\boldsymbol{\kappa}]_{QP} + \dots + \frac{(-1)^n}{n!} [\boldsymbol{\kappa}^n]_{QP} + \dots) \\ &= \sum_Q a_Q^\dagger \sum_{\alpha} \frac{(-1)^\alpha}{\alpha!} [\boldsymbol{\kappa}^\alpha]_{QP} \\ &= \sum_Q a_Q^\dagger [\exp(-\boldsymbol{\kappa})]_{QP} \end{aligned} \quad (2.30)$$

Which proves the substitution above to be valid, since:

$$\sum_Q a_Q^\dagger [\exp(-\boldsymbol{\kappa})]_{QP} = \sum_Q a_Q^\dagger U_{QP(\boldsymbol{\kappa})} = a_{P(\boldsymbol{\kappa})}^\dagger \quad (2.31)$$

The elegance of this redefinition becomes clear when applying the new creation operators on the vacuum state to create an orbital dependent expression for the state vector representing a Slater determinant:

$$\begin{aligned} \prod_i^{n_{ele}} a_{i(\boldsymbol{\kappa})}^\dagger | \rangle &= \prod_i^{n_{ele}} \exp(-\hat{\kappa}) a_i^\dagger \exp(\hat{\kappa}) | \rangle \\ &= \exp(-\hat{\kappa}) \prod_i^{n_{ele}} a_i^\dagger \exp(\hat{\kappa}) | \rangle \\ &= \exp(-\hat{\kappa}) \prod_i^{n_{ele}} a_i^\dagger | \rangle \end{aligned} \quad (2.32)$$

The reordering of the product is possible since  $\exp(-\hat{\kappa}) \exp(\hat{\kappa}) = 1$  and the last line above holds because one cannot annihilate an orbital in the vacuum state:

$$\exp(-\hat{\kappa}) | \rangle = \left( 1 - \hat{\kappa} + \frac{1}{2!} \hat{\kappa}^2 - \dots \right) | \rangle = | \rangle \quad (2.33)$$

The dependence of the wave function on orbital rotation parameters can then be expressed by the following general expression which holds for any determinant or linear combination of determinants  $|\Psi\rangle$ .

$$|\Psi(\boldsymbol{\kappa})\rangle = \exp(-\hat{\kappa}) |\Psi\rangle \quad (2.34)$$

To explain the connection of expectation values depicted in first and second quantization we may again use the density matrix. Let us note that the operator expressions in first and second quantization depend on the whole one particle basis and are unitary invariant. The dependency on the orbital rotation parameters therefore only remains in the wave function or the state

vector, respectively.

$$\begin{aligned}
\langle \Psi_{(\boldsymbol{\kappa})} | \hat{o}_1 | \Psi_{(\boldsymbol{\kappa})} \rangle &= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle \langle \Psi_{(\boldsymbol{\kappa})} | | \phi_P \rangle \langle \phi_Q | | \Psi_{(\boldsymbol{\kappa})} \rangle \\
&= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle \langle \Psi_{(\boldsymbol{\kappa})} | a_P^\dagger a_Q | \Psi_{(\boldsymbol{\kappa})} \rangle \\
&= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle \langle \Psi | \exp(\hat{\kappa}) a_P^\dagger a_Q \exp(-\hat{\kappa}) | \Psi \rangle \\
\langle \Psi_{(\boldsymbol{\kappa})} | \hat{o}_{12} | \Psi_{(\boldsymbol{\kappa})} \rangle &= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle \langle \Psi_{(\boldsymbol{\kappa})} | | \phi_P \phi_Q \rangle \langle \phi_R \phi_S | | \Psi_{(\boldsymbol{\kappa})} \rangle \\
&= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle \langle \Psi_{(\boldsymbol{\kappa})} | a_P^\dagger a_Q^\dagger a_S a_R | \Psi_{(\boldsymbol{\kappa})} \rangle \\
&= \sum_{PQRS} \langle \phi_P \phi_Q | \hat{o}_{12} | \phi_R \phi_S \rangle \\
&\quad \langle \Psi | \exp(\hat{\kappa}) a_P^\dagger a_Q^\dagger a_S a_R \exp(-\hat{\kappa}) | \Psi \rangle \tag{2.35}
\end{aligned}$$

As mentioned above, this work will deal with configuration interaction wave functions. For these wave functions the parameter space of  $(\boldsymbol{\kappa})$  can be further restricted. Since all configuration interaction coefficients are optimized, the energy does not change when applying unitary transformations within the occupied or the virtual space only. Only those transformations which couple these spaces increase the variational freedom of the wave function:

$$\begin{aligned}
\hat{\kappa} &= \sum_{PQ} \kappa_{PQ} a_P^\dagger a_Q \\
\Rightarrow \sum_{ai} \kappa_{ai} a_a^\dagger a_i + \sum_{ia} \kappa_{ia} a_i^\dagger a_a &= \sum_{ai} \kappa_{ai} (a_a^\dagger a_i - a_i^\dagger a_a) \tag{2.36}
\end{aligned}$$

Here, the last line results from the anti-symmetry of the matrix elements  $\kappa_{ia} = -\kappa_{ai}$ .

### 2.1.3 Spin Integration

In principle, the equations above are complete and the spin integration could be carried out after deriving the working equations. However, it is also possible to directly include the dependency on the spin coordinates into the second quantized equations. We may first explain the first quantized derivation using the expectation value of a one electron operator and the Hartree-Fock determinant. Let us define the spin orbitals  $\phi$  as a product of a spatial orbital  $\varphi$  and a spin function  $\sigma$ :

$$\phi(x_1) \rightarrow \varphi(r_1) \sigma(\omega_1) \tag{2.37}$$

The coordinates  $x_1$  are separated into spatial coordinates  $r_1$  and spin coordinates  $\omega_1$ , too. The spin functions represent the two possible states of an electron, namely  $\alpha(\omega_1)$  (spin up) and  $\beta(\omega_1)$  (spin down) and are orthonormal when integrated over  $\omega_1$ :

$$\langle \alpha | \alpha \rangle = 1 ; \langle \beta | \beta \rangle = 1 ; \langle \alpha | \beta \rangle = 0 ; \langle \beta | \alpha \rangle = 0 \quad (2.38)$$

For a closed shell determinant there is the same number of spin up and spin down spin orbitals. We may therefore split a sum over all spin orbitals  $\phi_i$  into two sums over spatial orbitals  $\varphi_i$ . Each of these sums extend over half the number of spin orbitals  $n_{occ} = n_{ele}/2$ , and one half depends on  $\alpha$  spin functions, the other half on  $\beta$  spin functions, respectively.

$$\begin{aligned} \langle \Phi | \hat{o}_1 | \Phi \rangle &= \sum_i^{n_{ele}} \langle \phi_i | \hat{o}_1 | \phi_i \rangle \\ &= \sum_i^{n_{occ}} \langle \varphi_i | \hat{o}_1 | \varphi_i \rangle \langle \alpha | \alpha \rangle + \sum_i^{n_{occ}} \langle \varphi_i | \hat{o}_1 | \varphi_i \rangle \langle \beta | \beta \rangle \\ &= 2 \sum_i^{n_{occ}} \langle \varphi_i | \hat{o}_1 | \varphi_i \rangle \end{aligned} \quad (2.39)$$

The Dirac notation hides a lot of information of variable dependencies. We may extend this notation here to explain this fact in more detail:

$$\begin{aligned} \langle \phi_i | \hat{o}_1 | \phi_i \rangle &= \int_{-\infty}^{\infty} \phi_{i(x_1)}^* \hat{o}_{(x_1)} \phi_{i(x_1)} dx_1 \\ &= \int_{-\infty}^{\infty} \varphi_{i(r_1)}^* \sigma_{(\omega_1)}^* \hat{o}_{(r_1, \omega_1)} \varphi_{i(r_1)} \sigma_{(\omega_1)} dr_1 d\omega_1 \end{aligned} \quad (2.40)$$

We may assume that the operator  $\hat{o}_{(r_1, \omega_1)}$  only depends on the electronic coordinate  $r_1$  and not on the spin coordinate  $\omega_1$  which is the case for all commonly considered operators in the electronic non-relativistic Hamiltonian of a molecule:

$$\begin{aligned} &\int_{-\infty}^{\infty} \varphi_{i(r_1)}^* \sigma_{(\omega_1)}^* \hat{o}_{(r_1)} \varphi_{i(r_1)} \sigma_{(\omega_1)} dr_1 d\omega_1 \\ &= \int_{-\infty}^{\infty} \varphi_{i(r_1)}^* \hat{o}_{(r_1)} \varphi_{i(r_1)} dr_1 \int_{-\infty}^{\infty} \sigma_{(\omega_1)}^* \sigma_{(\omega_1)} d\omega_1 \\ &= \langle \varphi_i | \hat{o}_1 | \varphi_i \rangle \langle \sigma | \sigma \rangle \end{aligned} \quad (2.41)$$

All these steps are commonly abbreviated by the short hand Dirac notation, which will be used throughout the rest of this work.

## Spin free operators

To illustrate the connection between first and second quantization we may again use the projection operator representation as defined in equation 2.9 and decompose the sums over orbitals into  $\alpha$  and  $\beta$  contributions.

$$\begin{aligned}
\hat{o}_1 &= \sum_{PQ} \langle \phi_P | \hat{o}_1 | \phi_Q \rangle | \phi_P \rangle \langle \phi_Q | \\
&= \sum_{PQ} \langle \varphi_P | \hat{o}_1 | \varphi_Q \rangle | \varphi_P \rangle \langle \varphi_Q | \sum_{\sigma\tau}^2 \langle \sigma | \tau \rangle | \sigma \rangle \langle \tau | \\
&= \sum_{PQ} \langle \varphi_P | \hat{o}_1 | \varphi_Q \rangle | \varphi_P \rangle \langle \varphi_Q | \sum_{\sigma}^2 | \sigma \rangle \langle \sigma | \\
&= \sum_{PQ} \langle \varphi_P | \hat{o}_1 | \varphi_Q \rangle | \varphi_P \rangle \langle \varphi_Q | (|\alpha\rangle \langle \alpha| + |\beta\rangle \langle \beta|)
\end{aligned} \tag{2.42}$$

Note that the sum over Greek letters extends over the two spin functions  $\alpha$  and  $\beta$ . To introduce the spin dependency at the level of second quantization, we have to generalize the creation and annihilation operators.

$$a_{P\sigma}^\dagger | \rangle = | \varphi_P \rangle | \sigma \rangle \tag{2.43}$$

Each creation operator now creates a product of a spatial orbital and a spin function. The operator representation in second quantization can be defined by the same decomposition of sums over spatial and spin functions as in equation 2.42:

$$\begin{aligned}
\hat{o}_1 &\rightarrow \sum_{PQ} \langle \varphi_P | \hat{o}_1 | \varphi_Q \rangle \sum_{\sigma\tau}^2 \langle \sigma | \tau \rangle a_{P\sigma}^\dagger a_{Q\tau} \\
&= \sum_{PQ} \langle \varphi_P | \hat{o}_1 | \varphi_Q \rangle (a_{P\alpha}^\dagger a_{Q\alpha} + a_{P\beta}^\dagger a_{Q\beta}) \\
&= \sum_{PQ} \langle \varphi_P | \hat{o}_1 | \varphi_Q \rangle E_{PQ}
\end{aligned} \tag{2.44}$$

Here, we defined the singlet excitation operator  $E_{PQ}$  as  $a_{P\alpha}^\dagger a_{Q\alpha} + a_{P\beta}^\dagger a_{Q\beta}$ . In the same way the two electron operator is given by:

$$\begin{aligned}
\hat{o}_{12} &\rightarrow \sum_{PQRS} \langle \varphi_P \varphi_Q | \hat{o}_{12} | \varphi_R \varphi_S \rangle (a_{P\alpha}^\dagger a_{Q\alpha}^\dagger a_{S\alpha} a_{R\alpha} + a_{P\alpha}^\dagger a_{Q\beta}^\dagger a_{S\beta} a_{R\alpha} \\
&\quad + a_{P\beta}^\dagger a_{Q\alpha}^\dagger a_{S\alpha} a_{R\beta} + a_{P\beta}^\dagger a_{Q\beta}^\dagger a_{S\beta} a_{R\beta}) \\
&= \sum_{PQRS} \langle \varphi_P \varphi_Q | \hat{o}_{12} | \varphi_R \varphi_S \rangle e_{PQRS}
\end{aligned} \tag{2.45}$$

The singlet two electron excitation operator  $e_{PQRS}$  may be rewritten by singlet excitation operators  $E_{PQ}$ :

$$e_{PQRS} = E_{PS}E_{QR} - E_{PS}\delta_{QR} \quad (2.46)$$

Note that this definition would change if the Mulliken notation is used for the integral representation. To derive working equations with these new creation operators we need to generalize the anti-commutator relations above (see equation 2.7):

$$a_{P\sigma}^\dagger a_{Q\tau} = \delta_{PQ}\delta_{\sigma\tau} - a_{Q\tau} a_{P\sigma}^\dagger \quad (2.47)$$

This modification results from overlap integrals which will not be derived here and can be obtained from literature [17].

### Spin operators

To discuss the spin properties of a wave function we have to define the spin operators  $\hat{S}_{z(1)}$  and  $\hat{S}_{(12)}^2$ . It is convenient to additionally define the Cartesian components  $\hat{S}_{x(1)}$  and  $\hat{S}_{y(1)}$  by the shift operators  $\hat{S}_{+(1)}$  and  $\hat{S}_{-(1)}$ .

$$\begin{aligned} \hat{S}_{x(1)} &= \frac{1}{2} (\hat{S}_{+(1)} + \hat{S}_{-(1)}) \\ \hat{S}_{y(1)} &= \frac{1}{2i} (\hat{S}_{+(1)} - \hat{S}_{-(1)}) \end{aligned} \quad (2.48)$$

The action of these operators on a spin function is defined as:

$$\begin{aligned} \hat{S}_{+(1)}\alpha_{(1)} &= 0 & ; & \quad \hat{S}_{+(1)}\beta_{(1)} = \alpha_{(1)} \\ \hat{S}_{-(1)}\alpha_{(1)} &= \beta_{(1)} & ; & \quad \hat{S}_{-(1)}\beta_{(1)} = 0 \\ \hat{S}_{z(1)}\alpha_{(1)} &= \frac{1}{2}\alpha_{(1)} & ; & \quad \hat{S}_{z(1)}\beta_{(1)} = -\frac{1}{2}\beta_{(1)} \end{aligned} \quad (2.49)$$

They allow us to represent the two-electron operator  $\hat{S}_{(12)}^2$  in a more compact form.

$$\begin{aligned} \hat{S}_{(12)}^2 &= (\hat{S}_{x(1)} - \hat{S}_{x(2)})^2 + (\hat{S}_{y(1)} - \hat{S}_{y(2)})^2 + (\hat{S}_{z(1)} - \hat{S}_{z(2)})^2 \\ &= [(\hat{S}_{+(1)} - \hat{S}_{+(2)}), (\hat{S}_{-(1)} - \hat{S}_{-(2)})] + (\hat{S}_{z(1)} - \hat{S}_{z(2)})^2 \\ &= [\hat{S}_{+(1)}, \hat{S}_{-(1)}] + [\hat{S}_{+(2)}, \hat{S}_{-(2)}] + (\hat{S}_{z(1)} - \hat{S}_{z(2)})^2 \end{aligned} \quad (2.50)$$

In second quantization a one-electron operator  $\hat{o}_1$  depending only on a spin coordinate can be represented as:

$$\begin{aligned} \hat{o}_1 &\rightarrow \sum_{\sigma\tau} \langle \sigma | \hat{o}_1 | \tau \rangle \sum_{PQ} \langle \varphi_P | \varphi_Q \rangle a_{P\sigma}^\dagger a_{Q\tau} \\ &= \sum_{\sigma\tau} \langle \sigma | \hat{o}_1 | \tau \rangle \sum_P a_{P\sigma}^\dagger a_{P\tau} \end{aligned} \quad (2.51)$$

For the two-electron operator  $\hat{o}_{12}$  depending only on spin coordinates holds:

$$\begin{aligned}\hat{o}_{12} &\rightarrow \sum_{\sigma\tau\epsilon\omega}^2 \langle \sigma\epsilon | \hat{o}_{12} | \tau\omega \rangle \sum_{PQRS} \langle \varphi_P | \varphi_R \rangle \langle \varphi_Q | \varphi_S \rangle a_{P\sigma}^\dagger a_{Q\epsilon}^\dagger a_{S\omega} a_{R\tau} \\ &= \sum_{\sigma\tau\epsilon\omega}^2 \langle \sigma\epsilon | \hat{o}_{12} | \tau\omega \rangle \sum_{PQ} a_{P\sigma}^\dagger a_{Q\epsilon}^\dagger a_{Q\omega} a_{P\tau}\end{aligned}\quad (2.52)$$

The second quantization one-electron operators  $\hat{S}_{+(1)}$ ,  $\hat{S}_{-(1)}$  and  $\hat{S}_{z(1)}$  are given by:

$$\begin{aligned}\sum_{\sigma\tau}^2 \langle \sigma | \hat{S}_{+(1)} | \tau \rangle \sum_P a_{P\sigma}^\dagger a_{P\tau} &= \sum_P a_{P\alpha}^\dagger a_{P\beta} = \hat{S}_+ \\ \sum_{\sigma\tau}^2 \langle \sigma | \hat{S}_{-(1)} | \tau \rangle \sum_P a_{P\sigma}^\dagger a_{P\tau} &= \sum_P a_{P\beta}^\dagger a_{P\alpha} = \hat{S}_- \\ \sum_{\sigma\tau}^2 \langle \sigma | \hat{S}_{z(1)} | \tau \rangle \sum_P a_{P\sigma}^\dagger a_{P\tau} &= \sum_P a_{P\alpha}^\dagger a_{P\alpha} - a_{P\beta}^\dagger a_{P\beta} = \hat{S}_z\end{aligned}\quad (2.53)$$

The two-electron operator  $\hat{S}_{(12)}^2$  can be represented using products of one-electron operators. Note that the right hand side of the expression below is completely defined in second quantization indicated by the missing electronic coordinates.

$$\sum_{\sigma\tau\epsilon\omega}^2 \langle \sigma\epsilon | \hat{S}_{(12)}^2 | \tau\omega \rangle \sum_{PQ} a_{P\sigma}^\dagger a_{Q\epsilon}^\dagger a_{Q\omega} a_{P\tau} = \hat{S}_+ \hat{S}_- + \hat{S}_z (\hat{S}_z - 1) = \hat{S}^2 \quad (2.54)$$

There are many other representations of this operator. The one defined above is however especially suitable for several derivations in the following section.

### Spin dependency of the wave function

The spin dependency of the operator expressions in second quantization is now defined and we may focus on the wave function. Since we consider configuration interaction in our work, we may rewrite the singly and doubly excited determinants. All higher excitations are accessible by the same rules as defined below. Assuming we only want to describe singlet states, which is the most common case, the whole wave function has to be a singlet state. Determinants are not necessarily a singlet state, but one can form linear combinations of certain subsets of them to create a singlet. If the normal set of

determinants is used, these linear combinations would be generated automatically during the optimization by the configuration interaction coefficients to ensure the correct spin properties of the wave function. An alternative way is to predefine the correct linear combinations of determinants without losing any variational freedom. These linear combinations of determinants with the correct spin properties are called configuration state functions (CSF) and represent the wave function in a more compact form since multiple determinants are replaced by one linear combination of them. The correct linear combinations of determinants have to fulfill the eigenvalue equations for a singlet state ( $M_S = 0$  and  $S = 0$ ). The value of  $2|S| + 1$  defines the spin multiplicity of the molecule and  $M_S$  ranges from  $-S, \dots, 0, \dots, S$ .

$$\begin{aligned}\hat{S}_z |\Psi\rangle &= M_S |\Psi\rangle \rightarrow 0 |\Psi\rangle \\ \hat{S}^2 |\Psi\rangle &= S(S+1) |\Psi\rangle \rightarrow 0 |\Psi\rangle\end{aligned}\quad (2.55)$$

Let us assume we use a closed shell Hartree-Fock wave function  $|\Phi\rangle$  as a reference. For this determinant the following equations hold:

$$\begin{aligned}\hat{S}_z |\Phi\rangle &= 0 |\Phi\rangle \\ \hat{S}^2 |\Phi\rangle &= 0 |\Phi\rangle\end{aligned}\quad (2.56)$$

since all electrons are paired. Acting upon this reference state with the single excitation operator  $\hat{T}_S$  and the double excitation operator  $\hat{T}_D$  generates a linear combination of all respective excitations.

$$\begin{aligned}\hat{T}_S &= \sum_{ai} \sum_{\sigma\tau}^2 A_{i\tau}^{a\sigma} a_{a\sigma}^\dagger a_{i\tau} \\ \hat{T}_D &= \sum_{abij} \sum_{\sigma\tau\epsilon\omega}^2 A_{i\tau j\omega}^{a\sigma b\epsilon} a_{a\sigma}^\dagger a_{b\epsilon}^\dagger a_{j\omega} a_{i\tau}\end{aligned}\quad (2.57)$$

Note that the sum over the Greek letters extends over the two spin functions  $\alpha$  and  $\beta$ . The action of the excitation operators upon the reference should not change the value of the eigenvalue equations 2.55 to maintain the correct spin state.

$$\begin{aligned}\hat{S}_z \hat{T}_S |\Phi\rangle &= 0 \hat{T}_S |\Phi\rangle \\ \hat{S}^2 \hat{T}_S |\Phi\rangle &= 0 \hat{T}_S |\Phi\rangle \\ \hat{S}_z \hat{T}_D |\Phi\rangle &= 0 \hat{T}_D |\Phi\rangle \\ \hat{S}^2 \hat{T}_D |\Phi\rangle &= 0 \hat{T}_D |\Phi\rangle\end{aligned}\quad (2.58)$$

We may first consider the action of  $\hat{S}_z$  upon  $\hat{T}_S |\Phi\rangle$ :

$$\begin{aligned}\hat{S}_z \hat{T}_S |\Phi\rangle &= \frac{1}{2} \sum_P (a_{P\alpha}^\dagger a_{P\alpha} - a_{P\beta}^\dagger a_{P\beta}) \sum_{ai} \sum_{\sigma\tau}^2 A_{i\tau}^{a\sigma} a_{a\sigma}^\dagger a_{i\tau} |\Phi\rangle \\ &= \frac{1}{2} \sum_{ai} \sum_{\sigma\tau}^2 A_{i\tau}^{a\sigma} a_{a\sigma}^\dagger a_{i\tau} |\Phi\rangle \sum_{\epsilon} (\delta_{\sigma\epsilon} - \delta_{\tau\epsilon}) + \hat{T}_S \hat{S}_z |\Phi\rangle\end{aligned}\quad (2.59)$$

To further simplify this equation we note that the last term vanishes.

$$\begin{aligned}\hat{S}_z |\Phi\rangle &= \frac{1}{2} \sum_P (a_{P\alpha}^\dagger a_{P\alpha} - a_{P\beta}^\dagger a_{P\beta}) |\Phi\rangle \\ &= \frac{1}{2} \sum_i (a_{i\alpha}^\dagger a_{i\alpha} - a_{i\beta}^\dagger a_{i\beta}) |\Phi\rangle \\ &= \frac{1}{2} (n_{occ} - n_{occ}) |\Phi\rangle \\ &= 0\end{aligned}\quad (2.60)$$

Only occupied orbitals can be annihilated from the reference, and the remaining sums annihilate and create a specific orbital  $n_{occ} = n_{ele}/2$  times for each spin function. To ensure that the eigenvalue equation 2.55 is fulfilled for a singlet state, we have to modify our definition of the singlet excitation operator  $\hat{T}_S \rightarrow \hat{T}'_S$ .

$$\hat{T}'_S = \sum_{ai} \sum_{\sigma\tau}^2 A_{i\tau}^{a\sigma} a_{a\sigma}^\dagger a_{i\tau} \delta_{\sigma\tau} = \sum_{ai} \sum_{\sigma}^2 A_{i\sigma}^{a\sigma} a_{a\sigma}^\dagger a_{i\sigma}\quad (2.61)$$

The eigenvalue equation is then given by:

$$\hat{S}_z \hat{T}'_S |\Phi\rangle = \frac{1}{2} \sum_{ai} \sum_{\sigma\tau}^2 A_{i\tau}^{a\sigma} a_{a\sigma}^\dagger a_{i\tau} |\Phi\rangle \sum_{\epsilon} (\delta_{\sigma\epsilon} - \delta_{\tau\epsilon}) \delta_{\sigma\tau} = 0 \hat{T}'_S |\Phi\rangle\quad (2.62)$$

We may now consider the action of  $\hat{S}^2$  upon  $\hat{T}'_S |\Phi\rangle$ , which is the second eigenvalue equation of 2.55 to be fulfilled. Note that the action of  $\hat{S}_z$  upon

$\hat{T}'_S |\Phi\rangle$  results in a zero as shown above.

$$\begin{aligned}
\hat{S}^2 \hat{T}'_S |\Phi\rangle &= (\hat{S}_+ \hat{S}_- + \hat{S}_z (\hat{S}_z - 1)) \hat{T}'_S |\Phi\rangle = \hat{S}_+ \hat{S}_- \hat{T}'_S |\Phi\rangle \\
&= \sum_P a_{P\alpha}^\dagger a_{P\beta} \sum_Q a_{Q\beta}^\dagger a_{Q\alpha} \sum_{ai} \sum_{\sigma}^2 A_{i\sigma}^{a\sigma} a_{a\sigma}^\dagger a_{i\sigma} |\Phi\rangle \\
&= (\hat{T}'_S - \sum_{ai} (A_{i\beta}^{a\beta} a_{a\alpha}^\dagger a_{i\alpha} + A_{i\alpha}^{a\alpha} a_{a\beta}^\dagger a_{i\beta})) |\Phi\rangle \\
&+ \sum_{ai} (A_{i\alpha}^{a\alpha} - A_{i\beta}^{a\beta}) a_{a\beta}^\dagger a_{i\alpha} \hat{S}_+ |\Phi\rangle \\
&- \sum_{ai} (A_{i\alpha}^{a\alpha} - A_{i\beta}^{a\beta}) a_{a\alpha}^\dagger a_{i\beta} \hat{S}_- |\Phi\rangle \\
&+ \hat{T}'_S \hat{S}_+ \hat{S}_- |\Phi\rangle
\end{aligned} \tag{2.63}$$

Here, all terms containing expressions such as  $S_- |\Phi\rangle$  and  $S_+ |\Phi\rangle$  vanish.

$$\begin{aligned}
\hat{S}_- |\Phi\rangle &= \sum_Q a_{Q\beta}^\dagger a_{Q\alpha} |\Phi\rangle \\
&= \sum_i a_{i\beta}^\dagger a_{i\alpha} |\Phi\rangle \\
&= 0
\end{aligned} \tag{2.64}$$

Only occupied orbitals can be annihilated from the reference determinant, and the remaining sum replaces a spin up electron by a spin down electron or vice versa. As a result there are two electrons with the same set of quantum numbers or in other words two identical rows or columns in the reference determinant. The remaining terms of the eigenvalue equation are given by:

$$\hat{S}^2 \hat{T}'_S |\Phi\rangle = (\hat{T}'_S - \sum_{ai} (A_{i\beta}^{a\beta} a_{a\alpha}^\dagger a_{i\alpha} + A_{i\alpha}^{a\alpha} a_{a\beta}^\dagger a_{i\beta})) |\Phi\rangle \tag{2.65}$$

To ensure that the second eigenvalue equation is fulfilled for a singlet state we have to modify our definition of the singlet excitation operator again  $\hat{T}'_S \rightarrow \hat{T}''_S$ .

$$\hat{T}''_S = \sum_{ai} A_i^a \sum_{\sigma}^2 a_{a\sigma}^\dagger a_{i\sigma} = \sum_{ai} A_i^a E_{ai} \tag{2.66}$$

The eigenvalue equation is then given by:

$$\begin{aligned}
\hat{S}^2 \hat{T}''_S |\Phi\rangle &= (\hat{T}''_S - \sum_{ai} A_i^a (a_{a\alpha}^\dagger a_{i\alpha} + a_{a\beta}^\dagger a_{i\beta})) |\Phi\rangle \\
&= (\hat{T}''_S - \hat{T}''_S) |\Phi\rangle = 0 \hat{T}''_S |\Phi\rangle
\end{aligned} \tag{2.67}$$

The excitation operator  $\hat{T}_S''$  now fulfills both eigenvalue equations 2.55 and is the appropriate ansatz to generate a wave function suited for singlet states  $\hat{T}_S^{singlet} = \hat{T}_S''$ . The configuration state functions  $|\tilde{\Phi}_i^a\rangle$  are then given by:

$$|\tilde{\Phi}_i^a\rangle = (|\Phi_{i\alpha}^{a\alpha}\rangle + |\Phi_{i\beta}^{a\beta}\rangle) \quad (2.68)$$

Using the same procedure, similar equations can be derived for the double excitation operator  $\hat{T}_D^{singlet}$  resulting in:

$$\hat{T}_D^{singlet} = \sum_{abij} A_{ij}^{ab} e_{abij} \quad (2.69)$$

The corresponding configuration state functions  $|\tilde{\Phi}_{ij}^{ab}\rangle$  are defined as:

$$|\tilde{\Phi}_{ij}^{ab}\rangle = (|\Phi_{i\alpha j\alpha}^{a\alpha b\alpha}\rangle + |\Phi_{i\alpha j\beta}^{a\alpha b\beta}\rangle + |\Phi_{i\beta j\alpha}^{a\beta b\alpha}\rangle + |\Phi_{i\beta j\beta}^{a\beta b\beta}\rangle) \quad (2.70)$$

All other possible combinations of spin functions are unnecessary for the description of the singlet wave function and the configuration interaction coefficients are independent of the spin labels resulting in a reduced parameter space while maintaining the full variational freedom. Since we also consider orbital optimization, we will have to redefine the orbital rotation operator of equation 2.36 as:

$$\hat{\kappa} = \sum_{ai} \kappa_{ai} (E_{ai} - E_{ia}) \quad (2.71)$$

With these definitions, all possible simplifications for singlet wave functions are defined and the working equations used below can be derived in a efficient way only using the elegant rules of second quantization.

## 2.2 Quasi-Newton Minimization

The search for extrema of functions depending on a huge number of variables is a very large field of research and cannot be covered completely here. However, since we have to minimize energy functionals with respect to coefficients of determinants or basis functions, we will shortly explain a very prominent procedure to solve such problems which is extensively used in electronic structure theory. The general procedure is based on expanding the energy in a Taylor series up to second order for a fixed set of parameters

$\mathbf{X}_n$  and a correction to these parameters  $\Delta\mathbf{X}$ .

$$\begin{aligned}
E_{(\mathbf{X}_n+\Delta\mathbf{X})} &\approx E_{(\mathbf{X}_n)}^{(0)} + (\Delta\mathbf{X})^T \mathbf{E}_{(\mathbf{X}_n)}^{(1)} + \frac{1}{2} \Delta\mathbf{X}^T \mathbf{E}_{(\mathbf{X}_n)}^{(2)} \Delta\mathbf{X} \\
E_p^{(1)} &= \frac{\partial}{\partial X_p} E_{(\mathbf{X}_n)} \\
E_{pq}^{(2)} &= \frac{\partial^2}{\partial X_p \partial X_q} E_{(\mathbf{X}_n)}
\end{aligned} \tag{2.72}$$

This ansatz already indicates an iterative procedure. The update  $\Delta\mathbf{X}$  for a new coefficient matrix  $\mathbf{X}_{n+1}$  can be calculated by a conjugate gradient[20] approach and additionally scaled by a step size:

$$\begin{aligned}
\mathbf{E}_{(\mathbf{X}_n)}^{(2)} \Delta\mathbf{X} &= -\mathbf{E}_{(\mathbf{X}_n)}^{(1)} \\
\mathbf{X}_{n+1} &= \mathbf{X}_n + \alpha \Delta\mathbf{X}
\end{aligned} \tag{2.73}$$

The correction  $\Delta\mathbf{X}$  vanishes as the gradient vector  $\mathbf{E}^{(1)}$  becomes zero and the procedure converges to an extremum.  $\mathbf{E}^{(2)}$  is the Hessian matrix. The first iteration needs a reasonable initial guess. For our problems  $\mathbf{X}_0 = 0$  is a good choice, since the final values of  $\mathbf{X}$  are mostly sufficiently close to zero, too. The step size  $\alpha$  can be varied using various criteria. Here, we use the simplest approach and roughly minimize the function value for the current correction  $\Delta\mathbf{X}$ . In more detail, we iteratively minimize the energy depending on the step size until the change of the energy becomes approximately an order of magnitude smaller than the change of the energy in the previous iteration.

So far, we only considered the conventional Newton method for minimization. Quasi-Newton minimization seeks to find reasonable approximations of the Hessian matrix  $\mathbf{E}^{(2)}$ . The matrix vector operation  $\mathbf{E}_{(\mathbf{X}_n)}^{(2)} \Delta\mathbf{X}$  scales quadratically, and the computation of this matrix is often the most time-consuming step of the optimization. For our minimization problems, the storage of this matrix is not feasible. Therefore, the recalculation of this matrix is mandatory for each matrix vector product formed. To reduce the time consumption of these steps one may assume that the matrix  $\mathbf{E}^{(2)}$  is diagonal. The matrix vector product becomes a linear operation and the computation of the diagonal elements is obviously much faster than the whole matrix computation. The elements of the correction vector  $\Delta X_p$  can be expressed by the analytical expression:

$$\Delta X_p = -E_{p(\mathbf{X}_n)}^{(1)} / E_{pp(\mathbf{X}_n)}^{(2)} \tag{2.74}$$

The diagonal approximation is broadly applied when minimizing the energy of pair correlation methods and shows fast and stable convergence.[21]

There are cases where an expansion of the function into a Maclaurin series is simpler. In this case, the parameters  $\mathbf{X}_n$  stay zero in each iteration and the correction is directly included into the functions  $E_{n(\mathbf{0})}^{(0)}$ . This way of expressing the optimization problem is of advantage if the gradient and Hessian equations are simpler in the limit  $\mathbf{X} \rightarrow 0$ . Note that  $\mathbf{X} = 0$  has to be the correct solution for the system of equations when using this ansatz.

$$\begin{aligned}
E_{n(\Delta\mathbf{X})} &\approx E_{n(\mathbf{0})}^{(0)} + (\Delta\mathbf{X})^T \mathbf{E}_{n(\mathbf{0})}^{(1)} + \frac{1}{2} \Delta\mathbf{X}^T \mathbf{E}_{n(\mathbf{0})}^{(2)} \Delta\mathbf{X} \\
E_p^{(1)} &= \lim_{\mathbf{X} \rightarrow 0} \frac{\partial}{\partial X_p} E_n(\mathbf{X}) \\
E_{pq}^{(2)} &= \lim_{\mathbf{X} \rightarrow 0} \frac{\partial^2}{\partial X_p \partial X_q} E_n(\mathbf{X})
\end{aligned} \tag{2.75}$$

The update  $\Delta\mathbf{X}$  for a new coefficient matrix  $\mathbf{X}_{n+1}$  can then be calculated by conjugate gradient [20] and additionally scaled by a step size:

$$\begin{aligned}
\mathbf{E}_{n(\mathbf{0})}^{(2)} \Delta\mathbf{X} &= -\mathbf{E}_{n(\mathbf{0})}^{(1)} \\
\mathbf{X}_{n+1} &= \alpha \Delta\mathbf{X}
\end{aligned} \tag{2.76}$$

This way of representing the series expansion is especially suitable for orbital optimization. As shown above, the energy depends on the orbital rotation parameters in an exponential form. In the limit  $\kappa \rightarrow 0$ , this exponential reduces to the identity operator and the resulting equations are greatly simplified:

$$\lim_{\kappa \rightarrow 0} \exp(\hat{\kappa}) = \lim_{\kappa \rightarrow 0} \hat{1} + \hat{\kappa} + \frac{1}{2} \hat{\kappa}^2 + \dots = \hat{1} \tag{2.77}$$

To fulfill all conditions stated above, we need to update the energy expression in each iteration. For the orbital rotation parameters of a given iteration  $n$  the new orbitals are defined as:

$$|\phi_{i(\kappa_n)}\rangle = \sum_p |\phi_p\rangle [\exp(-\kappa_n)]_{pi} \tag{2.78}$$

As soon as the orbitals are updated, the new function value  $E_{n(\mathbf{0})}^{(0)}$ , gradient vector  $\mathbf{E}_{n(\mathbf{0})}^{(1)}$  and hessian matrix  $\mathbf{E}_{n(\mathbf{0})}^{(2)}$  can be computed. For the subsequent iteration  $\kappa$  is set to zero to leave the updated orbitals as they are. Afterwards, a new correction can be computed at  $\kappa = 0$ . In the final iteration the correction and the resulting  $\kappa$  vanish.

## 2.3 Local Correlation Methods

Exploiting the restricted range of electron correlation effects to avoid the calculation of negligible energy contributions is a field of ongoing research, especially for accurate wave function based methods.[22] To use the locality of electron correlation, all modern local correlation methods use localized occupied molecular orbitals. There are different criteria available to define these, for example the Pipek-Mezey criterion.[23] Here, the number of basis functions over which a molecular orbital extends is minimized. Since these basis functions are atom centered, this also results in a minimal number atoms which are described by this orbital. In large molecular systems most atoms are distant to each other. Consequently, most of the localized orbitals are also distant. To utilize the distance of localized orbitals we may use the following definition of the correlation energy:

$$E_{CORR} = \sum_{ij} \epsilon_{ij} \quad (2.79)$$

Namely, the correlation energy of any correlation method can be decomposed into energy contributions from each pair of occupied orbitals. Note that any restrictions on the summation depend on the definition of the pair energies  $\epsilon_{ij}$ . If the occupied orbitals  $\phi_i$  and  $\phi_j$  are distant, there will be no correlation interaction and the corresponding pair energy will be approximately zero. There are quite efficient ways available to predict whether this is the case.[22] As a result all corresponding elements in the gradient matrix and the coefficient matrix will be approximately zero and do not have to be computed or saved.

The construction of the gradient vector  $\mathbf{E}^{(1)}$  is needed multiple times to minimize the energy (see section 2.2). This step is the computational most crucial operation of any correlation method. For pair correlation methods, the gradient vector is mapped on tensor elements, which will be called  $W_{ij}^{ab}$ . [2] As stated before, the construction of the matrix  $\mathbf{W}_{ij}^{MO}$  (with the dimensions of the virtual space) is not necessary in case the corresponding occupied orbital pair is distant. The tensor elements  $W_{ij}^{ab}$  can alternatively be defined in the atomic orbital basis  $W_{ij}^{\mu\nu}$ . The transition between these two bases for a given molecular orbital  $|a\rangle$  is given by a linear combination of atomic orbitals  $|\mu\rangle$ :

$$|a\rangle = \sum_{\mu} c_{\mu a} |\mu\rangle \quad (2.80)$$

Since the atomic orbital basis is much more local than the virtual space, the matrix  $\mathbf{W}_{ij}^{AO}$  is very sparse, and only the non-zero elements have to be

computed and stored. The transformation of the atomic orbital basis to the molecular orbital basis is defined as:

$$\mathbf{W}_{ij}^{MO} = \mathbf{C}^t \mathbf{W}_{ij}^{AO} \mathbf{C} \quad (2.81)$$

Here,  $\mathbf{C}$  is a rectangular matrix containing the virtual molecular orbital coefficient vectors resulting from a Hartree-Fock calculation (cf. equation 2.80). The whole construction of the tensor elements  $W_{ij}^{\mu\nu}$  can be reduced to multiplications of matrices in the atomic orbital or molecular orbital basis. Details are not important for an overview and can be found in reference [2].

To optimize the energy, a transformation of the atomic orbitals to molecular orbitals is necessary to retain the orthogonality relations assumed during the derivation (cf. section 2.1). However, there are compact working equations available where the orthogonality of the occupied space to the virtual space is sufficient.[2] In this case, the virtual space can be spanned by a more local basis than the virtual molecular orbitals resulting from a Hartree-Fock calculation. There are two commonly used ways to expand the virtual space in current local correlation methods.[22]

One possible way is to expand the virtual space into projected atomic orbitals (PAO's).[24] Here, we generate a new virtual space defined by  $\tilde{\mathbf{C}}$  which is orthonormal to the occupied space, but stays very local. The transformation results from the projection:

$$|\phi_{\mu}^{PAO}\rangle = (\hat{1} - \sum_i |\phi_i\rangle \langle \phi_i|) |\phi_{\mu}^{AO}\rangle \quad (2.82)$$

Here,  $\phi_{\mu}^{PAO}$  are the projected atomic orbitals,  $\phi_{\mu}^{AO}$  are the atomic orbitals and  $\phi_i$  are the localized occupied orbitals. As stated above local orbitals  $\phi_i$  extend only over a small subset of all atomic orbitals  $\phi_{\mu(i)}^{AO}$ . The same assignment is now possible for the projected atomic orbitals  $\phi_{\mu(i)}^{PAO}$ , which span our new virtual space. Considering the construction of the gradient tensor  $\mathbf{W}_{ij}^{MO} \rightarrow \mathbf{W}_{ij}^{PAO}$ , the union of the atomic- or projected atomic orbitals corresponding to the occupied orbitals  $i$  and  $j$  is sufficient. All other atomic- and projected atomic orbitals would lead to contributions which are approximately zero and do not have to be calculated or stored. This leads to a very compact representation of the gradient tensor.

Another possible way is to expand the virtual space into pair natural orbitals (PNO's), which are again very local.[25] Every occupied pair has a specific virtual space, defined by the orbital coefficient tensor  $\tilde{\mathbf{C}}_{ij}$ . These new coefficient matrices contain many elements which are approximately zero and can be neglected. This results in a very sparse and therefore compact representation of the corresponding gradient matrix  $\mathbf{W}_{ij}^{MO}$ .

$$\mathbf{W}_{ij}^{MO} = \tilde{\mathbf{C}}_{ij}^t \mathbf{W}_{ij}^{AO} \tilde{\mathbf{C}}_{ij} \quad (2.83)$$

There are multiple ways to define this virtual space, e.g. by diagonalizing the corresponding pair density matrix. Both approaches can in principle be combined leading to very efficient algorithms like DLPNO-CCSD.[22]

In our new contraction scheme all these transformations become unnecessary, since the conventional gradient tensor is contracted within the virtual space (cf. chapter 3.4).

# Chapter 3

## Contracted Pair Correlation Methods

Pair correlation methods are able to achieve highly accurate solutions for chemical problems. Unfortunately, their applicability is generally restricted to medium sized molecules due to storage requirements and computational costs. These restrictions can be partly overcome by local electron correlation methods. These methods use physical and mathematical criteria to decide which interactions are of such a long range that they do not have to be computed and saved. In our new ansatz, we define an alternative way towards local correlation. The range of interactions is strictly bound to the decay of integrals over Gaussian type geminals in the atomic orbital basis. The number of variables is reduced by orders of magnitude applying an efficient contraction scheme, leading to a naturally local representation of correlation effects.

### 3.1 Ansatz for the Wave Function

The basic ansatz for a pair correlation method can be reduced to configuration interaction doubles:

$$|\Psi_{(\mathbf{A})}\rangle = |\Phi\rangle + \sum_{ijab} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle = |\Phi\rangle + |\Psi_{D(\mathbf{A})}\rangle \quad (3.1)$$

Here,  $|\Phi\rangle$  is the Hartree Fock reference determinant and the sum contains all double excitations from the occupied to the virtual space ( $ij \rightarrow ab$ ). The part of the wave function containing all double excitations will be abbreviated by  $|\Psi_{D(\mathbf{A})}\rangle$  and  $\mathbf{A}$  denotes the dependence on all tensor elements  $A_{ij}^{ab}$ . With this wave function and a modified energy functional a whole class of

pair correlation methods arises.[2, 4, 3]. This ansatz for the wave function introduces a large set of parameters, namely the coefficients  $A_{ij}^{ab}$  which need to be stored. As a result, the application of this ansatz is limited to small and medium sized molecules. There are local correlation methods available which introduce a basis, where many of the coefficients are negligibly small, leading to a significant reduction of computation time and storage requirements (cf. section 3.4). Unfortunately, the computational benefit of these approaches is strongly dependent on the molecular structure. Namely, the computational benefits in chain like molecules will be much higher than in bulky molecules.[26, 27, 28, 29, 30] In our new ansatz, we introduce a contraction scheme which strongly reduces the number of coefficients to be stored independent of the molecular structure. The conventional coefficients are approximated by a linear combination of integrals which can be efficiently calculated and kept in the atomic orbital basis. The non-diagonal ansatz is given by:

$$\begin{aligned} A_{ij(B)}^{ab} &= \sum_{klg} B_{ij}^{klg} R_{kl}^{abg} \\ R_{kl}^{abg} &= \langle ab | \hat{f}_{12}^g | kl \rangle \end{aligned} \quad (3.2)$$

Alternatively, the even more compact diagonal ansatz is defined as:

$$A_{ij(B)}^{ab} = \sum_{klg} B_{ij}^{klg} R_{kl}^{abg} \delta_{ki} \delta_{lj} = \sum_g B_{ij}^g R_{ij}^{abg} \quad (3.3)$$

Since the relation between these two ansatzes is evident, we will proceed with the more general non-diagonal ansatz. In equation 3.2,  $\hat{f}_{12}^g$  are Gaussian type two-electron functions with different exponents and angular momenta. The coefficients  $B_{ij}^{klg}$  are optimized to minimize the energy.

## 3.2 Integral Evaluation

As mentioned above, our ansatz uses different angular momenta for the Gaussian geminals. Consequently, we have to elucidate the evaluation of the necessary integrals over atomic orbitals. The calculation of the integrals corresponding to s-type Gaussian geminals as an operator are given in the literature.[31] Let us define the following expressions for the basis functions leaving out all unnecessary parametric dependencies:

$$\begin{aligned} \mu_{(x_{1A})}^{iA} &= x_{1A}^{iA} \exp(-ax_{1A}^2) \\ \mu_{(r_{1A})}^{iA,jA,kA} &= \mu_{(x_{1A})}^{iA} \mu_{(y_{1A})}^{jA} \mu_{(z_{1A})}^{kA} \end{aligned} \quad (3.4)$$

Since the Gaussian functions decompose into products, we may integrate each Cartesian component separately.

$$I_{i_x}^{i_A i_B i_C i_D} = \langle \mu_{(x_{1A})}^{i_A} \mu_{(x_{2B})}^{i_B} | x_{12}^{i_x} \exp(-\alpha x_{12}^2) | \mu_{(x_{1C})}^{i_C} \mu_{(x_{2D})}^{i_D} \rangle \quad (3.5)$$

The entire integral is then given by:

$$I_{i_x j_y k_z}^{i_A i_B i_C i_D j_A j_B j_C j_D k_A k_B k_C k_D} = I_{i_x}^{i_A i_B i_C i_D} I_{j_y}^{j_A j_B j_C j_D} I_{k_z}^{k_A k_B k_C k_D} \quad (3.6)$$

For the calculation of p-type geminals with  $x_{12}^{i_x} = x_{12} = (x_1 - x_2)$ , we can now define:

$$\begin{aligned} (x_1 - x_2) &= (x_1 - x_A + x_A - x_2 + x_C - x_C) \\ &= (x_{1A} - x_{2C} + x_{AC}) \\ I_1^{i_A i_B i_C i_D} &= I_0^{(i_A+1) i_B i_C i_D} - I_0^{i_A i_B (i_C+1) i_D} + x_{AC} I_0^{i_A i_B i_C i_D} \end{aligned} \quad (3.7)$$

The same equations also hold for the other Cartesian components and higher angular momenta can be generated in a similar way. Furthermore, computation time can be saved by making use of the shell structure of the Gaussian geminals. If  $i_x$  is even, the operator of the Gaussian geminal is symmetric with respect to the permutation of electrons, while for odd  $i_x$  the operator is anti-symmetric. This has to be considered when symmetry is applied for the generation of integrals.

### 3.3 Energy Optimization

The energy functional for a whole class of pair correlation methods can be written in the following form:[2, 4, 3]

$$\begin{aligned} E_{(\mathbf{A})} &= E_0 + E_{Corr(\mathbf{A})} \\ &= E_0 + \frac{\langle \Psi_{(\mathbf{A})} | \hat{H} - E_0 | \Psi_{(\mathbf{A})} \rangle}{\langle \Phi | \Phi \rangle + c_D \langle \Psi_{D(\mathbf{A})} | \Psi_{D(\mathbf{A})} \rangle} \\ &= E_0 + N_{(\mathbf{A})}^{-1} \langle \Psi_{(\mathbf{A})} | \hat{H} - E_0 | \Psi_{(\mathbf{A})} \rangle \end{aligned} \quad (3.8)$$

Here,  $E_0$  denotes the energy of the reference determinant (Hartree Fock), and  $E_{Corr(\mathbf{A})}$  is the correlation energy of the corresponding energy functional. In equation 3.8 the choice of  $c_D = 1$  corresponds to CI(D),  $c_D = 0$  results in CEPA0(D) and  $c_D = \frac{1}{N_{el}}$  results in ACPF(D). Note that in the original definition of CEPA0 and ACPF the singles were also included.[4, 3] In our approach, all single excitations are omitted since we extend our ansatz to

orbital optimized pair correlation, where single excitations are included implicitly (cf. chapter 4). In this work, we will furthermore restrict the ansatz for the wave function to singlet states to simplify the working equations.[2]

To minimize the energy, it is convenient to expand the energy up to second order in a Taylor series around the coefficients  $\mathbf{A}_n$  of a given iteration  $n$ :

$$\begin{aligned} E_{(\mathbf{A}_n+\Delta\mathbf{A})} &\approx E_{(\mathbf{A}_n)}^{(0)} + (\Delta\mathbf{A})^T \mathbf{E}_{(\mathbf{A}_n)}^{(1)} + \frac{1}{2} \Delta\mathbf{A}^T \mathbf{E}_{(\mathbf{A}_n)}^{(2)} \Delta\mathbf{A} \\ E_{ijab}^{(1)} &= \frac{\partial}{\partial A_{ij}^{ab}} E_{(\mathbf{A}_n)} \\ E_{ijabklcd}^{(2)} &= \frac{\partial^2}{\partial A_{ij}^{ab} \partial A_{kl}^{cd}} E_{(\mathbf{A}_n)} \end{aligned} \quad (3.9)$$

The update  $\Delta\mathbf{A}$  of the coefficient matrix can be calculated by the conjugate gradient method.[20] Subsequently, the update is scaled by a step size  $\alpha$ :

$$\begin{aligned} \mathbf{E}_{(\mathbf{A}_n)}^{(2)} \Delta\mathbf{A} &= -\mathbf{E}_{(\mathbf{A}_n)}^{(1)} \\ \mathbf{A}_{n+1} &= \mathbf{A}_n + \alpha \Delta\mathbf{A} \end{aligned} \quad (3.10)$$

Here,  $\mathbf{E}^{(1)}$  denotes the electronic gradient vector which should be zero at the point of convergence.  $\mathbf{E}^{(2)}$  is the electronic Hessian matrix (or an approximation to it), and  $ijab$  is defined as a combined index. The variation of a specific coefficient  $A_{ij}^{ab}$  of the wave function is given by:

$$\frac{\partial}{\partial A_{ij}^{ab}} |\Psi_{(\mathbf{A})}\rangle = \sum_{klcd} \frac{\partial A_{kl}^{cd}}{\partial A_{ij}^{ab}} |\Phi_{kl}^{cd}\rangle = |\Phi_{ij}^{ab}\rangle \quad (3.11)$$

Analogously, for the transition to our contraction scheme the variation of a specific coefficient  $B_{ij}^{klg}$  is given by:

$$\frac{\partial}{\partial B_{ij}^{klg}} |\Psi_{(\mathbf{A})}\rangle = \sum_{ab} R_{kl}^{abg} |\Phi_{ij}^{ab}\rangle \quad (3.12)$$

The elements  $W_{ij(\mathbf{A})}^{ab}$  of the electronic gradient vector  $\mathbf{E}^{(1)}$  for conventional pair correlation methods are defined as:

$$W_{ij(\mathbf{A})}^{ab} = N_{(\mathbf{A})}^{-1} \langle \Phi_{ij}^{ab} | \hat{H} - E_0 - c_D E_{Corr} | \Psi_{(\mathbf{A})} \rangle \quad (3.13)$$

resulting in  $1/2 N_o(N_o + 1) N_v^2$  equations to be solved, while  $N_o$  and  $N_v$  denote the number of occupied and virtual orbitals, respectively. Using our new

ansatz of the wave function, this number can be significantly reduced to  $1/2N_o(N_o + 1)N_o^2N_g$  equations of the form:

$$\begin{aligned} X_{ij(\mathbf{A})}^{klg} &= N_{(\mathbf{A})}^{-1} \sum_{ab} R_{kl}^{abg} \langle \Phi_{ij}^{ab} | \hat{H} - E_0 - c_D E_{Corr} | \Psi_{(\mathbf{A})} \rangle \\ &= \sum_{ab} R_{kl}^{abg} W_{ij(\mathbf{A})}^{ab} \end{aligned} \quad (3.14)$$

Here,  $N_g$  denotes the number of geminals in the expansion. The storage requirements of the wave function and the electronic gradient are therefore drastically reduced. For the diagonal ansatz the storage requirements only scale with  $1/2N_o(N_o + 1)N_g$ .

For the calculation of the elements of the electronic gradient  $W_{ij}^{ab}$ , various efficient matrix oriented working equations exist[2], resulting in a straightforward transition to the contraction by the tensor elements of  $R_{kl}^{abg}$  (see section 3.4). The electronic Hessian  $\mathbf{E}^{(2)}$  is often approximated to be diagonal. Furthermore, these diagonal elements can be approximated by using differences of orbital energies:[21]

$$(\mathbf{E}_{(\mathbf{A}_n)}^{(2)} \Delta \mathbf{A})_{ijab} = N_{(\mathbf{A}_n)}^{-1} (\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j + c_{shift}) \Delta A_{ij}^{ab} \quad (3.15)$$

In the framework of our contracted pair correlation approach, we also use the approximated diagonal form of the conventional Hessian. In analogy to the gradient equations, this Hessian has to be contracted by integrals over Gaussian geminals. The following equations hold for the non-diagonal and the diagonal ansatz, respectively.

$$\begin{aligned} (\mathbf{E}_{(\mathbf{B}_n)}^{(2)} \Delta \mathbf{B})_{ijklg} &= N_{(\mathbf{A}_n)}^{-1} \sum_{ab} R_{kl}^{abg} (\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j) \Delta A_{ij}^{ab} \\ &\quad + c_{shift} B_{ij}^{klg} \\ (\mathbf{E}_{(\mathbf{B}_n)}^{(2)} \Delta \mathbf{B})_{ijg} &= N_{(\mathbf{A}_n)}^{-1} \sum_{ab} R_{ij}^{abg} (\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j) \Delta A_{ij}^{ab} \\ &\quad + c_{shift} B_{ij}^g \end{aligned} \quad (3.16)$$

The parameter  $c_{shift}$  is chosen to ensure that the Hessian is positive definite, which is a necessary condition for a faithful convergence. The step size  $\alpha$  is chosen to approximately minimize the energy. Of course this strategy is not unique, and there are other Quasi-Newton procedures such as the limited memory Broyden-Fletcher-Goldfarb-Shanno approach reported in the literature[32] which could reveal even better convergence. Since the resulting gradient and wave function dimensions are much smaller than in the conventional approaches, additional computation time and storage is saved for the linear algebra algorithms used to optimize the wave function.

### 3.4 Improved Scaling

A common and efficient way to construct the electronic gradient in pair correlation theory was defined by Meyer [33] and later refined by Pulay *et. al.* for singlet states.[2] The important idea is to construct the electronic gradient tensor  $\mathbf{W}_{ij(\mathbf{A})}^{AO}$  in the atomic orbital basis. The operator construction and the tensor contractions in the atomic orbital basis can be accelerated significantly by avoiding the calculation of negligibly small integrals using the Schwarz inequality and more advanced techniques such as the resolution of identity.[34] In order to optimize the energy the transformation from the atomic orbital to the molecular orbital basis is necessary and defined as:

$$\mathbf{W}_{ij(\mathbf{A})}^{MO} = \mathbf{C}^t \mathbf{W}_{ij(\mathbf{A})}^{AO} \mathbf{C} \quad (3.17)$$

Here, the orbital coefficient matrix  $\mathbf{C}$  is rectangular and contains the complete virtual space. The computationally demanding problem in the treatment of large molecules arises from the rapidly increasing size of the virtual space. This space is not easy to localize in general. One way to circumvent this problem is to define a pair specific virtual space spanned by the pair natural orbitals ( $\mathbf{C}_{ij}^{PNO}$ ).[35] These orbitals are relatively local and strongly reduce the number of non-zero elements of the electronic gradient. The transformation to the pair specific virtual space is defined by:[25]

$$\mathbf{W}_{ij(\mathbf{A})}^{PNO} = (\mathbf{C}_{ij}^{PNO})^t \mathbf{W}_{ij(\mathbf{A})}^{AO} \mathbf{C}_{ij}^{PNO} \quad (3.18)$$

The rectangular coefficient matrix  $\mathbf{C}_{ij}^{PNO}$  defines the virtual pair natural orbitals of the respective orbital pair. In our case, we sum over the complete virtual space making this transformation obsolete:

$$\begin{aligned} \sum_{ab} R_{kl}^{abg} W_{ij(\mathbf{A})}^{ab} &= Tr[\mathbf{R}_{klg}^{MO} \mathbf{W}_{ij(\mathbf{A})}^{MO}] \\ &= Tr[\mathbf{C}^t \mathbf{R}_{klg}^{AO} \mathbf{C} \mathbf{C}^t \mathbf{W}_{ij(\mathbf{A})}^{AO} \mathbf{C}] \\ &= Tr[\mathbf{D}_{virt}^{AO} \mathbf{R}_{klg}^{AO} \mathbf{D}_{virt}^{AO} \mathbf{W}_{ij(\mathbf{A})}^{AO}] \end{aligned} \quad (3.19)$$

Here, the coefficient matrix is again rectangular and contains the complete virtual space. The virtual density matrix is defined in complete analogy to the commonly used occupied density matrix:

$$\mathbf{D}_{virt}^{AO} = \mathbf{C} \mathbf{C}^t \quad (3.20)$$

This density matrix is invariant with respect to unitary transformations in the virtual subspace making pair natural orbitals obsolete for our ansatz. All tensor products can be evaluated in the local atomic orbital basis. The time

determining step reduces to sparse matrix multiplications, for which efficient algorithms are available. Note that the same reasoning as for PNO's, also holds for projected atomic orbitals (PAO's).[24] However, the projection is obsolete in our approach, since our working equations are defined in the intrinsically local atomic orbital basis.

Since we currently rely on our own integral package and pilot program, we are unfortunately unable to provide reasonable results with respect to scaling of computational time and storage requirements. Therefore, the scaling properties of our ansatz can only be demonstrated formally. Assuming that all necessary operators in the atomic orbital basis are computed on-the-fly, only the contracted gradient and the contraction coefficients need to be stored. This leads to drastic savings in storage requirements, namely about one or three orders of magnitude depending on the basis set size.[6] The computational scaling is more difficult to predict. In general, the computation of operators in the atomic orbital basis is not the time determining step in a local correlation treatment.[26] Since we only need operators in the atomic orbital basis, this might indicate a very good potential scaling behavior.

### 3.5 Results

To achieve reasonable accuracy with our contracted wave function we first need to define a reasonable expansion of the Gaussian geminal basis. In figure 3.1 we varied the exponents of two s- and p-type Gaussian geminals. All these calculations were performed using the averaged coupled pair functional (ACPF) correlation energy [3] of the methane molecule in the cc-pVTZ basis set [5]. The geometry was optimized with the MP2 method [36] and the 6-31+G\*\* basis set [37] using the Gaussian 09 program package.[38] Applying the exponents 0.75 and 0.1 we recover 95.6% of the conventional correlation energy of  $-0.219$  Hartree (see figure 3.1). This percentage can of course be increased by including more geminals in the expansion. Note that using only s-type Gaussian geminals resulted in a maximum percentage of 94,5% recovered, even when applying 24 geminals in the expansion. This indicates that higher angular momenta are necessary to recover a high percentage of the correlation energy.

In the following section, we discuss the extend of the correlation energy recovered by our new contraction scheme. All molecular geometries were again optimized on MP2 level using a 6-31+G\*\* basis. We use six s-type and six p-type Gaussian geminals with exponents ranging from 2.0 to 0.001 in an even tempered way for the diagonal ansatz and again two s- and two p-type geminals with the exponents 0.75 and 0.10 for the non-diagonal ansatz,

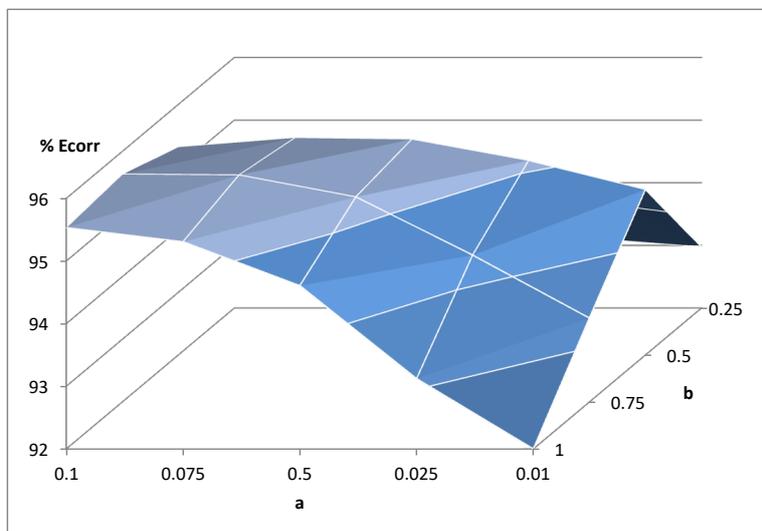


Figure 3.1: Percentage of the correlation energy recovered using two s- and p-type Gaussian geminals with respect to the variation of exponents  $a$  and  $b$ . Both angular momenta share the same exponent in each calculation. The correlation energy is computed with the ACPF[3] functional using methane and the cc-pVTZ basis set [5].

respectively.[6] Note that only double excitations are included into the correlation treatment due to the ansatz of the wave function. The occupied orbitals are localized using the Pipek-Mezey criterion.[23] For the diagonal ansatz, this influences the resulting energy while for the non-diagonal ansatz only the convergence is accelerated. This behavior is expected and can be explained by the unitary invariance of the respective ansatz with respect to orbital rotations within the space of occupied orbitals.[15]

In general, we recover more than 85% of the correlation energy by the diagonal ansatz (cf. Figure 3.2). However, significantly inferior results are obtained for larger systems and basis sets. This may be attributed to increasingly diffuse basis functions and longer ranges of interactions. In these cases, it might be necessary to add more diffuse Gaussian geminals and higher angular momenta, respectively.

The convergence of the diagonal ansatz to the limit of the full correlation energy was not investigated further since the unitary invariant non-diagonal ansatz already recovers more than 90% of the correlation energy in all cases and is virtually independent of the basis set and system size (cf. Figure 3.3). Therefore, the non-diagonal ansatz should be regarded as the superior

contraction scheme for the wave function.

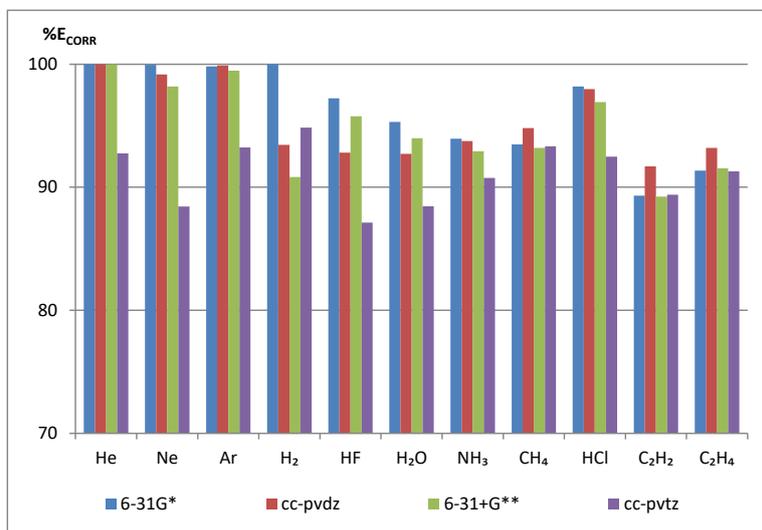


Figure 3.2: Percentage of the ACPF correlation energy recovered by the diagonal ansatz using six s- and six p-type Gaussian geminals.

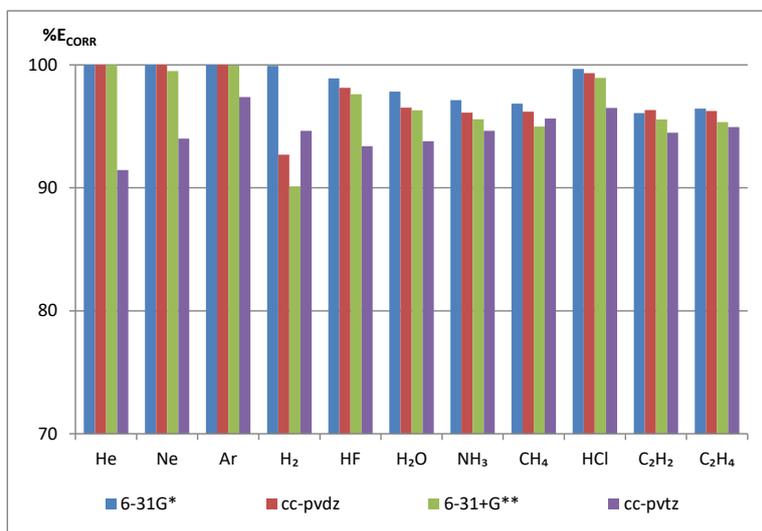


Figure 3.3: Percentage of the ACPF correlation energy recovered by the non-diagonal ansatz using two s- and two p-type Gaussian geminals.

For most problems of chemical interest, energy differences, such as reaction energies, are much more important than absolute energies. To describe energy differences sufficiently well, the error introduced by the contraction scheme has to be generally small or at least of a systematic nature. The diagonal ansatz already exhibits a dependency of the error with system size. As a consequence, the non-diagonal ansatz is far superior to the diagonal one (cf. Table 3.1). It recovers a sufficiently high percentage of the interaction energies with the given choice of Gaussian geminals. Actually, the energy variation by changing the correlation functional is often higher.

Table 3.1: Differences of contracted and conventional ACPF energies for some model reactions calculated in the 6-31+G\*\* basis.

	$\Delta E[kJ/mol]$		
	diagonal (6s6p)	non-diag. (2s2p)	conven- tional
$C_2H_2 + H_2 \rightarrow C_2H_4$	-238	-224	-220
$C_2H_4 + H_2 \rightarrow C_2H_6$	-186	-179	-168
$C_2H_6 + H_2 \rightarrow 2 CH_4$	-79	-73	-78
$N_2 + 3 H_2 \rightarrow 2 NH_3$	-179	-156	-153
RMS [conv.]	13	7	

Another important test for any new local correlation method is a faithful description of weak interactions. Along these lines, the dispersion interaction of two Neon atoms is investigated (cf. Figure 3.4). For the diagonal ansatz, we observe a repulsive potential energy curve leading to the conclusion that this ansatz cannot describe weak interactions. However, the non-diagonal ansatz recovers the energies of the conventional approach. In Table 3.2, results for two different choices of the Gaussian geminal basis are presented, revealing differences in the  $\mu$ Hartree regime. Thus, rather compact geminal expansions can be used in practice, provided that the exponents are carefully chosen. The deviation to coupled cluster calculations predominantly results from the choice of the pair correlation method as demonstrated in Figure 3.4. Extending our ansatz to coupled cluster doubles would significantly improve on these energies.[39]

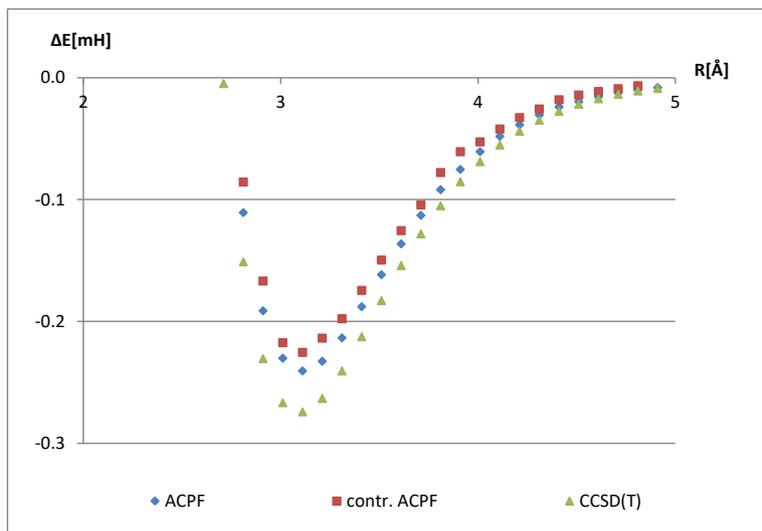


Figure 3.4: ACPF interaction energies of two neon atoms calculated with the 6-31+G\*\* basis set using six s- and six p-type Gaussian geminals

Table 3.2: Contracted ACPF interaction energies of the neon dimer at a distance of 3.11 Å using the 6-31+G\*\* basis. The conventional interaction ACPF energy is  $-0.241 mH$

	$\Delta E [mH]$ non-diag.
6s6p	-0.225
2s2p + diff.[6]	-0.232

### 3.6 Summary and Outlook

We have presented a new way for approaching local correlation methods introducing a novel contraction scheme which drastically reduces the storage requirements and recovers a large percentage of the correlation energy of a given pair correlation method. The error introduced is of a systematic nature such that energy differences are also reproduced with high accuracy. An important future development will be the extension of our ansatz to coupled cluster methods and thereby increase the accuracy of the correlation method itself. In principle, it is possible to extend the above ideas to lower or higher excitation types. Single excitations can be included by one electron

Gaussians or by summation over all occupied orbitals of one-electron in the integral expression of two electron Gaussians, like in the Coulomb or exchange operators. Triple excitations are available by forming products of the two electron Gaussian functions to build three electron Gaussian functions. This can of course also be extended up to the excitation level of full CI. Locality of correlation effects can further be exploited by describing important pairs with longer geminal expansions than the remainder. Further modifications are possible by changing the functional form of  $r_{12}$ . An inclusion of explicit correlation represents a natural extension and will be presented in chapter 5. We will also describe the inclusion of orbital optimization to treat multi-reference problems (cf. chapter 4).

# Chapter 4

## Orbital Optimized Contracted Pair Correlation Methods

The formation or breaking of chemical bonds represents a ubiquitous phenomenon in chemistry and often requires a multi-reference formulation of the wave function. We present the extension of our contraction scheme to orbital optimized pair correlation methods. These methods are able to describe the dissociation of an arbitrary number of single bonds within a molecule. A big advantage of these methods is the direct inclusion of dynamic correlation effects into a multi-reference calculation. Additionally, we can improve on the accuracy of molecular properties in general as shown by Bozkaya *et. al.* for OCEPA(0).[7] Using our contraction scheme, we included locality into the working equations and reduced the storage requirements by up to three orders of magnitude.[6] This may open a way to tackle large molecules with this class of methods.

### 4.1 Ansatz for the Wave Function

The basic ansatz for an orbital optimized pair correlation method can be written in the following form:

$$|\Psi_{(\mathbf{A},\boldsymbol{\kappa})}\rangle = |\Phi_{(\boldsymbol{\kappa})}\rangle + \sum_{klcd} A_{kl}^{cd} |\Phi_{kl(\boldsymbol{\kappa})}^{cd}\rangle \quad (4.1)$$

The coefficient tensor  $\mathbf{A}$  contains the standard configuration interaction coefficients  $A_{kl}^{cd}$  and the matrix  $\boldsymbol{\kappa}$  describes the orbital transformations. In the contracted pair correlation methods the coefficients are expressed by a sum over integrals and therefore also depend on the orbital rotation parameters

(cf. equation 4.2).

$$\begin{aligned}
A_{kl(\boldsymbol{\kappa}, \mathbf{B})}^{cd} &= \sum_{mng} B_{kl}^{mng} R_{mn(\boldsymbol{\kappa})}^{cdg} \\
R_{mn(\boldsymbol{\kappa})}^{cdg} &= \langle c_{(\boldsymbol{\kappa})} d_{(\boldsymbol{\kappa})} | \hat{f}_{12}^g | m_{(\boldsymbol{\kappa})} n_{(\boldsymbol{\kappa})} \rangle
\end{aligned} \tag{4.2}$$

## 4.2 Energy Optimization

Since the wave function now depends on the orbital rotation parameter  $\boldsymbol{\kappa}$ , the energy functional in equation 3.8 has to be generalized: [2]

$$\begin{aligned}
E_{(\mathbf{A}, \boldsymbol{\kappa})} &= E_{0(\boldsymbol{\kappa})} + E_{Corr(\mathbf{A}, \boldsymbol{\kappa})} \\
&= E_{0(\boldsymbol{\kappa})} + \frac{\langle \Psi_{(\mathbf{A}, \boldsymbol{\kappa})} | \hat{H} - E_{0(\boldsymbol{\kappa})} | \Psi_{(\mathbf{A}, \boldsymbol{\kappa})} \rangle}{\langle \Phi_{(\boldsymbol{\kappa})} | \Phi_{(\boldsymbol{\kappa})} \rangle + c_D \langle \Psi_{D(\mathbf{A}, \boldsymbol{\kappa})} | \Psi_{D(\mathbf{A}, \boldsymbol{\kappa})} \rangle} \\
&= E_{0(\boldsymbol{\kappa})} + N_{(\mathbf{A})}^{-1} \langle \Psi_{(\mathbf{A}, \boldsymbol{\kappa})} | \hat{H} - E_{0(\boldsymbol{\kappa})} | \Psi_{(\mathbf{A}, \boldsymbol{\kappa})} \rangle
\end{aligned} \tag{4.3}$$

The dependency of the wave function on orbital rotation parameters can be expressed by the following general expression which holds for any determinant or linear combination of determinants  $|\Psi\rangle$ . [18, 40]

$$\begin{aligned}
|\Psi_{(\boldsymbol{\kappa}, \mathbf{A})}\rangle &= \exp(-\hat{\kappa}) |\Psi_{(\mathbf{A})}\rangle \\
\hat{\kappa} &= \sum_{ai} \kappa_{ai} (E_{ai} - E_{ia})
\end{aligned} \tag{4.4}$$

Here,  $\kappa_{ai}$  are the matrix elements of the anti-symmetric matrix  $\boldsymbol{\kappa}$ , whose matrix exponential defines the unitary transformation matrix  $\mathbf{U} = \exp(-\boldsymbol{\kappa})$ . The singlet excitation operator  $E_{pq}$  is defined as  $a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta}$ . The symbols  $a^\dagger$  and  $a$  are the creation and annihilation operators for electrons with  $\alpha$  or  $\beta$  spin in the spatial orbitals  $p$  and  $q$ , respectively. The dependency of integral expressions on  $\boldsymbol{\kappa}$  is then given by:

$$\begin{aligned}
|p(\boldsymbol{\kappa})\rangle &= \sum_q |q\rangle U_{qp(\boldsymbol{\kappa})} = \sum_q |q\rangle [\exp(-\boldsymbol{\kappa})]_{qp} \\
R_{kl(\boldsymbol{\kappa})}^{cd} &= \langle c_{(\boldsymbol{\kappa})} d_{(\boldsymbol{\kappa})} | \hat{f}_{12}^g | k_{(\boldsymbol{\kappa})} l_{(\boldsymbol{\kappa})} \rangle \\
&= \sum_{pqrs} \langle pq | \hat{f}_{12}^g | rs \rangle U_{pc(\boldsymbol{\kappa})} U_{qd(\boldsymbol{\kappa})} U_{rk(\boldsymbol{\kappa})} U_{sl(\boldsymbol{\kappa})}
\end{aligned} \tag{4.5}$$

We may now define the reference and the correlation energy with respect to orbital and coefficient dependencies as:

$$\begin{aligned}
E_{0(\boldsymbol{\kappa})} &= \langle \Phi_{(\boldsymbol{\kappa})} | \hat{H} | \Phi_{(\boldsymbol{\kappa})} \rangle \\
&= \langle \Phi | \exp(\hat{\kappa}) \hat{H} \exp(-\hat{\kappa}) | \Phi \rangle \\
&= \langle \Phi | \hat{H}_{(\boldsymbol{\kappa})} | \Phi \rangle \\
E_{Corr(\boldsymbol{\kappa}, \mathbf{A})} &= \frac{\langle \Psi_{(\boldsymbol{\kappa}, \mathbf{A})} | \hat{H} - E_{0(\boldsymbol{\kappa})} | \Psi_{(\boldsymbol{\kappa}, \mathbf{A})} \rangle}{\langle \Phi_{(\boldsymbol{\kappa})} | \Phi_{(\boldsymbol{\kappa})} \rangle + c_D \langle \Psi_{D(\boldsymbol{\kappa}, \mathbf{A})} | \Psi_{D(\boldsymbol{\kappa}, \mathbf{A})} \rangle} \\
&= \frac{\langle \Psi_{(\mathbf{A})} | \hat{H}_{(\boldsymbol{\kappa})} - E_{0(\boldsymbol{\kappa})} | \Psi_{(\mathbf{A})} \rangle}{\langle \Phi | \Phi \rangle + c_D \langle \Psi_{D(\mathbf{A})} | \Psi_{D(\mathbf{A})} \rangle} \\
&= N_{(\mathbf{A})}^{-1} \langle \Psi_{(\mathbf{A})} | \hat{H}_{(\boldsymbol{\kappa})} - E_{0(\boldsymbol{\kappa})} | \Psi_{(\mathbf{A})} \rangle \tag{4.6}
\end{aligned}$$

We left out the dependency of  $\mathbf{A}$  on orbital transformations in the above expression for readability. The optimization of the coefficients was already discussed in detail in chapter 3. An orbital optimization is slightly more difficult than the optimization of the CI-coefficients due to the exponential dependence of the wave function on the variational parameters. Along these lines, derivatives of the exponential result in the following non-linear equations:

$$\begin{aligned}
\frac{d}{d\kappa_{ai}} \exp(-\boldsymbol{\kappa}) &= -(\boldsymbol{\delta}_{ai} - \boldsymbol{\delta}_{ia}) \exp(-\boldsymbol{\kappa}) \\
\frac{d}{d\kappa_{ai}} \exp(-\hat{\kappa}) &= -(E_{ai} - E_{ia}) \exp(-\hat{\kappa}) \\
\frac{d}{d\kappa_{ai}} \hat{H}(\boldsymbol{\kappa}) &= [(E_{ai} - E_{ia}), \hat{H}(\boldsymbol{\kappa})] \tag{4.7}
\end{aligned}$$

Here,  $\boldsymbol{\delta}_{pq}$  is a zero matrix with element  $pq$  set to one. To simplify the non-linear equations, we expand the energy up to second order around  $\boldsymbol{\kappa} = 0$  in a McLaurin series for an arbitrary iteration  $n$ :

$$\begin{aligned}
E_{n(\Delta\boldsymbol{\kappa})} &\approx E_{n(0)}^{(0)} + (\Delta\boldsymbol{\kappa})^T \mathbf{E}_{n(0)}^{(1)} + \frac{1}{2} \Delta\boldsymbol{\kappa}^T \mathbf{E}_{n(0)}^{(2)} \Delta\boldsymbol{\kappa} \\
E_{ai}^{(1)} &= \lim_{\boldsymbol{\kappa} \rightarrow 0} \frac{d}{d\kappa_{ai}} E_{n(\boldsymbol{\kappa})} \\
E_{aibj}^{(2)} &= \lim_{\boldsymbol{\kappa} \rightarrow 0} \frac{d^2}{d\kappa_{ai} d\kappa_{bj}} E_{n(\boldsymbol{\kappa})} \tag{4.8}
\end{aligned}$$

The update  $\Delta\boldsymbol{\kappa}$  of the transformation matrix can be calculated by the conjugate gradient method. [20] Subsequently, the update is scaled by a step size parameter  $\alpha$ :

$$\begin{aligned}
\mathbf{E}_{n(0)}^{(2)} \Delta\boldsymbol{\kappa} &= -\mathbf{E}_{n(0)}^{(1)} \\
\boldsymbol{\kappa}_{n+1} &= \alpha \Delta\boldsymbol{\kappa} \tag{4.9}
\end{aligned}$$

Here,  $\mathbf{E}_{n(0)}^{(1)}$  denotes the electronic gradient vector constructed from the orbitals belonging to the iteration  $n$  and  $\mathbf{E}_{n(0)}^{(2)}$  is the electronic Hessian matrix (or an approximation to it), such that  $ai$  is defined as a combined index. We now only need to define  $E_{ai}^{(1)}$  within the limit  $\boldsymbol{\kappa} \rightarrow 0$ . A very general variation of the energy using the product rule is given by:

$$\begin{aligned}
\delta E_{(\boldsymbol{\kappa}, \mathbf{A})} &= \delta E_{0(\boldsymbol{\kappa})} + \delta E_{Corr(\boldsymbol{\kappa}, \mathbf{A})} \\
&= \delta E_{0(\boldsymbol{\kappa})} + N_{(\mathbf{A})}^{-1} (\delta \langle \Psi_{(\mathbf{A})} | \hat{H}_{(\boldsymbol{\kappa})} - E_{0(\boldsymbol{\kappa})} | \Psi_{(\mathbf{A})} \rangle \\
&\quad - c_D E_{Corr(\boldsymbol{\kappa}, \mathbf{A})} \delta \langle \Psi_{D(\mathbf{A})} | \Psi_{D(\mathbf{A})} \rangle)
\end{aligned} \tag{4.10}$$

Accordingly, the intermediate quantities are defined as follows:

$$\begin{aligned}
\lim_{\boldsymbol{\kappa} \rightarrow 0} \frac{d}{d\kappa_{ai}} \hat{H}_{(\boldsymbol{\kappa})} &= [(E_{ai} - E_{ia}), \hat{H}] \\
&= \hat{H}_{ai} \\
\lim_{\boldsymbol{\kappa} \rightarrow 0} \frac{d}{d\kappa_{ai}} \langle \Phi | \hat{H}_{(\boldsymbol{\kappa})} | \Phi \rangle &= \langle \Phi | \hat{H}_{ai} | \Phi \rangle \\
&= F_{ai}^{\Phi} \\
\lim_{\boldsymbol{\kappa} \rightarrow 0} \frac{d}{d\kappa_{ai}} \langle \Psi_{(\mathbf{A})} | \hat{H}_{(\boldsymbol{\kappa})} | \Psi_{(\mathbf{A})} \rangle &= \langle \Psi_{(\mathbf{A})} | \hat{H}_{ai} | \Psi_{(\mathbf{A})} \rangle \\
&= F_{ai(\mathbf{A})}^{\Psi} \\
\lim_{\boldsymbol{\kappa} \rightarrow 0} \frac{d}{d\kappa_{ai}} U_{pq(\boldsymbol{\kappa})} &= [-(\boldsymbol{\delta}_{ai} - \boldsymbol{\delta}_{ia})]_{pq} \\
&= -\delta_{pa} \delta_{qi} + \delta_{pi} \delta_{qa} \\
\lim_{\boldsymbol{\kappa} \rightarrow 0} \frac{d}{d\kappa_{ai}} R_{kl(\boldsymbol{\kappa})}^{cdg} &= -R_{al}^{cdg} \delta_{ki} - R_{ka}^{cdg} \delta_{li} \\
&\quad + R_{kl}^{idg} \delta_{ca} + R_{kl}^{cig} \delta_{da}
\end{aligned} \tag{4.11}$$

Using these definitions and the general energy variation we get:

$$\begin{aligned}
E_{ai}^{(1)} &= F_{ai}^{\Phi} + N_{(\mathbf{A})}^{-1} (F_{ai(\mathbf{A})}^{\Psi} - F_{ai}^{\Phi} (1 + \langle \Psi_{D(\mathbf{A})} | \Psi_{D(\mathbf{A})} \rangle)) \\
&\quad + 2 \langle \Psi_{(\delta \mathbf{A})} | \hat{H}_{(\boldsymbol{\kappa})} - E_{0(\boldsymbol{\kappa})} | \Psi_{(\mathbf{A})} \rangle
\end{aligned} \tag{4.12}$$

For the last part of equation 4.12 we may additionally define the matrices:

$$\begin{aligned}
Y_{ai(\mathbf{A})} &= 2 \sum_{klc} \left( \sum_{mng} B_{lk}^{mng} R_{mn}^{icg} + B_{kl}^{mng} R_{mn}^{cig} \right) W_{kl(\mathbf{A})}^{ac} \\
&= 4 \sum_{klc} A_{kl}^{ci} W_{kl(\mathbf{A})}^{ca} \\
Z_{ai(\mathbf{A})} &= 2 \sum_{klcd} \sum_{mg} (B_{kl}^{img} R_{am}^{cdg} + B_{kl}^{mig} R_{ma}^{cdg}) W_{kl(\mathbf{A})}^{cd} \\
&= 4 \sum_{klmg} B_{kl}^{img} X_{kl(\mathbf{A})}^{amg}
\end{aligned} \tag{4.13}$$

Finally, the following working equations are obtained:

$$\begin{aligned}
E_{ai}^{(1)} &= F_{ai}^{\Phi} + N_{(\mathbf{A})}^{-1} (F_{ai}^{\Psi} - F_{ai}^{\Phi} (1 + \langle \Psi_{D(\mathbf{A})} | \Psi_{D(\mathbf{A})} \rangle)) \\
&+ Y_{ai(\mathbf{A})} - Z_{ai(\mathbf{A})}
\end{aligned} \tag{4.14}$$

Note that the terms depending on the generalized and the normal Fock matrix ( $F_{ai}^{\Psi}$  and  $F_{ai}^{\Phi}$ ) belong to the conventional orbital optimized pair correlation methods. There are efficient working equations available for this part of equation 4.14.[7] The other terms can also be calculated efficiently by the equations resulting from the coefficient optimization (cf. chapter 3).

To optimize the energy, the electronic Hessian has to be defined and is currently approximated by the Hessian of the Hartree-Fock ansatz.[7]

$$\begin{aligned}
(\mathbf{E}_{n(0)}^{(2)} \Delta \boldsymbol{\kappa})_{ai} &= 2 \sum_{bj} (\delta_{ij} F_{ab}^{\Phi} - \delta_{ab} F_{ij}^{\Phi} \\
&+ (2G_{ij}^{ab} - G_{ij}^{ba}) - (2G_{ib}^{ja} - G_{ib}^{aj}) \\
&+ \delta_{ij} \delta_{ab} c_{shift}) \kappa_{bj} \\
G_{pq}^{rs} &= \langle pq | \hat{g}_{12} | rs \rangle
\end{aligned} \tag{4.15}$$

The parameter  $c_{shift}$  and the step size  $\alpha$  are chosen as explained in section 3.3.

To improve the storage requirements of the contracted orbital optimized pair correlation methods as far as possible, we will have to calculate operators in the atomic orbital basis on-the-fly. Any sum over virtual orbitals can again be replaced by tensor products in the atomic orbital basis (cf. section 3.4 for further details).

## 4.3 Results

In order to investigate the ability of orbital optimized pair correlation methods to describe static correlation, the homolytic bond breaking of small

molecules is studied.[10] This process needs two reference determinants for single bond breaking, namely the optimized ground state determinant and a doubly excited determinant, where the orbital describing the bond is replaced by the anti-bonding one. These determinants obviously occur in our methods, since we include all doubly excited determinants in our expansion. The remaining question is if the contraction scheme is also able to reproduce the correct dissociation energies. In the following investigations, we use two s- and p-type Gaussian geminals with the exponents 0.75 and 0.10, respectively (cf. chapter 3).[6] All the following calculations were performed using the 6-31+G\*\* basis set[37] and the frozen core approximation was applied by neglecting all electrons of the inner shells.

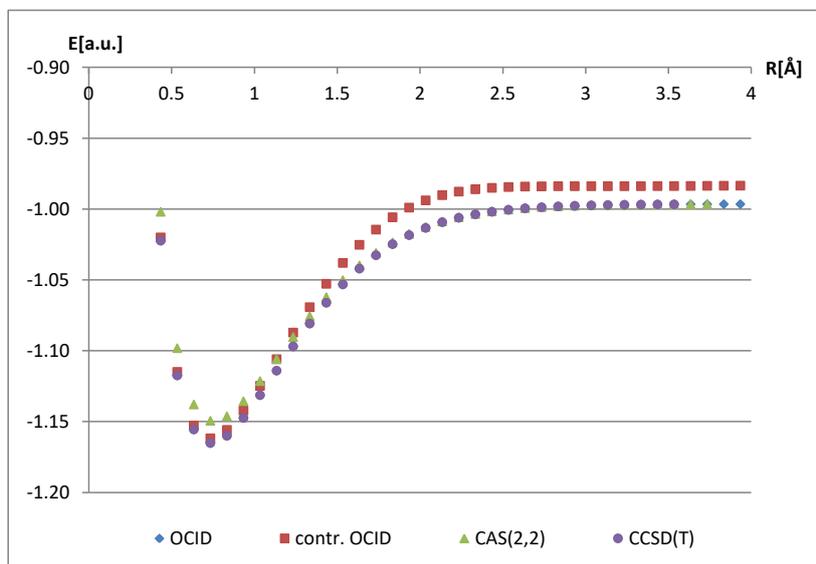


Figure 4.1: Bond dissociation of the hydrogen molecule

The most simple example of a homolytic bond breaking occurs in the hydrogen molecule (cf. Figure 4.1). It is demonstrated, that the orbital optimized configuration interaction doubles (OCID) perfectly reproduces the correct dissociation limit and also adds a small correlation contribution which was not recovered by the minimal CAS(2,2) ansatz.[41] The small difference between the contracted and the conventional ansatz for larger distances is due to the rather steep geminal exponents and can be further reduced by adding more diffuse Gaussian geminals.

Another prominent test example is the dissociation of lithium fluoride. According to the results demonstrated in Figure 4.2, the contracted and

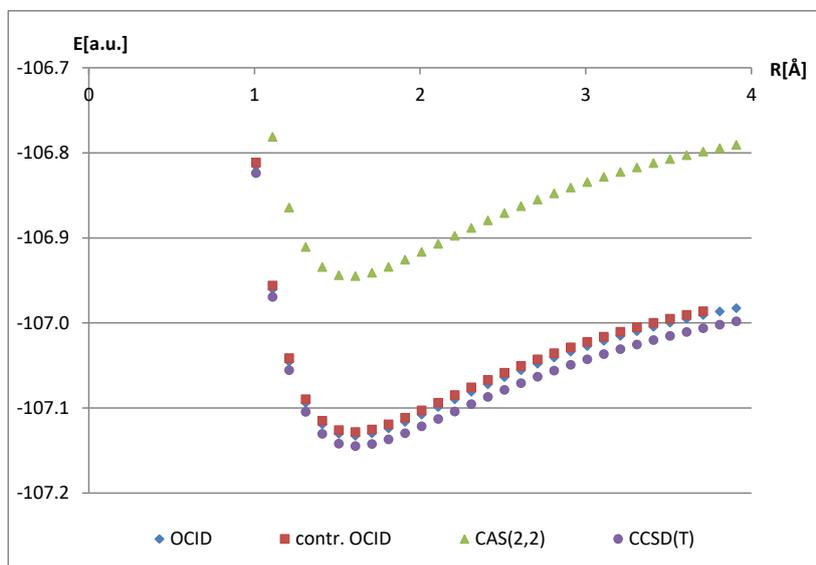


Figure 4.2: Bond dissociation of the lithium fluoride molecule

the conventional orbital optimized methods perfectly reproduce the topology of the CAS(2,2) calculation, while including correlation effects with similar accuracy as CCSD(T).[42] Note that the CCSD(T) method tends to diverge for large bond distances since a single determinant represents a poor reference due to pronounced static correlation.

The dissociation of the nitrogen molecule is even more challenging (cf. Figure 4.3), since additionally two quadruply excited determinants and one hexuply excited determinant are necessary to describe this process correctly. Consequently, the orbital optimized pair correlation method must converge into a higher excited state as compared to the CAS(6,6) calculation. This behavior is expected and should not be regarded as a general failure of our approach. Extending the orbital optimized contraction scheme to higher excitation levels will allow a correct contracted complete active space. A conventional or contracted CASSCF[41] like wave function can therefore be constructed and would lead to multi-reference approaches directly including dynamic correlation. In the geometric minimum, we recover a large fraction of the correlation energy given by the coupled cluster method which quickly diverges for larger bond distances.

Note that so far the size-extensive orbital optimized pair correlation methods OCEPA0[7] and OACPF diverged when increasing the bond distance in the test molecules above. Orbital optimized coupled cluster seems to per-

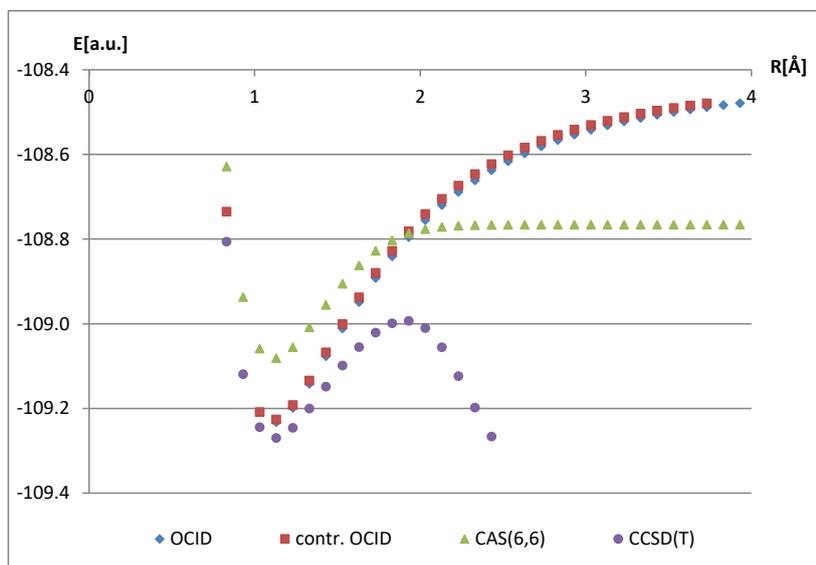


Figure 4.3: Bond dissociation of the nitrogen molecule

form more stable in these cases and should be tested in future work.[10] Future developments should also focus on including additional excitations in form of a complete active space.[41] This will lead to methods which are able to describe any multi-reference problem while directly including dynamic correlation in a locally contracted form. Note that the description of static correlation is different from size-extensive multi-reference treatments like multi-configurational CEPA.[43] In this method, the static correlation is described by a CASSCF reference function, and only the dynamic correlation is described by a CEPA energy functional.

We may now investigate if reaction energies can also be improved by orbital optimization. For this purpose, we first compare the uncontracted definition of orbital optimized CEPA0 to CCSD(T). As shown in table 4.1, all reaction energies are closer to this reference energy indicating that including orbital optimization is indeed a reasonable approach. Note that Bozkaya *et al.* already found this behavior for a much larger test set of reactions.[7] According to our previous findings, further investigations reveal to which extent accuracy is affected by the contraction scheme (cf. table 4.2). Again, orbital optimization improves on the accuracy. Further improvements are in principle possible by including more Gaussian geminals into the expansion of the wave function.

Table 4.1: Conventional CEPA0 and OCEPA0 electronic reaction energies for some model reactions calculated in the 6-31+G\*\* basis

	$\Delta E[kJ/mol]$		
	CEPA0	OCEPA0	CCSD(T)
$C_2H_2 + H_2 \rightarrow C_2H_4$ [6]	-218	-217	-212
$C_2H_4 + H_2 \rightarrow C_2H_6$	-165	-161	-163
$C_2H_6 + H_2 \rightarrow 2 CH_4$	-77	-76	-75
$N_2 + 3 H_2 \rightarrow 2 NH_3$	-146	-139	-136
RMS [CCSD(T) [38]]	6	3	—

Table 4.2: Contracted CEPA0 and OCEPA0 electronic reaction energies for some model reactions calculated using the 6-31+G\*\* basis

	$\Delta E[kJ/mol]$	
	contr. CEPA0	contr. OCEPA0
$C_2H_2 + H_2 \rightarrow C_2H_4$ [6]	-222	-219
$C_2H_4 + H_2 \rightarrow C_2H_6$	-178	-175
$C_2H_6 + H_2 \rightarrow 2 CH_4$	-72	-71
$N_2 + 3 H_2 \rightarrow 2 NH_3$	-150	-143
RMS [CCSD(T) [38]]	12	8

## 4.4 Summary and Outlook

Orbital optimized pair correlation methods lead to very accurate molecular properties, such as energy differences and potential energy surfaces. We presented a contraction scheme, which includes the use of locality into orbital optimized pair correlation methods. This may pave the way towards the application of these methods to rather large molecules. The effects of the contraction on the accuracy are negligibly small, even when static correlation plays a major role for the system under investigation.

## Chapter 5

# Explicitly Correlated Orbital Optimized Contracted Pair Correlation Methods

Explicitly correlated wave functions are part of extensive research since they offer a way to drastically improve the otherwise poor convergence of the correlation energy with respect to the size of the one-particle basis. In chapter 3, we presented a way to efficiently contract pair correlation wave functions, resulting in very compact working equations and thereby reducing the storage requirements by up to three orders of magnitude.[6] The extension to orbital optimization was also shown. The contraction coefficients in these ansatzes are given by integrals over Gaussian type geminals. Since these integrals are the basis for explicitly correlated wave functions, the resulting extension is obvious and will be presented here. Improved convergence towards the basis set limit can be shown for the contracted pair correlation methods.

### 5.1 Ansatz for the Wave Function

In our previous ansatz [6] the configuration interaction coefficients were represented by a sum of integrals over Gaussian geminals. To introduce explicit correlation, we will rewrite these equations implying a complete one-particle basis. We will reintroduce the basis used for the actual computations, when-

ever it is reasonable.

$$\begin{aligned}
|\Psi_{(\mathbf{A})}\rangle &= |\Phi\rangle + \sum_{ijab} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle \rightarrow |\Phi\rangle + \sum_{ijAB} A_{ij}^{AB} |\Phi_{ij}^{AB}\rangle \\
A_{ij(\mathbf{B})}^{AB} &= \sum_{klg} B_{ij}^{klg} R_{kl}^{ABg} \\
R_{kl}^{ABg} &= \langle AB | \hat{f}_{12}^g | kl \rangle
\end{aligned} \tag{5.1}$$

The indices  $a, b$  label the virtual space of the basis used for the calculation, while  $A, B$  label the virtual space of the complete one-particle basis. Moreover,  $|\Phi\rangle$  is the reference determinant and the sum contains all double excitations from the occupied to the complete virtual space ( $ij \rightarrow AB$ ). In order to introduce the way of including explicit correlation, it is convenient to reorder the summation of the wave function.

$$\begin{aligned}
|\Psi_{(\mathbf{B})}\rangle &= |\Phi\rangle + \sum_{ijklg} B_{ij}^{klg} |\Xi_{ij}^{klg}\rangle \\
|\Xi_{ij}^{klg}\rangle &= \sum_{AB} R_{kl}^{ABg} |\Phi_{ij}^{AB}\rangle
\end{aligned} \tag{5.2}$$

Let us first introduce the projection operators:

$$\begin{aligned}
\hat{\mathbf{1}}_{(1)} &= \sum_P^\infty |P_{(1)}\rangle \langle P_{(1)}| ; \hat{P}_{(1)} = \sum_p^{n_{bas}} |p_{(1)}\rangle \langle p_{(1)}| \\
\hat{\mathbf{O}}_{(1)} &= \sum_i^{n_{occ}} |i_{(1)}\rangle \langle i_{(1)}| ; \hat{V}_{(1)} = \sum_a^{n_{virt}} |a_{(1)}\rangle \langle a_{(1)}|
\end{aligned} \tag{5.3}$$

Note that the indices  $PQRS$  denote the complete one-particle basis,  $pqrs$  denote the computational basis,  $ijkl$  denote the occupied space and  $abcd$  label the virtual space. We can now combine these projection operators in the following form.[44]

$$\begin{aligned}
\hat{Q}_{12} &= (\hat{\mathbf{1}}_{(1)} - \hat{\mathbf{O}}_{(1)})(\hat{\mathbf{1}}_{(2)} - \hat{\mathbf{O}}_{(2)})(\hat{\mathbf{1}}_{(1)}\hat{\mathbf{1}}_{(2)} - \hat{V}_{(1)}\hat{V}_{(2)}) \\
&\approx (\hat{\mathbf{1}}_{(1)}\hat{\mathbf{1}}_{(2)} - \hat{P}_{(1)}\hat{P}_{(2)})
\end{aligned} \tag{5.4}$$

Note, that the last approximation only holds if we assume the one-particle basis used for calculations to be complete for integrals including more than two electrons ( $\hat{\mathbf{1}}_{(1)}\hat{\mathbf{O}}_{(2)} = \hat{P}_{(1)}\hat{\mathbf{O}}_{(2)}$ ). This approximation allows for a direct replacement of summations over the complete virtual space  $A, B$  by summations over the complete one-particle basis  $P, Q$  and the computational basis  $p, q$ . This last ansatz for the projection operator is not used and normally refined by introducing an auxiliary basis.[44, 12, 13, 14] Nevertheless,

we will demonstrate in the results section that quite accurate results can be obtained by introducing an additional approximation for the one-electron integrals. Along these lines, we can use equation 5.4 to modify our present ansatz in the following way:

$$\begin{aligned}
|\Xi_{ij}^{klg}\rangle &= \sum_{AB} \langle AB | (\hat{V}_{(1)} \hat{V}_{(2)} + \frac{1}{2} \hat{Q}_{12}) \hat{f}_{12}^g | kl \rangle |\Phi_{ij}^{AB}\rangle \\
&= \sum_{ab} R_{kl}^{abg} |\Phi_{ij}^{ab}\rangle \\
&+ \frac{1}{2} \left( \sum_{AB} R_{kl}^{ABg} |\Phi_{ij}^{AB}\rangle - \sum_{ab} R_{kl}^{abg} |\Phi_{ij}^{ab}\rangle \right) \\
&\approx \sum_{ab} R_{kl}^{abg} |\Phi_{kl}^{ab}\rangle \\
&+ \frac{1}{2} \left( \sum_{PQ} R_{kl}^{PQg} |\Phi_{ij}^{PQ}\rangle - \sum_{pq} R_{kl}^{pqg} |\Phi_{ij}^{pq}\rangle \right) \tag{5.5}
\end{aligned}$$

The expressions in parenthesis can be interpreted as a correction to the wave function. This correction is a manifestation of the incompleteness of the one-particle basis. If the basis becomes complete, i.e.  $\{p\} \rightarrow \{P\}$ , the correction vanishes.

We have demonstrated in chapter 3 that it is necessary to include p-type Gaussian geminals into the contracted wave function to achieve sufficient accuracy.[6] These functions do not improve the description of the Coulomb cusp due to their nodal structure, and the corresponding explicitly correlated integrals would be expensive and complicated to calculate. The modification of the ansatz for the wave function is therefore only applied for the subset of s-type geminals denoted by  $(\mathbf{s})$  in the functions  $|\Xi_{ij}^{klg(\mathbf{s})}\rangle$ . The p-type geminals denoted by  $(\mathbf{p})$  will be augmented with the projection operator  $\hat{V}_{(1)} \hat{V}_{(2)}$  resulting in the conventional contribution.

$$\begin{aligned}
|\Xi_{ij}^{klg(\mathbf{s})}\rangle &= \sum_{ab} R_{kl}^{abg} |\Phi_{kl}^{ab}\rangle \\
&+ \frac{1}{2} \left( \sum_{PQ} R_{kl}^{PQg} |\Phi_{ij}^{PQ}\rangle - \sum_{pq} R_{kl}^{pqg} |\Phi_{ij}^{pq}\rangle \right) \\
|\Xi_{ij}^{klg(\mathbf{p})}\rangle &= \sum_{ab} R_{kl}^{abg} |\Phi_{kl}^{ab}\rangle \tag{5.6}
\end{aligned}$$

In principle, a more general ansatz for the wave function could be defined as well.

$$|\Psi_{(\mathbf{B})}\rangle = |\Phi\rangle + \sum_{ijab} A_{ij(\mathbf{B})}^{ab} |\Phi_{ij}^{ab}\rangle + \sum_{ijAB} A_{ij(\mathbf{B})}^{AB} |\Phi_{ij}^{AB}\rangle \tag{5.7}$$

Here, the coefficients  $A_{ij(\mathbf{B})}^{ab}$  and  $A_{ij(\mathbf{B})}^{AB}$  are still defined by contraction. This ansatz is more reminiscent to explicitly correlated wave functions from literature, except for the contracted representation of  $A_{ij(\mathbf{B})}^{ab}$ . [12, 13, 14] Since in our approach both sets of coefficients are contracted by the same integral types it seems reasonable to summarize them. Note that the difference between these two ansatzes is given by one set of contraction coefficients  $B_{ij}^{klg}$  in equation 5.6 and two sets of contraction coefficients for equation 5.7, since the conventional and the explicitly correlated contraction are not identical.

Our ansatz for the wave function can be understood as a combination and extension of the studies reported previously by Höfener *et. al.* [45] and Valeev [46]. Höfener *et. al.* [45] applied a similar ansatz to the coupled cluster approximation CC2. In this study, the conventional amplitudes are replaced by contraction, but the linear combination of the Gaussian geminals was fixed and only s-type Gaussian geminals were used.

$$\begin{aligned} A_{ij(\mathbf{B})}^{ab} &= \sum_{klg} B_{ij}^{kl} R_{kl}^{ab} \\ R_{kl}^{ab} &= \sum_g C_g \langle ab | \hat{f}_{12}^g | kl \rangle \end{aligned} \quad (5.8)$$

Valeev [46] did not contract the working equations, but chose a more flexible ansatz for the geminal expansion which was still restricted to s-type Gaussian geminals. The numerical observations of this study will in fact be important for the justification of our approximation for the one-particle integral contributions as will be demonstrated in the result section.

## 5.2 Modified Working Equations

In our previous approaches, the conventional electronic gradient elements  $W_{ij}^{AB}(\mathbf{A})$  were contracted:

$$X_{ij(\mathbf{A})}^{klg} = \sum_{AB} R_{kl}^{ABg} W_{ij}^{AB} \quad (5.9)$$

This expression contains various sums over products of integrals which can be reduced to two-, three- and four-electron integrals within a complete one-particle basis. Integrals containing more than two electrons cannot be solved analytically in an efficient way. To avoid these integrals, we assume that the one-particle basis used for calculation is complete, i.e.  $\{p\} \rightarrow \{P\}$ . Therefore, the conventional expressions from the working equations apply for

these contributions. We currently do not employ an auxiliary basis set[44] for these integrals, since the virtual space is contracted anyway. In case the used basis set is very small, the assumption that the computational basis is complete may introduce large errors due to the approximation of integrals explained above. It is recommended to use sufficiently large basis sets of at least triple zeta quality or specialized double zeta basis sets.[47] As already discussed above, we only want to modify the contributions corresponding to two electron integrals with s-type Gaussian geminals. Therefore, we will first isolate the corresponding summations in the working equations and then explain the necessary modifications. The two electron operators containing  $\hat{g}_{12}$  are given by the following contributions to the electronic gradient elements  $X_{ij(\mathbf{A})}^{klg}$ :

$$X_{ij(\mathbf{A})}^{klg} = \sum_{AB} R_{kl}^{ABg} G_{ij}^{AB} + \sum_{ABCD} R_{kl}^{ABg} G_{AB}^{CD} A_{ij}^{CD} + \dots \quad (5.10)$$

The contributions above hold for the conventional contraction scheme. The explicitly correlated wave function now introduces additional contributions, as shown in the equations 5.5 and 5.6. Note that only s-type Gaussian geminals denoted by ( $\mathbf{s}$ ) are considered in the correction of the wave function. The contributions of p-type geminals will remain as defined in chapter 3.

$$\begin{aligned} \sum_{AB} R_{kl}^{ABg(\mathbf{s})} G_{ij}^{AB} &\rightarrow \sum_{ab} R_{kl}^{abg(\mathbf{s})} G_{ij}^{ab} \\ + \frac{1}{2} \left( \sum_{PQ} R_{kl}^{PQg(\mathbf{s})} G_{ij}^{PQ} - \sum_{pq} R_{kl}^{pqg(\mathbf{s})} G_{ij}^{pq} \right) \end{aligned} \quad (5.11)$$

$$\sum_{AB} R_{kl}^{ABg(\mathbf{p})} G_{ij}^{AB} \rightarrow \sum_{ab} R_{kl}^{abg(\mathbf{p})} G_{ij}^{ab} \quad (5.12)$$

The quadratic contributions are given by:

$$\begin{aligned}
& \sum_{ABCD} R_{kl}^{ABg(s)} G_{AB}^{CD} A_{ij}^{CD(s)} \rightarrow \sum_{abcd} R_{kl}^{abg(s)} G_{ab}^{cd} A_{ij}^{cd(s)} \\
& + \frac{1}{2} \left( \sum_{PQcd} R_{kl}^{PQg(s)} G_{PQ}^{cd} A_{ij}^{cd(s)} - \sum_{pqcd} R_{kl}^{pqg(s)} G_{pq}^{cd} A_{ij}^{cd(s)} \right) \\
& + \frac{1}{2} \left( \sum_{abRS} R_{kl}^{abg(s)} G_{ab}^{RS} A_{ij}^{RS(s)} - \sum_{abrs} R_{kl}^{abg(s)} G_{ab}^{rs} A_{ij}^{rs(s)} \right) \\
& + \frac{1}{4} \left( \sum_{PQRS} R_{kl}^{PQg(s)} G_{PQ}^{RS} A_{ij}^{RS(s)} \right. \\
& - \sum_{pqRS} R_{kl}^{pqg(s)} G_{pq}^{RS} A_{ij}^{RS(s)} - \sum_{PQrs} R_{kl}^{PQg(s)} G_{PQ}^{rs} A_{ij}^{rs(s)} \\
& \left. + \sum_{pqrs} R_{kl}^{pqg(s)} G_{pq}^{rs} A_{ij}^{rs(s)} \right) \tag{5.13}
\end{aligned}$$

$$\begin{aligned}
& \sum_{ABCD} R_{kl}^{ABg(\mathbf{p})} G_{AB}^{CD} A_{ij}^{CD(\mathbf{p})} \rightarrow \sum_{abcd} R_{kl}^{abg(\mathbf{p})} G_{ab}^{cd} A_{ij}^{cd(\mathbf{p})} \\
& + \frac{1}{2} \left( \sum_{PQcd} R_{kl}^{PQg(\mathbf{p})} G_{PQ}^{cd} A_{ij}^{cd(\mathbf{p})} - \sum_{pqcd} R_{kl}^{pqg(\mathbf{p})} G_{pq}^{cd} A_{ij}^{cd(\mathbf{p})} \right) \tag{5.14}
\end{aligned}$$

$$\begin{aligned}
& \sum_{ABCD} R_{kl}^{ABg(\mathbf{p})} G_{AB}^{CD} A_{ij}^{CD(s)} \rightarrow \sum_{abcd} R_{kl}^{abg(\mathbf{p})} G_{ab}^{cd} A_{ij}^{cd(s)} \\
& + \frac{1}{2} \left( \sum_{abRS} R_{kl}^{abg(\mathbf{p})} G_{ab}^{RS} A_{ij}^{RS(s)} - \sum_{abrs} R_{kl}^{abg(\mathbf{p})} G_{ab}^{rs} A_{ij}^{rs(s)} \right) \tag{5.15}
\end{aligned}$$

$$\sum_{ABCD} R_{kl}^{ABg(\mathbf{p})} G_{AB}^{CD} A_{ij}^{CD(\mathbf{p})} \rightarrow \sum_{abcd} R_{kl}^{abg(\mathbf{p})} G_{ab}^{cd} A_{ij}^{cd(\mathbf{p})} \tag{5.16}$$

The equations above are derived by including the projection operators into the definition of the integral expression  $R_{kl}^{ABg}$  as defined in the ansatz for the contraction functions (see equations 5.5 and 5.6). The coefficients are decomposed according to:

$$\begin{aligned}
A_{ij}^{ab(s)} &= \sum_{kl} \sum_g^{s\text{-type}} B_{ij}^{klg} R_{kl}^{abg} \\
A_{ij}^{ab(\mathbf{p})} &= \sum_{kl} \sum_g^{p\text{-type}} B_{ij}^{klg} R_{kl}^{abg} \tag{5.17}
\end{aligned}$$

Parts of the equations above contain a summation over the complete one-particle basis. These parts of the equations can be modified further to introduce explicitly correlated integrals (cf. section 5.3). All contributions to the electronic gradient elements  $X_{ij(A)}^{klg}$  which must be decomposed as explained above since they resolve in two-electron integrals are given by the following expressions:

$$\begin{aligned} & \sum_{AB} R_{kl}^{ABg} A_{ij}^{AB} ; \quad \sum_{AB} R_{kl}^{ABg} G_{ij}^{AB} ; \quad \sum_{ABCD} R_{kl}^{ABg} G_{AB}^{CD} A_{ij}^{CD} \\ & \sum_{ABC} R_{kl}^{ABg} F_A^C A_{ij}^{CB} + R_{kl}^{ABg} F_B^C A_{ij}^{AC} \end{aligned} \quad (5.18)$$

For orbital optimization, additional sums in the orbital gradient equations have to be replaced accordingly:

$$\sum_{AB} A_{kl}^{AB} A_{ij}^{AB} ; \quad \sum_{AB} G_{pq}^{AB} A_{ij}^{AB} ; \quad \sum_A F_A^p A_{ij}^{Aq} \quad (5.19)$$

To derive the replacement for the expressions above, the projection operators need to be applied to the integral expressions  $R_{kl}^{ABg}$  as already shown in the beginning of this section (cf. equations 5.11 - 5.16). The explicit expressions are derived in the appendix (cf. section 8.2). Due to integral symmetry the expressions above might occur with slightly different ordered indices, depending on the working equations in use.

### 5.3 Explicitly Correlated Two-Electron Integrals

In order to introduce explicitly correlated integrals into our working equations, it is convenient to replace the projection operators over the complete one-particle basis  $\{P\}$  by the identity operator.[12]

$$\sum_P |P_1\rangle \langle P_1| = \hat{1}_1 \quad (5.20)$$

We may now proceed to replace all summations over the complete one particle basis in the expressions of section 5.2 by the identity operator. The following expression can then be reduced to the integral:

$$\begin{aligned} & \sum_{PQRS} R_{ij}^{PQg} G_{PQ}^{RS} R_{kl}^{RSg} \\ & = \sum_{PQRS} \langle ij | \hat{f}_{12}^g | PQ \rangle \langle PQ | \hat{g}_{12} | RS \rangle \langle RS | \hat{f}_{12}^g | kl \rangle \\ & = \langle ij | \hat{f}_{12}^g \hat{g}_{12} \hat{f}_{12}^g | kl \rangle \end{aligned} \quad (5.21)$$

Similar equations can be derived for all other operators:

$$\begin{aligned}\sum_{PQ} R_{ij}^{PQg} A_{kl}^{PQ} &= \sum_{mnh} \langle ij | \hat{f}_{12}^g \hat{f}_{12}^h | mn \rangle B_{kl}^{mnh} \\ &= RA_{ij}^{klg}\end{aligned}\quad (5.22)$$

$$\begin{aligned}\sum_{PQ} R_{ij}^{PQg} G_{pq}^{PQ} &= \langle ij | \hat{f}_{12}^g \hat{g}_{12} | pq \rangle \\ &= RG_{ij}^{pqg}\end{aligned}\quad (5.23)$$

$$\begin{aligned}\sum_{PQ} G_{pq}^{PQ} A_{ij}^{PQ} &= \sum_{klh} \langle pq | \hat{g}_{12} \hat{f}_{12}^h | kl \rangle B_{ij}^{klh} \\ &= GA_{pq}^{ij}\end{aligned}\quad (5.24)$$

$$\begin{aligned}\sum_{PQRS} R_{ij}^{PQg} G_{PQ}^{RS} A_{kl}^{RS} \\ &= \sum_{mnh} \langle ij | \hat{f}_{12}^g \hat{g}_{12} \hat{f}_{12}^h | mn \rangle B_{kl}^{mnh} \\ &= RGA_{ij}^{klg}\end{aligned}\quad (5.25)$$

$$\begin{aligned}\sum_{PQR} R_{kl}^{PQg} F_P^R A_{ij}^{RQ} + R_{kl}^{PQg} F_Q^R A_{ij}^{PR} \\ &= \sum_{mnh} \langle ij | \hat{f}_{12}^g (\hat{f}_1 + \hat{f}_2) \hat{f}_{12}^h | mn \rangle B_{kl}^{mnh} \\ &= RFA_{ij}^{klg}\end{aligned}\quad (5.26)$$

In case orbital optimization should be performed, we need to replace some additional summations in the orbital gradient equations:

$$\begin{aligned}\sum_{PQ} A_{ij}^{PQ} A_{kl}^{PQ} &= AA_{ij}^{kl} \\ \sum_{PQ} G_{pq}^{PQ} A_{ij}^{PQ} &= GA_{pq}^{ij} \\ \sum_P F_P^p A_{ij}^{Pq} &= FA_{pq}^{ij}\end{aligned}\quad (5.27)$$

Note that the actual dimensions of the tensors used will have to be modified in case of orbital optimization (cf. chapter 4).

There are multiple ways to calculate the contribution of the explicitly correlated integrals containing the Fock operator  $\hat{f}_1$ . [12, 13, 14] The definitions above do not impose any restrictions on this choice. We currently use a special kind of approximation explained and tested in the results section.

## 5.4 Results

In the following section, we discuss the results obtained by applying the explicitly correlated ansatz of equation 5.6. All considered molecular geometries were optimized on the MP2 level of theory using a 6-31+G\*\* basis set.

To explain the transition from explicitly correlated theory in literature to our ansatz we may first show a short comparison of different approximations for the coefficients. The so called non-diagonal ansatz [15] is closely related to our approach and will be labeled analogously. The only difference to our ansatz consists in the application of a fixed linear combination of geminals for each orbital pair, reducing the flexibility of the wave function:

$$A_{ij(\mathbf{B})}^{AB} = \sum_{klg} B_{ij}^{klg} R_{kl}^{ABg} \rightarrow A_{ij(\mathbf{B})}^{AB} = \sum_{kl} B_{ij}^{kl} \sum_g c_g R_{kl}^{ABg} \quad (5.28)$$

The diagonal ansatz further reduces the number of coefficients by the definition  $B_{ij}^{kl} = B_{ij} \delta_{ik} \delta_{jl}$  or in our case  $B_{ij}^{klg} = B_{ij}^g \delta_{ik} \delta_{jl}$ . In modern explicitly correlated ansatzes, the coefficients  $B_{ij}$  are fixed to constant predefined values which fit the cusp conditions for S and P partial waves.[16] The most commonly used ansatz for the description of the two-electron integral contributions resulting from the Fock-operator is labeled as approximation C.[48] Since we previously showed that for our contraction scheme a high variational freedom is crucial to recover a high percentage of correlation energy by contraction, we would like to adopt this principle for the explicitly correlated wave function as well. Our studies reveal numerical instabilities for the hydrogen fluoride molecule in a cc-pVTZ basis[5] when combining approximation C with the non-diagonal ansatz using one or more s-type Gaussian geminals. The same holds for the diagonal ansatz using two Gaussian geminals. Only the diagonal approximation C using one geminal leads to faithful results (cf. Table 5.1). Note that these numerical instabilities were also reported earlier by Adler *et al.*[49] and Valeev [46]. One possible way to resolve this problem would be the introduction of large auxiliary basis sets [44] or removing the instabilities by a Singular Value Decomposition. Alternatively, we pursued a different approach labeled approximation D. In this approximation, we describe all contributions from the Fock-operator completely by the resolution of the identity in the basis set used for the calculation. Approximation D leads to stable results for the non-diagonal ansatz with one or more geminals. The fraction of correlation energy recovered compared to the (most possibly close to perfect)  $f_{12}$ -CCSD(T)[50] energy is already quite high for such a small number of Gaussian geminals and can be further increased by a larger geminal basis. This approximation may slightly impair the convergence to the basis set limit, but the dominant contributions are still described.[11] We

assume that the numerical instabilities can be removed this way, since the contributions of the kinetic energy operator to matrix elements is by orders of magnitude higher than for  $1/r_{12}$ . Small errors in the description of these explicitly correlated integrals may therefore lead to unstable results when the variational freedom is too large.

Table 5.1: Correlation energies of the hydrogen fluoride molecule using the explicitly correlated diagonal and non-diagonal ansatz and different approximations of the one-electron contributions in the *cc*-pVTZ basis set. The  $f_{12}$ -CCSD(T) correlation energy is given by  $-0.329 E_H$

	$E_{CORR} [E_H]$			
	Approx. C	Approx. D	Approx. D	Approx. D
	diag.	diag.	non-diag.	non-diag.
	1s1p	1s1p	1s1p	2s2p
ACPF	-0.345	-0.225	-0.260	-0.275
CEPA0	-0.351	-0.233	-0.262	-0.278
CID	-0.327	-0.225	-0.253	-0.267

In figure 5.1 and 5.2 we studied the basis set convergence for the helium atom and the hydrogen molecule, respectively. In both cases we clearly see an improvement with respect to the contracted approach without explicit correlation and a smooth convergence towards the basis set limit. Note that the limit of the contracted and the conventional approach are only equivalent for a sufficiently large geminal basis set.

In figure 5.3 we depicted the percentage of the correlation energy recovered by different CEPA0 versions for a test set of molecules. The  $f_{12}$ -CCSD(T)[49] results, calculated with the ORCA program package[50] and a *cc*-pVDZ- $f_{12}$  basis[47], were taken as a reasonable reference energy. The fraction of the correlation energy recovered by CCSD(T) and CEPA0 is similar (86 and 84%), while we loose 8% when using our contraction scheme. However, the contracted explicitly correlated ansatz increases the fraction of the correlation energy drastically and recovers the  $f_{12}$ -CCSD(T) correlation energies quite well with about 92%. Similar results are obtained for the orbital optimized version (cf. Figure 5.4). Here, the percentage of the correlation energy recovered increases by about 2% for all ansatzes compared to the non-orbital optimized version. Note, that further improvements are again possible by including more geminals into the expansion.

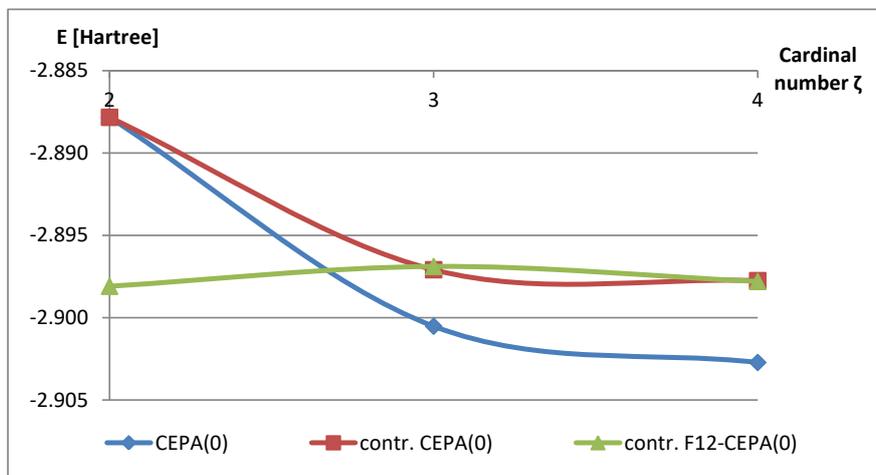


Figure 5.1: Total energy of the helium atom calculated by the CEPA0 variants using two s- and p- Gaussian geminals and the cc-pV $\zeta$ Z basis sets with different cardinal numbers  $\zeta$ .

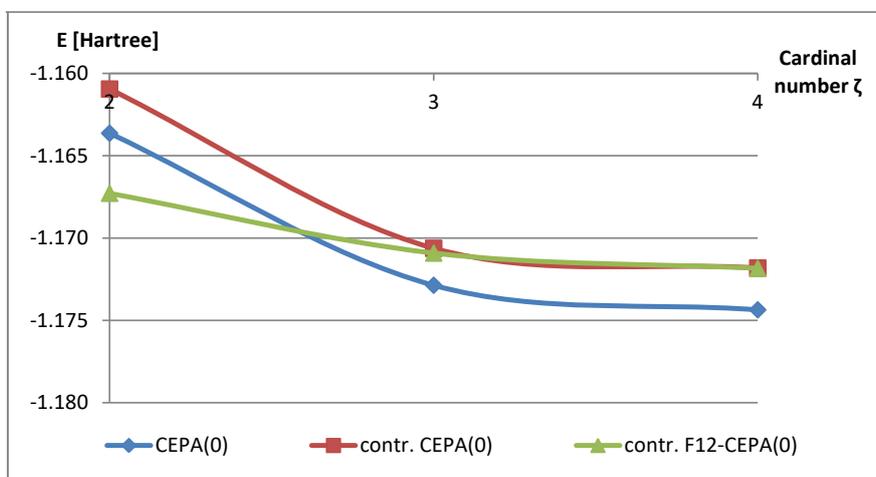


Figure 5.2: Total energy of the hydrogen molecule calculated by the CEPA0 variants using two s- and p- Gaussian geminals and the cc-pV $\zeta$ Z basis sets with different cardinal numbers  $\zeta$ .

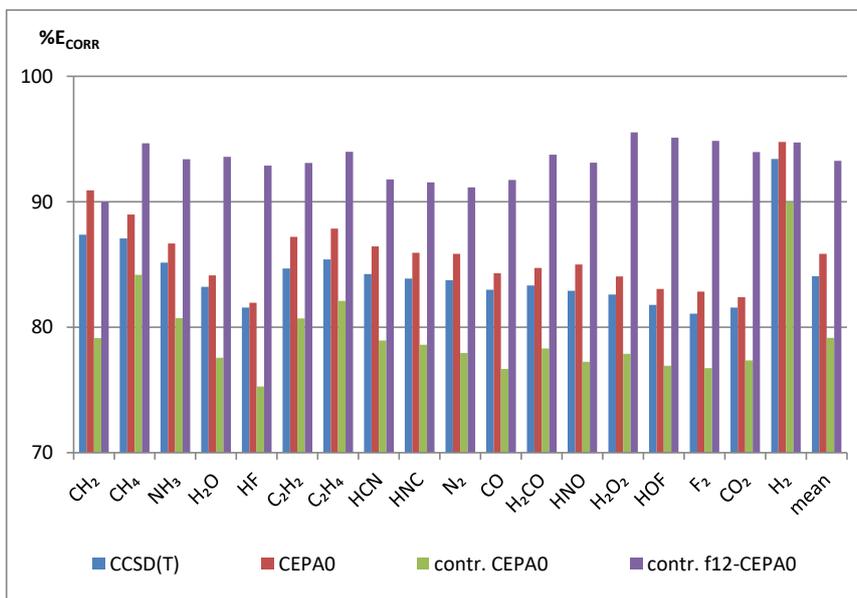


Figure 5.3: Percentage of the  $f_{12}$ -CCSD(T) correlation energy recovered by the CEPA0 variants using two s- and p- Gaussian geminals and the cc-pVDZ- $f_{12}$  basis set.

Additionally, we would like to present a comparison of reaction energies computed by our new ansatz to experimental data (cf. Table 5.2 and 5.3).[51] The RMS (root mean square) deviation, the MAPE (mean absolute percentage error), and the number of reactions, whose reaction energies deviate by more than 10% from the experimental reference value, are used as reasonable quantities for analysis. Note that  $f_{12}$ -CCSD(T) leads to the values  $RMS = 5 \frac{kJ}{mol}$ ,  $MAPE = 2\%$  and no reaction with larger deviations than 10%. Comparing the contracted and the conventional ansatz without explicit correlation, we found that all of the above quantities are slightly improved by contraction, even though the percentage of the correlation energy recovered decreased. This holds with and without including orbital optimization, and might be regarded as an error cancellation. A comparison between the conventional wave function with and without orbital optimization reveals virtually identical results regarding these quantities. The same arguments hold when including contraction and explicit correlation. This is somewhat unexpected, since all variational parameters are optimized and the percentage of the correlation energy recovered is improved. The explicitly correlated contracted reaction energies improve the results compared to the conventional

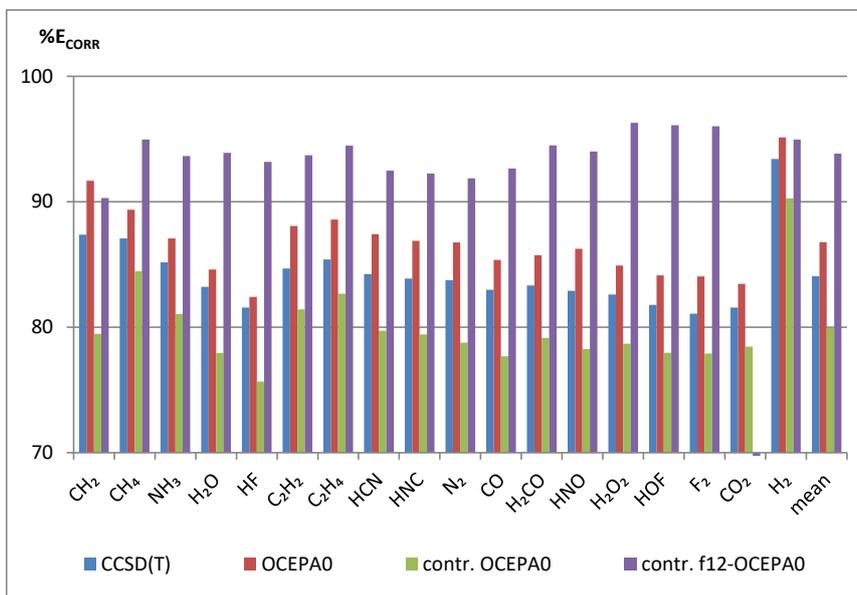


Figure 5.4: Percentage of the  $f_{12}$ -CCSD(T) correlation energy recovered by the OCEPA0 variants using two  $s$ - and  $p$ - Gaussian geminals and the cc-pVDZ- $f_{12}$  basis set.

ansatz with and without orbital optimization. In this case, the percentage of the correlation energy recovered is also increased. Comparing the reaction energies of the explicitly correlated contracted with the contracted ansatz, no significant improvement can be seen, even though the percentage of the correlation energy recovered is drastically improved by 14% with and without orbital optimization.

The remaining deviations might be attributed to the small orbital and geminal basis set and the pair correlation method itself. Further studies on these possible sources of error and an extension to orbital optimized coupled cluster doubles[10] are necessary to provide further information on the potential accuracy for the explicitly correlated contraction approach.

Table 5.2: CEPA0 electronic reaction energies using the cc-pVDZ- $f_{12}$  basis[47].

	$\Delta E[\frac{kJ}{mol}]$			
	exp.[51]	CEPA0	contr. CEPA0	contr. $f_{12}$ - CEPA0
$CH_2 + H_2 \rightarrow CH_4$	-544	-518	-548	-557
$C_2H_2 + H_2 \rightarrow C_2H_4$	-203	-207	-213	-215
$C_2H_2 + 3H_2 \rightarrow 2CH_4$	-446	-448	-466	-463
$N_2 + 3H_2 \rightarrow 2NH_3$	-164	-144	-163	-180
$F_2 + H_2 \rightarrow 2HF$	-563	-546	-539	-534
$H_2CO + 2H_2 \rightarrow CH_4 + H_2O$	-251	-249	-254	-250
$H_2O_2 + H_2 \rightarrow 2H_2O$	-365	-364	-358	-340
$CO + 3H_2 \rightarrow CH_4 + H_2O$	-272	-259	-274	-288
$HCN + 3H_2 \rightarrow CH_4 + NH_3$	-320	-312	-333	-339
$HNO + 2H_2 \rightarrow H_2O + NH_3$	-444	-429	-448	-444
$H_2O + F_2 \rightarrow HOF + HF$	-129	-108	-108	-117
$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$	-244	-259	-237	-238
$2CH_2 \rightarrow C_2H_4$	-844	-795	-844	-867
RMS [REF]		20	12	17
MAPE [REF]		5	3	5
Number of Deviations > 10%		2	1	0

Table 5.3: OCEPA0 electronic reaction energies using the cc-pVDZ- $f_{12}$  basis[47].

	$\Delta E[\frac{kJ}{mol}]$			
	exp.[51]	OCEPA0	contr. OCEPA0	contr. $f_{12}$ - OCEPA0
$CH_2 + H_2 \rightarrow CH_4$	-544	-517	-548	-558
$C_2H_2 + H_2 \rightarrow C_2H_4$	-203	-206	-212	-213
$C_2H_2 + 3H_2 \rightarrow 2CH_4$	-446	-444	-462	-460
$N_2 + 3H_2 \rightarrow 2NH_3$	-164	-138	-157	-175
$F_2 + H_2 \rightarrow 2HF$	-563	-534	-527	-520
$H_2CO + 2H_2 \rightarrow CH_4 + H_2O$	-251	-242	-249	-245
$H_2O_2 + H_2 \rightarrow 2H_2O$	-365	-358	-351	-334
$CO + 3H_2 \rightarrow CH_4 + H_2O$	-272	-252	-267	-282
$HCN + 3H_2 \rightarrow CH_4 + NH_3$	-320	-306	-328	-335
$HNO + 2H_2 \rightarrow H_2O + NH_3$	-444	-418	-439	-436
$H_2O + F_2 \rightarrow HOF + HF$	-129	-106	-106	-114
$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$	-244	-248	-224	
$2CH_2 \rightarrow C_2H_4$	-844	-795	-847	-869
RMS [REF]		23	15	20
MAPE [REF]		6	4	5
Number of Deviations > 10%		2	1	1

## 5.5 Summary and Outlook

We presented an efficient scheme to include explicit correlation into contracted orbital optimized pair correlation methods. The results reveal improved convergence of the correlation energy with respect to the basis set size. Reaction energies including the explicitly correlated contracted ansatz improve the conventional ansatz with and without orbital optimization. The orbital optimization itself does not lead to higher accuracy for the current test set of reactions. Further studies with larger orbital and geminal basis sets and higher geminal angular momenta are necessary to confirm this result.

As a perspective, the inclusion of auxiliary basis sets[44] and the extension to orbital optimized coupled cluster doubles will be an important future development to finally provide a method which can describe chemical reactions including homolytic bond breakings with very high accuracy.[10]

# Chapter 6

## Summary

The main purpose of theoretical chemistry is to predict chemical properties and reactions by simulation. For this purpose a huge number of methods are available, which mainly differ in accuracy and computational costs.

In this thesis a new scheme to make highly accurate pair correlation methods computationally more efficient was presented. This scheme introduces a contraction of the wave function which strongly reduces the number of variables which have to be optimized and is applicable to any pair correlation method[2, 3, 4] (cf. chapter 3). A large percentage of the correlation energy is recovered and the introduced error is of systematic nature, such that reaction energies are reproduced with high accuracy. Even weak dispersion interactions can be described accurately.

The extension to orbital optimization allows the description of single bond breaking with high accuracy. Contraction recovers this property perfectly (see chapter 4). An improvement on reaction energies by orbital optimization as stated by other authors[7] has not been found so far (see chapter 5). Further studies with different pair correlation methods have to be performed for conclusive statements.[8, 9, 10] Nevertheless, the presented ansatz paves the way towards multi-reference methods including dynamic correlation with applicability to large molecules.

Explicitly correlated methods improve the convergence of chemical properties with respect to the basis set size. The usage of large and therefore computationally expensive basis sets can be avoided.[12, 13, 14] An extension of our approach to explicitly correlated orbital optimized contracted pair correlation methods is shown in chapter 5. The results show improved convergence towards the basis set limit of the corresponding contracted wave function. A comparison of reaction energies calculated by the explicitly correlated contracted ansatz to the energies of the conventional ansatz exhibits improved accuracy with and without orbital optimization. In combination

with full exploitation of locality one may thereby decrease simulation time and increase the accuracy significantly.

An extension of our ansatz to coupled cluster methods will be an important future development to increase the accuracy of the underlying correlation method.[39, 52] Additionally, higher excitation levels are also accessible by contraction. In combination with orbital optimization this will lead to efficient and accurate methods, which are able to describe excited states and the dissociation of multiple bonds. A parallel implementation using the short range properties of the atomic orbital basis is still missing and would lead to the ability of calculating huge systems (see chapter 2.3).[22, 24, 25] The short range of correlation effects can be further exploited by describing close pairs with longer geminal expansions than the remainder.

In conclusion, a new approach towards the simulation of large molecular systems was presented, which is applicable to accurate pair-correlation methods, can be extended with orbital optimization to describe multi-reference problems, and includes the effects of explicit correlation for improved basis set convergence.

# Chapter 7

## German Summary

Das Hauptaufgabengebiet der theoretischen Chemie beinhaltet die Vorhersage von chemischen Eigenschaften und Reaktionen durch Simulationen. Für diese Simulationen stehen verschiedene Verfahren zur Verfügung, welche sich hauptsächlich in ihrer Genauigkeit und dem Rechenaufwand unterscheiden.

In dieser Arbeit wurde ein neues Konzept vorgestellt, welches den rechnerischen Aufwand hochgenauer Paar-Korrelationsverfahren deutlich reduziert. Hierbei wird die Wellenfunktion kontrahiert, wodurch sich die Anzahl der Variablen, welche optimiert werden müssen, deutlich reduziert. Diese Kontraktion ist auf jedes Paar-Korrelationsverfahren anwendbar (siehe Kapitel 3).[2, 3, 4] Dabei bleibt ein großer Prozentsatz der Korrelationsenergie erhalten und der verbleibende Fehler ist von systematischer Natur. Dementsprechend werden Reaktionsenergien ebenfalls mit hoher Genauigkeit reproduziert. Sogar kleine dispersive Wechselwirkungsenergien werden korrekt wiedergegeben.

Die Erweiterung des Ansatzes auf eine zusätzliche Orbitaloptimierung ermöglicht die Vorhersage von (Einfach-)Bindungsbrüchen mit hoher Genauigkeit. Die Kontraktion reproduziert diese Eigenschaft perfekt. Eine verbesserte Vorhersage von Reaktionsenergien, wie sie von anderen Autoren[7] gefunden wurde, konnte bisher nicht bestätigt werden. Weitere Studien mit verschiedenen orbitaloptimierten Paar-Korrelationsverfahren sind notwendig, um dieses Ergebnis zu validieren. [8, 9, 10] Nichtsdestotrotz stellt der präsentierte Ansatz für die Wellenfunktion einen guten Ausgangspunkt für Multi-Referenz Verfahren dar, welche statische und dynamische Korrelation gleichzeitig mit hoher Genauigkeit auch für große Moleküle berechnen können.

Explizit korrelierte Verfahren verbessern die Konvergenz chemischer Eigenschaften mit der Basissatzgröße. Dadurch kann die Verwendung großer und damit rechnerisch aufwändiger Basissätze verhindert werden. Die Ergebnisse dieser Erweiterung unseres Ansatzes zeigen eine beschleunigte Konver-

genz auf das Basissatzlimit der jeweiligen kontrahierten Wellenfunktion (siehe Kapitel 5). Ein Vergleich der mit und ohne Orbitaloptimierung berechneten Reaktionsenergien mit experimentellen Werten zeigt eine erhöhte Vorhersagegenauigkeit. Unter Ausnutzung der Lokalität des Ansatzes kann also potentiell die Genauigkeit erhöht und gleichzeitig die Simulationszeit reduziert werden.

Die Erweiterung unseres Ansatzes auf Coupled-Cluster Wellenfunktionen wird eine wichtige zukünftige Entwicklung darstellen. [39, 52] Dadurch würde die Genauigkeit des zugrundeliegenden Paar-Korrelationsverfahrens weiter erhöht. Desweiteren könnten höhere Anregungen ebenfalls durch Kontraktion dargestellt werden. In Verbindung mit der Orbitaloptimierung können so auch angeregte Zustände und die Dissoziation von Mehrfachbindungen genau und effizient beschrieben werden. Eine parallele Implementierung, welche die kurze Reichweite der Atomorbitale ausnutzt, fehlt ebenfalls und würde die Simulation großer Moleküle ermöglichen. [22, 24, 25]

Zusammengefasst wurde ein neuer Ansatz für die Simulation großer molekularer Systeme vorgestellt, welcher auf alle Paar-Korrelationsverfahren angewendet werden kann. Dieser Ansatz wurde für Multi-Referenz Probleme erweitert und beinhaltet die beschleunigte Basissatzkonvergenz explizit korrelierter Verfahren.

# Chapter 8

## Appendix

### 8.1 Generalised Transformation Rules

As demonstrated in chapter 3, the summations over virtual orbitals can be replaced by matrix multiplications in the local atomic orbital basis. Accordingly, operator products arise in the working equations, which can be reformulated as matrix products as follows. A sum over the virtual orbital index  $a$  of two arbitrary matrices  $\mathbf{M}_1^{MO}$  and  $\mathbf{M}_2^{MO}$  in the molecular orbital basis may be rewritten as a matrix product in the following way:

$$M^{pq} = \sum_a M_1^{pa} M_2^{aq} \Rightarrow \mathbf{M} = \mathbf{M}_1^{MO} \mathbf{M}_2^{MO} \quad (8.1)$$

The transformation of an operator from the atomic- to the molecular orbital basis is defined as:

$$\mathbf{M}^{MO} = \mathbf{C}^t \mathbf{M}^{AO} \mathbf{C} \quad (8.2)$$

Consequently, the whole expression in atomic orbital basis can be formulated as:

$$\mathbf{M}_1^{MO} \mathbf{M}_2^{MO} = \mathbf{C}^t \mathbf{M}_1^{AO} \mathbf{D}_{virt}^{AO} \mathbf{M}_2^{AO} \mathbf{C} \quad (8.3)$$

The same transformation applies for a sum over the computational basis  $p$ , while the definition of the density has to be changed. The corresponding modifications lead to working equations completely defined in the local atomic orbital basis.

## 8.2 Additional Contributions for Explicit Correlation

In this part of the appendix we present explicit definitions of the remaining contributions of section 5.2. For the coefficient gradient the following contributions are necessary. To derive the replacement for:

$$\sum_{AB} R_{kl}^{ABg} A_{ij}^{AB} \quad (8.4)$$

a simple substitution of  $G_{ij}^{AB}$  by  $A_{ij}^{AB}$  in the equations 5.11 and 5.12 is sufficient. To derive the replacement for:

$$\sum_{ABC} R_{kl}^{ABg} F_A^C A_{ij}^{CB} + R_{kl}^{ABg} F_B^C A_{ij}^{AC} \quad (8.5)$$

we need to extend the sum by:

$$\sum_{ABCD} R_{kl}^{ABg} (F_A^C \delta_{BD} + F_B^D \delta_{AC}) A_{ij}^{CD} \quad (8.6)$$

And can again substitute  $G_{AB}^{CD}$  by  $(F_A^C \delta_{BD} + F_B^D \delta_{AC})$  in the equations 5.13 to 5.16.

For the orbital gradient the following contributions are necessary. To derive the replacement for:

$$\sum_{AB} A_{kl}^{AB} A_{ij}^{AB} \quad (8.7)$$

a simple substitution of  $R_{kl}^{ABg}$  by  $A_{kl}^{AB}$  and  $G_{ij}^{AB}$  by  $A_{ij}^{AB}$  in the equations 5.11 and 5.12 is sufficient. To derive the replacement for:

$$\sum_{AB} G_{pq}^{AB} A_{ij}^{AB} \quad (8.8)$$

a simple substitution of  $R_{kl}^{ABg}$  by  $G_{pq}^{AB}$  and  $G_{ij}^{AB}$  by  $A_{ij}^{AB}$  in the equations 5.11 and 5.12 is sufficient. To derive the replacement for:

$$\sum_A F_A^p A_{ij}^{Aq} \quad (8.9)$$

we need to extend the sum by:

$$\sum_{AB} F_A^p \delta_{AB} A_{ij}^{Aq} \quad (8.10)$$

And can again substitute  $R_{kl}^{ABg}$  by  $F_A^p \delta_{AB}$  and  $G_{ij}^{AB}$  by  $A_{ij}^{Aq}$  in the equations 5.11 and 5.12.

# Chapter 9

## List of Abbreviations

### General:

PAO	<u>p</u> rojected <u>a</u> tomical <u>o</u> rbital[24]
PNO	<u>p</u> air <u>n</u> atural <u>o</u> rbital[25]
BCH	<u>B</u> aker- <u>C</u> ampel- <u>H</u> ausdorff expansion[19]
CSF	<u>c</u> onfiguration <u>s</u> tate <u>f</u> unction
BFGS	<u>B</u> royden- <u>F</u> letcher- <u>G</u> oldfarb- <u>S</u> hanno[32]
RMS	<u>r</u> oot <u>m</u> ean <u>s</u> quare deviation
MAPE	<u>m</u> ean <u>a</u> bsolute <u>p</u> ercentage <u>e</u> rror

### Methods:

CI	<u>C</u> onfiguration <u>I</u> nteraction[2]
CID	<u>C</u> I including only <u>d</u> ouble excitations[2]
CISD	<u>C</u> I including only <u>s</u> ingle and <u>d</u> ouble excitations[2]
ACPF	<u>a</u> veraged <u>c</u> oupled <u>p</u> air <u>f</u> unctional[3]
CEPA0	<u>c</u> oupled <u>e</u> lectron <u>p</u> air <u>a</u> pproximation[4]
OCID	<u>o</u> rbital optimized <u>C</u> ID
OACPF	<u>o</u> rbital optimized <u>A</u> CPF
OCEPA0	<u>o</u> rbital optimized <u>C</u> EPA0[7]

CCSD(T)	<u>C</u> oupled <u>C</u> luster <u>s</u> ingles <u>d</u> oubles with approximate <u>t</u> riples[42]
CC2	second-order approximate <u>C</u> oupled <u>C</u> luster <u>s</u> ingles <u>d</u> oubles[53]
MP2	second-order <u>M</u> øller- <u>P</u> lesset perturbation theory[36]
CASSCF	<u>c</u> omplete <u>a</u> ctive <u>s</u> pace <u>s</u> elf <u>c</u> onsistent <u>f</u> ield theory[41]
CAS( $N_{ele}$ , $N_{orb}$ )	<u>c</u> omplete <u>a</u> ctive <u>s</u> pace with $N_{ele}$ electrons and $N_{orb}$ orbitals[41]
DLPNO-Method	<u>d</u> omain- <u>b</u> ased <u>l</u> ocal <u>p</u> air <u>n</u> atural <u>o</u> rbital correlation method[22]
f12-Method	explicitly correlated version of the correlation method[12, 13, 14]
Basissets:	
cc-pVDZ	<u>c</u> orrelation <u>c</u> onsistent <u>p</u> olarized <u>v</u> alence <u>d</u> ouble <u>z</u> eta[5]
cc-pVTZ	<u>c</u> orrelation <u>c</u> onsistent <u>p</u> olarized <u>v</u> alence <u>t</u> riple <u>z</u> eta[5]
6-31G*	pople type basis with polarization functions for atoms beginning from the second period[54]
6-31+G**	pople type basis with polarization functions for all atoms and diffuse functions for atoms beginning from the second period[37]
Basisset-f12	specialized variant of the basis set optimized for explicit correlation[47]

# Chapter 10

## Bibliography

- [1] Deglmann, P.; Schäfer, A.; Lennartz, C. *International Journal of Quantum Chemistry* **2015**, *115*(3), 107–136.
- [2] Pulay, P.; Saebø, S.; Meyer, W. *The Journal of Chemical Physics* **1984**, *81*(4), 1901–1905.
- [3] Gdanitz, R. J.; Ahlrichs, R. *Chemical Physics Letters* **1988**, *143*(5), 413 – 420.
- [4] Meyer, W. *Theoretica chimica acta* **1974**, *35*(4), 277–292.
- [5] Dunning Jr., T. H. *The Journal of Chemical Physics* **1989**, *90*(2), 1007–1023.
- [6] Lasar, C.; Klüner, T. *The Journal of Physical Chemistry A* **2017**, *121*(24), 4707–4711.
- [7] Bozkaya, U.; Sherrill, C. D. *The Journal of Chemical Physics* **2013**, *139*(5), 054104.
- [8] Scuseria, G. E.; Schaefer III, H. F. *Chemical Physics Letters* **1987**, *142*(5), 354 – 358.
- [9] Sherrill, C. D.; Krylov, A. I.; Byrd, E. F. C.; Head-Gordon, M. *The Journal of Chemical Physics* **1998**, *109*(11), 4171–4181.
- [10] Krylov, A. I.; Sherrill, C. D.; Byrd, E. F. C.; Head-Gordon, M. *The Journal of Chemical Physics* **1998**, *109*(24), 10669–10678.
- [11] Kutzelnigg, W. *Theoretica chimica acta* **1985**, *68*(6), 445–469.

- [12] Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. *Chemical Reviews* **2012**, *112*(1), 4–74.
- [13] Kong, L.; Bischoff, F. A.; Valeev, E. F. *Chemical Reviews* **2012**, *112*(1), 75–107.
- [14] Ten-no, S.; Noga, J. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*(1), 114–125.
- [15] Klopper, W. *Chemical Physics Letters* **1991**, *186*(6), 583 – 585.
- [16] Ten-no, S. *Chemical Physics Letters* **2004**, *398*(13), 56 – 61.
- [17] Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley, 2014.
- [18] Dalgaard, E.; Jørgensen, P. *The Journal of Chemical Physics* **1978**, *69*(8), 3833–3844.
- [19] Achilles, R.; Bonfiglioli, A. *Archive for History of Exact Sciences* **2012**, *66*(3), 295–358.
- [20] Hestenes, Magnus R.; Stiefel, E. *Journal of Research of the National Bureau of Standards* **1952**, *49*(6), 409–436.
- [21] Ahlrichs, R. *Computer Physics Communications* **1979**, *17*(1), 31 – 45.
- [22] Riplinger, C.; Neese, F. *The Journal of Chemical Physics* **2013**, *138*(3), 034106.
- [23] Pipek, J.; Mezey, P. G. *The Journal of Chemical Physics* **1989**, *90*(9), 4916–4926.
- [24] Saebø, S.; Pulay, P. *Chemical Physics Letters* **1985**, *113*(1), 13 – 18.
- [25] Meyer, W. *The Journal of Chemical Physics* **1973**, *58*(3), 1017–1035.
- [26] Schütz, M.; Hetzer, G.; Werner, H.-J. *The Journal of Chemical Physics* **1999**, *111*(13), 5691–5705.
- [27] Hetzer, G.; Schütz, M.; Stoll, H.; Werner, H.-J. *The Journal of Chemical Physics* **2000**, *113*(21), 9443–9455.
- [28] Schütz, M. *The Journal of Chemical Physics* **2000**, *113*(22), 9986–10001.
- [29] Schütz, M.; Werner, H.-J. *The Journal of Chemical Physics* **2001**, *114*(2), 661–681.

- [30] Schütz, M. *The Journal of Chemical Physics* **2002**, *116*(20), 8772–8785.
- [31] Persson, B. J.; Taylor, P. R. *Theoretical Chemistry Accounts* **1997**, *97*(1), 240–250.
- [32] Byrd, R. H.; Nocedal, J.; Schnabel, R. B. *Mathematical Programming* **1994**, *63*(1), 129–156.
- [33] Meyer, W. *The Journal of Chemical Physics* **1976**, *64*(7), 2901–2907.
- [34] Vahtras, O.; Almlöf, J.; Feyereisen, M. *Chemical Physics Letters* **1993**, *213*(5), 514 – 518.
- [35] Pulay, P.; Saebø, S. *Annual Review of Physical Chemistry* **1993**, *44*(1), 213–236.
- [36] Møller, C.; Plesset, M. S. *Physical Review* **1934**, *46*, 618–622.
- [37] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *The Journal of Chemical Physics* **1972**, *56*(5), 2257–2261.
- [38] Gaussian 09 Revision B.01. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E. **2009**.
- [39] Scuseria, G. E.; Janssen, C. L.; Schaefer III, H. F. *The Journal of Chemical Physics* **1988**, *89*(12), 7382–7387.
- [40] Thouless, D. *Nuclear Physics* **1960**, *21*(Supplement C), 225 – 232.
- [41] Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. *Chemical Physics* **1980**, *48*(2), 157 – 173.
- [42] Watts, J. D.; Gauss, J.; Bartlett, R. J. *The Journal of Chemical Physics* **1993**, *98*(11), 8718–8733.
- [43] Fink, R.; Staemmler, V. *Nov* **1993**, *87*(1), 129–145.
- [44] Klopper, W.; Samson, C. C. M. *The Journal of Chemical Physics* **2002**, *116*(15), 6397–6410.
- [45] Höfener, S.; Tew, D. P.; Klopper, W.; Helgaker, T. *Chemical Physics* **2009**, *356*(13), 25 – 30.
- [46] Valeev, E. F. *The Journal of Chemical Physics* **2006**, *125*(24), 244106.
- [47] Peterson, K. A.; Adler, T. B.; Werner, H.-J. *The Journal of Chemical Physics* **2008**, *128*(8), 084102.

- [48] Kedžuch, S.; Milko, M.; Noga, J. *International Journal of Quantum Chemistry* **2005**, *105*(6), 929–936.
- [49] Adler, T. B.; Knizia, G.; Werner, H.-J. *The Journal of Chemical Physics* **2007**, *127*(22), 221106.
- [50] Neese, F. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*(1), 73–78.
- [51] Klopper, W.; Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T. *Journal of Physics B: Atomic, Molecular and Optical Physics* **1999**, *32*(13), R103.
- [52] Čížek, J. *The Journal of Chemical Physics* **1966**, *45*(11), 4256–4266.
- [53] Christiansen, O.; Koch, H.; Jørgensen, P. *Chemical Physics Letters* **1995**, *243*(5), 409 – 418.
- [54] Hariharan, P. C.; Pople, J. A. *Theoretica chimica acta* **1973**, *28*(3), 213–222.

# Chapter 11

## Acknowledgments

I would like to thank Prof. Dr. Klüner and Robert Röhse for providing an interesting starting point for this work and the scientific discussions during my time as a PhD student.

Additionally, i would like to thank Prof. Dr. Staemmler and Prof. Dr. Klopper for their scientific support, which helped a lot to improve this work and Prof. Dr. Holthaus for kindly undertaking the job of the second referee.

Of course i would like to acknowledge my colleagues and the students in our work group, which always ensured a nice atmosphere during work- and free time. Especially, i would like to mention Hendrik Spieker, who had a lot of potential to become a great scientist and had to leave this world way to early.

Last but not least, i would like to thank my girlfriend and my family, who always supported me during this demanding time.

# Chapter 12

## Declaration

Ich versichere hiermit, dass ich diese Arbeit selbständig verfasst habe und nur die angegebenen Quellen und Hilfsmittel benutzt habe. Die Dissertation hat weder in Teilen noch in Gänze einer anderen wissenschaftlichen Hochschule zur Begutachtung in einem Promotionsverfahren vorgelegen.

Die Leitlinien guter wissenschaftlicher Praxis an der Carl von Ossietzky Universität Oldenburg sind befolgt worden. Im Zusammenhang mit dem Promotionsvorhaben wurden keine kommerziellen Vermittlungs- oder Beratungsdienste (Promotionsberatung) in Anspruch genommen.