UNIVERZITA KARLOVA V PRAZE Přírodovědecká fakulta

Studijní program: Fyzikální chemie



Mgr. Jan Šmydke

Analytický gradient multireferenční metody vázaných klastrů

An Analytical Gradient of a Multireference Coupled Cluster Method

Disertační práce

Školitel: Mgr. Jiří Pittner, Dr. rer. nat. Praha 2011

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze 1.6.2011

Jan Šmydke

Abstract

The main objective of this work was to implement an analytical gradient of energy for a multireference coupled cluster method (MRCC).

Multireference quantum chemical methods play a fundamental role for studying non-equilibrium molecular geometries, dissociation of molecules into fragments or generally quasidegenerate systems. In all such cases the singlereference methods fail and the multireference description thus becomes inevitable.

While the standard single-reference CC methods have been widely used for almost 40 years as the most accurate, yet computationally feasible methods for calculating the correlation energy, the MRCC theories are still the subject of an active development and are only slowly accepted for a routine quantum chemical use.

Until recently, there was only one implementation of an analytical gradient for a MRCC method which, however, specialized on a two-reference low-spin singlet systems only. Because of the key importance of the analytical gradient for the whole quantum chemistry, an analytical gradient implementation for a general MRCC method is therefore highly desirable.

A gradient of the state-specific multireference Brillouin–Wigner CCSD method with the iterative correction of size-extensivity has been chosen since the method had already been implemented in the ACES II [1] program package, which is provided with an efficient single-reference CCSD analytical gradient implementation so that its code could conveniently be reused.

As an inevitable development tool, a more general Full-CI based program has been used, which enabled a pilot implementation of the target MR BWCC gradient with and without the iterative size-extensivity correction and also a gradient of the state-universal MRCC method. The pilot implementation of the three analytical gradient variants has been successfully tested on CH_2 and SiH₂ systems and also published [2].

Finally, the analytical gradient of the iteratively corrected MR BWCC method has also been implemented in the ACES II package, although the orbital response contribution requires an explicit solution of the CPHF equations, which is performed by an external code.

In addition, as a testing application of a newly developed state-specific multireference Mukherjee $CCSD(T_u)$ method, an investigation of the singlettriplet energy separation of the tetramethyleneethane molecule has been performed. The obtained potential energy curves were smooth without singularities and the calculated results very well agreed with an experiment. Since the method does not require an iterative solution of the approximate \hat{T}_3 amplitudes and has proved to remain resistant to intruders, it is thus a perspective alternative among highly accurate multireference approaches. This work is also a part of a publication [3].

Abstrakt

Hlavním cílem této práce bylo naimplementovat analytický gradient energie pro multireferenční metodu vázaných klastrů (MRCC).

Multireferenční kvantově chemické metody jsou klíčové pro studium molekul v nerovnovážných konfiguracích, rozpadu molekul na fragmenty nebo obecně kvazidegenerovaných systémů. Ve všech těchto případech běžné jednoreferenční metody selhávají a multireferenční popis je pro ně tedy nezbytný.

Zatímco jednoreferenční metody vázaných klastrů (CC) se již téměř 40 let používají pro přesné výpočty korelační energie, jejich multireferenční verze jsou stále předmětem aktivního vývoje a do běžné praxe se dostávají jen pomalu.

Až donedávna existovala pouze jediná implementace analytického gradientu multireferenční metody vázaných klastrů, ta se však týká jen speciálního případu dvoureferenčních singletových stavů. Poněvadž analytický gradient hraje klíčovou roli ve všech oblastech kvantové chemie, jeho implementace pro obecnou MRCC metodu je tedy zásadní.

Pro implementaci byl zvolen analytický gradient stavově specifické multireferenční Brillouinovy–Wignerovy metody CCSD s iterativní korekcí sizeextenzivity. Tato metoda je totiž implementována do programu ACES II [1], který již nabízí efektivní analytický gradient jednoreferenční CCSD metody a jeho kód lze tedy s výhodou využít.

Během práce bylo nutné použít obecnější program, založený na rozvoji do Full-CI báze, který však umožnil pilotní implementaci analytického gradientu pro MR BWCC metodu s iterativní korekcí a bez korekce size-extenzivity a též stavově univerzální MRCC metody. Pilotní implementace všech tří variant byla úspěšně otestována na karbenu CH_2 a silylenu SiH_2 a tato práce byla opublikována [2].

Analytický gradient energie MR BWCCSD metody s iterativní korekcí size-extenzivity byl nakonec též naimplementován do programu ACES II, ačkoliv tato implementace vyžaduje explicitní řešení CPHF rovnic pro vyjádření derivací LCAO koeficientů, což je prováděno externím programem.

Jako poslední část práce byla testována nově vyvinutá stavově specifická multireferenční Mukherjeeho metoda s poruchovým zahrnutím \hat{T}_3 amplitud v aproximaci nespřažených amplitudových rovnic, MR MkCCSD(T_u). Bylo studováno siglet-tripletové štěpení molekuly tetrametylenetanu. Získané křivky potenciální energie byly hladké a výsledky velmi přesně souhlasily s experimentem. Tato metoda nevyžaduje iterativní řešení \hat{T}_3 rovnic a přesto je imunní vůči intruder stavům, což potvrdil i průběh spočtených křivek. Mezi přesnými multireferenčními metodami je tato metoda proto velmi perspektivní. Tato práce je též součástí publikace [3].

Contents

1	Motivation				
	1.1	Aims of the Thesis	1		
	1.2	Analytical Gradient of Energy in Quantum Chemistry	1		
	1.3 Quantum Chemical Methods				
	1.4	Analytical Gradient in the Multireference Coupled Cluster			
		Methods	10		
	1.5	Perturbative Triples Contribution in the MRCC Methods	11		
2	The	eory	13		
	2.1	Generally Applied Approximations	13		
	2.2	The Hartree–Fock Method	14		
	2.3	Second Quantization	18		
		2.3.1 Creation and Annihilation Operators	19		
		2.3.2 Normal-Ordered Operators and the Wick's Theorem .	22		
		2.3.3 The Particle–Hole Formalism	24		
	2.4	The Coupled Cluster Method	26		
		2.4.1 The Connected Cluster Expansion	26		
		2.4.2 Size-Extensivity and Size-Consistency	28		
	2.5	The Hilbert Space Multireference Coupled Cluster Methods .	29		
		2.5.1 The State-Specific Multireference Brillouin–Wigner			
		Coupled Cluster method (MR BWCC)	30		
		2.5.2 The State-Universal Multireference Coupled Cluster			
		method (SU MRCC) \ldots	32		
		2.5.3 The State-Specific Multireference Mukherjee Coupled			
		Cluster method (MR MkCC) $\ldots \ldots \ldots \ldots \ldots$	32		
	2.6	Analytical Gradient of the Hartree–Fock Energy	33		
		2.6.1 Coupled Perturbed Hartree–Fock Equations (CPHF) .	36		
	2.7	Analytical Gradient of the Coupled Cluster Energy	38		
		2.7.1 The λ Equation	39		
	2.8	Analytical Gradient of the Hilbert Space MRCC Methods	40		
		2.8.1 Analytical Gradient of the MR BWCC with the Itera-			
		tive Correction of Size-Extensivity	42		
		2.8.2 Analytical Gradient of the MR BWCC without Cor-			
		rection of Size-Extensivity	43		

		Conci	usion	00		
	4.3	Conch	usion	00		
	4.2	Result	s and Discussion	64		
	4.1	Comp	utational	63		
	Gap) Inves	tigation of Tetramethyleneethane (TME)	62		
4	Application of the MR $MkCCSD(T_u)$ on a Singlet-Triplet					
		MR B	WCCSD with the Iterative Correction of Size–Extensivity	56		
	3.3	Efficie	nt Implementation of the Analytical Gradient of the			
	3.2	Testin	g of the Analytical Gradient Implementation	52		
	3.1	Pilot I	mplementation	51		
3	Implementation of the Analytical Gradient					
		2.9.2	$MR \ MkCCSD(T_u) \ Approximation \ . \ . \ . \ . \ . \ . \ . \ . \ . \ $	49		
		2.9.1	$MR \ MkCCSD(T_i) \ Method \ \ldots \ $	49		
	2.9	Pertur	bative Triples Contribution in the MR MkCC Method .	47		
		2.8.3	Analytical Gradient of the SU MRCC	45		

Chapter 1

Motivation

1.1 Aims of the Thesis

The principal aim of this thesis is to implement an analytical gradient of energy for a general multireference coupled cluster method. The computer implementation should be tested by comparing the analytical gradient with a gradient calculated numerically and finally applied to a geometry optimization of a molecular system with a multireference character.

Specifically, the state-specific multireference Brillouin–Wigner coupled cluster method with the iterative correction of size-extensivity, the MR BWCCSD, has been chosen since its amplitude equations at convergence are uncoupled and thus enable a relatively easy analytical gradient derivation. The method has been implemented in the ACES II program package, therefore the gradient implementation could also make use of the available single-reference analytical gradient code.

Another goal of the thesis is to test a newly developed perturbative triples correction to the state-specific multireference Mukherjee coupled cluster method in the uncoupled approximation, the MR MkCCSD(T_u) on an investigation of a singlet-triplet energy separation of the tetramethyleneethane molecule.

1.2 Analytical Gradient of Energy in Quantum Chemistry

Gradient of energy -a vector of the first partial derivatives of energy with respect to a set of coordinates

$$grad E = \left(\frac{\partial E}{\partial x_1}, \frac{\partial E}{\partial x_2}, \dots\right)$$
 (1.1)

or generally a partial derivative of energy with respect to a general perturbation χ : $\frac{\partial E}{\partial \chi}$ plays a fundamental role in all branches of quantum chemistry.

For instance, many physical properties of molecules can be expressed as first and higher derivatives of energy with respect to a set of specific coordinates. The simplest example can be spectroscopic constants – the electric dipole moment and polarizability, which is a vector of first derivatives and a matrix of second derivatives of energy with respect to external electric field strength coordinates. However, the most common quantum chemical tasks like prediction of equilibrium molecular geometry, vibrational frequencies calculation, transition state search or molecular dynamics, require evaluation of the energy derivatives with respect to position of nuclei.

First of all, chemists are interested in equilibrium molecular geometries. That is such a configuration of atoms in a molecule having the lowest energy of all possible atomic configurations. Finding a local minimum of energy with respect to several atomic coordinates, referred to as **geometry optimization**, can be numerically very well performed by a Newton-Raphson function minimization technique. This is an analog to the famous Newton method of finding roots of equations. It works by locally approximating the potential energy surface (PES) by a paraboloid (expressed as a quadratic form) and with the knowledge of a gradient and a Hessian (a matrix of all second partial derivatives) it reaches the minimum of the paraboloid in just one step. For a real PES, however, it requires several steps to reach the minimum by the formula

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{H}^{-1}(\mathbf{x}_n) \operatorname{grad} E(\mathbf{x}_n)$$
(1.2)

where $\mathbf{H}^{-1}(\mathbf{x}_n)$ is an inverse of the Hessian matrix calculated at the *n*th step.

Because of the expensive evaluation of the Hessian matrix, which scales quadratically with the number of coordinates and what's more, the second analytical derivatives are often unavailable for many quantum chemical methods, the Newton-Raphson technique is usually replaced by quasi-Newton methods (for instance the variant by Broyden, Fletcher, Goldfarb and Shanno — BFGS [4]) which need just the first derivatives and an approximated inverse of the Hessian which is being improved at each step.

The other most common application of the gradient in quantum chemistry is the calculation of **harmonic vibrational frequencies**. This is usually performed as an additional step after the geometry has been optimized to ensure the predicted geometry is really at a minimum and not in a saddle point of the PES. But besides is needed as a first approach for assigning infrared and Raman spectra, obtaining zero point vibrational energies and subsequent thermodynamics property calculations.

The frequencies are eigenvalues of a vibrational Hamiltonian which approximates the vicinity of the energetical minimum by a multidimensional parabolic potential

$$\hat{V} = E^{eq} + \frac{1}{2} \sum_{ij}^{3 \# Nuclei} q_i f_{ij} q_j$$
(1.3)

where the E^{eq} is the energy in the equilibrium geometry,

$$q_i = \sqrt{m_i} \left(x_i - x_i^{eq} \right) \tag{1.4}$$

is the mass weighted coordinate of the ith nucleus in terms of displacements from the equilibrium geometry and

$$f_{ij} = \left(\frac{\partial^2 E}{\partial q_i \,\partial q_j}\right)_{\mathbf{q}=\mathbf{0}} \tag{1.5}$$

are harmonic force constants. The potential is thus inherently described by a Hessian, which is, as already noted, often unavailable in an analytical form for many quantum chemical methods. Therefore it has to be evaluated numerically either from energies or, preferably, from first derivatives.

A transition state search is another example of an inevitable gradient use. It is a search for a state with the highest energy along a reaction path. The state lies in a saddle point connecting the valleys of the PES where the initial and final states of a reaction are located. The methods for finding the transition state makes use of the fact that the gradient is zero in a saddle point. One such a technique just minimizes a gradient norm about its most expected position on the PES. Another technique — a *hill climbing* predicts the reaction path by the least steep gradient corrected by minimizing the energy along all other directions. Similar but more accurate technique is based on a modified quasi-Newton method.

And lastly, the **classical molecular dynamics**, which studies the time evolution of a molecular system, is strongly dependent on the gradient evaluation. It solves the Newton equations of motion for atoms on the PES, where the gradient drives the atomic movement. A common method for integrating the Newton equations in the classical molecular dynamics is the *velocity Verlet* algorithm [5,6]:

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1}{2}a_i(t)\Delta^2 t$$
 (1.6)

$$v_i(t + \Delta t) = v_i(t) + \langle a_i \rangle \Delta t$$
(1.7)

$$\langle a_i \rangle = \frac{a_i(t) + a_i(t + \Delta t)}{2} \tag{1.8}$$

where the forces are given by the potential energy gradient

$$a_i(t) = -\frac{1}{m_i} \frac{\partial V(\mathbf{r})}{\partial r_i}$$
(1.9)

Let's now look at the formulas used for numerical evaluation of the gradient and the Hessian matrix and consider **the advantage of the analytical gradient** formula over the numerical differentiation. The symmetrized numerical gradient formula reads

$$\frac{\partial f(x,y)}{\partial x} = \lim_{h \to 0} \frac{f(x+h,y) - f(x-h,y)}{2h} \tag{1.10}$$

and the symmetrized formulas for the second partial derivatives using an analytical gradient

$$\frac{\partial^2 f(x,y)}{\partial x^2} = \lim_{h \to 0} \frac{\frac{\partial f(x+h,y)}{\partial x} - \frac{\partial f(x-h,y)}{\partial x}}{2h}$$
(1.11)

$$\frac{\partial^2 f(x,y)}{\partial x \partial y} = \lim_{h \to 0} \frac{\frac{\partial f(x,y+h)}{\partial x} - \frac{\partial f(x,y-h)}{\partial x}}{2h}$$
(1.12)

Finally the symmetrized second partial derivatives if an analytical gradient is not available

$$\frac{\partial^2 f(x,y)}{\partial x^2} = \lim_{h \to 0} \frac{f(x+h,y) - 2f(x,y) + f(x-h,y)}{h^2}$$
(1.13)

$$\frac{\partial^2 f(x,y)}{\partial x \partial y} = \lim_{h \to 0} \frac{1}{4h^2} \left[f(x+h,y+h) - f(x+h,y-h) - f(x-h,y+h) + f(x-h,y-h) \right]$$
(1.14)

The complexity of a full gradient and Hessian evaluation is presented in Table 1.1. We can see that as long as the complexity of the analytical gradi-

	anal. grad. available	anal. grad. not available
gradient	1G	2NE
Hessian	2NG	$(1+2N^2)E$

Table 1.1: A comparison of complexity of an N-dimensional gradient and Hessian matrix computation if an analytical gradient formula is available or absent. E stands for the complexity of energy evaluation and G for the complexity of an analytical gradient evaluation.

ent evaluation is comparable to the evaluation of the energy, the numerical calculation of the gradient from energies scales linearly with the number of degrees of freedom (e.g. number of atoms) whereas the analytical calculation is just constant. Similarly the numerical Hessian calculation using energies scales quadratically, while its evaluation from gradients scales only linearly.

This seems to be a solid argument for preferring the analytical gradient, but it would completely fail if the analytical gradient evaluation strongly depended on the number of perturbations (degrees of freedom). Fortunately, the energy gradient formulas can be factored to a dominant perturbationindependent part and a minor perturbation-dependent contributions so that the total evaluation time of the analytical gradient is of the same order as that of the energy.

Such formula factorization is usually referred to as the Z-vector technique and for the coupled cluster methods it leads to solving the λ -equations from which the set of perturbation-independent coefficients, called λ -amplitudes, is obtained.

The analytical gradient evaluation is also more accurate than its numerical calculation from finite differences and, moreover, avoids possible convergence difficulties of energy computations that usually happen at displacements which reduce the symmetry of the investigated molecular system.

1.3 Quantum Chemical Methods

Quantum chemistry, a quantum mechanical description of molecular systems, has to inherently deal with the many-body (many-electron) problem. It does so by expanding the electronic Hamiltonian \hat{H}_{el}^{-1} in a finite basis of antisymmetric N-electron functions and finding some of its eigenvalues (energies) and eigenvectors (usually only the ground state energy and wavefunction).

The antisymmetric N-electron basis functions are usually given by Slater determinants 2 (or their linear combination to satisfy properly a spin symmetry, for instance):

$$|\Phi_{0}\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_{1}(1) & \dots & \chi_{1}(N) \\ \chi_{2}(1) & \dots & \chi_{2}(N) \\ \vdots & \ddots & \vdots \\ \chi_{N}(1) & \dots & \chi_{N}(N) \end{vmatrix}$$
(1.15)

These are constructed from a set of one-electron basis functions, known as molecular spinorbitals $\chi_i(k)$ (*i* is a function index and *k* denotes a *k*-th electron), consisting of a three-dimensional spatial function multiplied formally by a spin function $|\alpha\rangle$ or $|\beta\rangle$.

The spatial component of the one-electron basis is mostly formed from a basis of atomic orbitals (usually Gaussian functions placed at centers of individual atoms) by solving the Schrödinger equation with an assumption that the molecular wavefunction is given by just one N-electron Slater determinant. Such a way of choosing the one-electron basis set is called the Hartree–Fock method (HF, SCF) and results in a set of orthonormal molecular orbitals, where the first N with the lowest orbital energy, called occupied orbitals, form the Hartree–Fock wavefunction (the single Slater determinant) and the remaining eigenfunctions are called virtual orbitals.

¹i.e. a Hamiltonian in the Born–Oppenheimer approximation

 $^{^{2}\}mathrm{a}$ determinantal form of an $N\text{-}\mathrm{electron}$ function ensures the antisymmetry of a fermionic wavefunction

The number of N-electron Slater determinants which can be constructed from a set of M orbitals is, however, extremely large. For example a full Nelectron basis set of a water molecule with 10 electrons in a closed-shell singlet electronic state using a moderately large one-electron basis of 19 functions (e.g. cc-pVTZ) consists of 135,210,384 Slater determinants. We can see that the size of the N-electron basis thus represents a principal obstacle in solving the many-body Schrödinger equation.

The advantage of the Hartree–Fock method for choosing the molecular orbitals is that since the N occupied orbitals provide the best possible description of the molecular system for the Hartree–Fock wavefunction (i.e. the only Slater determinant), the Slater determinants in the N-electron basis which are formed mostly by the occupied Hartree-Fock molecular orbitals play the significant role in the expansion of the proper N-electron wavefunction. And truly, the proper ground state N-electron wavefunction is mostly given by the reference Slater determinant (the Hartree–Fock one) and singly, doubly (and triply) excited configurations (i.e. such Slater determinants derived from the reference determinant where one, two or three occupied orbitals are replaced by a corresponding number of virtual orbitals). This fact helps to significantly reduce the size of the N-electron basis and makes the solution of the many-body problem feasible.

The Hartree–Fock method assumes a model of almost independent electrons ³. The post-Hartree–Fock methods described below thus try to involve the correlation of electronic motion in the proper wavefunction. Hence the difference between the proper energy \mathcal{E}_0 and the Hartree–Fock energy E_0 is called the *correlation energy*

$$E^{corr} = \mathcal{E}_0 - E_0 \tag{1.16}$$

The straightforward way of finding the spectrum of the many-body Hamiltonian by employing the complete space of N-electron Slater determinants formed from a finite set of molecular orbitals is the **full configuration interaction** method (Full-CI, FCI). The method is simply formulated but is extremely demanding on computational resources due to the enormous size of the N-electron basis and is therefore practically limited to calculations of systems of only a few atoms.

Nevertheless, using the property of the Hartree–Fock orbitals mentioned above, it is possible to dramatically reduce the N-electron space by involving only the reference determinant and all singly, doubly or more excited configurations without a significant loss of the wavefunction quality while using

³In the Hartree–Fock model, the electrons are actually not fully independent as each electron moves in an averaged field of the other electrons. The averaged field is calculated iteratively until a self-consistency is reached so that it no longer improves (it is why the Hartree–Fock method is also known as the SCF — a self-consistent field method). Moreover, two electrons of the same spin cannot occupy the same location at the same time, due to the fermionic character of electrons, which is assured by the determinantal form of the wavefunction

the same simple CI algorithm. This approach is known as the **truncated CI** (for instance CISD or CISDT) and can be applied to systems of up to several tens of atoms.

As an example you can see a CISD expansion of a wavefunction

$$|\Psi_{\text{CISD}}\rangle = |\Phi_0\rangle + \sum_{\substack{i \\ a}} c_i^a |\Phi_i^a\rangle + \sum_{\substack{i < j \\ a < b}} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle$$
(1.17)

using a common notation for the reference (Hartree–Fock) determinant $|\Phi_0\rangle$, singly $|\Phi_i^a\rangle$ and doubly $|\Phi_{ij}^{ab}\rangle$ excited configurations where *i* and *j* denote indices of occupied orbitals which are replaced by virtual orbitals with indices *a* and *b*. The expansion coefficient of the reference determinant is equal to one due to the usual *intermediate normalization* of a wavefunction

$$\langle \Psi | \Phi_0 \rangle = 1 \tag{1.18}$$

The great advantage of CI methods is that they are formulated by a *variational principle*, which is a simple consequence of linear algebra and is given by a relation

$$\mathcal{E}_{0} \leq E\left[|\Phi\rangle\right] = \frac{\left\langle \Phi \left| \hat{H} \right| \Phi \right\rangle}{\left\langle \Phi | \Phi \right\rangle} \tag{1.19}$$

where \mathcal{E}_0 is the true lowest eigenvalue of the Hamiltonian operator and $|\Phi\rangle$ is an arbitrary ket-vector of the operator space. The main benefit is that by truncated CI methods (but also by any other variational method) an upper bound limit to the true Full-CI energy is obtained.

However, the truncated CI is not a size-extensive method, which means that the energy does not scale correctly with the number of particles in a system. The reason why it fails for the truncated CI while the Full-CI is sizeextensive will be shown in the next chapter. To partially compensate this deficiency, the Davidson correction of size-extensivity [7,8] is always applied to truncated CI calculations.

A rather different approach for expanding the wavefunction in excited configurations represents the **coupled cluster** (CC) method which is based on the *exponential ansatz*

$$|\Psi_{\rm CC}\rangle = e^{\hat{T}}|\Phi_0\rangle = e^{\hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_N}|\Phi_0\rangle$$
 (1.20)

where the $\hat{T}_1, \hat{T}_2, \ldots$ are operators (called *cluster operators*) formally generating a series of all possible single, double and up to *N*-tuple excitations from a reference function upon which they act. For example

$$\hat{T}_1 |\Phi_0\rangle = \sum_{\substack{i\\a}} t_i^a |\Phi_i^a\rangle \tag{1.21}$$

$$\hat{T}_2 |\Phi_0\rangle = \sum_{\substack{i < j \\ a < b}} t_{ij}^{ab} |\Phi_{ij}^{ab}\rangle \tag{1.22}$$

where the coefficients t_i^a and t_{ij}^{ab} are called the *cluster amplitudes*.

By expanding the exponential operator in the Taylor series

$$e^{\hat{T}} = 1 + \frac{\hat{T}}{1!} + \frac{\hat{T}^2}{2!} + \dots$$
 (1.23)

we can see how a CI-like expansion of a wavefunction in configurations is formed. Let's take an example of the CCSD method, in which the excitation operator is truncated to involve just singles and doubles $\hat{T} = \hat{T}_1 + \hat{T}_2$

$$e^{\hat{T}_1+\hat{T}_2} = 1 + \hat{T}_1 + \hat{T}_2 + \frac{\hat{T}_1^2}{2!} + \frac{2\hat{T}_1\hat{T}_2}{2!} + \frac{\hat{T}_2^2}{2!} + \dots$$
 (1.24)

It is obvious that a CCSD expansion of a wavefunction includes not only singly and doubly, but partially also triply and highly excited configurations due to coupling of the excitation operators. And this is just one of the main points which make the coupled cluster method superior to the configuration interaction. The other is that the coupled cluster method is size-extensive regardless the level of truncation of the \hat{T} operator.

Nevertheless, it is not a variational method, which means that the calculated CC energy may be below the proper Full-CI value and that the cluster amplitudes (which define the wavefunction expansion) have to be calculated iteratively from a set of nonlinear equations. This not only makes the codes more complicated but also can lead to convergence problems. In spite of these disadvantages, the coupled cluster method is regarded as the most reliable in the family of *ab initio* quantum chemical methods.

It is also necessary to briefly mention the **perturbation theory**. Its principal idea is based on a partitioning of the Hamiltonian \hat{H} into a main (unperturbed) part \hat{H}_0 that is relatively easy to solve, and a small perturbation \hat{W}

$$\hat{H} = \hat{H}_0 + \hat{W} \tag{1.25}$$

If the perturbation term is small enough, the eigenstate of the full Hamiltonian can be obtained as a convergent series of correction terms to the unperturbed solution.

Usually the Møller–Plesset partitioning of the electronic Hamiltonian is applied so that the unperturbed solution is the Hartree–Fock wavefunction. The second-order correction, known as the MP2 method, is a very cheap and popular way of involving about 80% of electronic correlation.

Schematically the perturbation series for the correlation energy can be expressed as

$$E^{corr} = \sum_{m=0}^{\infty} \left\langle \Phi_0 \left| \hat{W} \left[\frac{\hat{1} - |\Phi_0\rangle \langle \Phi_0|}{z - \hat{H}_0} (z - \mathcal{E}_0 + \hat{W}) \right]^m \right| \Phi_0 \right\rangle$$
(1.26)

where the fractional expression is a resolvent, which depends on an arbitrary number z. The two most common variants used are the Rayleigh–Schrödinger perturbation theory (RSPT), where the z is set to the lowest eigenvalue of the unperturbed Hamiltonian \hat{H}_0 and the Brillouin–Wigner perturbation theory in which the z is equal to the true lowest eigenvalue \mathcal{E}_0 of the full (perturbed) Hamiltonian. The RSPT variant is a fully size-extensive method. The BWPT, by contrast, is not, but its formalism is much simpler.

The post-Hartree–Fock methods described above (apart from the Full-CI) involve properly only a specific part of the Full-CI N-electron space by starting from one reference configuration and incorporating in some specific way the other configurations by excitations of the reference determinant. However, in many cases this approach becomes insufficient and can completely fail to describe a chemical species even though a great portion of dynamical electronic correlation is included by expanding the wavefunction in a very large Full-CI subspace. This happens when a single determinant cannot correctly describe the spin symmetry (e.g. $^{1}\Delta$ state of O₂ molecule needs a two-determinantal description) or when a dissociation of a molecule is studied, in which the molecular wavefunction is continuously split in two parts, or generally if quasi-degeneracies occur. In all these cases, involvement of certain configurations is inevitable for a proper system description and its effect is usually referred to as inclusion of non-dynamical electronic correlation in a wavefunction. Such systems have to be treated with **multireference** methods.

The basic method which describes a system with more Slater determinants is the *multiconfigurational SCF method* (MCSCF) or its special case — the *complete active space SCF* (CASSCF), in which an appropriate combination of configurations is set and both their coefficients and the orbitals are optimized at the same time.

The multireference post-Hartree–Fock methods usually take the appropriate combination of determinants (references) using either the HF or MCSCF optimized orbitals and generate excitations for each reference analogously to the single-reference variant of the method. So there is the *multireference configuration interaction* method (MRCI), the *multireference perturbation theory* (such as the common CASPT2 method) and also the *multireference coupled cluster* method (MRCC).

In the context of this thesis, the Hilbert-space MRCC methods are used, namely the state-specific multireference Brillouin–Wigner CC (MR BWCC), the state-specific multireference Mukherjee CC (MR MkCC) and the stateuniversal multireference CC (SU MRCC). The state-specific methods concentrate on calculation of a single particular state. The state-universal method, by contrast, computes several states in one step, but is vulnerable to convergence problems.

1.4 Analytical Gradient in the Multireference Coupled Cluster Methods

Many quantum chemical methods have been provided with analytical gradients long time ago and since then the gradients are routinely used in everyday productive calculations.

While the analytical gradient of the single-reference coupled cluster methods have already been introduced in the 1980s and 1990s, there are still few analytical gradient implementations for the multireference coupled cluster theory. When the work on this thesis started, only a gradient of a special case of the two-determinantal open shell low spin singlet coupled cluster had been implemented [9] and gradients of another two MRCC variants had been developed but without an implementation [10]. There were analytical gradients of the Fock space MRCC formulation [11–14], but no general Hilbert space MRCC analytical gradient had actually been implemented. Because of the superiority of the CC method and inevitable capabilities of the multireference approach, an efficient implementation of a general Hilbert space MRCC method was therefore highly desirable.

In this thesis the state-specific MR BWCCSD method with the iterative correction of size-extensivity was chosen for the analytical gradient implementation. Based on the Brillouin–Wigner perturbation theory, the method is not size-extensive a priori, but is provided with a correction. This has, however, a convenient consequence in the form of its final MRCC equations, which lead to a relatively easy gradient derivation. Although the BW formalism avoids intruder state problems ⁴ while solving the CC equations, the iterative size-extensivity correction slowly transforms the formulas into a Rayleigh–Schrödinger form which may however give rise to intruder state problems and can lead to convergence difficulties. The method itself and also the gradient scale only linearly with the number of reference configurations so it might be applied to investigation of larger molecular systems. The gradient has been derived in such a way that much of the existing efficient single-reference CCSD analytical gradient code of the ACES II [1] program package could be reused since it has also an efficient MR BWCC implementation.

During the work, however, it turned out to be necessary to first code a pilot implementation of the gradient as an inevitable tool that helps with the efficient code reuse. And since this pilot implementation was performed in a generic Full-CI based way, another two analytical gradient variants could be

 $^{{}^{4}}$ Intruder states are numerical inconveniences in the perturbation theory formalism when the denominator of the resolvent becomes too close to zero

implemented as well. In the end a pilot analytical gradient implementation of the state-specific MR BWCCSD with the iterative size-extensivity correction and without the correction and a state-universal MRCCSD in the formulation by Kucharski and Bartlett [15] were completed and published [2].

It has to be noted that only after the pilot implementation had been published, an efficient implementation of the state-specific multireference Mukherjee coupled cluster (MR MkCC) energy gradient appeared, coded by Prochnow *et al.* [16], which is however a more general case of the iteratively corrected MR BWCC gradient.

Nevertheless, the analytical energy gradient of the iteratively corrected MR BWCCSD variant has finally been efficiently implemented as well, although the orbital response contribution requires an explicit solution of the CPHF equations.

1.5 Perturbative Triples Contribution in the MRCC Methods

While in many cases the coupled cluster calculations with contributions from the singles \hat{T}_1 and doubles \hat{T}_2 are sufficient enough for a correct description of a molecular system, for precise quantitative results also contributions from the triples \hat{T}_3 have to be incorporated.

However, while the relatively cheap CCSD calculation, which scales with $\mathcal{O}(\mathcal{N}^6)$, where \mathcal{N} means the number of orbitals, the full CCSDT method is two magnitudes more expensive, scaling with $\mathcal{O}(\mathcal{N}^8)$. Therefore an approximate triples contribution is usually employed which is known as the CCSD(T) method and is commonly referred to as "the gold standard" for quantitative quantum chemical calculations. In most cases it provides results comparable with experimental values while remaining computationally feasible. In this method, the equations for solving the CCSD amplitudes are iterated and after the convergence has been reached, a few other terms (which are chosen by inspecting the formulas by the perturbation theory) are added to approximate the triples contribution. The method thus requires only one additional step that scales with $\mathcal{O}(\mathcal{N}^7)$.

In the multireference coupled cluster theory, it was first developed for the SU MRCC method [17]. The MR BWCCSD(T) approximation was implemented a few years ago [18]. Recently, the MR MkCC method was provided with a non-iterative triples variant, denoted by MR MkCCSD(T_n), which however introduces intruder state problems [19]. Later a different approach, denoted here as MR MkCCSD(T_i), has been suggested which eliminates the intruder state problem, but requires an iterative solution for the triples amplitude equation [20]. Another formulation finally appeared, the MR MkCCSD(T_u), which is based on an uncoupled approximation of the MR MkCC method and avoids both the intruder state problem and the iterative solution for the triples equation [3]. The method has already been successfully tested on the BeH_2 system, but for testing on a real chemical application, a singlet-triplet gap investigation of the tetramethyleneethane molecule (TME) has been suggested, which is actually the topic of the last part of this thesis.

Chapter 2

Theory

2.1 Generally Applied Approximations

Let us introduce the physical constraints upon which all the theory of the thesis is built.

First of all, only a **non-relativistic physics** is applied since in most quantum chemical calculations, which investigate elements of the first three rows of the periodic table, the relativistic effects are still almost negligible in comparison with the calculated properties of the main interest. Thus only the Schrödinger equation is employed, which makes the situation much easier than using the relativistic Dirac equation.

The molecular system is thus fully described by a non-relativistic Hamiltonian

$$\hat{H} = \underbrace{-\frac{1}{2}\sum_{A} m_{A}\Delta_{A}}_{\hat{T}_{n}} + \underbrace{-\frac{1}{2}\sum_{i}\Delta_{i}}_{\hat{T}_{e}} + \underbrace{\sum_{A < B} \frac{Z_{A}Z_{B}}{r_{AB}}}_{\hat{V}_{nn}} + \underbrace{\sum_{iA} \frac{Z_{A}}{r_{iA}}}_{\hat{V}_{ne}} + \underbrace{\sum_{i < j} \frac{1}{r_{ij}}}_{\hat{V}_{ee}} \quad (2.1)$$

where the usual notation for kinetic \hat{T} and potential energy \hat{V} is used as well as symbols e and n for electrons and nuclei respectively. This and all the other equations in the thesis also assume the use of atomic units.

Nevertheless, even with the Schrödinger equation, the relativistic effects could be in some extent involved by using the *effective core potentials*, describing an effective field of the core atomic electrons.

Because of the relatively large ratio between the weight of nuclei and electrons, also the **Born–Oppenheimer approximation** (BOA) is applied in most cases. It simply removes the nuclear kinetic energy operator \hat{T}_n from the total Hamiltonian, forming the so-called *electronic Hamiltonian* \hat{H}_{el}

$$\hat{H}_{el} = \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ne} + \hat{V}_{ee}$$
(2.2)

This actually describes an electronic system in the field of fixed nuclei, which obeys the Schrödinger equation

$$\hat{H}_{el}(\mathbf{R})|\Phi(\mathbf{r};\mathbf{R})\rangle = \mathcal{E}(\mathbf{R})|\Phi(\mathbf{r};\mathbf{R})\rangle$$
(2.3)

in which the **r** denotes electronic coordinates and the **R** means that the electronic Hamiltonian and its eigenvalues and eigenfunctions depend on the nuclear coordinates parametrically. The eigenvalue $\mathcal{E}(\mathbf{R})$ thus represents a potential energy (hyper)surface (PES) on which the nuclei move.

Solution of the Born–Oppenheimer Hamiltonian is usually sufficient enough. It may, however, fail in situations where two potential energy surfaces of the same symmetry come close together or even cross, forming the *conical intersection*.

In the rest of this work, because of the implicit use of the Born– Oppenheimer approximation, the electronic Hamiltonian will be further denoted with just the symbol \hat{H} .

And lastly, although the *N*-electron Hilbert space is infinite, for computational reasons, only **finite basis sets** can be employed. For molecular systems, however, they can be chosen in a way that is satisfactory enough for comparison with experimental values. Moreover, special basis sets also exist that successfully give physical properties as if an infinite basis were used. It is achieved by extrapolation of results obtained from calculations using several finite basis sets of an increasing size.

2.2 The Hartree–Fock Method

As already noted earlier, the Hartree–Fock method [21, 22] is used to obtain an appropriate set of one-electron basis functions, the molecular spinorbitals $\chi_i(\mathbf{x})$.

The spinorbital consists of a spatial orbital $\phi_i(\mathbf{r})$ and a spin component $|\alpha\rangle$ or $|\beta\rangle$

$$\chi_i(\mathbf{x}) = \begin{cases} \phi_i(\mathbf{r}) |\alpha\rangle \\ \phi_i(\mathbf{r}) |\beta\rangle \end{cases}$$
(2.4)

The spinorbital variable \mathbf{x} thus involves both the spatial and the spin variables altogether.

The spin components are only formal, since they are used to construct a proper spin of the N-electron wavefunction, which is given and fixed. Parameters of the spatial molecular orbital $\phi_i(\mathbf{r})$ is actually what is being optimized in the Hartree–Fock procedure.

The spatial molecular orbitals (MO) are usually formed from a fixed set of atomic orbitals (AO) as their linear combination (LCAO)

$$\phi_i(\mathbf{r}) = \sum_{\mu}^{AO} |\mu(\mathbf{r})\rangle C_{\mu i}$$
(2.5)

where the LCAO coefficients $C_{\mu i}$ are just the Hartree–Fock optimization variables.

Intuitively, the atomic orbitals are functions located at nuclei, which should mimic the eigenfunctions of a Hydrogen-type atoms (i.e. systems of a nucleus and only one electron). Such simplified functions are the Slater-type orbitals (STO)

$$\phi_{ijk}^{STO}(\mathbf{r}) = x^i y^j z^k e^{-\zeta |\mathbf{r} - \mathbf{r_0}|} \tag{2.6}$$

which are, however, rarely used because of their too expensive electron repulsion integral evaluation. Instead, the Gaussian-type orbitals (GTO) are used

$$\phi_{ijk}^{GTO}(\mathbf{r}) = x^i y^j z^k e^{-\zeta(\mathbf{r}-\mathbf{r_0})^2} \tag{2.7}$$

which can be integrated much more easily, though they do not describe the one-electron wavefunction so well. To compensate this deficiency, the contracted GTOs are used, which is just a fixed linear combination of several primitive GTOs (of the same polynomial part, but differing in the exponent)

$$\phi_{ijk}^{CGTO}(\mathbf{r}) = x^i y^j z^k \sum_m K_m e^{-\zeta_m (\mathbf{r} - \mathbf{r_0})^2}$$
(2.8)

which describe the spatial orbitals better and remain computationally feasible. There are numerous atomic basis sets of various size, optimized for calculation of a variety of physical properties (mainly the energy) with the exponents and contraction coefficients tabelized.

The functions shown above (2.7), (2.8) are called cartesian Gaussian-type orbitals and are characterized by a shell number l = i + j + k. All functions in shells l = 0, 1 are eigenfunctions of the orbital momentum operator \hat{L}^2 (i.e. s and p functions respectively). The cartesian GTO functions of shells with $l \ge 2$ however mix also functions of lower orbital momentum. For example there are 6 cartesian GTOs of l = i + j + k = 2. But they can actually be reduced to 5 d and one s function. If a particular quantum chemical program allows, a spherical GTO basis set can be constructed from a cartesian one, decreasing the size of the one-electron basis a little.

The Hartree–Fock molecular orbitals are obtained by the variational principle by taking the N-electron wavefunction $|\Phi\rangle$ as a single Slater determinant only and varying the LCAO coefficients while keeping the orbitals orthonormal so that the minimum of the wavefunction energy functional is reached.

$$E^{\rm HF} = \left\langle {\rm HF} \left| \hat{H} \right| {\rm HF} \right\rangle = \min E \left[\Phi \right] = \min \left\langle \Phi \left| \hat{H} \right| \Phi \right\rangle \tag{2.9}$$

This way, within a given AO basis set, no better single determinantal N-electron wavefunction can be obtained.

The simplest case is the closed-shell restricted Hartree–Fock method (RHF) in which each spatial molecular orbital is shared between a pair of an $|\alpha\rangle$ and $|\beta\rangle$ spinorbitals.

To calculate also the open-shell systems (i.e. with unpaired electrons), this restriction can be removed to form the unrestricted Hartree–Fock method (UHF) which leads to two coupled sets of equations for $|\alpha\rangle$ and $|\beta\rangle$ spinorbitals. Despite still remaining quite simple, the UHF method actually does not give a wavefunction of a pure spin. In contrast to the RHF method, it can however qualitatively well describe a dissociation of a molecule.

The restricted open-shell Hartree-Fock method (ROHF) uses a spinadapted linear combination of a few Slater determinants to provide a proper eigenfunction of the spin (\hat{S}^2) operator, but its formalism is the most complicated.

To reach the Hartree–Fock energy functional minimum (2.9), the functional variation must be zero

$$\delta E\left[\Phi\right] = \delta \left\langle \mathrm{HF} \left| \hat{H} \right| \mathrm{HF} \right\rangle = 0 \tag{2.10}$$

With an additional requirement for the molecular orbitals to be orthonormal

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \tag{2.11}$$

the method of Lagrange multipliers directly leads to equations of the Fock operator \hat{f} (specified below)

$$\hat{f}|\phi_i\rangle = \sum_{j=1}^N \lambda_{ij} |\phi_j\rangle \tag{2.12}$$

Because this operator is invariant under a unitary transformation, such a transformation can be found which diagonalizes the matrix of the multipliers and gives the *canonical Hartree–Fock equations*

$$\hat{f}|\phi_i\rangle = \varepsilon_i |\phi_i\rangle \tag{2.13}$$

from which the set of canonical molecular orbitals $|\phi_i\rangle$ is obtained where ε_i is called the energy of the *i*-th orbital.

The Fock operator is an effective one-electron energy operator defined (for the simplest closed-shell RHF method) as

$$\hat{f}(1) = \hat{h}(1) + \sum_{j=1}^{N} \left(2\hat{J}_j - \hat{K}_j \right)$$
(2.14)

where

$$\hat{h}(1) = -\frac{1}{2}\Delta_1 - \sum_{A}^{nuclei} \frac{Z_A}{r_{1A}}$$
(2.15)

is the one-electron operator for the electronic kinetic energy and the interaction between the electron and the nuclei,

$$\hat{J}_{j}(1) |\phi_{i}(1)\rangle = \int d\mathbf{r}_{2} \,\phi_{j}^{*}(2)\phi_{j}(2) \frac{1}{r_{12}}\phi_{i}(1)$$
(2.16)

is the *Coulomb operator* and

$$\hat{K}_{j}(1) |\phi_{i}(1)\rangle = \int d\mathbf{r}_{2} \,\phi_{j}^{*}(2)\phi_{i}(2) \frac{1}{r_{12}}\phi_{j}(1)$$
(2.17)

is the exchange operator. The last two operators just represent an averaged field of the other electrons. From their definition it follows that the Fock operator itself depends on the molecular orbitals. The Hartree–Fock equations (2.13) are thus pseudo-eigenvalue equations which have to be solved iteratively. From an initial orbital guess the Fock operator calculates the averaged field of the electrons and a new set of orbitals is generated. This procedure is repeated until the self-consistency is reached, which means that the averaged field as well as the orbitals no longer change.

As stated earlier, the molecular orbitals are expanded in the set of atomic orbitals by LCAO coefficients (2.5). Substituting (2.5) in (2.13) one obtains a set of *Hartree–Fock–Roothaan* equations that are actually used in practical calculations

$$\sum_{\nu}^{AO} F_{\mu\nu} C_{\nu i} = \varepsilon_i \sum_{\nu}^{AO} S_{\mu\nu} C_{\nu i}$$
(2.18)

where $F_{\mu\nu}$ and $S_{\mu\nu}$ are the Fock and overlap matrix in the AO basis. The equations are often written in a matrix form

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \tag{2.19}$$

where the $\boldsymbol{\varepsilon}$ is a diagonal matrix of orbital energies.

After the solution converges to self-consistency, the N occupied and M - N virtual orbitals are produced, where M is the size of the AO basis. In post-HF methods, which calculate the electron correlation, an N-electron function space is generated by excitations of electrons from the occupied orbitals to the virtual ones, as described in the previous part. Excitations from the occupied orbitals that represent the core (non-valence) shell electrons however contribute only minimally to the total electronic correlation and are often *frozen* (i.e. excluded from excitations). Freezing the core orbitals as well as their corresponding counterparts in the virtual orbital set further significantly reduces the size of the N-electron configuration space.

At the end of this section and before going further in theoretical formalism, a note about the common notation used in quantum-chemical formulas should be briefly presented.

Orbital indices i, j, k, l denote the occupied, while a, b, c, d the virtual molecular orbitals. Indices p, q, r, s are used for general (i.e. either occ. or virt.) molecular orbitals. Greek letters $\mu, \nu, \lambda, \sigma$, however, represent the atomic orbital indices. For example a matrix element of the one-electron part of the Hamiltonian in the AO basis is

$$h_{\mu\nu} = \left\langle \phi_{\mu} \left| \hat{h} \right| \phi_{\nu} \right\rangle \tag{2.20}$$

The two-electron integrals (i.e. the electron repulsion integrals) in the molecular spinorbital basis are denoted by

$$\langle pq|rs\rangle = \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_p^*(\mathbf{x}_1) \chi_q^*(\mathbf{x}_2) \frac{1}{r_{12}} \chi_r(\mathbf{x}_1) \chi_s(\mathbf{x}_2)$$
(2.21)

and a common shorthand for their antisymmetrized form

$$\langle pq || rs \rangle = \langle pq | rs \rangle - \langle pq | sr \rangle$$
 (2.22)

However, integrals over spatial orbitals (either atomic or molecular) use braces instead of brackets and gather orbitals with the same electronic variable

$$(pq|rs) = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_p^*(\mathbf{r}_1) \phi_q(\mathbf{r}_1) \frac{1}{r_{12}} \phi_r^*(\mathbf{r}_2) \phi_s(\mathbf{r}_2)$$
(2.23)

Common Coulomb and exchange integrals are thus

$$J_{ij} = (ii|jj) \tag{2.24}$$

$$K_{ij} = (ij|ji) \tag{2.25}$$

Using this notation, the closed shell RHF energy can be expressed as

$$E^{\text{RHF}} = 2\sum_{i}^{N/2} (i|\hat{h}|i) + \sum_{ij}^{N/2} [2(ii|jj) - (ij|ji)]$$
(2.26)

or in the spinorbital basis

$$E^{\text{RHF}} = \sum_{i}^{N} h_{ii} + \frac{1}{2} \sum_{ij}^{N} \langle ij || ij \rangle \qquad (2.27)$$

And the widely used Fock matrix elements read

$$f_{pq} = h_{pq} + \sum_{i}^{N} \langle pi || qi \rangle$$
(2.28)

2.3 Second Quantization

Second quantization [21, 23–25] is a notion coming from quantum electrodynamics, which however (along with the diagrammatic technique) became very popular with the many-body perturbation theory and related topics. It makes investigation of methods in the field much easier than the classical approach since the second-quantized operators of physical observables (like \hat{H}_{el}) are independent of the number of electrons while the antisymmetry of the electronic wavefunction is assured in an elegant way without the need to directly manipulate Slater determinants.

A brief introduction to the second quantization technique is presented because for explanation of the key concepts of the coupled cluster theory, the use of the second quantized formalism cannot be avoided. Otherwise the description as well as the algebraic tractability of the formulas would be extremely difficult.

2.3.1 Creation and Annihilation Operators

The key concept represents the normalized vacuum state $|\rangle$ which is just a wavefunction of zero electrons. Electrons can be created in the vacuum state by the creation operators a_p^+

$$a_p^+ \mid \rangle = \mid p \rangle \tag{2.29}$$

and annihilated again by the annihilation operators a_p

$$a_p \left| p \right\rangle = \left| \right\rangle \tag{2.30}$$

which are mutually related as Hermitian conjugates

$$a_p = \left(a_p^+\right)^\dagger \tag{2.31}$$

The operators obey the following anticommutation rules

$$\left[a_{p}^{+}, a_{q}^{+}\right]_{+} = 0 \tag{2.32}$$

$$[a_p, a_q]_+ = 0 (2.33)$$

$$\left[a_p^+, a_q\right]_+ = \delta_{pq} \tag{2.34}$$

which actually imply the N-electron wavefunction to be inherently antisymmetric. The application of the operators can thus be summarized

$$a_p^+ | q \dots s \rangle = | pq \dots s \rangle \tag{2.35}$$

$$a_p | pq \dots s \rangle = | q \dots s \rangle \tag{2.36}$$

$$a_p^+ a_q^+ \dots a_s^+ |\rangle = |pq \dots s\rangle \tag{2.37}$$

$$|pq\ldots\rangle = a_p^+ a_q^+ |\ldots\rangle$$

= $-a_q^+ a_p^+ |\ldots\rangle$
= $-|qp\ldots\rangle$ (2.38)

$$a_q |pqr...\rangle = -a_q |qpr...\rangle$$

= -|pr...\rangle (2.39)

$$|ab\dots\rangle = a_a^+ a_b^+ a_j a_i |ij\dots\rangle$$
(2.40)

$$a_p^+|p\rangle = 0 \tag{2.41}$$

$$a_p|\rangle = 0 \tag{2.42}$$

$$a_{p} |q\rangle = a_{p}a_{q}^{+} |\rangle$$

$$= \delta_{pq} - a_{q}^{+}a_{p}|\rangle$$

$$= \delta_{pq} - 0 \qquad (2.43)$$

and one can see that an N-electron wavefunction $|pq...s\rangle$ can be represented by a Slater determinant formed from orbitals pq...s each occupied by an electron.

Using the creation and annihilation operators, one can, for instance, conveniently express a wavefunction expansion in the configuration space. For example the CI expansion is given by

$$|\Psi_{\rm CI}\rangle = \left(1 + \hat{C}_1 + \hat{C}_2 + \dots + \hat{C}_N\right)|\Phi_0\rangle \tag{2.44}$$

where

$$\hat{C}_1 = \sum_{\substack{i \\ a}} c_i^a a_a^+ a_i \tag{2.45}$$

$$\hat{C}_{2} = \sum_{\substack{i < j \\ a < b}}^{\infty} c_{ij}^{ab} a_{a}^{+} a_{b}^{+} a_{j} a_{i}$$
(2.46)

etc. are excitation operators ¹ generating all singly, doubly and up to Ntuply excited configurations from the reference determinant $|\Phi_0\rangle$ together with their CI expansion coefficients.

And the CC exponential ansatz can be easily written as

$$\begin{aligned} |\Psi_{\rm CC}\rangle &= e^{\hat{T}} |\Phi_0\rangle \\ &= e^{\hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_N} |\Phi_0\rangle \\ &= \left[1 + \left(\hat{T}_1 + \dots + \hat{T}_N\right) + \frac{1}{2}\hat{T}_1^2 + \hat{T}_1\hat{T}_2 + \dots \right] |\Phi_0\rangle \quad (2.47)
\end{aligned}$$

¹Note the order of annihilators in the \hat{C}_2 operator; compare with (2.40)

where

$$\hat{T}_{1} = \sum_{\substack{i \\ a}} t_{i}^{a} a_{a}^{+} a_{i} \tag{2.48}$$

$$\hat{T}_{2} = \sum_{\substack{i < j \\ a < b}} t_{ij}^{ab} a_{a}^{+} a_{b}^{+} a_{j} a_{i}$$
(2.49)

etc. are similar excitation operators, but in the CC expansion (2.47), they are commonly referred to as the *cluster operators* and the expansion coefficients $t_i^a, t_{ij}^{ab}, \ldots$ are called the *cluster amplitudes*.

Note an important feature of the cluster operators. Since a cluster operator contains annihilators of occupied and creators of virtual orbitals, from (2.34)

$$\left[a_{a}^{+}, a_{i}\right]_{+} = \delta_{ai} = 0 \tag{2.50}$$

we see that all second quantization operators in the cluster operators anticommute. And because there is always an even number of creators or annihilators in the cluster operators, all the cluster operators thus commute.

To do physics, however, a second-quantized operators of physical observables should be introduced. It can be shown that the second-quantized form of a one-particle operator \hat{O}_1 is given by

$$\hat{O}_1 = \sum_{i}^{N} \hat{h}(i) = \sum_{pq} \left\langle p \left| \hat{h} \right| q \right\rangle a_p^+ a_q$$
(2.51)

whereas a two-particle operator \hat{O}_2 becomes

$$\hat{O}_{2} = \sum_{i < j}^{N} \hat{v}(i, j) = \frac{1}{2} \sum_{pqrs} \langle pq | \hat{v} | rs \rangle \ a_{p}^{+} a_{q}^{+} a_{s} a_{r}$$
(2.52)

A second-quantized electronic Hamiltonian thus reads

$$\hat{H}_{el} = \sum_{pq} h_{pq} a_p^+ a_q + \frac{1}{2} \sum_{pqrs} \langle pq | rs \rangle \ a_p^+ a_q^+ a_s a_r$$
(2.53)

$$= \sum_{pq} h_{pq} a_p^+ a_q + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle a_p^+ a_q^+ a_s a_r$$
(2.54)

A matrix element of any operator (which is actually a string of creation and annihilation operators between the vacuum bra and ket) can be straightforwardly evaluated using the anticommutation rules by moving the annihilators to the right. Because annihilating an electron from a vacuum state gives zero, such procedure would finally result in Kronecker deltas only. This can be demonstrated on an overlap of two mutually singly-excited two-electron Slater determinants

$$|\Phi_0\rangle = |ij\rangle = a_i^+ a_j^+|\rangle \qquad (2.55)$$

$$|\Phi_i^a\rangle = |aj\rangle = a_a^+ a_j^+|\rangle; \ \langle\Phi_i^a| = \langle |a_j a_a \tag{2.56}$$

Their overlap reads

Matrix elements of more complicated operator strings like Hamiltonian between various determinants would be however very tedious to evaluate in this way. A more convenient method for evaluation of matrix elements is described in the following subsection.

2.3.2 Normal-Ordered Operators and the Wick's Theorem

It has just been shown that in order to evaluate a matrix element of an operator string, the annihilation operators had to be moved to the right of the creation operators by the anticommutation rules so that they give zero by acting on a vacuum state and also to reduce one operator pair of the string to a Kronecker delta. An operator string in which all the annihilators are to the right of all the creators is said to be *normal ordered* and its matrix element is obviously zero. The *Wick's theorem* gives a rule how to easily transform an arbitrary operator string into Kronecker deltas and normal ordered strings so that matrix element evaluation becomes simple even for very complicated operator strings.

Let us first introduce a contraction of an operator pair (see Fig. 2.1) by connecting two operators with a line. The braces around a string $\{\ldots\}$ mean that the string inside is permuted to be normal ordered with a sign corresponding to a parity of such permutation. So a contraction of an already normal ordered string gives zero, otherwise a Kronecker delta is produced.

The Wick's theorem (Fig. 2.2) says that an arbitrary string of operators can be expanded in normal ordered strings with zero and all single, double and up to N-tuple (full) contractions. The great consequence of the theorem is that a matrix element of an operator string must be just the sum of only the fully contracted normal ordered strings (see Fig. 2.3), because the nonfully contracted normal ordered operators would annihilate the ket vacuum.

$$\overline{AB} = AB - \{AB\}$$

$$\overline{a_p^+ a_q^+} = 0$$

$$\overline{a_p^- a_q} = 0$$

$$\overline{a_p^+ a_q} = 0$$

$$\overline{a_p^- a_q^+} = a_p a_q^+ - (-a_q^+ a_p) = \delta_{pq}$$

$$ABC \dots XYZ = \{ABC \dots XYZ\} + \\ + \sum_{singles} \{\overline{ABC} \dots XYZ\} + \\ + \sum_{doubles} \{\overline{ABC} \dots XYZ\} + \\ + \dots$$

Figure 2.2: The Wick's theorem

$$\langle |ABC \dots XYZ| \rangle = \sum_{\substack{full\\contractions}} \{ ABC \dots XYZ \}$$

Figure 2.3: A matrix element evaluation by the Wick's theorem

$$\left\langle \left| a_{j}a_{a}a_{i}^{+}a_{j}^{+} \right| \right\rangle = \left\{ \overline{a_{j}a_{a}a_{i}^{+}a_{j}^{+}} \right\} + \left\{ \overline{a_{j}a_{a}a_{i}^{+}a_{j}^{+}} \right\}$$
$$= -\delta_{ji}\delta_{aj} + \delta_{jj}\delta_{ai}$$
$$= \delta_{ai}$$

Figure 2.4: An example of a matrix element evaluation by the Wick's theorem

The example of the overlap of two determinants that was shown earlier (2.57) can be evaluated by the Wick's theorem very easily (see Fig. 2.4) since there are only two fully contracted strings, one of which gives immediately zero.

2.3.3 The Particle–Hole Formalism

Although everything necessary for the second-quantized matrix element evaluation has already been introduced, the formalism can further be improved so that formula manipulations is even more convenient.

In the post-Hartree–Fock methods the wavefunction expansion usually starts from the reference configuration $|\Phi_0\rangle$ rater than a true vacuum state. It means that the operator strings in many matrix elements contain similar substrings of creation and (annihilation) operators corresponding to the ket (and bra) configurations, differing from the reference state by only a few excitations.

By redefining the meaning of the creation and annihilation operators slightly, those common substrings can be reduced to only a few operators which describe the difference between the particular ket (or bra) configuration and the reference state. Such operator elimination leads to much shorter strings that have to be evaluated and further simplifies the formalism.

Let us first introduce the *Fermi vacuum* $|\Phi_0\rangle$, which represents the reference configuration (i.e. a vacuum state filled with N electrons in a set of occupied orbitals)

$$|\Phi_0\rangle = a_i^+ a_j^+ \dots |\rangle \tag{2.58}$$

So far the creation operators generated an electron by acting on a vacuum state. From now on, however, a creator means an operator that either generates a *particle* (i.e. an electron in a virtual orbital) or creates a *hole* (i.e. annihilates an electron from an occupied orbital). The annihilators are redefined analogously. They produced zero by acting on the true vacuum but now the annihilators give zero by acting on the Fermi vacuum. It means that they either annihilate a particle (i.e. an electron from a virtual orbital) or annihilate a hole (i.e. create an electron in an occupied orbital). The redefinition is schematically summarized in Table 2.1 with the usual notation for hole indices i, j, \ldots (i.e. occupied orbital), particle indices a, b, \ldots (i.e. virtual orbitals) and general orbital indices p, q, \ldots .

Creates	Annihilates	
a_a^+	a_a	particle
a_i	a_i^+	hole

Table 2.1: Definition of creation and annihilation operators in the particlehole formalism

An example of a particle creation

$$a_a^+|ijk\rangle = |aijk\rangle \tag{2.59}$$

a hole creation

$$a_i |ijk\rangle = |jk\rangle \tag{2.60}$$

and a hole annihilation

$$a_i^+|jk\rangle = |ijk\rangle \tag{2.61}$$

Similarly, the redefined normal ordered string is such that puts the redefined annihilation operators to the right of all creators. For example

$$a_a^+ a_i^+ |ijk\rangle = 0 \tag{2.62}$$

demonstrates how a redefined normal ordered string gives zero by acting on the Fermi vacuum.

The same applies to the contraction definition and finally to the Wick's theorem too.

The \hat{H}_{el} is usually written in a normal ordered form using the particlehole formalism. By applying the Wick's theorem to the Hamiltonian

$$\hat{H}_{el} = \sum_{pq} h_{pq} \, a_p^+ a_q + \frac{1}{4} \sum_{pqrs} \langle pq \, || \, rs \rangle \, a_p^+ a_q^+ a_s a_r \tag{2.63}$$

and considering all the cases where the general orbital indices run over particles and holes separately, one finally gets an interesting result

$$\hat{H}_{el} = \underbrace{\sum_{pq} f_{pq} \{a_p^+ a_q\}}_{\hat{F}_{N}} + \underbrace{\frac{1}{4} \sum_{pqrs} \langle pq \mid \mid rs \rangle \ \{a_p^+ a_q^+ a_s a_r\}}_{\hat{V}_{N}} + \underbrace{\left\langle \Phi_0 \left| \hat{H}_{el} \right| \Phi_0 \right\rangle}_{E^{\text{HF}}} \quad (2.64)$$

or

$$\hat{H}_{\rm N} = \hat{H}_{el} - \left\langle \Phi_0 \left| \hat{H}_{el} \right| \Phi_0 \right\rangle \tag{2.65}$$

so the normal ordered electronic Hamiltonian $\hat{H}_{\rm N}$ can also be considered as a correlation operator.

2.4 The Coupled Cluster Method

The origin of the Coupled Cluster method dates back to 1958 where the principles were first introduced in nuclear physics [26,27]. In 1966 [28] it also appeared in quantum chemistry as the *Coupled Pair Many Electron Theory* (CPMET)² and only in the late 1970s [29,30] the computer implementations allowed the CC method to be applied also to systems of real chemical interest. The exponential ansatz (2.47) of the wavefunction has an analogy in the cluster expansion of the configuration integral in statistical physics and hence the method got its name.

The coupled cluster correlation energy expression can be obtained from the Schrödinger equation by using the normal ordered electronic Hamiltonian. The Schrödinger equation thus becomes

$$\hat{H}_{\rm N} e^{\hat{T}} |\Phi_0\rangle = \underbrace{\left(E^{\rm CC} - E^{\rm HF}\right)}_{\Delta E^{\rm CC}} e^{\hat{T}} |\Phi_0\rangle \qquad (2.66)$$

and by multiplying it from left by the $e^{-\hat{T}}$ operator and projecting by the Fermi vacuum, the energy yields

$$\Delta E^{\rm CC} = \left\langle \Phi_0 \left| e^{-\hat{T}} \hat{H}_{\rm N} e^{\hat{T}} \right| \Phi_0 \right\rangle \tag{2.67}$$

In a similar way, projecting by excited configurations leads to sets of equations for the cluster amplitudes. Thus the \hat{T}_1 equations read

$$\left\langle \Phi_{i}^{a} \left| e^{-\hat{T}} \hat{H}_{\mathrm{N}} e^{\hat{T}} \right| \Phi_{0} \right\rangle = 0$$
(2.68)

 \hat{T}_2 equations

$$\left\langle \Phi_{ij}^{ab} \left| e^{-\hat{T}} \hat{H}_{\mathrm{N}} e^{\hat{T}} \right| \Phi_0 \right\rangle = 0 \tag{2.69}$$

and analogically also the equations for higher cluster amplitudes.

2.4.1 The Connected Cluster Expansion

The reason why the energy expression (2.67) and the equations for the cluster amplitudes (2.68) and (2.69) are derived in the presented way is that not only the amplitude equations are uncoupled from the energy, but also that the similarity transformed Hamiltonian can conveniently be expanded by the Baker–Campbell–Hausdorff formula (BCH) and naturally truncated.

The BCH expansion reads

²an equivalent of the CCD method in the modern terminology

$$e^{-\hat{T}}\hat{H}_{N}e^{\hat{T}} = \hat{H}_{N} + \left[\hat{H}_{N},\hat{T}\right] \\ + \frac{1}{2!}\left[\left[\hat{H}_{N},\hat{T}\right],\hat{T}\right] \\ + \frac{1}{3!}\left[\left[\left[\hat{H}_{N},\hat{T}\right],\hat{T}\right],\hat{T}\right] \\ + \frac{1}{4!}\dots$$
(2.70)

By application of the Wick's theorem and evaluating the commutators it can be shown that this expansion is truncated after the fourth nested commutator and results in the following form

$$e^{-\hat{T}}\hat{H}_{N}e^{\hat{T}} = \left(\hat{H}_{N} + \hat{H}_{N}\hat{T} + \frac{1}{2!}\hat{H}_{N}\hat{T}^{2} + \frac{1}{3!}\hat{H}_{N}\hat{T}^{3} + \frac{1}{4!}\hat{H}_{N}\hat{T}^{4}\right)_{C}$$

$$= \left(\hat{H}_{N}e^{\hat{T}}\right)_{C}$$
(2.71)

where the symbol C stands for *connected* terms only. The connected term means that there must be at least one contraction line between the $\hat{H}_{\rm N}$ and each of the cluster operators to the right of it, otherwise the commutators would produce zero. And because there are at most four operators in the Hamiltonian (in its two-electron part), there can be at most four contraction lines leading from it and thus the expansion naturally truncates after the $\hat{H}_{\rm N}\hat{T}^4$ term.

The total cluster operator \hat{T} is usually truncated to a certain level of excitation. For instance the mostly used CCSD approximation involves only the single and double excitation cluster operators $(\hat{T} = \hat{T}_1 + \hat{T}_2)$, while the CCSDT method includes also the triples (\hat{T}_3) in addition.

Nevertheless, after substituting (2.71) into (2.67) and evaluating, the final coupled cluster energy expression becomes

$$\Delta E^{\rm CC} = \sum_{\substack{i \\ a}} f_{ia} t^a_i + \frac{1}{4} \sum_{\substack{ij \\ ab}} \langle ij \mid \mid ab \rangle t^{ab}_{ij} + \frac{1}{2} \sum_{\substack{ij \\ ab}} \langle ij \mid \mid ab \rangle t^a_i t^b_j \qquad (2.72)$$

which depends only on the \hat{T}_1 and \hat{T}_2 amplitudes regardless the level of truncation of the total cluster operator \hat{T} . The contribution of higher excitations to the coupled cluster energy is thus only indirect via the equations in which the \hat{T}_1 and \hat{T}_2 amplitudes are determined.

Analogically also the amplitude equations (2.68) and (2.69) for a given level of truncation of the cluster operator \hat{T} can be evaluated to final algebraic formulas by using the connected cluster expansion (2.71) and the Wick's theorem. However, the additional string of the bra operators creates so many contraction possibilities to consider that even the Wick's theorem application becomes too complicated and the diagrammatic technique has to be employed instead.

2.4.2 Size-Extensivity and Size-Consistency

At this point, the meaning of the two entities should be briefly explained since size-extensivity underlines the importance of the coupled cluster method while size-consistency gives rise to multireference theories.

These terms are usually best explained on an example of two noninteracting systems like a dimer of two infinitely distant water molecules A and B and comparing energy additivity for CC and CI methods.

Molecule A has the following CC wavefunction

$$|\mathbf{A}^{\mathrm{CC}}\rangle = e^{\hat{T}_{\mathbf{A}}} |\mathbf{A}^{\mathrm{HF}}\rangle \tag{2.73}$$

and the CI wavefunction

$$|\mathbf{A}^{\mathrm{CI}}\rangle = \left(1 + \hat{C}_{\mathrm{A}}\right)|\mathbf{A}^{\mathrm{HF}}\rangle \tag{2.74}$$

The wavefunctions of the B molecule are analogous. The CC wavefunction of the noninteracting dimer would be

$$|AB^{CC}\rangle = e^{\hat{T}_{A} + \hat{T}_{B}} |A^{HF}\rangle |B^{HF}\rangle$$

$$= e^{\hat{T}_{A}} |A^{HF}\rangle e^{\hat{T}_{B}} |B^{HF}\rangle$$

$$= |A^{CC}\rangle |B^{CC}\rangle \qquad (2.75)$$

which directly shows the additivity of CC energy for noninteracting systems

$$E_{\rm AB}^{\rm CC} = E_{\rm A}^{\rm CC} + E_{\rm B}^{\rm CC} \tag{2.76}$$

where the Hamiltonian is given as

$$\hat{H}_{\rm AB} = \hat{H}_{\rm A} + \hat{H}_{\rm B} \tag{2.77}$$

By contrast, the CI wavefunction

$$|AB^{CI}\rangle = \left(1 + \hat{C}_{A} + \hat{C}_{B}\right)|A^{HF}\rangle|B^{HF}\rangle \qquad (2.78)$$

is not factorizable (except for the Full-CI expansion, which is proper from its nature) and hence the truncated CI method cannot give proper energy of the supersystem. The CC method is thus said to be size-consistent (regardless the level of the \hat{T} operator truncation) while the truncated CI method is not.

It has been shown how energy of the CC and CI methods scale when a system with noninteracting components is studied and such proper scaling was called size-consistency. A method should, however, provide right scaling of
the energy regardless the system being investigated. That means not only to properly describe the noninteracting subsystems, but just any system should scale appropriately with the number of electrons. This is a mathematical feature of each particular method and is referred to as size-extensivity.

The truncated CI method is not size-extensive since the formulas from which the truncated CI coefficients can be obtained are coupled with energy (which is dependent on the number of electrons) but do not involve higher excitations that are present in the corresponding (size-extensive) Full-CI formulas and which cancel the incorrect scaling. Therefore an approximate Davidson correction for the size-extensivity [7,8] has to be applied.

The terms which are not compensated for the scaling error in the truncated CI method come from disconnected terms (diagrams). The CC method, however, contains only connected terms (diagrams) in the energy expression for any level of the cluster operator truncation and because the CC amplitude equations are uncoupled from energy, it is thus guaranteed that the coupled cluster method is size-extensive.

The term size-consistency has, however, a broader meaning. A sizeconsistent method must be size-extensive and must also properly describe separation (dissociation) into components. This depends on the particular molecular system. Studying a dissociation of a dimer of Fluorine molecules by the coupled cluster method would lead to size-consistent results while a dissociation of a single Fluorine molecule into atoms would not. For these cases a treatment of multireference methods is needed.

2.5 The Hilbert Space Multireference Coupled Cluster Methods

In multireference post-Hartree–Fock methods the several reference configurations enable investigation of systems in which the non-dynamic correlation plays role in qualitatively correct description of the potential energy surface. This involves cases like dissociation of molecules into fragments or nonequilibrium geometries (twisted ethylene). But generally other quasi-degenerate states like a singlet carbene molecule where single-reference methods completely fail to describe a realistic potential energy surface and often even fail to converge require the use of multireference methods. Another example of a necessary multireference treatment are systems with spin that cannot be described by a single Slater determinant like the ${}^{1}\Delta$ state of O₂ molecule. Moreover, the multireference description can in effect partially compensate the size-extensivity error of the truncated single-reference CI method due to the presence of higher excitations.

There is no unique multireference generalization of the exponential coupled cluster ansatz and thus many various approaches exist, which include the Hilbert space [31] and Fock space methods [32–35], reduced multireference CC scheme [36–38] or tailored CC [39,40] or even methods employing only single Fermi vacuum [41–49].

In contrast to the Fock space formalism which defines only one wave operator acting on several sectors of the Fock space, the Hilbert space multireference coupled cluster methods expand the wavefunction similarly like the single-reference method but from several reference configurations so that each reference determinant has its own cluster operator. Unlike the singlereference CC, the MRCC energy is then obtained as an eigenvalue of an effective Hamiltonian. In state-universal methods, all the eigenstates have physical meaning of the ground and the excited states. In state-specific methods, by contrast, only one specific state is physical and the other eigensolutions are artificial. The main advantage of the state-specific methods over the state-universal or even the Fock space formalism is their resistance to intruder states.

In the following, the theory of the Hilbert space multireference coupled cluster methods that concern this thesis is presented.

2.5.1 The State-Specific Multireference Brillouin– Wigner Coupled Cluster method (MR BWCC)

The several reference configurations $|\Phi_{\mu}\rangle$ form a model space \mathcal{P} with a projection operator

$$\hat{P} = \sum_{\mu=1}^{M} |\Phi_{\mu}\rangle \langle \Phi_{\mu}| \qquad (2.79)$$

and its orthogonal complement Q = 1 - P.

Within the model space a model wavefunction

$$|\Psi^{\mathcal{P}}_{\alpha}\rangle = \sum_{\mu=1}^{M} c^{\alpha}_{\mu} |\Phi_{\mu}\rangle \qquad (2.80)$$

is constructed where the coefficients c^{α}_{μ} are not known *a priori*. The exact wavefunction $|\Psi_{\alpha}\rangle$ is obtained from the model wavefunction by the state-specific wave operator $\hat{\Omega}_{\alpha}$

$$|\Psi_{\alpha}\rangle = \hat{\Omega}_{\alpha}|\Psi_{\alpha}^{\mathcal{P}}\rangle \tag{2.81}$$

and is required to fulfill the intermediate normalization

$$\left\langle \Psi_{\alpha} | \Psi_{\alpha}^{\mathcal{P}} \right\rangle = 1 \tag{2.82}$$

The wave operator is taken in the form of Jeziorski–Monkhorst ansatz [31]

$$\hat{\Omega}_{\alpha} = \sum_{\mu=1}^{M} e^{\hat{T}(\mu)} |\Phi_{\mu}\rangle \langle \Phi_{\mu}| \qquad (2.83)$$

The index μ in the cluster operators $T(\mu)$ denotes that the excitations are done with respect to μ -th reference configuration as a Fermi vacuum and that each reference configuration has its own set of independent amplitudes. The amplitudes that correspond to mutual excitations between the reference configurations are by definition set to zero.

The exact energy of the α th state E_{α} is obtained as the α th eigenvalue of the effective Hamiltonian \hat{H}^{eff} on the model space

$$\hat{H}^{\text{eff}}|\Psi^{\mathcal{P}}_{\alpha}\rangle = E_{\alpha}|\Psi^{\mathcal{P}}_{\alpha}\rangle \tag{2.84}$$

which is a non-Hermitian operator defined as

$$\hat{H}^{\text{eff}} = \hat{P}\hat{H}\hat{\Omega}_{\alpha}\hat{P} \tag{2.85}$$

As a state-specific method, only the α th state has physical meaning.

The coupled cluster amplitude equations are obtained by inserting the wave operator into the generalized Bloch equation [50]

$$\hat{H}\hat{\Omega}_{\alpha}\hat{P} - \eta E_{\alpha}\hat{\Omega}_{\alpha}\hat{P} - (1-\eta)\hat{\Omega}_{\alpha}\hat{H}_{\alpha}^{\text{eff}} = 0$$
(2.86)

In this equation, η is an arbitrary parameter between zero and one, with $\eta = 0$ corresponding to the Rayleigh-Schrödinger perturbation theory and $\eta = 1$ to the Brillouin–Wigner theory. The η -scaled term in (2.86), which is characteristic for the Brillouin-Wigner theory, corresponds to unlinked diagrams, leading to the size-inextensivity of the MR BWCC method. Therefore, a correction for size-extensivity is necessary. Two corrections were suggested which can be obtained from the above continuous transition from which, however, the $(1 - \eta)$ -scaled coupling terms are omitted. This leads to the following CC amplitude equations

$$\left\langle \Phi_{\vartheta} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle_{C} + \eta \left\langle \Phi_{\vartheta} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle_{DC} - \eta E_{\alpha} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle = 0$$
 (2.87)

in which the C and DC labels denote the connected and disconnected diagrams, respectively. In the a posteriori correction [51], after converging the equations (2.87) with $\eta = 1$, one additional iteration of cluster equations is performed while setting $\eta = 0$. In the iterative correction, however, these terms are gradually scaled to zero by successive iterations of the cluster equations and the converged result corresponds to $\eta = 0$.

Despite the unpleasant consequence of the Brillouin–Wigner formalism that the method is not size-extensive *a priori*, two advantages also arise. Since the Brillouin–Wigner resolvent has the true energy in its denominator, intruder state problems, in which the denominator goes to zero and makes numerical difficulties, are avoided. This is however not true when the iterative correction is being applied, which often faces to convergence problems. The other advantage is that the CC amplitude equations (2.87) of different reference configurations are coupled only via the energy so the method scales linearly with the size of the reference space.

The method has been used to study many chemically interesting systems at CCSD level [52–60]. Later it was also provided with connected triples contribution to involve more dynamical correlation at an approximate iterative [61], non-iterative [18] and finally full iterative MR BWCCSDT level [62].

2.5.2 The State-Universal Multireference Coupled Cluster method (SU MRCC)

The Hilbert space MRCC theory is similar to the method above, except that the wave operator is state-universal and the transition of the Bloch equation [50] to the Kucharski-Bartlet formulation of the state-universal MRCC [15] leads to the amplitude equations

$$\left\langle \Phi_{\vartheta} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle_{C} + \eta \left\langle \Phi_{\vartheta} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle_{DC} - \eta E_{\alpha} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle - (1 - \eta) \sum_{\nu \in \mathcal{P}, \nu \neq \mu} \left\{ \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\nu)} \right| \Phi_{\nu} \right\rangle \hat{H}_{\nu\mu}^{\text{eff}} \right\}_{C} = 0$$
(2.88)

which are coupled between different reference configurations.

2.5.3 The State-Specific Multireference Mukherjee Coupled Cluster method (MR MkCC)

In contrast to the Brillouin–Wigner MRCC theory, the derivation of MR MkCC method starts by the following resolution of identity

$$1 = e^{\hat{T}(\mu)} e^{-\hat{T}(\mu)} = e^{\hat{T}(\mu)} (\mathcal{P} + \mathcal{Q}) e^{-\hat{T}(\mu)} = \sum_{\nu}^{\mathcal{P}} e^{\hat{T}(\mu)} |\nu\rangle \langle \nu| e^{-\hat{T}(\mu)} + e^{\hat{T}(\mu)} \mathcal{Q} e^{-\hat{T}(\mu)}$$
(2.89)

which is inserted from the left to the Schrödinger equation

$$\hat{H}\hat{\Omega}_{\alpha}|\Psi^{\mathcal{P}}_{\alpha}\rangle = \mathcal{E}_{\alpha}\hat{\Omega}_{\alpha}|\Psi^{\mathcal{P}}_{\alpha}\rangle \tag{2.90}$$

Subsequent interchange of the summation indices μ and ν in the first term yields to

$$\sum_{\mu}^{\mathcal{P}} \left\{ \sum_{\nu}^{\mathcal{P}} e^{\hat{T}(\nu)} |\Phi_{\mu}\rangle H_{\mu\nu}^{\text{eff}} c_{\nu}^{\alpha} + e^{\hat{T}(\mu)} \mathcal{Q} e^{-\hat{T}(\mu)} \hat{H} e^{\hat{T}(\mu)} |\Phi_{\mu}\rangle c_{\mu}^{\alpha} - \mathcal{E}_{\alpha} e^{\hat{T}(\mu)} |\Phi_{\mu}\rangle c_{\mu}^{\alpha} \right\} = 0$$

$$(2.91)$$

The sufficiency conditions are then applied to resolve the redundancy of the Jeziorski–Monkhorst ansatz in the state-specific context, requiring that each μ th contribution of the above summation is equal to zero. The equations for the cluster amplitudes are finally obtained by multiplying the resulting equation from the left by $e^{-\hat{T}(\mu)}$ and projecting to excited configurations, which gives

$$\left\langle \Phi_{\vartheta} \left| e^{-\hat{T}(\mu)} \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle c_{\mu}^{\alpha} + \sum_{\nu \neq \mu}^{\mathcal{P}} H_{\mu\nu}^{\text{eff}} c_{\nu}^{\alpha} \left\langle \Phi_{\vartheta} \left| e^{-\hat{T}(\mu)} e^{\hat{T}(\nu)} \right| \Phi_{\mu} \right\rangle = 0$$
(2.92)

Uncoupled Approximation to MR MkCC

Mukherjee *et al.* [63,64] suggested to approximate the cluster operator $\hat{T}(\nu)$ in the coupling term $\left\langle \Phi_{\vartheta} \left| e^{-\hat{T}(\mu)} e^{\hat{T}(\nu)} \right| \Phi_{\mu} \right\rangle$ of (2.92) by $\hat{T}'_{\nu}(\mu)$ which is actually a subset of $\hat{T}(\mu)$ containing only such excitation operators that give nonzero when acting on $|\Phi_{\nu}\rangle$. Thus the amplitudes are given by

$$t_{\nu \, i...}^{\prime \, a...} = \begin{cases} t_{i...}^{a...} & \text{if } i \cdots \in occ(\nu); a \cdots \in virt(\nu) \\ 0 & \text{otherwise} \end{cases}$$
(2.93)

By further defining a complement cluster operator

$$\bar{T}_{\nu}(\mu) = \hat{T}(\mu) - \hat{T}_{\nu}'(\mu)$$
(2.94)

the coupling term becomes $\left\langle \Phi_{\vartheta} \left| e^{-\bar{T}_{\nu}(\mu)} \right| \Phi_{\mu} \right\rangle$.

In the context of the triples, the theory has been recently further investigated by Demel *et al.* [65].

2.6 Analytical Gradient of the Hartree–Fock Energy

Although not directly necessary for derivation of the coupled cluster gradient formulas, the example of the Hartree–Fock energy derivative helps to introduce the basic principles and terms like skeletons and U matrices which are commonly used in the analytical gradient theory. Before differentiating the Hartree–Fock electronic energy formula

$$E^{\text{RHF}} = 2\sum_{i}^{N/2} h_{ii} + \sum_{ij}^{N/2} \left[2\left(ii|jj\right) - \left(ij|ji\right) \right]$$
(2.95)

let us first show how derivatives of the LCAO coefficients are usually expressed. Starting from an expansion of a molecular orbital $|i\rangle$ in the basis of atomic orbitals by the LCAO coefficients

$$|i\rangle = \sum_{\mu}^{AO} |\mu\rangle C_{\mu i}$$
(2.96)

and differentiating with respect to a general perturbation χ

$$\frac{\partial |i\rangle}{\partial \chi} = \sum_{\mu}^{AO} |\mu\rangle \frac{\partial C_{\mu i}}{\partial \chi}$$
(2.97)

the RHS of (2.97) is expanded to the MO basis in terms of U matrices

$$\frac{\partial |i\rangle}{\partial \chi} = \sum_{m}^{MO} |m\rangle U_{mi}^{\chi}$$
$$= \sum_{m}^{MO} \sum_{\mu}^{AO} |\mu\rangle C_{\mu m} U_{mi}^{\chi}$$
(2.98)

By comparing the RHS of (2.97) and the right-hand side of (2.98), the expression for the LCAO coefficient derivatives in terms of the U matrices is finally obtained and reads

$$\frac{\partial C_{\mu i}}{\partial \chi} = \sum_{m}^{MO} C_{\mu m} U_{m i}^{\chi}$$
(2.99)

The U matrices can be calculated from the coupled perturbed Hartree– Fock equations (CPHF), which will be described at the end of this section. But now that the U matrices have been introduced, the differentiation of the Hartree–Fock energy expression in MO basis (2.95) can be performed and formally gives

$$\frac{\partial E^{\text{RHF}}}{\partial \chi} = 2\sum_{i}^{occ} \frac{\partial h_{ii}}{\partial \chi} + \sum_{ij}^{occ} \left[2 \frac{\partial (ii|jj)}{\partial \chi} - \frac{\partial (ij|ij)}{\partial \chi} \right]$$
(2.100)

where

$$\frac{\partial h_{ii}}{\partial \chi} = h_{ii}^{\chi} + 2 \sum_{m}^{\text{MO}} U_{mi}^{\chi} h_{mi}$$
(2.101)

$$\frac{\partial (ii|jj)}{\partial \chi} = (ii|jj)^{\chi} + 2\sum_{m}^{MO} \left[U_{mi}^{\chi} (mi|jj) + U_{mj}^{\chi} (ii|mj) \right]$$
(2.102)

$$\frac{\partial (ij|ij)}{\partial \chi} = (ij|ij)^{\chi} + 2\sum_{m}^{MO} \left[U_{mi}^{\chi}(mj|ij) + U_{mj}^{\chi}(im|ij) \right]$$
(2.103)

where the terms with AO integral derivatives

$$h_{ij}^{\chi} = \sum_{\mu\nu}^{AO} C_{\mu i} \frac{\partial h_{\mu\nu}}{\partial \chi} C_{\nu j}$$
(2.104)

and

$$(ij|kl)^{\chi} = \sum_{\substack{\mu\nu\\\varrho\sigma}}^{AO} \frac{\partial \ (\mu\nu|\varrho\sigma)}{\partial \chi} C_{\mu i} C_{\nu j} C_{\varrho k} C_{\sigma l}$$
(2.105)

are commonly referred to as the *skeletons*.

After some algebra, the (2.100) however turns out to be independent of the U matrices and reads

$$\frac{\partial E^{\text{RHF}}}{\partial \chi} = 2\sum_{i}^{occ} h_{ii}^{\chi} + \sum_{ij}^{occ} \left[2\left(ii|jj\right)^{\chi} - (ij|ij)^{\chi} \right] - 2\sum_{i}^{occ} S_{ii}^{\chi} \varepsilon_i \qquad (2.106)$$

where

$$S_{ij}^{\chi} = \sum_{\mu\nu}^{AO} C_{\mu i} \frac{\partial S_{\mu\nu}}{\partial \chi} C_{\nu j}$$
(2.107)

is a skeleton term of the overlap matrix.

The gradient formula (2.106) is then expressed in the AO basis by substituting the skeletons and finally yields

$$\frac{\partial E^{\text{RHF}}}{\partial \chi} = 2 \sum_{\mu\nu}^{\text{AO}} D_{\mu\nu} h_{\mu\nu}^{\chi} + \sum_{\substack{\mu\nu\\\varrho\sigma}}^{\text{AO}} \left\{ 2D_{\mu\nu}D_{\varrho\sigma} - D_{\mu\varrho}D_{\nu\sigma} \right\} (\mu\nu|\varrho\sigma)^{\chi} -2 \sum_{\mu\nu}^{\text{AO}} W_{\mu\nu} S_{\mu\nu}^{\chi}$$
(2.108)

where

$$D_{\mu\nu} = \sum_{i}^{occ} C_{\mu i} C_{\nu i}$$
 (2.109)

is the Hartree–Fock one-electron density matrix and

$$W_{\mu\nu} = \sum_{i}^{occ} C_{\mu i} C_{\nu i} \varepsilon_i \qquad (2.110)$$

is the energy weighted density matrix while using the common notation for the AO integral derivatives

$$S^{\chi}_{\mu\nu} = \frac{\partial S_{\mu\nu}}{\partial \chi} \tag{2.111}$$

$$h_{\mu\nu}^{\chi} = \frac{\partial h_{\mu\nu}}{\partial \chi} \tag{2.112}$$

$$(\mu\nu|\rho\sigma)^{\chi} = \frac{\partial \ (\mu\nu|\rho\sigma)}{\partial \chi} \tag{2.113}$$

The final U-matrix-free gradient formula (2.108) is in accordance to the Wigner's 2n + 1 rule which says that from a wavefunction exact to the nth order of a perturbation expansion, an energy of the (2n + 1)th order can be obtained. And thus from the Hartree–Fock wavefunction (zeroth order) the first energy derivative can be calculated without the need to differentiate the LCAO coefficients.

2.6.1 Coupled Perturbed Hartree–Fock Equations (CPHF)

It has been shown that U matrices are not necessary for evaluation of the Hartree–Fock energy gradient. Nevertheless, it is not the case of higher derivatives and the U matrices also play an important role in the analytical gradient theory of the post-HF methods. To obtain them, a set of CPHF equations [66] has to be solved for each gradient coordinate χ .

But before introducing the equations, an important relation for the U matrices should be derived from the orthonormality condition for the molecular orbitals

$$\langle j|i\rangle = \delta_{ji}$$

$$\sum_{\mu\nu}^{AO} \langle \mu| C_{\mu j} C_{\nu i} |\nu\rangle = \delta_{ji}$$

$$\sum_{\mu\nu}^{AO} C_{\mu j} S_{\mu\nu} C_{\nu i} = \delta_{ji} \qquad (2.114)$$

Differentiating (2.114) with respect to χ gives

$$\sum_{\mu\nu}^{AO} \frac{\partial C_{\mu i}}{\partial \chi} S_{\mu\nu} C_{\nu j} + C_{\mu i} \underbrace{\frac{\partial S_{\mu\nu}}{\partial \chi}}_{S_{\mu\nu}^{\chi}} C_{\nu j} + C_{\mu i} S_{\mu\nu} \frac{\partial C_{\nu j}}{\partial \chi} = 0$$
(2.115)

By substituting the U matrices, the derivative of the condition finally becomes

$$U_{ji}^{\chi} + S_{ij}^{\chi} + U_{ij}^{\chi} = 0 (2.116)$$

where

$$S_{ij}^{\chi} = \sum_{\mu\nu}^{AO} C_{\mu i} S_{\mu\nu}^{\chi} C_{\nu j} = S_{ji}^{\chi}$$
(2.117)

From (2.116), for the diagonal elements of the U matrices it then follows

$$U_{ii}^{\chi} = -\frac{1}{2}S_{ii}^{\chi} \tag{2.118}$$

Without showing the derivation (by differentiating the Fock matrix), the CPHF equations read

$$\left(\varepsilon_{j} - \varepsilon_{i}\right) U_{ij}^{\chi} - \sum_{k}^{virt} \sum_{l}^{occ} U_{kl}^{\chi} A_{ij,kl} = B_{ij}^{\chi}$$
(2.119)

where

$$A_{ij,kl} = 4 (ij|kl) - (ik|jl) - (il|jk)$$
(2.120)

$$B_{ij}^{\chi} = F_{ij}^{\chi} - \varepsilon_j S_{ij}^{\chi} - \sum_{kl}^{occ} S_{kl}^{\chi} \left\{ 2 \left(ij|kl \right) - \left(ik|jl \right) \right\}$$
(2.121)

$$F_{ij}^{\chi} = h_{ij}^{\chi} + \sum_{k}^{occ} \left\{ 2 \, (ij|kk)^{\chi} - (ik|jk)^{\chi} \right\}$$
(2.122)

The (2.119) is actually a set of linear equations for $\#occ \times \#virt$ variables from which the *occ-virt* blocks of the U matrices are obtained. The transposed U matrix elements are then calculated from (2.116) and the diagonals from (2.118). The remaining *occ-occ* and *virt-virt* blocks are finally calculated explicitly from (2.119) by using the *occ-virt* blocks

$$U_{ij}^{\chi} = \frac{1}{\varepsilon_j - \varepsilon_i} \left\{ B_{ij}^{\chi} + \sum_{k}^{virt} \sum_{l}^{occ} U_{kl}^{\chi} A_{ij,kl} \right\}$$
(2.123)

2.7 Analytical Gradient of the Coupled Cluster Energy

The analytical energy derivatives of the single-reference coupled cluster methods were developed in several groups [67–83]. It was shown that despite being a non-variational method, the CC amplitude derivatives need not be determined and only a single, perturbation independent equation must be solved. Instead of a direct differentiation of the coupled cluster energy formula, which would inevitably lead to the derivatives of the cluster amplitudes for each gradient component, it is advantageous to construct and differentiate a Lagrangian which directly results in solving only a set of linear λ equations, which are independent of the perturbation parameter and the resulting λ amplitudes are thus common for all gradient coordinates [10,84,85].

The Lagrangian consists of the CC energy expression and the CC amplitude equations as the constraints with the unspecified coefficients λ_{Φ} called the λ amplitudes

$$\mathcal{L}(t_{\Phi}, \lambda_{\Phi'}) = \left\langle \Phi_0 \left| \left[\hat{H}_{\mathrm{N}} e^{\hat{T}} \right]_C \right| \Phi_0 \right\rangle + \sum_{\Phi'} \lambda_{\Phi'} \left\langle \Phi' \left| \left[\hat{H}_{\mathrm{N}} e^{\hat{T}} \right]_C \right| \Phi_0 \right\rangle \quad (2.124)$$

Here the t_{Φ} means a CC amplitude in the sense of

$$\hat{T} = \sum_{\Phi} t_{\Phi} \hat{\tau}_{\Phi} \tag{2.125}$$

$$\hat{\tau}_{\Phi}|\Phi_0\rangle = \pm|\Phi\rangle \tag{2.126}$$

where the $\hat{\tau}_{\Phi}$ is its corresponding excitation operator (the sign corresponds to the parity of permutation necessary to bring the string of the excited determinant to a canonical order).

Minimizing the Lagrangian with respect to λ and CC amplitudes leads to solving the CC amplitude equations

$$\frac{\partial \mathcal{L}}{\partial \lambda_{\Phi'}} = 0 \tag{2.127}$$

and the λ equations

$$\frac{\partial \mathcal{L}}{\partial t_{\Phi}} = 0 \tag{2.128}$$

The CC energy gradient formula can then be expressed as a partial derivative of the Lagrangian and has thus a simple form

$$\frac{d\Delta E^{\rm CC}}{d\chi} = \frac{d\mathcal{L}}{d\chi}
= \frac{\partial\mathcal{L}}{\partial\chi} + \sum_{\Phi} \underbrace{\frac{\partial\mathcal{L}}{\partial t_{\Phi}}}_{0} \frac{\partial t_{\Phi}}{\partial\chi} + \sum_{\Phi'} \underbrace{\frac{\partial\mathcal{L}}{\partial\lambda_{\Phi'}}}_{0} \frac{\partial\lambda_{\Phi'}}{\partial\chi}
= \frac{\partial\mathcal{L}}{\partial\chi}
= \left\langle \Phi_{0} \left| \left[\hat{H}_{\rm N}^{\chi} e^{\hat{T}} \right]_{C} \right| \Phi_{0} \right\rangle
+ \sum_{\Phi'} \lambda_{\Phi'} \left\langle \Phi' \left| \left[\hat{H}_{\rm N}^{\chi} e^{\hat{T}} \right]_{C} \right| \Phi_{0} \right\rangle$$
(2.129)

where

$$\hat{H}_{\rm N}^{\chi} = \frac{\partial \, \hat{H}_{\rm N}}{\partial \, \chi} \tag{2.130}$$

which actually corresponds to the generalized Hellman–Feynman theorem, since the perturbation parameter χ appears explicitly in the Hamiltonian only.

For algebraic manipulations it is profitable to define the λ_{Φ} coefficients in terms of a deexcitation operator $\hat{\Lambda}$

$$\lambda_{\Phi} = \left\langle \Phi_0 \left| \hat{\Lambda} \right| \Phi \right\rangle \tag{2.131}$$

where

$$\hat{\Lambda} = \sum_{\substack{a \\ i}} \lambda_a^i \, a_i^+ a_a + \frac{1}{4} \sum_{\substack{ab \\ ij}} \lambda_{ab}^{ij} \, a_i^+ a_j^+ \, a_b \, a_a + \dots$$
(2.132)

The final gradient formula thus reads

$$\frac{d\Delta E^{\rm CC}}{d\chi} = \left\langle \Phi_0 \left| \left[\hat{H}_{\rm N}^{\chi} e^{\hat{T}} \right]_C \right| \Phi_0 \right\rangle + \left\langle \Phi_0 \left| \hat{\Lambda} \left[\hat{H}_{\rm N}^{\chi} e^{\hat{T}} \right]_C \right| \Phi_0 \right\rangle$$
(2.133)

2.7.1 The λ Equation

Starting from its definition equation (2.128) the differentiation gives

$$\frac{1}{\left\langle \Phi_{0} \left| -\hat{\tau}_{\Phi}e^{-\hat{T}}\hat{H}_{N}e^{\hat{T}} \right| \Phi_{0} \right\rangle} + \left\langle \Phi_{0} \left| e^{-\hat{T}}\hat{H}_{N}e^{\hat{T}}\hat{\tau}_{\Phi} \right| \Phi_{0} \right\rangle + \sum_{\Phi'} \lambda_{\Phi'} \left\langle \Phi' \left| -\hat{\tau}_{\Phi}e^{-\hat{T}}\hat{H}_{N}e^{\hat{T}} \right| \Phi_{0} \right\rangle + \sum_{\Phi'} \lambda_{\Phi'} \left\langle \Phi' \left| e^{-\hat{T}}\hat{H}_{N}e^{\hat{T}}\hat{\tau}_{\Phi} \right| \Phi_{0} \right\rangle = 0 \qquad (2.134)$$

After some manipulations and using the $\hat{\Lambda}$ operator definition (2.131) the final λ equation becomes

$$\left\langle \Phi_{0} \left| \left[\hat{\Lambda} \left[\hat{H}_{N} e^{\hat{T}} \right]_{C} \right]_{C} \right| \Phi \right\rangle + \sum_{\Phi'} \left\langle \Phi_{0} \left| \left[\hat{H}_{N} e^{\hat{T}} \right]_{C} \right| \Phi' \right\rangle \left\langle \Phi' \left| \hat{\Lambda} \right| \Phi \right\rangle + \left\langle \Phi_{0} \left| \left[\hat{H}_{N} e^{\hat{T}} \right]_{C} \right| \Phi \right\rangle = 0$$

$$(2.135)$$

2.8 Analytical Gradient of the Hilbert Space MRCC Methods

The analytical gradient of energy of the Hilbert space MRCC methods [2] is derived analogically to the single-reference CC gradient by using the Lagrangian technique.

For simplicity, let us define a general MRCC amplitude equation as

$$\mathcal{Q}_{\vartheta}(\mu) = 0 \tag{2.136}$$

where μ denotes the particular Fermi vacuum from the model space and ϑ represents the bra- configuration corresponding to the $\mathcal{Q}(\mu)$ subspace ³. The energy can conveniently be expressed as the eigenvalue of the effective Hamiltonian

$$E_{\alpha} = \sum_{\mu\nu}^{\mathcal{P}} \tilde{c}_{\nu}^{\alpha} H_{\nu\mu}^{\text{eff}} c_{\mu}^{\alpha}$$
(2.137)

where the \tilde{c}^{α}_{ν} and c^{α}_{μ} are components of the left and right eigenvectors of the \hat{H}^{eff} .

The Lagrangian is then given by

³That is the subspace of $\mathcal Q$ which corresponds to excitations from the μ th Fermi vacuum

$$\mathcal{L} = \sum_{\mu,\nu}^{\mathcal{P}} \tilde{c}_{\nu}^{\alpha} c_{\mu}^{\alpha} H_{\nu\mu}^{\text{eff}} + \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \mathcal{Q}_{\vartheta}(\mu)$$
(2.138)

The $\lambda_{\vartheta}(\mu)$ multipliers are expressed in terms of the $\hat{\Lambda}(\mu)$ deexcitation operator specific for each reference configuration

$$\lambda_{\vartheta}(\mu) \equiv \left\langle \Phi_{\mu} \left| \hat{\Lambda}(\mu) \right| \Phi_{\vartheta} \right\rangle \tag{2.139}$$

for $\vartheta \in \mathcal{Q}(\mu)$, while for $\nu \in \mathcal{P}$ the internal λ amplitudes are defined to be zero

$$\left\langle \Phi_{\mu} \left| \hat{\Lambda}(\mu) \right| \Phi_{\nu} \right\rangle = 0$$
 (2.140)

similarly like the internal MRCC amplitudes.

The total derivative of the Lagrangian becomes

$$\frac{d\mathcal{L}}{d\chi} = \frac{\partial \mathcal{L}}{\partial \chi} + \sum_{\nu}^{\mathcal{P}} \frac{\partial \mathcal{L}}{\partial \tilde{c}_{\nu}^{\alpha}} \frac{\partial \tilde{c}_{\nu}^{\alpha}}{\partial \chi} + \sum_{\mu}^{\mathcal{P}} \frac{\partial \mathcal{L}}{\partial c_{\mu}^{\alpha}} \frac{\partial c_{\mu}^{\alpha}}{\partial \chi} + \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \frac{\partial \mathcal{L}}{\partial \lambda_{\vartheta}(\mu)} \frac{\partial \lambda_{\vartheta}(\mu)}{\partial \chi} + \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \frac{\partial \mathcal{L}}{\partial t_{\vartheta}(\mu)} \frac{\partial t_{\vartheta}(\mu)}{\partial \chi} \qquad (2.141)$$

The fourth term in the above equation (2.141) vanishes for the converged MRCC amplitudes since

$$\frac{\partial \mathcal{L}}{\partial \lambda_{\vartheta}(\mu)} = \mathcal{Q}_{\vartheta}(\mu) = 0 \qquad (2.142)$$

and the fifth term vanishes as well, provided the λ equations are solved

$$\frac{\partial \mathcal{L}}{\partial t_{\vartheta}(\mu)} = 0 \tag{2.143}$$

If the eigenvectors of the \hat{H}^{eff} are biorthonormal

$$\sum_{\mu}^{\mathcal{P}} \tilde{c}^{\alpha}_{\mu} c^{\beta}_{\mu} = \delta_{\alpha\beta} \tag{2.144}$$

the second and third terms of (2.141) give zero and the derivative of the MRCC energy becomes

$$\frac{d E_{\alpha}}{d \chi} = \frac{d \mathcal{L}}{d \chi} = \frac{\partial \mathcal{L}}{\partial \chi}$$

$$= \frac{\partial E_{\alpha}}{\partial \chi} + \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \frac{\partial \mathcal{Q}_{\vartheta}(\mu)}{\partial \chi}$$

$$= \sum_{\mu,\nu}^{\mathcal{P}} \tilde{c}_{\nu}^{\alpha} c_{\mu}^{\alpha} \frac{\partial H_{\nu\mu}^{\text{eff}}}{\partial \chi} + \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \frac{\partial \mathcal{Q}_{\vartheta}(\mu)}{\partial \chi} \qquad (2.145)$$

2.8.1 Analytical Gradient of the MR BWCC with the Iterative Correction of Size-Extensivity

Since the MRCC amplitude equation has the simple uncoupled form after the iterative correction converges to zero

$$\mathcal{Q}_{\vartheta}(\mu) \equiv \left\langle \Phi_{\vartheta} \left| e^{-\hat{T}(\mu)} \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle = 0$$
 (2.146)

the gradient derivation of this MRCC variant is relatively easy.

It is convenient to augment the lambda operator also for internal lambda amplitudes

$$\left\langle \Phi_{\mu} \left| \tilde{\Lambda}(\mu) \right| \Phi_{\nu} \right\rangle = \tilde{c}_{\nu}^{\alpha} c_{\mu}^{\alpha} \quad [\nu \in \mathcal{P}],$$

$$\left\langle \Phi_{\mu} \left| \tilde{\Lambda}(\mu) \right| \Phi_{\vartheta}(\mu) \right\rangle = \lambda_{\vartheta}(\mu) \quad [\vartheta \in \mathcal{Q}(\mu)]$$

$$(2.147)$$

By taking into account that for a complete model space,

$$H_{\nu\mu}^{\text{eff}} = \left\langle \Phi_{\nu} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
$$= \left\langle \Phi_{\nu} \left| e^{-\hat{T}(\mu)} \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
(2.148)

the λ equation thus reads

$$\frac{\partial}{\partial t_{\zeta}(\varrho)} \sum_{\mu}^{\mathcal{P}} \left\langle \Phi_{\mu} \left| \tilde{\Lambda}(\mu) e^{-\hat{T}(\mu)} \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle = 0$$
 (2.149)

Expressing the cluster operator $\hat{T}(\mu)$ in analogy to (2.125), the λ equation becomes

$$\left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) e^{-\hat{T}(\varrho)} \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle - \left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \tau_{\zeta}(\varrho) e^{-\hat{T}(\varrho)} \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\varrho} \right\rangle = 0 \quad (2.150)$$

The second term can be simplified by inserting the resolution of identity $1 = P + Q(\rho)$ after the $\tau_{\zeta}(\rho)$ operator and taking into account that the Q-space projection vanishes for converged t amplitudes, which yields

$$\left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) e^{-\hat{T}(\varrho)} \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle - \sum_{\nu}^{\mathcal{P}} \left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \tau_{\zeta}(\varrho) \right| \Phi_{\nu} \right\rangle H_{\nu\varrho}^{\text{eff}} = 0 \quad (2.151)$$

The $\nu = \rho$ term can be moved from the sum to the first term, which yields the final form of the lambda equation

$$\left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \left[e^{-\hat{T}(\varrho)} \hat{H} e^{\hat{T}(\varrho)} - H_{\varrho\varrho}^{\text{eff}} \right] \right| \Phi_{\zeta} \right\rangle - \sum_{\nu \in \mathcal{P}, \nu \neq \varrho} \left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \tau_{\zeta}(\varrho) \right| \Phi_{\nu} \right\rangle H_{\nu\varrho}^{\text{eff}} = 0 \qquad (2.152)$$

Since Φ_{ν} is an internal excitation from Φ_{ϱ} and $\hat{\tau}_{\zeta}(\varrho)$ is an excitation operator with respect to Φ_{ϱ} , $\hat{\tau}_{\zeta}(\varrho)|\Phi_{\nu}\rangle$ is a semi-internal excitation from Φ_{ρ} and the second term can be nonzero only if its excitation rank does not exceed the deexcitation level included in $\tilde{\Lambda}$. Note that the lambda equations for amplitudes of different reference configurations are uncoupled.

Because of the convenient definition of the Λ operator in (2.147), the final gradient formula (2.145) can be written in a compact form

$$\frac{d E_{\alpha}}{d \chi} = \sum_{\mu,\nu}^{\mathcal{P}} \tilde{c}_{\nu}^{\alpha} c_{\mu}^{\alpha} \left\langle \Phi_{\nu} \left| e^{-\hat{T}(\mu)} \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle
+ \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \left\langle \Phi_{\vartheta} \left| e^{-\hat{T}(\mu)} \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle
= \sum_{\mu}^{\mathcal{P}} \left\langle \Phi_{\mu} \left| \tilde{\Lambda}(\mu) e^{-\hat{T}(\mu)} \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle = 0 \qquad (2.153)$$

The gradient thus consists of independent contributions from each reference configuration.

2.8.2 Analytical Gradient of the MR BWCC without Correction of Size-Extensivity

For the uncorrected MR BWCC the situation is a bit more complicated, since the amplitude equations are now coupled via the total energy E_{α}

$$\mathcal{Q}_{\vartheta}(\mu) \equiv \left\langle \Phi_{\vartheta} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle - E_{\alpha} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle = 0$$
(2.154)

However, it turns out that this coupling leads to separable expressions and terms mixing t or λ amplitudes of different reference configurations never arise.

Similarly like in the iteratively corrected variant, the term with the $H_{\nu\mu}^{\text{eff}}$ derivative can be moved to the modified $\tilde{\Lambda}$ operator term which leads to the following λ equation

$$\left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle - \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \times \left[\frac{\partial E_{\alpha}}{\partial t_{\zeta}(\varrho)} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle + E_{\alpha} \frac{\partial}{\partial t_{\zeta}(\varrho)} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle \right] = 0 (2.155)$$

By differentiating the energy from (2.137) using the biorthonormal eigenvectors (2.144) and considering that in a complete model space,

$$\left\langle \Phi_{\nu} \left| e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle = \delta_{\nu\mu}$$
 (2.156)

the λ equation yields

$$\left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle$$
$$-\omega \sum_{\sigma}^{\mathcal{P}} \tilde{c}_{\sigma}^{\alpha} c_{\varrho}^{\alpha} \left\langle \Phi_{\sigma} \left| \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle$$
$$-E_{\alpha} \left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle = 0 \qquad (2.157)$$

where

$$\omega = \sum_{\mu}^{\mathcal{P}} \left\langle \Phi_{\mu} \left| \tilde{\Lambda}(\mu) e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
(2.158)

By further redefining the $\tilde{\Lambda}$ operator

$$\left\langle \Phi_{\mu} \left| \bar{\Lambda}(\mu) \right| \Phi_{\nu} \right\rangle = (1 - \omega) \tilde{c}_{\nu}^{\alpha} c_{\mu}^{\alpha} \quad [\nu \in \mathcal{P}], \\ \left\langle \Phi_{\mu} \left| \bar{\Lambda}(\mu) \right| \Phi_{\vartheta}(\mu) \right\rangle = \lambda_{\vartheta}(\mu) \quad [\vartheta \in \mathcal{Q}(\mu)]$$
(2.159)

the λ equation finally becomes

$$\left\langle \Phi_{\varrho} \left| \bar{\Lambda}(\varrho) \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle - E_{\alpha} \left\langle \Phi_{\varrho} \left| \bar{\Lambda}(\varrho) e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle = 0$$
(2.160)

which couples different reference configurations only via the energy.

The energy gradient reads

$$\frac{d E_{\alpha}}{d \chi} = \sum_{\mu\nu}^{\mathcal{P}} \tilde{c}_{\nu}^{\alpha} c_{\mu}^{\alpha} \left\langle \Phi_{\nu} \left| \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle
+ \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \left\langle \Phi_{\vartheta} \left| \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle
- \frac{\partial E_{\alpha}}{\partial \chi} \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
(2.161)

The sum in the last term is actually equal to the definition of the factor ω and since the partial energy derivative gives

$$\frac{\partial E_{\alpha}}{\partial \chi} = \sum_{\mu\nu}^{\mathcal{P}} \tilde{c}^{\alpha}_{\nu} c^{\alpha}_{\mu} \left\langle \Phi_{\nu} \left| \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
(2.162)

the final gradient formula yields

$$\frac{d E_{\alpha}}{d \chi} = \sum_{\mu}^{\mathcal{P}} \left\langle \Phi_{\mu} \left| \bar{\Lambda}(\mu) \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
(2.163)

Again, the gradient sums contributions of different reference configurations with coupling given only by the \hat{H}^{eff} eigenvector coefficients and the factor ω .

2.8.3 Analytical Gradient of the SU MRCC

In contrast to the MR BWCC method, the amplitude equations of the Kucharski–Bartlett formulation of the state-universal MRCC are coupled between different reference configurations

$$\mathcal{Q}_{\vartheta}(\mu) \equiv \left\langle \Phi_{\vartheta} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle - \sum_{\nu}^{\mathcal{P}} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\nu)} \right| \Phi_{\nu} \right\rangle H_{\nu\mu}^{\text{eff}} = 0 \qquad (2.164)$$

The λ equation thus reads

$$\sum_{\mu\nu}^{\mathcal{P}} \tilde{c}_{\nu}^{\alpha} c_{\mu}^{\alpha} \frac{\partial H_{\nu\mu}^{\text{eff}}}{\partial t_{\zeta}(\varrho)} + \\ + \sum_{\mu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\vartheta}(\mu) \left[\frac{\partial}{\partial t_{\zeta}(\varrho)} \left\langle \Phi_{\vartheta} \left| \hat{H} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle \right. \\ \left. - \sum_{\nu}^{\mathcal{P}} H_{\nu\mu}^{\text{eff}} \frac{\partial}{\partial t_{\zeta}(\varrho)} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\nu)} \right| \Phi_{\nu} \right\rangle \\ \left. - \sum_{\nu}^{\mathcal{P}} \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\nu)} \right| \Phi_{\nu} \right\rangle \frac{\partial H_{\nu\mu}^{\text{eff}}}{\partial t_{\zeta}(\varrho)} \right] = 0 \qquad (2.165)$$

By using the $\tilde{\Lambda}$ operator from (2.147) and treating the derivatives of the last two terms like in the MR BWCC, the λ equation yields

$$\left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle - \sum_{\mu\nu}^{\mathcal{P}} \sum_{\vartheta}^{\mathcal{Q}(\mu)} \lambda_{\zeta}(\mu) \left[\left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\nu)} \right| \Phi_{\zeta} \right\rangle \delta_{\varrho\nu} H_{\nu\mu}^{\text{eff}} + \left\langle \Phi_{\vartheta} \left| e^{\hat{T}(\nu)} \right| \Phi_{\nu} \right\rangle \delta_{\varrho\mu} \left\langle \Phi_{\nu} \left| \hat{H} e^{\hat{T}(\nu)} \right| \Phi_{\zeta} \right\rangle \right] = 0 \qquad (2.166)$$

After introducing an "effective lambda" matrix elements as

$$L_{\varrho\mu}^{\text{eff}} = \left\langle \Phi_{\varrho} \left| \hat{\Lambda}(\varrho) e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
(2.167)

the final λ equation becomes

$$\left\langle \Phi_{\varrho} \left| \tilde{\Lambda}(\varrho) \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle - \sum_{\mu}^{\mathcal{P}} \left[H_{\varrho\mu}^{\text{eff}} \left\langle \Phi_{\mu} \left| \hat{\Lambda}(\mu) e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle + L_{\varrho\mu}^{\text{eff}} \left\langle \Phi_{\mu} \left| \hat{H} e^{\hat{T}(\varrho)} \right| \Phi_{\zeta} \right\rangle \right] = 0 \qquad (2.168)$$

The gradient of energy by (2.145) reads

$$\frac{dE}{d\chi} = \sum_{\mu\nu}^{\mathcal{P}} \tilde{c}^{\alpha}_{\nu} c^{\alpha}_{\mu} \left\langle \Phi_{\nu} \left| \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle
+ \sum_{\mu}^{\mathcal{P}} \left\langle \Phi_{\mu} \left| \hat{\Lambda}(\mu) \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle
- \sum_{\mu\nu}^{\mathcal{P}} \left\langle \Phi_{\mu} \left| \hat{\Lambda}(\mu) e^{\hat{T}(\nu)} \right| \Phi_{\nu} \right\rangle \left\langle \Phi_{\nu} \left| \hat{H}^{\chi} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle$$
(2.169)

which can be rewritten using the above definitions of $\tilde{\Lambda}$ and $L_{\varrho\mu}^{\rm eff}$ to the final compact form

$$\frac{dE}{d\chi} = \sum_{\mu}^{\mathcal{P}} \left[\langle \Phi_{\mu} | \tilde{\Lambda}(\mu) - \sum_{\nu}^{\mathcal{P}} L_{\mu\nu}^{\text{eff}} \langle \Phi_{\nu} | \right] \hat{H}^{\chi} e^{\hat{T}(\mu)} | \Phi_{\mu} \rangle$$
(2.170)

Both the λ equation and the gradient formula thus mix CC and λ amplitudes of different reference configurations.

2.9 Perturbative Triples Contribution in the MR MkCC Method

The perturbative triples correction in MRCC methods was first introduced in the SU MRCCSD(T) approach [17] and recently also in the MR BWCCSD(T) method with the *a posteriori* size-extensivity correction [18]. The first suggested perturbative triples contribution in the multireference Mukherjee coupled cluster theory, denoted here as MR MkCCSD(T_u), has been derived in an analogous way [19]. A brief introduction to these approaches is presented in the following.

After the MR CCSD amplitude equations are solved, the approximate \hat{T}_3 amplitudes are calculated from

$$t_{ijk}^{abc}(\mu) = \frac{\left\langle \left(\Phi_{\mu}\right)_{ijk}^{abc} \left| \hat{V}_{N}(\mu) \, \hat{T}_{2}(\mu) \right| \, \Phi_{\mu} \right\rangle_{C} - \sum_{\nu \neq \mu}^{\mathcal{P}} \left[\left\langle \left(\Phi_{\mu}\right)_{ijk}^{abc} \left| \hat{T}_{2}(\nu) \right| \, \Phi_{\nu} \right\rangle H_{\mu\nu}^{\text{eff}} \right]_{C}}{D_{ijk}^{abc}(\mu)}$$

$$(2.171)$$

for the SU MRCC method or from an uncoupled equation

$$t_{ijk}^{abc}(\mu) = \frac{\left\langle (\Phi_{\mu})_{ijk}^{abc} \left| \hat{V}_{N}(\mu) \, \hat{T}_{2}(\mu) \right| \, \Phi_{\mu} \right\rangle_{C}}{D_{ijk}^{abc}(\mu)} \tag{2.172}$$

for the MR BWCC method, in which the \hat{T}_2 amplitudes were already provided with the *a posteriori* size-extensivity correction. The effect of non-diagonal Fock matrix elements is neglected so that the \hat{T}_3 equation does not require an iterative solution.

The perturbative triples correction enters the effective Hamiltonian where its diagonal elements are analogous to the single-reference CCSD(T) energy correction

$$H_{\mu\mu}^{\text{eff}}(T) = H_{\mu\mu}^{\text{eff}}(\text{CCSD}) + E_{T}^{[4]}(\mu) + E_{ST}^{[5]}(\mu) + E_{ST}^{[4]}(\mu)$$
(2.173)

where the fourth and fifth order terms are given by

$$E_{\rm T}^{[4]}(\mu) = \sum_{\substack{ijk\\abc}} \left\langle (\Phi_{\mu})_{ijk}^{abc} \left| \hat{V}_{\rm N}(\mu) \, \hat{T}_{2}(\mu) \right| \, \Phi_{\mu} \right\rangle_{C} \, t_{ijk}^{abc}(\mu) \tag{2.174}$$

$$E_{\rm ST}^{[5]}(\mu) = \sum_{i} s_i^a(\mu) t_i^a(\mu)$$
(2.175)

$$E_{\rm ST}^{[4]}(\mu) = \frac{1}{4} \sum_{\substack{ijk\\abc}} f_{kc}(\mu) t_{ij}^{ab}(\mu) t_{ijk}^{abc}(\mu)$$
(2.176)

where

$$s_i^a(\mu) = \frac{1}{4} \sum_{\substack{jk\\bc}} \langle bc \mid \mid jk \rangle \ t_{ijk}^{abc}(\mu)$$

$$(2.177)$$

Due to the lack of the coupling terms in (2.172), the $E_{\rm T}^{[4]}$ term in the MR BWCCSD(T) method becomes symmetric and resembles the single-reference CCSD(T)

$$E_{\rm T}^{[4]}(\mu) = \frac{1}{36} \sum_{\substack{ijk\\abc}} t_{ijk}^{abc}(\mu) D_{ijk}^{abc}(\mu) t_{ijk}^{abc}(\mu)$$
(2.178)

The off-diagonal elements of the effective Hamiltonian are computed at the CCSDT-1 level with linear \hat{T}_3 contribution to both \hat{T}_1 and \hat{T}_2 equations

$$H_{\nu\mu}^{\text{eff}}(T) = \left\langle \Phi_{\nu} \left| H_{N}(\mu) e^{\hat{T}_{123}(\mu)} \right| \Phi_{\mu} \right\rangle_{C}^{\text{CCSDT}-1} \\ = \left\langle \Phi_{\nu} \left| H_{N}(\mu) \left(e^{\hat{T}_{12}(\mu)} + \hat{T}_{3}(\mu) \right) \right| \Phi_{\mu} \right\rangle_{C}$$
(2.179)

The MR CCSD(T) energy is finally obtained by diagonalization of the (T)-corrected effective Hamiltonian.

In the MR MkCCSD(T_n) approximation, the coupling terms of the \hat{T}_3 equation, which mix the $t_{ijk}^{abc}(\mu)$ and $t_{ijk}^{abc}(\nu)$ amplitudes, are neglected so that the equation actually becomes identical to (2.172) and thus does not require an iterative solution. The $E_{\rm T}^{[4]}(\mu)$ term is also symmetric like in the MR BWCCSD(T). However an intruder state problem can come up if the denominator of the (2.172) equation approaches zero.

2.9.1 MR MkCCSD (T_i) Method

Later, a quite different derivation appeared [20], denoted here as MR MkCCSD(T_i), which does not suffer from intruder states. It is based on a perturbative analysis of the MR MkCC Lagrangian [20,86,87] and represents a multireference generalization of the A-CCSD(T) method [88–91]. In contrast to the MR CCSD(T) approaches described above in which the (T)-correction entered the effective Hamiltonian before diagonalization, the perturbative triples correction to the energy is here calculated explicitly. Thus instead of the "perturb then diagonalize", the "diagonalize then perturb" scheme is used.

The triples equation contains the linear coupling terms

$$\left\langle \Phi_{\vartheta} \left| e^{-\hat{T}(\mu)} e^{\hat{T}(\nu)} \right| \Phi_{\mu} \right\rangle_{\text{linear}} = t^{abc}_{ijk}(\nu/\mu) - t^{abc}_{ijk}(\mu)$$
(2.180)

where

$$t_{ijk}^{abc}(\nu/\mu) = \begin{cases} t_{ijk}^{abc}(\nu) & \text{if } i, j, k \in \text{occ}(\mu) \cap \text{occ}(\nu); a, b, c \in \text{virt}(\mu) \cap \text{virt}(\nu) \\ 0 & \text{otherwise} \end{cases}$$
(2.181)

The $t_{ijk}^{abc}(\mu)$ terms in the \hat{T}_3 amplitude equation (2.92) are moved to the lefthand side and after an application of the \hat{H}^{eff} eigenvalue equation, the final \hat{T}_3 equation will contain the Brillouin–Wigner type denominator, yielding

$$t_{ijk}^{abc}(\mu) = \frac{\left\langle \left(\Phi_{\mu}\right)_{ijk}^{abc} \left| \hat{V}_{N}(\mu) \, \hat{T}_{2}(\mu) \right| \, \Phi_{\mu} \right\rangle + \sum_{\nu \neq \mu}^{\mathcal{P}} t_{ijk}^{abc}(\nu/\mu) \, H_{\mu\nu}^{\text{eff}} \, \frac{c_{\nu}^{\alpha}}{c_{\mu}^{\alpha}}}{D_{ijk}^{abc}(\mu) + (E_{\alpha} - H_{\mu\mu}^{\text{eff}})} \qquad (2.182)$$

The energy corrections are then calculated explicitly to the energy obtained from the MR MkCCSD effective Hamiltonian. The $E_{\rm T}^{[4]}(\mu)$ term is nonsymmetric because of the couplings. Although this method is resistant to intruders because of the BW-type denominator in (2.182), the \hat{T}_3 equations have to be solved iteratively with the complexity $\mathcal{O}(\mathcal{N}^6)$ due to the presence of the coupling terms.

2.9.2 MR MkCCSD (T_u) Approximation

In order to avoid the need to iteratively solve the \hat{T}_3 equation like in the MR MkCCSD(T_i) method, an approach based on the uncoupled approximation to MR MkCC has been suggested [3], denoted by MR MkCCSD(T_u). The triples equation thus becomes

$$t_{ijk}^{abc}(\mu) = \frac{\left\langle \left(\Phi_{\mu}\right)_{ijk}^{abc} \left| \hat{V}_{N}(\mu) \, \hat{T}_{2}(\mu) \right| \, \Phi_{\mu} \right\rangle_{C} - \sum_{\nu \neq \mu}^{\mathcal{P}} \left\langle \left(\Phi_{\mu}\right)_{ijk}^{abc} \left| e^{\bar{T}_{\nu}(\mu)} \right| \, \Phi_{\mu} \right\rangle H_{\mu\nu}^{\text{eff}} \frac{c_{\nu}^{\alpha}}{c_{\mu}^{\alpha}}}{D_{ijk}^{abc}(\mu)}$$

$$(2.183)$$

which does not mix amplitudes of different references and can thus be solved explicitly. If all the linear \hat{T}_3 terms are moved to the left-hand side, the equation yields

$$t_{ijk}^{abc}(\mu) = \frac{1}{D_{ijk}^{abc}(\mu) + \sum_{\nu \neq \mu}^{\dagger} H_{\mu\nu}^{eff} \frac{c_{\nu}^{\alpha}}{c_{\mu}^{\alpha}}} \left\{ \left\langle (\Phi_{\mu})_{ijk}^{abc} \left| \hat{V}_{N}(\mu) \, \hat{T}_{2}(\mu) \right| \Phi_{\mu} \right\rangle_{C} - \sum_{\nu \neq \mu}^{\mathcal{P}} \left[P(i/jk) \, P(a/bc) \, \bar{t}_{\nu \, i}^{a}(\mu) \, \bar{t}_{\nu \, jk}^{bc}(\mu) - P(ijk) \, \bar{t}_{\nu \, i}^{a}(\mu) \, \bar{t}_{\nu \, j}^{b}(\mu) \, \bar{t}_{\nu \, k}^{c}(\mu) \right] H_{\mu\nu}^{eff} \frac{c_{\nu}^{\alpha}}{c_{\mu}^{\alpha}} \right\}$$
(2.184)

where the \dagger sign at the sum means that only such terms are included in which at least one of the orbitals i, j, k, a, b or c has different occupation in the μ th and ν th reference.

By comparing with equation (2.182) of MR MkCCSD(T_i) variant, it is obvious that the appropriate \hat{T}_3 amplitudes can be obtained from (2.184) in just one step. Moreover, because of the denominator shift, the equation can be resistant to intruders and has thus an advantage over the MR MkCCSD(T_n) method. Although the denominator shift is in general different from the one in equation (2.182), which is a complete BW-type shift, its analysis for the most common CAS(2,2) reference space [3] showed that the shift in (2.184) should be sufficient to avoid the intruder state problems.

The method performs the "perturb then diagonalize" scheme with the asymmetric form of the $E_{\rm T}^{[4]}(\mu)$ term. The "diagonalize then perturb" variant is also possible, but test calculations showed that the corrected energy values differ only in the order of $10^{-6}a.u$.

Chapter 3

Implementation of the Analytical Gradient

3.1 Pilot Implementation

The pilot implementation of all three described MRCC analytical gradient variants at the CCSD level has been coded in a program package called TINY, which is mainly written and maintained by Dr. Jiří Pittner as a reference, benchmarking and debugging tool used during development of various multireference coupled cluster models. It features its own modules for integral and integral derivatives evaluation, SCF, CPHF, integral transformation, full-CI and MRCC. It is a C++ code widely based on a free ¹ open-source C++ library interface to the BLAS [93] and LAPACK [94] linear algebra library routines, called LA [95]. That enables convenient coding of various vector and matrix operations while keeping the high efficiency of the program.

The core of the implementation is based on the Knowles–Handy algorithm [96,97] for action of a Full-CI expanded Hamiltonian on a trial vector without the need to construct the expanded Hamiltonian explicitly. This is also used for the Hamiltonian derivative \hat{H}^{χ} and generalized for $\hat{T}(\mu)$ and $\hat{\Lambda}(\mu)$ operators which take into account the particular Fermi vacua with respect to which the amplitudes are defined. This algorithm is used for straightforward evaluation of various terms in the amplitude or λ equations as well as in the gradient formulae.

The amplitude equation is solved iteratively using a residual vector that vanishes when the respective coupled cluster equations are converged, similarly as in Ref. [98]. The amplitude update is performed by the formula

$$t_{\vartheta}^{\text{new}}(\mu) = t_{\vartheta}^{\text{old}}(\mu) + \frac{\left\langle \Phi_{\mu} \left| \hat{\tau}_{\vartheta}^{\dagger}(\mu) \right| \Phi_{\vartheta} \right\rangle \mathcal{Q}_{\vartheta}(\mu)}{E_{\alpha} - H_{\mu\mu}^{\text{eff}} + D_{\vartheta}(\mu)}$$
(3.1)

¹the library is distributed under the Gnu General Public License (v3) [92]

where the first term in the numerator is just a sign factor of a given excitation and $D_{\vartheta}(\mu)$ is the standard CC denominator used in the single-reference amplitude equation updates, which consists of the diagonal elements of the Fock matrix with respect to the Fermi vacuum $|\Phi_{\mu}\rangle$.

Since $\mathcal{Q}_{\vartheta}(\mu)$ vanishes at convergence, the choice of the denominator is in principle arbitrary and influences only the convergence rate. A dynamic denominator shift is thus implemented, ensuring that the denominator never approaches zero.

The λ equation is first conjugated so that it has similar structure like the MRCC amplitude equation and then is solved analogously. Only the denominator is set to $D_{\vartheta}(\mu) + C$ where the constant C is set so that all denominators are in absolute value greater than one.

The effect of molecular orbital relaxation in response to the perturbation is involved in terms of the U matrices in the integral transformation formulas

$$\frac{\partial h_{ij}}{\partial \chi} = \sum_{\mu\nu}^{AO} \frac{\partial h_{\mu\nu}}{\partial \chi} C_{\mu i} C_{\nu j} + \sum_{\mu\nu}^{AO} \sum_{m}^{MO} h_{\mu\nu} \left[C_{\mu m} U_{mi}^{\chi} C_{\nu j} + C_{\mu i} C_{\nu m} U_{mj}^{\chi} \right]$$
(3.2)

$$\frac{\partial (ij|kl)}{\partial \chi} = \sum_{\substack{\mu\nu\\\varrho\sigma}}^{AO} C_{\mu i} C_{\nu j} C_{\varrho k} C_{\sigma l} \frac{\partial (\mu\nu|\varrho\sigma)}{\partial \chi}
+ \sum_{\substack{\mu\nu\\\varrho\sigma}}^{AO} \sum_{m}^{MO} (\mu\nu|\varrho\sigma)
\times \left[C_{\mu m} U_{mi}^{\chi} C_{\nu j} C_{\varrho k} C_{\sigma l} + C_{\mu i} C_{\nu m} U_{mj}^{\chi} C_{\varrho k} C_{\sigma l}
+ C_{\mu i} C_{\nu j} C_{\varrho m} U_{mk}^{\chi} C_{\sigma l} + C_{\mu i} C_{\nu j} C_{\varrho k} C_{\sigma m} U_{ml}^{\chi} \right]$$
(3.3)

which thus requires to solve the CPHF equations for each gradient component.

3.2 Testing of the Analytical Gradient Implementation

Testing of the implemented analytical gradients was performed on methylene CH_2 and silylene SiH_2 molecules in the ${}^{1}A_1$ state. These systems have almost single-reference character in their equilibrium geometry, which has the

bond angle about 100° in the C_{2v} symmetry, but by opening the bond angle the multireference description increases up to two equally weighted reference configurations

$$|\Phi_0\rangle = (\operatorname{core}) (a_1)^2 (b_1)^2 (a_1)^2 (b_2)^0$$
 (3.4)

$$|\Phi_1\rangle = (\operatorname{core}) (a_1)^2 (b_1)^2 (a_1)^0 (b_2)^2$$
 (3.5)

when the molecules become linear. The possible monoexcited configurations within the two active orbitals do not contribute due to the C_{2v} symmetry.

First, the energy and its analytical gradient of the CH_2 molecule were calculated for different values of the bond angle, ranging from 100° to 179° describing thus the transition from an almost single-reference to the exactly two-reference system, while keeping a fixed bond length. The molecule was positioned in such a way (see Figure 3.1) that the gradient consisted only of three non-zero non-equivalent components C_z , H_z and H_y . Two basis



Figure 3.1: Position of the CH_2 molecule in the coordinate system and its three non-zero non-equivalent gradient components

sets were employed, the 6–31G and the frozen core 6–31G^{*}. Figure 3.2 shows the energy and analytical gradient dependence on the bond angle for a fixed bond length 1.11 Å in the 6–31G basis for the iteratively corrected BWCCSD method. The other two methods do not differ neither in energy nor its gradient from the iteratively corrected BWCC results by more than 10^{-3} hartree (hartree/bohr) and provide pictures indistinguishable from one another in the same scale.

At each point, a numerical gradient was also calculated by finite energy differences with the numerical step set to 10^{-4} bohr. At every displaced geometry the Hartree–Fock orbitals were first reoptimized before the MRCC



Figure 3.2: Iteratively corrected 2R BWCCSD energy and gradient components of CH_2 as a function of bond angle with fixed C–H bond length 1.11 Å in the 6–31G basis

energy was computed so that the full orbital relaxation was taken into account. All numerical BWCC gradient calculations were done by the ACES II program since it has an efficient MR BWCCSD implementation.

The analytical gradient was compared to the numerical differentiation by an average error Δ , given by the formula

$$\Delta = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left[\left(\frac{\partial E}{\partial x_i} \right)_{\text{analytical}} - \left(\frac{\partial E}{\partial x_i} \right)_{\text{numerical}} \right]^2}$$
(3.6)

where M ranges the non-zero nonequivalent gradient components.

The dependence of the average error Δ on the bond angle together with a square of the expansion coefficient of the first reference configuration describing the varying multireference character are shown in Figure 3.3. The most accurate gradient was provided by the uncorrected BWCC method that gave error values below 10^{-8} in the whole range of bond angles. Results from both basis sets also nicely resembled each other. In contrast, the average errors of the other two methods varied more significantly, especially the iteratively corrected BWCC that reached almost 10^{-6} in two points. Values obtained from the two basis sets also considerably differed in many cases. Nevertheless, for any method the errors never exceeded 10^{-6} and such result is completely satisfactory.



Figure 3.3: Gradient error Δ of the CH₂ molecule in a logarithmic scale and the expansion coefficient of the first reference configuration as a function of the bond angle with fixed bond lengths 1.11 Å, compared for 6–31G and frozen core 6–31G* basis sets. The expansion coefficient curves in both basis sets are almost indistinguishable in the scale of this figure and only one is shown.

A geometry optimization of the two molecules in two basis sets was then employed, starting from a shape that was far from the equilibrium geometry so as to apply the whole range of the multireference character. All three methods performed very well and the λ amplitudes converged quickly at each point of the optimization path. Since the pilot implementation didn't allow to use larger basis set than 6–31G^{*} with frozen core orbitals on the available computer resources so that the results were comparable with experiments, a Full-CI geometry optimization has thus been performed as a reference. The results for each system together with the Full-CI calculation and a comparison with experimental values are summarized in Table 3.1 and Table 3.2.

Method	Basis	Energy	R	Φ
BW Uncorrected	6-31G	-38.942637	1.129	103.7
	$6-31G^*$ fzc	-38.996567	1.116	102.0
BW Iter. corr.	6-31G	-38.943014	1.128	103.8
	$6-31G^*$ fzc	-38.997243	1.116	102.1
SUMRCC	6-31G	-38.942991	1.129	103.7
	$6-31G^*$ fzc	-38.997052	1.116	102.0
Full-CI	6-31G	-38.944209	1.130	103.6
	$6-31G^*$ fzc	-38.999103	1.118	101.8
Experiment [99]			1.107 ± 0.002	102.4 ± 0.4

Table 3.1: ${}^{1}A_{1}$ CH₂ optimal energy and geometry obtained from all three MRCC methods in 6–31G and 6–31G^{*} frozen core basis sets compared to the Full-CI and an experiment. Energy is in atomic units, bond length in Å and bond angle in degrees.

All the methods proved to give results very close to the Full-CI values, differing by the order of 10^{-3} Å and 0.1° . The largest difference from the Full-CI geometries represents the SiH₂ bond angle, which was overestimated by 0.4° by the SU MRCCSD method using the smaller basis set. In the larger basis, the iteratively corrected MR BWCCSD underestimated the Si-H bond length by 0.003 Å and overestimated the bond angle of both molecules by 0.3°.

3.3 Efficient Implementation of the Analytical Gradient of the MR BWCCSD with the Iterative Correction of Size–Extensivity

For the efficient analytical gradient implementation of the MR BWCCSD method with the iterative correction of size-extensivity, it was intended to make use of the efficient analytical gradient code of the single-reference CCSD method, which is available in the ACES II program package. The single-

Method	Basis	Energy	R	Φ
BW Uncorrected	6-31G fzc	-290.034835	1.575	92.8
	$6\text{-}31\mathrm{G}^*$ fzc	-290.092095	1.530	92.5
BW Iter. corr.	6-31G fzc	-290.034932	1.575	92.8
	$6\text{-}31\mathrm{G}^*$ fzc	-290.092553	1.529	92.6
SUMRCC	6-31G fzc	-290.035033	1.576	93.1
	$6\text{-}31\mathrm{G}^*$ fzc	-290.092417	1.530	92.5
Full-CI	6-31G fzc	-290.035669	1.576	92.7
	$6\text{-}31\mathrm{G}^*$ fzc	-290.094287	1.532	92.3
Experiment [100]			1.51402	91.9830
Experiment [101]			1.5141	92.0
Experiment [102]			1.516	92.08

Table 3.2: ${}^{1}A_{1}$ SiH₂ optimal energy and geometry obtained from all three MRCC methods in 6–31G frozen core and 6–31G* frozen core basis sets compared to the Full-CI and an experiment. Energy is in atomic units, bond length in Å and bond angle in degrees.

reference formalism was developed by Salter *et al.* [71] using the normal ordered operators which leads to the formulas (2.135) for the λ equation and (2.133) for the gradient.

The multireference λ equation and gradient were thus derived in a similar way, resulting in formulas that look a little different from those used in the pilot implementation. The MR λ equation thus yields

$$\left\langle \Phi_{\mu} \left| \tilde{\Lambda}(\mu) \left[\hat{H}_{N}(\mu) e^{\hat{T}(\mu)} \right]_{C} \right| \Phi_{\zeta} \right\rangle_{C} + \sum_{\substack{ext.\\ \vartheta(\mu)}}^{\text{int.}} \left\langle \Phi_{\mu} \left| \left[\hat{H}_{N}(\mu) e^{\hat{T}(\mu)} \right]_{C} \right| \Phi_{\vartheta} \right\rangle \left\langle \Phi_{\vartheta} \left| \tilde{\Lambda}(\mu) \right| \Phi_{\zeta} \right\rangle + \tilde{c}_{\mu}^{\alpha} c_{\mu}^{\alpha} \left\langle \Phi_{\mu} \left| \left[\hat{H}_{N}(\mu) e^{\hat{T}(\mu)} \right]_{C} \right| \Phi_{\zeta} \right\rangle - \sum_{\nu \neq \mu}^{\mathcal{P}} \sum_{\vartheta(\mu)}^{\text{ext.}} \lambda_{\vartheta}(\mu) \left\langle \Phi_{\vartheta} \left| \hat{\tau}_{\zeta}(\mu) \right| \Phi_{\nu} \right\rangle H_{\nu\mu}^{\text{eff}} = 0 \quad (3.7)$$

and the MR gradient becomes

$$\frac{d E_{\alpha}}{d \chi} = \sum_{\mu}^{\mathcal{P}} \tilde{c}^{\alpha}_{\mu} c^{\alpha}_{\mu} \left\langle \Phi_{\mu} \left| \hat{H}^{\chi}_{N} e^{\hat{T}(\mu)} \right| \Phi_{\mu} \right\rangle_{C} + \sum_{\mu}^{\mathcal{P}} \left\langle \Phi_{\mu} \left| \tilde{\Lambda}(\mu) \left[\hat{H}^{\chi}_{N}(\mu) e^{\hat{T}(\mu)} \right]_{C} \right| \Phi_{\mu} \right\rangle \quad (3.8)$$

in which the redefined Λ operator (2.147) was used involving also the internal λ amplitudes. The formulas are very close to their single-reference analogs. In the λ equation, the third term is just scaled by the expansion coefficients of the particular reference in the model wavefunction and the additional fourth term actually vanishes for references mutually at most biexcited, which is always valid in the MR BWCCSD implementation available in the ACES II program. Note also that the λ equations of different references remain uncoupled. The multireference gradient formula is a sum of independent contributions from each reference configuration, which differ from the single-reference gradient formula only by the scaling factor at the first term.

The efficient single-reference gradient implementation has been described by Gauss *et al.* [72]. They gather all terms with derivatives of the Fock matrix elements and remaining terms with derivatives of the two-electron integrals, which results in the formula

$$\frac{d E}{d \chi} = \sum_{pq} D_{pq} \frac{\partial f_{pq}}{\partial \chi} + \sum_{pqrs} \Gamma_{pqrs} \frac{\partial \langle pq \mid \mid rs \rangle}{\partial \chi}$$
(3.9)

where the D_{pq} and Γ_{pqrs} are the one and two-particle relaxed density matrices. The derivatives can be divided in skeletons and terms containing the U matrices

$$\frac{dE}{d\chi} = \sum_{pq} D_{pq} f_{pq}^{(\chi)} + \sum_{pqrs} \Gamma_{pqrs} \langle pq \mid \mid rs \rangle^{\chi} - 2 \sum_{pq} I'_{pq} U^{\chi}_{pq}$$
(3.10)

where the skeleton part of the Fock matrix is defined as

. ...

$$f_{pq}^{(\chi)} = h_{pq}^{\chi} + \sum_{m} \langle pm \mid \mid qm \rangle^{\chi}$$
(3.11)

The U matrices are then eliminated by the Z-vector technique using the fact that the single-reference CC energy is invariant with respect to rotations among occupied (or virtual) orbitals. The efficient single-reference CC gradient formula thus finally becomes

$$\frac{dE}{d\chi} = \sum_{pq} D_{pq} f_{pq}^{(\chi)} + \sum_{pqrs} \Gamma_{pqrs} \langle pq \mid \mid rs \rangle^{\chi} + \sum_{pq} I_{pq} S_{pq}^{\chi}$$
(3.12)

which consists only of the relaxed density matrices, the skeleton terms and the I intermediate and does not require the solution of the CPHF equations for each gradient component.

In the multireference case, such U matrix elimination cannot be performed in a general way, but the MR variants of the (3.9) and (3.10) formulas are always valid and can thus be used. Since the MR gradient formula (3.8) just sums contributions from each reference configuration, the MR relaxed density matrices can be computed as a sum of relaxed density matrices of a particular reference that is calculated by a modified single-reference code.

The suggested multireference implementation that is based on the singlereference CCSD analytical gradient code in the ACES II program package but which needs to solve the CPHF equations to account for the orbital response thus requires the following principal steps:

- provide the appropriate storage for the λ amplitudes, density matrices and other variables which correspond to a particular reference configuration and introduce loops over all references in the reference-dependent parts of the code (solving the λ equation, formation of the density matrices, ...)
- setting the internal λ amplitudes and identifying the terms in the λ equation which have to be scaled
- identify and scale the terms corresponding to the first term in (3.8)
- calculate also the D_{ia} blocks of the one-particle density matrix since the single-reference code presumes the use of the Hartree–Fock orbitals for which the f_{ia} elements are zero, which is however not the case in the multireference generalization
- gather the total D and Γ relaxed density matrices
- form the I' intermediate and solve the U matrices
- evaluate the gradient (3.10) by the modified single-reference code for (3.12)

While it was relatively easy to generalize the single-reference code for the multireference calculation and to identify all the terms that had to be modified in the λ equations and the relaxed density matrices, an enormous effort was put in an attempt to finish the gradient evaluation in the ACES II package. For efficiency reasons, all the intermediates calculated by the program scale elements corresponding to a particular occupied or virtual orbital index combination by various factors or add another terms. Such efficiency improvements have been very poorly documented and often had to be decoded element by element. A special program had to be designed, based on the code of the pilot gradient implementation, which calculated the unmodified form of such terms so that the values could be compared. Because of such difficulties, it was finally suggested to export the full form of the calculated total D and Γ matrices to an external code, which solves the CPHF equations and evaluates the MO integral derivatives and finally to calculate the gradient by

$$\frac{d E_{\alpha}}{d \chi} = \sum_{pq} D_{pq} \frac{\partial h_{pq}}{\partial \chi} + \sum_{pqrs} \Gamma_{pqrs} \frac{\partial \langle pq \mid \mid rs \rangle}{\partial \chi}$$
(3.13)

which is a variant of (3.9), but instead of the Fock matrix derivatives uses derivatives of the one-electron operator \hat{h} matrix elements. This required a minor modification of the total Γ matrix, but was easier to implement.

The analytical gradient implemented in this way was first tested on the CH_2 molecule by comparing with values obtained from the pilot implementation. The largest difference was below 10^{-8} hartree/bohr which is a completely satisfactory result.

A geometry optimization of the two benchmark molecules CH_2 and SiH_2 was also performed with results summarized in Table 3.3. The CH_2 results in the 6-31G basis which was used also in the testing of the pilot implementation agree with the result obtained earlier (Table 3.1). The other calculations however employed larger basis sets.

System	Basis	Energy	R	Φ
CH ₂	6-31G	-38.943015	1.128	103.8
	cc- $pVDZ$	-39.025637	1.124	101.1
Experiment [99]			1.107 ± 0.002	102.4 ± 0.4
SiH ₂	6-31G	-290.038273	1.575	92.8
	6-311G	-290.180587	1.563	93.4
Experiment [100]			1.51402	91.9830

Table 3.3: A geometry optimization of the ${}^{1}A_{1}$ state of the CH₂ and SiH₂ molecules by the efficient analytical gradient implementation of the MR BWCCSD method with the iterative correction of size-extensivity. Energy is in atomic units, bond length in Å and bond angle in degrees.

To roughly illustrate the gradient evaluation time improvement gained by the ACES II implementation over the pilot implementation, the Table 3.4 presents average real times of calculation of the CH_2 gradient on an AMD Opteron CPU at 2.8GHz with 16 GB RAM for the various basis set sizes used. Only the 6-31G basis was used in both implementations, since the ACES II code cannot calculate the gradient using frozen orbitals and larger basis sets lead to insufficient memory for the pilot implementation. At the 6-31G basis, the real time of calculation reduced almost by the factor 10^3 . The analytical gradient computed by the ACES II implementation was also always faster than the numerical gradient evaluation, although the gain did not exceed the factor 2 for this molecule. The ACES II implementation can thus be considered as efficient.

Although the calculations of the benchmark molecules performed very well, application of the implemented gradient to optimization of larger molecular systems like N₂O₂ or a Cyclobutadiene (C₄H₄) turned out to be problematic since in many points the multireference λ equations failed to converge

Basis	# MOs	TINY	ACES II
6-31G	13	5930	8
$6-31G^*$ fzc	18	10800	
cc-pVDZ	25		44
6-311G**	31		105

Table 3.4: Average real computer time of a CH_2 analytical gradient evaluation calculated by either the program TINY (pilot implementation) or by the ACES II implementation in various basis sets. The calculations were performed on an AMD Opteron CPU at 2.8 GHz / 16 GB RAM. The time is in seconds.

and thus the optimization procedure did not finish. This could be explained by the fact that the iterative size-extensivity correction of the method reintroduces the intruder state problem.

Since an analytical gradient of the MR MkCCSD method has recently been implemented by Prochnow *et al.*, which is a superior method to the iteratively corrected MR BWCC variant and because of the convergence difficulties of the λ equations experienced in calculations of larger molecules, it was suggested that the more efficient orbital response contribution to the gradient is not going to be further developed.

Chapter 4

Application of the MR MkCCSD (T_u) on a Singlet-Triplet Gap Investigation of Tetramethyleneethane (TME)

The newly developed MR MkCCSD(T_u) method has already been successfully tested on the BeH₂ molecule and proved to provide smooth potential energy curve that is free of singularities while closely resembling the results from the more expensive MR MkCCSD(T_i) approximation [3]. To test the method also in a real chemical application, an investigation of the singlettriplet separation of the tetramethyleneethane molecule (TME) has been suggested, because of a multireference character of its singlet state.

Since 1970 that the first EPR spectrum of the system was published [103], a number of experimental and theoretical studies followed, trying to assign the ground state with either singlet or triplet multiplicity [104–113]. A gas phase negative ion photoelectron spectroscopy experiment by Clifford *et al.* [110] found that the ground state is singlet, being about 2 kcal mol⁻¹ more stable than the triplet state. This was however in contradiction with previously reported matrix isolation EPR studies. An explanation has been suggested that the matrix could actually fix the molecule at a triplet equilibrium geometry, at which the singlet energy was above the triplet. Calculations by spin-restricted open-shell Kohn–Sham (ROKS) and spin-restricted ensemblereferenced Kohn–Sham (REKS) methods [111] predicted the singlet ground state at \mathcal{D}_{2d} geometry while the triplet being about 3 kcal mol⁻¹ above. Also the difference dedicated configuration interaction calculations [112] confirmed the singlet as the ground state for all conformations, although the energy difference was only 0.29 kcal mol⁻¹ at the equilibrium triplet geometry

The structure of the TME molecule (C_6H_8) is pictured in Figure 4.1. Each end of the ethylene skeleton forms a plane together with its bonded methylene groups and the two planes are mutually twisted.



Figure 4.1: Geometry of the TME molecule is formed by two planes rotated mutually by a twist angle

4.1 Computational

The potential energy surface has been scanned in the following way. For each value of the twist angle (from 0° to 90°), a restricted geometry optimization was performed (keeping the molecular symmetry) using the CASPT2(6,6) method and the cc-pVDZ basis set for both the singlet and the triplet state so that also a dynamical electronic correlation was partially involved in the prediction of the geometries. In these optimized geometries, CASSCF(2,2) molecular orbitals were employed for further correlation treatment. Calculations using CASSCF(6,6) orbitals were also performed in some points, but the results were almost identical. Two basis sets were employed, the spherical cc-pVDZ and a modified spherical cc-pVTZ basis (denoted as cc-pVTZ'), from which the d-functions at Hydrogen and f-functions at Carbon atoms were removed so that the system remained computationally feasible. The six core orbitals at Carbon atoms and the six highest virtual orbitals were kept frozen during the following coupled cluster calculations.

The potential energy curves were calculated at the CCSD, CCSD(T) and in the smaller basis set also at the approximate iterative CCSDT-1 level. The singlet state employed the two-reference BWCC and MkCC methods. All MR BWCC calculations were provided with the *a posteriori* correction of size-extensivity. The perturbative triples in the MkCC method used both the standard "perturb then diagonalize" scheme, MR MkCCSD(T_u) and also the other variant with an explicit energy correction evaluation, denoted as MR MkCCSD(T_u)e. The two variants gave values which differed less than $2 \times 10^{-6} a.u.$ so only the MR MkCCSD(T_u) results are presented. The triplets were calculated by the standard single-reference CC methods. All the coupled cluster calculations were performed by the ACES II program [1], while the CASSCF orbitals and CASPT2 geometries were obtained using the MOL-PRO package [114].

4.2 Results and Discussion

The calculated potential energy curves are illustrated in Figure 4.2, which compares both the effect of increasing level of CC approximation (in rows) and the two basis sets used (in columns). The MR BWCC method predicted the singlet as the ground state in all cases and every point of the calculated curves. The MR MkCCSD calculations put the singlet curve above the triplet in both basis sets and only the inclusion of approximate triples in the MkCC method moved the singlet curve below the triplet. The low energy of the MR BWCC singlet curves even at the CCSD level can be put down to the fact that the *a posteriori* size-extensivity correction tends to overestimate the correlation energy.

The difference between the singlet and the triplet curves for various levels of approximation and the two basis sets used is shown in Figure 4.3. All the curves are very similar, with a minimum at about 45° , but differ in the position on the vertical axis. At the CCSD level, the MR BWCC result lies between 1 and 4 kcal mol^{-1} being in a good agreement with both the gas phase experiment and other theoretical predictions. On the other hand, the MR MkCCSD method gave even a different order of the curves. However, by inclusion of the perturbative triples, the MR MkCCSD (T_u) method provided very good results between 2 and 6 kcal mol^{-1} in both basis sets, which also closely resembled each other. The triples inclusion in MR BWCC method resulted in rather large energy gap between 11 and 14 kcal mol^{-1} , which can be explained by the overestimated correlation energy due to the *a posteriori* size-extensivity correction. It can be concluded that the relatively good result of the MR BWCCSD calculation could be obtained due to the compensation of this energy overestimation by the insufficient correlation description at the CCSD level. The approximate iterative triples in MR MkCCSDT-1 method gave results very close to the MR $MkCCSD(T_u)$ values. At the curve minimum, the MR MkCCSDT-1 method gave 2.4 kcal mol^{-1} and the MR MkCCSD(T_{u}) 2.6 kcal mol⁻¹ in the cc-pVDZ basis. These results are thus in a close agreement with the gas phase negative ion photoelectron spectroscopy experiment [110].


Figure 4.2: Potential energy curves of TME as a function of the twist angle. Comparison of the singlet state (2R-BWCC and 2R-MkCC) and the triplet state (SR-CC) calculations at different levels of CC approximation. Graphs in the left column were obtained from the cc-pVDZ basis, the right column corresponds to the cc-pVTZ' basis.



Figure 4.3: The singlet-triplet energy difference in TME for the CCSD (blue), CCSD(T) (green) and CCSDT-1 (red) level of approximation. The label ^{*a*} denotes the cc-pVDZ and the label ^{*b*} the cc-pVTZ' basis set.

4.3 Conclusion

The singlet-triplet energy separation of the tetramethyleneethane molecule has been studied by the newly developed state-specific multireference $MkCCSD(T_u)$ method, which is based on the uncoupled MR MkCC approximation. The calculated potential energy curves were smooth and the results well agreed with an experiment.

By comparison with the MR MkCCSD and MR MkCCSDT-1 approximations, it was shown that even for obtaining proper qualitative results, the inclusion of the triples was inevitable. The results calculated from the perturbative triples corrections closely resembled those from the MR MkCCSDT-1 level of approximation and thus for obtaining quantitative results, the MR MkCCSD(T_u) approximation is sufficient enough.

For comparison, the MR BWCC method with the *a posteriori* correction of size-extensivity was also employed, but gave rather poor results at the CCSD(T) level, which could be explained by its tendency to overestimate the correlation energy.

The perturbative triples correction in the uncoupled approximation to the MR MkCC method was able to provide quantitative results. Unlike the MR MkCCSD(T_i) method, the MR MkCCSD(T_u) approximation does not need to solve the \hat{T}_3 equation iteratively and is thus less expensive while remaining resistant to intruder states.

Chapter 5 Conclusion

Three Hilbert space multireference coupled cluster methods, the state-specific MR BWCCSD with the iterative correction of size-extensivity and without a correction and the state-universal MR CC method in the formulation by Kucharski and Bartlett, have been provided with a pilot implementation of an analytical gradient of energy, which is based on a Full-CI expansion.

The implementation was tested by comparing the analytical gradients with gradients calculated numerically at various geometries of different multireference character. At all points, all three gradient variants performed very well, providing results of sufficient accuracy.

The pilot implementation was then applied to a geometry optimization of the singlet CH_2 and SiH_2 molecules, which have a two-reference character. The results were compared both to experimental values and Full-CI calculations. All the methods differed from the Full-CI values by the order of 10^{-3} Å and 0.1° using double-zeta basis sets and such results are satisfactory enough. The pilot implementation of the three MRCC analytical gradient variants has been published [2].

The analytical gradient of the MR BWCCSD method with the iterative correction of size-extensivity has also been implemented in the ACES II quantum chemical program package. It makes use of the efficient single-reference CCSD analytical gradient code since the formulas of this MRCC gradient variant are very similar. The orbital response contribution to the gradient however requires solving the CPHF equations for each gradient component. A more efficient orbital response evaluation like in the single-reference case cannot be applied to the multireference formalism and further development is needed. The ACES II program thus calculates the multireference relaxed density matrices, which are then contracted with MO integral derivatives by an external code.

This implementation has been successfully tested by comparing with the pilot implementation and it has also been applied to the CH_2 and SiH_2 geometry optimization. The time of the gradient evaluation reduced almost by a factor 10^3 in the 6-31G basis comparing to the pilot implementation. In

spite of the inefficient orbital response calculation, the implemented gradient is still faster than the numerical evaluation and the implementation is thus efficient.

However, applications to larger systems than the benchmark molecules like N_2O_2 or cyclobutadiene failed since at most geometries the multireference λ equations did not converge. This could be explained by the fact that the iterative correction of the MR BWCC method reintroduces the intruder state problems.

After the pilot implementation has been published, an analytical gradient of the state-specific multireference Mukherjee coupled cluster method has been implemented by Prochnow *et al.* [16], which is however superior to the iteratively corrected BWCC method. Because of this fact and for the severe convergence problems of the BWCC λ equations, it has been decided that the efficient orbital response contribution for the BWCC gradient is not going to be further developed.

Finally, the newly developed MR MkCCSD(T_u) method has been successfully tested on the investigation of the singlet-triplet energy separation of the tetramethyleneethane molecule. The calculated results nicely agreed with an experiment, suggesting the ground state to be singlet for all values of the twist angle with 2.6 kcal mol⁻¹ as the minimum in the cc-pVDZ basis.

It has been shown that even for the qualitatively correct description, the inclusion of the triples was inevitable. However, since the MR MkCCSD(T_u) results closely resembled those from the MR MkCCSDT-1 level of approximation, the perturbative triples inclusion is sufficient enough.

The tested MR MkCCSD(T_u) method, which is based on the uncoupled approximation to the MR MkCC method, has thus proved to be a less expensive alternative to the MR MkCCSD(T_i) approximation that has to solve the \hat{T}_3 equations iteratively, yet remains resistant to intruder states. This part of the thesis has also been published [3].

Bibliography

- ACES II is a program product of the Quantum Theory Project, University of Florida. Authors: J.F. Stanton, J. Gauss, J.D. Watts, M. Nooijen, N. Oliphant, S.A. Perera, P.G. Szalay, W.J. Lauderdale, S.A. Kucharski, S.R. Gwaltney, S. Beck, A. Balková D.E. Bernholdt, K.K. Baeck, P. Rozyczko, H. Sekino, C. Hober, and R.J. Bartlett. Integral packages included are VMOL (J. Almlöf and P.R. Taylor); VPROPS (P. Taylor) ABACUS; (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, J. Olsen, and P.R. Taylor).
- [2] J. Pittner and J. Šmydke, J. Chem. Phys. **127**, 114103 (2007).
- [3] K. Bhaskaran-Nair, O. Demel, J. Smydke, and J. Pittner, J. Chem. Phys. 134, 154106 (2011).
- [4] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical Recipes in C (Cambridge University Press, New York, 1992).
- [5] L. Verlet, PhysRev **159**, 98 (1967).
- [6] L. Verlet, PhysRev **165**, 201 (1967).
- [7] E. R. Davidson, in *The World of Quantum Chemistry*, edited by R. Daudel and B. Pullman (D. Reidel, Dordrecht, 1974), pp. 17–30.
- [8] S. R. Langhoff and E. R. Davidson, Int. J. Quant. Chem. 8, 61 (1974).
- [9] P. G. Szalay and R. J. Bartlett, J. Chem. Phys. **101**, 4936 (1994).
- [10] P. G. Szalay, Int. J. Quant. Chem. Symp. 55, 151 (1995).
- [11] J. F. Stanton and J. Gauss, J. Chem. Phys. **101**, 8938 (1994).
- [12] S. Pal, Phys. Rev. A **39**, 39 (1991).
- [13] S. Pal, Int. J. Quant. Chem. **41**, 443 (1992).
- [14] N. Vaval, R. Roy, and S. Pal, Phys. Rev. A 49, 1623 (1994).
- [15] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **95**, 8227 (1991).

- [16] E. Prochnow, F. A. Evangelista, H. F. Schaefer III, W. D. Allen, and J. Gauss, J. Chem. Phys. **131**, 064109 (2009).
- [17] A. Balková and R. J. Bartlett, J. Chem. Phys. 101, 8972 (1994).
- [18] O. Demel and J. Pittner, J. Chem. Phys. **124**, 144112 (2006).
- [19] K. Bhaskaran-Nair, O. Demel, and J. Pittner, J. Chem. Phys. 129, 184105 (2008).
- [20] F. A. Evangelista, E. Prochnow, J. Gauss, and H. F. Schaefer III, J. Chem. Phys. 132, 074107 (2010).
- [21] A. Szabo and N. S. Ostlund, Modern Quantum Chemistry (Mineola, N.Y.: Dover Publications, New York, 1996).
- [22] R. E. Christoffersen, Basic Principles and Techniques of Molecular Quantum Mechanics (Springer-Verlag, Berlin, 1989).
- [23] P. Jørgensen and J. Simons, Second Quantization Based Methods in Quantum Chemistry (Academic Press, New York, 1981).
- [24] P. R. Surján, Second Quantized Approach to Quantum Chemistry (Springer-Verlag, Berlin, 1989).
- [25] T. D. Crawford and H. F. Schaefer III, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2000), Vol. 14.
- [26] F. Coester, NuclPhys 7, 421 (1958).
- [27] F. Coester and H. Kümmel, NuclPhys **17**, 477 (1960).
- [28] J. Cížek, J. Chem. Phys. 45, 4256 (1966).
- [29] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quant. Chem. Symp. 14, 545 (1978).
- [30] R. J. Bartlett and G. D. Purvis, Int. J. Quant. Chem. 14, 561 (1978).
- [31] B. Jeziorski and H. J. Monkhorst, Phys. Rev. A 24, 1668 (1981).
- [32] J.-P. Malrieu, P. Durand, and J.-P. Daudey, J. Phys. A 18, 809 (1985).
- [33] L. Meissner, J. Chem. Phys. **108**, 9227 (1998).
- [34] A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. 113, 9905 (2000).
- [35] N. Vaval, S. Pal, and D. Mukherjee, Theor. Chem. Acc. 99, 100 (1998).

- [36] X. Li and J. Paldus, J. Chem. Phys. **107**, 6257 (1997).
- [37] X. Li and J. Paldus, Int. J. Quant. Chem. 80, 743 (2000).
- [38] X. Li and J. Paldus, J. Chem. Phys. **124**, 174101 (2006).
- [39] T. Kinoshita, O. Hino, and R. J. Bartlett, J. Chem. Phys. 123, 074106 (2005).
- [40] O. Hino, T. Kinoshita, G. K.-L. Chan, and R. J. Bartlett, J. Chem. Phys. 124, 114311 (2006).
- [41] N. Oliphant and L. Adamowicz, J. Chem. Phys. **95**, 6645 (1991).
- [42] P. Piecuch, N. Oliphant, and L. Adamowicz, J. Chem. Phys. 99, 1875 (1993).
- [43] P. Piecuch, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. 110, 6103 (1999).
- [44] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).
- [45] P. Piecuch, S. A. Kucharski, K. Kowalski, and M. Musial, Comp. Phys. Commun. 149, 71 (2002).
- [46] P. Piecuch, K. Kowalski, I. S. O. Pimienta, P. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski, T. Kus, and M. Musial, Theor. Chem. Acc. 112, 349 (2004).
- [47] K. Kowalski and P. Piecuch, J. Chem. Phys. **122**, 074107 (2005).
- [48] M. Wloch, J. R. Gour, K. Kowalski, and P. Piecuch, J. Chem. Phys. 122, 214107 (2005).
- [49] P. Piecuch and M. Wloch, J. Chem. Phys. **123**, 224105 (2005).
- [50] J. Pittner, J. Chem. Phys. **118**, 10876 (2003).
- [51] I. Hubač, J. Pittner, and P. Čársky, J. Chem. Phys. **112**, 8779 (2000).
- [52] J. Pittner, J. Šmydke, P. Čársky, and I. Hubač, J. Mol. Struct. (Theochem) 547, 239 (2001).
- [53] J. Pittner, O. Demel, P. Čársky, and I. Hubač, Int. J. Mol. Sci. 2, 281 (2001).
- [54] M. Tobita, S. A. Perera, M. Musial, R. J. Bartlett, M. Nooijen, and J. S. Lee, J. Chem. Phys. **119**, 10713 (2003).
- [55] J. C. Sancho-García, J. Pittner, P. Cársky, and I. Hubač, J. Chem. Phys. **112**, 8785 (2000).

- [56] J. Pittner, P. Nachtigall, P. Cársky, and I. Hubač, J. Phys. Chem. A 105, 1354 (2001).
- [57] O. Rey-Puiggros, J. Pittner, P. Cársky, P. Stampfuß, and W. Wenzel, Collect. Czech. Chem. Commun. 68, 2309 (2003).
- [58] O. Demel, J. Pittner, P. Cársky, and I. Hubač, J. Phys. Chem. A 108, 3125 (2004).
- [59] I. S. K. Kerkines, P. Cársky, and A. Mavridis, J. Phys. Chem. A 109, 10148 (2005).
- [60] J. Brabec and J. Pittner, J. Phys. Chem. A **110**, 11765 (2006).
- [61] J. Pittner and O. Demel, J. Chem. Phys. **122**, 181101 (2005).
- [62] O. Demel and J. Pittner, J. Chem. Phys. **128**, 104108 (2008).
- [63] S. Das, N. Bera, S. Ghosh, and D. Mukherjee, J. Mol. Struct. (Theochem) 771, 79 (2006).
- [64] S. Das, D. Datta, R. Maitra, and D. Mukherjee, Chem. Phys. 349, 115 (2008).
- [65] O. Demel, K. Bhaskaran-Nair, and J. Pittner, J. Chem. Phys. 133, 134106 (2010).
- [66] Y. Yamaguchi, J. D. Goddard, Y. Osamura, and H. F. Schaefer III., A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory (Oxford University Press, New York, 1994).
- [67] L. Adamowicz, W. D. Laidig, and R. J. Bartlett, Int. J. Quant. Chem. Symp. 18, 245 (1984).
- [68] R. J. Bartlett, in Geometrical Derivatives of Energy Surfaces and Molecular Properties, edited by P. Joergensen and J. Simons (D. Reidel Publishing Company, Dordrecht, 1986), pp. 35–61.
- [69] G. Fitzgerald, R. J. Harrison, and R. J. Bartlett, J. Chem. Phys. 85, 5143 (1986).
- [70] A. C. Scheiner, G. E. Scuseria, J. E. Rice, and T. J. Lee, J. Chem. Phys. 87, 5361 (1987).
- [71] E. A. Salter, G. W. Trucks, and R. J. Bartlett, J. Chem. Phys. 90, 1752 (1989).
- [72] J. Gauss, J. F. Stanton, and R. J. Bartlett, J. Chem. Phys. 95, 2623 (1991).

- [73] J. Gauss, J. F. Stanton, and R. J. Bartlett, J. Chem. Phys. 95, 2639 (1991).
- [74] J. Gauss, W. J. Lauderdale, J. F. Stanton, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. 182, 207 (1991).
- [75] J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. 98, 8718 (1993).
- [76] G. E. Scuseria, J. Chem. Phys. **94**, 442 (1991).
- [77] J. D. Watts, J. Gauss, and R. J. Bartlett, Chem. Phys. Lett. 200, 1 (1992).
- [78] J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. 98, 8718 (1993).
- [79] K. K. Baeck, J. D. Watts, and R. J. Bartlett, J. Chem. Phys. 107, 3853 (1997).
- [80] J. F. Stanton, J. Chem. Phys. **99**, 8840 (1993).
- [81] J. F. Stanton and J. Gauss, J. Chem. Phys. **100**, 4695 (1994).
- [82] T. J. Lee and A. P. Rendell, J. Chem. Phys. **94**, 6229 (1991).
- [83] J. Gauss and J. F. Stanton, J. Chem. Phys. **116**, 1773 (2002).
- [84] P. Jørgensen and T. Helgaker, J. Chem. Phys. 89, 1560 (1988).
- [85] T. Helgaker and P. Jørgensen, Theor. Chim. Acta 75, 111 (1989).
- [86] F. A. Evangelista, A. C. Simmonett, H. F. Schaefer III, D. Mukherjee, and W. D. Allen, Phys. Chem. Chem. Phys. 11, 4728 (2009).
- [87] P. Jørgensen and T. Helgaker, J. Chem. Phys. 89, 1560 (1988).
- [88] J. F. Stanton, Chem. Phys. Lett. **281**, 130 (1997).
- [89] T. D. Crawford and J. F. Stanton, Int. J. Quantum Chem. 70, 601 (1998).
- [90] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. 108, 5243 (1998).
- [91] A. G. Taube and R. J. Bartlett, J. Chem. Phys. **128**, 044110 (2008).
- [92] GNU, http://www.gnu.org/.
- [93] BLAS, http://www.netlib.org/blas/.
- [94] LAPACK Linear Algebra PACKage, http://www.netlib.org/lapack/.

- [95] LA library: C++ interface to BLAS, LAPACK and CUBLAS linear algebra libraries, http://www.pittnerovi.com/la/.
- [96] P. J. Knowles and N. C. Handy, Chem. Phys. Lett. **111**, 315 (1984).
- [97] P. J. Knowles and N. C. Handy, J. Chem. Phys. **91**, 2396 (1989).
- [98] J. Gauss, in *The Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollmann, H. F. Schaefer III, and P. R. Scheiner (Wiley, Chichester, 1998), pp. 615–636.
- [99] H. Petek, D. J. Nesbitt, D. C. Darwin, P. R. Ogilby, and C. B. Moore, J. Chem. Phys. 91, 6566 (1989).
- [100] R. Escribano and A. Campargue, J. Chem. Phys. **108**, 6249 (1998).
- [101] W. D. Allen and H. F. Schaefer III, Chem. Phys. **108**, 243 (1986).
- [102] I. Dubois, Can. J. Phys. 46, 2485 (1968).
- [103] P. Dowd, J. Am. Chem. Soc. **92**, 1066 (1970).
- [104] W. Borden and E. R. Davidson, J. Am. Chem. Soc. **99**, 4587 (1977).
- [105] P. Dowd, W. Chang, and Y. H. Paik, J. Am. Chem. Soc. 108, 7416 (1986).
- [106] P. Du and W. T. Borden, J. Am. Chem. Soc. **109**, 930 (1987).
- [107] P. Nachtigall and K. D. Jordan, J. Am. Chem. Soc. 114, 4743 (1992).
- [108] P. Nachtigall and K. D. Jordan, J. Am. Chem. Soc. **115**, 270 (1993).
- [109] P. Dowd, W. Chang, C. J. Partian, and W. Zhang, J. Phys. Chem. 97, 13408 (1993).
- [110] E. P. Clifford, P. G. Wenthold, W. C. Lineberger, G. B. Ellison, C. X. Wang, J. J. Grabowski, F. Vila, and K. D. Jordan, J. Chem. Soc., Perkin Trans. 2, 1015 (1998).
- [111] M. Filatov and S. Shaik, J. Phys. Chem. A **103**, 8885 (1999).
- [112] E. Rodriguez, M. Reguero, and R. Caballol, J. Phys. Chem. A 104, 6253 (2000).
- [113] J. Pittner, P. Nachtigall, P. Čársky, and I. Hubač, J. Phys. Chem. A 105, 1354 (2001).

[114] H.-J. Werner, P. J. Knowles, F. R. Manby, M. Schütz, P. Celani, G. Knizia, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, and A. Wolf, MOLPRO, version 2006.1, a package of ab initio programs, 2006, see http://www.molpro.net.