NTNU
Norwegian University of
Science and Technology

Reduced orbital space coupled cluster calculations using a multilevel Hartree-Fock wave function

Sandra Sæther

Chemical Engineering and Biotechnology
Submission date: June 2017
Supervisor: Ida-Marie Høyvik, IKJ
Co-supervisor: Henrik Koch, IKJ

Norwegian University of Science and Technology
Department of Chemistry

Abstract

Embedding, or multilevel, schemes have become exceedingly popular in recent years. The idea is to divide the system into an active part, which is treated by an accurate quantum mechanical method, and then treat the rest of the system, referred to as the inactive part, with a less accurate, more computationally efficient method. In this work, recent developments in multilevel Hartree-Fock theory (MLHF) is connected with coupled cluster singles and doubles (CCSD) to enable computation of accurate size-intensive properties for large molecular systems. The method needs no a priori orbital assignment, no bonds are broken and one wavefunction describes the whole system. One part of the wave function is described at CCSD level, one part at HF level and the other kept constant after an initial density start guess. The objective is to explore whether a full HF calculation is needed for a reduced space CC model, or if only a region around the CCSD space needs to be optimized. Specifically, will the reduced space CCSD results be compromised when embedded in a MLHF wave function rather than a HF wave function.

The MLHF method makes use of a molecular orbital (MO) density matrix driven optimization procedure to optimize only a subset of the MOs in the Slater determinant by using an orthogonal parametrization that ensures conserved symmetry, trace and idempotency properties for both the active and inactive density matrix. This is connected with CCSD by generating a set of canonical MOs from the active density, and using these MOs and the orbital energies as a starting point for the CCSD calculation. Since the reference wave function is not a HF wave function, this introduces some changes in the CCSD working equations. All changes enters through a modified Fock matrix.

Some benchmarking calculations on core excitations have been carried out on smaller systems to establish the effect of only having a partly optimized HF state as a reference state, and the method is showing promising results. The errors are small, but still large enough to compromise the CCSD calculation. However, as this is for small systems, these preliminary results are promising for the application on larger molecular systems. From the results, it is expected that local CC calculations on large systems can be carried out in a MLHF framework without compromising accuracy.

Samandrag

Multilevel metodar har blitt meir og meir poplære dei siste åra. Ideen er å dele systemet inn i ein aktive del, som blir behandla med ei nøyaktig kvante mekanisk metode, og så behandle resten av systemet, som blir referert til som den inaktive delen, med ei mindre nøyaktig, meir berekningseffektiv metode. I dette arbeidet er nylege utviklingar i multilevel Hartree-Fock teori (MLHF) kopla saman med coupled cluster singles doubles (CCSD) for å gjere det mogleg å berekne nøyaktige itensive eigenskapar for store molekylære system. Metoda treng ingen a priori orbital tildeling, bryt ingen bindingar og har ein bølgefunksjon som beskriv heile systemet. Ein del av bølgefunksjonen er beskriven på CCSD nivå, ein del på HF nivå og den siste delen blir heldt konstant etter eit initielt startgjett. Objektivet er å undersøke om ei full HF berekning er naudsynt for ein redusert rom CC modell, eller om berre ein region rundt CCSD rommet treng å bli optimalisert. Spesifikt, vil resultatet frå ei redusert rom CCSD berekning bli påverka når metoda blir kopla med ein MLHF bølgefunksjon istadenfor ein HF bølgefunskjon.

MLHF metoda brukar ein molekylorbital (MO) tettleiksmatrise styrt optimalisering til å optimere berre eit utval av MOane i Slater determinanten ved å bruke ei ortogonal parametrisering som sikrar at symmetri, spor og idempotent eigenskapane er oppfylt for både den aktive og den inaktive tettleiksmatrisa. Dette er kopla med CCSD ved å generere eit sett av kanoniske MOar frå den aktive tettleiken, og desse pluss orbital energiane er brukt som startpunkt i CCSD berekninga. Sidan referansefunksjonen ikkje er ein HF bølgefunksjon vil dette føre til endringar i CCSD likningane. Alle endringar oppstår i form av ei modifisert Fock matrise.

Nokre referansekalkulasjonar på kjerneeksitasjonar har blitt gjennomført på mindre system for å undersøke effekta av å ikkje ha ein fullstendig optimert HF tilstand som referanse, og metoda gir lovande resultat. Feila er små, men fortsatt av ein slik størrelse at dei vil kompromittere CCSD berekninga. Sidan dette er for små system, er dette lovande for berekningar på større molekylære system. Frå desse resultata er det forventa at lokale CC berekningar på større system kan bli gjort inne i eit MLHF rammeverk utan at det vil komprimere nøyaktigheita.

Preface

This thesis presents the conclusion of my study in the Chemical Engineering and Biotechnology integrated master program at the Norwegian University of Science and Technology.

I would like to thank my awesome supervisor Ida-Marie Høyvik for all her help, advice and support throughout this process. I would also like to thank my co-supervisor, Henrik Koch, for his helpful input in this project.

Furthermore my thanks goes to the people at my study hall, for helping keeping each other more or less sane these last weeks. It has been a bumpy ride.

Last, but not least, my love goes to Cato Seljevold, my parents Thor-Eldar and Gunnhild, and my sister Malin for their endless support and encouragement.

Sandra Sæther,
Trondheim, June 18, 2017

Contents

1 Introduction **1**

2 Theory **4**
2.1 The Hartree-Fock approximation 5
2.1.1 Exponential parametrization of the density matrix in the full MO basis 6
2.1.2 Multilevel Hartree-Fock formulation 8
2.1.3 Start guess for the total density 10
2.1.4 Partition of the total density matrix 10
2.2 Basis sets and post-Hartree-Fock methods 11
2.3 Coupled cluster theory 13
2.4 Excitation energies in Coupled Cluster 15

3 Method **16**
3.1 Preparation for enabling CCSD in MLHF 17
3.1.1 Terminology for the division of the system 17
3.1.2 The Fock matrix 17
3.1.3 Obtaining the embedding terms 18
3.2 CCSD procedure 20
3.2.1 Working equations 20

4 Proof of concept calculations **23**
4.1 Computational details 23
4.2 Lysine 24
4.2.1 Oxygen 24
4.2.2 Nitrogen 29
4.3 β-peptide 33
4.3.1 Oxygen 33
4.3.2 Nitrogen 36

5 Summary and concluding remarks **39**

6 Further work **40**

1 Introduction

In time-independent electronic structure theory, the goal is to obtain molecular properties such as geometry, total energy and excitation energies. This is done by solving the electronic Schrödinger equation. Unfortunately, the solution can not be found analytically for a system containing more than one electron, and approximations has to be made. The *ab initio* methods offers a hierarchy of approximated wave functions, which permits a systematic way to increase the accuracy. The starting point in this hierachy is the *Hartree-Fock* (HF) wave function. HF is a useful method in its own right as a qualitative treatment of larger systems, as the HF energy usually deviates by 1 % from the true energy [1], but it falls short when a more accurate description is desired. The reason is the neglection of the *electron correlation energy*, which is defined as the difference between the HF-limit and the true non-relativistic ground state energy for a complete basis [2]. *Post HF methods*, like *coupled cluster* (CC) and *configuration interaction* (CI) are commonly using HF as a starting point, and tries to recover the correlation energy. The end point of this hierarchy is a full configuration interaction (FCI) calculation, in which the exact solution of the Schrödiger equation for a given basis set is obtained. However, this method scales exponentially with the size of the system, and is only feasible for small systems and basis sets [1]. Thereby, the method has to be truncated to reduce the scaling, and CI loses some of its appeal as it is no longer *size-extensive* [3]. Truncated CC is however size-extensive at every level of approximation, and includes more of the lost correlation energy than CI at the same level of correlation. Because of this, CC is usually considered to be the better method of these two. However, even though CC captures more of the correlation energy than CI, its most common truncation, *coupled cluster singles doubles* (CCSD) [4], does not include enough of these effects to obtain chemical accuracy [5]. The accuracy obtained by including the triple excitations therefore often is required, and again the problem of scaling arises. *Coupled cluster singles doubles triplets*, CCSDT [6], scales as n^8, where n is the number of basis functions, so it is only applicable on small systems. There exist methods that approximates the effect of the triplets, e.g. CCSD(T) [7] and CC3 [8], but these methods still scales as n^7 and will be expensive on larger systems.

As can be seen from the discussion above, the dilemma of accuracy vs computational cost is a prominent problem. The more accurate the method, the more expensive the calculations will be, and the smaller the system they can be performed on becomes. Systems that typically are of chemical interest, such as biomolecules, liquid phases and surfaces, are simply too big to be described by the higher level methods, such as the CC hierarchy of methods, but methods such as HF, which uses significantly less computer time, will not provide the level of chemical accuracy needed. A good compromise, that have become exceedingly popular, are so-called *multi-level*, or *frozen density embedded* schemes. If the property in question is only affected by a small part of the system, the idea is to combine an accurate quantum method that describe the local event in the electronic system, and then use a lower-level, more computationally efficient method on the environment. The first

multi-level methods were QM/MM methods, in which a semi-empirical quantum mechanical method were combined with a molecular mechanics description of the environment. These methods were introduced by Warshel and Lewitt [9], [10], and they received the Nobel prize in Chemistry in 2013 along with Karplus for their work. A more recent QM/MM model, ONIOM, was developed by Svenson et al. in 1996 and is still very much in use [11]. The MM-layer can also be modified to use an inexpensive ab intio method such as HF or a *density-functional theory* (DFT) calculation. Other methods might be added, as ONIOM can have several layers [12]. For a review of these QM/MM methods, see Senn and Thiel [13]. Also Continuum Solvation Models are popular. In one of these models, the *polarizable continuum method* (PCM), the environment is incorporated as a dielectric polarizable continuum [14]. For a more detailed review of the different continuum models, see the review by Tomasi et. al. [15]

In recent years, methods involving embedding various level of DFT have also been developed. DFT is an alternative route to the ab intio hierachy, as, in the Kohn-Sham density-functional theory (KS-DFT), the functionals of the molecular electronic probability density is used instead of the wave function to calculate properties of the system. DFT has the advantage of allowing for correlation effects to be included in a calculation that takes roughly the same time as a HF calculation. However, DFT also has some problems. Conventional DFT is a ground state theory, meaning that, in general, excitations are beyond its scope. It is not *variational*, meaning that it can yield an energy below the true ground state energy. It is also a problem that, even though the Hohenberg-Kohn theorems prove that there is a one-to-one mapping between the 3N dimensional many-electron wave function and the three dimensional electron density, the exact physics is not known. The unknown term is the exchange-correlation functional, E_{xc}, which have been approximated in many ways, but the true functional form is not found [16]. In DFT-in-DFT embedding schemes, an embedding operator is constructed, and the goal is to obtain how this operator interacts with the region of interest. Several approaches exist, such as subsystem DFT [17], partition DFT [18] and the subsystem embedding methods [19] [20]. The drawbacks of these methods in general is the unknown exchange-correlation functional, and, as pointed out by Fornace et al. [19], the requirement of a priori knowledge of the chemical bonding. They introduce embedded mean-field theory to solve this problem, and the method preforms as good as ONIOM, but as pointed out by Kallay et al. [21], this method only enables DFT-in-DFT-type embedding. In wave function (WF)-DFT, DFT is used as embedding for more accurate methods [21–24]. DFT is not treated in this work, and for a review of DFT-in-DFT embedding, see the review by Wesolowski et al. [25]. For a comparison of DFT-in-DFT and WF-in-DFT methods, see Kallay et. al [21].

The method presented in this thesis will be a WF-in-WF method, meaning that we will combine the recently developed method multilevel HF (MLHF) with CCSD. In MLHF the system is divided into an active part which is treated at HF level, and an inactive part which is kept constant after an initial start guess. There has not been

much development of such a scheme until recently. There exists several WF-in-WF embedding schemes, such as MP2 in HF by Godvin et al.[26], CC-in-CC by Höfener and Visscher[27], the hybrid scheme of Mata et. al[28], and the cluster in molecule approach by Li and Piechuch[29], but all of these depends on having solved either the HF or Kohn-Sham problem in advance. In 1996 an embedded HF method was used by Shukla et al.[30], in which the Roothan-Hall equations are solved for molecular orbtials in a reference cell in the basis of the reference and neighbouring cell in the AO basis. The environment field is included through two-electron terms from the neighbouring cells. As solving the Roothan-Hall equations requires a diagonalization of the Fock matrix, which is now the time critical step, this will be problematic with large reference cells. The density based embedding scheme by Godvind et al.[26] could also be used for HF calculations, but what differentiates their approach from the MLHF used here is that the two electron interactions in the environment, and between the environment and the active part, is neglected. One of the most recent method was developed in the master thesis by Dundas[31] under the supervison of Ida-Marie Høyvik, in which an exponential parametrization developed by Helgaker et al.[32] is used to do MLHF in an Atomic Orbital (AO) basis. The diagonalization step is avoided by optimizing the density matrix directly. In the approach developed by Sæther et. al[33] (to be submitted), a molecular orbital (MO) basis is used instead, enabling a reduction of dimensions as it is possible to work only in the active molecular orbital space. The method needs no a priori orbital assignment, no bonds are broken and one wavefunction describes the whole system, where one part is optimized and the other kept constant. This is all done in LSDALTON[34]. For this thesis, a subpart of the HF optimized wavefunction is then optimized further in the CCSD code in DALTON[34]. All interactions between the active and inactive parts is kept through inactive two-electrons integrals that interacts with the optimized wavefunction at each iteration. For size-itensive properties, this approach should yield good results. However, as already mentioned, the CCSD method often need the correlating effects from the triplets, so the end goal of this method would be to couple the MLHF method with the multilevel CC (MLCC) method by Myhre et al.[35], in which CC3 will be added on top as the highest level. CC3 is here preferred over CCSD(T) as molecular properties are to be examined, and CC3 is then a better choice as it has a wavefunction linked to its approximation, unlike CCSD(T) which is a two-step procedure that adds the triplets effects by perturbation[1].

This thesis is structured as follows. In Section 2 some essential theory will be introduced, including an outline of the MLHF method. In Section 3 an outline of the implementation is shown, and the fusion of the two methods are explained. In Section 4 the results of some proof of concept and benchmarking calculations are shown and discussed. In Section 5 some concluding remarks are presented, and finally further work is discussed in Section 6.

2 Theory

All the properties of a quantum mechanical system is stored in the system's wave function, ψ. To obtain it, we have to solve the Schrödinger equation, which in its time independent form is given as

$$\hat{H}\psi = E\psi. \tag{2.1}$$

By using the *Born-Oppenheimer approximation*, which states that due to the large mass difference between the nuclei and the electrons of the system, the electrons will instantaneously respond to any movement in the nuclei, the separation of the system's wavefunction in to a nuclear and an electronic contribution is justified. This means that in the electronic Hamiltonian, the nuclei can be viewed as stationary [2], and in the molecular orbital basis of a system it can be written as

$$\hat{H} = \sum_{pq} h_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} e_{pqrs} + h_{nuc}, \tag{2.2}$$

in the second quantization formalism. h_{nuc} is then the constant nuclear repulsion term. E_{pq} denotes the singlet excitation operators,

$$E_{pq} = a^{\dagger}_{p\alpha} a_{q\alpha} + a^{\dagger}_{p\beta} a_{q\beta} \tag{2.3}$$

and e_{pqrs} the two electron excitation operator,

$$e_{pqrs} = \sum_{\sigma\tau} a^{\dagger}_{p\sigma} a^{\dagger}_{r\tau} a_{s\tau} a_{q\sigma}. \tag{2.4}$$

$a^{\dagger}$ and a are the creation and annihilation operator, respectively, and combined $a^{\dagger}_{p\alpha} a_{q\alpha}$ represent an excitation of an electron with spin α from spin-orbital $\phi_{\alpha q}$ to $\phi_{\alpha p}$. A spin orbital is the product state of a spatial orbital ϕ_p and the spin state $\sigma = \{\alpha, \beta\}$. The term

$$h_{pq} = \int \phi^*_p(r)(-\frac{1}{2}\nabla^2 - \sum_I \frac{Z_I}{r_I})\phi_q(r)dr \tag{2.5}$$

and

$$g_{pqrs} = \int dr_1 \int \frac{\phi^*_p(r_1)\phi_q(r_1)\phi^*_r(r_2)\phi_s(r_2)}{|r_1 - r_2|} dr_2 \tag{2.6}$$

are the one- and two-electron integrals, respectively, in atomic units. ϕ denotes a molecular orbital, ϕ^* its complex conjugate, Z_I the nuclear charge of atom I, and r_I the nuclear-electron separation in equation 2.5 [1]. In this thesis, only real orbitals will be used. Because of the electron interaction term the Schrödinger equation is not separable and it cannot be solved exactly when more than two particles interact. For larger systems only approximate solutions are feasible.

2.1 The Hartree-Fock approximation

In this section, the HF formulation and some of the approaches to solve it will be explained. Both the traditional Roothan-Hall approach, the more unconventional exponential parametrization introduced by Helgaker et. al [32] and the newly developed MLHF in the molecular orbital basis will be presented.

The most basic ab initio wave function model is the HF wave function. The starting point of the HF method is to assume that the wave function for a closed shell N_e-electron state can be represented by a single Slater determinant. A Slater determinant is an antisymmetrized product of spinorbitals. As only closed shell states will be considered in this thesis, this will be assumed throughout the next sections. This version of HF is called *closed-shell restricted* HF (RHF), where the wave function is restricted to be an eigenfunction of the total spin of the system [1]. All orbitals will either be unoccupied orbitals, called virtual orbitals and denoted a,b,c and d, or doubly occupied orbtials, called inactive orbitals denoted i,j,k and l. Finally general orbitals will be denoted p,q,s and r throughout the thesis. The Slater determinant for such a system may be defined as

$$|\Phi_{SD}\rangle = \prod_{i=1}^{N_e/2} a_{i\alpha}^\dagger a_{i\beta}^\dagger |vac\rangle, \tag{2.7}$$

where vac is the vacuum state, $a_p|vac\rangle = 0$. The energy of the optimized system is found by calculating the *expectation value* of the Hamiltonian of the system, given in equation 2.2,

$$E_{HF} = \langle \Phi_{HF} \mid \hat{H} \mid \Phi_{HF} \rangle = 2 \sum_{i}^{N_e/2} h_{ii} + \sum_{ij}^{N_e/2} \left(2g_{iijj} - g_{ijji}\right) + h_{nuc}. \tag{2.8}$$

The first term is the one-electron term that includes the kinetic energy of one electron and its interaction with the nuclei. The second is the Coulomb term, which includes the potential energy of the electron we are looking at interacting with an electrostatic field created by the average of the other electrons. This is why HF often is referred to as the mean-field approximation. The third term is called the exchange term, and has no physical analogue, but falls out from the Slater-determinant and fulfills the Pauli principle [2]. To find the HF wave function, the energy is minimized under the constraint that the spin orbitals should stay orthonormal. As HF is a *variational method*, the resulting energy will be an upper bound for the true ground state energy of the system [16]. Thus, the set of orbitals that minimizes the energy will be the best possible single determinant wave function. To find these spin orbitals, an effective one-electron Schrödinger equation is solved, where the Fock operator, $\hat{F}$, replaces the Hamiltonian. Thus the equations to solve become

$$\hat{F} a_{p\sigma}^\dagger |vac\rangle = \varepsilon_p a_{p\sigma}^\dagger |vac\rangle. \tag{2.9}$$

As the Hamiltonian, the Fock operator needs to be Hermetian and totally symmetric in spin-space, and therefore has the form

$$\hat{F} = \sum_{pq} F_{pq} E_{pq}, \tag{2.10}$$

where F_{pq} is the elements of the symmetric Fock matrix, $\mathbf{F}$, and E_{pq} the one-electron excitation operator. In the basis where the Fock matrix is diagonal, the eigenvalues ε_p in equation 2.9 are interpreted as the orbital energies, and the spin orbitals are called canonical orbitals. However, to satisfy the variational conditions, any orbitals that ensure that the Fock matrix has a block diagonal form with vanishing occupied-virtual elements , $F_{ai} = F_{ia} = 0$ is sufficient [1].

The classical way to solve the HF equations is by the Roothan-Hall method, in which the molecular orbitals are expanded in terms of one-electron basis functions, also called atomic orbitals (AO),

$$\phi_p = \sum_{\mu} C_{p\mu} \chi_{\mu}. \tag{2.11}$$

The set of all AOs employed in a particular calculation is referred to as the AO basis. This expansion is then introduced into the energy expression in equation 2.8, which is optimized with the MO coefficient $C_{\mu p}$ as variational parameters. The set of equations obtained after introducing this and using a canonical basis is the Roothan-Hall HF equations,

$$\mathbf{F}^{\mathrm{AO}} \mathbf{C} = \mathbf{S} \mathbf{C} \boldsymbol{\varepsilon} \tag{2.12}$$

where $\mathbf{F}^{\mathrm{AO}}$ is the Fock matrix in the AO basis, $\mathbf{S}$ is the atomic orbital overlap matrix, $\boldsymbol{\varepsilon}$ contains the orbital energies and $\mathbf{C}$ is a matrix containing the expansion coefficients for each orbital expressed as a linear combination of the basis functions. This can be viewed as a pseudo-eigenvalue equation, and solved iteratively by diagonalizing $\mathbf{F}$ [1]. This method, especially combined with some sort of acceleration scheme, such as Direct Inversion in the Iterative Subspace (DIIS) [36] [37], is very successful. However, as the diagonalization of the Fock matrix scales cubically with the system, this is not feasible in very large systems [1]. An alternative method, that sidesteps this problem, was developed by Helgaker et al. [32]. Here, the HF equations are solved by a direct optimization of the one-electron AO density matrix. An analogue to this scheme will be used in this thesis, in which an exponential parametrization is used to induce changes in the molecular orbitals. In the following subsections, the main points will be highlighted.

2.1.1 Exponential parametrization of the density matrix in the full MO basis

Equation 2.8 can be reformulated in terms of density matrices, and for a closed shell molecular orbital basis, the energy expression is given as

$$E_{\mathrm{HF}} = 2 \sum_{ij} h_{ij} D_{ij} + \sum_{ijkl} (2 g_{ijkl} - g_{ikjl}) D_{ij} D_{kl} + h_{nuc}. \tag{2.13}$$

By introducing the matrices $\mathbf{h}$ for the one-electron interactions and

$$[\mathbf{G}(\mathbf{D})]_{ij} = \sum_{kl}(2g_{ijkl} - g_{ikjl})D_{kl} \tag{2.14}$$

for the two-electron interactions, the energy expression becomes

$$E_{HF} = 2\mathrm{Tr}[\mathbf{hD}] + \mathrm{Tr}[\mathbf{DG}(\mathbf{D})] + h_{nuc}, \tag{2.15}$$

where the molecular orbital density matrix is defined as

$$\mathbf{D} = \begin{pmatrix} \mathbf{1}_{oo} & \mathbf{0}_{ov} \\ \mathbf{0}_{vo} & \mathbf{0}_{vv} \end{pmatrix}. \tag{2.16}$$

for a set of orthonormal MOs. For the density matrix to be a valid representation of a single Slater determinant wave function, it has to fulfill the following requirements for a closed shell system with N_e electrons

$$\mathbf{D} = \mathbf{D}^T, \tag{2.17}$$

$$\mathrm{Tr}[\mathbf{DS}] = \frac{1}{2}N_e, \tag{2.18}$$

$$\mathbf{D}^2 = \mathbf{D}, \tag{2.19}$$

where the first is called the symmetry condition, the second the trace condition and the last the idempotency condition [1].

Analogue to the scheme developed by Helgaker et. al, rotations among the molecular orbitals is introduced by defining $\mathbf{W}$ as

$$\mathbf{W} = \exp(-\boldsymbol{\kappa}), \tag{2.20}$$

where $\boldsymbol{\kappa}$ is a matrix that carries out rotations among the molecular orbitals. By keeping it anti-symmetric the metric will be conserved, meaning that if the starting orbitals are orthonormal, they will stay that way throughout the optimization process. The orbital transformation in equation 2.20 is used to parametrize changes on the density matrix, obtaining

$$\mathbf{D}(\boldsymbol{\kappa}) = \exp(-\boldsymbol{\kappa})\mathbf{D}\exp(\boldsymbol{\kappa}). \tag{2.21}$$

For some non-zero $\boldsymbol{\kappa}$ the density matrix does not change. These parameters are not needed for the optimization, and are therefore referred to as redundant. These parameters can be excluded from our optimization without doing harm, and keeping them can in fact lead to problems in the optimization due to singularities in the Hessian [1]. For a closed-shell wavefunction, these redundant rotations will be rotations within the virtual and occupied space. By constraining $\boldsymbol{\kappa}$ to have a block structure, where the only non-zero parameters are in the occupied-virtual blocks, the redundant rotations can be kept out of the optimization processes without the need of the projections used in the AO basis [32].

The parametrization in equation 2.21 does not break the symmetry, trace and idempotency properties of the MO density matrix [32], and thus the changed density matrix is still a valid representation. By using the symmetric *Baker-Campbell-Hausdorff* (BCH)-expansion [1], 2.21 becomes

$$\mathbf{D}(\boldsymbol{\kappa}) = \mathbf{D} + [\mathbf{D}, \boldsymbol{\kappa}] + \frac{1}{2!}[[\mathbf{D}, \boldsymbol{\kappa}], \boldsymbol{\kappa}] + ... \tag{2.22}$$

The energy in equation 2.15 can now be viewed as a function of the the orbital rotation parameter matrix, $\boldsymbol{\kappa}$, and by inserting the expression for the transformed MO denisty matrix and expand to second order around $\boldsymbol{\kappa} = 0$, the gradient and the linear transformation of the Hessian on a vector,

$$\boldsymbol{\sigma} = \mathbf{E}^{[2]}\boldsymbol{\kappa}, \tag{2.23}$$

is obtained, namely

$$\mathbf{E}^{[1]} = \mathbf{F}_{\mathrm{vo}} - \mathbf{F}_{\mathrm{ov}} \tag{2.24}$$

and

$$\boldsymbol{\sigma} = \boldsymbol{\kappa}(\mathbf{f}_{\mathrm{oo}} - \mathbf{f}_{\mathrm{vv}}) + (\mathbf{f}_{\mathrm{oo}} - \mathbf{f}_{\mathrm{vv}})\boldsymbol{\kappa} + \mathbf{G}_{\mathrm{ov}}([\mathbf{D}, \boldsymbol{\kappa}]) + \mathbf{G}_{\mathrm{vo}}([\mathbf{D}, \boldsymbol{\kappa}]) \tag{2.25}$$

respectively.

Here, both expressions have been divided by four, and the projector $\mathbf{PXQ}^T + \mathbf{QXP}^T$, $\mathbf{X} = \{\mathbf{E}^{[1]}, \boldsymbol{\sigma}\}$, previously used by Helgaker et al. [32], have been used to make sure that no redundant terms has entered the expressions. $\mathbf{P} = \mathbf{D}$ projects onto the occupied MO space, while $\mathbf{Q} = \mathbf{1} - \mathbf{D}$ projects onto the unoccupied space, and due to the simple structure of $\mathbf{D}$ and $\mathbf{1} - \mathbf{D}$ in the MO basis, these projections will simply extract the relevant block of the matrix. E.g., $\mathbf{F}_{\mathrm{oo}}$ is a matrix with the only non-zero elements in the occupied-occupied block, and similarly for the other matrices.

2.1.2 Multilevel Hartree-Fock formulation

As already mentioned, the MLHF method is the theme of an upcoming paper [33], but will here be explained in some detail as the approach is essential in this thesis.

In MLHF, the system is divided in to an active and an inactive part. The total density matrix is divided into an active and an inactive/rest part, $\mathbf{D}_a$ and $\mathbf{D}_r$, where both are required to separately fulfill the symmetry, trace and idempotency conditions. The total density matrix can then be written as

$$\mathbf{D} = \mathbf{D}_a + \mathbf{D}_r. \tag{2.26}$$

The structure of these matrices are illustrated in Figure 1, where the labels o_a and o_r indicates the occupied active and inactive space, respectively, and v_a and v_r the virtual active and inactive space.

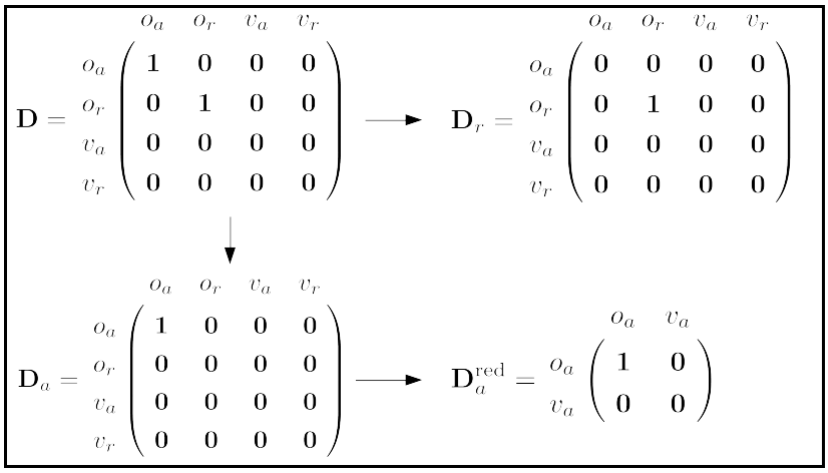


Figure 1: An illustration of the structure of the total, active and inactive MO denisty matrices, $\mathbf{D}, \mathbf{D}_a$ and $\mathbf{D}_r$, where the labels o_r and o_v indicates the occupied active and inactive space, respectively, and v_a and v_r the virtual active and inactive space. The structure of the active density matrix in the reduced dimension of the active MOs, $\mathbf{D}_a^{\mathrm{red}}$, is also illustrated.

The energy expression for the MLHF method is obtained by inserting equation 2.26 into the energy expression for the full system given in equation 2.15. The resulting energy expression then becomes

$$E = E(\mathbf{D}_a) + E(\mathbf{D}_r) + 2\mathrm{Tr}[\mathbf{D}_a\mathbf{G}(\mathbf{D}_r)] + h_{\mathrm{nuc}}, \tag{2.27}$$

$$E(\mathbf{D}_i) = 2\mathrm{Tr}[\mathbf{h}\mathbf{D}_i] + \mathrm{Tr}[\mathbf{D}_i\mathbf{G}(\mathbf{D}_i)]. \tag{2.28}$$

The energy can be viewed as a function of the two different densities. It will be optimized by imposing changes in $\mathbf{D}_a$ through the exponential parametrization of orbital rotation explained in section 2.1.1 for a fixed $\mathbf{D}_r$. This means that the only non-zero elements of $\boldsymbol{\kappa}$ will be in the active occupied-active virtual blocks. As a consequence of this, the block structure of the matrices will remain unchanged under the optimization process. It is therefore possible to write the terms that contain $\mathbf{D}_a$ in terms of a reduced set of dimensions, $N_{\mathrm{red}} = N_a \times N_a$, where N_a is the total number of active molecular orbitals, $N_a = o_a + v_a$. In theory the reduced matrices will contain contribution from the AOs of the entire system. This means that no information is lost by doing this reduction of dimensions. The structure of the MO density matrix in the reduced basis is illustrated in Figure 1.

The energy expression for the reduced dimensions will, in analogue with equation 2.27, take the form

$$E = E(\mathbf{D}_a^{\mathrm{red}}) + E(\mathbf{D}_r) + 2\mathrm{Tr}[\mathbf{D}_a^{\mathrm{red}}\mathbf{G}(\mathbf{D}_r)^{\mathrm{red}}]. \tag{2.29}$$

$E(\mathbf{D}_r)$ will be a constant and is needed if the goal is to obtain energy differences. On the other hand, if the purpose is to generate starting orbitals for a post-HF

calculation, this can be omitted. $\mathbf{G}(\mathbf{D}_r)^{\text{red}}$ will be computed only once, but needs to be kept through the iterations as it interacts with the active part through the third term.

By projecting onto the active occupied and active virtual space, where $\mathbf{P} = \mathbf{D}_a^{\text{red}}$ and $\mathbf{Q} = \mathbf{1} - \mathbf{D}_a^{\text{red}}$, we obtain the gradient

$$\mathbf{E}^{[1],\text{red}} = \mathbf{F}^{\text{red}}_{\text{eff vo}} - \mathbf{F}^{\text{red}}_{\text{eff ov}} \tag{2.30}$$

and the linear transformations,

$$\begin{aligned}\boldsymbol{\sigma}^{\text{red}} = (\mathbf{F}^{\text{red}}_{\text{eff oo}} - \mathbf{F}^{\text{red}}_{\text{eff vv}})\boldsymbol{\kappa}^{\text{red}} + \boldsymbol{\kappa}^{\text{red}}(\mathbf{F}^{\text{red}}_{\text{eff oo}} - \mathbf{F}^{\text{red}}_{\text{eff vv}}) \\ + \mathbf{G}^{\text{red}}_{\text{ov}}([\mathbf{D}_a^{\text{red}}, \boldsymbol{\kappa}^{\text{red}}]) + \mathbf{G}^{\text{red}}_{\text{vo}}([\mathbf{D}_a^{\text{red}}, \boldsymbol{\kappa}^{\text{red}}])\end{aligned} \tag{2.31}$$

were the effective Fock matrix,

$$\mathbf{F}^{\text{red}}_{\text{eff}} = \mathbf{F}^{\text{red}}(\mathbf{D}_a) + \mathbf{G}^{\text{red}}(\mathbf{D}_r), \tag{2.32}$$

have been used. By comparing the gradient given here with the gradient for the full system, 2.24, we see that they have the same structural form, but the gradient for the multilevel system have two additional terms stemming from the division of the system into an active and inactive part. Where the gradient for the full system will have contributions from all of the molecular orbitals in the two first terms, the multilevel system will only have contributions from the active molecular orbitals in $\mathbf{F}^{\text{red}}(\mathbf{D}_a)$. The $\mathbf{G}^{\text{red}}(\mathbf{D}_r)$-terms will therefore be a constant contribution from the inactive molecular orbitals. The same situation appears in the linear transformation of the Hessian, in which the difference between the full system, equation 2.25, and equation 2.31 also stems from the division of the full system into an active, changing part and a constant rest-part.

2.1.3 Start guess for the total density

It is important that the start guess for the orbital densities is reasonable for any type of calculation, but even more so for this type of an embedded scheme, as the rest-part of the matrix will not be optimized further. Here, the initial density will be found through the *superposition of atomic densities* (SAD) approach by Lenthe et al.[38]. The density matrix obtained is not idempotent, so to solve this, the Fock matrix is created by using the non-idempotent AO densities and then a diagonalization is performed. The resulting orbitals can then be used to construct an idempotent matrix. The diagonalization step makes a good starting density, as it comes from essentially having carried out one SCF iteration. This step is usually quite costly as it scales with n^3, however the SAD density is diagonal, meaning that there is sparsity in the Fock matrix which can be taken advantage of to reduce the cost.

2.1.4 Partition of the total density matrix

To partition the initial density matrix, Cholesky decomposition is used, previously used by Myhre et al.[35] in their MLCC scheme. This is not a black-box method,

as we have to specify which atoms that belong to the active part. The active density is given by decomposition of the diagonal elements of the initial density corresponding to active AOs. Not all of the active part is included. Only diagonal elements that are larger than a given threshold is selected. This provides active MOs, and an inactive AO density.

2.2 Basis sets and post-Hartree-Fock methods

There are four sources of error in ab initio molecular electronic structure theory: (1) neglect of incomplete treatment of electron correlation, (2) incompleteness of the basis set, (3) relativistic effects and (4) deviations from the Born-Oppenheimer approximation. For ground state molecules, and molecules without heavy atoms, the first two terms are the main sources of error [16].

As already mentioned, the molecular orbitals are commonly expressed as a linear combination of atomic orbitals, as can be seen in equation 2.11. These are referred to as the basis set. As a basis set consists of a finite number of functions, it is not complete and will introduce errors in the wave function. This is called the basis set truncation error. The larger the basis set, the smaller the error, but also the more computationally demanding the calculation becomes [2]. Two common choices to represent the radial one-electron functions in quantum chemistry is Slater type orbitals, STO's, or Gaussian type orbitals, GTO's. The Slater type orbitals has the advantage that they have a physical interpretation, as they give the correct radial distributions, but there is no analytically solution to the two electron integral when this representation is used. The GTO's on the other hand, do not display the correct radial behaviour as they give no cusp at the nucleus and falls off to quickly. This means that more GTO's are needed to display the correct behaviour in comparison to the STO's. However, this is more than made up for by the faster evolution of the molecular integrals, and these functions are almost universally preferred [39].

A much used Gaussian basis set in quantum chemistry is the basis set developed by Dunning [40] to be used in correlated calculations. Here a correlating orbital is represented by a single primitive Gaussian, chosen to maximize its contribution to the correlation. The notation of these basis sets are e.g. cc-pVDZ, cc-pVTZ etc., where cc-p stands for correlation-consistent polarized meaning that, at each stage, functions that contributes the same to the correlation energy are included, and D, T means respectively that the double or triple the amount of the number of basis functions in the minimal basis are included. The V denotes that they are a valence-only basis set. Augmented functions can be added, which are indicated by adding aug- to the beginning of the notation. The basis set is then augmented with one extra function that has a smaller exponent for each angular momentum to better describe the electron density far away from the nucleus. Finally, if core-excitations are to be calculated, it is possible to add functions with large exponents, known as tight functions. The acronym will then be cc-pCVXZ [39].

If a HF calculation were done with a complete basis set, the HF limit is obtained, which is not the exact ground-state energy of the molecule. This is because HF neglects the electron correlation effects, often called dynamical correlation, which is instantaneous Coulombic interactions between electrons, as its basis assumption is that the electrons behaves as independent particles, and thus do not interact. This motivates the definition of electron correlation energy,

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}. \tag{2.33}$$

The HF wave function is, as already mentioned, the simplest wave function in the ab initio hierarchy of wave functions, but it is the starting point of other methods, which are trying to recover the lost correlation energy from HF. In the orbital picture, the correlated motion of the electrons manifest themselves as excitations from occupied orbitals into virtual orbitals. The first method that applies this strategy to obtain the correlation energy is CI, in which the wave function is constructed as a linear combination of Slater determinants, Φ,

$$\Psi_{\text{CI}} = \sum_I C_I \Phi_I \tag{2.34}$$

The expansion coefficients C_I are found by variational optimization, whereas the spin orbitals are generated from a preceding HF calculation and held fixed during the optimization. Ψ_{CI} can be written as excitation from the HF reference determinant, i.e.

$$\Psi_{\text{CI}} = c_0\Phi_0 + \sum_{ia} c_a^i \Phi_a^i + \sum_{a<b,i<j} c_{ab}^{ij}\Phi_{ab}^{ij} + \sum_{i<j<k,a<b<c} c_{abc}^{ijk}\Phi_{abc}^{ijk} + \dots, \tag{2.35}$$

where Φ_0 is the HF reference determinant, Φ_a^i is the singly excited Slater determinant arising from promoting a single electron from spin orbital a to spin orbital i and so on. In FCI the wave function is a linear combination of all possible Slater determinants, and the exact wave function is obtained [2]. Since FCI is only applicable for the simplest systems, the method has to be truncated so that only a small subset of the full set of determinants are considered. The most common is *CID*, configuration-interaction doubles, which only includes double excitations, and configuration-interaction singles and doubles, *CISD*, in which single excitations are also included through interactions via the doubles. *CIS*, only singles, will on the other hand not cause an improvement over the HF state. The reason for this is Brillouin's theorem,

$$\langle \Phi_0 | \hat{H} E_{ai} | \Phi_o \rangle = 0, \tag{2.36}$$

which states that the closed-shell HF state will not interact with singly excited states [1].

Even though the truncated CI method recovers much of the lost correlation energy, there is a problem with size-extensivity, meaning that it does not yield the same energy for two infinitely separated systems and the sum of the energy computed for the two systems in separate calculations [1]. A method that overcomes this problem is the CC method.

2.3 Coupled cluster theory

CC theory was first introduced in quantum chemistry by Čížek[41] and Paldus et al.[42]. Here, the wavefunction is obtained through the exponential ansatz,

$$\Psi_{\mathrm{CC}} = \exp(\hat{T})|\Phi_0\rangle, \tag{2.37}$$

in which Φ_0 traditionally is the reference HF determinant. The cluster operator is a linear combination of excitation operators, and may be arranged into classes consisting of all single excitations, all double excitations and so on. This can be expressed as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + ...\hat{T}_N, \tag{2.38}$$

and introduces a hierarchy of approximations[1]. By truncating CC after the first operator, $\hat{T}_1$, we obtain CC singles, CCS, by truncating after the second operator, $\hat{T}_2$, we obtain CCSD, and so on. Without truncation, the FCI wave function is obtained, and one would prefer the linear CI method over the non linear CC. However, as already mentioned, this is not feasible. Truncation is necessary, and it is in these cases where the advantages of this non-linear parametrization becomes apparent. By focusing our attention on the CCSD wavefunction,

$$\Psi_{\mathrm{CCSD}} = \exp(\hat{T}_1 + \hat{T}_2)|\Phi_0\rangle, \tag{2.39}$$

we will see the advantages truncated CC has over truncated CI in more detail. In the second quantization formalism, the excitation operators $\hat{T}_1$ and $\hat{T}_2$ is defined as

$$\hat{T}_1 = \sum_{ai} t_i^a E_{ai} \tag{2.40}$$

and

$$\hat{T}_2 = \sum_{a>b,i>j} t_{ij}^{ab} E_{ai} E_{bj}. \tag{2.41}$$

where the single excitation operator from equation 2.3 is used. By expanding the exponential operator in equation 2.39,

$$\exp(\hat{T}_1 + \hat{T}_2) = 1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_1\hat{T}_2 + \frac{1}{2}\hat{T}_2^2 + \cdots, \tag{2.42}$$

the advantages of the exponential operator can be seen. Even though the cluster operator is truncated at the double excitation level, higher order excitations are still included as a result of the independent interactions between the included cluster operators (so called *disconnected clusters*). This means that at the same truncation level, CC will recover more of the lost correlation energy than the equivalent truncation in CI as it samples from all of the determinants in the FCI state. This gives rise to an observation of the difference between the CI weights, c, in equation 2.35, and the CC *amplitudes*, t, in equation 2.40. Whereas the CI weight represent how important a determinant is in the FCI wavefunction, the CC amplitudes can be interpreted as how important an excitation is. Another advantage truncated CC has over truncated CI is, as already mentioned, size extensivity.

However, the CC method is not variational. Due to the non linear form of the wave function, the usual procedure with the variational principle becomes too complicated. Instead, the wavefunction and amplitudes are found through projecting the similarity transformed Schrödinger equation onto the reference state, $\langle\Phi_0|$, and the excitation manifolds,

$$\langle\mu| = \langle\Phi_0|\hat{\tau}_\mu, \tag{2.43}$$

where $\hat{\tau}_\mu$ is an excitation operator. As a result, the truncated CC energy is not a upper bound to the exact energy. However, as the CC energy is a rather accurate, small correction to the reference state's electronic energy, which is usually HF and variationally determined, and in addition approaches the variational FCI to higher order excitation, this does not matter that much in practice [1].

The equations to be solved are thereby

$$E_{CC} = \langle\Phi_0|\exp(-\hat{T})\hat{H}\exp(\hat{T})|\Phi_0\rangle, \tag{2.44}$$

and

$$\Omega_{\mu_i} = \langle\mu_i|\exp(-\hat{T})\hat{H}\exp(\hat{T})|\Phi_0\rangle = 0, \tag{2.45}$$

where μ_i represent the different excitation manifold. In the CCSD approximation, $i = 1, 2$ and one approach is to use the T_1- transformed Hamiltonian, $\tilde{\hat{H}} = \exp(-\hat{T}_1)\hat{H}\exp(T_1)$, to simplify the complexity of the CCSD equation to that of the CCD equations. The cost of this is reduced symmetry [43].

Equation 2.44 can be further simplified by noting that

$$\langle\Phi_0|\exp(-\hat{T}) = \langle\Phi_0| \tag{2.46}$$

and recognizing that the Hamiltonian is a two particle operator, meaning that any cluster operator higher than doubles will not contribute (directly) to the energy. The resulting energy is thereby

$$\langle\Phi_0|\hat{H}(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2)|\Phi_0\rangle = E_{CC}. \tag{2.47}$$

Also, if the reference state, $|\Phi_0\rangle$, is the closed shell HF state, the contribution from the pure $\hat{T}_1$ is zero due to Brillouin's theorem, as given in equation 2.36.

The fact that only $\hat{T}_1$ and $\hat{T}_2$ contribute directly to the energy does of course not mean that CCSD is exact, as the higher order excitation will contribute indirectly through the coupling of the amplitudes in the projected equations. However, the CCSD method does reduce the error in both molecular properties and the electronic energy by a factor of typically three or four relative to the HF description [5]. This is a significant improvement, and stems from the fact that at most two electrons with opposite spins may coincide in space, meaning that the dominant correlating motion is captured in the pair clusters. As already mentioned, CCS will, as CIS,

not impose an improvment over HF in terms of total energy calculations if the reference state is a HF state [2]. However, they do contribute through interactions with the doubles, and more importantly, they represent a relaxation of the spin orbitals that happens as a consequence of the many-electron excitations that are now being included. This is important when obtaining molecular properties [1]. The CCSD wavefunction also does a good job in capturing most of the quadruple excitation effects in the case of a good reference state, but the triple effect is not represented well by disconnected clusters. In the cases where this description is needed, the triplet's effects has to be included in some manner, either approximate or exact [1].

2.4 Excitation energies in Coupled Cluster

CC linear response theory (LR) [44] [45] offers a way to calculate excitations energy within the CC approach. The projected Schrödinger equation, equation 2.44, now has to be time-dependent. Excitation energies and left and right excitation vectors (originating from the fact that the similarity transformed Hamiltonian is non-Hermitian [46]) are obtained through the asymmetric eigenvalue equations

$$\begin{aligned} \mathbf{A}\mathbf{R}_k &= \omega_k \mathbf{R}_k \\ \mathbf{L}_k\mathbf{A} &= \omega_k \mathrm{L}_k \end{aligned} \tag{2.48}$$

under the biorthonormality condition $\mathbf{L}_j\mathbf{R}_k = \delta_{rk}$. $\mathbf{A}$ is the Jacobian matrix, defined as

$$\frac{\partial \Omega_\mu}{t_v} = \langle \mu | \exp(-\hat{T})[\hat{H}, \hat{\tau}_v] \exp(\hat{T}) | \Phi_0 \rangle. \tag{2.49}$$

In most cases, equation 2.48 is solved via some scheme based on the iterative Davidson algorithm [47]. However, this approach is not feasible if core excitations is to be evaluated, as the traditional eigenvalue solvers apply a bottom-up approach, meaning that the valence excitations, which are much lower in energy, are obtained before the core excitations are targeted. In an approach suggested by Coriani et. al [48], the so-called core-valence separation (CVS) is used to to effectively reduce the excitation space by using a projector that removes all vector elements not referring to at least one, or a predetermined set, of core orbitals. This approach have been implemented in DALTON [34], and a perturbation correction is also added that incorporates the effect of the excluded part of the excitation space [48].

3 Method

The goal of this thesis to connect multilevel HF and CCSD, both described in the theory section. These methods are already implemented, but in two different programs, CCSD in DALTON [34], and MLHF in LSDALTON [34]. In this section, the main points of the fusion of these to procedures is explained. This includes the division on the system into a SAD, HF and CC active part, the construction of the Fock matrix and the working equations in the CCSD calculation.

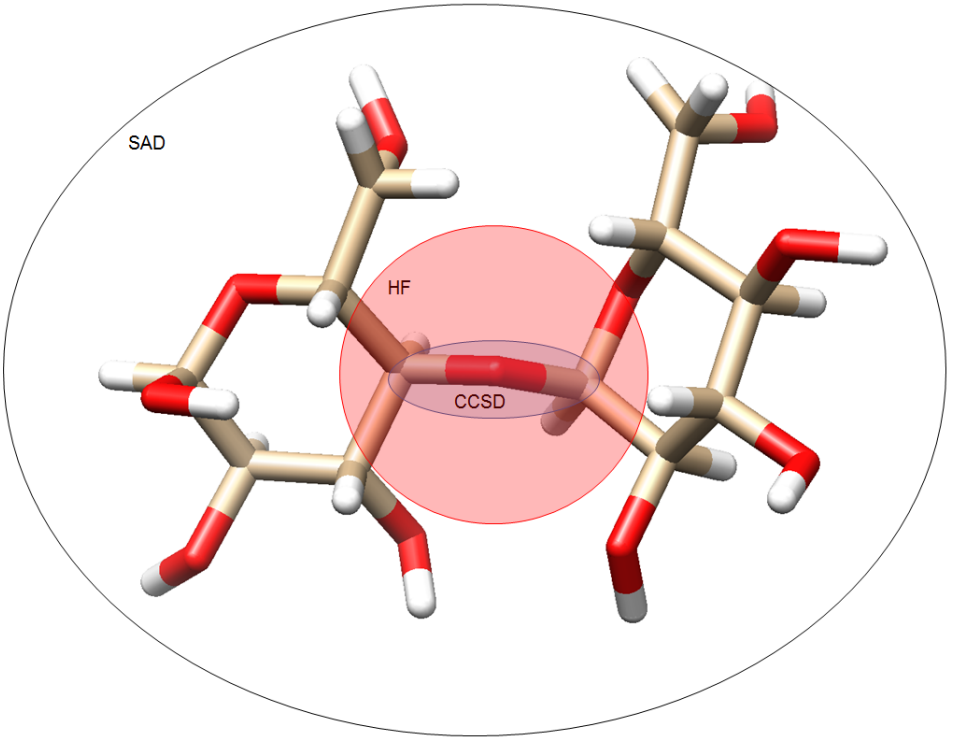


Figure 2: An illustration of the division of a amylose molecule into a SAD, HF and CCSD part.

3.1 Preparation for enabling CCSD in MLHF

In this subsection, the steps taken to prepare for the CCSD calculation in the MLHF scheme is highlighted.

3.1.1 Terminology for the division of the system

An illustrative picture of the division of a system is shown in Figure 2. Before a calculation on a specific system is carried out, some background knowledge of the system is needed to select the active atoms. This is the region in which you want to carry out the coupled cluster calculation. These will hereby be referred to as the "CC-active" atoms, from which CC-active MOs will later be generated. They will represent the active space in the CCSD calculation. In the MLHF calculation, a distance criteria will be used to define the "HF-active" atoms. The atoms (and thus orbitals from decomposition) active in the MLHF calculation will therefore be the CC-orbitals and the additional HF-active orbitals. For simplicity in later equations, we denote HF-active orbitals as orbitals optimized in MLHF but not included in the CC calculation, whereas CC-orbitals are optimized in MLHF and used in the CC calculation. When referring to the "active MO" basis or "MLHF active MOs" we refer to the union of the CC and HF active spaces.

3.1.2 The Fock matrix

After the MLHF density is optimized, the effective Fock matrix in the active MO basis, $\mathbf{F}_{\text{eff}} = \mathbf{F}(\mathbf{D}_a) + \mathbf{G}(\mathbf{D}_r)$, will have vanishing elements in the occupied-virtual and virtual-occupied blocks, following from equation 2.30. The superscript "red" is dropped for simplicity as we throughout the sections only work in the active MO space. The effective Fock matrix obtained in the last iteration is transformed from active MO basis to the AO basis as

$$\mathbf{F}^{\text{AO}} = \mathbf{S}\mathbf{C}\mathbf{F}_{\text{eff}}\mathbf{C}^{\text{T}}\mathbf{S}. \tag{3.1}$$

An occupied density matrix, $\mathbf{C}_{\text{occ}}\mathbf{C}_{\text{occ}}^{\text{T}}$, and a virtual density matrix, $\mathbf{C}_{\text{virt}}\mathbf{C}_{\text{virt}}^{\text{T}}$ are generated from the active MO space and a Cholesky decomposition is performed on both density matrices. This results in a new MO coefficient matrix, $\mathbf{C}_{\text{CC-active}}$, where occupied orbitals are generated from the occupied density and virtual orbitals from the virtual density. When these are used to transform the effective Fock matrix from the AO basis to the CC-active MO basis, the resulting Fock matrix will have the same block-diagonal structure, as illustrated in Figure 3.

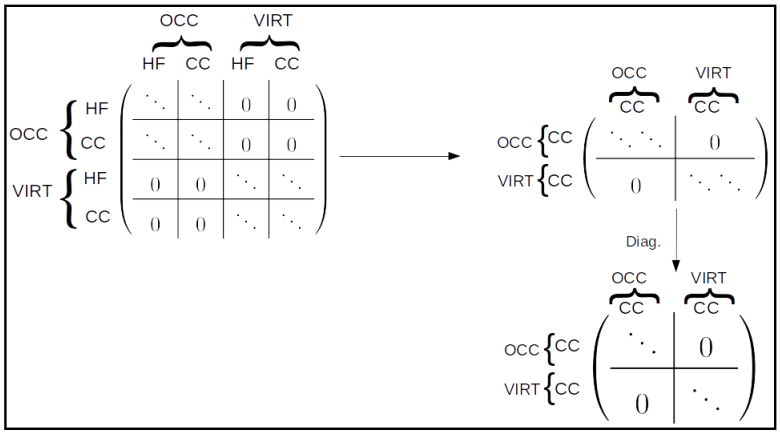


Figure 3: An illustration of the MO Fock matrices in the active basis MO basis ($\mathrm{HF}_{\mathrm{active}}$ + $\mathrm{CC}_{\mathrm{active}}$) and in the CC-active MO basis. The Fock matrix in the CC MO basis is then diagonalized.

The reason for this is that the Cholesky decomposition generates MOs which will be related to the canonical MOs by an orthogonal transformation and thus satisfy the variational condition [49]. The matrix is then diagonalized, and we obtain a set of canonical MOs, $\mathbf{C}_{\mathrm{can}}$, and the orbital energies. The canonical MOs and their orbital energies are written to file and will be read in DALTON.

3.1.3 Obtaining the embedding terms

In the CCSD calculation, there will be two embedding orbital spaces. As the Fock matrix is built from occupied orbitals, only that part is needed. First we find $\mathbf{D}^{\mathrm{HF}}_{\mathrm{occ}}$ by subtracting the CC-active occupied density matrix obtained from the Cholesky decomposition from the active occupied density matrix (which is contains both the HF-active and the CC-active orbitals),

$$\mathbf{D}^{\mathrm{HF}}_{\mathrm{occ}} = \mathbf{D}_{\mathrm{occ}} - \mathbf{D}^{\mathrm{CC}}_{\mathrm{occ}}. \tag{3.2}$$

A subroutine in LSDALTON is then called to calculate the two electron repulsion integrals and build $\mathbf{G}(\mathbf{D}^{\mathrm{HF}})$. The sum of $\mathbf{G}(\mathbf{D}_r)$ and $\mathbf{G}(\mathbf{D}^{\mathrm{HF}})$ is written to a file.

An illustration of the optimization process in MLHF is shown Figure 4.

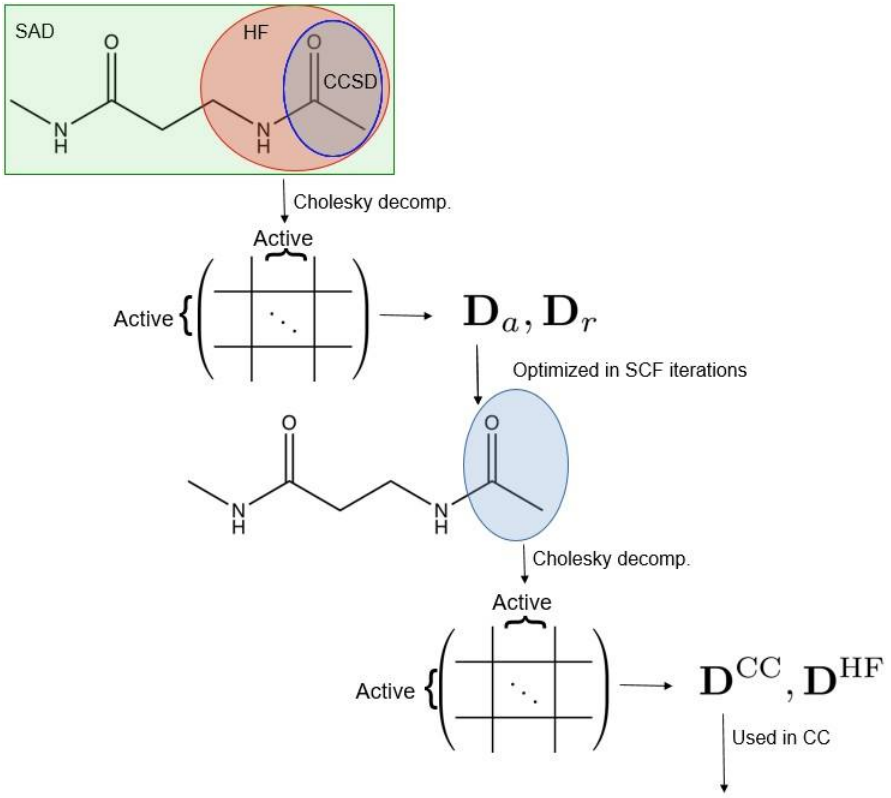


Figure 4: An illustration of the different steps in the MLHF calculation.

3.2 CCSD PROCEDURE

In this subsection, the modified working equations in CCSD will be presented, and some main points of the implementation is given.

3.2.1 Working equations

The systems wavefunction, $|R\rangle$ is part Hartree-Fock, part non-Hartree Fock,

$$|R\rangle = \prod_{i=1}^{\mathrm{O_{CC}}} a_{i\alpha}^{\dagger} a_{i\beta}^{\dagger} \prod_{I=1}^{\mathrm{O_{HF}+O_r}} a_{I\alpha}^{\dagger} a_{I\beta}^{\dagger} |vac\rangle. \tag{3.3}$$

Here, lowercase letters indicates CC-active orbitals, whereas capital letters indicates CC-inactive orbitals, in which some orbitals are Hartree-Fock optimized, and the rest is only treated by SAD start guess. Otherwise, the notation is as usual, i,j,k denotes occupied orbitals, a,b,c denotes virtual orbitals and p,q,r general orbitals, and similarly for capital letters. All orbitals, both the active and inactive, are orthogonal to each other.

The cluster operator, $\hat{T}$, is defined so that it only works within the CC-active orbital space. For CCSD we thus have $\hat{T} = \hat{T}_1 + \hat{T}_2$ as,

$$\hat{T} = \sum_{ai} t_i^a E_{ai} + \frac{1}{2} \sum_{ab,ij} t_{ij}^{ab} E_{ai} E_{bj}. \tag{3.4}$$

By using a BCH-expansion of the energy expression given in equation 2.44, we obtain the energy as,

$$\langle R|\hat{H}|R\rangle + \langle R|[\hat{H},\hat{T}]|R\rangle + \frac{1}{2}\langle R|[[\hat{H},\hat{T}],\hat{T}]|R\rangle = E_{\mathrm{R}}. \tag{3.5}$$

The Hamiltonian can be divided into three, $\hat{H}_{\mathrm{act}}$, where all the indicies are CC-active, $\hat{H}_{\mathrm{inact}}$, where all the indicies are CC-inactive, and $\hat{H}_{\mathrm{mix}}$, where the indicies are mixed, meaning that the operators refer to both CC-active and inactive orbitals.

$$\hat{H} = \hat{H}_{\mathrm{act}} + \hat{H}_{\mathrm{inact}} + \hat{H}_{\mathrm{mix}} + h_{\mathrm{nuc}} \tag{3.6}$$

The active Hamiltonian will give rise to the standard CC equations. The inactive Hamiltonian will only contribute to the first term in equation 3.5, which we identify as the MLHF energy, E_{MLHF}, as it commutes with the cluster operator. This means that the only non-standard contribution to the CC energy can come from the mix terms,

$$\begin{aligned}\hat{H}_{\mathrm{mix}} = &\frac{1}{2}\sum_{pqRS} g_{pqRS} e_{pqRS} + \frac{1}{2}\sum_{PQrs} g_{PQrs} e_{PQrs} + \frac{1}{2}\sum_{pQRs} g_{pQRs} e_{pQRs} \\ &+ \frac{1}{2}\sum_{PqrS} g_{PqrS} e_{PqrS} + \frac{1}{2}\sum_{PqRs} g_{PqRs} e_{PqRs} + \frac{1}{2}\sum_{pQrS} g_{pQrS} e_{pQrS}.\end{aligned} \tag{3.7}$$

Here, only the operators with even numbers of CC-active and inactive annhilation/creation operators are included. The reason for this is that, as the cluster operator only excites CC-active orbitals, all operators with odd numbers of CC-inactive excitations will end up being zero as a consequence of the orthogonality of the states. I.e, $\hat{H}_{\text{mix}}$ is effectively a one-electron operator in the CC-active space.

By expanding the two electron excitation operator, e_{pqrs}, in terms of the annihilation and creation operators, as shown in equation 2.4, and using the standard commutation relation between these operators and the permutation symmetry of the two electron integrals [1], equation 3.7 becomes

$$\begin{aligned}\hat{H}_{\text{mix}} = &\sum_{pqRS} g_{pqRS} e_{pqRS} + \sum_{PqrS} g_{PqrS} e_{PqrS} \\ &+ \frac{1}{2} \sum_{PqRs} g_{PqRs} e_{PqRs} + \frac{1}{2} \sum_{pQrS} g_{pQrS} e_{pQrS}.\end{aligned} \tag{3.8}$$

By evaluating the terms in equation 3.5 which only contain $\hat{H}_{\text{mix}}$, we obtain

$$\langle R|[\hat{H}_{\text{mix}}, \hat{T}_1]|R\rangle + \langle R|[\hat{H}_{\text{mix}}, \hat{T}_2]|R\rangle + \frac{1}{2}\langle R|[[\hat{H}_{\text{mix}}, \hat{T}_1], \hat{T}_1]|R\rangle = E_{\text{R}_{\text{mix}}} \tag{3.9}$$

The only non-zero contribution from this will be from the first term, since $\hat{H}_{\text{mix}}$ is a one-electron operator in the CC-active space. We obtain,

$$\begin{aligned}\langle R|[\hat{H}_{\text{mix}}, \hat{T}_1]|R\rangle = &\sum_{ai} \sum_{pqRS} t_i^a g_{pqRS} \langle R|[e_{pqRS}, E_{ai}]|R\rangle \\ &+ \sum_{ai} \sum_{PqrS} t_i^a g_{PqrS} \langle R|[e_{PqrS}, E_{ai}]|R\rangle \\ &+ \frac{1}{2} \sum_{ai} \sum_{PqRs} t_i^a g_{PqRs} \langle R|[e_{PqRs}, E_{ai}]|R\rangle \\ &+ \frac{1}{2} \sum_{ai} \sum_{pQrS} t_i^a g_{pQrS} \langle R|[e_{pQrS}, E_{ai}]|R\rangle,\end{aligned} \tag{3.10}$$

where the definition of $\hat{T}_1$ from equation 3.4 have been used. These terms are then evaluated by using the standard commutation expressions for the singlet and two electron excitation operator [1] and manipulated by using the commutators and anticommutators of the creation and annihilation operators. We obtain

$$\begin{aligned}\langle R|[\hat{H}_{\text{mix}}, \hat{T}_1]|R\rangle &= 2 \sum_{\mu\nu\rho\sigma} \sum_{I} \sum_{ai} t_i^a (2 g_{\rho\sigma\mu\nu} - g_{\rho\nu\mu\sigma}) C_{\mu I} C_{\nu I} C_{\rho i} C_{\sigma a} \\ &= 2 \sum_{ai} t_i^a G_{ia}(\mathbf{D}^{\text{I}}).\end{aligned} \tag{3.11}$$

where the definition of the two-electron integral matrix from equation 2.14 have been used. $\mathbf{D}^{\text{I}}$ is the CC-inactive density matrix which consists of two terms,

$$\mathbf{D}^{\text{I}} = \mathbf{D}^{\text{HF}} + \mathbf{D}_r \tag{3.12}$$

where $\mathbf{D}^{\mathrm{HF}}$ is the density of the HF-active orbitals and $\mathbf{D}_r$ is the inactive HF density from equation 2.26. I.e,

$$\langle R|[\hat{H}_{\mathrm{mix}}, \hat{T}_1]|R\rangle = 2\sum_{ai} t_i^a(\mathrm{G}_{ia}(\mathbf{D}^{\mathrm{HF}}) + \mathrm{G}_{ia}(\mathbf{D}_r) \tag{3.13}$$

$\hat{H}_{\mathrm{act}}$ will give rise to the standard terms,

$$\langle R|[\hat{H}_{\mathrm{act}}, \hat{T}_1]|R\rangle = 2\sum_{ai} t_i^a F_{ia}, \tag{3.14}$$

$$\langle R|[\hat{H}_{\mathrm{act}}, \hat{T}_2|R\rangle + \frac{1}{2}\langle R|[[\hat{H}_{\mathrm{act}}, \hat{T}_1], \hat{T}_1]|R\rangle = \sum_{aibj}(t_{ij}^{ab} + t_i^a t_j^b)(2g_{iajb} - g_{ibja}), \tag{3.15}$$

where

$$\mathrm{F}_{ia} = \mathrm{h}_{ia} + \mathrm{G}_{ia}(\mathbf{D}^{\mathrm{CC}}). \tag{3.16}$$

Equation 3.14 will no longer be zero due to Brillouin's theorem as the reference state is not a pure Hartree-Fock state. However, the combination of the term in equation 3.14 with the extra contribution stemming from the mixed Hamiltonian will be, since

$$\begin{aligned}\langle R|[\hat{H}_{\mathrm{act}} + \hat{H}_{\mathrm{mix}}, \hat{T}_1]|R\rangle &= 2\sum_{ai} t_i^a F_{ia} + 2\sum_{ai} t_i^a G_{ia}(\mathbf{D}^{\mathrm{I}}) \\ &= 2\sum_{ia} t_i^a(\mathrm{h}_{ia} + \mathrm{G}_{ia}(\mathbf{D}^{\mathrm{CC}}) + \mathrm{G}_{ia}(\mathbf{D}^{\mathrm{HF}}) + \mathrm{G}_{ia}(\mathbf{D}_r))\end{aligned} \tag{3.17}$$

where we recognize the active MLHF Fock matrix in the CC-active basis as,

$$\mathrm{F}_{ia}^{\mathrm{HF}}(\mathbf{D}_a) = \mathrm{h}_{ia} + \mathrm{G}_{ia}(\mathbf{D}_a) = \mathrm{h}_{ia} + \mathrm{G}_{ia}(\mathbf{D}^{\mathrm{CC}}) + \mathrm{G}_{ia}(\mathbf{D}^{\mathrm{HF}}), \tag{3.18}$$

where we have introduced the superscript "HF" to distinguish $\mathrm{F}_{ia}^{\mathrm{HF}}$ from the active CC Fock matrix, F_{ia}, in equation 3.16. We use this to reformulate equation 3.17 and obtain

$$\langle R|[\hat{H}_{\mathrm{act}} + \hat{H}_{\mathrm{mix}}, \hat{T}_1]|R\rangle = 2\sum_{ia} t_i^a\{\mathrm{F}_{ia}^{\mathrm{HF}}(\mathbf{D}_a) + \mathrm{G}_{ia}(\mathbf{D}_r)\}, \tag{3.19}$$

where we recognizes this as the effective Fock matrix of MLHF, equation 2.32, in the basis of the CC active orbitals. From the optimization criterion in MLHF, we know that these elements are zero by the gradient, given in equation 2.30. Thus, there are no extra terms in the CCSD energy, despite not having a fully optimized HF wave function.

Similar manipulations as the ones shown above were done on the amplitude equations, and the results were that the same modified "effective" Fock matrix,

$$\mathbf{F}^{\mathrm{eff}}(\mathbf{D}) = \mathbf{F}(\mathbf{D}^{\mathrm{CC}}) + \mathbf{G}(\mathbf{D}^{\mathrm{HF}}) + \mathbf{G}(\mathbf{D}_r). \tag{3.20}$$

enters the amplitude equations in place of the standard Fock matrix. The first term on the right will be the active term, optimized with the CCSD calculations, and the two last will be a constant interaction which is not further optimized. These are read from file in the places were the Fock matrix is constructed in the DALTON calculations, and added to the one-electron term as they stem from the effective one-electron operator, $\hat{H}_{\mathrm{mix}}$.

4 Proof of concept calculations

In this section some proof of concept calculations are presented. The MLHF code in LSDALTON optimizes the density matrix in the reduced dimensions of the active MOs, but the DALTON CC code runs partly in the AO basis. This constrains the size of the system tested. As this method is developed to be used on large systems, a CCSD code that offers to work in the MO basis should be used to test the method further. However, the goal of this thesis is to investigate the effect of not having a fully optimized HF wavefunction for core excitation energies in a reduced CC space. The small systems considered will be able to give some preliminary results on how sensitive the CCSD space is to its surrounding wavefunction. The core excitation energies are chosen because they are local, and will thus offer a suitable starting point for testing the CC in MLHF method. Note that the objective of these calculations are not to evaluate the performance for a given CCSD space to full CCSD, or evaluate the errors of MLCC. Ref. [35, 50] discusses the quality of MLCC excitation energies. The calculations will be evaluated from the criteria that the CCSD results for a given space should stay approximately the same for a full HF and MLHF.

4.1 Computational details

All calculations are done with a aug-cc-pCVDZ basis set. The molecule images in this section were obtained by using the Chimera package [51]. The molecules tested are

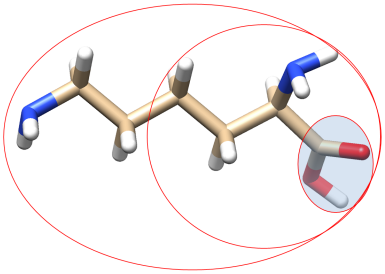
- Lysine, $C_6H_{14}N_2O_2$
- β-dipeptide, $C_6H_{12}N_2O_2$

4.2 Lysine

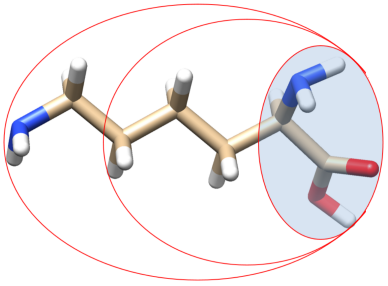
Here, some proof of concept calculations on the lysine molecule are presented. The excitation energy from the core orbital of the -OH oxygen and nitrogen within the active spaces are calculated, and the effect of changing the number of optimized HF orbitals within a given CC-active space is investigated.

4.2.1 Oxygen

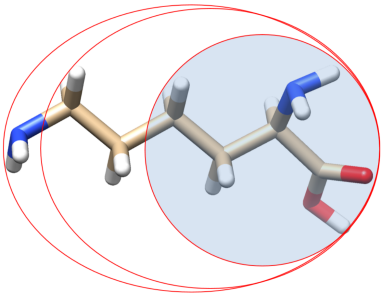
The three CCSD active spaces for lysine are highlighted in blue in Figure 5a - c. For each CCSD space, three HF optimized spaces are tested, the smallest being the CCSD space itself and the largest the full molecule.



(a) Lysine system A with four CC active atoms and three different HF layers. The CC-space is marked in blue and the HF space in red.

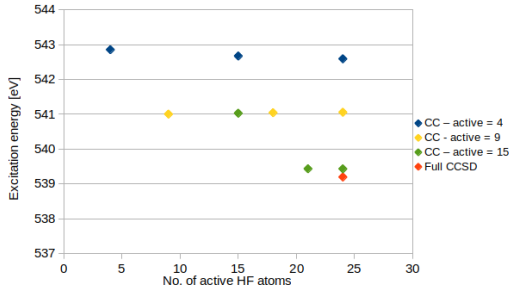


(b) Lysine system B with nine CC active atoms and three different HF layers. The CC-space is marked in blue and the HF space in red.

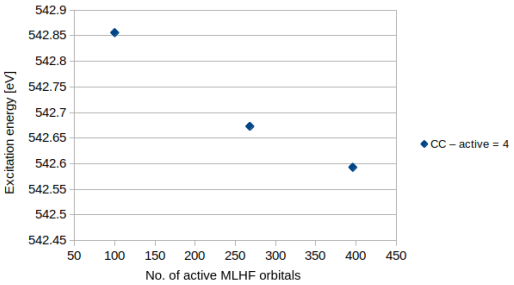


(c) Lysine system C with fifteen CC active atoms and three different HF layers. The CC-space is marked in blue and the HF space in red.

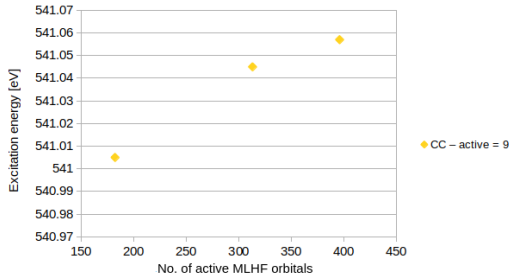
Figure 5: The three different lysine systems in the -OH oxygen core excitation calculation.



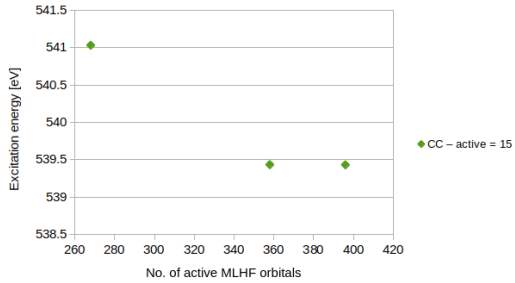
(a) Excitation energies for the different lysine systems, A, B and C displayed in Figure 5 as a function of the increasing HF-layer. A full CCSD calculation is also included.



(b) The excitation energies for lysine system A, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.



(c) The excitation energies for lysine system B, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.



(d) The excitation energies for lysine system C, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.

Figure 6: The excitation energies for the three different lysine systems A, B and C.

Figure 6 gives an overview of the excitation energy calculated plotted as a function of the number of active HF atoms within a given CC-active space. Figure 6b-6d takes a closer look at the individual CC spaces, and display the excitation energy as a function of the number of active MLHF orbitals. All information from the three systems, and also a full CCSD calculation computed for comparison, can be found in Table 1. In Figure 7 the number of CC-active orbitals are plotted against the active MLHF orbitals with the excitation energy as a superscript.

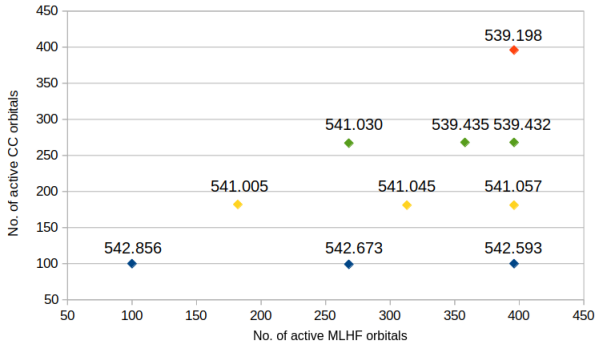


Figure 7: The number of CC-active orbitals plotted against the number of MLHF-active orbitals with the excitation energies for each given calculation.

From Figure 6b, we see that there are large differences in the excitation energy for the different number of MLHF orbitals optimized for the small CC room of only 100 optimized CC orbitals. The energy difference is 0.183 eV between the calculation with 100 and 268 optimized MLHF orbitals, and 0.080 eV between the calculation of 268 and 396 optimized MLHF orbitals, meaning that there is a substantial error, and that the excitation energy has not converged with respect to changes in the MLHF space. For the medium CC space there are 180/181 optimized CC orbitals. The difference in the number of CC-active orbitals most likely originates from the Cholesky decomposition and threshold value, meaning that there is an orbital just at the threshold limit. The differences are 0.04 eV and 0.012 eV between the layers of 182 and 313, and 313 and 396 optimized MLHF orbitals respectively. This implies reduction of 83 optimized MLHF orbitals (which is 20.96 % of the total system), for a error of 0.012 eV. This is a substantial reduction of the optimized orbitals, and a promising result, even though the error affects the excitation energy. In MLCC for core excitation energies, the errors in the CCSD/CC3 calculations are on the orders of 0.01 - 0.1 eV approximately [50]. Thus, the acceptable error for using MLHF rather than HF environment should be less than this to not compromise the accuracy of the CC calculation. It is worth noting that the excitation energy for this subsystem seems to converge to a different value than the excitation energy for the full CCSD system, as can be seen in Table 1, but the source of this error is more likely introduced by the reduced CC room, and not the partly optimized HF wave function.

From Figure 6d we see that there is an outlier with an excitation energy of 541.03 eV. The reason for this is that the excitation is different than the one targeted, as has been seen by examining the cluster amplitudes in DALTON. This is a common problem in CC calculations of excitation from low energy states, and a solution to this could be to calculate the two lowest excitation energies.

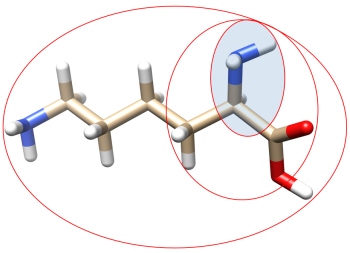
The difference between the two remaining states of 358 and 396 optimized MLHF orbitals is 0.003 eV in the CC optimized space of 267/268 orbitals. This error is small, however, as the difference in optimized orbitals is only 38, this might not be a true convergence with respect to the MLHF space.

Table 1: Excitation energies for the different lysine systems, A, B and C displayed in Figure 5 with the number of CC-active and MLHF-active occupied and virtual orbitals. The absolute error relative to the full HF calculation within the given CC space is calculated, but the outlier for system C is excluded. The full CCSD excitation energy is included for comparison.

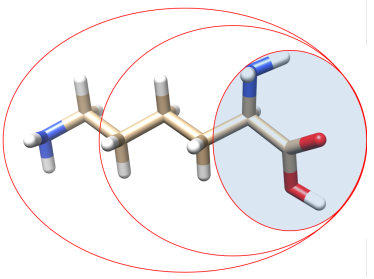
	# $\text{MLHF}^{\text{occ}}_{\text{act}}$	# $\text{MLHF}^{\text{virt}}_{\text{act}}$	# $\text{CC}^{\text{occ}}_{\text{act}}$	# $\text{CC}^{\text{virt}}_{\text{act}}$	Excit. eng [eV]	Absolute error
	13	87	13	87	542.856	0.263
A	28	240	13	87	542.673	0.080
	40	356	13	87	542.593	0.000
	21	161	21	161	541.005	0.052
B	32	281	20	161	541.045	0.012
	40	356	20	161	541.057	0.000
	28	240	27	240	541.03	-
C	36	322	28	240	539.435	0.003
	40	356	28	240	539.432	0.000
Full CCSD	40	356	40	356	539.198	-

4.2.2 Nitrogen

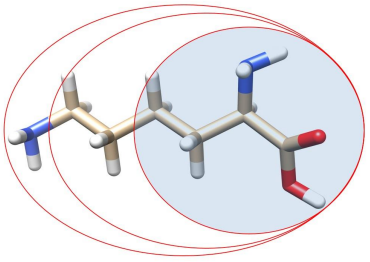
The same procedure as explained above is performed for nitrogen core excitation in lysine. Three different CC-active spaces are investigated, each with three different optimized MLHF active spaces, the smallest only including the CCSD basis itself, and the largest includes the full molecule. The three different subsystem investigated is illustrated in Figure 8 a-c, where the CCSD active space is highlighted in blue, and the HF-limits in red.



(a) Lysine system A with five CC active atoms and three different HF layers. The CC space is marked in blue and the HF space in red.



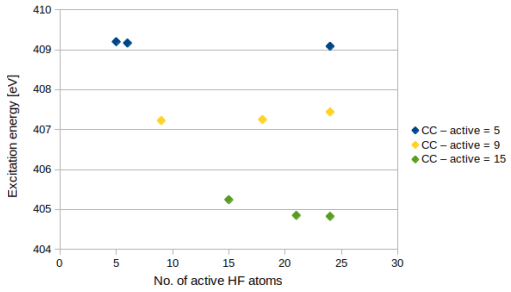
(b) Lysine system B with nine CC active atoms and three different HF layers. The CC space is marked in blue and the HF space in red.



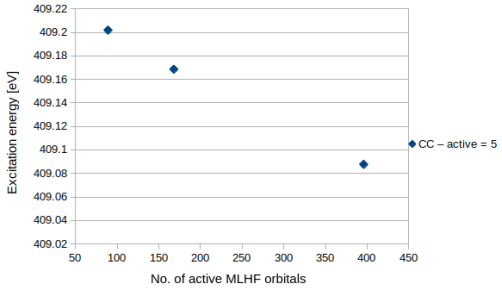
(c) Lysine system C with fifteen CC active atoms and three different HF layers. The CC space is marked in blue and the HF space in red.

Figure 8: The three different lysine systems in the nitrogen excitation calculation.

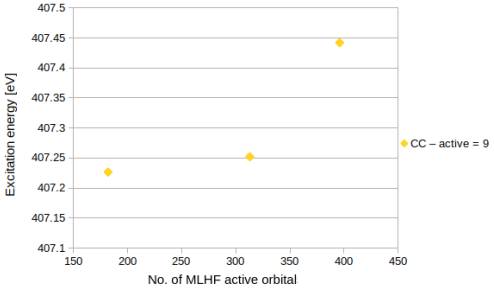
Figure 9 gives an overview of the excitation energy calculated plotted as a function of the number of active HF atoms within a given CC-active space. Figure 9b-9d takes a closer look at the individual CC spaces, and display the excitation energy as a function of the number of active MLHF orbitals.



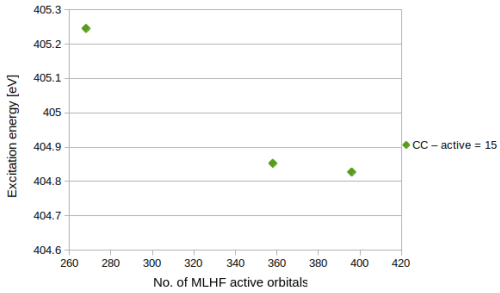
(a) Excitation energies for the different lysine systems, A, B and C displayed in Figure 8 as a function of the increasing HF-layer.



(b) The excitation energies for lysine system A, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.



(c) The excitation energies for lysine system B, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.



(d) The excitation energies for lysine system C, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.

Figure 9: The excitation energies for the three different lysine systems, A, B and C in the nitrogen core excitation calculation.

In Figure 10 the number of CC-active orbitals are plotted against the active MLHF orbitals with the excitation energy as a superscript. The energy from the fully optimized HF wave function for System B, and the MLHF wavefunction with only the CC-active orbitals optimized in system C will be excluded from this discussion. By looking to the CC-amplitudes in the CCSD calculation, these values were found to be different excitations than the one investigated, and should therefore not be included in this validation.

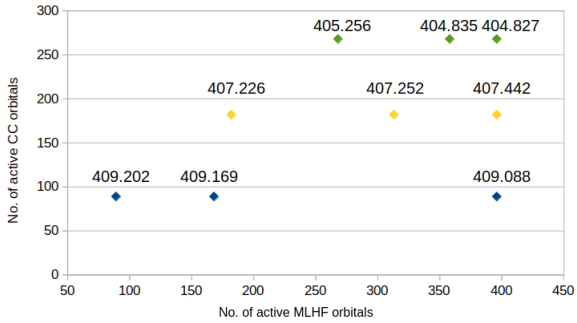


Figure 10: The number of CC-active orbitals plotted against the number of MLHF-active orbitals with the excitation energies for each given calculation.

As for the calculation on the oxygen's core orbital, we see that the absolute difference between the excitation energies within system A is large, 0.114 eV and 0.081 eV for 89 and 182 variationaly optimized orbitals respectively, and the excitation energy does not seem to converge within this space. For the B system, the fully optimized HF wave function converged to the wrong excitation energy, so the data is inconclusive. For system C the respective values are 0.003 eV for a difference of 48 optimized orbitals. A summary of the calculation on these systems can be found in Table 2.

Table 2: Excitation energies for the different lysine systems, A, B and C displayed in Figure 8 with the number of CC-active and MLHF-active occupied and virtual orbitals. The absolute error relative to the full HF calculation within the given CC space is calculated, but the outlier for the B and C system is excluded. As the outlier in the B is the full HF state, the absolute error is not calculated.

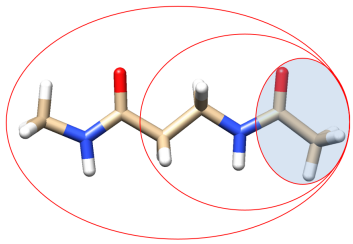
	$\#\text{MLHF}^{\text{occ}}_{\text{act}}$	$\#\text{MLHF}^{\text{virt}}_{\text{act}}$	$\#\text{CC}^{\text{occ}}_{\text{act}}$	$\#\text{CC}^{\text{virt}}_{\text{act}}$	Excit. eng [eV]	Absolute error
	10	79	10	79	409.202	0.114
A	21	161	13	87	409.169	0.081
	40	356	13	87	409.088	0.000
	21	161	21	161	407.226	-
B	32	281	20	161	407.252	-
	40	356	20	161	407.442	-
	28	240	27	240	405.246	-
C	36	322	28	240	404.853	0.026
	40	356	28	240	404.827	0.000

4.3 β-PEPTIDE

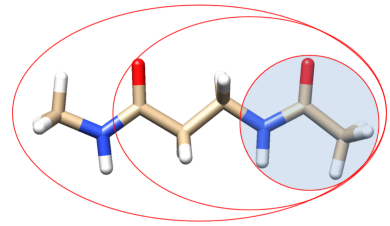
Here, some proof of concept calculations on the β peptide molecule are presented. The excitation energy from the core orbital of the oxygen and nitrogen within the active spaces are calculated, and the effect of changing the number of optimized HF orbitals within a given CC-active space is investigated.

4.3.1 Oxygen

The same procedure as previously used on the lysine molecule is performed. Here, only two CC-active spaces are investigated, each with three different optimized HF active spaces, the smallest only including the CCSD basis itself, and the largest includes the full molecule. The two different subsystem investigated is illustrated in Figure 11 a-b, where the CCSD active space is highlighted in blue, and the HF-limits in red.



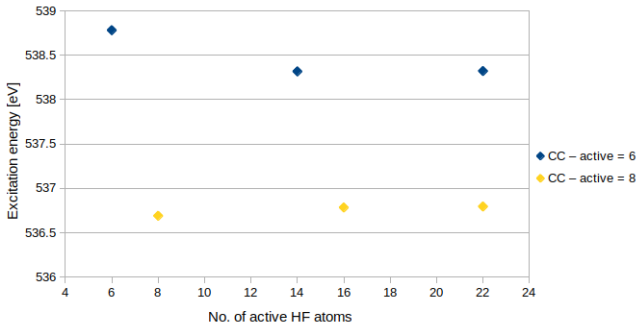
(a) Peptide system A with six CC active atoms and three different HF layers. The CC space is marked in blue and the HF space in red.



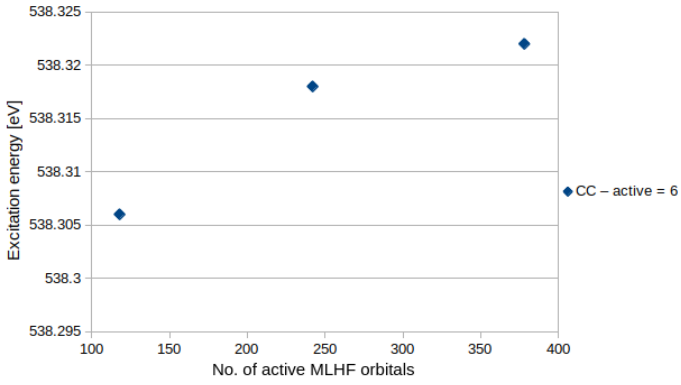
(b) Peptide system B with eight CC active atoms and three different HF layers. The CC space is marked in blue and the HF space in red.

Figure 11: The two different peptide systems in the oxygen core excitation calculation.

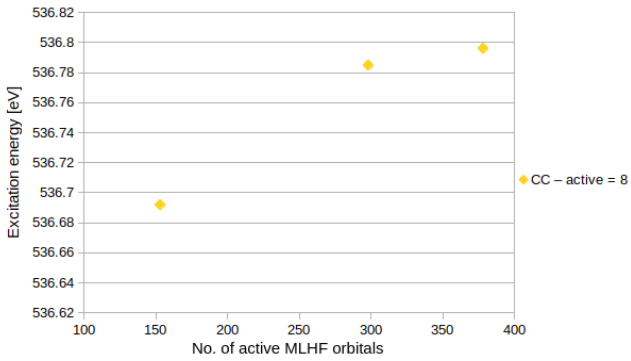
Figure 12 gives an overview of the excitation energy calculated plotted as a function of the number of active HF atoms within a given CC-active space. Figure 12b and 12c takes a closer look at the individual CC spaces, and display the excitation energy as a function of the number of active MLHF orbitals.



(a) Excitation energies for the different peptide systems, A, and B displayed in Figure 8 as a function of the increasing HF-layer.



(b) The excitation energies for peptide system A, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.



(c) The excitation energies for peptide system B, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.

Figure 12: The excitation energies for the two different peptide systems in the oxygen core excitation.

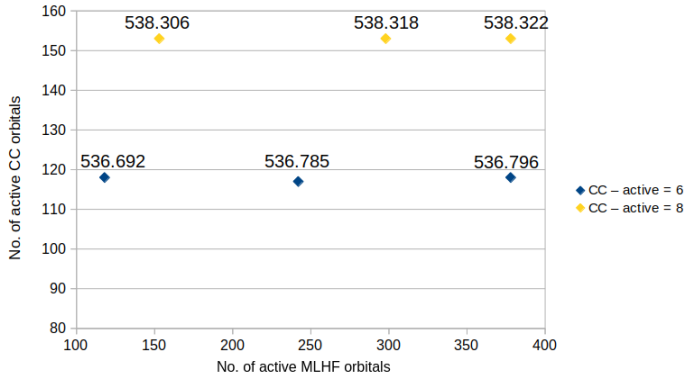


Figure 13: The number of CC-active orbitals plotted against the number of MLHF-active orbitals with the excitation energies for each given calculation.

Within the smallest CC-active space, where the number of CC-active orbitals are 117/118, the excitation energy increases as can be seen in Figure 13, and the relative error goes down as can be seen from Table 3. The excitation energy seems to converge with the changes in the MLHF space. The relative error between the state with 242 and 378 optimized MLHF orbitals is 0.004 eV. This is a promising result, as it is within the acceptable error with a reduce of 136 orbitals, which is approximately 36% of the orbital space.

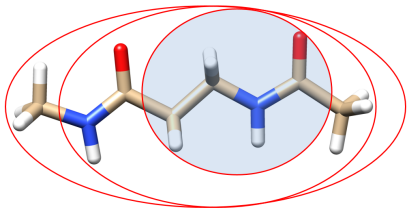
For the larger CC-active space, the energy decreases as the number of optimized MLHF orbitals goes up, if the outlier at 538.318 eV is excluded. This was found to be a different excitation by examining the CC-amplitudes in DALTON. The relative error between the state with 153 and 378 optimized MLHF orbitals is at 0.104 eV, and there is not enough data to conclude if the series converged or not.

Table 3: Excitation energies for the different lysine-systems, A and B displayed in Figure 11 with the number of CC-active and MLHF-active occupied and virtual orbitals. The absolute error relative to the full HF calculation within the given CC space is calculated, but the outlier for system B is excluded.

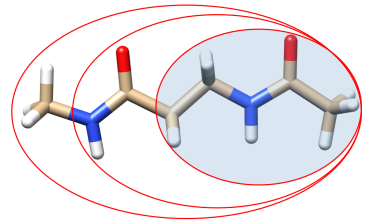
	$\#\text{MLHF}^{\text{occ}}_{\text{act}}$	$\#\text{MLHF}^{\text{virt}}_{\text{act}}$	$\#\text{CC}^{\text{occ}}_{\text{act}}$	$\#\text{CC}^{\text{virt}}_{\text{act}}$	Excit. eng [eV]	Absolute error
	13	105	13	105	538.306	0.016
A	25	217	13	104	538.318	0.004
	39	339	13	105	538.322	0.000
	16	137	16	137	536.692	0.104
B	33	265	16	137	536.785	-
	39	339	16	137	356.796	0.000

4.3.2 Nitrogen

For the studying of the excitation energy from the nitrogen's core orbital, two CC-active spaces are investigated, each with three different optimized MLHF active spaces, the smallest only including the CCSD basis itself, and the largest includes the full molecule. The two different systems investigated are illustrated in Figure 14 a-b, where the CCSD active space is highlighted in blue, and the HF-limits in red.



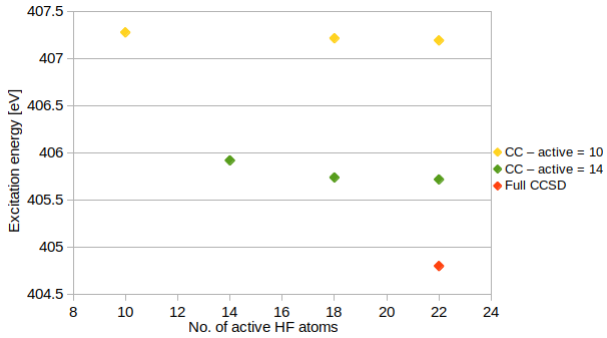
(a) Peptide system A with ten CC active atoms and three different HF layers. The CC space is marked in blue and the HF space in red.



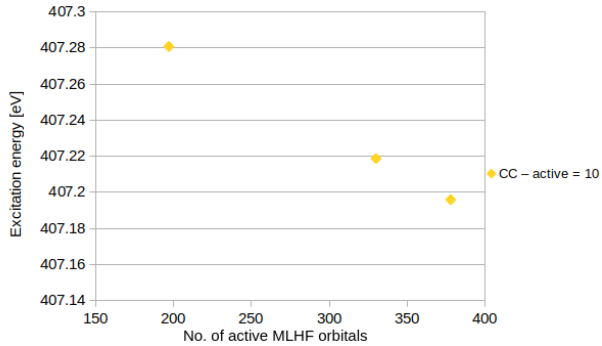
(b) Peptide system B with fourteen CC active atoms and three different HF layers. The CC space is marked in blue and the HF space in red.

Figure 14: The two different peptide systems in the nitrogen core excitation calculation.

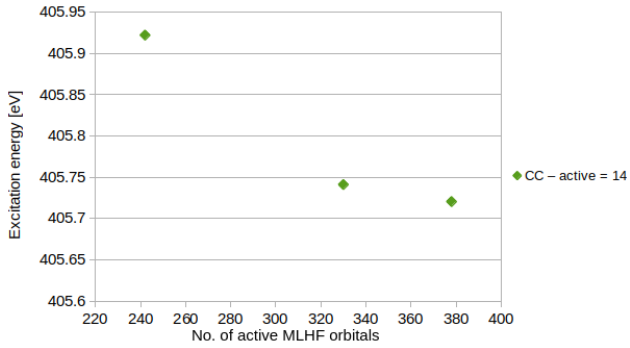
Figure 12 gives an overview of the excitation energy calculated plotted as a function of the number of active HF atoms within a given CC-active space. Figure 12b and 12c takes a closer look at the individual CC spaces, and display the excitation energy as a function of the number of active MLHF orbitals.



(a) Excitation energies for the different peptide systems, A and B displayed in Figure 14 as a function of the increasing HF-layer. A full CCSD calculation is also included.



(b) The excitation energies for peptide system A, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.



(c) The excitation energies for peptide system B, where the excitation energy is displayed as a function of the number of optimized MLHF orbitals.

Figure 15: The excitation energies for the two different peptide systems in the nitrogen core excitation calculation.

Figure 16 displays the number of CC-active orbitals plotted against the active MLHF orbitals with the excitation energy as a superscript. After examining the CC-amplitudes in DALTON, both the A system and B systems wavefunction of HF orbitals only optimized in the active CC-spcae converged to the wrong excitation energy and will therefore be excluded from the discussion. From table 4, we see that a difference in 48 in both states gives a relative total error of 0.023 eV and 0.02 eV, indicating that the size of the CC-active space has a role to play in the total error within the CC-active space.

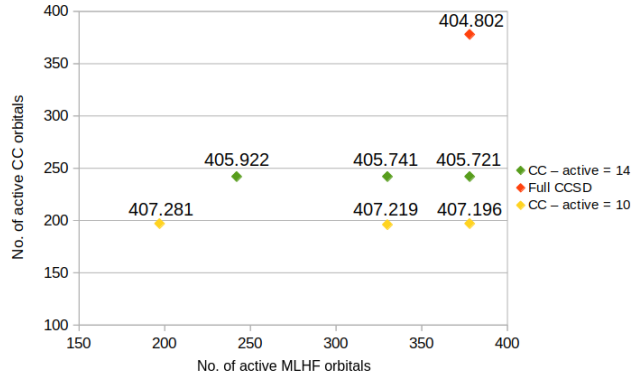


Figure 16: The number of CC-active orbitals plotted against the number of MLHF-active orbitals with the excitation energies for each given calculation.

Table 4: Excitation energies for the different lysine-systems, A and B displayed in Figure 14 with the number of CC-active and MLHF-active occupied and virtual orbitals. The absolute error relative to the full HF calculation within the given CC space is calculated, but the outlier for system A and B is excluded. The full CCSD excitation energy is included for comparison.

	$\#\text{MLHF}^{\text{occ}}_{\text{act}}$	$\#\text{MLHF}^{\text{virt}}_{\text{act}}$	$\#\text{CC}^{\text{occ}}_{\text{act}}$	$\#\text{CC}^{\text{virt}}_{\text{act}}$	Excit. eng [eV]	Absolute error
	23	174	23	174	407.281	-
A	35	295	22	174	407.219	0.023
	39	339	23	174	407.200	0.000
	25	217	25	217	405.922	-
B	35	295	25	217	405.741	0.020
	39	339	25	217	405.721	0.000
Full CCSD	39	339	39	339	404.801	-

5 Summary and concluding remarks

In this work a method enabling reduced orbital space CC calculations in a MLHF wave function have been developed. The whole system is described by one wave function, and the approach needs no a priori orbital assignment and no bonds are broken. The starting point is to define a set of CC-active atoms, and by using a threshold criteria, the active atoms in MLHF are defined. After a Cholesky decomposition over the active atoms is preformed, the resulting active density is optimized through SCF iterations in the active MO basis by using an orthogonal parametrization. The rest of the density is kept constant after an initial start guess. After the optimization procedure has converged, a new Cholesky decomposition is performed over the previously defined CC-active atoms, and the resulting Fock matrix built from these are then diagonalized, so that the input to the CCSD calculation is a set of canonical MOs and orbitals energies. As the reference wave function is not a HF wave function, this introduces some changes in the CCSD working equations. All changes enters through a modified Fock matrix.

Some benchmarking calculations on core excitations have been carried out on smaller systems to establish the effect of only having a partly optimized HF state as a reference state. The errors are small, but mostly not within the acceptable range of 0.01-0.001 eV, which means that using a MLHF wave function for these systems would compromise the CCSD calculation. There seems to be trend of faster convergence if the CC-active space is larger, but there is not enough data to conclude on this. Also, as the CC-active space tested grows larger, the differences between the HF layers tested also goes down, so this could also be a side effect.

This method is not developed for small systems such as this. Firstly, for such small systems a more accurate method is affordable and will give better results. Secondly, we may not reach the part of the wave function that becomes less important to optimize as a local property such as a core excitation energy still is affected by the nearest surroundings, and for a small system this may be most of the molecule. With that said, the results obtained are promising for this approach as errors are already small for the small systems investigated here. For larger systems it is expected that a relatively small MLHF active space will reproduce the effect of a fully optimized HF embedding wave function.

6 Further work

Throughout the calculations there have been a recurring problem of the CCSD calculation not finding the lowest excitation energy. This problem was attempted solved by adding start guesses and calculating the two lowest excitation energies instead, and were successful in the calculations that converged within the deadline. The approach of calculating the two lowest excitation energies should have been used from the start, and should be used in further testing of the method.

Further work should include proper benchmarking of the method and testing on larger systems which is what this method is developed for. The effect the size of the CC-active space has on the convergence of the different HF layers tested should be further investigated. One test calculation that would be interesting is seeing how well the method would describe the effect of e.g. a molecule surrounded by water molecules. This was attempted tested on the caffeine molecule, with and without 11 water molecules surrounding it, to see if the method would catch the effect the surrounding water has on the excitation energy without optimizing that part of the wave function. Sadly, these calculations did not converge within the deadline. We would however expect this method to catch some of this effect. MLHF does a god job in describing the solvent effect even if only a subset is variationally optimized [33], and as long as the CCSD space is large enough, this method should be able to give a good description. Further developments of this method therefore relies on a CCSD code that operates in the MO basis to be able to take advantage of the MO space reduction to run large systems. Also CCSD is not satisfactory in most cases to obtain chemical accuracy. The next step will be to include the effects of the triplets. This will be done by adding a CC3 layer in a subset of the CCSD space, and the model will thereby have four layers, with four different levels of theory; SAD, HF, CCSD and CC3.

References

[1] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen. John Wiley & Sons, 2014.

[2] Peter W Atkins and Ronald S Friedman. Oxford university press, 2011.

[3] Isaiah Shavitt. *Mol. Phys.*, 94(1):3–17, 1998.

[4] George D Purvis III and Rodney J Bartlett. *J. Chem. Phys.*, 76(4):1910–1918, 1982.

[5] Trygve Helgaker, Torgeir A Ruden, Poul Jørgensen, Jeppe Olsen, and Wim Klopper. *J. Phys. Org. Chem.*, 17(11):913–933, 2004.

[6] Jozef Noga and Rodney J Bartlett. *J. Chem. Phys.*, 86(12):7041–7050, 1987.

[7] Krishnan Raghavachari, Gary W Trucks, John A Pople, and Martin Head-Gordon. *Chem. Phys. Lett.*, 157(6):479–483, 1989.

[8] Henrik Koch, Ove Christiansen, Poul Jørgensen, Alfredo M Sanchez de Merás, and Trygve Helgaker. *J. Chem. Phys.*, 106(5):1808–1818, 1997.

[9] A Warshel and M Karplus. *J. Am .Chem. Soc*, 94(16):5612–5625, 1972.

[10] Arieh Warshel and Michael Levitt. *J. Mol. Biol.*, 103(2):227–249, 1976.

[11] Mats Svensson, Stephane Humbel, Robert DJ Froese, Toshiaki Matsubara, Stefan Sieber, and Keiji Morokuma. *J. Phys. Chem.*, 100(50):19357–19363, 1996.

[12] Thom Vreven and Keiji Morokuma. *Annu. Rep. Comput. Chem.*, 2:35–51, 2006.

[13] Hans Martin Senn and Walter Thiel. *Angew. Chem. Int. Ed.*, 48(7):1198–1229, 2009.

[14] Andrew R Leach. Pearson education, 2001.

[15] Jacopo Tomasi, Benedetta Mennucci, and Roberto Cammi. *Chem. Rev.*, 105(8):2999–3094, 2005.

[16] Ira N Levine. Prentice-Hall, 2000.

[17] P. Cortona. *Phys. Rev. B*, 44:8454–8458, 1991.

[18] P. Elliott, K. Bruker, MH. Cohen, and A. Wasserman. *Phys. Rev. A*, 82:024501, 2010.

[19] Mark E Fornace, Joonho Lee, Kaito Miyamoto, Frederick R Manby, and Thomas F Miller III. *J. Chem. Theory Comput.*, 11(2):568–580, 2015.

[20] Christoph R. Jacob and Johannes Neugebauer. *WIREs Comput. Mol. Sci.*, 4(4):325–362, 2014.

[21] Bence Hégely, Péter R Nagy, György G Ferenczy, and Mihály Kállay. *J. Chem. Phys.*, 145(6):064107, 2016.

[22] Csaba Daday, Carolin König, Johannes Neugebauer, and Claudia Filippi. *Chem. Phys. Chem.*, 15(15):3205–3217, 2014.

[23] Florian Libisch, Chen Huang, and Emily A Carter. *Acc. Chem. Res.*, 47(9):2768–2775, 2014.

[24] Simon J. Bennie, Martina Stella, Thomas F. Miller III, and Frederick R. Manby. *J. Chem. Phys.*, 143(2):024105, 2015.

[25] Tomasz A Wesolowski, Sapana Shedge, and Xiuwen Zhou. *Chem. Rev.*, 115(12):5891–5928, 2015.

[26] Niranjan Govind, Yan Alexander Wang, and Emily A Carter. *J. Chem. Phys.*, 110(16):7677–7688, 1999.

[27] Sebastian Hofener and Lucas Visscher. *J. Chem. Phys.*, 137(20), 2012.

[28] RA. Mata, HJ. Werner, and M. Schütz. *J. Chem. Phys.*, 128:144106, 2008.

[29] W. Li and P. Piecuch. *J. Phys. Chem. A*, 114:6721, 2010.

[30] Alok Shukla, Michael Dolg, Hermann Stoll, and Peter Fulde. *Chem. Phys. Lett.*, 262(3):213–218, 1996.

[31] Dundas Karen. Master's thesis, Norwegian University of Science and Technology, Norway, 2016.

[32] Trygve Helgaker, Helena Larsen, Jeppe Olsen, and Poul Jørgensen. *Chem. Phys. Lett.*, 327(5):397–403, 2000.

[33] Sandra Sæther, Thomas Kjærgaard, Henrik Koch, and Ida-Marie Høyvik. *To be submitted*, 2017.

[34] Kestutis Aidas, Celestino Angeli, Keld L. Bak, Vebjørn Bakken, Radovan Bast, Linus Boman, Ove Christiansen, Renzo Cimiraglia, Sonia Coriani, Pål Dahle, Erik K. Dalskov, Ulf Ekström, Thomas Enevoldsen, Janus J. Eriksen, Patrick Ettenhuber, Berta Fernández, Lara Ferrighi, Heike Fliegl, Luca Frediani, Kasper Hald, Asger Halkier, Christof Hättig, Hanne Heiberg, Trygve Helgaker, Alf Christian Hennum, Hinne Hettema, Eirik Hjertenæs, Stinne Høst, Ida-Marie Høyvik, Maria Francesca Iozzi, Branislav Jansík, Hans Jørgen Aa. Jensen, Dan Jonsson, Poul Jørgensen, Joanna Kauczor, Sheela Kirpekar, Thomas Kjærgaard, Wim Klopper, Stefan Knecht, Rika Kobayashi, Henrik Koch, Jacob Kongsted, Andreas Krapp, Kasper Kristensen, Andrea Ligabue, Ola B Lutnæs, Juan I. Melo, Kurt V. Mikkelsen, Rolf H. Myhre, Christian

Neiss, Christian B. Nielsen, Patrick Norman, Jeppe Olsen, Jógvan Magnus H. Olsen, Anders Osted, Martin J. Packer, Filip Pawlowski, Thomas B. Pedersen, Patricio F. Provasi, Simen Reine, Zilvinas Rinkevicius, Torgeir A. Ruden, Kenneth Ruud, Vladimir V. Rybkin, Pawel Sałek, Claire C. M. Samson, Alfredo Sánchez de Merás, Trond Saue, Stephan P. A. Sauer, Bernd Schimmelpfennig, Kristian Sneskov, Arnfinn H. Steindal, Kristian O. Sylvester-Hvid, Peter R. Taylor, Andrew M. Teale, Erik I. Tellgren, David P. Tew, Andreas J. Thorvaldsen, Lea Thøgersen, Olav Vahtras, Mark A. Watson, David J. D. Wilson, Marcin Ziolkowski, and Hans Ågren. *WIREs Comput. Mol. Sci.*, 4(3):269–284, 2014.

[35] Rolf H Myhre, Alfredo MJ Sánchez de Merás, and Henrik Koch. *J. Chem. Phys.*, 141(22):224105, 2014.

[36] Peter Pulay. *Chem. Phys. Lett.*, 73(2):393–398, 1980.

[37] Peter Pulay. *J. Comput. Chem.*, 3(4):556–560, 1982.

[38] JH Van Lenthe, R Zwaans, Huub JJ Van Dam, and MF Guest. *J. Comput. Chem*, 27(8):926–932, 2006.

[39] Frank Jensen. John wiley & sons, 2007.

[40] Thom H Dunning Jr. *J. Chem. Phys.*, 90(2):1007–1023, 1989.

[41] Jiří Čížek. *J. Chem. Phys.*, 45(11):4256–4266, 1966.

[42] J Paldus, J Čížek, and I Shavitt. *Phys. Rev. A.*, 5(1):50, 1972.

[43] Henrik Koch, Ove Christiansen, Rika Kobayashi, Poul Jørgensen, and Trygve Helgaker. *Chem. Phys. Lett.*, 228(1-3):233–238, 1994.

[44] Henrik Koch and Poul Jørgensen. *J. Chem. Phys.*, 93(5):3333–3344, 1990.

[45] Thomas Bondo Pedersen and Henrik Koch. *J. Chem. Phys.*, 106(19):8059–8072, 1997.

[46] Rodney J Bartlett and Monika Musiał. *Rev. Mod. Phys.*, 79(1):291, 2007.

[47] Ernest R Davidson. *J. Comput. Phys.*, 17(1):87–94, 1975.

[48] Sonia Coriani and Henrik Koch. *J. Chem. Phys.*, 2015.

[49] Francesco Aquilante, Thomas Bondo Pedersen, Alfredo Sánchez de Merás, and Henrik Koch. *J. Chem. Phys.*, 125(17):174101, 2006.

[50] Ida-Marie Høyvik, Rolf Heilemann Myhre, and Henrik Koch. *J. Chem. Phys.*, 146(14):144109, 2017.

[51] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. *J. Comput. Chem.*, 25(13):1605–1612, 2004.