



**NTNU – Trondheim**  
Norwegian University of  
Science and Technology

# Development and implementation of extended CC2 models

**Rolf Heilemann Myhre**

Chemical Engineering and Biotechnology

Submission date: July 2013

Supervisor: Henrik Koch, IKJ

Norwegian University of Science and Technology  
Department of Chemistry



# Declaration

I, Rolf HEILEMANN MYHRE, declare that the work presented in this thesis has been carried out independently and in agreement with *Reglement for sivilarkitekt- og sivilingeniøreksamen* at the Norwegian University of Science and Technology (NTNU)

---

Trondheim, July 21, 2013



# Preface

This thesis is submitted to the Department of Chemistry, Faculty of Natural Sciences and Technology, Norwegian University of Science and Technology (NTNU) as a partial fulfillment of the degree of Master of Technology (M. Tech.). This thesis concludes the five year Master's degree program in Chemical Engineering and Biotechnology, leading to a M. Tech. degree in Physical Chemistry at NTNU.

The present work has been carried out during the spring semester of 2013, between January 10<sup>th</sup> and July 20<sup>th</sup>. The project was supervised by Professor Henrik Koch at the Department of Chemistry.

The project is part of an ongoing research project into the development and implementation of multi-level coupled cluster theory.



# Abstract

Multi-level coupled cluster theory is presented with special focus on the extended CC2 model. Combined with Cholesky molecular orbitals, these models makes it possible to treat different subsystems with different levels of coupled cluster theory giving potentially large reductions in computational complexity. Total energy and excitation energies using ECC2 are presented for several different molecular systems. ECC2 can reproduce CCSD results when using appropriate parameters.



# Sammendrag

Denne oppgaven er en del av et pågående forskningsprosjekt som ser på utvikling og implementering av *multi-level coupled cluster* (MLCC) teori. I MLCC, blir orbitalrommet delt opp i to eller flere underrom. Basert på hvilket underrom orbitalene involvert tilhører blir eksitasjonsoperatorene i *cluster*-operatoren delt inn i forskjellige operatører. Ved å behandle delene av *cluster*-operatoren med forskjellige nivå av *coupled cluster* teori kan man få store besparelser i beregningstid med minimale tap av nøyaktighet.

I Cholesky dekomposisjon blir en positiv semidefinit matrise dekomponert i en nedre triangulær matrise og dens transponerte. For glisse, diagonaldominerte matriser er denne dekomposisjonen svært effektiv. Én-elektrons tetthetsmatrisen i atomorbitalbasen er både positiv semidefinit, diagonaldominert og glisse og ved å dekomponere den vil Cholesky-vektorene tilsvare parametriseringen til lokaliserte molekylorbitaler (MO). Diagonalelementene i tetthetsmatrisen korresponderer til atomorbitaler sentrert på atomer og et lokalt, aktivt rom kan oppnås ved å definere et sett med atomer som aktive.

Pilotkode for total energi og eksitasjonsenergi har blitt implementert for MLCC modellen *extended CC2* (ECC2) i programvarepakken DALTON. I ECC2 deles orbitalene inn i to rom som behandles med CC2 og CCSD. Ved å inkludere HOMO og LUMO i det aktive rom unngikk man problemer med nær degenerering som oppstod i dissosiasjonsprosesser med CC2. Med et lokalt aktivt rom ble CCSD-dissosiasjonsenergien for eten reproduert til kjemisk nøyaktighet. ECC2 viste også forbedringer i beregninger av elektronisk dipolmoment og polariserbarhet for systemer hvor CC2 var unøyaktig i forhold til CCSD.

Ved å bruke lokale aktive rom kunne ECC2 reproducere CCSD-verdier for eksitasjonsenergi. ECC2 var mest vellykket dersom alle orbitalene involvert i eksitasjonen var inkludert i det aktive rommet, men dette var ikke absolutt nødvendig for høy nøyaktighet.

Ved å implementere et rom beskrevet med SCF langt fra det aktive rom vil beregninger skalere lineært med systemstørrelse for molekylsystemer over en viss størrelse fordi elektronkorrelasjon bare blir beregnet for et fast, aktivt rom. Ved hjelp av Cholesky dekomposisjon kan en generere en *auxiliary* basis for beregning av MO-integraler. Dette gjør det i prinsippet mulig å effektivt implementere en modell med størrelsesintensiv kompleksitet som kan konkurrere med DFT i skalering for store systemer.

# Acknowledgements

First, I would like to thank my supervisor, Professor Henrik Koch, and Professor Alfredo M. J. Sánchez de Merás, who has acted as co-supervisor, for letting me work on this project. It has been very interesting and challenging and they have provided patience and guidance when things have appeared too challenging.

A special thanks goes to Andrew Dibbs and Eirik Hjertenæs for reading through and correcting spelling and grammar in this thesis, as well as helping with practical things and L<sup>A</sup>T<sub>E</sub>X.

I would also like to thank the groups of Applied Theoretical Chemistry at NTNU and ICMOL at the University of Valencia. I have felt very welcome in both groups and working with researchers have encouraged me to seek out an academic career.

Finally, I would like to thank my family and friends who have supported and encouraged me throughout my studies and when working with this project.



*To my mother Nina*



# Contents

<b>Declaration</b>	<b>i</b>
<b>Preface</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Sammendrag</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xvii</b>
<b>Abbreviations</b>	<b>xix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Scope of the work . . . . .	2
1.3 Outline of the thesis . . . . .	2
<b>2 Theory</b>	<b>5</b>
2.1 Coupled cluster theory . . . . .	5
2.1.1 Cluster operator and exponential ansatz . . . . .	5
2.1.2 Coupled cluster equations . . . . .	7
2.1.3 Truncated cluster operator . . . . .	9
2.1.4 Approximate coupled cluster models . . . . .	11
2.2 Multi-level coupled cluster theory . . . . .	12
2.2.1 Cholesky decomposition . . . . .	12
2.2.2 Extended CC2 . . . . .	14
2.2.3 Beyond ECC2 . . . . .	17

2.3	MLCC response theory . . . . .	19
2.3.1	Quasi-energy response method . . . . .	19
2.3.2	Linear response function for ECC2 . . . . .	22
<b>3</b>	<b>Implementation</b>	<b>25</b>
3.1	ECC2 energy calculations . . . . .	25
3.2	ECC2 excitation energies . . . . .	26
<b>4</b>	<b>Results</b>	<b>29</b>
4.1	ECC2 energy calculations . . . . .	29
4.1.1	Abstraction processes . . . . .	29
4.1.2	Local geometry . . . . .	32
4.1.3	Static properties . . . . .	34
4.2	Excitation energies . . . . .	37
4.2.1	Functional groups . . . . .	37
4.2.2	Solvent effects . . . . .	42
4.2.3	Conjugate systems . . . . .	46
<b>5</b>	<b>Discussion</b>	<b>51</b>
5.1	Linear and sublinear scaling . . . . .	51
5.2	Extensive properties . . . . .	52
5.3	Intensive properties . . . . .	53
5.4	Strategy . . . . .	55
5.5	Variations . . . . .	56
<b>6</b>	<b>Conclusion</b>	<b>59</b>
6.1	Development of MLCC . . . . .	59
6.2	Results from calculations . . . . .	60
<b>7</b>	<b>Future work</b>	<b>61</b>
<b>A</b>	<b>ECC2 response derivation</b>	<b>63</b>
<b>B</b>	<b>Initial geometries</b>	<b>69</b>
	<b>Bibliography</b>	<b>71</b>

# List of Figures

2.1	Example of an active space of 1,3-butadiene. I denotes internal, SE semi-external and E external excitations with respect to the active CCSD space. . . . .	15
4.1	Total energy curves for dissociation of lithium hydride using the basis set aug-cc-pVDZ. . . . .	30
4.2	Total energy curves for dissociation of a sodium dimer using the basis set aug-cc-pVDZ. . . . .	31
4.3	Total energy curves for dissociation of a hydrogen atom from ethene using the basis set aug-cc-pVDZ. . . . .	32
4.4	Total energy curves for dissociation of a hydrogen atom from 1,3-butadiene using the basis set aug-cc-pVDZ. . . . .	33
4.5	Decanal A . . . . .	39
4.6	Decanal B . . . . .	39
4.7	Decanal C . . . . .	39
4.8	Decanal D . . . . .	39
4.9	Trans-ethyl-i-butyl-diazene A . . . . .	42
4.10	Trans-ethyl-i-butyl-diazene B . . . . .	42
4.11	Tert-butyl hydroperoxide A . . . . .	43
4.12	Tert-butyl hydroperoxide B . . . . .	43
4.13	Acetone A . . . . .	45
4.14	Acetone B . . . . .	45
4.15	(2E,4E,6E,8E)-2,4,6,8-decatetraene . . . . .	47
4.16	2,4,6,8-decatetraenal A . . . . .	48
4.17	2,4,6,8-decatetraenal B . . . . .	48
4.18	1,3-octadiene . . . . .	49
4.19	6-methyl-2,4-heptanedione . . . . .	49
4.20	2,4-octadienal . . . . .	49



# List of Tables

4.1	Electronic dipole moments and polarisability in a.u. of hydrogen fluoride the along $C_\infty$ axis with the basis set aug-cc-pVTZ. . . . .	34
4.2	Electronic dipole moment and polarisability in a.u. of ozone the along $C_2$ axis with the basis set aug-cc-pVTZ. . . . .	35
4.3	Polarisability in a.u. of ethyne and ethene along the C-C bonds with the basis set aug-cc-pVTZ. . . . .	36
4.4	Polarisability in a.u. of benzene along a $C_2$ axis going through two hydrogen atoms and 1-3-butadiene along the C-C single bond with the basis set aug-cc-pVTZ. . . . .	36
4.5	Excitation energies of decanal in eV using cc-pVDZ. . . . .	40
4.6	Norm of the amplitudes of the different types of single excitations for decanal using cc-pVDZ. . . . .	40
4.7	Excitation energies of trans-ethyl-i-butyl-diazene in eV using cc-pVDZ above and aug-cc-pVDZ below. . . . .	41
4.8	Norm of the amplitudes of the different types of single excitations for trans-ethyl-i-butyl-diazene using cc-pVDZ. . . . .	41
4.9	Excitation energies of tert-butyl hydroperoxide in eV using cc-pVDZ. . . . .	42
4.10	Norm of the amplitudes of the different types of single excitations for tert-butyl hydroperoxide using cc-pVDZ. . . . .	43
4.11	Excitation energies of acetone with water in eV using cc-pVDZ above and aug-cc-pVDZ below. . . . .	44
4.12	Norm of the amplitudes of the different types of single excitations for acetone with water using cc-pVDZ above and aug-cc-pVDZ below. . . . .	45
4.13	Excitation energies of (2E,4E,6E,8E)-2,4,6,8-decatetraene in eV using cc-pVDZ. . . . .	46
4.14	Norm of the amplitudes of the different types of single excitations for (2E,4E,6E,8E)-2,4,6,8-decatetraene using cc-pVDZ. . . . .	46
4.15	Excitation energies of 2,4,6,8-decatetraenal in eV using cc-pVDZ. . . . .	47
4.16	Norm of the amplitudes of the different types of single excitations for 2,4,6,8-decatetraenal using cc-pVDZ. . . . .	47
B.1	Initial geometry of ethene ( $\text{\AA}$ ). . . . .	69
B.2	Initial geometry of 1,3-butadiene ( $\text{\AA}$ ). . . . .	69
B.3	Geometry of acetone with three water molecules ( $\text{\AA}$ ). . . . .	70



# Abbreviations

<b>AO</b>	<b>A</b> tomic- <b>O</b> rbital
<b>BCH</b>	<b>B</b> aker- <b>C</b> ampbell- <b>H</b> ausdorff (expansion)
<b>CC</b>	<b>C</b> oupled <b>C</b> luster
<b>CI</b>	<b>C</b> onfiguration <b>I</b> nteraction
<b>CCS</b>	<b>C</b> oupled <b>C</b> luster <b>S</b> ingles
<b>CCSD</b>	<b>C</b> oupled <b>C</b> luster <b>S</b> ingles and <b>D</b> oubles
<b>CCSDT</b>	<b>C</b> oupled <b>C</b> luster <b>S</b> ingles, <b>D</b> oubles and <b>T</b> riples
<b>DFT</b>	<b>D</b> ensity <b>F</b> unctional <b>T</b> heory
<b>ECC</b>	<b>E</b> xtended <b>C</b> oupled <b>C</b> luster
<b>FCI</b>	<b>F</b> ull <b>C</b> onfiguration <b>I</b> nteraction
<b>HF</b>	<b>H</b> artree- <b>F</b> ock
<b>HOMO</b>	<b>H</b> ighest <b>O</b> ccupied <b>M</b> olecular <b>O</b> rbital
<b>LUMO</b>	<b>L</b> owest <b>U</b> noccupied <b>M</b> olecular <b>O</b> rbital
<b>MLCC</b>	<b>M</b> ulti- <b>L</b> evel <b>C</b> oupled <b>C</b> luster
<b>MO</b>	<b>M</b> olecular <b>O</b> rbital
<b>MP</b>	<b>M</b> øller- <b>P</b> lesset
<b>QE</b>	<b>Q</b> uasi- <b>E</b> nergy
<b>SCF</b>	<b>S</b> elf- <b>C</b> onsistent <b>F</b> ield
<b>SIC</b>	<b>S</b> ize- <b>I</b> ntensive <b>C</b> omplexity



# Chapter 1

## Introduction

### 1.1 Motivation

Modern quantum chemistry calculations can be performed for small molecules to obtain more accurate results than is currently possible through experimental methods[1]. However, for larger systems, wave function based theories encounter a computational barrier as all interactions between all electrons in all orbitals have to be calculated. One solution is to truncate the expressions and neglect terms that are considered less important. However, scaling is still a challenge. For example, the most expensive term in coupled cluster singles and doubles[2–4] (CCSD) scales as  $V^4O^2$  where  $V$  is the number of virtual and  $O$  is the number of occupied orbitals.

An alternative is density functional theory[5–7] (DFT) where the energy is determined as a functional of the electron density. In principle, DFT, can yield exact results, however, in practice, an approximate density functional must be employed. This makes the method less reliable.

Coupled cluster (CC) theory is arguably the most successful wave function based theory[1, 4] in use today. Much of current development is focused on speeding up calculations without reducing their accuracy. Many of the techniques developed make use of the fact that electron correlation is quite localised in non-conducting

systems[8–10]. Using pair natural orbitals and projected atomic orbitals, Riplinger and Neese obtained near-linear scaling for CCSD[11]. More indirectly, the local nature of the interactions can be exploited by noting that two-electron matrices are sparse and positive semi-definite. Linear dependencies can then be removed using Cholesky decomposition[12–18].

Multi-level CC (MLCC) theory uses a different approach to simplifying the calculations based on active spaces. Active spaces have been combined with CC theory before[19, 20], however, MLCC gives the opportunity to describe successive active spaces with increasing levels of accuracy. Combined with Cholesky molecular orbitals[21, 22] (MO) this makes it possible to treat distinct parts of a molecular system with different levels of theory, reminiscent of the ONIOM model[23]. Unlike similar schemes like embedding[24] and QM/MM[25], MLCC is fully antisymmetric across levels of approximation. By carefully assigning the active spaces, MLCC should be able to produce highly accurate results while computational complexity scales sublinearly.

## 1.2 Scope of the work

In this work, the equations for the extended CC2 (ECC2) model[26] and equations for more general MLCC models are presented. From these, the linear response functions for ECC2 is developed using the quasi-energy (QE) Lagrangian approach[27–32]. The ECC2 energy code[26] is tested on a number of systems and an excitation energy solver is implemented and tested.

## 1.3 Outline of the thesis

Chapter 2 starts with a summary of CC theory before presenting the equations of some MLCC models. This is followed by a discussion of the QE approach to response functions and derivation of these for ECC2. Cholesky MOs are also

discussed. Chapter 3 outlines the implementation of ECC2 in the Dalton software package[33, 34]. Results from ECC2 calculations are presented and compared with CC2 and CCSD in Chapter 4 and discussed in Chapter 5. Chapter 6 summarises the results and further work is discussed in Chapter 7.



# Chapter 2

## Theory

### 2.1 Coupled cluster theory

#### 2.1.1 Cluster operator and exponential ansatz

This section assumes the reader is familiar with the basics of quantum chemistry and electronic structure theory. For a review, see Szabo and Ostlund[35].

Coupled cluster theory can be viewed as a correction that introduces electron correlation to Hartree-Fock (HF) theory. The HF state is used as the reference state and thus necessitates that the HF state is a relatively good description of the system. Consequently, the ground state must be dominated by a single electron configuration.

Interactions between electrons is described by the simultaneous excitation of two or more electrons from occupied orbitals to virtual orbitals in the HF ground state. Using the indices  $i, j, k, l$  for occupied;  $a, b, c, d$  for virtual and  $p, q, r, s$  for general spin-orbitals, an interaction between two electrons is written as

$$\tau_{ij}^{ab} = a_a^\dagger a_i a_b^\dagger a_j. \quad (2.1)$$

where  $a_p^\dagger$  is the second quantisation creation operator that creates an electron in spin-orbital  $p$  while  $a_p$  is the annihilation operator that removes an electron from the spin-orbital. For a thorough discussion of second quantisation and coupled cluster theory, see the monograph by Helgaker, Jørgensen and Olsen[1]. To ensure that the wave function is antisymmetric, the creation and annihilation operators obey the anticommutation relations

$$[a_p^\dagger, a_q^\dagger]_+ = [a_p, a_q]_+ = 0 \quad (2.2)$$

$$[a_p^\dagger, a_q]_+ = \delta_{pq} \quad (2.3)$$

The Hamiltonian in second quantisation is

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} (pq|rs) a_p^\dagger a_q^\dagger a_r a_s + h_{\text{nuc}} \quad (2.4)$$

where

$$h_{pq} = \int \phi_p^*(\mathbf{x}) \left( -\frac{1}{2} \nabla^2 - \sum_I \frac{Z_I}{r_I} \right) \phi_q(\mathbf{x}) d\mathbf{x} \quad (2.5)$$

$$(pq|rs) = \int \int \frac{\phi_p^*(\mathbf{x}_1) \phi_r^*(\mathbf{x}_2) \phi_q(\mathbf{x}_1) \phi_s(\mathbf{x}_2)}{r_{12}} d\mathbf{x}_1 d\mathbf{x}_2 \quad (2.6)$$

$$h_{\text{nuc}} = \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{R_{IJ}} \quad (2.7)$$

$\phi$  refers to spin-orbitals,  $I$  and  $J$  to atomic nuclei and  $Z$  to nuclear charge.  $R$  is internuclear distances and  $r$  is the distance between an electron and a nucleus.

A general excitation is denoted  $\tau_\mu$  and can, in exact theory, involve all electrons in the system. If the excitation is a single replacement, only involving one electron, it is considered a relaxation of the orbitals as a response to the change in the electric field due to interactions.

The exact CC wave function is obtained by taking all such excitations into account. This is done by associating an amplitude,  $t_\mu$ , with every excitation operator and

writing the CC wave function as

$$|\text{CC}\rangle = \left[ \prod_{\mu} (1 + t_{\mu} \tau_{\mu}) \right] |\text{HF}\rangle. \quad (2.8)$$

This form of the wave function is a non-linear parametrisation of the wave function and is in itself not very useful. However, using the commutation relations of the excitation operators

$$[\tau_{\mu}, \tau_{\nu}] \quad (2.9)$$

and

$$\tau_{\mu}^2 = 0, \quad (2.10)$$

the CC wave function can be written as

$$|\text{CC}\rangle = \exp(X)|\text{HF}\rangle \quad (2.11)$$

where  $X$  is the cluster operator usually denoted  $T$  in the literature.

$$X = \sum_{\mu} t_{\mu} \tau_{\mu} \quad (2.12)$$

Eq. (2.11) is known as the exponential ansatz. Writing out the exponential expansion, one obtains contributions from all possible configurations of the spin orbitals.

## 2.1.2 Coupled cluster equations

In configuration interaction (CI) models, the paramtrisation is linear and taking the derivative with respect to the variational parameters is easy. This is not the case for CC models. Taking the derivative of the wave function in Eqs. (2.8) or (2.11) with respect to the amplitudes give expressions that depend on the amplitudes themselves. As a result, finding the wave function by minimising the energy expectation value of the Hamiltonian,

$$\langle E \rangle = \frac{\langle \text{CC} | H | \text{CC} \rangle}{\langle \text{CC} | \text{CC} \rangle} \quad (2.13)$$

leads to a set of nonlinear equations that each involve all the state determinants and all combinations of the amplitudes.

To avoid the minimisation problem, the projected CC equations are used.

$$E = \langle \text{HF} | H | \text{CC} \rangle \quad (2.14)$$

$$\langle \mu | \text{CC} \rangle E = \langle \mu | H | \text{CC} \rangle \quad (2.15)$$

In Eqs. (2.14) and (2.15), known as the unlinked CC equations,  $E$  is the CC energy. They hold because the overlap between the reference state and the wave function is

$$\langle \text{HF} | \text{CC} \rangle = 1 \quad (2.16)$$

and  $|\text{CC}\rangle$  satisfies the Schrödinger equation

$$H|\text{CC}\rangle = E|\text{CC}\rangle \quad (2.17)$$

The unlinked CC equations can be reformulated as the equivalent linked CC equations.

$$E = \langle \text{HF} | \exp(-X) H \exp(X) | \text{HF} \rangle \quad (2.18)$$

$$0 = \langle \mu | \exp(-X) H \exp(X) | \text{HF} \rangle \quad (2.19)$$

In the exact formulation the equivalence is apparent because

$$\exp(-X) H \exp(X) | \text{HF} \rangle = E \exp(-X) \exp(X) | \text{HF} \rangle = E | \text{HF} \rangle \quad (2.20)$$

and

$$\langle \text{HF} | \exp(-X) = \langle \text{HF} | \quad (2.21)$$

For a truncated cluster operator, the equivalence is more involved, but it holds if the excitation manifold,  $\{\langle \mu | \}$ , is closed under de-excitation[1].

While both the linked and unlinked formulations can be used to formulate CC

theory, the linked equations have the advantage that they are easily expanded using the Baker-Campbell-Hausdorff (BCH) expansion.

$$\exp(\mathbf{A})\mathbf{B}\exp(-\mathbf{A}) = \mathbf{B} + [\mathbf{A}, \mathbf{B}] + \frac{1}{2!}[\mathbf{A}, [\mathbf{A}, \mathbf{B}]] + \frac{1}{3!}[\mathbf{A}, [\mathbf{A}, [\mathbf{A}, \mathbf{B}]]] + \dots \quad (2.22)$$

The Hamiltonian contains one and two electron operators, so each time it is commuted with an excitation operator, one of its creation or annihilation operators are fixed. Consequently, all commutator terms involving more than four commutators vanish.

$$\begin{aligned} \exp(-X)H\exp(X) = \\ H + [H, X] + \frac{1}{2!}[[H, X], X] + \frac{1}{3!}[[[H, X], X], X] + \frac{1}{4!}[[[[H, X], X], X], X] \end{aligned} \quad (2.23)$$

One of the advantages of CC theory is that it is size-extensive. This can be shown using the fact that the cluster operators for two non-interacting systems will commute with each other and the Hamiltonian of the other system. The linked formulation is even termwise size-extensive. This is very useful when truncating the cluster operator[1].

### 2.1.3 Truncated cluster operator

If all excitations are included in the cluster operator in Eq. (2.12), the CC wave function is equivalent to the full CI (FCI) wave function with a more complicated formulation and different normalisation. To make a practically feasible model, the cluster operator is truncated. In the standard CC models, the cluster operator is expanded by excitation levels

$$\begin{aligned} X &= X_1 + X_2 + X_3 + \dots \\ &= \sum_{ai} t_i^a \tau_i^a + \frac{1}{2!} \sum_{abij} t_{ij}^{ab} \tau_{ij}^{ab} + \frac{1}{3!} \sum_{abcijk} t_{ijk}^{abc} \tau_{ijk}^{abc} + \dots \end{aligned} \quad (2.24)$$

The doubles cluster operator,  $X_2$ , describes the interactions between pairs of electrons while  $X_3$  describes the simultaneous interaction between three electrons and so on. Higher level interactions are much less likely and the cluster operator can be truncated by leaving out excitations above a set level. For example, in the CCSD model developed by Purvis and Bartlett[2], only single and double excitations are included. Introducing the  $X_1$ -transformed Hamiltonian

$$\hat{H} = \exp(-X_1)H \exp(X_1), \quad (2.25)$$

the CCSD equations for a closed shell system become

$$\langle \mu_1 | \hat{H} + [\hat{H}, X_2] | \text{HF} \rangle = 0 \quad (2.26)$$

$$\langle \mu_2 | \hat{H} + [\hat{H}, X_2] + \frac{1}{2} [[\hat{H}, X_2], X_2] | \text{HF} \rangle = 0 \quad (2.27)$$

In Eqs. (2.26) and (2.27),  $\langle \mu_1 |$  and  $\langle \mu_2 |$  refers to single and double excited Slater determinants, respectively.

Even though  $X_4$ , with the excitations called the connected quadruples, is not included, the quadruply excited Slater determinants will still contribute to the CCSD wave function through disconnected excitations. A disconnected excitation is the product of lower connected excitations. For example,  $X_2^2$  is a quadruple excitation and can be interpreted as the interactions between two distinct electron pairs. Depending on the system, the disconnected terms can give a highly accurate description of the contributions from the highly excited Slater determinants[1].

CCSD combined with an approximate, non-iterative treatment of the triple excitations, called CCSD(T)[36], is often referred to as the gold standard of computational chemistry[37] due to the high accuracy obtained compared to computational cost. The simplest CC model is CC singles (CCS) which only includes the single excitations and is equivalent to HF theory. In CCSDT, the triples as well as the singles and doubles are included. Higher level models like CCSDTQ[38] and up to CCSDTQ567[39] have also been implemented. However, despite very high

accuracy, their computational cost is prohibitive for applications beyond benchmarking.

### 2.1.4 Approximate coupled cluster models

Many approximate CC models have been developed[4, 36, 40, 41]. In CC2[40], the CCSD equations are expanded using perturbation theory. The single excitations are considered zero order while the double excitations are considered first order.

Without explicitly writing out the nuclear potential, the Hamiltonian is separated into the zero order Fock operator,  $F$ , and the first order fluctuation potential,  $U$ .

$$H = F + U \quad (2.28)$$

In the canonical representation, the Fock operator is diagonal,

$$F = \sum_p \varepsilon_p a_p^\dagger a_p, \quad (2.29)$$

so the commutator between  $F$  and a cluster operator is

$$[F, X] = \sum_\mu \varepsilon_\mu t_\mu \tau_\mu \quad (2.30)$$

and higher commutators vanish.

As the single excitations describe approximate orbital relaxation and their equations are relatively few and simple, they are treated to infinite order. The doubles, on the other hand, are only solved to first order. This leads to the approximate CCSD equations

$$\langle \mu_1 | \hat{H} + [\hat{H}, X_2] | \text{HF} \rangle = 0 \quad (2.31)$$

$$\langle \mu_2 | \hat{H} + [F, X_2] | \text{HF} \rangle = 0 \quad (2.32)$$

In the similar CC3 model[41, 42], the singles and doubles are included to infinite order while the triples are treated perturbatively. The CCS, CC2, CCSD, CC3 and CCSDT models form a hierarchy of increasing accuracy and computational complexity, scaling as  $N^4$ ,  $N^5$ ,  $N^6$ ,  $N^7$ , and  $N^8$  respectively where  $N$  is the number of orbitals[40].

## 2.2 Multi-level coupled cluster theory

### 2.2.1 Cholesky decomposition

Multi-level coupled cluster models treat different parts of the system with different models from the CC hierarchy. This makes it possible to treat the most important parts at a higher level of accuracy without paying the full computational cost for the whole system. To achieve this, the spin-orbitals of the system must be assigned to different subspaces of the orbital space. For example, in the ECC2[26], model, a number of occupied and virtual orbitals are assigned to an active space, while the rest are considered an inactive space. There are several methods to assign orbitals. The simplest is to use orbital energies and assign the highest energy occupied (HOMO) and lowest energy unoccupied orbital (LUMO) to the active space. Another method is to use Cholesky decomposition to generate localised orbitals[21, 22] and use these to assign localised active spaces.

The one-electron density matrix in the non-orthogonal atomic orbital (AO) basis, assuming real orbitals, is positive semi-definite and symmetrical and the Cholesky decomposition is well defined[21]. The Cholesky decomposition of a matrix  $\mathbf{A}$  is written as the matrix product of a lower triangular matrix  $\mathbf{L}$  and its transpose[43]

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T. \tag{2.33}$$

The elements in  $\mathbf{L}$ , when  $\mathbf{A}$  has dimension  $N$ , are given by

$$L_{ii} = \left( a_{ii} - \sum_{k=1}^{i-1} L_{ik}^2 \right)^{\frac{1}{2}} \quad (2.34)$$

and

$$L_{ji} = \frac{1}{L_{ii}} \left( a_{ij} - \sum_{k=1}^{i-1} L_{ik} L_{jk} \right)^{\frac{1}{2}}, j = i + 1, i + 2, \dots, N. \quad (2.35)$$

The set of vectors  $\mathbf{L}_i$  generated this way are referred to as Cholesky vectors.

The one-electron density matrix in the AO basis, written in terms of the MO coefficients is

$$D_{\alpha\beta} = \sum_i^{occ} C_{\alpha i} C_{\beta i} \quad (2.36)$$

where  $\alpha$  and  $\beta$  refers to AOs and  $i$  refers to occupied MOs. A pseudo-density matrix is defined for the virtual orbitals as

$$D_{\alpha\beta}^V = \sum_a^{virt} C_{\alpha a} C_{\beta a} \quad (2.37)$$

For positive definite matrices, the number of Cholesky vectors will be equal to  $N$ . However, the one-electron density matrix will almost always be positive semi-definite. In these cases, a threshold is defined and the decomposition algorithm is stopped when there are no diagonal elements greater than this threshold. This will generate less than  $N$  vectors.

The density matrix in the AO basis is inherently local and as a result, it is sparse. When performing Cholesky decomposition, this property is retained and elements of the Cholesky vectors  $\tilde{C}_{\alpha i}$  can be viewed as the coefficients of a new set of localised, occupied MOs.

$$D_{\alpha\beta} = \sum_i^{occ} \tilde{C}_{\alpha i} \tilde{C}_{\beta i} \quad (2.38)$$

Performing the decomposition on the pseudo-density matrix generates the equivalent virtual MOs. By construction, these MOs will be orthonormal.

To avoid mathematical instability, pivoting is performed based on the diagonal elements. Usually, the pivoting is based on the size of the diagonal elements with the largest element selected for pivoting. For positive semi-definite matrices, unlike positive definite ones, the decomposition is not unique but depends on the pivoting. Assuming no ghost orbitals, each diagonal element corresponds to an AO centred on an atom. To generate a localised active space, a number of atoms considered to be of special interest are designated as active. The density matrix is then decomposed using the diagonals corresponding to AOs centred on active atoms for pivoting until none of these elements greater than the threshold remains. The resulting MOs will form the occupied orbitals in a localised active space. The same procedure on the pseudo-density matrix generates the virtual orbitals. Performing the decomposition on the residual matrix will generate another set that will form one or more inactive spaces.

A possible disadvantage with this method is that it may generate an unbalanced set as orbitals describing bonds between atoms may be assigned entirely to one atom. To counter this, one may limit the number of orbitals in the active space or perform the decomposition on an atom-by-atom basis[22].

Unlike other localisation schemes like Boys[44], Edmiston-Ruedenberg[45] and Pipek-Mezey[46], Cholesky decomposition is noniterative and can be made to scale linearly. However, the orbitals obtained are not as local as those obtained using the trust region minimisation procedure developed by Høyvik et al.[47].

### **2.2.2 Extended CC2**

Assuming the orbitals have been assigned to spaces, the excitation manifold can be split into sub-manifolds. Which sub-manifold an excitation is assigned to, depends on which orbitals are involved in the excitation. Figure 2.1 demonstrates the classification of excitations. If an excitation only involves orbitals from one space, it is considered internal in that space while an excitation involving no

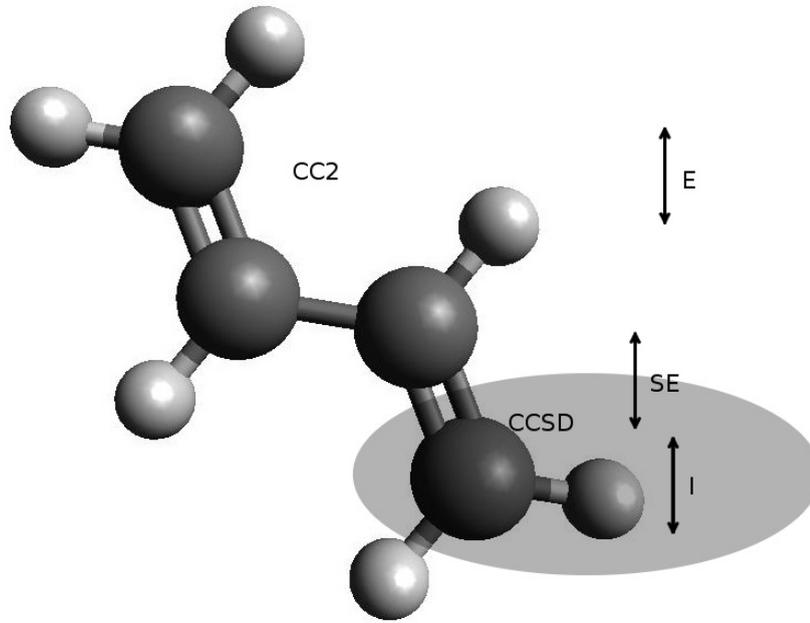


FIGURE 2.1: Example of an active space of 1,3-butadiene. I denotes internal, SE semi-external and E external excitations with respect to the active CCSD space.

orbitals from a space is external to that space. If it involves orbitals from several spaces, it is referred to as semi-external to those spaces.

In the ECC2 model[26], only two spaces are assigned, one active to be treated with CCSD and one inactive to be treated with CC2. The excitation manifold is then divided into two submanifolds  $\{\langle\mu^S|\}$ , which contains excitations internal and semi-external to the inactive space, and  $\{\langle\mu^T|\}$  with the excitations internal to the active space. The cluster operator,  $X$ , is then split in two.

$$|\text{CC}\rangle = \exp(X)|\text{HF}\rangle = \exp(T + S)|\text{HF}\rangle. \quad (2.39)$$

Giving the CC equations

$$\langle\mu^T| \exp(-T - S)H \exp(T + S)|\text{HF}\rangle = 0 \quad (2.40)$$

$$\langle\mu^S| \exp(-T - S)H \exp(T + S)|\text{HF}\rangle = 0 \quad (2.41)$$

Amplitudes in  $T$  are treated to infinite order, while those in  $S$  will be treated in an approximate fashion. Expanding  $S$  in orders of the perturbation yields

$$S = S^{(0)} + S^{(1)} + S^{(2)} + \dots = \sum_{\mu} s_{\mu}^{(0)} \tau_{\mu} + \sum_{\mu} s_{\mu}^{(1)} \tau_{\mu} + \sum_{\mu} s_{\mu}^{(2)} \tau_{\mu} + \dots \quad (2.42)$$

where  $\mu$  runs over the excitations in  $\{\langle \mu^S \rangle\}$ . Splitting the Hamiltonian as in Eq. (2.28), the amplitude equations for  $S$  becomes

$$\varepsilon_{\mu} s_{\mu}^{(0)} = 0 \quad (2.43)$$

$$\varepsilon_{\mu} s_{\mu}^{(1)} = \langle \mu | \exp(-T) U \exp(T) | \text{HF} \rangle \quad (2.44)$$

$$\varepsilon_{\mu} s_{\mu}^{(2)} = \langle \mu | \exp(-T) [U, S^{(1)}] \exp(T) | \text{HF} \rangle \quad (2.45)$$

In both CC2 and CCSD, the single excitations are treated to infinite order, so the single excited determinants in ECC2 are referred to as  $\langle \mu_1 |$ , regardless of which submanifold they belong to. The CC equations for the singles amplitudes are the same as for CCSD.

$$\langle \mu_1 | \hat{H} + [\hat{H}, X_2] | \text{HF} \rangle = 0 \quad (2.46)$$

For the same reason, the equations for the amplitudes in  $T_2$  are also the same as in CCSD

$$\langle \mu_2^T | \hat{H} + [\hat{H}, X_2] + \frac{1}{2} [[\hat{H}, X_2], X_2] | \text{HF} \rangle = 0 \quad (2.47)$$

Excitations in  $\{\langle \mu_2^S \rangle\}$  must involve at least one orbital not in the active space, so

$$\langle \mu_2^S | [[\hat{H}, T_2], T_2] | \text{HF} \rangle = 0. \quad (2.48)$$

In standard CC perturbation theory, all double excitations are considered first order in the perturbation. This is not the case in ECC2, where the excitations in  $T_2$  are considered zero order. This leads to an additional commutator term,  $[\hat{H}, T_2]$ , in the  $S_2$ -amplitude equations

$$\langle \mu_2^S | [F, S_2] + \hat{H} + [\hat{H}, T_2] | \text{HF} \rangle = 0 \quad (2.49)$$

compared to the standard CC2 equations(2.32). The additional computational complexity from the commutator term is low due to the relatively small number of  $T_2$  amplitudes, but it leads to an increased accuracy as demonstrated in Chapter 4.

The advantage of the ECC2 model compared with CCSD is the reduced computational scaling while retaining a comparable accuracy (see Chapter 4). In CCSD, the most demanding term, the B-term, scales as  $V^4O^2$  where  $V$  is the number of virtual and  $O$  the number of occupied orbitals[1, 48]. In ECC2, the B-term scales simply as  $kV^2$ , where  $k$  is a prefactor scaling as  $V_A^2O_A^2$ .  $V_A$  and  $O_A$  denotes virtual and occupied orbitals in the active space. As the active space should contain only a fraction of the total number of orbitals, this greatly reduces computational complexity. For systems above a certain size, ECC2 will scale as CC2.

### 2.2.3 Beyond ECC2

ECC2 is a relatively simple model that only contains two spaces and two levels of theory. More advanced models can be formulated that contain several spaces and levels of theory. In these models, more orbital spaces are defined and the same scheme for assignment of excitations is used as the one in ECC2. Internal excitations are assigned to their space, while semi-external excitations are assigned to to the lowest level space they include orbitals from.

Including a space at the CCS level is the easiest extension of the ECC2 model. In all three standard models, CCS, CC2 and CCSD, the singles amplitudes are treated to infinite order. This means that the singles in a CCS space are treated the same way as in the other spaces and Eq. (2.46) is still valid. The double excitations internal and semi-external to the CCS space are simply set to zero. In addition to a CCS space, an SCF space can also be included where all cluster amplitudes are set to zero.

To include triple excitations, the best models to use are CCSD(T)[36] or CC3[41]. In such a model, up to five orbital spaces, denoted  $P, Q, R, S$  and  $T$  are required. The  $P$ -space is only treated at the SCF level, so all amplitudes in the corresponding

part of the cluster operator are set to zero. Eqs. (2.50-2.53) summarise the splitting of the cluster operator.

$$X = X_1 + X_2 + X_3 \quad (2.50)$$

$$X_1 = Q_1 + R_1 + S_1 + T_1 \quad (2.51)$$

$$X_2 = R_2 + S_2 + T_2 \quad (2.52)$$

$$X_3 = T_3 \quad (2.53)$$

In the standard CCSD(T) model, the amplitude equations are the same as for the CCSD and this is also true for the MLCC formulation, giving the following amplitude equations.

$$\langle \mu_1 | \hat{H} + [\hat{H}, X_2] | \text{HF} \rangle = 0 \quad (2.54)$$

$$\langle \mu_2^R | \hat{H} + [F, R_2] + [\hat{H}, S_2 + T_2] | \text{HF} \rangle = 0 \quad (2.55)$$

$$\langle \mu_2^S | \hat{H} + [\hat{H}, X_2] + [[\hat{H}, X_2] X_2] | \text{HF} \rangle = 0 \quad (2.56)$$

$$\langle \mu_2^T | \hat{H} + [\hat{H}, X_2] + [[\hat{H}, X_2] X_2] | \text{HF} \rangle = 0 \quad (2.57)$$

Perturbation theory is used to calculate the  $T_3$  amplitudes. As the CCSD double excitations are considered zero order, only the first order triple amplitudes are needed. The double commutator term between the double excitation operators and  $U$  is not included in current implementations due to its high computational complexity. This term could be important and the effect of including it should be investigated. Computational simplification will still be achieved as the projection manifold will be restricted to the T-manifold.

$$\varepsilon_{\mu_3^T} t_{\mu_3^T}^{(1)} = \langle \mu_3^T | [U, S_2 + T_2] | \text{HF} \rangle \quad (2.58)$$

In the resulting energy correction,  $E_{corr}^{CCSD(T)}$ , only the first order terms are retained. This is another difference from standard CCSD(T) where both fourth and

fifth order terms are retained.

$$E_{corr}^{CCSD(T)} = \sum_{\mu_1^T, \mu_2^T, \mu_2^S} t_\mu \langle \mu | [U, T_3] | HF \rangle \quad (2.59)$$

In the CC3 case, the singles and doubles are included unperturbatively while the triples are included perturbatively, analogously to doubles in CC2. Only the singles equations from the T-space will be affected as the Hamiltonian does not contain higher than two-electron operators. The  $R_2$  and  $T_3$  operators does not appear in each others amplitude equations as they are first order and the equations only solved to first order.

$$0 = \langle \mu_1^{Z \neq T} | \hat{H} + [\hat{H}, X_2] | HF \rangle \quad (2.60)$$

$$0 = \langle \mu_1^T | \hat{H} + [\hat{H}, X_2] + [\hat{H}, T_3] | HF \rangle \quad (2.61)$$

$$0 = \langle \mu_2^R | \hat{H} + [F, R_2] + [\hat{H}, S_2 + T_2] | HF \rangle \quad (2.62)$$

$$0 = \langle \mu_2^S | \hat{H} + [\hat{H}, X_2] + [[\hat{H}, X_2], X_2] | HF \rangle \quad (2.63)$$

$$0 = \langle \mu_2^T | \hat{H} + [\hat{H}, X_2] + [[\hat{H}, X_2], X_2] | HF \rangle \quad (2.64)$$

$$0 = \langle \mu_3^T | [\hat{H}, S_2 + T_2] + [[\hat{H}, S_2 + T_2], S_2 + T_2] + [F, T_3] | HF \rangle \quad (2.65)$$

## 2.3 MLCC response theory

### 2.3.1 Quasi-energy response method

As amplitudes in approximate CC models are not obtained through projection, it is not possible to obtain response functions with the generalised Hellmann-Feynmann theorem[42]. However, it is possible to derive them as derivatives of the QE Lagrangian[30, 42, 49, 50]. The QE method assumes that the Hamiltonian can be written as

$$H = H_0 + V^t \quad (2.66)$$

where  $V^t$  is a periodic, time-dependent perturbation with period  $\mathcal{T}$ . Furthermore, to ensure hermicity,  $V^t$  has the Fourier transform

$$V^t = \sum_{j=-N}^N \sum_A A \epsilon_A(\omega_j) \exp(-i\omega_j t) \quad (2.67)$$

where  $A = A^\dagger$  and is a real frequency-independent operator. Furthermore,  $\omega_{-j} = -\omega_j$  and  $(\epsilon_A(\omega_j))^* = \epsilon_A(-\omega_j)$ .

In time-dependent theory, the CC ansatz is somewhat modified.

$$|\text{CC}(t)\rangle = \exp(X(t))|\text{HF}\rangle \exp(i\alpha(t)) \quad (2.68)$$

Where  $\alpha$  is a generally complex phase factor and the time-dependent cluster operator  $X(t)$  is written as

$$X(t) = \sum_{\mu} t_{\mu}(t) \tau_{\mu} \quad (2.69)$$

The corresponding dual type state[29, 41] is

$$\langle \Lambda | = \left( \langle \text{HF} | + \sum_{\mu} \bar{t}_{\mu}(t) \langle \mu | \exp(-T(t)) \right) \exp(-\alpha(t)) \quad (2.70)$$

where  $\{\bar{t}_{\mu}(t)\}$  are the Lagrange multipliers corresponding to the  $t_{\mu}$  amplitudes.

If the CC wave function is a solution of the time dependent Schrödinger equation, it satisfies

$$\langle \Lambda | \left[ H_0 + V^t - i \frac{d}{dt} \right] |\text{CC}(t)\rangle = 0 \quad (2.71)$$

The QE Lagrangian is defined as

$$L(t) = \text{Re} \left( \langle \tilde{\Lambda} | \left[ H_0 + V^t - i \frac{d}{dt} \right] |\tilde{\text{CC}}(t)\rangle \right) \quad (2.72)$$

with

$$|\tilde{\text{CC}}(t)\rangle = \exp(X(t))|\text{HF}\rangle \quad (2.73)$$

and

$$\langle \tilde{\Lambda} | = \langle \text{HF} | + \sum_{\mu} \bar{t}_{\mu}(t) \langle \mu | \exp(-T(t)) \quad (2.74)$$

The QE Lagrangian is not variational, but the term differing from Eq. (2.71) is proportional to  $\dot{\alpha}$ . Because the time-dependent perturbation is periodic, so is  $\alpha$  and the time averaged QE Lagrangian is variational.

$$\delta \{L(t)\}_{\mathcal{T}} = \delta \left( \frac{1}{\mathcal{T}} \int_{t_0}^{t_0+\mathcal{T}} L(t) dt \right) = 0 \quad (2.75)$$

Expanding the CC amplitudes

$$t_{\mu}(t) = t_{\mu}^{(0)} + t_{\mu}^{(1)}(t) + t_{\mu}^{(2)}(t) + \dots \quad (2.76)$$

with

$$\begin{aligned} t_{\mu}^{(1)}(t) &= \sum_{j=-N}^N t_{\mu}^{(1)}(\omega_j) \exp(-i\omega_j t) \\ &= \sum_{j=-N}^N \sum_A t_{\mu}^A(\omega_j) \epsilon_A(\omega_j) \exp(-i\omega_j t) \end{aligned} \quad (2.77)$$

The Lagrange multipliers can be expanded in a similar manner. Due to the  $2n+1$  and  $2n+2$  rules[51, 52], higher order amplitudes and multipliers does not affect the linear response functions as they are second order. Expanding the QE Lagrangian similarly

$$L = L^{(0)} + L^{(1)} + L^{(2)} + \dots \quad (2.78)$$

the variational condition is then expressed as

$$\frac{\partial}{\partial t_{\mu}^{(m)}(t)} \{L^{(n)}(t)\}_{\mathcal{T}} = \frac{\partial}{\partial \bar{t}_{\mu}^{(m)}(t)} \{L^{(n)}(t)\}_{\mathcal{T}} = 0, \quad m \leq n \quad (2.79)$$

The linear response function can then be written as the derivative of the the time-averaged QE Lagrangian with respect to the perturbation strength parameters[29, 30].

$$\langle\langle A, B \rangle\rangle_{\omega_j} = \frac{\partial^2 \{L^{(2)}(t)\}_{\mathcal{T}}}{\partial \epsilon_A(-\omega_j) \partial \epsilon_B(\omega_j)} \quad (2.80)$$

### 2.3.2 Linear response function for ECC2

To introduce time dependent perturbation theory in the ECC2 model, the Hamiltonian is split into three parts.

$$H = F + U + V^t \quad (2.81)$$

The  $S$ -operator is first order in both  $U$  and  $V^t$ , while  $T$  is still zero order in  $U$ , but first order in  $V^t$ [40, 42].

Inserting the ECC2 equations (2.39) into the definition (2.72), the explicit QE Lagrangian is

$$\begin{aligned} L(t) = & \langle HF | H \exp(T + S) | HF \rangle \\ & + \sum_{\mu_1} \bar{t}_{\mu_1} \left( \langle \mu_1 | \hat{H} + [\hat{H}, X_2] | HF \rangle - i \frac{dt_{\mu_1}}{dt} \right) \\ & + \sum_{\mu_2^T} \bar{t}_{\mu_2^T} \left( \langle \mu_2^T | \hat{H} + [\hat{H}, X_2] + \frac{1}{2} [[\hat{H}, X_2], X_2] | HF \rangle - i \frac{dt_{\mu_2^T}}{dt} \right) \\ & + \sum_{\mu_2^S} \bar{t}_{\mu_2^S} \left( \langle \mu_2^S | [F + \hat{V}, S_2] + \hat{H} + [\hat{H}, T_2] | HF \rangle - i \frac{dt_{\mu_2^S}}{dt} \right) \end{aligned} \quad (2.82)$$

Derivation of the amplitudes and multipliers from the Lagrangian variational principle is tedious and the reader is referred to Appendix A for details. Only the results are presented here. The zero order amplitudes are those from the time-independent theory, so only the first order amplitudes and zero order multipliers are needed for the linear response function.

$$\bar{\mathbf{t}}^0 \mathbf{A} = \boldsymbol{\eta}^0 \quad (2.83)$$

$$(\omega \mathbf{1} - \mathbf{A}) \mathbf{t}^A(\omega) = \boldsymbol{\xi}^A \quad (2.84)$$

In Eqs. (2.83) and (2.84),  $\bar{\mathbf{t}}^0$  is a row vector with the zero order Lagrange multipliers while  $\mathbf{t}^A(\omega)$  is a column vector with the first order amplitudes defined in

Eq. (2.77). The elements in  $\boldsymbol{\eta}^0$  are given by

$$\eta_{\nu_i}^0 = \langle HF | [\hat{H}_0, \tau_{\nu_i}] | HF \rangle \quad (2.85)$$

and

$$\boldsymbol{\xi}^A = \begin{pmatrix} \langle \mu_1 | \hat{A} + [\hat{A}, X_2^{(0)}] | HF \rangle \\ \langle \mu_2^T | [\hat{A}, X_2^{(0)}] | HF \rangle \\ \langle \mu_2^S | [\hat{A}, X_2^{(0)}] | HF \rangle \end{pmatrix} \quad (2.86)$$

where  $\hat{A}$  is the  $T_1$ -transformed one-electron operators from Eq. (2.67).  $\mathbf{A}$  is the ECC2 Jacobian.

$$\mathbf{A} = \begin{pmatrix} \langle \mu_1 | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_1] | HF \rangle & \langle \mu_1 | [\hat{H}_0, \nu_2^T] | HF \rangle & \langle \mu_1 | [\hat{H}_0, \nu_2^S] | HF \rangle \\ \langle \mu_2^T | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_1] | HF \rangle & \langle \mu_2^T | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_2^T] | HF \rangle & \langle \mu_2^T | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_2^S] | HF \rangle \\ \langle \mu_2^S | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_1] | HF \rangle & \langle \mu_2^S | [\hat{H}_0, \nu_2^T] | HF \rangle & \langle \mu_2^S | [F, \nu_2^S] | HF \rangle \end{pmatrix} \quad (2.87)$$

In Eq. (2.87),  $\tau_{\nu_i^Z}$  is written as  $\nu_i^Z$ . The ECC2 Jacobian is very similar to that in CC2[40], but it has additional terms in the elements corresponding to the  $T_2$  amplitudes and the additional commutator term in the  $S_2$  amplitude equations appears in the  $(\mu_2^S, \nu_1)$  and  $(\mu_2^S, \nu_2^T)$  elements.

Eq. (2.84) will be singular if  $\omega$  is an eigenvalue of  $\mathbf{A}$  and the first order amplitudes will go to infinity. These are the poles in the response function, so the eigenvalues of  $\mathbf{A}$  correspond to the excitation energies of the system[40].

The linear response function is

$$\begin{aligned}
\langle\langle A, B \rangle\rangle_{\omega_i} &= \frac{\partial\{L^{(2)}\}_{\mathcal{T}}}{\partial\epsilon_A(-\omega_i)\partial\epsilon_B(\omega_i)} \\
&= P(A(-\omega_i), B(\omega_i)) \left\{ \langle HF | [\hat{A}, T_1^{(B)}] + \frac{1}{2} [[\hat{H}_0, T_1^{(A)}], T_1^{(B)}] | HF \rangle \right. \\
&\quad + \sum_{\mu_1} \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | [\hat{A}, X^{(B)}] + \frac{1}{2} [[\hat{H}_0, X^{(A)}], X^{(B)}] | HF \rangle \\
&\quad + \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [\hat{A}, X_2^{(B)}] + \frac{1}{2} [[\hat{H}_0, T_1^{(A)}], T_1^{(B)}] + [[\hat{A}, T_1^{(B)}], T_2^{(0)}] \\
&\quad + [[\hat{H}_0, T_1^{(A)}], T_2^{(B)}] + \frac{1}{2} [[[\hat{H}_0, T_1^{(A)}], T_1^{(B)}], T_2^{(0)}] | HF \rangle \\
&\quad + \sum_{\mu_2^T} \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | [\hat{A}, X_2^{(B)}] + \frac{1}{2} [[\hat{H}_0, T_1^{(A)}], T_1^{(B)}] \\
&\quad + [[\hat{A}, T_1^{(B)}], X_2^{(0)}] + [[\hat{H}_0, T_1^{(A)}], X_2^{(B)}] \\
&\quad \left. + \frac{1}{2} [[[\hat{H}_0, T_1^{(A)}], T_1^{(B)}], X_2^{(0)}] + \frac{1}{2} [[\hat{H}_0, X_2^{(A)}], X_2^{(B)}] | HF \rangle \right\} \tag{2.88}
\end{aligned}$$

where  $T^{(A)}(\omega_i) = \partial T^{(1)}(\omega_i) / \partial \epsilon_A(\omega_i)$  is the derivative of the frequency dependent first order cluster operator

$$T^{(1)}(t) = \sum_{i=-N}^N T^{(1)}(\omega_i) \exp(-i\omega_i t) \tag{2.89}$$

and  $P(x, y)f(x, y) = f(x, y) + f(y, x)$ . In Eq. (2.88), the frequency dependence of the cluster operators has been suppressed.

Compared to the CC2 response function[40], additional terms appear in Eq. (2.88). This is due to the appearance of double excitations that are zero order in  $U$ . In standard theory, all terms containing  $U$  and  $T_2$  are at least second order and disappear. In ECC2, this is no longer the case and these terms are retained as well as terms containing two double commutators.

# Chapter 3

## Implementation

### 3.1 ECC2 energy calculations

To further study the viability of MLCC methods, the ECC2 model has been implemented in a pre-release version of the DALTON 2013 software package[33, 34]. The focus of the implementation is proof of principle and to investigate the accuracy of the ECC2 model compared to CC2 and CCSD. The pilot code is quite primitive and options for frozen core[53] and symmetry are not yet implemented. Thus, for the time being, reduced computational complexity cannot be investigated. In fact, the implementation uses more time than equivalent CCSD calculations.

The basis for the ECC2 pilot code is the CCSD implementation in DALTON by Koch[54, 55]. As noted previously, the  $T_1$  and  $T_2$  amplitude equations, Eqs. (2.46) and (2.47), are the same as for CCSD. The CC equations are solved in an iterative fashion and in each iteration,  $\langle \mu | \exp(-X) H \exp(X) | \text{HF} \rangle$  is initially calculated as in CCSD.

Inspecting Eq. (2.49), the terms involving  $S_2$ -amplitudes are the same that would appear in CC2. To obtain these, the variable indicating which model to use is set to be CC2 and the  $T_2$ -amplitudes temporarily set to be zero before calling the subroutine that calculates value of the terms. The commutator term,  $[\hat{H}, T_2]$ , is calculated by restoring the  $T_2$ -amplitudes, setting the  $S_2$ -amplitudes to zero and

resetting the model variable to CCSD. Adding this to the previous calculation gives a  $\langle \mu_S^2 | \hat{H} | \text{HF} \rangle$ -term too much. This term is then calculated alone by setting all amplitudes to zero. Subtracting this from the previous results gives the correct value. The terms calculated using only CCSD are then overwritten and the amplitudes restored before continuing the iteration in the usual fashion.

Unless otherwise stated, ECC2 refers to the model described above, but two alternative models are also implemented. In the ECC2-II, as opposed to ECC2-I described above, excitations semi-external to the active space are included in  $T_2$ . This is done by setting the semi-external excitation amplitudes to zero and restoring them together with the  $T_2$ -amplitudes. In the second variation, called ECC2a, the commutator term,  $[\hat{H}, T_2]$  is not included and the iteration proceeds normally after calculating the CC2 amplitudes for  $S_2$ . If the commutator is included, the model is referred to as ECC2b.

The ECC2 model is also implemented with a CCS extension. This is achieved by setting double excitations corresponding to the CCS space to zero for all calculations.

## 3.2 ECC2 excitation energies

As discussed in Chapter 2, the excitation energies of a system can be found by computing the eigenvalues of the Jacobian matrix. In DALTON, originally implemented by Koch[56] and the current version was by Christiansen[57], the Davidson algorithm[58] is employed to find the lowest eigenvalues. This algorithm works by solving the eigenvalues of the  $n \times n$  matrix  $\mathbf{A}$  in a reduced space  $K$  spanned by the  $k$  orthonormal vectors  $\{v_m\}$ . The interaction matrix  $H_k$  is then defined as

$$H_k = V_k^T \mathbf{A} V_k \quad (3.1)$$

where  $V_k = [v_1, v_2, \dots, v_k]$  is a  $n \times k$  matrix.

For a sparse and diagonally dominant matrix, such as the CC Jacobian,  $k$  will be much smaller than  $n$ , so finding the eigenvalues  $\tilde{\lambda}_k$  and eigenvectors  $y_k$  of  $H_k$  is a much simpler task than for the full matrix. Spanning  $y$  out in the whole space gives an approximate eigenvector  $\tilde{u}$  for  $\mathbf{A}$

$$\tilde{u} = V_k y_k \quad (3.2)$$

Convergence is tested by calculating the norm of the residual

$$r_k = \left( \tilde{\lambda}_k \mathbf{I} - \mathbf{A} \right) \tilde{u} \quad (3.3)$$

If the algorithm is not converged, a new vector  $t_{k+1}$  is generated

$$t_{k+1} = \mathbf{M}^{-1} r_k \quad (3.4)$$

and orthonormalised with respect to  $\{v_m\}$  before being added as a column in  $V_{k+1}$ . The preconditioner  $\mathbf{M}$  is an easily invertible approximation of  $\left( \tilde{\lambda}_k \mathbf{I} - \mathbf{A} \right)$ , usually  $\left( \tilde{\lambda}_k \mathbf{I} - \mathbf{D} \right)$ , where  $\mathbf{D}$  is the diagonal of  $\mathbf{A}$ . For the ECC2 implementation, the subroutine for generating new vectors was not changed, so the preconditioner used contained the CCSD diagonal which might have affected the number of iterations.

In DALTON, the Davidson algorithm is initiated by a set of  $k$  guess vectors  $\{v_m\}$  where  $k$  is the number of excitation energies to be calculated. All the elements of  $v_k$  are zero except the element corresponding to the  $k$ th largest diagonals in the Jacobian  $\mathbf{A}$ . This is a good start guess due to sparsity and diagonal dominance. Again, the CCSD diagonal is used for the ECC2 model, however, the diagonal elements corresponding to single excitations are likely to be the largest and these are the same for both models.

$\mathbf{A}$  appears in Eqs. (3.1) and (3.3), both times on the form  $\mathbf{A}v$  and the transformed vector will be referred to as  $\rho$ .

$$\begin{pmatrix} \rho_1 \\ \rho_{T_2} \\ \rho_{S_2} \end{pmatrix} = \mathbf{A} \begin{pmatrix} v_1 \\ v_{T_2} \\ v_{S_2} \end{pmatrix} \quad (3.5)$$

The CCSD[28] and CC2[40] Jacobians are

$$\mathbf{A}_{CCSD} = \begin{pmatrix} \langle \mu_1 | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_1] | HF \rangle & \langle \mu_1 | [\hat{H}_0, \nu_2] | HF \rangle \\ \langle \mu_2 | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_1] | HF \rangle & \langle \mu_2 | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_2] | HF \rangle \end{pmatrix} \quad (3.6)$$

$$\mathbf{A}_{CC2} = \begin{pmatrix} \langle \mu_1 | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \nu_1] | HF \rangle & \langle \mu_1 | [\hat{H}_0, \nu_2] | HF \rangle \\ \langle \mu_2 | [\hat{H}_0, \nu_1] | HF \rangle & \langle \mu_2 | [F, \nu_2] | HF \rangle \end{pmatrix} \quad (3.7)$$

Comparing Eqs. (2.87) and (3.6), the equations for  $\rho_1$  and  $\rho_{T_2}$  are the same for ECC2 and CCSD. Thus, as for the energy, the iteration starts with calculating the CCSD terms. Calculating  $\rho_{S_2}$  is more complicated and is done in three separate steps. First,  $v_1$  and  $v_{T_2}$  is set to zero and the model to CC2 to obtain the contribution from the lower right part of  $\mathbf{A}$ . For the next steps, the model is reset to CCSD. The lower left part contribution is obtained by setting  $v_{T_2}$  and  $v_{S_2}$  and all  $S_2$ -amplitudes to zero. Finally,  $v_1$  and  $v_{S_2}$  and all amplitudes are set to zero to obtain the commutators in the  $T_2, S_2$  part of  $\mathbf{A}$ .

In total, the transformation routine must be called four times for each iteration step. In addition, the amplitudes,  $v$  and  $\rho$  are read and written to disc several times and intermediates calculated and saved during the energy calculation must be recalculated. The result is an implementation that often takes ten times longer than normal CCSD. This greatly reduces the range of feasible applications, both in terms of system size and number of excitations.

# Chapter 4

## Results

### 4.1 ECC2 energy calculations

#### 4.1.1 Abstraction processes

Multi-level coupled cluster models are intended to calculate local properties in systems that are too large for a more conventional CC approach. In a typical system, the active space will only contain a fraction of the orbitals of the system and MLCC is not expected to perform particularly well with size-extensive properties such as total energy and polarisability. An exception is cases where perturbation theory diverges.

During abstraction processes, HOMO and LUMO will typically move closer in energy and become quasi-degenerate. As a result, the correction terms in perturbative models like CC2 and second-order Møller-Plesset theory (MP2) will approach singularities and the models fail[1, 4]. If HOMO, LUMO and possibly some of the orbitals closest in energy are treated with CCSD, the HOMO-LUMO gap observed by CC2 may become larger and singularities may be avoided.

Figure 4.1 presents a minimal example of such a system. In the figure, the energy of lithium hydride is calculated using various models is plotted against bond length.

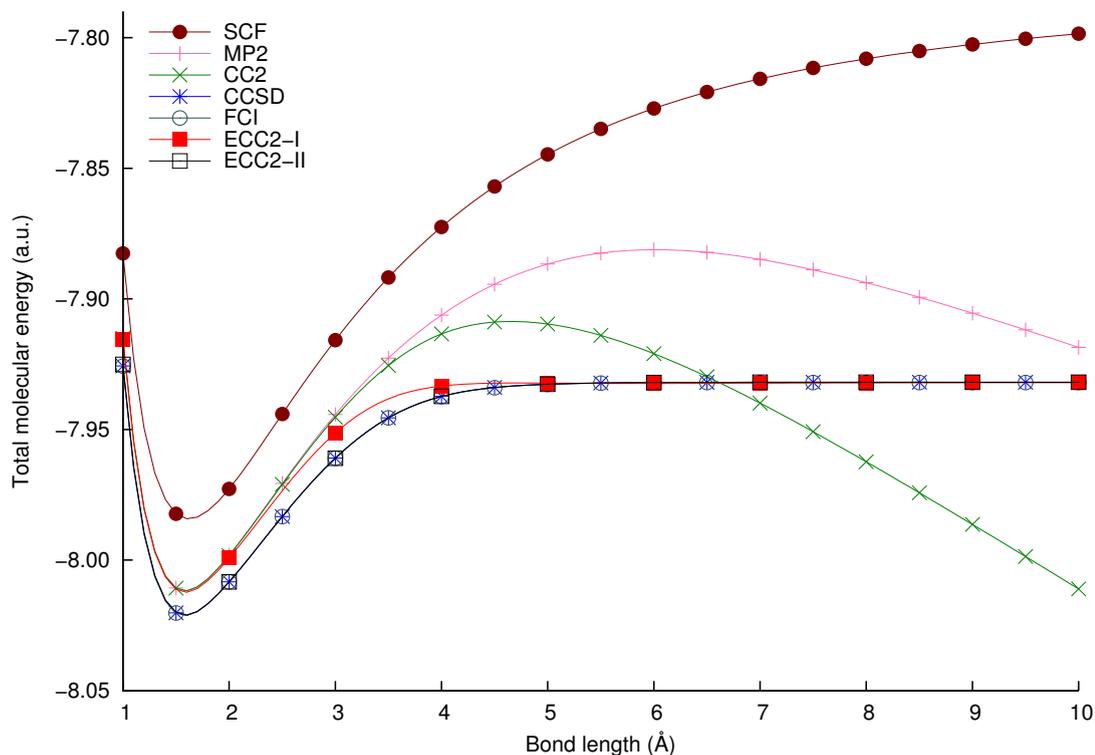


FIGURE 4.1: Total energy curves for dissociation of lithium hydride using the basis set aug-cc-pVDZ.

SCF far overestimates the energy because a single configuration state can only describe the dissociation into  $\text{Li}^+$  and  $\text{H}^-$  rather than  $\text{Li}$  and  $\text{H}$ [59]. For such a small system with only four electrons, linked triple and quadruple excitations are negligible, so CCSD and FCI are identical, consistent with results reported elsewhere[60]. At equilibrium geometry, the perturbative models MP2 and CC2 perform reasonably well, however, as the bond becomes stretched, the performance is drastically reduced.

Two ECC2 variations are included in Figure 4.1, ECC2-I and ECC2-II, described in Chapter 3. In both, the active space consists of HOMO and LUMO. As only excitations of the electrons in the lithium s1-orbital to orbitals higher in energy than LUMO are treated at the CC2 level, the ECC2-II model is almost identical with CCSD and FCI. At the equilibrium geometry, ECC2-I performs comparably to the perturbative models. At increased bond distances, however, both the ECC2 models converge to the FCI value.

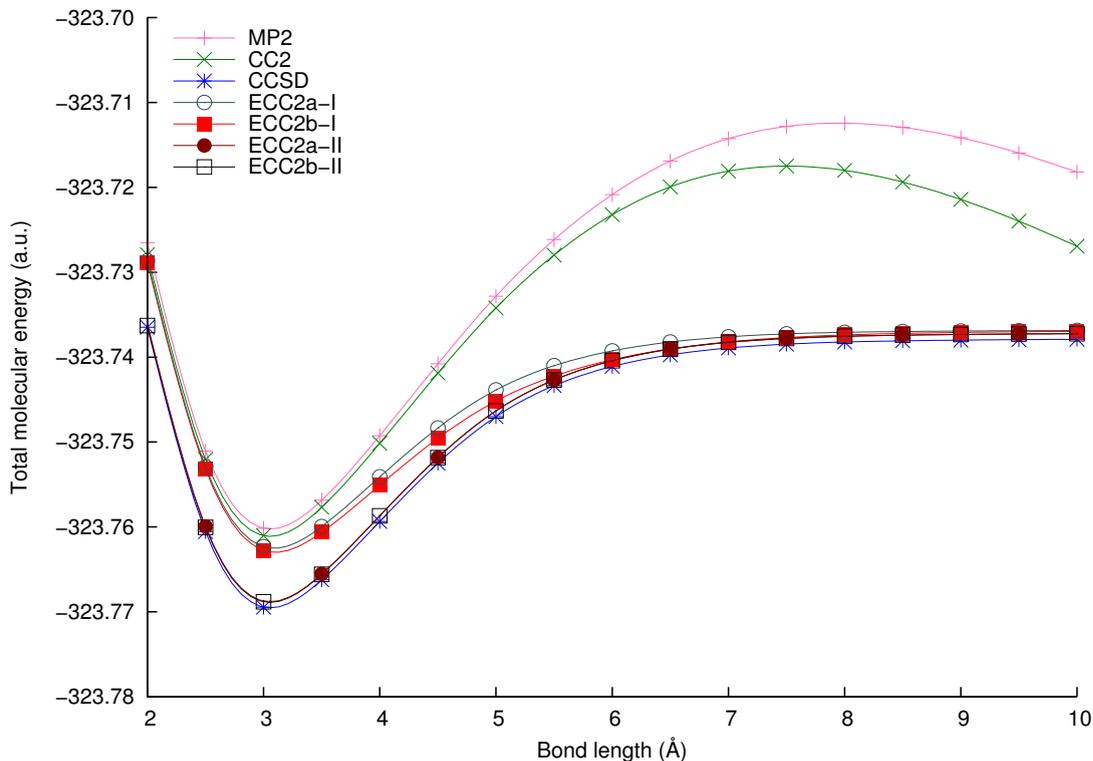


FIGURE 4.2: Total energy curves for dissociation of a sodium dimer using the basis set aug-cc-pVDZ.

A similar behaviour is observed for the sodium dimer in Figure 4.2. The active space consists of HOMO and LUMO and again ECC2-I performs comparably to the perturbative methods in the equilibrium geometry, but converges to the CCSD value when the perturbative methods fail for stretched bonds. The ECC2-II model is very close to the CCSD value for all geometries.

In addition to the I and II variations of ECC2 described above, Figure 4.2 also shows the performance of the ECC2a and ECC2b models. The description provided by ECC2b is slightly better for all geometries.

Figure 4.3 shows the energy curves of ethene during abstraction of one of the hydrogen atoms. From the initial geometry (Table B.1), a hydrogen atom was moved along the direction of its bond while the other atoms were held in place. During the abstraction process, the nature of HOMO and LUMO change from the  $\pi$ -orbitals of the carbon-carbon double bond to the bonding and anti-bonding  $\sigma$ -orbitals of the carbon-hydrogen bond. Consequently, only including HOMO and

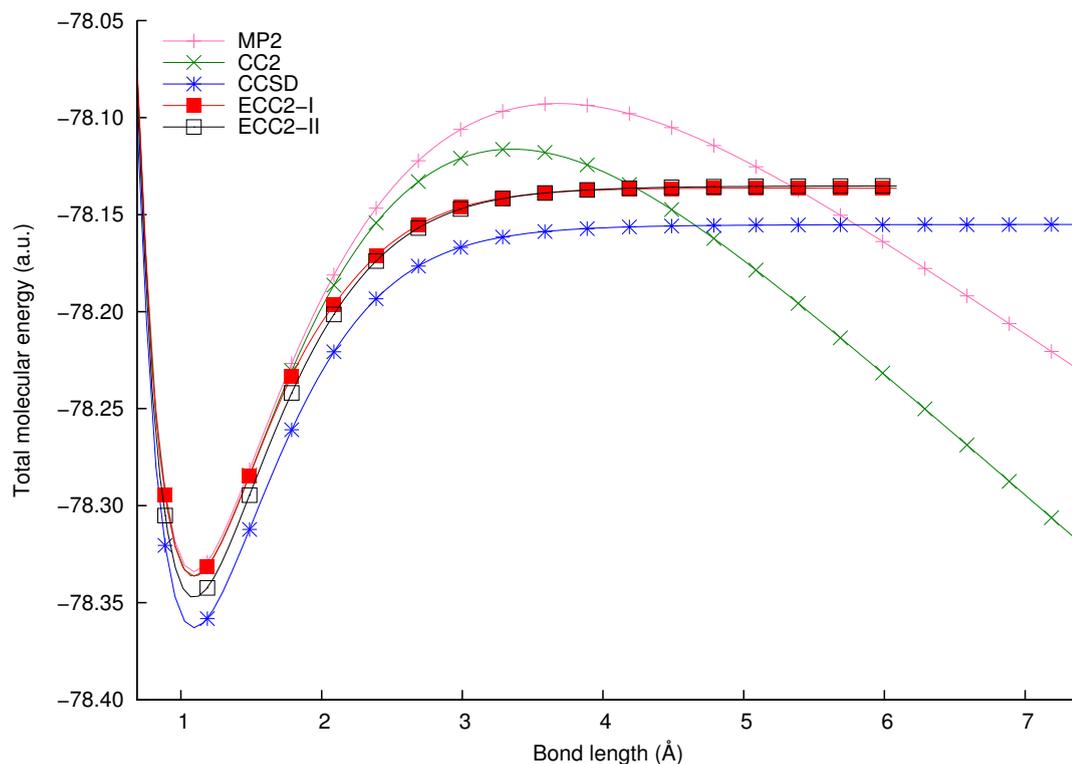


FIGURE 4.3: Total energy curves for dissociation of a hydrogen atom from ethene using the basis set aug-cc-pVDZ.

LUMO in the active space results in an unphysical drop in the total energy during abstraction. To remedy this problem, the two highest occupied and two lowest virtual orbitals are included in the active space in Figure 4.3.

### 4.1.2 Local geometry

Another possible application for MLCC is the optimisation of local geometries. Making use of the localised orbitals described in Section 2.2, one can define a local active space to be accurately modelled. One particular application is the modelling of chemical reactions in large systems where only a small part of the system is directly involved in the reaction[61].

Figure 2.1 illustrates the active space used for the modelling of an abstraction process where a hydrogen atom is dissociated from 1,3-butadiene. This active space contains 5 occupied and 26 virtual orbitals compared to 10 occupied and

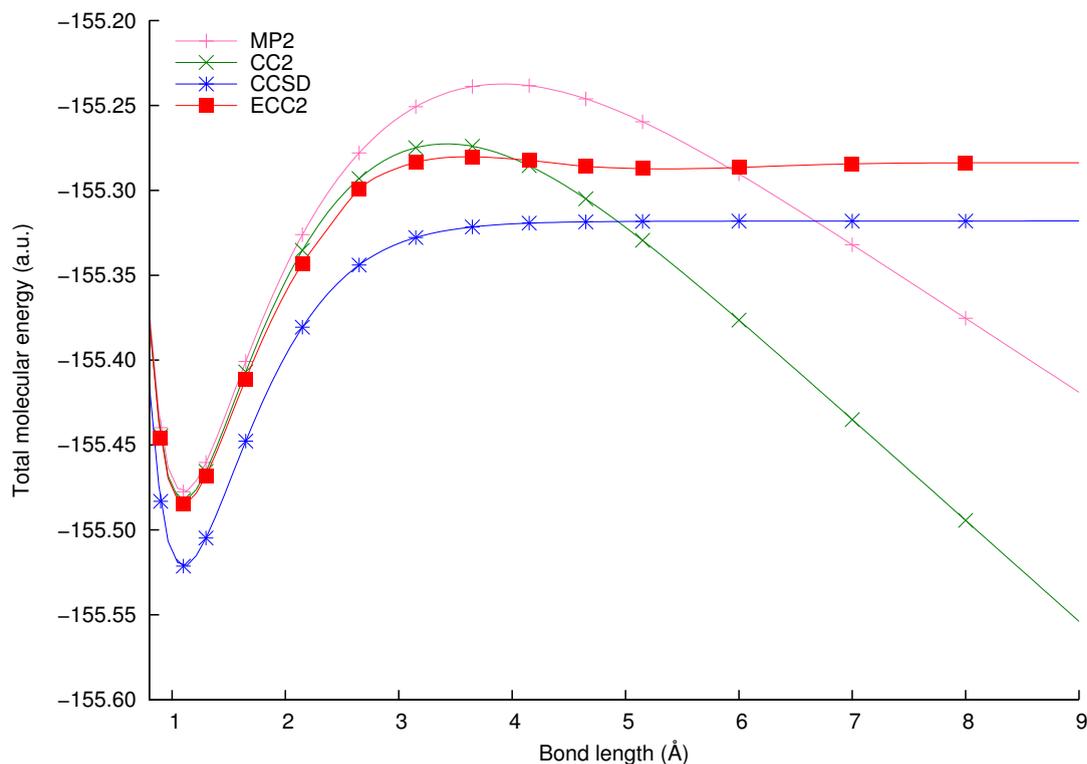


FIGURE 4.4: Total energy curves for dissociation of a hydrogen atom from 1,3-butadiene using the basis set aug-cc-pVDZ.

131 virtual orbitals in the inactive space. From the initial geometry (Table B.2), a hydrogen atom was abstracted along the bond direction. The total energy curves calculated with various models are plotted in Figure 4.4. The accuracy of ECC2 is about the same as for CC2 in the equilibrium geometry. However, the error in ECC2 with respect to CCSD is almost constant throughout the process and the dissociation energy of ECC2 is 531.5 kJ/mol compared to 535.1 kJ/mol for CCSD.

A problem with this model is the unphysical reduction in energy when increasing the bond length beyond 3.5 Å. This is most likely due to a change in the Cholesky decomposition during the abstraction process which again leads to a change in the number of active and inactive orbitals.

TABLE 4.1: Electronic dipole moments and polarisability in a.u. of hydrogen fluoride the along  $C_\infty$  axis with the basis set aug-cc-pVTZ.

Active space		Dipole moment		Polarisability	
o	v <sup>1</sup>	ECC2a	ECC2b	ECC2a	ECC2b
CC2		0.1679		6.73	
1	1	0.1679	0.1679	6.72	6.73
2	2	0.1675	0.1685	6.72	6.74
3	3	0.1605	0.1590	6.54	6.50
4	4	0.1607	0.1592	6.49	6.45
4	5	0.1603	0.1580	6.46	6.42
CCSD		0.1527		6.35	

<sup>1</sup> o - number of occupied orbitals,  
v - number of virtual orbitals

### 4.1.3 Static properties

Static properties such as dipole moments and polarisabilities can be calculated by taking the numerical derivatives of the energy with respect to field strength. As these properties are size-extensive, localised active space are generally not expected to greatly improve their accuracy. However, electrons in higher energy orbitals are more mobile than the core electrons[62], and thus, treating these electrons with higher accuracy in an MLCC model may give a better accuracy to computation complexity ratio.

All static property calculations presented used experimental geometry from the CRC handbook[63], except for 1,3-butadiene where an SCF optimised geometry was used[64]. Numerical differentiation was employed to obtain electronic dipole moments and polarisabilities, as these are not implemented for the ECC2 model. Both analytical and numerical results are available for CC2 and the difference between them was in the order of  $10^{-4}$  a.u. for dipole moments and  $10^{-2}$  a.u. for polarisabilities.

Dipole moment and polarisability for hydrogen fluoride along the H-F bond are presented in Table 4.1. Using CC2, the errors in dipole moment and polarisability is about 10% and 6% respectively with error defined for property  $p$  as  $|(p_{CCSD} - p_{CC2})/p_{CCSD}|$ . With the ECC2b model and the largest active space, 4

TABLE 4.2: Electronic dipole moment and polarisability in a.u. of ozone the along  $C_2$  axis with the basis set aug-cc-pVTZ.

Active space		Dipole moment		Polarisability	
o	v <sup>1</sup>	ECC2a	ECC2b	ECC2a	ECC2b
CC2		0.1474		14.71	
1	1	0.2091	0.2037	14.17	13.99
2	2	0.2123	0.2044	14.14	13.99
3	3	0.2113	0.2047	14.08	13.96
4	4	0.1868	0.1866	14.06	13.97
5	5	0.1858	0.1837	13.99	13.92
6	6	0.1802	0.1821	13.84	13.82
7	7	0.1828	0.1838	13.74	13.76
8	8	0.1808	0.1824	13.74	13.76
CCSD		0.1876		13.63	

<sup>1</sup> o - number of occupied orbitals,  
v - number of virtual orbitals

occupied and 5 virtual orbitals, the error is reduced to 4% and 1%. In total, there are 76 orbitals in the system, so the number treated with CCSD is greatly reduced compared to the full model.

Due to its resonance structure, large electron correlation effects are expected for ozone. The calculated values are presented in Table 4.2 and errors for CC2 is 21% and 8%. With the ECC2b model and an active space of 8 occupied and 8 virtual orbitals out of a total of 12 occupied and 126 virtual, this is reduced to 3% and 1%. For most of the systems tested, the ECC2 models gave an intermediate value between CCSD and CC2. Ozone appears to be a special case, as the smallest active spaces give an ECC2 dipole moment higher than the CCSD dipole moment while the CC2 moment is smaller.

Table 4.3 presents the polarisabilities calculated for ethene and ethyne along the C-C bond. For these systems, CC2 performs considerably better than the previous examples with errors of 3.3% and 4.4%. Because CC2 performs so well, there is little to be gained by employing the ECC2 model. Including all occupied orbitals except the 1s orbitals of carbon and 8 virtual orbitals in the active space approximately halves the errors to 1.6% and 2.7%.

TABLE 4.3: Polarisability in a.u. of ethyne and ethene along the C-C bonds with the basis set aug-cc-pVTZ.

Active space		C2H2		C2H4	
o	v <sup>1</sup>	ECC2a	ECC2b	ECC2a	ECC2b
CC2		31.30		35.73	
1	1	31.30	31.30	35.72	35.72
2	2	31.29	31.30	35.72	35.73
3	3	31.22	31.22	35.70	35.71
4	4	31.17	31.11	35.69	35.70
5	5	31.14	31.08	35.34	35.29
5	6	31.00	31.06	-	-
5	7	31.07	30.99	-	-
5	8	30.91	30.79	-	-
6	6	-	-	35.27	35.19
6	7	-	-	35.26	35.17
6	8	-	-	35.24	35.15
CCSD		30.30		34.21	

<sup>1</sup> o - number of occupied orbitals,  
v - number of virtual orbitals

TABLE 4.4: Polarisability in a.u. of benzene along a C<sub>2</sub> axis going through two hydrogen atoms and 1-3-butadiene along the C-C single bond with the basis set aug-cc-pVTZ.

Active space		C6H6		C4H6	
o	v <sup>1</sup>	ECC2a	ECC2b	ECC2a	ECC2b
CC2		86.31		82.59	
1	1	86.32	86.32	82.56	82.56
2	2	86.35	86.31	82.55	82.55
3	3	86.33	86.30	82.53	82.55
4	4	86.30	86.30	82.51	82.54
5	5	86.29	86.29	82.46	82.52
6	6	86.26	86.25	82.44	82.49
7	7	85.77	85.44	80.38	80.17
8	8	85.18	84.94	80.04	79.84
9	9	85.07	84.82	79.97	79.73
10	10	84.92	84.62	79.34	78.85
11	11	-	-	79.34	78.85
CCSD		82.38		76.81	

<sup>1</sup> o - number of occupied orbitals,  
v - number of virtual orbitals

For the two larger conjugated systems, benzene and 1,3-butadiene, the errors are reduced from 4.8% to 2.7% and 7.5% to 2.6% respectively. For both the systems, the initial improvement is slow and the results worse than for CC2 in some cases. However, there is a strong improvement when the seventh and eighth occupied and virtual orbitals for both systems are included. Inspection of the output reveals that these orbitals introduce new symmetries in the active space.

## 4.2 Excitation energies

### 4.2.1 Functional groups

Excitation energies are size-intensive properties and often have a highly local nature[65, 66]. In particular, systems with functional groups may have excitations involving only orbitals localised on the atoms in the functional group. Decanal, shown in Figure 4.5-4.8, is such a system with a long, inert carbon chain with an aldehyde group at the end. The two lowest excitation energies calculated with CCSD, CC2 and ECC2 with various active spaces using a geometry from PubChem[67] is presented in Table 4.5.  $\Delta_{model}$  is the excitation energy calculated using CCSD minus the excitation energy of the model.

CC2 performs quite well for the lowest excitation, considering that CCSD excitation energies typically have an error in the order of 0.1 eV with respect to FCI[56]. Defining oxygen, the closest carbon and the attached hydrogen as active atoms halves the error compared to CC2, however, the norms of the amplitudes in Table 4.6 reveals that the excitation has a considerable semi-external character from the active space. In the table, T refers to the CCSD space, while S refers to the CC2 space. Only single excitations are included because most of the examples are almost entirely single excitation and none have more than 10% doubles contribution. As only excitations with an amplitude greater than 0.161467 are printed, the total printed norm is less than 0.7 in some cases. This makes the analysis problematic

and a future implementation should print out the character of the excitations by default.

By expanding the active space to contain the next carbon and its hydrogen, the excitation becomes entirely internal in the active space and the error is reduced to 0.001 eV. The next two models use the same CCSD space, however, they also use a CCS space to describe the carbon chain at the other end of the molecule. As the lowest excitation is still internal to the CCSD space, the excitation energy is still more accurate than CC2, even though half the molecule is treated with CCS in the ECC2 D model.

The second excitation energy is considerably less accurate for all models and CC2 is more than 0.4 eV lower than CCSD. ECC2 A and ECC2 B does not perform particularly well because all the single excitations are either external or semi-external to the active space. ECC2 C perform surprisingly well, however, this is due to two opposing effects. CCS excitation energies are generally a lot higher than those of CC2 and CCSD, so the excitation found using CC2 and CCSD becomes much higher in energy. Instead, the algorithm finds an excitation with a greater internal character. As CC2 lowers the energy, the sum is quite close to the CCSD value. In ECC2 D, the excitation found in the previous model is also pushed up, so the reported excitation is a high energy excitation mostly internal to the CCSD space.

The number of orbitals in the active space is proportional to the number of active atoms. For decanal using cc-pVDZ, ECC2 A uses an active space of 8 occupied and 27 virtual orbitals compared to a total of 44 occupied and 210 virtual orbitals. In ECC2 B, this is increased to 12 and 49 orbitals. ECC2 C and ECC2 D has the same number of CCSD orbitals as ECC2 B and in addition a CCS space. In the first of these, 8 occupied and 41 virtual orbitals are treated with CCS while 20 occupied and 101 virtual are in the CCS space in ECC2 D.

A similar pattern appears for trans-ethyl-i-butyl-diazene (Table 4.7). The active spaces used in the ECC2 models can be found in Figures 4.9 and 4.10 and the geometry obtained from PubChem[68]. Note that the hydrogen atoms are not

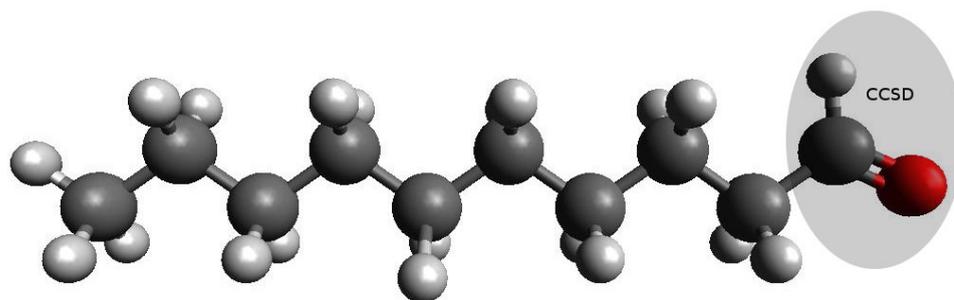


FIGURE 4.5: Decanal A

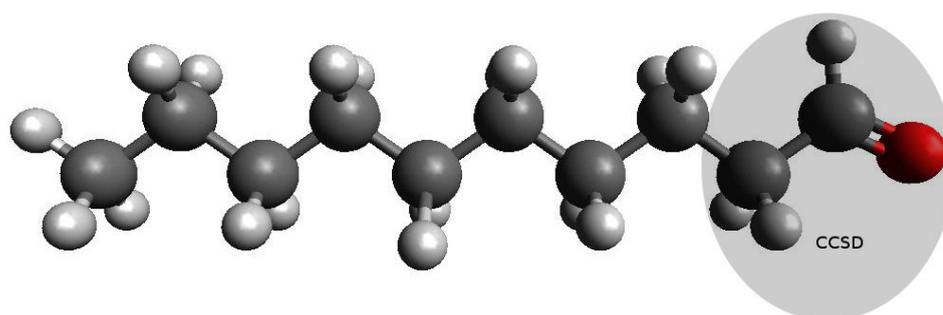


FIGURE 4.6: Decanal B

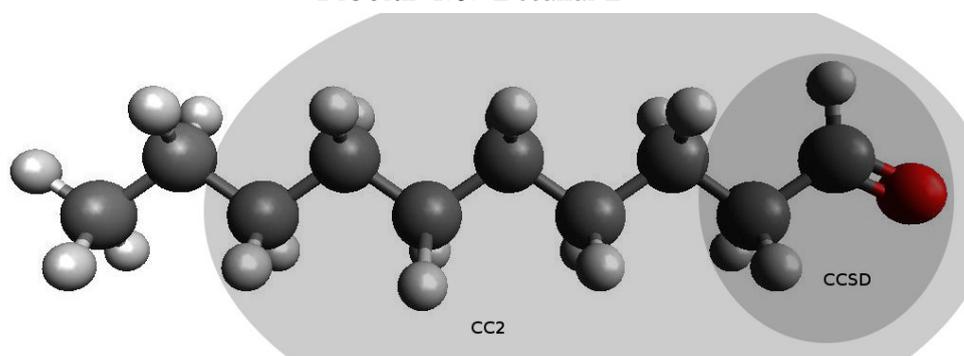


FIGURE 4.7: Decanal C

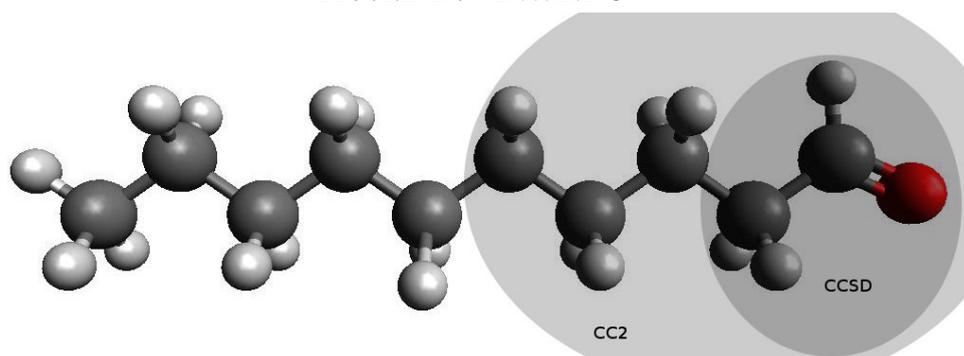


FIGURE 4.8: Decanal D

TABLE 4.5: Excitation energies of decanal in eV using cc-pVDZ.

	CCSD	$\Delta$ CC2	$\Delta$ ECC2 A	$\Delta$ ECC2 B	$\Delta$ ECC2 C	$\Delta$ ECC2 D
1	4.08	-0.08	0.04	0.00	-0.05	-0.05
2	8.52	0.43	0.22	0.17	-0.04	-0.21

TABLE 4.6: Norm of the amplitudes of the different types of single excitations for decanal using cc-pVDZ.

Model	Exci.	T $\rightarrow$ T	T $\rightarrow$ S	S $\rightarrow$ T	S $\rightarrow$ S	tot.
ECC2 A	1	0.68	0.54	-	-	0.87
	2	-	0.23	-	0.89	0.92
ECC2 B	1	0.87	-	-	-	0.87
	2	-	0.21	-	0.85	0.88
ECC2 C	1	0.88	-	-	-	0.88
	2	0.20	0.24	-	0.82	0.88
ECC2 D	1	0.88	-	-	-	0.88
	2	0.82	-	0.23	-	0.85

active in ECC2 B, so the virtual orbitals centred on hydrogen are considered inactive. For the lowest excitations, CC2 performs well, however, at higher energies it performs progressively worse. ECC2 with an active space only including the nitrogen atoms does not perform very well either and considerably worse than CC2 for the second excitation. Analysis of the output, summarised in Table 4.8, reveals that the first excitation is dominated by internal excitations, but also contains semi-external excitations from the T-space to S-space. Excitations 2-4 have a smaller internal character and also include S  $\rightarrow$  T excitations while excitation five is dominated by T  $\rightarrow$  S excitations. As a result, they are all quite inaccurate.

Expanding the active space to also include the carbon atoms neighbouring the nitrogen atoms greatly reduces the errors in the excitation energies. The internal character of excitations 1-4 is considerably increased and the semi-external excitations into the active space seems to become internal. In the last excitation, the internal character is reduced even though the accuracy is increased. This may be due to chance, but all these excitations have quite low total norms. As a result, there might be considerable internal character that is not accounted for.

TABLE 4.7: Excitation energies of trans-ethyl-i-butyl-diazene in eV using cc-pVDZ above and aug-cc-pVDZ below.

	CCSD	$\Delta$ CC2	$\Delta$ ECC2 A	$\Delta$ ECC2 B
1	3.40	0.02	-0.03	-0.01
2	7.41	-0.08	-0.16	-0.01
3	8.14	0.20	-0.22	-0.00
4	8.31	0.29	-0.22	-0.04
5	8.46	0.43	-0.20	0.01
1	3.39	0.04	-0.03	-0.01
2	6.24	0.52	0.03	0.03
3	6.55	0.53	0.04	0.03
4	6.60	0.54	0.03	0.03
5	6.88	0.59	0.06	0.04

TABLE 4.8: Norm of the amplitudes of the different types of single excitations for trans-ethyl-i-butyl-diazene using cc-pVDZ.

Model	Exci.	T $\rightarrow$ T	T $\rightarrow$ S	S $\rightarrow$ T	S $\rightarrow$ S	tot.
ECC2 A	1	0.78	0.41	-	-	0.88
	2	0.65	0.24	0.27	-	0.75
	3	0.50	0.25	0.39	-	0.69
	4	0.54	0.20	0.46	-	0.73
	5	0.38	0.73	-	-	0.82
ECC2 B	1	0.87	0.22	-	-	0.89
	2	0.80	0.19	-	-	0.82
	3	0.78	0.21	-	-	0.81
	4	0.81	0.20	-	-	0.84
	5	0.21	0.76	-	-	0.79

When using an augmented basis set, the character of the higher excitations change considerably. While the character of the first excitation remains more or less the same, dominated by internal excitations and some semi-external, the higher excitations become completely semi-external from the CCSD space to the CC2 space. However, they remain quite accurate.

For tert-butyl hydroperoxide (Table 4.9) ECC2 performs very well, even though CC2 performs poorly. The active spaces are described in Figures 4.11 and 4.12 and the geometry obtained from PubChem[69]. As the highest energy electrons are situated in the peroxide part of the molecule, these are always described by CCSD. Consequently, there are no excitations completely external to the active

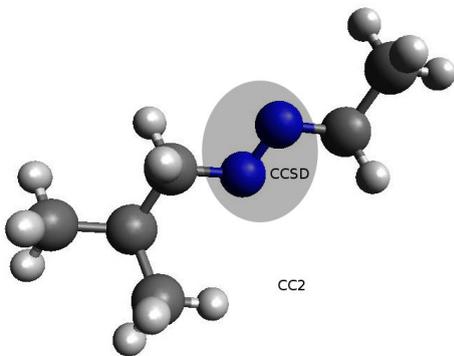


FIGURE 4.9: Trans-ethyl-i-butyl-diazene A

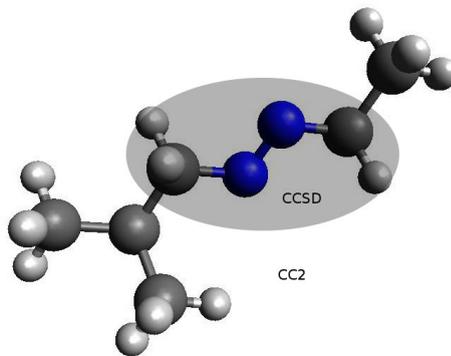


FIGURE 4.10: Trans-ethyl-i-butyl-diazene B

TABLE 4.9: Excitation energies of tert-butyl hydroperoxide in eV using cc-pVDZ.

	CCSD	$\Delta$ CC2	$\Delta$ ECC2 A	$\Delta$ ECC2 B
1	6.00	0.26	0.00	-0.03
2	6.79	0.53	0.02	0.02
3	6.96	0.22	0.00	-0.02
4	7.57	0.70	0.03	0.06
5	7.75	0.60	0.03	0.03

space involved in any of the excitations (Table 4.10). Excitation 1 and 3 are almost entirely internal in ECC2 A, while the others have substantial semi-external character. In ECC2 B, there is a higher semi-external character for all the excitations except number four. In addition, several of the excitations gain external character, but the results remain quite accurate. The total norm is around 0.7 for this system, therefore a significant proportion of the excitations are not accounted for.

## 4.2.2 Solvent effects

When describing molecules in solution, one is rarely interested in the properties of solvent molecules, so such a system is a natural candidate for MLCC. By treating the solute with a high level method and the solvent with a low level, one can obtain an accurate description of the solvent effect.

TABLE 4.10: Norm of the amplitudes of the different types of single excitations for tert-butyl hydroperoxide using cc-pVDZ.

Model	Exci.	T $\rightarrow$ T	T $\rightarrow$ S	S $\rightarrow$ T	S $\rightarrow$ S	tot.
ECC2 A	1	0.74	0.19	-	-	0.76
	2	0.35	0.68	-	-	0.77
	3	0.66	0.19	-	-	0.68
	4	-	0.71	-	-	0.71
	5	0.19	0.68	-	-	0.71
ECC2 B	1	0.65	0.34	-	-	0.73
	2	0.34	0.65	-	0.21	0.76
	3	0.59	0.33	-	-	0.68
	4	0.23	0.58	-	0.35	0.71
	5	0.29	0.63	-	0.20	0.72

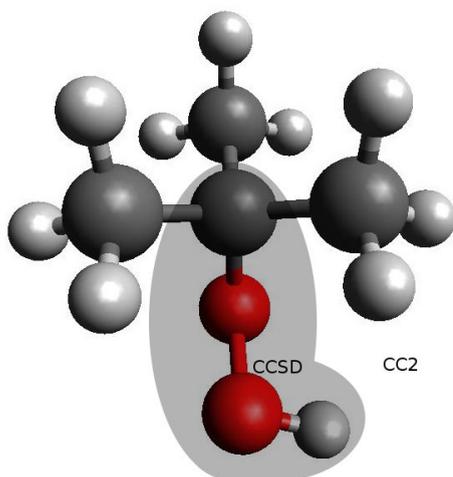


FIGURE 4.11: Tert-butyl hydroperoxide A

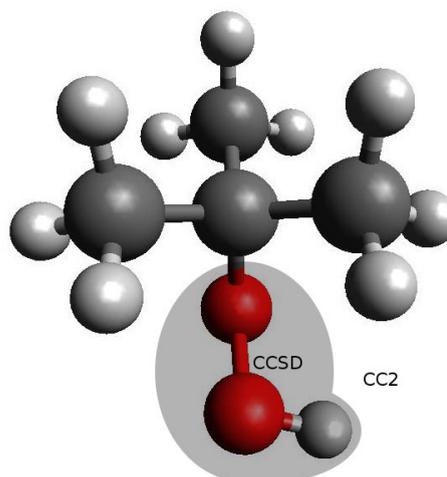


FIGURE 4.12: Tert-butyl hydroperoxide B

Table 4.11 presents the excitation energies of acetone with three water molecules. The geometry was obtained through optimisation using the MMFF94s force field in Avogadro[70] (Table B.3) while the active spaces are presented in Figures 4.13 and 4.14. Using CC2 with a cc-pVDZ basis yields quite accurate results, however, describing acetone with CCSD and the water molecules with CC2 gives the same first excitation energy as CCSD on the whole system. Even when two of the water molecules are treated with CCS, the error is only 0.04 eV. The first excitation energy calculated with CCSD without water is 0.13 eV higher than with water.

The norms are presented in Table 4.12. Accuracy of the first excitation is a result

TABLE 4.11: Excitation energies of acetone with water in eV using cc-pVDZ above and aug-cc-pVDZ below.

	CCSD	$\Delta$ CC2	$\Delta$ ECC2 A	$\Delta$ ECC2 B
1	4.52	-0.06	0.00	-0.04
2	8.30	0.15	0.06	-0.63
3	8.57	0.39	0.08	-0.43
1	4.51	0.00	0.00	-0.10
2	6.92	0.65	0.02	-0.29
3	7.41	0.42	0.37	-0.66

of the excitations being almost completely internal. The next two excitations in ECC2 A are external to the active space and it is thus surprising that they differ from the CC2 excitations to such an extent. In the ECC2 B model, CCS will increase the excitation energies of water above those of acetone, resulting in the excitations becoming internal and semi-external to the T-space.

The lower half of Table 4.11 contains the excitations for the system using aug-cc-pVDZ. ECC2 A again reproduces the CCSD value for the lowest excitation. In addition, the second excitation is now situated mostly on acetone and considerably more accurate. The third, however, is still on water and is not much better than CC2. The same pattern can be observed for the two lowest excitations in ECC2 B, but it is not very accurate. As the diffuse basis functions reach quite far, it is possible that orbitals that should be treated with CCSD are instead treated with CCS. Again, CCS increases the energy of the water excitations, so the third excitation is entirely internal in the CCSD space on acetone.

Similar calculations on propenal, treated with CCSD, and water, with CC2, gave similar results. Using cc-pVDZ the errors for the two lowest excitations were 0.03 and 0.07 eV compared to 0.00 and 0.16 eV using CC2. With the augmented basis set, ECC2 deviated 0.02 and 0.00 eV from CCSD compared to 0.04 and 0.16 eV with CC2.

TABLE 4.12: Norm of the amplitudes of the different types of single excitations for acetone with water using cc-pVDZ above and aug-cc-pVDZ below.

Model	Exci.	T $\rightarrow$ T	T $\rightarrow$ S	S $\rightarrow$ T	S $\rightarrow$ S	tot.
ECC2 A	1	0.93	-	-	-	0.93
	2	-	-	-	0.94	0.94
	3	-	-	-	0.94	0.94
ECC2 B	1	0.93	-	-	-	0.93
	2	0.47	0.78	-	-	0.91
	3	0.78	0.49	-	-	0.92
ECC2 A	1	0.83	-	-	-	0.83
	2	0.60	0.47	-	-	0.76
	3	-	-	-	0.90	0.90
ECC2 B	1	0.82	-	-	-	0.82
	2	0.73	0.44	-	-	0.86
	3	0.89	-	-	-	0.89

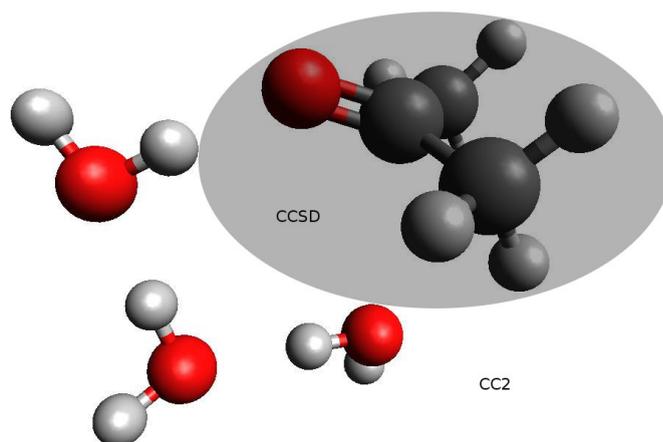


FIGURE 4.13: Acetone A

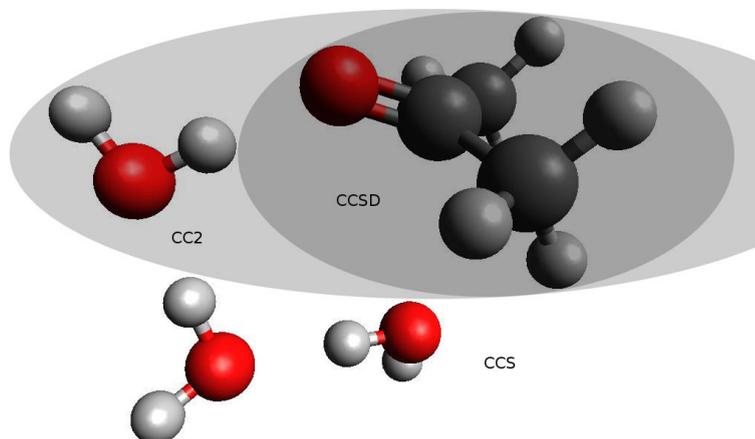


FIGURE 4.14: Acetone B

TABLE 4.13: Excitation energies of (2E,4E,6E,8E)-2,4,6,8-decatetraene in eV using cc-pVDZ.

	CCSD	$\Delta$ CC2	$\Delta$ ECC2
1	5.09	0.36	-0.03
2	6.15	0.21	-0.28
3	6.93	0.12	-0.07
4	7.10	0.27	0.01
5	7.15	0.05	-0.50
6	7.54	0.02	-0.19
7	7.63	0.12	-0.20
8	7.75	0.20	-0.12
9	8.05	0.16	-0.10
10	8.07	0.17	-0.09

TABLE 4.14: Norm of the amplitudes of the different types of single excitations for (2E,4E,6E,8E)-2,4,6,8-decatetraene using cc-pVDZ.

Model	Exci.	T $\rightarrow$ T	T $\rightarrow$ S	S $\rightarrow$ T	S $\rightarrow$ S	tot.
	1	0.47	0.58	0.34	0.42	0.76
	2	0.59	0.35	0.38	0.43	0.89
ECC2	3	0.64	0.39	0.36	0.29	0.87
	4	0.34	0.50	0.47	0.43	0.88
	5	0.23	0.53	0.53	0.20	0.83

### 4.2.3 Conjugate systems

Conjugate systems are highly delocalised. Therefore, using MLCC might become problematic if parts of the conjugate system are treated with different levels of theory. The ten lowest excitation energies of (2E,4E,6E,8E)-2,4,6,8-decatetraene are presented in Table 4.13. As there are many nearly degenerate excitations, it is difficult to determine if the excitations found by the different methods actually corresponds to each other. The geometry were obtained from PubChem[71] and the four central carbons with attached hydrogens were considered active to maintain the same point group symmetry (Figure 4.15). Neither CC2 nor ECC2 performs particularly well in this case and ECC2 is for several excitations worse than CC2. Due to the delocalised nature of the excitations, all of them have a large external and semi-external character (Table 4.14).

A similar behaviour is observed for 2,4,6,8-decatetraenal in Table 4.15. Active

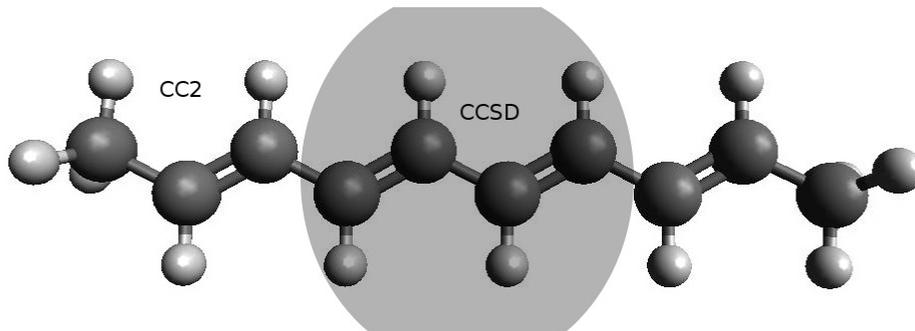


FIGURE 4.15: (2E,4E,6E,8E)-2,4,6,8-decatetraene

TABLE 4.15: Excitation energies of 2,4,6,8-decatetraenal in eV using cc-pVDZ.

	CCSD	$\Delta$ CC2	$\Delta$ ECC2 A	$\Delta$ ECC2 B
1	3.78	0.10	-0.01	0.01
2	4.70	0.37	0.02	0.06
3	6.01	0.24	-0.19	-0.16
4	6.50	0.31	0.04	0.02
5	6.96	0.65	0.05	0.09

TABLE 4.16: Norm of the amplitudes of the different types of single excitations for 2,4,6,8-decatetraenal using cc-pVDZ.

Model	Exci.	T $\rightarrow$ T	T $\rightarrow$ S	S $\rightarrow$ T	S $\rightarrow$ S	tot.
ECC2 A	1	0.86	0.27	-	-	0.90
	2	0.35	0.33	0.36	0.66	0.89
	3	-	0.52	0.33	0.64	0.89
	4	0.51	-	0.30	0.61	0.85
	5	0.63	0.59	-	-	0.86
ECC2 B	1	0.78	0.48	-	-	0.92
	2	0.28	0.41	0.22	0.74	0.92
	3	-	0.54	0.20	0.71	0.92
	4	0.48	0.25	0.34	0.60	0.88
	5	0.52	0.74	-	-	0.90

spaces can be found in Figures 4.16 and 4.17 and the geometry was obtained from PubChem[72]. Note that the hydrogens closest to oxygen are not included in the active space in ECC2 B. While most of the excitations are quite accurate, excitation three has no internal contribution, resulting in a large error.

On the other hand, if the entire conjugated system is included in the active space, such as for 1,3-octadiene[73], 6-methyl-2,4-heptanedione[74] and 2,4-octadienal[75] in Figures 4.18-4.20, ECC2 appears to perform very well. For the five lowest

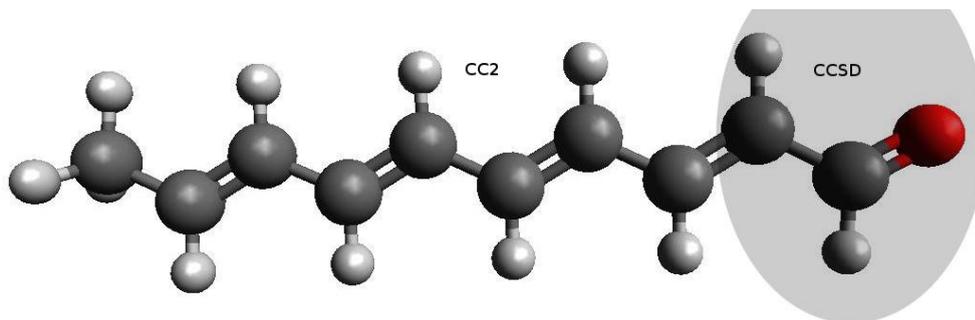


FIGURE 4.16: 2,4,6,8-decatetraenal A

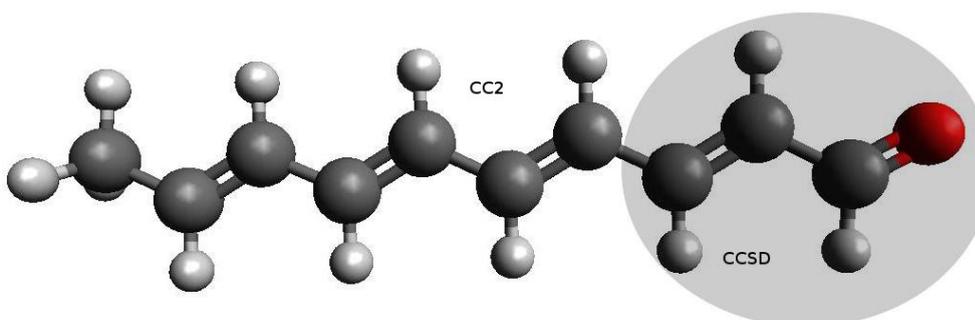


FIGURE 4.17: 2,4,6,8-decatetraenal B

excitations in these systems, there were no errors with respect to CCSD greater than 0.03 eV, even though the errors for CC2 varied widely and were greater than 0.70 eV in some cases. All these excitations were dominated by internal excitations with small semi-external contributions.

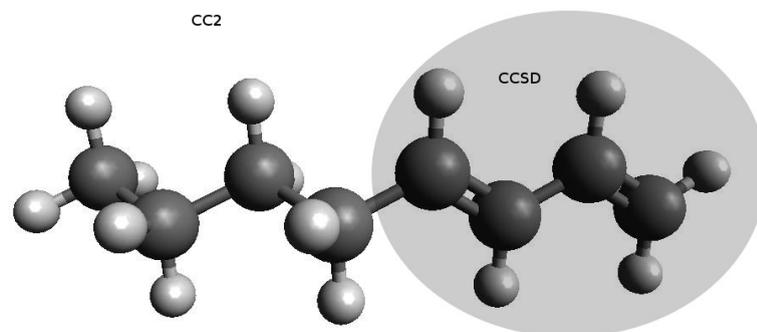


FIGURE 4.18: 1,3-octadiene

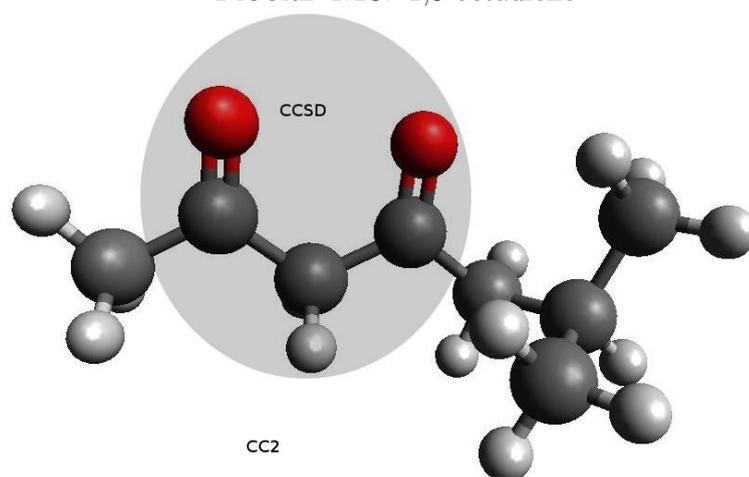


FIGURE 4.19: 6-methyl-2,4-heptanedione

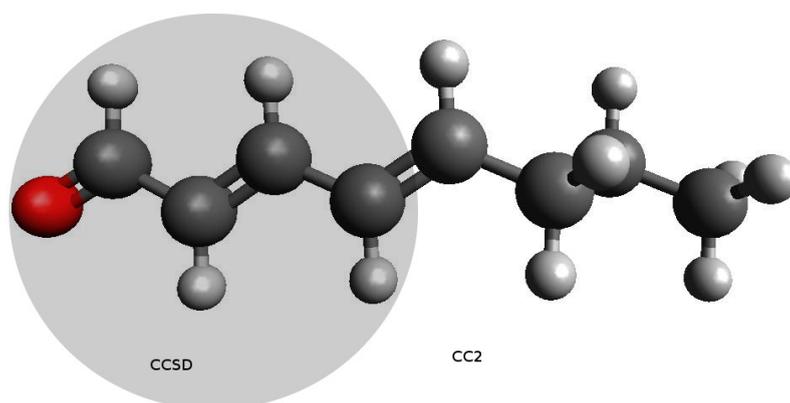


FIGURE 4.20: 2,4-octadienal



# Chapter 5

## Discussion

### 5.1 Linear and sublinear scaling

Linear scaling techniques are computational techniques whose complexity scales linearly with the size of the system. The first such technique was implemented for DFT by Yang in 1991[76] and have since been developed for wave function based methods like SCF and CC theory[77, 78]. Considerable reductions in computational complexity has been achieved. For example, Riplinger and Neese recently performed DLPNO-CCSD calculations on a system with over 450 atoms[11].

The long-term goal of MLCC is to develop models that can determine local properties with size-intensive complexity (SIC). ECC2, as outlined here, will result in significant reductions in computational complexity, but it will not achieve sublinear scaling because it uses more or less the standard CC2 equations in the inactive space. If the system is above a certain size it can be treated at the SCF level far away from the active space. The SCF calculation, like a DFT calculation, will scale linearly with the size of the system. However, if the size of the active space remains constant, the electron correlation calculation can be made size-intensive by creating an auxiliary basis for the active MO integrals using Cholesky decomposition[18, 79].

## 5.2 Extensive properties

If a size-extensive property, like total energy or polarisability, is desired, one can at best achieve linear scaling. In such cases, a more traditional active space based on the orbital energies will be the most useful as the orbitals close to HOMO and LUMO contribute disproportionately to electron correlation[80].

In Chapter 4, several examples of energy calculations using ECC2 were presented. As the current implementation is not optimal, calculations have only been performed on quite small systems. The results demonstrate the effect of different active spaces, but performance should be tested for larger systems.

At equilibrium geometry, the total energy from ECC2 was about the same as CC2 in Figures 4.1-4.3. The active spaces chosen were the smallest possible with only HOMO and LUMO for lithium hydride and the sodium dimer and two occupied and two virtual for ethene. As the total energy of a system is rarely of much interest, the focus should be on how well relative energies like the dissociation energy are reproduced. Far from the equilibrium geometry, ECC2 greatly outperformed CC2 and MP2 as HOMO and LUMO becomes near degenerate. While the CC2 energy approached a singularity, ECC2 produced physically reasonable dissociation energies.

For the ethene calculations, more orbitals had to be included because HOMO and LUMO changed nature. In a large system, this is likely to be problematic as the changing geometry might cause large changes in orbital energies. Identifying which orbitals that will become near degenerate and must be included can prove challenging and require a large active space. For these reasons, a localised active space will be more appropriate for a large system.

The results from the polarisability calculations were of mixed quality. For ozone and hydrogen fluoride, the CC2 polarisability, and electronic dipole moment, were quite inaccurate. Including most of the occupied orbitals, but only a fraction of the virtual orbitals brought the error in the polarisability down to only one percent. For the organic molecules, CC2 performed quite well and ECC2 gave small error

reductions. The convergence towards the CCSD value were not linear and some orbitals led to much greater gains than others. To establish a strategy, more testing is required, but including very low energy occupied or very high energy virtual orbitals is unlikely to lead to great improvements[62]. The calculations on 1,3-butadiene and benzene indicate that the symmetry of the active orbitals may also be important.

When more accuracy is required, but using a more accurate model is too costly when calculating static properties, using MLCC and an orbital energy based active space might be a good compromise. The size of the required active space is likely to scale with the size of the system and one will not obtain SIC. The results from Chapter 4 indicate that there must be a large difference between the models employed to achieve significant improvements in accuracy.

### 5.3 Intensive properties

Intensive properties, like local geometry and excitation energies[81] relies mostly on the local environment and are almost independent of the electronic structure further away. Such properties can then be obtained with high accuracy by only treating the local environment with advanced models and more approximate models further away.

The energy curves for the dissociation of hydrogen from 1,3-butadiene described by ECC2 in Figure 4.4 is very encouraging. Even though the equilibrium geometry energy is about as accurate as CC2, the abstraction process is accurately described and the difference with respect to CCSD remains almost constant. As a result, the dissociation energy can be obtained to within chemical accuracy. When scaling up the size of the system, it is unlikely that it will be necessary to increase the size of the active space to describe the same process with the same accuracy. This would make it possible to implement SIC models.

As noted previously, the energy goes through an unphysical minimum at about 5 Å. This is most likely an artefact of the Cholesky decomposition. As the hydrogen atom moves away, the density matrix and therefore, the number of active Cholesky MOs changes. To avoid this, one can use the same pivoting diagonal elements for each geometry. This will ensure that the MOs for each calculation correspond to each other and that the number of active MOs remains constant. Another possibility is to increase the size of the active space so that the atoms changing position are further from the active space boundary.

Only one example of local geometry calculation is presented here and the method needs further testing on several different systems. Particularly the effect of larger system size and how well reactions can be described must be more thoroughly investigated. However, this requires an efficient implementation of ECC2 which is beyond the scope of this thesis.

Upon calculating excitation energies, ECC2 was able to reproduce CCSD results in the majority of cases. For non-conjugated organic compounds it seems sufficient to only treat the functional groups and the nearest two to three carbon atoms at the full CCSD level and the rest with CC2. In cases where the excitation is completely internal in the active space, one often obtains the same results as with CCSD. Excitation energies with some semi-external character tend to be reasonably accurate with errors less than 0.05 eV though this varies from system to system. The calculations on *trans*-ethyl-*i*-butyl-diazene indicate that the method is more tolerant towards semi-external excitations from the active space to the inactive than the opposite, but this needs more testing. Describing solvents with less accurate models does not seem to affect excitations on the solute much, but many excitations also involve orbitals located on solvent molecules.

Excitation energies for molecules with conjugate systems can be obtained with CCSD accuracy, though they seem to require larger active spaces that encompass the entire conjugate system. For the calculations on decatetraene and decatetraenal the accuracy of the results becomes unpredictable as the excitations are highly

delocalised and may involve orbitals from the inactive space. For these systems, CC2 performs poorly with larger errors with respect to CCSD.

One must be careful when introducing a third space described by CCS. As excitation energies calculated using CCS are usually much higher than those from CC2 and CCSD, contributions from CCS-space amplitudes tends to be smaller than they should be. This can in turn increase the excitation energy or shift the order of excitations. To avoid these problems, one should make sure there is a wide CC2 space to act as buffer between the CCS space and the region of interest. This seems to be particularly important when using augmented basis sets, as these can reach further from the atom they are centred on.

The character of the excitations may vary a lot with the basis set and often becomes more semi-external with an augmented basis set. The effect of this on accuracy varies and it seems to occur more often for higher excitations.

## 5.4 Strategy

Unlike standard CC theory and many other quantum chemistry methods, MLCC is not black box. To make effective use of the method, one has to use knowledge of the system and chemical intuition to assign an active space. For an unfamiliar system, this may not be straightforward and a systematic strategy would be advantageous. Determining the lowest excitation energy of decanal will be used as an example in the following outline.

If an ECC2 calculation is desired using CC2 and CCSD, a starting point could be a standard CC2 calculation and an ECC2 calculation with a minimal active space. In the case of decanal, a minimal active space is ECC2 A. When the calculation is finished, one observes a quite large difference of 0.12 eV between the two models. In addition, inspection of the ECC2 output reveals a substantial semi-external character. Both of these factors indicate that the active space should be expanded. Running the calculation again using the active space from ECC2 B,

the resulting excitation energy is reduced by only 0.04 eV. More importantly, the lowest excitation is now completely internal in the active space, a strong indication that the result is trustworthy. To be certain, one may include a third carbon in the active space to observe the energy change.

To include CCS, one must be more careful as discussed above. More testing has to be performed to develop general rules, but the decanal calculation seems to suggest that a three carbon chain is a big enough CC2 buffer to obtain excitation energies with reasonably good accuracy if they are internal when using cc-pVDZ. A possible strategy in the above example would be to first perform the calculations using CCS and CC2 spaces of varying size and observe how the changing spaces affect the excitation energies. If they are more or less unchanged, one can proceed to add a CCSD space. In a calculation of solvent effects, the innermost solvent shell should probably be treated with CC2 while CCS can be used farther from the solute molecule.

Reduced computational complexity depends on the size of the active space as only the active space is treated with the highest level of theory. The size of the active space scales approximately linearly with the number of active atoms and one should include enough atoms get accurate results, but no more.

## 5.5 Variations

A few variations of the ECC2 model were discussed in Chapter 3 and tested in Chapter 4. As expected, the ECC2-II version performed better than the standard version, especially in equilibrium geometries. In this model, semi-external excitations are treated at the CCSD level. This will greatly increase the number of CCSD amplitudes that has to be calculated and decrease the computational gain. Using ECC2-I with a larger active space is a better method to obtain accurate results. The effect of intermediate formulations should be investigated. For example, one could include doubles involving three active orbitals in the  $T$ -commutator.

Including the commutator in ECC2b results in a small gain in accuracy and may be important for describing semi-external excitations. As the size of the active space is limited, the commutator does not greatly increase computational cost and should be included.

The active space types tested so far are energy based or Cholesky based. More localised active spaces can be achieved using more localised orbitals like those obtained using trust region minimisation[47] and should be tested. Both with canonical and localised orbitals only a few, high energy occupied orbitals are involved in excitations. A possibility is to combine energy and localised active spaces. First, a localised active space can be defined using a suitable localisation scheme. A second active space can then be defined based on the orbital energies in the localised space. This could be particularly useful for computationally demanding methods like the non-orthogonal coupled cluster theory that are in development[82, 83].



# Chapter 6

## Conclusion

### 6.1 Development of MLCC

In MLCC theory, the cluster operator of standard CC theory is split into two or more parts. Each of the cluster operators are associated with a subspace of the orbital space and assigned the excitations in that space. By treating each space with different levels of theory and assigning excitations between spaces to the lowest level space it involves orbitals from, great reductions in computational cost can be achieved. For example, in ECC2, the costly B-term that scales as  $V^4O^2$  in CCSD only has to be computed for a small number of orbitals and the overall scaling becomes that of CC2.

One method to assign the spaces and the splitting of the cluster operator is using the orbital energy of the canonical MOs with the lowest energy virtual and highest energy occupied orbitals described by higher levels of accuracy. For larger systems, this method will scale with the size of the system and may prove problematic when changing the geometry as the energy of the orbitals change. An alternative is to generate local Cholesky MOs and assign them to different spaces depending on which atom they are localised on. The Cholesky decomposition scales as  $N^3$  and is noniterative. While this method may also generate artefacts when changing the geometry, fixing the decomposition pivoting should alleviate the problem.

Both the ECC2 Jacobian and linear response function appear as combinations of CCSD and CC2. However, using a small active space, the CCSD part is correspondingly small and computational complexity should scale as CC2.

## 6.2 Results from calculations

An ECC2 pilot code has been implemented in the DALTON software package. By repeating CCSD calculations with various combinations of CC amplitudes set to zero, the ECC2 energy and Jacobian can be calculated and total energy and excitation energy calculations can be evaluated.

Using the orbital energy based active spaces, energy curves for abstraction processes were obtained. This demonstrated that an active space of two or four orbitals is enough to avoid singularities that occur in CC2 and MP2 in small systems. ECC2 improved the results for electronic dipole moments and polarisability in cases where CC2 deviated far from CCSD.

ECC2 with a Cholesky localised active space obtained the dissociation energy for a hydrogen atom from 1,3-butadiene to within chemical accuracy, indicating that the model can be used to obtain the local geometry in a subsystem. With appropriate active spaces, ECC2 reproduced CCSD excitation energies for some systems. However, care is required when assigning an active space, especially if a third CCS space is used.

# Chapter 7

## Future work

Multi-level coupled cluster theory is still in the early stages of development and much work is required before it can be put to regular use. At the time of writing, pilot code exists for ECC2 total energy and excitation energies with and without a CCS space. A version with CCSD(T) has also been implemented, but it has not been tested yet, so the next step is to explore this model.

The intention of MLCC is to efficiently describe local properties of large systems. This is not possible with the pilot code because it is too primitive. For the same reason, it is not suitable for higher level CC models like CC3 or CCSDT. A proper implementation should solve these issues and therefore have a high priority.

To achieve SIC, Cholesky decomposition can be used to generate auxiliary basis sets for MO integral calculations. Implementing an efficient model that describes parts of the system with SCF, MLCC should be able to scale as DFT. Furthermore, combining MLCC with non-orthogonal CC theory should lead to an efficient multireference model.



# Appendix A

## ECC2 response derivation

To determine the amplitude and multiplier equations, the zero- and first order expansions of the QE Lagrangian(2.82) are required.

$$\begin{aligned}
L^{(0)} &= \langle HF | H_0 \exp(X^{(0)}) | HF \rangle \\
&+ \sum_{\mu_1} \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] | HF \rangle \\
&+ \sum_{\mu_2^T} \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] + \frac{1}{2} [[\hat{H}_0, X_2^{(0)}], X_2^{(0)}] | HF \rangle \\
&+ \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [F, S_2^{(0)}] + \hat{H}_0 + [\hat{H}_0, T_2^{(0)}] | HF \rangle
\end{aligned} \tag{A.1}$$

Taking the derivative of the zero order Lagrangian with respect to the zero order multipliers results in the time-independent ECC2 amplitude equations. Note that the zero order Lagrangian is time-independent, so  $\{L^{(0)}\}_{\mathcal{T}} = L^{(0)}$ .

$$\frac{\partial}{\partial \bar{t}_{\mu_1}^{(0)}} L^{(0)} = \langle \mu_1 | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] | HF \rangle \tag{A.2}$$

$$\frac{\partial}{\partial \bar{t}_{\mu_2^T}^{(0)}} L^{(0)} = \langle \mu_2^T | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] + \frac{1}{2} [[\hat{H}_0, X_2^{(0)}], X_2^{(0)}] | HF \rangle \tag{A.3}$$

$$\frac{\partial}{\partial \bar{t}_{\mu_2^S}^{(0)}} L^{(0)} = \langle \mu_2^S | [F, S_2^{(0)}] + \hat{H}_0 + [\hat{H}_0, T_2^{(0)}] | HF \rangle \tag{A.4}$$

The zero order Lagrange multipliers are determined taking the derivative of the zero order Lagrangian with respect to the amplitudes.

$$\frac{\partial}{\partial t_{\nu_1}^{(0)}} \sum_{\mu_1} \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] | HF \rangle = \sum_{\mu_1} \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \tau_{\nu_1}] | HF \rangle \quad (\text{A.5})$$

$$\frac{\partial}{\partial t_{\nu_2}^{(0)}} \sum_{\mu_1} \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] | HF \rangle = \sum_{\mu_1} \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | [\hat{H}_0, \tau_{\nu_2}] | HF \rangle \quad (\text{A.6})$$

$$\begin{aligned} \frac{\partial}{\partial t_{\nu_1}^{(0)}} \sum_{\mu_2^T} \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] + \frac{1}{2} [[\hat{H}_0, X_2^{(0)}], X_2^{(0)}] | HF \rangle \\ = \sum_{\mu_2^T} \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \tau_{\nu_1}] | HF \rangle \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} \frac{\partial}{\partial t_{\nu_2}^{(0)}} \sum_{\mu_2^T} \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | \hat{H}_0 + [\hat{H}_0, X_2^{(0)}] + \frac{1}{2} [[\hat{H}_0, X_2^{(0)}], X_2^{(0)}] | HF \rangle \\ = \sum_{\mu_2^T} \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | [\hat{H}_0 + [\hat{H}_0, X_2^{(0)}], \tau_{\nu_2}] | HF \rangle \end{aligned} \quad (\text{A.8})$$

$$\begin{aligned} \frac{\partial}{\partial t_{\nu_1^T}^{(0)}} \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [F, S_2^{(0)}] + \hat{H}_0 + [\hat{H}_0, T_2^{(0)}] | HF \rangle \\ = \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [\hat{H}_0 + [\hat{H}_0, T_2^{(0)}], \tau_{\nu_1}] | HF \rangle \end{aligned} \quad (\text{A.9})$$

$$\frac{\partial}{\partial t_{\nu_2^T}^{(0)}} \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [F, S_2^{(0)}] + \hat{H}_0 + [\hat{H}_0, T_2^{(0)}] | HF \rangle = \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [\hat{H}_0, \tau_{\nu_2^T}] | HF \rangle \quad (\text{A.10})$$

$$\frac{\partial}{\partial t_{\nu_2^S}^{(0)}} \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [F, S_2^{(0)}] + \hat{H}_0 + [\hat{H}_0, T_2^{(0)}] | HF \rangle = \sum_{\mu_2^S} \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [F, \tau_{\nu_2^S}] | HF \rangle \quad (\text{A.11})$$

Writing Eqs. (A.5-A.11) in matrix form gives Eq. (2.83).

In the first order Lagrangian, the dependence on  $\exp(\omega t)$  ensures that the time averaged QE Lagrangian  $\{L^{(1)}\}_{\mathcal{T}} = 0$ . Consequently, the second order Lagrangian is required to determine the first order amplitudes. Introducing the frequency dependent perturbation  $V^t(\omega_i)$

$$V^t(t) = \sum_i V^t(\omega_i) \exp(-i\omega_i t) \quad (\text{A.12})$$

Terms involving second order amplitudes or multipliers does not affect the linear response function, nor the first order amplitude equations. Disregarding these, the terms in the second order Lagrangian are.

$$\begin{aligned} \{L_0^{(2)}\}_{\mathcal{T}} = & \left\{ \sum_{i,j=-N}^N \langle HF | [\hat{V}(\omega_i), T_1^{(1)}(\omega_j)] \right. \\ & \left. + \frac{1}{2} [[\hat{H}_0, T_1^{(1)}(\omega_i)], T_1^{(1)}(\omega_j)] | HF \rangle e^{-i(\omega_i + \omega_j)t} \right\}_{\mathcal{T}} \end{aligned} \quad (\text{A.13})$$

$$\begin{aligned} = & \sum_{i=-N}^N \langle HF | [\hat{V}(-\omega_i), T_1^{(1)}(\omega_i)] \\ & + \frac{1}{2} [[\hat{H}_0, T_1^{(1)}(-\omega_i)], T_1^{(1)}(\omega_i)] | HF \rangle \end{aligned}$$

$$\begin{aligned}
\{L_1^{(2)}\}_\mathcal{T} = & \left\{ \sum_{i,j=-N}^N \left( \sum_{\mu_1} \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | [\hat{V}(\omega_i), X^{(1)}(\omega_j)] \right. \right. \\
& + \frac{1}{2} [[\hat{H}_0, X^{(1)}(\omega_i)], X^{(1)}(\omega_j)] |HF\rangle \\
& + \sum_{\mu_1} \bar{t}_{\mu_1}^{(1)}(\omega_i) \left[ \langle \mu_1 | \hat{V}(\omega_j) + [\hat{H}_0, T_1^{(1)}(\omega_j)] + [\hat{V}(\omega_j), X_2^{(0)}] \right. \\
& \left. \left. + \frac{1}{2} [[\hat{H}_0, X^{(1)}(\omega_j)], X_2^{(0)}] |HF\rangle - \omega_j t_{\mu_1}^{(1)}(\omega_j) \right] \right) e^{-i(\omega_i + \omega_j)t} \left. \right\}_\mathcal{T} \quad (\text{A.14})
\end{aligned}$$

$$\begin{aligned}
= & \sum_{i=-N}^N \sum_{\mu_1} \left( \bar{t}_{\mu_1}^{(0)} \langle \mu_1 | [\hat{V}(-\omega_i), X^{(1)}(\omega_i)] \right. \\
& + \frac{1}{2} [[\hat{H}_0, X^{(1)}(\omega_i)], X^{(1)}(-\omega_i)] |HF\rangle \\
& + \bar{t}_{\mu_1}^{(1)}(-\omega_i) \left[ \langle \mu_1 | \hat{V}(\omega_i) + [\hat{H}_0, T_1^{(1)}(\omega_i)] \right. \\
& \left. \left. + \frac{1}{2} [[\hat{H}_0, X^{(1)}(\omega_i)], X_2^{(0)}] |HF\rangle - \omega_i t_{\mu_1}^{(1)}(\omega_i) \right] \right)
\end{aligned}$$

$$\begin{aligned}
\{L_{T_2}^{(2)}\}_\tau &= \left\{ \sum_{i,j=-N}^N \sum_{\mu_2^T} \left( \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | [\hat{V}(\omega_i), X_2^{(1)}(\omega_j)] + \frac{1}{2} [[\hat{H}_0, T_1^{(1)}(\omega_i)], T_1^{(1)}(\omega_j)] \right. \right. \\
&\quad + [[\hat{V}(\omega_i), T_1^{(1)}(\omega_j)], X_2^{(0)}] + [[\hat{H}_0, T_1^{(1)}(\omega_i)], X_2^{(1)}(\omega_j)] \\
&\quad + \frac{1}{2} [[[\hat{H}_0, T_1^{(1)}(\omega_i)], T_1^{(1)}(\omega_j)], X_2^{(0)}] + \frac{1}{2} [[\hat{H}_0, X_2^{(1)}(\omega_i)], X_2^{(1)}(\omega_j)] | HF \rangle \\
&\quad + \bar{t}_{\mu_2^T}^{(1)}(\omega_i) \left[ \langle \mu_2^T | [\hat{H}_0, X^{(1)}(\omega_j)] + [\hat{V}(\omega_j), X_2^{(0)}] \right. \\
&\quad + [[\hat{H}_0, T_1^{(1)}(\omega_j)], X_2^{(0)}] + [[\hat{H}_0, X_2^{(0)}], X_2^{(1)}(\omega_j)] | HF \rangle \\
&\quad \left. \left. - \omega_j t_{\mu_2^T}^{(1)}(\omega_j) \right) e^{-i(\omega_i + \omega_j)t} \right\}_\tau \\
&= \sum_{i=-N}^N \sum_{\mu_2^T} \left( \bar{t}_{\mu_2^T}^{(0)} \langle \mu_2^T | [\hat{V}(-\omega_i), X_2^{(1)}(\omega_i)] + \frac{1}{2} [[\hat{H}_0, T_1^{(1)}(-\omega_i)], T_1^{(1)}(\omega_i)] \right. \\
&\quad + [[\hat{V}(-\omega_i), T_1^{(1)}(\omega_i)], X_2^{(0)}] + [[\hat{H}_0, T_1^{(1)}(-\omega_i)], X_2^{(1)}(\omega_i)] \\
&\quad + \frac{1}{2} [[[\hat{H}_0, T_1^{(1)}(-\omega_i)], T_1^{(1)}(\omega_i)], X_2^{(0)}] + \frac{1}{2} [[\hat{H}_0, X_2^{(1)}(-\omega_i)], X_2^{(1)}(\omega_i)] | HF \rangle \\
&\quad + \bar{t}_{\mu_2^T}^{(1)}(-\omega_i) \left[ \langle \mu_2^T | [\hat{H}_0, X^{(1)}(\omega_i)] + [\hat{V}(\omega_i), X_2^{(0)}] \right. \\
&\quad \left. \left. + [[\hat{H}_0, T_1^{(1)}(\omega_i)], X_2^{(0)}] + [[\hat{H}_0, X_2^{(0)}], X_2^{(1)}(\omega_i)] | HF \rangle - \omega_i t_{\mu_2^T}^{(1)}(\omega_i) \right) \right) \\
&\hspace{15em} (A.15)
\end{aligned}$$

$$\begin{aligned}
\{L_{S2}^{(2)}\}_{\mathcal{T}} &= \left\{ \sum_{i,j=-N}^N \sum_{\mu_2^S} \left( \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [\hat{V}(\omega_i), X_2^{(1)}(\omega_j)] + \frac{1}{2} [[\hat{H}_0, T_1^{(1)}(\omega_i)], T_1^{(1)}(\omega_j)] \right. \right. \\
&\quad + [[\hat{V}(\omega_i), T_1^{(1)}(\omega_j)], T_2^{(0)}] + [[\hat{H}_0, T_1^{(1)}(\omega_i)], T_2^{(1)}(\omega_j)] \\
&\quad + \frac{1}{2} [[[\hat{H}_0, T_1^{(1)}(\omega_i)], T_1^{(1)}(\omega_j)], T_2^{(0)}] | HF \rangle \\
&\quad + \bar{t}_{\mu_2^S}^{(1)}(\omega_i) \left[ \langle \mu_2^S | [\hat{H}_0, T_1^{(1)}(\omega_j) + T_2^{(1)}(\omega_j)] + [\hat{V}(\omega_j), X_2^{(0)}] \right. \\
&\quad \left. \left. + [[\hat{H}_0, T_1^{(1)}(\omega_j)], T_2^{(0)}] | HF \rangle - \omega_j t_{\mu_2^S}^{(1)}(\omega_j) \right] \right) e^{-i(\omega_i + \omega_j)t} \Bigg\}_{\mathcal{T}} \\
&= \sum_{i=-N}^N \sum_{\mu_2^S} \left( \bar{t}_{\mu_2^S}^{(0)} \langle \mu_2^S | [\hat{V}(-\omega_i), X_2^{(1)}(\omega_i)] + \frac{1}{2} [[\hat{H}_0, T_1^{(1)}(-\omega_i)], T_1^{(1)}(\omega_i)] \right. \\
&\quad + [[\hat{V}(-\omega_i), T_1^{(1)}(\omega_i)], T_2^{(0)}] + [[\hat{H}_0, T_1^{(1)}(-\omega_i)], T_2^{(1)}(\omega_i)] \\
&\quad + \frac{1}{2} [[[\hat{H}_0, T_1^{(1)}(-\omega_i)], T_1^{(1)}(\omega_i)], T_2^{(0)}] | HF \rangle \\
&\quad + \bar{t}_{\mu_2^S}^{(1)}(-\omega_i) \left[ \langle \mu_2^S | [\hat{H}_0, T_1^{(1)}(\omega_j) + T_2^{(1)}(\omega_i)] + [\hat{V}(\omega_i), X_2^{(0)}] \right. \\
&\quad \left. \left. + [[\hat{H}_0, T_1^{(1)}(\omega_i)], T_2^{(0)}] | HF \rangle - \omega_i t_{\mu_2^S}^{(1)}(\omega_i) \right] \right) \\
\end{aligned} \tag{A.16}$$

Summing the terms and taking the derivative with respect to the first order multipliers results in Eq. (2.84). Due to the  $2n + 2$  rule, the first order multipliers does not contribute to the linear response theory, so the terms involving them can be disregarded when solving for the response function. Noting that  $\partial \hat{V}^t(\omega_i) / \partial \epsilon_A(\omega_i) = \hat{A}$ , inserting  $\{L^{(2)}\}$  into Eq. (2.80) results in Eq. (2.88).

# Appendix B

## Initial geometries

TABLE B.1: Initial geometry of ethene ( $\text{\AA}$ ).

atom	x	y	z
C	0.0	0.0	0.665647
C	0.0	0.0	-0.665647
H	0.0	-0.923658	1.239711
H	0.0	0.923658	-1.239711
H	0.0	-0.923658	-1.239711
H	0.0	0.923658	1.239711

TABLE B.2: Initial geometry of 1,3-butadiene ( $\text{\AA}$ ).

atom	x	y	z
C	0.0	0.0	0.0
C	-0.524786	3.619893	0.0
C	-0.575773	1.160600	0.0
C	0.050986	2.459292	0.0
H	-1.729024	3.619893	0.0
H	-0.394695	-0.938866	0.0
H	-0.130090	4.558760	0.0
H	1.138098	2.500263	0.0
H	-1.662884	1.119629	0.0
H <sup>1</sup>	1.204238	0.0	0.0

<sup>1</sup> The abstracted hydrogen atom.

TABLE B.3: Geometry of acetone with three water molecules ( $\text{\AA}$ ).

atom	x	y	z
C	1.45893	1.01303	-0.72862
C	0.18412	0.26419	-0.44739
O	0.18705	-0.94482	-0.19858
C	-1.09931	1.04517	-0.53780
H	1.49786	1.27765	-1.78806
H	1.50510	1.91653	-0.11651
H	2.31761	0.38199	-0.48331
H	-1.02606	1.95677	0.05946
H	-1.29656	1.29860	-1.58221
H	-1.92478	0.44126	-0.15099
O	0.40206	1.76801	2.51905
H	0.08973	0.91990	2.91556
H	0.81688	2.17521	3.29534
O	-0.26980	-0.53566	3.79206
H	-0.83081	-1.00476	4.43233
H	-0.07087	-1.27753	3.16677
O	0.07542	-2.49924	1.96898
H	0.33305	-3.36278	1.60350
H	0.08734	-1.95179	1.14527

# Bibliography

- [1] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen. *Molecular Electronic-Structure Theory*. John Wiley & Sons, LTD, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, England, May 2004.
- [2] George D. Purvis and Rodney J. Bartlett. A full coupledcluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.*, 76:1910, 1982.
- [3] Hendrik J. Monkhorst. Calculation of properties with the coupled-cluster method. *Int. J. Quantum Chem.*, S11:421, 1977.
- [4] Rodney J. Bartlett and Monika Musial. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.*, 79:291–352, 2007.
- [5] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864B871, 1964.
- [6] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133A1138, 1965.
- [7] Kieron Burke. Perspective on density functional theory. *J. Chem. Phys.*, 136:150901, 2012.
- [8] Marcin Ziólkowski, Branislav Jansik, Thomas Kjærgaard, and Poul Jørgensen. Linear scaling coupled cluster method with correlation energy based error control. *J. Chem. Phys.*, 133:014107, 2010.

- [9] Hans-Joachim Werner and Martin Schütz. An efficient local coupled cluster method for accurate thermochemistry of large systems. *J. Chem. Phys.*, 135:144116, 2011.
- [10] Martin Schütz, Jun Yang, Garnet Kin-Lic Chan, Frederic R. Manby, and Hans-Joachim Werner. The orbital-specific virtual local triples correction: OSV-L(T). *J. Chem. Phys.*, 138:054109, 2013.
- [11] Christoph Riplinger and Frank Neese. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.*, 138:034106, 2013.
- [12] Henrik Koch and Alfredo Sánchez de Merás. Size-intensive decomposition of orbital energy denominators. *J. Chem. Phys.*, 113:508, 2000.
- [13] Francesco Aquilante, Linus Boman, Jonas Boström, Henrik Koch, Roland Lindh, Alfredo Sánchez de Merás, and Thomas Bondo Pedersen. *Linear-Scaling Techniques in Computational Chemistry and Physics*, chapter Cholesky Decomposition Techniques in Electronic Structure Theory. Springer Netherlands, P.O. Box 989, 3300 AZ Dordrecht, The Netherlands, 2011.
- [14] Francesco Aquilante, Thomas Bondo Pedersen, and Roland Lindh. Low-cost evaluation of the exchange fock matrix from cholesky and density fitting representations of the electron repulsion integrals. *J. Chem. Phys.*, 126:194106, 2007.
- [15] Henrik Koch, Alfredo Sánchez de Merás, and Thomas Bondo Pedersen. Reduced scaling in electronic structure calculations using cholesky decompositions. *J. Chem. Phys.*, 118:9481–9484, 2003.
- [16] Thomas Bondo Pedersen, Alfredo Sánchez de Merás, and Henrik Koch. Polarizability and optical rotation calculated from the approximate coupled cluster singles and doubles cc2 linear response theory using cholesky decompositions. *K. Chem. Phys.*, 120:8887–8897, 2004.

- [17] I. G. Cuesta, T. B. Pedersen, H. Koch, and A. M. J. Sánchez de Merás. Polarizabilities of small annulenes from cholesky cc2 linear response theory. *Chem. Phys. Lett.*, 390:170–175, 2004.
- [18] Linus Boman, Henrik Koch, and Alfredo Sánchez de Merás. Method specific cholesky decomposition: Coulomb and exchange energies. *J. Chem. Phys.*, 129:134107, 2008.
- [19] Andreas Köhn and Jeppe Olsen. Coupled-cluster with active space selected higher amplitudes: Performance of seminatural orbitals for ground and excited state calculations. *J. Chem. Phys.*, 125:174110, 2006.
- [20] Piotr Piecuch. Active-space coupled-cluster methods. *Mol. Phys.*, 108:2987–3015, 2010.
- [21] Francesco Aquilante, Thomas Bondo Pedersen, Alfredo Sánchez de Merás, and Henrik Koch. Fast noniterative orbital localization for large molecules. *J. Chem. Phys.*, 125:174101, 2006.
- [22] A. M. J. Sánchez de Merás, H. Koch, I. G. Cuesta, and L. Boman. Cholesky decomposition-based definition of atomic subsystems in electronic structure calculations. *J. Chem. Phys.*, 132:204105, 2010.
- [23] Keiji Morokuma. Oniom and its applications to material chemistry and catalyses. *Bull. Korean Chem. Soc.*, 24:797, 2003.
- [24] André Severo Pereira Gomes and Christoph R. Jacob. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 108:222–277, 2012.
- [25] Andrew R. Leach. *Molecular Modelling Principles and Applications*. Pearson Education Limited, Edinburgh Gate, Harlow, Essex CM20 2JE, England, 2nd edition, 2001.

- [26] Rolf Heilemann Myhre, Alfredo M. J. Sánchez de Merás, and Henrik Koch. The extended CC2 model ECC2. *J. Mol. Phys.*, 2013. doi: 10.1080/00268976.2013.798435.
- [27] Esper Dalgaard and Hendrik J. Monkhorst. Some aspects of the time-dependent coupled-cluster approach to dynamic response functions. *Phys. rev. A*, 28:1217, 1983.
- [28] Henrik Koch, Hans Jørgen Aa. Jensen, Poul Jørgensen, and Trygve Helgaker. Excitation energies from the coupled cluster singles and doubles linear response function (CCSDLR). Applications to Be, CH<sup>+</sup>, CO, and H<sub>2</sub>O. *J. Chem. Phys.*, 93:3345, 1990.
- [29] H. Koch and P. Jørgensen. Coupled cluster response functions. *J. Chem. Phys.*, 93:3333–3344, 1990.
- [30] T. B. Pedersen and H. Koch. Coupled cluster response functions revisited. *J. Chem. Phys.*, 106:8059–8072, 1997.
- [31] Ove Christiansen, Poul Jørgensen, and Christof Hättig. Response functions from fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quantum Chem.*, 68:1–52, 1997.
- [32] Trygve Helgaker, Sonia Coriani, Poul Jørgensen, Kasper Kristensen, Jeppe Olsen, and Kenneth Ruud. Recent advances in wave function-based methods of molecular property calculations. *Chem. Rev.*, 112:543–631, 2012.
- [33] Celestino Angeli, Keld L. Bak, Vebjørn Bakken, Gao Bin, Ove Christiansen, Renzo Cimiraglia, Sonia Coriani, Pål Dahle, Erik K. Dalskov, Thomas Enevoldsen, Berta Fernandez, Lara Ferrighi, Heike Fliegl, Luca Frediani, Christof Hättig, Kasper Hald, Asger Halkier, Hanne Heiberg, Trygve Helgaker, Hinne Hettema, Stinne Høst, Branislav Jansik, Hans Jørgen Aagaard Jensen, Dan Jonsson, Poul Jørgensen, Joanna Kauczor, Sheela Kirpekar, Thomas Kjærgaard, Wim Klopper, Stefan Knecht, Rika Kobayashi, Jacob Kongsted, Henrik Koch, Andreas Krapp, Kasper Kristensen, Andrea Ligabue, Ola B. Lutnæs, Kurt V. Mikkelsen, Christian Neiss, Christian B.

Nielsen, Patrick Norman, Jeppe Olsen, Anders Osted, Martin J. Packer, Filip Pawlowski, Thomas B. Pedersen, Simen Reine, Zilvinas Rinkevicius, Elias Rudberg, Torgeir A. Ruden, Kenneth Ruud, Vladimir Rybkin, Pawel Salek, Claire C. M. Samson, Alfredo M. J. Sánchez de Merás, Trond Saue, Stephan P. A. Sauer, Bernd Schimmelpfennig, Arnfinn H. Steindal, Kristian O. Sylvester-Hvid, Peter R. Taylor, David P. Tew, Andreas J. Thorvaldsen, Lea Thøgersen, Olav Vahtras, Mark Watson, David J. Wilson, and Hans gren. Dalton2011 Quantum Chemistry Suit, July 2013. URL <http://daltonprogram.org/>.

- [34] Celestino Angeli, Keld L. Bak, Vebjørn Bakken, Gao Bin, Ove Christiansen, Renzo Cimiraglia, Sonia Coriani, Pål Dahle, Erik K. Dalskov, Thomas Enevoldsen, Berta Fernandez, Lara Ferrighi, Heike Fliegl, Luca Frediani, Christof Hättig, Kasper Hald, Asger Halkier, Hanne Heiberg, Trygve Helgaker, Eirik Hjertenæs, Hinne Hettema, Stinne Høst, Maria Francesca Iozzi, Branislav Jansik, Hans Jørgen Aagaard Jensen, Dan Jonsson, Poul Jørgensen, Joanna Kauczor, Sheela Kirpekar, Thomas Kjærgaard, Wim Klopper, Stefan Knecht, Rika Kobayashi, Jacob Kongsted, Henrik Koch, Andreas Krapp, Kasper Kristensen, Andrea Ligabue, Ola B. Lutnæs, Kurt V. Mikkelsen, Rolf Heilemann Myhre, Christian Neiss, Christian B. Nielsen, Patrick Norman, Jeppe Olsen, Anders Osted, Martin J. Packer, Filip Pawlowski, Thomas B. Pedersen, Simen Reine, Zilvinas Rinkevicius, Elias Rudberg, Torgeir A. Ruden, Kenneth Ruud, Vladimir Rybkin, Pawel Salek, Claire C. M. Samson, Alfredo M. J. Sánchez de Merás, Trond Saue, Stephan P. A. Sauer, Bernd Schimmelpfennig, Arnfinn H. Steindal, Kristian O. Sylvester-Hvid, Peter R. Taylor, David P. Tew, Andreas J. Thorvaldsen, Lea Thøgersen, Olav Vahtras, Mark Watson, David J. Wilson, and Hans gren. The DALTON quantum chemistry program system. *WIREs Comput. Mol. Sci.*, 3, 2013. Submitted.
- [35] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. McGraw-Hill Publishing Company, 1221 Avenue of the Americas, New York, NY 10020, 1989.

- [36] Krishnan Raghavachari, Gary W. Trucks, John A. Pople, and Martin Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.*, 157:479–483, 1989.
- [37] Jan Řezáč and Pavel Hobza. Describing noncovalent interactions beyond the common approximations: How accurate is the gold standard, CCSD(T) at the complete basis set limit? *J. Chem. Theory Comput.*, 9:2151–2155, 2013.
- [38] Stanislaw A. Kucharski and Rodney J. Bartlett. The coupled-cluster single, double, triple, and quadruple excitation method. *J. Chem. Phys.*, 97:4282, 1992.
- [39] Lea Thøgersen. *Optimization of Densities in Hartree-Fock and Density-functional Theory Atomic Orbital Based Response Theory and Benchmarking for Radicals*. PhD thesis, University of Aarhus, 2005.
- [40] O. Christiansen, H. Koch, and P. Jørgensen. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.*, 243:409–418, 1995.
- [41] Henrik Koch, Ove Christiansen, Poul Jørgensen, Alfredo M. J. Sánchez de Merás, and Trygve Helgaker. The CC3 model: An iterative coupled cluster approach including connected triples. *J. Chem. Phys.*, 106:1808, 1997.
- [42] Ove Christiansen, Henrik Koch, and Poul Jørgensen. Response functions in the CC3 iterative triple excitation model. *J. Chem. Phys.*, 103:7429, 1995.
- [43] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, editors. *Numerical Recipes in FORTRAN*. Press Syndicate of the University of Cambridge, The Pitt Building, Trumpington Street, Cambridge CB2 1RP, 2nd edition, 1992.
- [44] S. F. Boys. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.*, 32:296–299, 1960.

- [45] Clyde Edmiston and Klaus Ruedenberg. Localized atomic and molecular orbitals. *Rev. Mod. Phys.*, 35:457–464, 1963.
- [46] János Pipek and Paul G. Mezey. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.*, 90:4916, 1989.
- [47] Ida-Marie Høyvik, Branislav Jansik, and Poul Jørgensen. Trust region minimization of orbital localization functions. *J. Chem. Theory Comput.*, 8:3137–3146, 2012.
- [48] Henrik Koch, Alfredo Sánchez de Merás, Trygve Helgaker, and Ove Christiansen. The integral-direct coupled cluster singles and doubles model. *J. Chem. Phys.*, 104:4157, 1996.
- [49] Kotoku Sasagane, Fumihiko Aiga, and Reikichi Itoh. Higher-order response theory based on the quasienergy derivatives: The derivation of the frequency-dependent polarizabilities and hyperpolarizabilities. *J. Chem. Phys.*, 99:3738, 1993.
- [50] Trygve Helgaker, Sonia Coriani, Poul Jørgensen, Kasper Kristensen, Jeppe Olsen, and Kenneth Ruud. Recent advances in wave function-based methods of molecular-property calculations. *Chem. Rev.*, 112:543–631, 2012.
- [51] Trygve Helgaker and Poul Jørgensen. Analytical calculation of geometrical derivatives in molecular electronic structure theory. *Adv. Quantum Chem.*, 19:183–245, 1988.
- [52] Trygve Helgaker and Poul Jørgensen. Configuration-interaction energy derivatives in a fully variational formulation. *Theor. Chim. Acta*, 75:111–127, 1989.
- [53] U. von Barth and C. D. Gelatt. Validity of the frozen-core approximation and pseudopotential theory for cohesive energy calculations. *Phys. Rev. B*, 21:2222–2228, Mar 1980. doi: 10.1103/PhysRevB.21.2222. URL <http://link.aps.org/doi/10.1103/PhysRevB.21.2222>.

- [54] Henrik Koch, Ove Christiansen, Rika Kobayashi, Poul Jørgensen, and Trygve Helgaker. A direct atomic orbital driven implementation of the coupled cluster singles and doubles (ccsd) model. *Chemical Physics Lett.*, 228:233–238, 1994.
- [55] Henrik Koch, Alfredo Sánchez de Merás, Trygve Helgaker, and Ove Christiansen. The integraldirect coupled cluster singles and doubles model. *J. Chem. Phys.*, 104:4157, 1996.
- [56] Henrik Koch, Ove Christiansen, Poul Jørgensen, and Jeppe Olsen. Excitation energies of BH, CH<sub>2</sub> and Ne in full configuration interaction and the hierarchy CCS, CC2, CCSD and CC3 of coupled cluster models. *Chemical Physics Lett.*, 244:75 – 82, 1995.
- [57] Ove Christiansen, Henrik Koch, Asger Halkier, Poul Jørgensen, Trygve Helgaker, and Alfredo Sánchez de Merás. Largescale calculations of excitation energies in coupled cluster theory: The singlet excited states of benzene. *J. Chem. Phys.*, 105:6921, 1996.
- [58] M. Crouzeix, B. Philippe, and M Sadkane. The Davidson method. *SIAM J. Sci. Comput.*, 15:62–76, 1994.
- [59] Anna I. Krylov, C. David Sherrill, Edward F. C. Byrd, and Martin Head-Gordon. Size-consistent wave functions for nondynamical correlation energy: The valence active space optimized orbital coupled-cluster doubles model. *J. Chem. Phys.*, 109:10669, 1998.
- [60] Arteum D. Bochevarova and Richard A. Friesnerb. The densities produced by the density functional theory: Comparison to full configuration interaction. *J. Chem. Phys.*, 128:034102, 2008.
- [61] Vincent Thry, Daniel Rinaldi, Jean-Louis Rivail, Bernard Maignet, and Gyrgy G. Ferenczy. Quantum mechanical computations on very large molecular systems: The local self-consistent field method. *J. Comput. Chem.*, 15: 269282, 1994.

- [62] Robert J. Hinde. Interaction-induced dipole moment of the Ar-H<sub>2</sub> dimer: Dependence on the H<sub>2</sub> bond length. *J. Chem. Phys.*, 124:154309, 2006.
- [63] William M. Haynes, editor. *CRC Handbook of Chemistry and Physics*. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742, 93 edition, June 2012.
- [64] P. G. Szalay, A. Karpfen, and H. Lischka. SCF and electron correlation studies on structures and harmonic in-plane force fields of ethylene, trans 1,3-butadiene, and all-trans 1,3,5-hexatriene. *J. Chem. Phys.*, 87:3530, 1987.
- [65] Michael J. G. Peach, Peter Benfield, Trygve Helgaker, and David J. Tozer. Excitation energies in density functional theory: An evaluation and a diagnostic test. *J. Chem. Phys.*, 128:044118, 2008.
- [66] Reinhold F. Fink, Johannes Pfister, Hong Mei Zhao, and Bernd Engels. Assessment of quantum chemical methods and basis sets for excitation energy transfer. *Chem. Phys.*, 346:275 – 285, 2008.
- [67] Decanal - PubChem, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?sid=85230839>.
- [68] trans-ethyl-i-butyl-diazene - pubchem, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=524240>.
- [69] tert-butylhydroperoxide - pubchem, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=6410>.
- [70] Avogadro - free cross-platform molecule editor, July 2013. URL [http://avogadro.openmolecules.net/wiki/Main\\_Page](http://avogadro.openmolecules.net/wiki/Main_Page).
- [71] (2e,4e,6e,8e)-2,4,6,8-decatetraene - pubchem, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=13095348>.
- [72] 2,4,6,8-decatetraenal - pubchem, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=5283353>.

- [73] 1,3-pubchem - octadiene, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=517653>.
- [74] 6-methyl-2,4-heptanedione - pubchem, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?sid=87572623>.
- [75] 2,4-octadienal - pubchem, July 2013. URL <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=5283329>.
- [76] Weitao Yang. Direct calculation of electron density in density-functional theory. *Phys. Rev. Lett.*, 66:1438–1441, 1991.
- [77] Robert Zalesny, Manthos G. Papadopoulos, Paul G. Mezey, and Jerzy Leszczynski, editors. *Linear-Scaling Techniques in Computational Chemistry and Physics*. Springer Netherlands, P.O. Box 989, 3300 AZ Dordrecht, The Netherlands, 2011.
- [78] Kenny B. Lipkowitz, Thomas R. Cundari, Christian Ochsenfeld, Jörg Kussmann, and Daniel S. Lambrecht. *Linear-Scaling Methods in Quantum Chemistry*, volume 23, chapter Linear-Scaling Methods in Quantum Chemistry. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2007.
- [79] Henrik Koch. Multi-level coupled cluster models. In *Very Accurate and Large Computations and Applications*, 2013.
- [80] Björn O. Roos, Kerstin Andersson, Markus P. Fülcher, Per åke Malmqvist, Luis Serrano-Andrés, Kristin Pierloot, and Manuela Merchán. *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*, volume 93, pages 219–331. John Wiley & Sons, Inc., 1996.
- [81] C. E. Dykstra and B. Kirtman. Local quantum chemistry. *Annu. Rev. Phys. Chem.*, 41:155–174, 1990.
- [82] Henrik Koch and Esper Dalgaard. Linear superposition of optimized non-orthogonal slater determinants for singlet states. *Chem. Phys. Lett.*, 212:193–200, 1993.

- [83] Eirik Hjertenæs. Part c: Development of ab initio computational methods based on non-orthogonal slater determinants. Master's thesis, NTNU, 2011.