

# Multireference Approximate Variational Coupled Cluster Theories

- A thesis submitted for the degree of Doctor of Philosophy -

- by -

Andrew Jenkins

Cardiff University

December 2013



## **Abstract**

Multireference Configuration Interaction (MRCI) is widely and successfully used for the accurate computation of molecular energies and properties when chemical bonds are broken, and in other cases of quasi-degeneracy. However, it suffers from a lack of extensivity, which introduces significant errors for large molecules. Currently feasible Multireference Coupled Cluster (MRCC) approaches, which address the extensivity issue, are complicated and expensive. In this work we investigate simple approximations to MRCC, generalising the single-reference linked pair functional theories to the multi-reference case. We show how to develop a simple extension to MRCI that gives approximate extensivity, retains orbital invariance and acts as an approximation to a multireference variational Coupled Cluster theory.



# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>1</b>  |
| 1.1      | Quantum Chemistry Basics . . . . .                                   | 2         |
| 1.1.1    | The Schrödinger Equation . . . . .                                   | 2         |
| 1.1.2    | Slater Determinants . . . . .  | 3         |
| 1.1.3    | Hartree-Fock Theory . . . . .  | 4         |
| 1.2      | Electron Correlation . . . . .                                       | 7         |
| 1.2.1    | Dynamic Correlation . . . . .  | 7         |
| 1.2.2    | Static Correlation . . . . .   | 9         |
| 1.3      | Capturing Correlation: Single-Reference Methods . . . . .            | 14        |
| 1.3.1    | Configuration Interaction and Size-Extensivity . . . . .             | 14        |
| 1.3.2    | Coupled Cluster . . . . .  | 18        |
| 1.3.3    | CI Corrections and Pair Theories . . . . .                           | 20        |
| 1.3.4    | Variational Coupled Cluster and the Linked Pair Functional . . . . . | 28        |
| 1.4      | Capturing Correlation: Multi-Reference Methods . . . . .             | 36        |
| 1.4.1    | MRCC . . . . .   | 38        |
| 1.4.2    | MRCI . . . . .   | 40        |
| 1.4.3    | Multireference Corrections . . . . .                                 | 42        |
| <b>2</b> | <b>A Multireference Approximate VCC Theory</b>                       | <b>55</b> |
| 2.1      | Spin-Orbital Derivation of MR3VCC . . . . .                          | 56        |
| 2.1.1    | Isolation of Unlinked Parts of VCCD . . . . .                        | 56        |
| 2.1.2    | Development in Cumulants . . . . .                                   | 59        |
| 2.1.3    | Approximate MRVCC 3rd Order Energy Terms . . . . .                   | 66        |
| 2.1.4    | Two Electron Case . . . . .  | 68        |
| 2.1.5    | Positivity of the Transformation Matrix . . . . .                    | 74        |

---

|          |   |            |
|----------|---|------------|
| 2.2      | Spin Free Formulation . . . . .                       | 81         |
| 2.3      | Redundancies and Conclusions . . . . .                | 86         |
| <b>3</b> | <b>A Multireference LPF+1 Theory</b>                  | <b>89</b>  |
| 3.1      | Preliminaries . . . . .                               | 89         |
| 3.1.1    | Extending LPF . . . . .                               | 91         |
| 3.1.2    | Operator Considerations . . . . .                     | 95         |
| 3.1.3    | Redundancy Removal . . . . .                          | 100        |
| 3.2      | Construction of the Transformation . . . . .          | 102        |
| 3.2.1    | Form of $\eta$ using Orthogonal Pairs . . . . .       | 102        |
| 3.2.2    | $\eta$ in the Limiting Cases . . . . .                | 104        |
| 3.2.3    | The Transformation Matrix . . . . .                   | 105        |
| 3.2.4    | Diagonals of $\mathbf{U}$ . . . . .                   | 106        |
| 3.2.5    | Positivity of the Transformation Matrix . . . . .     | 106        |
| 3.2.6    | Complexity of Making $\mathbf{U}$ . . . . .           | 107        |
| 3.2.7    | Transformation Model . . . . .                        | 109        |
| 3.2.8    | As an Approximation to VCC . . . . .                  | 114        |
| 3.3      | Internals Energy . . . . .                            | 116        |
| 3.3.1    | Exactness for 2-Electrons . . . . .                   | 117        |
| 3.3.2    | Single Reference Limit . . . . .                      | 120        |
| 3.4      | Full Energy Expression . . . . .                      | 122        |
| 3.4.1    | Full Energy Evaluation for the Model System . . . . . | 122        |
| 3.5      | Direct CI . . . . .                                   | 125        |
| 3.5.1    | Pair Residual . . . . .                               | 126        |
| 3.5.2    | Internals Residual . . . . .                          | 133        |
| 3.5.3    | Subspace Method . . . . .                             | 135        |
| 3.5.4    | Singles . . . . .                                     | 139        |
| 3.6      | Proposed Tests . . . . .                              | 140        |
| 3.6.1    | Size Consistency and Extensivity . . . . .            | 140        |
| 3.6.2    | Potential Energy Curves . . . . .                     | 141        |
| 3.7      | Comments and Future Development . . . . .             | 143        |
| <b>4</b> | <b>Conclusions and Comments</b>                       | <b>145</b> |

|   |            |
|---|------------|
| <b>Appendix</b>                                   | <b>149</b> |
| 2-Electron Multi-Reference Model System . . . . . | 149        |
| List of Acronyms . . . . .                        | 151        |
| <b>Bibliography</b>                               | <b>153</b> |



# Chapter 1

## Introduction

*“The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed”*

– P. A. M. Dirac, 1929

Quantum chemistry involves the application of the principles of quantum mechanics to problems in chemistry and chemical physics. This application offers the possibility to describe real chemical processes and phenomena, complementing existing experimental techniques and in many cases providing data on systems as yet inaccessible to experiments.

To describe a molecule on a quantum mechanical level, the relativistic-Schrödinger (Dirac) equation must be solved. However, for practicality, approximations must be made such as use of the Born-Oppenheimer (clamped nucleus) approximation, and neglect of relativistic effects. The standard methods in electronic structure theory therefore attempt to solve the non-relativistic time-independent Schrödinger equation for the electrons in an atom or molecule only.

The simplest and least computationally demanding methods use a single Slater determinant to form their electronic wavefunction, with methods such as Configuration Interaction (CI) and Coupled Cluster designed to further improve upon the single Slater

determinant theory. These methods have been successful at predicting equilibrium geometries, and properties of molecules near equilibrium. However, these simple methods have been shown to fail in circumstances corresponding to areas of chemical interest. As such, more demanding multireference methods have been developed to describe bond breaking processes, excited states, biradicals and photochemical processes, but the most popular multireference methods still contain systematic errors and unwanted features. These errors can be both quantitative and qualitative, leading to flawed and invalid conclusions.

The work presented in this thesis attempts to correct errors in the widely used Multi-Reference Configuration Interaction (MRCI) method, while simultaneously extending the Linked Pair Functional (LPF) theory to the multireference domain. In order to explain and give context to this work, a review of common methods currently used in electronic structure theory is given, highlighting the defects and discussing methods relevant to the work presented here.

## 1.1 Quantum Chemistry Basics

### 1.1.1 The Schrödinger Equation

The central equation of quantum mechanics is the Schrödinger equation. It is shown here in its time-independent form.

$$\hat{H} \psi = E \psi \quad (1.1)$$

To solve this equation for molecules, the Hamiltonian operator is defined,

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^n \nabla_i^2 - \frac{1}{2} \sum_{A=1}^M \frac{1}{m_A} \nabla_A^2 - \sum_{i=1}^n \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^n \sum_{j=1}^{i-1} \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B=1}^{A-1} \frac{Z_A Z_B}{R_{AB}}. \quad (1.2)$$

$n$  number of electrons,  $M$  number of nuclei,  $m_A$  relative mass of A,  $Z_A$  charge of nucleus A.

Solution of this equation gives the molecular wavefunction,  $\psi$ , and the total energy of the molecule.

## Born-Oppenheimer Approximation

The time-independent Schrödinger equation using the Hamiltonian defined above requires the solution of a total molecular wavefunction. To simplify this, the clamped nucleus condition may be used, which fixes the nuclear geometry, and solves only the electronic Schrödinger equation, where the nuclei-nuclei interaction can easily be computed as the nuclei are considered as fixed point charges simply defining a potential in which the electrons move. This approximation may be rationalised by considering the relative motion of the nuclei and electrons; the nuclei are more massive and slow moving than the electrons, so as an approximation may be considered stationary. This simplifies the solution of the Schrödinger equation, which is solved under this approximation using the modified Hamiltonian,

$$\hat{H}_{e^-} = -\frac{1}{2} \sum_{i=1}^n \nabla_i^2 - \sum_{i=1}^n \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^n \sum_{j=1}^{i-1} \frac{1}{r_{ij}}. \quad (1.3)$$

$$\hat{H}_{e^-} \psi_{e^-} = E_{e^-} \psi_{e^-} \quad (1.4)$$

The total energy of the system is the sum of the electronic energy and the nuclear repulsion term.

$$E_{\text{Tot}} = E_{e^-} + \sum_{A=1}^M \sum_{B=1}^{A-1} \frac{Z_A Z_B}{R_{AB}} \quad (1.5)$$

### 1.1.2 Slater Determinants

The wavefunction of two separated electrons can be written as a product of the wavefunctions of the individual electrons.

$$\Psi(1, 2) = \phi_x(1) \phi_y(2) \quad (1.6)$$

Electrons belong to the class of particle known as Fermions, which possess the property of Fermionic antisymmetry, where interchange of electrons results in a change in the sign of the wavefunction.

$$\Psi(1, 2) = -\Psi(2, 1) \quad (1.7)$$

This has no effect on the probability density  $\Psi^* \Psi$ .

However, it is not possible to distinguish between two electrons, electron “1” may be described by  $\phi_x$  and “2” by  $\phi_y$  or vice versa, so the total wavefunction must be written as a linear combination of the two cases

$$\Psi(1, 2) = \phi_x(1)\phi_y(2) - \phi_y(1)\phi_x(2) \quad (1.8)$$

As a result of this property, the total wavefunction can be written as a determinant

$$\Psi(1, 2) = \begin{vmatrix} \phi_x(1) & \phi_y(1) \\ \phi_x(2) & \phi_y(2) \end{vmatrix} \quad (1.9)$$

This can be generalised to any number of electrons (with correct normalisation),

$$\Psi(1, 2, \dots, n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \cdots & \phi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \cdots & \phi_n(n) \end{vmatrix} \quad (1.10)$$

Linear algebra tells us that interchange of two rows or columns of a determinant changes the sign of the determinant, which makes them ideal for expressing electronic wavefunctions, as satisfaction of the constraint of Fermionic antisymmetry is built in.

### 1.1.3 Hartree-Fock Theory

In Hartree-Fock theory[1, 2] the independent particle model introduced by use of a Slater Determinant is used as the basis to describe n interacting electrons.

The energy associated with the Slater determinant (Hartree-Fock) wavefunction  $|\Psi_0\rangle$  can be found via minimisation of the standard quantum mechanical energy.

$$E_{HF} = \frac{\langle \Psi_0 | \hat{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \quad (1.11)$$

The central aim of Hartree-Fock theory is to construct the optimal set of 1-electron spinorbitals  $\{\phi_i\}$  that define the Slater determinant, to order to minimise the standard quantum mechanical energy. The spinorbitals are required to solve the eigenvalue equation, giving the energy of each spinorbital.

$$\hat{f}\phi_i = \epsilon_i\phi_i \quad (1.12)$$

The Fock-operator is defined as the core-Hamiltonian plus the Hartree-Fock potential,

$$\hat{f} = \hat{h}(1) + \sum_j \hat{J}_j(1) - \sum_j \hat{K}_j(1), \quad (1.13)$$

The one electron Hamiltonian is given by,

$$\hat{h}(1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}}, \quad (1.14)$$

and the Coulomb operators  $\hat{J}_i(1)$  and Exchange operators  $\hat{K}_i(1)$  are defined,

$$\hat{J}_i(1)\phi_j(1) = \phi_j(1) \int \phi_i^*(2) \frac{1}{r_{12}} \phi_i(2) d\mathbf{x}_2, \quad (1.15)$$

$$\hat{K}_i(1)\phi_j(1) = \phi_i(1) \int \phi_i^*(2) \frac{1}{r_{12}} \phi_j(2) d\mathbf{x}_2. \quad (1.16)$$

The action of the Coulomb operator gives the Coulomb integral, which averages the Coulomb interaction between electrons over all space and spin coordinates. Here, this gives an average potential between an electron in the  $i^{\text{th}}$  spin orbital and an electron in the  $j^{\text{th}}$  orbital. The sum in the Fock operator means each electron only feels the average field of the other  $n - 1$  electrons in the other spinorbitals. The Exchange operator exchanges the electron indices, accounting for the antisymmetric nature of the wavefunction which is neglected by the Coulomb term. This operator builds in *exchange* or *Fermi correlation*, where the motion of electrons of the same spin is correlated, meaning they cannot occupy the same point in space. In Hartree-Fock, the motion of electrons of opposite spin is not correlated.

After some manipulation (see [3] for details), the problem of finding the set of orbitals reduces the solution of the integro-differential equation

$$\hat{f} |\phi_i\rangle = \epsilon_i |\phi_i\rangle \quad (1.17)$$

Solution of this set of equations yields the set of canonical orbitals. Aside from the case of the hydrogen atom, the form of the orbitals is unknown, so they are approximated as a linear combination of known functions, the basis functions.

$$\phi_i = \sum_{j=1}^{x < \infty} c_{ij} \chi_j \quad (1.18)$$

If the set of  $x$  basis functions (the basis set) is complete then the exact orbitals can be found. However, we are of course limited to a finite set of basis functions, meaning

the orbitals are not exact. Re-casting the canonical equation above (eq. 1.17 via introduction of a basis converts the integro-differential equation problem to a problem of linear algebra, which allows for efficient computation. The recast equations (with a few more manipulations) give the Roothaan-Hall equations [4, 5], giving the optimal linear combination of basis functions.

$$\mathbf{F} \mathbf{C} = \mathbf{S} \mathbf{C} \epsilon \quad (1.19)$$

This matrix equation requires the evaluation of the Fock matrix elements, and an overlap matrix arising from the non-orthogonal nature of the basis functions.

$$F_{ij} = \langle \chi_i | \hat{f} | \chi_j \rangle \quad (1.20)$$

$$S_{ij} = \langle \chi_i | \chi_j \rangle \quad (1.21)$$

Solution of the Roothaan-Hall equations gives the coefficients in the linear expansion, defining the orbitals. It can be seen that the Fock matrix itself depends on the orbitals, meaning that the equations must be solved iteratively until convergence is achieved, i.e. when the input orbitals are the same as the output. For this reason the method is often referred to as the Self-Consistent Field (SCF) method.

The orbitals obtained via the SCF procedure are dependent on the basis functions in the basis set. Larger basis sets in general give more accurate orbitals, but convergence to the exact answer is slow. Therefore sets must also be chosen to suit the specific situation. For all but the smallest basis sets, the number of basis functions exceeds the number of electrons. This generates many more orbitals than electrons, hence some orbitals are occupied and make up the Slater determinant, while the remaining orbitals (usually greater in number) are left unoccupied as virtual orbitals.

The Hartree-Fock method solves the electronic Schrödinger equation, with the total energy of the system being found after adding the nuclear-nuclear Coulomb term (as is done in eq. 1.5), for the fixed nuclear geometry.

## 1.2 Electron Correlation

“The correlation energy is the difference between the Hartree-Fock energy and Full Configuration Interaction”

Statements such as the one above are often repeated and remembered. However, it does not explain explicitly where correlation energy comes from or what it actually is, and furthermore is not technically correct. An overview of electron correlation is now given.

The Hartree-Fock approximation provides a mean field description of the electronic structure and, despite the mean-field approximation, gives 99% of the energy of the molecule. By utilising the mean-field approximation, the correlated nature of the motion of electrons is mostly neglected, causing the Hartree-Fock energy to be erroneously higher than the total energy of the molecule. The Hartree-Fock method does include some correlation energy by design, as the wavefunction is written as a Slater determinant and therefore is antisymmetric with respect to the interchange of two electrons (Fermionic antisymmetry) and as a result obeys the Pauli Exclusion Principle. This is known as the Fermi correlation and is captured by Hartree-Fock, but it does not capture the Coulomb correlation relating to the instantaneous Coulomb interaction of the electrons. The description of the correlation energy can be split into descriptions of dynamic and non-dynamic(static) correlation.

### 1.2.1 Dynamic Correlation

The Hartree-Fock theory is a mean-field theory and only contains information about average positions and potentials. Therefore an electron is only feeling the mean-field of all the other electrons, i.e. each electron “knows” only the average position of the other electrons, therefore being unaffected by whether other electrons are instantaneously far away or close. Hartree-Fock therefore over-estimates the probability of electrons being close together.

In another way, Hartree-Fock assumes that the probability of finding electron 1 at a certain position and electron 2 at another position is simply the product of the two

1-electron probability densities. However, this is not the case because the position of electron 2 is correlated with that of electron 1.

Lack of dynamic correlation overestimates the probability that electrons are near to each other. As mentioned, Hartree-Fock contains some correlation as a result of obeying the Pauli Exclusion Principle, so electrons of the same spin are less likely to be near to each other, hence dynamic correlation effects are most important for electrons of opposite spin. Dynamic correlation is therefore greatest in doubly occupied orbitals that, for one reason or another, are spatially small, such an example would be in  $F_2$  where the molecular orbitals are smaller than the atomic ones, increasing the dynamic correlation energy.

Since the probability of finding 2 electrons close together is overestimated in a mean field theory such as Hartree-Fock, the repulsion energy is overestimated and the resulting total energy is higher than that of the exact answer. It is this difference that results in the common definition of correlation energy being the difference between Hartree-Fock and the exact answer. A lack of dynamic correlation causes Hartree-Fock to overestimate bond lengths and underestimate binding.

## FCI

The neglect of correlation energy means the electrons spend too much time close together, but by allowing electrons to occupy the virtual orbitals, electrons avoid each other and correlation energy is recouped. Removing an electron from an occupied orbital and allowing it to occupy a virtual orbital is an “excitation” in quantum chemistry and is used as a mathematical tool to produce excited Slater determinants. Moving a single electron is called a “single excitation”, and so on. Allowing all electrons to be excited to all orbitals, i.e. producing all excited determinants, and linearly combining them with the Hartree-Fock determinant is known as Full Configuration Interaction (FCI). FCI captures all correlation energy, giving the exact solution of the non-relativistic time-independent Schrödinger equation within the Born-Oppenheimer approximation and within the confines of the basis set.

The FCI wavefunction can be written in terms of the excited determinants, weighted by coefficients (determination of these coefficients is described later).

$$|\Psi_{FCI}\rangle = C_0|\Psi_0\rangle + \sum_{\substack{i \\ a}} C_i^a |\psi_i^a\rangle + \sum_{\substack{ij \\ ab}} C_{ij}^{ab} |\psi_{ij}^{ab}\rangle + \dots + \sum_{\substack{ij..n_o \\ ab..n_v}} C_{ij..n_o}^{ab..n_v} |\psi_{ij..n_o}^{ab..n_v}\rangle \quad (1.22)$$

where  $n_o$  ( $n_v$ ) represent the number of occupied (virtual) orbitals and  $|\psi_i^a\rangle$  represents an “excitation” of the reference determinant with the electron moving from orbital  $i$  to orbital  $a$ . FCI calculations are not feasible beyond small molecules due to their computational demand. The cost of the calculation scaling as  $n$  factorial ( $n!$ ). Perturbation theory shows the hierarchy of excitations needed. At first order only determinants corresponding to double excitations from the reference function are included, due to the Hamiltonian only containing at most interactions of two bodies. Second order perturbation theory says to include excitations from single up to quadruple. Since these excitations are deemed to be the most important, methods which use “truncated excitation operators”, that only excite to certain levels have developed, the most widely known being truncated Configuration Interaction (CI) and Coupled Cluster (CC). These methods are known to perform well at capturing the dynamic correlation energy, which is the dominant form of correlation energy around the equilibrium geometry, but they perform less well, and are very poor approximation of FCI when other forms of correlation energy become important.

## 1.2.2 Static Correlation

When a molecule forms a bond, the atomic valence orbitals overlap and molecular orbitals are formed. There exists a gap between the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO). When this gap is large, the wavefunction is dominated by a single configuration, i.e. the FCI expansion gives a dominant weighting to a specific determinant. In this case the valence electrons seem to move in a mean-field potential of the other electrons. This can be adequately described qualitatively by a theory based on a single reference determinant, and in many cases by a single reference mean field approximation such as Hartree-Fock. When the gap between HOMO and LUMO is small, close to degeneracy, the wavefunction is no longer dominated by a single configuration and is instead a super-

position of several configurations similar in energy. This correlation of electrons is called static correlation.

The effects of static correlation are most visible when studying reaction pathways and bond stretching. In most cases when atoms are placed at long bond lengths, states become nearly degenerate, leading to multi-reference behaviour. It is for this reason that static correlation is also known as “long range” correlation. A simple and much quoted example of a situation of static correlation is that of dissociation of  $\text{H}_2$  to open-shell products. The Hartree-Fock wavefunction for the  $X^1\Sigma_g^+$  ground state can be written,

$$\psi_G = \hat{\mathcal{A}} \sigma_g^\alpha(1) \sigma_g^\beta(2) \quad (1.23)$$

$$\sigma_g = Z_{\sigma_g} (\chi_A + \chi_B) \quad (1.24)$$

where  $\chi_X$  is a basis function centred on atom  $X$ ,  $Z_{\sigma_g}$  is the normalisation factor,  $\hat{\mathcal{A}}$  is the antisymmetrizing operator and  $\alpha$  and  $\beta$  indicate the spin.

Asymptotically becoming

$$\psi_G \sim Z_{\sigma_g}^2 \hat{\mathcal{A}} \left( 1s_A^\alpha 1s_B^\beta + 1s_A^\beta 1s_B^\alpha + 1s_A^\alpha 1s_A^\beta + 1s_B^\alpha 1s_B^\beta \right) \quad (1.25)$$

The last two terms have both electrons centred on one atom, representing homolytic fission of the diatomic. This causes an overestimation of the energy and a qualitatively incorrect potential energy surface (PES).

A way of capturing static correlation upon dissociation of  $\text{H}_2$  to ensure a qualitatively correct description of the process must be found. Looking at the  $^1\Sigma_g^+$  excited state of  $\text{H}_2$  where both electrons are in the anti-bonding orbital, the wavefunction can be written

$$\psi_E = \hat{\mathcal{A}} \sigma_u^\alpha(1) \sigma_u^\beta(2) \quad (1.26)$$

$$\sigma_u = Z_{\sigma_u} (\chi_A - \chi_B) \quad (1.27)$$

with normalisation factor  $Z_{\sigma_u}$ .

Asymptotically becoming

$$\psi_E \sim Z_{\sigma_u}^2 \hat{\mathcal{A}} \left( 1s_A^\alpha 1s_B^\beta + 1s_A^\beta 1s_B^\alpha - 1s_A^\alpha 1s_A^\beta - 1s_B^\alpha 1s_B^\beta \right) \quad (1.28)$$

Which contains covalent and ionic terms as before for  $\psi_G$ . It can be seen that the linear combination of the two configurations,  $\psi_G - \psi_E = \sigma_g^2 - \sigma_u^2$ , cancels the ionic terms exactly and correctly describes the dissociation of  $H_2$  into 2 equal fragments in a qualitatively correct way. The exchange integral for the two electrons is small, indicating a low probability of exchange of the electrons, hence the phrase “static correlation”.

In general, for all bond lengths, a linear combination of the determinants can describe the state of the  $H_2$  molecule.

$$\Psi = c_G\psi_G + c_E\psi_E \quad (1.29)$$

The wavefunction is a linear combination of Slater determinants, this is a form of Configuration Interaction (CI) mentioned in the previous subsection, but importantly, to describe the  $H_2$  molecule qualitatively correctly at all bond lengths, more than 1 Slater determinant is required. A linear combination of similarly weighted configurations is the key way to capture static correlation, and give a qualitatively correct description of the wavefunction.

### Single- and Multi-reference Situations

It is perhaps simpler to talk about single reference situations and multireference situations. Looking at the exact Hamiltonian of the system as the Hartree-Fock (self-consistent field (SCF)) Hamiltonian plus a perturbation operator to define the missing contributions,

$$\hat{H} = \hat{H}_0 + \hat{V} \quad (1.30)$$

$$\hat{H}_0 = \sum_i \hat{f}_i \quad (1.31)$$

When the perturbation is small, i.e. the norm of the  $V$  matrix is small, then the wavefunction is dominated by a single determinant, that of the Hartree-Fock wavefunction and singly and doubly excited states will have a small amplitude. If the perturbation is large, then the wavefunction is not dominated by a single determinant and the exact wavefunction contains two or more configurations with similarly large amplitudes; this can be known as configurational quasi-degeneracy (CQD). This is considered a multireference case because multiple configurations are needed for a qualitatively correct description of the wavefunction. This behaviour arises most commonly upon bond

stretching (at long bond lengths), with excited states and biradicals [6]. In the multireference situations, where static correlation is significant and the minimal qualitatively correct wavefunction consists of 2 or more similarly weighted configurations, the dynamic correlation is also important and must also be found, using the multiconfigurational wavefunction as the reference state from which to make excitations.

## Methods

Static correlation cannot be captured by a Hartree-Fock calculation. Hartree-Fock uses only a single reference determinant as its wavefunction, so therefore cannot include other configurations of importance, therefore lacking the freedom required. The post-Hartree-Fock methods, such as CCSD and CISD, do form a wavefunction based on many configurations and therefore might be thought to capture static correlation, however, the configurations in the CCSD and CISD wavefunction are those configurations reached by single and double excitations of a single reference determinant only. To capture the static correlation a method that spans more of the space by making more excitations (i.e. approaches FCI) could be used, or, a method that spans different areas of the space to provide a more balanced wavefunction. The first approach is a continuation of the CCSD and CISD idea, to include higher excitations in the excitation operator (CCSDT,CCSDTQ, CCSDTQ...). Including higher and higher excitations will approach FCI, accounting for all correlation, however it is very slow to converge to FCI and to capture static correlation (but very accurately captures dynamic correlation). Including higher order excitations also becomes prohibitively expensive, with scaling of  $n^8$  for CCSDT,  $n^{10}$  for CCSDTQ and  $n^{12}$  for CCSDTQP.

The second approach encompasses the multi-configurational nature of the wavefunction by including different configurations in the reference function, therefore spanning different areas of the space. This approach therefore can give equal weightings to different determinants in the wavefunction and really embodies the idea that the state is not simply 1 configuration but a combination. This method of capturing static correlation is mentioned above in the section on static correlation. Multi-Configurational Self-Consistent Field (MCSCF) theory is one such approach where a linear combination of several configurations is used as the zeroth order wavefunction, with the orbitals and expansion coefficients being optimised. Further excitations can be made from the multiple configurations of MCSCF in order to capture more dynamic correlation en-

ergy, which can be done using existing techniques mentioned previously (CI and CC to form Multi-Reference CI (MRCI) and Multi-Reference Coupled Cluster (MRCC)). But importantly, these excitations are now being made from a broader range of configurations that were deemed to be important for the zeroth order approximation of the state. MCSCF itself and methods using the multi-configurational MCSCF wavefunction as the 0th order reference from which to make excited configurations are called multireference methods. Application of the method of configuration interaction to a multi-configurational reference function yielded MRCI, and attempts to apply Coupled Cluster theories to a multi-configurational reference function to give an MRCC are on-going.

A review of methods designed to capture the correlation energy and accurately describe the many body effects is now given, dealing first with methods based on a single reference determinant (1.3) before describing multireference methods (1.4)

## 1.3 Capturing Correlation: Single-Reference Methods

The Hartree-Fock approximation does not offer exact results, and in many cases such as bond breaking, can offer qualitatively wrong results. However, this approximation does provide a basis for other methods to go further and capture electron correlation.

### 1.3.1 Configuration Interaction and Size-Extensivity

#### Configuration Interaction

The FCI wavefunction can be written as linear combination of the reference and all possible excited states.

$$|\Psi_{FCI}\rangle = C_0|\Psi_0\rangle + \sum_{\substack{i \\ a}} C_i^a |\psi_i^a\rangle + \sum_{\substack{ij \\ ab}} C_{ij}^{ab} |\psi_{ij}^{ab}\rangle + \dots + \sum_{\substack{ij..n_o \\ ab..n_v}} C_{ij..n_o}^{ab..n_v} |\psi_{ij..n_o}^{ab..n_v}\rangle \quad (1.32)$$

This expansion uses the general notation of  $|\psi_{ij..n_o}^{ab..n_v}\rangle$  indicating electrons in orbitals  $i,j$  have been excited to the orbitals  $a,b$ . Orbitals  $i,j$  belong to the set of  $n_o$  occupied orbitals in the reference and  $a,b$  belong to the set of  $n_v$  virtual orbitals. The Full-CI limit is reached when all electrons have been excited to all virtual orbitals.

The coefficients in the linear combination can be determined variationally, via minimisation of the *Rayleigh quotient* with respect to the coefficients.

$$\begin{aligned} \rho &= \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = E \\ \frac{\partial \rho}{\partial \mathbf{c}} &= 0 \end{aligned} \quad (1.33)$$

As stated previously, the FCI calculation can only be performed on small molecules with small basis sets due to its  $n!$  computational scaling. Truncated forms of CI are performed where the excitation operator is truncated to include only specific excitation levels. The typical truncations are to include only doubles or to include single and double excitations, giving the methods CID and CISD.

#### General Size-extensivity Error

Size-extensivity is defined as the correct linear scaling of a property with system size as opposed to an intensive property that does not change with increasing system size. With regard to quantum chemical methods, the energy must scale correctly with system size, otherwise calculations cannot provide a constant level of accuracy across systems, therefore providing unbalanced and inaccurate descriptions of processes.

The linked diagram theorem states that a many-body theory is extensive if its energy expression contains linked diagram contributions only [7, 8] (for the energy expression this is equivalent to “connected” diagrams). This means that the energy terms must be composed of parts that are all linked/connected by sharing indices, for example the fictitious term,

$$\langle ij||ab\rangle X_{cd}^{ij} Y_{kl}^{cd} Z_{ab}^{kl} \quad (1.34)$$

is linked and connected because indices are shared across all parts, each part is linked and connected to another. A fictitious example of a disconnected and unlinked term is given below,

$$\langle ij||ab\rangle X_{ab}^{ij} Y_c^{kl} Z_{kl}^{cd} \quad (1.35)$$

This energy contribution consists of a product of connected parts but importantly they are not connected to each other, therefore one part can grow irrespective of the other, causing erroneous scaling. There also exist disconnected diagrams that are linked.

$$\langle ij||ab\rangle X_c^k \quad (1.36)$$

This example is linked because multiplication by another object can make it connected, it doesn't contain closed disconnected parts like the previous example. This leads us to the conclusion that “linked” diagrams or terms should be called “linkable” [9].

For the energy to scale correctly with system size, only linked/connected diagrams must be included. It is perhaps easiest to see if a term is connected/linked or disconnected/unlinked via diagrammatic representation, for which the reader is directed to the review of Crawford and Schaefer [9].

The size-extensivity error can be alternatively viewed for the CI case. Equations to determine the CI coefficients can be obtained by minimisation of the Rayleigh quotient

or by projection of excited determinants on the Schrödinger equation. We will look at the CI size extensivity/consistency errors in the projected equations, as this is most commonly done due to its ease of comparison with Coupled Cluster theory and other corrective methods that cannot be formulated in a functional form. To determine the amplitude of a double excitation in FCI, the equation below must be solved.

$$\begin{aligned}
& \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \Psi_0 \rangle + \sum_{k,c} C_k^c \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_k^c \rangle \\
& + \sum_{\substack{k>l \\ c>d}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle \\
& + \sum_{\substack{k>l>m \\ c>d>e}} C_{klm}^{cde} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{klm}^{cde} \rangle \\
& + \sum_{\substack{k>l>m>n \\ c>d>e>f}} C_{klmn}^{cdef} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{klmn}^{cdef} \rangle = \Delta E C_{ij}^{ab} \quad (1.37)
\end{aligned}$$

where the correlation energy can be written as a sum of the pair energies

$$\Delta E = C_{kl}^{cd} \langle \psi_{kl}^{cd} | \hat{H} - E_0 | \Psi_0 \rangle \quad (1.38)$$

The correlation energy  $\Delta E$  grows with system size  $n$ . For the size-extensivity rules to be obeyed, the left hand side of 1.37 must also scale with  $n$ . It can be seen that the left hand side term contains disconnected parts, i.e. the amplitude is independent of the 2-electron integrals. This causes erroneous scaling, as for a given orbital  $i$ , the number of non-zero coefficients involving orbitals  $i, j$ , etc, increases with  $n$ . This acts to balance the  $n$  scaling of the right hand side of 1.37. This happens at all orders, leaving the method size-extensive. This is the case in Full-CI.

Truncated-CI methods, where the excitation operator is truncated to include only lower excitations, excitations of higher rank are not included meaning the cancellation of errors to balance the equation does not occur at any order. Upon truncation to CID, the amplitude equation no longer contains the disconnected quadruple terms.

$$\langle \psi_{ij}^{ab} | \hat{H} - E_0 | \Psi_0 \rangle + \sum_{\substack{k>l \\ c>d}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle = \Delta E C_{ij}^{ab} \quad (1.39)$$

As a result, the cancellation of the disconnected terms does not occur, the right hand side of 1.39 still scales with  $n$ , but the left hand side does not. The method is therefore no longer size-extensive. A related property to that of size-extensivity is the property of size-consistency.

### Size Consistency Error

Size-consistency is the property that the simultaneous calculation of the energy of two separated subsystems must be the same as the sum of the energies calculated in separate calculations. The wavefunction for the super-system must be the product of the wavefunctions of the two separated subsystems.

A truncated CI calculation, for example CID, on the separated monomers will contain double excitations within the monomers. A CID calculation on the super-system will not allow a double excitation in monomer A simultaneous with a double excitation in monomer B because this represents a quadruple excitation in the super-system. The CID wavefunction and energy for the super-system are therefore different to those of the sum of the monomers. CID applied to the super-system does not have enough flexibility to give twice the monomer energy. When increasing the level of excitation of the monomers, the super-system must contain even higher excitations.

$$\hat{C} = \hat{C}_A + \hat{C}_B \quad (1.40)$$

$$|\Psi_0\rangle = |\psi_{0A}\rangle \cdot |\psi_{0B}\rangle \quad (1.41)$$

$$(1 + \hat{C})|\Psi_0\rangle = (1 + \hat{C}_A + \hat{C}_B)|\psi_{0A}\rangle|\psi_{0B}\rangle \quad (1.42)$$

$$\neq (1 + \hat{C}_A)|\psi_{0A}\rangle \cdot (1 + \hat{C}_B)|\psi_{0B}\rangle \quad (1.43)$$

The wavefunction of the supersystem is not a product of the subsystems as it should be, because CI represents a linear parameterisation. As a result the CI wavefunction of the supersystem does not separate into the wavefunctions of the subsystems. The lack of size-consistency has many consequences, for example the reaction  $AB \rightarrow A + B$ , the calculation of the monomer energies in effect contains disconnected quadruple excitations giving a better description of the monomers than the compound AB whose CID wavefunction contains only double excitations. For this example, endothermic reaction barriers are underestimated and that of exothermic overestimated, leading to incorrect conclusions.

To restate, CID is not size-consistent because its wavefunction is not separable. Methods to correct the size-consistency error of the popular truncated CI forms of CID and CISD focus on inclusion of quadruple excitations, as this can treat the consistency problem for CID.

### 1.3.2 Coupled Cluster

Coupled cluster theory replaces the linear parameterisation of the wavefunction present in CI with an exponential parameterisation (exponential ansätze).

$$|\Psi_{CC}\rangle = e^{\hat{T}}|\Psi_0\rangle = \sum_{n=0}^{\infty} \frac{\hat{T}^n}{n!} |\Psi_0\rangle \quad (1.44)$$

The exponential parameterisation ensures the correct behaviour on separation of the system into non-interacting subsystems, i.e. the wavefunction of the supersystem is a product of the wavefunctions of the subsystems.

$$\begin{aligned} \hat{T} &= \hat{T}_A + \hat{T}_B \\ |\Psi_0\rangle &= |\Psi_{0A}\rangle |\Psi_{0B}\rangle \\ e^{\hat{T}}|\Psi_0\rangle &= e^{\hat{T}_A + \hat{T}_B} |\Psi_{0A}\rangle |\Psi_{0B}\rangle \\ &= e^{\hat{T}_A} e^{\hat{T}_B} |\Psi_{0A}\rangle |\Psi_{0B}\rangle \\ &= e^{\hat{T}_A} |\Psi_{0A}\rangle \cdot e^{\hat{T}_B} |\Psi_{0B}\rangle, \end{aligned} \quad (1.45)$$

Coupled cluster theory is therefore size-consistent, irrespective of the level of truncation of  $\hat{T}$ .

The exponential ansätze also partially introduces the effects of higher levels of excitation through products of lower order operators. For example, including only double excitations in  $\hat{T}$  ( $\hat{T} = \hat{T}_2$ ) introduces quadruple excitations via  $\frac{1}{2!} \hat{T}_2^2$ .

We now look at the extensivity of the Coupled Cluster method. The energy and amplitude equations can be found via left projection of the reference or excited determinants.

$$\langle \Psi_0 | \hat{H}_N e^{\hat{T}} |\Psi_0\rangle = E_{corr} \quad (1.46)$$

$$\langle \psi_{ij}^{ab} | \hat{H}_N e^{\hat{T}} |\Psi_0\rangle = E_{corr} T_{ij}^{ab} \quad (1.47)$$

using the normal-ordered Hamiltonian operator  $\hat{H}_N$ , which is equivalent to  $\hat{H} - E_0$ . For simplicity, truncating the excitation operator to double excitations ( $\hat{T} = \hat{T}_2$ ),

$$\langle \Psi_0 | \hat{H}_N e^{\hat{T}} |\psi_{kl}^{cd}\rangle T_{cd}^{kl} = E_{corr} \quad (1.48)$$

$$\langle \psi_{ij}^{ab} | \hat{H}_N \left( 1 + \hat{T}_2 + \frac{1}{2} \hat{T} \right) |\Psi_0\rangle = E_{corr} T_{ij}^{ab} \quad (1.49)$$

It is easy to see that unlike the CID case, there are now disconnected quadruple term on both sides of equation 1.49. The disconnected parts of  $\hat{T}_2|\Psi_0\rangle$  cancel  $E_{corr} T_{ij}^{ab}$ , leaving a fully connected equation.

$$\langle \psi_{ij}^{ab} | \hat{H}_N \left( 1 + \hat{T}_2 + \frac{1}{2} \hat{T}^2 \right) | \Psi_0 \rangle_c = 0 \quad (1.50)$$

This results in CCD being size-extensive, and similar cancellations occur for all orders, meaning Coupled Cluster is extensive irrespective of the truncation of  $\hat{T}$ .

The form of the Coupled Cluster equations above is not that of the general formalism, and was written as such above to show similarities to other methods developed as corrections to CI, described later. The standard method of solving the Coupled Cluster equations is to define a similarity transformed Hamiltonian, in order to decouple the energy and amplitude equations, then project on the left with either the reference to form the energy equations, or an excited determinant to form the amplitude equations. For example, the CCD equations show the projective nature of traditional Coupled Cluster.

$$\langle \Psi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle = E_{\text{CCD}} \quad (1.51)$$

$$\langle \psi_{ij}^{ab} | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle = 0. \quad (1.52)$$

Traditional Coupled Cluster also has a favourable scaling of  $n_o^2 n_v^4$  where  $n_o$  and  $n_v$  represent the number of occupied and virtual orbitals respectively. The favourable scaling can be seen to be a result of the truncation of the Campbell-Baker-Hausdorff (CBH) expansion[9]. The CBH expansion takes the similarity-transformed Hamiltonian and expands it as a series of commutators[10].

$$\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}} = \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2!} [[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{3!} [[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}] + \dots \quad (1.53)$$

This expansion truncates at the fourth power of  $\hat{T}$  due to the Hamiltonian being a 2-body operator, reducing the  $n!$  complexity.

Coupled cluster is size-extensive and size-consistent, orbital invariant and exact when the excitation operator is not truncated, i.e. is of the order of the number of electrons. This has led to its widespread use, but as noted later, it is known to fail for geometries away from the equilibrium[11, 12].

### 1.3.3 CI Corrections and Pair Theories

It has been shown above that the extensivity and consistency error of CI can be seen as a result of the neglect of disconnected quadruples. Therefore a simple approach to correct the error is the inclusion of disconnected quadruples.

Quadratic Configuration Interaction(QCI) [13] simply includes the quadratic term of Coupled Cluster in the CI equations. For the case of QCID i.e.  $\hat{T} = \hat{T}_2$ , the equations become,

$$\langle \Psi_0 | \hat{H}_N \hat{T}_2 | \Psi_0 \rangle = E_{corr} \quad (1.54)$$

$$\langle \psi_{ij}^{ab} | \hat{H}_N \left( 1 + \hat{T}_2 + \frac{1}{2} \hat{T}^2 \right) | \Psi_0 \rangle = E_{corr} T_{ij}^{ab} \quad (1.55)$$

These equations are the same as the CCD equations described earlier (eq. 1.49), as initially noted by Pople [13], meaning there is no independent QCID method. When including other excitations, the QCI method differs from Coupled Cluster. A comparison of CISD, QCISD and CCSD is now given.

$$\text{CISD :} \quad \langle \Psi_0 | \hat{H}_N \hat{T}_2 | \Psi_0 \rangle = E_{corr} \quad (1.56)$$

$$\langle \psi_i^a | \hat{H}_N \left( \hat{T}_1 + \hat{T}_2 \right) | \Psi_0 \rangle = E_{corr} T_i^a \quad (1.57)$$

$$\langle \psi_{ij}^{ab} | \hat{H}_N \left( 1 + \hat{T}_1 + \hat{T}_2 \right) | \Psi_0 \rangle = E_{corr} T_{ij}^{ab} \quad (1.58)$$

$$\text{QCISD :} \quad \langle \Psi_0 | \hat{H}_N \hat{T}_2 | \Psi_0 \rangle = E_{corr} \quad (1.59)$$

$$\langle \psi_i^a | \hat{H}_N \left( \hat{T}_1 + \hat{T}_2^2 + \hat{T}_1 \hat{T}_2 \right) | \Psi_0 \rangle = E_{corr} T_i^a \quad (1.60)$$

$$\langle \psi_{ij}^{ab} | \hat{H}_N \left( 1 + \hat{T}_1 + \hat{T}_2^2 + \frac{1}{2} \hat{T}_2^2 \right) | \Psi_0 \rangle = E_{corr} T_{ij}^{ab} \quad (1.61)$$

$$\text{CCSD :} \quad \langle \Psi_0 | \hat{H}_N \left( \hat{T}_2 + \hat{T}_1^2 \right) | \Psi_0 \rangle = E_{corr} \quad (1.62)$$

$$\langle \psi_i^a | \hat{H}_N \left( \hat{T}_1 + \hat{T}_2 + \hat{T}_1 \hat{T}_2 - \frac{1}{3} \hat{T}_1^3 \right) | \Psi_0 \rangle = E_{corr} T_i^a \quad (1.63)$$

$$\langle \psi_{ij}^{ab} | \hat{H}_N \left( 1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 + \frac{1}{2} \hat{T}_2^2 + \hat{T}_1 \hat{T}_2 - \frac{1}{3} \hat{T}_1^3 - \frac{1}{12} \hat{T}_1^4 \right) | \Psi_0 \rangle = E_{corr} T_{ij}^{ab} \quad (1.64)$$

The variational property of CISD is sacrificed in order to gain size-extensivity as QCISD. QCISD has shown to perform well, but not as well as CCSD, for example,

in the calculation of spectroscopic properties of BeO[14]. This error is a result of the missing terms in QCISD compared to CCSD (as shown above). These terms are relatively simple to compute, the most demanding scaling like  $\mathcal{O}(n^6)$ , not increasing the complexity of the method. As a result, QCISD is seen as an approximation to CCSD, with CCSD is generally used and QCISD seldom used. In addition, extending to include triple excitations in the  $\hat{T}$  operator, as is often required to produce chemically accurate results, breaks the extensivity of QCISDT [15]. This further adds to the overwhelming reasons to use Coupled Cluster over QCI, as CCSDT (and above) remains extensive.

The QCI method modifies CI, to make it size-extensive, approximating Coupled Cluster. However, before Coupled Cluster theory was practical, many method to correct CI were formulated, to sit on efficient CI code available, and in some cases retain the variational property.

### CEPA Methods

To look at methods designed to correct configuration interaction errors or go beyond CI, we must first look at the full-CI equations. The equation to determine the doubles amplitude in FCI is as follows.

$$\begin{aligned}
 & \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \Psi_0 \rangle + \sum_{k,c} C_k^c \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_k^c \rangle \\
 & + \sum_{\substack{k>l \\ c>d}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle \\
 & + \sum_{\substack{k>l>m \\ c>d>e}} C_{klm}^{cde} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{klm}^{cde} \rangle \\
 & + \sum_{\substack{k>l>m>n \\ c>d>e>f}} C_{klmn}^{cdef} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{klmn}^{cdef} \rangle = \Delta E C_{ij}^{ab} \quad (1.65)
 \end{aligned}$$

As noted, full-CI is too computationally demanding for all but small systems, so the excitation operator is truncated, most commonly to include only single and double excitations. The equation to determine the doubles amplitudes in the resulting CISD method is shown,

$$\begin{aligned}
 & \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \Psi_0 \rangle + \sum_{k,c} C_k^c \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_k^c \rangle \\
 & + \sum_{\substack{k>l \\ c>d}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle = \Delta E C_{ij}^{ab} \quad (1.66)
 \end{aligned}$$

where the pairs energy is given,

$$\Delta E = \sum_{i>j, a>b} \langle \Psi_0 | \hat{H} | \psi_{ij}^{ab} \rangle C_{ij}^{ab} = \sum_{i>j} \epsilon_{ij} \quad (1.67)$$

It was shown above that the lack of extensivity of CI is a result of the energy dependent term  $\Delta E C_{ij}^{ab}$ , which contains disconnected quadruple terms. In the case of QCI the solution to the extensivity error was to introduce the quadruple terms as a product of double excitations. The methods described below do not include all quadruples, they include only those needed to cancel the  $\Delta E C_{ij}^{ab}$  term, i.e. the disconnected terms. The disconnected terms can be included via the use of the Cluster condition:

$$C_{ijkl}^{abcd} \approx C_{ij}^{ab} C_{kl}^{cd} \quad (1.68)$$

Including the approximate quadruples, the amplitude equation becomes,

$$\begin{aligned} & \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \Psi_0 \rangle + \sum_{k,c} C_k^c \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_k^c \rangle \\ & + \sum_{\substack{k>l \\ c>d}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle \\ & + \sum_{\substack{k>l \\ c>d \\ \neq ij, \neq ab}} C_{ij}^{ab} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle = \Delta E C_{ij}^{ab} \end{aligned} \quad (1.69)$$

Inserting the correlation energy from 1.67 as a sum of pair energies,

$$\begin{aligned} & \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \Psi_0 \rangle + \sum_{k,c} C_k^c \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_k^c \rangle \\ & + \sum_{\substack{k>l \\ c>d}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle \\ & + \sum_{\substack{k>l \\ c>d \\ \neq ij, \neq ab}} C_{ij}^{ab} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle = \sum_{k>l, c>d} C_{ij}^{ab} C_{kl}^{cd} \langle \Psi_0 | \hat{H} | \psi_{kl}^{cd} \rangle \end{aligned} \quad (1.70)$$

The last term on the left-hand-side looks almost identical to the right-hand-side, apart from the restricted summations indicating the neglect of Exclusion Principle Violating (EPV) terms on the left hand side. The Exclusion Principle Violating (EPV) terms are those where the same electron is created/annihilated more than once, which violates the Pauli exclusion principle. These terms therefore partially cancel.

For the moment we assume these terms cancel exactly. The result is equations are those of CEPA(0) [16], giving the amplitude and energy equations (including double

excitations only in the excitation operator)

$$\langle \psi_{ij}^{ab} | (\hat{H} - E_0) (1 + \hat{C}_2) | \Psi_0 \rangle = 0 \quad (1.71)$$

$$\langle \Psi_0 | \hat{H} - E_0 \hat{C}_2 | \Psi_0 \rangle + \langle \Psi_0 | \hat{C}_2^\dagger \hat{H} - E_0 \hat{C}_2 | \Psi_0 \rangle = \Delta E \quad (1.72)$$

It is clear from these equations that CEPA(0) is equivalent to a linearised Coupled Cluster method (LCCD) [17] (also known as Linear-Coupled Pair Many Electron Theory (LCPMET) [18]), where the exponential series of the cluster operator is truncated at the linear term. CEPA(0) is size-extensive as a result of inclusion of disconnected quadruples and resulting removal of the energy dependence of the amplitude equation. This method is not exact, even for 2-electrons and is known to overestimate the effect of higher excitations (linearised CC methods often do [19]). Assuming the full cancellation of the terms noted above is too much of an approximation.

We now look again at the partial cancellation of the quadruple terms in 1.70. the partial cancellation leaves only disconnected EPV terms. In simplified terms, where  $(C_4)_D$  symbolises the disconnected quadruple terms (with no EPVs) and  $(C_4)_D^{EPV}$  the disconnected EPV terms, the cancellation between the two terms occurs,

$$((C_4)_D) - ((C_4)_D + (C_4)_D^{EPV}) = -(C_4)_D^{EPV} \quad (1.73)$$

The doubles amplitude equation can then be written,

$$\begin{aligned} & \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \Psi_0 \rangle + \sum_{k,c} C_k^c \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_k^c \rangle + \sum_{\substack{k>l \\ c>d}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle \\ & - \sum_{klcd}^{\text{“restr”}} C_{ij}^{ab} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle = 0 \end{aligned} \quad (1.74)$$

$$\langle \psi_{ij}^{ab} | (\hat{H} - E_0) (1 + \hat{C}_1 + \hat{C}_2) | \Psi_0 \rangle - C_{ij}^{ab} \sum_{klcd}^{\text{“restr”}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle = 0 \quad (1.75)$$

where “restr” indicates at least one of the indices  $klcd$  must match  $ijab$ , giving an EPV term. If the EPVs are included exactly, there is full cancellation of the two terms. The problem lies in finding the connected EPV terms, with many methods designed to approximate them in different ways. The language of the literature can be confusing

when referring to “inclusion” and “exclusion” of EPV terms. To restate, EPV terms remain in equation 1.75. These are often neglected but for a better description of the correlation energy, they should be calculated. The Coupled Electron-Pair Approximation (CEPA) methods were originally designed as size-extensivity corrections for single reference CI that also improved the correlation energy, before Coupled Cluster seemed feasible. The different CEPA methods aim to capture the EPV terms (mentioned above) either completely or partially via different adjustments of the equations.

$$\text{EPV Terms:} \quad C_{ij}^{ab} \sum_{klcd}^{\text{“restr”}} C_{kl}^{cd} \langle \psi_{ij}^{ab} | \hat{H} - E_0 | \psi_{kl}^{cd} \rangle = C_{ij}^{ab} R_{ij}^{ab} \quad (1.76)$$

The simplest CEPA approach is that of CEPA(0) described above. This method actually neglects EPV terms completely as they are quadratic (but also implicitly includes them in order for the cancellation of disconnected quadruples to occur in equation 1.75). Further CEPA methods attempt to capture the EPV terms in order to better approximate the quadruple excitations, negating CEPA(0)’s overshoot of the effects of higher excitations, therefore give a more accurate energy. Several CEPA approaches exist [20–24], with references [25, 26] providing a detailed review. The CEPA(0) approach was outlined here in the single reference formalism due to connection to its multireference version and the work carried out in this thesis which looks to extend a method, with a CEPA(0)-like energy, to the multireference case.

### The Functional Form of the Energy

The description of the extensivity error in CI and the corrective methods has focused on projected equations. Alternatively, the cause of the CI extensivity error can be clearly and most easily seen from the functional form of the energy.

$$E = E_0 + \frac{\langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle}{1 + \langle \Psi_c | \Psi_c \rangle} \quad (1.77)$$

$$= E_0 + \frac{\langle \Psi_0 | (1 + \hat{C})^\dagger (\hat{H} - E_0) (1 + \hat{C}) | \Psi_0 \rangle}{1 + \langle \Psi_0 | (1 + \hat{C})^\dagger (1 + \hat{C}) | \Psi_0 \rangle} \quad (1.78)$$

The right hand side of 1.66 contained disconnected terms, which can be seen in the functional form because the denominator is totally independent of the numerator, sharing no indices, hence the denominator and numerator are not connected/linked, meaning the denominator can grow independently of the numerator leading to disconnected

energy contributions that do not scale correctly with system size. The numerator scales with  $n$ , whereas the denominator scales as  $1 + n$ . By expanding the denominator in a binomial series the disconnected nature of the energy can be seen.

$$E = E_0 + \langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle (1 - \langle \Psi_c | \Psi_c \rangle + \mathcal{O}(\langle \Psi_c | \Psi_c \rangle^2)) \quad (1.79)$$

The second term on the right hand side can grow independently of the first term.

$$\Delta E = \langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle \langle \Psi_c | \Psi_c \rangle \quad (1.80)$$

From this outline of the cause for CI extensivity errors, it is not a surprise that methods for correction focus on inclusion of quadruples (including capture of EPV terms to better approximate the quadruples), removal of energy dependence in the amplitude equations and introducing a linked denominator for the functional form of the energy.

As stated, when looking at the functional form of the CI energy, it is easy to see that the denominator contains a  $\langle \hat{C}^\dagger \hat{C} \rangle$  term that is not linked to the numerator, causing size extensivity problems. Accordingly there has been effort to form a functional where the denominator is linked to the numerator, leading to correct scaling. The following methods are most clearly viewed in their functional form as opposed to their projected equations, so are discussed here in that form.

The Coupled Pair Functional (CPF) method of Ahlrichs *et al* [27, 28] uses a partial normalisation procedure. The denominator of the energy functional contains a topological factor,  $T_{PQ}$ , that scales the contributions to pair part of the denominator.

$$F[\Psi_c] = 2 \sum_P \frac{\langle \Psi_0 | \hat{H} - E_0 | \psi_P \rangle}{N_P} + \sum_{P,Q} \frac{\langle \psi_P | \hat{H} - E_0 | \psi_Q \rangle}{M_P M_Q} \quad (1.81)$$

$$N_P = 1 + \sum_Q T_{PQ} \langle \psi_Q | \psi_Q \rangle \quad (1.82)$$

$$M_P = \sqrt{N_P} \quad (1.83)$$

The case of  $T_{PQ} = 1$  returns the CI(SD) energy functional and a choice of  $T_{PQ} = 0$  gives CEPA(0) or Linearised-CCSD. For the general case, the factors scale the  $\langle \psi_Q | \psi_Q \rangle$  contribution based on how much the orbitals in Q are related to the orbitals in P. If the orbitals occupied in P are denoted  $i,j$  and the spin coupling denoted  $p$ , and

similarly klq for state Q, the topological factors can be found via,

$$T_{PQ} = \frac{\delta_{ik} + \delta_{il}}{2n_i} + \frac{\delta_{jk} + \delta_{jl}}{2n_j} \quad (1.84)$$

Where  $n_i$  indicate the occupation number of orbital i. The more “related” the orbitals in P are to the orbitals in Q, then the greater  $T_{PQ}$  and therefore this contributes more to the denominator. When there is little or no relation, then  $T_{PQ}$  is very small or 0. In CI as system size increases the denominator grows because all the  $\langle \psi_Q | \psi_Q \rangle$  are included in the sum for the denominator for  $\psi_P$ , whereas CPF doesn’t include (or includes to a lesser extent) the cases where  $\langle \psi_Q | \psi_Q \rangle$  is unrelated to  $\psi_P$ , i.e. the disconnected cases. Hence, the denominator of the energy functional scales more reasonably, because the energy is approximately linked/connected. The form of topological factors was chosen to make the method correct for separated electron pairs and make it invariant to unitary transformations of equivalent orbitals of identical subsystems. The idea utilised here is similar to that used by the linked pair functional[29, 30] theories described later, that form the basis for the work in this thesis.

### ACPF

Further work looked at simply scaling the  $\langle \hat{T}^\dagger \hat{T} \rangle$  term in the denominator to make the method extensive, even though it still contains disconnected/unlinked terms. The Averaged Coupled Pair Functional (ACPF) [31] was developed from CPF [27, 28] and CI.

$$F[\Psi_c] = \frac{\langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle}{1 + g \langle \Psi_c | \Psi_c \rangle} \quad (1.85)$$

If the factor,  $g$ , is 1 then CI the functional reduces to CI. The norm of the correlation part of the wavefunction in the denominator is the cause of the lack of size extensivity in CI, due to unlinked diagram contributions. It is apt to set the  $g$  factor to scale the norm to counter the norms unphysical increasing size due to unlinked contributions. A choice of  $g$  proportional to  $1/n$ , where  $n$  is the number of correlated electrons is used. Adjusting the method so it is correct for separated pairs of electrons,  $g$  is chosen to be  $2/n$ . This method therefore takes into account the EPV terms in an average way, as it distributes the correlation energy equally among all  $n/2$  non-interacting pairs, giving an average pair energy. In the language of the previous section on projected equations, the EPV terms  $C_{ij}^{ab} R_{ij}^{ab}$  are approximated by the average pair energy  $\frac{\Delta E}{(\frac{n}{2})}$ .

Again looking at the EPVs in the projected equations, ACPF averages out the exclusion principle violation terms, by using an averaged pair energy.

$$\bar{\epsilon} = \frac{\Delta E}{\frac{n}{2}} \quad (1.86)$$

The EPV terms are then approximately averaged out,

$$C_{ij}^{ab} R_{ij}^{ab} = C_{ij}^{ab} \bar{\epsilon} \quad (1.87)$$

leaving the method approximately size-extensive.

### AQCC

Average Quadratic Coupled Cluster (AQCC) theory [32] accounts for the EPV terms in a similar way to ACPF, but uses a different averaged pair energy and distributes the electron correlation differently over all pairs. The AQCC average pair energy

$$\bar{\epsilon} = \frac{\Delta E}{\binom{n}{2}} \quad (1.88)$$

The EPV terms are found by summing over electron pairs with shared indices.

$$C_{ij}^{ab} R_{ij}^{ab} = C_{ij}^{ab} \left( \binom{n}{2} - \binom{n-2}{2} \right) \frac{\Delta E}{\binom{n}{2}} = C_{ij}^{ab} \left( 1 - \frac{(n-2)(n-3)}{n(n-1)} \right) \Delta E \quad (1.89)$$

Again a functional form exists for this method, where the only difference to ACPF is the choice of  $g$  in the denominator. ACPF's choice of  $g$  is to make it exact for non-interacting pairs of electron pairs. AQCC is not exact for non-interacting pairs but is exact for 2 electrons (where there are no unlinked diagrams) and for 3-electrons (where no EPVs exist). AQCC distributes the correlation energy over more electron pairs than ACPF, it distributes over all pairs while ACPF only distributes over non-interacting pairs. ([32], Gives detailed comparison of the ACPF and AQCC energy functional and choice of  $G$ ).

### 1.3.4 Variational Coupled Cluster and the Linked Pair Functional

The Knowles research group has in recent years developed a series of methods that starting from a CEPA(0) like energy expression, introduces a linked denominator to ensure exact size extensivity while simultaneously approximating Variational Coupled Cluster (VCC) to infinite order. Successive methods approximate VCC exactly to higher order. To explain the Linked electron Pair Functional (LPF) methods [29, 30, 33], a short review of VCC is now given.

#### VCC

In Coupled Cluster theory an “exponential ansatz” is used where the excitation operator ( $e^{\hat{T}}$ ) is expanded as an exponential series  $1 + \hat{T} + \frac{\hat{T}^2}{2!} + \dots$ . This is then inserted into the Schrödinger equation to give:

$$\hat{H}e^{\hat{T}}|\Psi_0\rangle = Ee^{\hat{T}}|\Psi_0\rangle \quad (1.90)$$

In traditional Coupled Cluster, the energy is usually determined from the equation below where multiplication by  $e^{-\hat{T}}$  has occurred to help decouple the energy and amplitude equations.

$$E = \langle \Psi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle \quad (1.91)$$

The expansion coefficients are determined via projection with the excited determinants,

$$\langle \Psi_{ab}^{ij} | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle = 0 \quad (1.92)$$

This method has proved to be very accurate around equilibrium geometries and is the most widely used correlated method. However, the traditional Coupled Cluster formalism that relies on a projective method is poor when stretching bonds, i.e. when static/non-dynamical correlation energy is important, giving both quantitatively and qualitatively incorrect results. A side note is to mention that in recent years there has been much research into MRCC in order to correct bond-stretching errors etc but there are several different formalisms and each are very computationally demanding [34, 35]. The failures of traditional Coupled Cluster have been seen as a result of the truncation of the excitation operator, however, it has also been shown [12] that the

projective method itself is to blame and using a different set of equations, at the same truncation level, will yield much more reliable results. The formalism suggested is that of a Variational Coupled Cluster. A VCC is considered to be advantageous because the energy of the variational method is known to be an upper bound to the exact energy, meaning it cannot suffer from the same unphysical behaviour as Traditional Coupled Cluster (TCC) at long bond lengths where the energy can tend to minus infinity. A VCC also allows the energy to satisfy the generalised Hellmann-Feynman theorem, giving an energy functional allowing properties of the system to be computed [12]. In traditional Coupled Cluster the bra and ket of the energy equation are not complex conjugates of each other ( $e^{-\hat{T}}$  and  $e^{\hat{T}}$ ) and therefore the method will never be variational, however in VCC, instead of multiplying by  $e^{-\hat{T}}$ , the adjoint of the excitation operator is used. The VCC energy functional:

$$E_{vcc} = \frac{\langle \Psi_0 | e^{\hat{T}^\dagger} \hat{H} e^{\hat{T}} | \Psi_0 \rangle}{\langle \Psi_0 | e^{\hat{T}^\dagger} e^{\hat{T}} | \Psi_0 \rangle} \quad (1.93)$$

It is often inconvenient to have a denominator, hence the denominator is removed via complete factorisation of the numerator, causing cancellation between the numerator and the denominator. All disconnected terms in the denominator cancel exactly with the disconnected terms in the numerator, leaving a fully-connected energy expression:

$$E_{vcc} = \langle \Psi_0 | e^{\hat{T}^\dagger} \hat{H} e^{\hat{T}} | \Psi_0 \rangle_c \quad (1.94)$$

The quotient form of the energy is extensive. However, upon carrying out the series expansions of the exponential operators, the series only terminates when the level of excitation reaches the  $n$ -electron excitation term, meaning all-electrons have been replaced. This gives the method factorial complexity for all  $\hat{T}$ , irrespective of the truncation level of the excitation operator itself.

This is different to traditional coupled cluster where the expansions terminate after the product of four  $\hat{T}$  operators, due to the Campbell-Baker-Hausdorff expansion [10]. This gives the method scaling problems and is much more computationally expensive than a projective Coupled Cluster method.

The connected form of the energy is extensive but does not terminate after a product of four  $\hat{T}$  operators or even at the  $n$ -electron limit, it has an infinite series. The differences

described between the quotient and connected forms of the VCC energy functional can be seen by analysis of the expressions themselves.

For demonstrative purposes and for ease of reading, it is simplest to truncate the exponential at  $1 + \hat{T}$ , doing this shows the first terms of the true expansion. The connected form of the energy can be expanded out as follows:

$$E_{vcc} = \langle 0 | e^{\hat{T}^\dagger} \hat{H} e^{\hat{T}} | 0 \rangle_c \quad (1.95)$$

$$= \langle \hat{H} \rangle + 2 \langle \hat{H} \hat{T} \rangle_c + \langle \hat{T} \hat{H} \hat{T} \rangle_c + \mathcal{O}(\hat{T}^3) \quad (1.96)$$

$$= \langle \hat{H} \rangle + \begin{array}{c} \text{diagram 1} \\ \text{diagram 2} \\ \text{diagram 3} \\ \text{diagram 4} \end{array} + \dots \quad (1.97)$$

The quotient form of the energy can be expanded using the binomial theorem to expand the denominator (shown below), noting that the expression contains disconnected terms.

$$\frac{\langle (1 + \hat{T} + \dots)^\dagger \hat{H} (1 + \hat{T} + \dots) \rangle}{1 + \langle \hat{T}^\dagger \hat{T} \rangle} \quad (1.98)$$

$$= \frac{\langle \hat{H} \rangle + 2 \langle \hat{H} \hat{T} \rangle + \langle \hat{T}^\dagger \hat{H} \hat{T} \rangle + \dots}{1 + \langle \hat{T}^\dagger \hat{T} \rangle} \quad (1.99)$$

The numerator:

$$= \langle \hat{H} \rangle + \begin{array}{c} \text{diagram 1} \\ \text{diagram 2} \\ \text{diagram 3} \\ \text{diagram 4} \\ \text{diagram 5} \\ \text{diagram 6} \\ \dots \end{array} + \dots \quad (1.100)$$

The binomial expansion of the denominator

$$(1 + \langle \hat{T}^\dagger \hat{T} \rangle)^{-1} = (1 - \langle \hat{T}^\dagger \hat{T} \rangle + \dots) \quad (1.101)$$

$$= 1 - \begin{array}{c} \text{diagram 1} \\ \text{diagram 2} \end{array} + \dots \quad (1.102)$$

The numerator and the expansion of the denominator can then be multiplied together (note: the identity:  $\langle \hat{T}^\dagger \hat{H} \hat{T} \rangle_c = \langle \hat{T}^\dagger \hat{H} \hat{T} \rangle - \langle \hat{T}^\dagger \hat{T} \rangle \langle \hat{H} \rangle$ )

$$\left( \langle \hat{H} \rangle + 2 \langle \hat{H} \hat{T} \rangle + \langle \hat{T}^\dagger \hat{H} \hat{T} \rangle + \dots \right) \left( 1 - \langle \hat{T}^\dagger \hat{T} \rangle + \dots \right) \quad (1.103)$$

So multiplying the numerator by each part of the denominator:

(Numerator)x(1)

$$= \langle \hat{H} \rangle + \text{diagram} + \langle \hat{H} \rangle \text{diagram} + \text{diagram} + \text{diagram} + \text{diagram} + \dots$$

(Numerator)x(-⟨T†T⟩)

$$= - \langle \hat{H} \rangle \text{diagram} - \text{diagram} \text{diagram} - \text{diagram} \text{diagram} - \left( \text{diagram} + \text{diagram} + \text{diagram} \right) \left( \text{diagram} \right) + \dots \quad (1.104)$$

Adding,

$$= \langle \hat{H} \rangle + \text{diagram} - \text{diagram} \text{diagram} - \text{diagram} \text{diagram} + \left( \text{diagram} + \text{diagram} + \text{diagram} \right) \left( 1 - \text{diagram} \right) \quad (1.105)$$

By adding these parts together, it can be seen that cancellation occurs of the disconnected  $\langle \hat{T} \hat{H} \hat{T} \rangle$  term. The cancelling term came from  $\langle \hat{H} \rangle$  in the numerator, multiplied by  $-\langle \hat{T}^\dagger \hat{T} \rangle$  from the binomial expansion of the denominator. Only one multiplication

of  $-\langle \hat{T}^\dagger \hat{T} \rangle$  is shown, for simplicity, but cancellation of this kind happens at all orders, this example is illustrative and does not show all products. The disconnected part is cancelled by the connected part, this trend continues for higher orders. This shows the EPV terms in the quotient form disappear via cancellation of disconnected and connected terms. An apt question would be “With a quotient energy already existing that is extensive and truncates, why the need for the connected form of the energy?” It can often be inconvenient to have a denominator in an expression. Hence the numerator of the quotient form is factorised completely in order to cancel the denominator (Cizek 1969). However, the complete factorisation of the numerator requires summations over all indices, meaning that some unphysical diagrams are included in the sum, these are Exclusion Principle violating terms (EPV terms). The connected form of the VCC energy contains EPV terms that cause the expansion to be non-terminating because multiple creation and annihilation operations can now be applied to the same hole state. Truncation of the expansion must be carried out in order to obtain an expectation value, however, this value is no longer exact. The quotient form of the VCC energy does not contain EPV terms due to cancellation of connected and disconnected terms, as shown above. On the non-terminating nature of the connected form of the energy, the cause can be thought of as a result of the binomial expansion of the reciprocal of the denominator,  $(1 + X)^{-1}$ .

Due to the scaling problems the projective Coupled Cluster method has been widely used as it does not have the same flaw. As mentioned, it is however advantageous to have a variational CC method. Many attempts have been made to mimic or include the variational VCC structure into methods, such as the CEPA methods [36] or quadratic coupled cluster (QCC) [37]. In these methods the power series expansion of the excitation operator is truncated, in CEPA(0) it is truncated to the linear term (linearised Coupled Cluster), QCC gets its name from including up to the quadratic terms in the series expansion. The Knowles group have recently developed a set of new approximations to variational Coupled Cluster, based on the LPF method [29].

## LPF

The LPF method starts from the CEPA(0) energy equation, but LPF theory contains only connected diagrams in the evaluation of its energy. Also, instead of the standard amplitudes, it uses transformed amplitudes. The transformation of these amplitudes

is constructed as to introduce a local(linked) denominator, similar to the idea of CPF, which gives the CI denominator for 2-electrons, which is exact for that case.

$$E_{LPF} = E_0 + \langle \Psi_0 | \hat{H}_2 \hat{T} | \Psi_0 \rangle + \langle \Psi_0 |_1 \hat{T} (\hat{H} - E_0)_1 \hat{T} | \Psi_0 \rangle_c \quad (1.106)$$

Where the amplitudes used are the transformed amplitudes:

$${}_q \hat{T} | \Psi_0 \rangle = \frac{1}{4} {}_q T_{ab}^{ij} | \Psi_{ij}^{ab} \rangle \quad (1.107)$$

The transformed amplitudes are generated from the old amplitudes using the U matrix

$${}_q T_{ab}^{ij} = (U^{-\frac{q}{2}} T)_{ab}^{ij} = \frac{1}{2} (U^{-\frac{q}{2}})_{kl}^{ij} T_{ab}^{kl} \quad (1.108)$$

where U is defined using hole-density matrix elements

$$U_{kl}^{ij} = \delta_{kl}^{ij} + \Delta_{kl}^{ij} \quad (1.109)$$

$$\Delta_{kl}^{ij} = \delta_j^i \eta_l^j + \delta_l^j \eta_k^i - \delta_l^i \eta_j^k - \delta_k^j \eta_l^i - \eta_{kl}^{ij} \quad (1.110)$$

$$\eta_j^i = \frac{1}{2} T_{ab}^{ik} T_{jk}^{ab} = \eta_{jk}^{ik} \quad \eta_{kl}^{ij} = \frac{1}{2} T_{ab}^{ij} T_{kl}^{ab} \quad (1.111)$$

$$U_{kl}^{ij} = \delta_{kl}^{ij} + \Delta_{kl}^{ij} \quad (1.112)$$

$$\Delta_{kl}^{ij} = -\lambda \eta_{kl}^{ij} + \frac{(1-\lambda)}{2} \left( \delta_j^i \eta_l^j + \delta_l^j \eta_k^i - \delta_l^i \eta_j^k - \delta_k^j \eta_l^i \right) \quad (1.113)$$

$$\eta_j^i = \frac{1}{2} T_{ab}^{ik} T_{jk}^{ab} = \eta_{jk}^{ik} \quad \eta_{kl}^{ij} = \frac{1}{2} T_{ab}^{ij} T_{kl}^{ab} \quad (1.114)$$

The matrix inverse introduces division by the parts of the CI denominator that are local to a specific part of the numerator, they are coupled to the correlation of 2 specific reference orbitals. The size of the coupling between  $C_{ij}$  in the denominator to terms involving a pair of reference orbitals in the numerator ( $C_{kl}$ ) is given by the magnitude of the matrix element  $\Delta_{kl,ij}$ . From this point of view it is seen that the LPF method introduces a local denominator. Although the energy functional looks the same as CEPA(0), the transformed amplitudes are key. Binomial expansion of the local denominator introduced via the amplitude transformations shows that transformations introduce an infinite series. This transformations are designed so that the infinite series approximates the infinite series of connected VCC, and captures specific 3rd order VCC terms to make the theory exact for 2 electrons.

The  $\lambda$  parameter specifies the terms captured from 3rd order VCC, required for exactness for 2-electrons.

For 2 electrons, the following combinations of terms are true

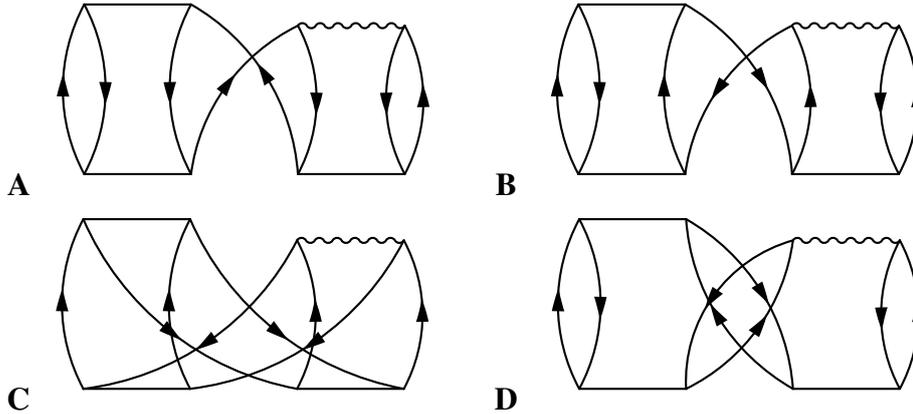


Figure 1.1: 3rd Order VCCD diagrams

$$A + D = 0 \quad (1.115)$$

$$B + 2C = 0 \quad (1.116)$$

$$A + B + C + D = B + C = -C = \frac{1}{2}B \quad (1.117)$$

Using the definition of  $\Delta$  (above) it can be seen that the  $\lambda = +1$  theory (LPF+1) gives diagram  $-C$ ,  $\lambda = 0$  (LPF0) gives  $\frac{1}{2}B$  and  $\lambda = -1$  gives  $B+C$  (LPF-1). Any choice of  $\lambda$  gives the correct result for 2-electrons. Recent research [11] has shown that the original LPF U matrix, defined above, where  $\lambda=-1$ , cannot be guaranteed to give a positive definite matrix, and hence the matrix powers or inverses are not guaranteed to exist. A choice of  $\lambda=+1$  is considered to be the most suitable choice as through that choice the U matrix is guaranteed to be positive definite, so the inverse matrix is guaranteed to exist. This choice of  $\lambda$  also makes the theory also exact for 2 holes.

The series expansion of the transformed amplitude equation gives terms identical to a subset of terms in CCD. For the 2-electron case, the expansion is identical to CID/CCD because all are exact. For that case the U matrix has the CID norms as its diagonals.

Therefore LPF approximates the connected form of the VCC energy, but does not suffer from termination problems like the connected form of the VCC energy functional, as it is written in a closed form.

The LPF method is exact for 2 electrons systems, it is size-extensive as it includes linked-diagrams only, its energy is invariant to transformations of the orbitals spaces and although it is not an upper-limit to the exact energy, it does satisfy the generalised Hellmann-Feynman theorem [38]. The LPF-doubles method also has accuracy matching VCCD [29, 30]. The Knowles group have developed a set of approximate variational coupled cluster theories based around transformations similar to LPF, including approximate variational coupled cluster theory (AVCC) [11, 39] and quasi-variational coupled cluster theory [11, 39].

## 1.4 Capturing Correlation: Multi-Reference Methods

### MR or not MR

Since different methods capture, and are optimised to capture, different forms of correlation energy, diagnostics have arisen in order to indicate if one's method of choice is appropriate, and whether a multireference method should be used. The T1 diagnostic[40] used in coupled cluster, simply takes the norm of the singles vector of CCSD, normalised by the number of correlated electrons, and uses its magnitude to tell whether a multireference method should be used, and logically, whether the result obtained with single reference coupled cluster can be trusted, or whether a multireference method should be used. This diagnostic is used after the calculation has finished, by which time a large amount of time may have elapsed, hence tools have been developed to tell whether a multireference method should be used, or not, from the outset. One such method is that of Boguslawski et al.[41] who look at the measure of orbital entanglement to determine the single- or multi-reference nature of the problem, the major correlation energy contributions and the most appropriate *ab initio* method for the problem. However, in general cases, it is often known beforehand that a single Slater determinant does not provide a qualitatively correct description of the wavefunction, in those cases, a multireference method must be used.

When the wavefunction is poorly approximated by a single determinant, e.g. Hartree-Fock, the wavefunction can be written as a linear combination of several configurations. This may resemble CI, but the difference is key. In CI, the configurations in the linear combination are those which can be reached by, for example, double excitations from the Hartree-Fock determinant. The configurations in a MCSCF wavefunction are any configurations deemed to be important for the basic qualitative description of the wavefunction. These generally consist of low-lying excited states, or configurations that are degenerate or near degenerate to the Hartree-Fock determinant. Although there is freedom to select specific configurations, a commonly used and more theoretically attractive approach is to split the orbitals of the reference determinant into inactive, active and virtual orbitals, as opposed to just occupied and virtual. The inactive orbitals are generally doubly occupied, and the virtuals unoccupied, but the active set of orbitals form an active space, where excitations are made to distribute the electrons in

the space. This generates many configurations, and if all possible ways to distribute the electrons are included then the active space is complete and the method is then referred to as the Complete Active Space Self Consistent Field (CASSCF) method. The use of a complete active space rather than selecting a subset of configurations provides a better description of the wavefunction but also importantly use of a complete active space means the method retains the invariance of the energy with respect to unitary rotations of the orbital spaces. The CASSCF method can be viewed as a FCI within an active space, with simultaneous orbital optimisation.

Considering the wavefunction as a linear combination of configurations

$$\Psi_{mcscf} = \sum_I c_I |I\rangle \quad (1.118)$$

the equation to determine the coefficients of the configurations is given as

$$\langle I | \hat{H} | J \rangle c_J = E_{corr} c_I \quad (1.119)$$

under the constraint of  $\sum_I c_I^2 = 1$ . The orbitals are optimised by parameterising the orthogonal rotations between orbitals using  $\mathbf{U} = e^{\mathbf{R}}$ , where the  $\mathbf{R}$  matrix is antisymmetric. The change in the orbitals must be stationary with respect to the change in the  $\mathbf{R}$  matrix.

$$0 = \frac{\partial E}{\partial R_{rs}} = 2(1 - \tau_{rs}) F_{rs} \quad (1.120)$$

$$F_{rs} = \sum_u D_{su} h_{ru} + \sum_{uvw} D_{suvw} (ru|vw) \quad (1.121)$$

where  $D_{su}$  and  $D_{suvw}$  are the 1 and 2 particle density matrices. These two sets of equations must be solved simultaneously to yield the MCSCF energy and orbitals. The reader is directed to references [42–44] for a review of MCSCF theory and the theoretical and computational techniques employed in its implementation.

The MCSCF/CASSCF method captures static correlation and is capable of describing processes that single-reference methods struggle with such as bond breaking. However, MCSCF/CASSCF does not capture enough dynamic correlation. Ways of capturing dynamic correlation in the single-reference case can be modified and implemented to use a MCSCF/CASSCF wavefunction as the reference function. This has been done (to differing degrees of success) for Coupled Cluster, Configuration Interaction and Perturbation Theory, yielding the MRCC methods, MRCI and MR-perturbation theory

methods such as CASPT2 [45] (MR-perturbation theories are not dealt with again in this work, and are generally not extensive). The work in this thesis is based on extensivity corrections for MRCI and approximating a multireference variational Coupled Cluster theory, we therefore must look at the current methods used in this area.

### 1.4.1 MRCC

In light of traditional single reference coupled cluster's failure when static/non-dynamical correlation energy become important and in cases of near degeneracy, a multireference implementation of a coupled cluster theory is desirable. The multireference nature would allow the correct qualitative description of the state in question and the coupled cluster ansätze would capture the dynamic correlation. However, over the last 30 years there has been much research into the formulation of a MRCC method, with no method yet satisfying all the criteria required of a multireference coupled cluster theory.

A satisfactory MRCC method would: be size-extensive, be size-consistent, be exact for 2-electrons, reduce to the single reference theory, have tractable equations, be invariant to orbital rotations.

Initial MRCC methods were multi root methods in that they computed several electronic states at once in the same computation. These methods can be split into Valence Universal (Fock space)[46] and State Universal (Hilbert space) methods. Fock space methods use an exponential operator that can generate states with different numbers of electrons; it acts in the Fock space which is the union of the Hilbert spaces of different particle number. The other main multi-root methods are the State Universal methods that are particle conserving and consider a manifold of states that have a fixed number of electrons. These are often based on the Jeziorski-Monkhorst (JM) [47] ansätze. All multi-root calculations suffer from the same flaws; they are expensive and hard to converge. These methods suffer from intruder states which are states energetically close to model space on which the excitation operators act, causing divergence in the iterative solution of the equations. This is caused by their being multiple solutions to the equations. The problems of the multi-root methods have led to the development of State Specific (SS) methods that attempt to alleviate the intruder state problem by solving for one specific state of interest only. These are in general also based on the JM

ansätze[47]. Many of these methods exist, such as Brillouin-Wigner (BW-MRCC)[48], and Mk-MRCC and its many subsequent adjustments[49–51]. In general, the effective Hamiltonian is diagonalised,

$$\sum_{\mu} \langle \Phi_{\nu} | \hat{H} e^{\hat{T}_{\mu}} | \Phi_{\mu} \rangle = E c_{\nu} \quad (1.122)$$

Where  $T_{\mu}$  and  $c_{\mu}$  are reference specific cluster operators and coefficients.

Since JM based methods use reference specific operators, there exist redundancies in the method because a specific excited determinant may possibly be reached by excitations from multiple model space configurations. Therefore the number of equations is less than the number of unknowns. The SS-MRCC methods differ in their handling of the redundancies via introduction of “sufficiency conditions“. These sufficiency conditions cause the Schrödinger equation to not be satisfied because some of the residual equations are not equal to zero. This is termed the “proper residual“ problem. Adjusting the equations to solve this problem can cause a lack of extensivity such as in BW-MRCC. Of the SS-MRCC methods Mk-MRCC has fulfilled more of the requirements of a multireference coupled cluster method. Mk-MRCC is size-extensive and consistent when localised orbitals are used, and has been shown to be accurate. However it does not fulfil the proper residual condition and is not invariant to orbital rotations[6, 52]. All JM based methods have an increased complexity compared to MRCI. As stated, these methods use reference specific cluster operators which immediately increases the number of variables by M, the size of the reference model space. This causes a computational scaling of A factorial (A!), where A is the number of active orbitals.

There has been much recent work in the groups of Gauss and Köhn [53–57] on an alternative MRCC approach which has much improved computational scaling. These methods use internal contraction, like that of icMRCI[58], where the excitation operators act on the whole reference function. This means that, unlike the JM approaches, the number of amplitudes does not increase with the size of the reference space. The several different formulations from each of the authors give chemically accurate results for traditionally multireference problems, albeit small ones. These methods are also formulated to be invariant to orbital rotations. Of particular relevance to the work presented in this thesis is the method of removal of redundancies in the excitation

operator. Linear redundancies are also present in icMRCI[58] and are removed by simple deletion of redundant pairs, however has been proved that this cannot be used for a multireference coupled cluster theory because it destroys the property of orbital invariance[53]. The authors instead use a singular value decomposition to remove the redundancies. The icMRCC methods are a basis for comparison for the work presented within this thesis, which aims to form an icMRCI size extensivity correction while simultaneously approximating a MRCC method.

For a very detailed review of multireference coupled cluster theories, the reader is directed to a review by Bartlett *et al* [6] and a review of state specific MRCC[59], and all articles in [60] and [61].

Apart from genuine multireference methods that treat a set of reference configurations on an equal footing, there have also been many mixed methods and methods that try to incorporate configurations deemed to be important to describe cases of static/non-dynamical correlation, from within frameworks that are less computationally complex than existing MRCC theories. The author points to the relevant references, [62–67].

As no MRCC method obeys all the requirements, the multireference implementation of MRCI method used since the 1980s is still used as the method of choice for multireference problems.

## 1.4.2 MRCI

The MRCI methods such as MRCISD and MRCID use the configuration interaction technique to capture dynamic correlation, but they use a reference consisting of multiple configurations, in order to capture static correlation and give a qualitatively correct description of the wavefunction. Excitations are made from all reference configurations

The MRCI wavefunction can then be written analogously to the single reference CI

wavefunction.

$$|\Psi_{MRCI}\rangle = \sum_{\mu} c_{\mu} |\mu\rangle + \sum_{\mu} \sum_{i,a} C_i^{a(\mu)} |\mu_i^a\rangle + \sum_{\mu} \sum_{i>j, a>b} C_{ij}^{ab(\mu)} |\mu_{ij}^{ab}\rangle + \dots \quad (1.123)$$

Where  $\mu$  is each reference configuration.

Alternatively, in notation that fits the rest of the thesis.

$$|\Psi_{MRCI}\rangle = \sum_I c^I |\Psi_I\rangle + \sum_S \sum_a c_a^S |\Psi_S^a\rangle + \sum_P \sum_{ab} C_{ab}^P |\Psi_P^{ab}\rangle + \dots \quad (1.124)$$

Where S and P denote internal  $n - 1$  and  $n - 2$  hole states. The external orbitals are denoted  $ab$ .

The coefficients can again be found via minimisation of the Rayleigh quotient, or equivalently by inserting the definition of the MRCI wavefunction into the Schrödinger equation and left projecting with the excited determinants.

### Internally Contracted MRCI

In internally contracted MRCI (icMRCI)[58][68] the excitation operators are not applied to each individual internal configuration, they are instead applied to the whole reference function. The name “internally contracted” applies because using this technique is equivalent to a linear combination of the states  $\Psi_P^{ab}$  with differing internal P. The number of contracted internal states is independent of the the number of internal configurations, they only depend on the number of possible  $n - 1$  and  $n - 2$  states that can be generated, which depends on the number of correlated orbitals, not explicitly on the number of reference/internal configurations. This reduces the number of coefficients in the CI expansion by 1 or 2 orders of magnitude, making more MRCI calculations feasible. icMRCI introduces a contraction error, so has an error compared to MRCI but this has been shown to be relatively small when compared to the difference to FCI[58].

$$\begin{aligned} |\Psi_{icMRCI}\rangle = & \sum_{tuvw} C^{tuvw} \hat{E}_{tu,vw} |\Psi_0\rangle + \sum_{tuva} C_a^{tuv} \hat{E}_{at,uv} |\Psi_0\rangle \\ & + \sum_p \sum_{ab} \sum_{t \geq u} C_{ab}^{tup} \frac{1}{2} \left( \hat{E}_{at,bu} + p \hat{E}_{au,bt} \right) |\Psi_0\rangle \end{aligned} \quad (1.125)$$

Using second quantised excitation operator and splitting the external sum into  $p=1$  or  $p=-1$  contributions depending on whether the configuration  $\Phi_P^{ab}$  is singlet or triplet coupled in the external space.

### 1.4.3 Multireference Corrections

Analogously to the single reference case, we look at multireference corrections to MRCI. The reader is pointed to [69] for a very good overview.

The full-CI equation can be partitioned into separate spaces. The P space contains the reference functions, the Q space contains the single and double excitations out of the P space and the R space contains all other configurations consisting of higher order excitations from the P space.

$$\frac{\langle \Psi_P + \Psi_Q + \Psi_R | \hat{H} - E_0 | \Psi_P + \Psi_Q + \Psi_R \rangle}{\langle \Psi_P + \Psi_Q + \Psi_R | \Psi_P + \Psi_Q + \Psi_R \rangle} = E_c \quad (1.126)$$

Varying the coefficients, the stationary equation must be solved.

$$\langle \Phi_i | \hat{H} - E | \Psi_P + \Psi_Q + \Psi_R \rangle = 0 \quad i \in \{P, Q, R\} \quad (1.127)$$

As discussed earlier, due to computation costs the excitation operators are normally truncated to include only single and double excitations, giving the truncated CISD equation, but an additional term can be defined that uses lower order amplitudes to approximate the higher order excitations. This section now uses the notation Bartlett[69] in defining the  $K_i$  term.

$$\langle \Phi_i | \hat{H} - E | \Psi_P + \Psi_Q \rangle + K_i \langle \Phi_i | \Psi_P + \Psi_Q \rangle = 0 \quad (1.128)$$

$$K_i = \frac{\langle \Phi_i | \hat{H} | \Psi_R \rangle}{C_i} \quad (1.129)$$

Using the Cluster Condition,  $c_j^R = c_i c_k$ , we include the disconnected parts of the higher excitations, just as in QCI and TCC described above, which are responsible for extensivity errors. The expression for  $K_i$  can then be simplified,

$$K_i = \sum_{k \in \mathcal{D}} c_k \langle \Phi_0 | \hat{H} | \Phi_k^Q \rangle \quad (1.130)$$

where  $\mathcal{D}$  denotes the disconnected set where  $i$  does not equal  $k$ . The FCI correlation energy can be written,

$$E_c = \sum_i c_i \langle \Phi_0 | \hat{H} | \Phi_i^Q \rangle \quad (1.131)$$

$K_i$  can be further simplified,

$$K_i = E_c - \sum_{i \notin \mathcal{D}} c_k \langle \Phi_0 | \hat{H} | \Phi_k^Q \rangle \quad (1.132)$$

This is a restatement of the problem of capturing EPV terms in the previous section. Using the definition of pair energies,

$$\epsilon_{ij} = \sum_{ab} C_{ij}^{ab} \langle \Psi_0 | \hat{H} | \Phi_{ij}^{ab} \rangle \quad (1.133)$$

The relationship with the previous section can be seen, where the second term the equation above are the EPV terms which the previous section aims to find in a single reference setting. Until this point the definitions are valid for both single- and multi-reference cases, depending on definitions of  $E_0$  and the reference, and can be used as an alternative way of thinking about the methods defined in the previous section. For example, setting  $K_i = E_c$  for  $i \in \mathcal{D}$  and 0 else, can be seen to give CEPA(0), where energy dependence of the equations is cancelled and the EPVs are neglected.

The definitions of the reference, the reference energy and the correlation correction will now be specific to the multireference case.

The wave function is thought of as a reference function and correlation correction. These partitions can be defined in different ways. The  $\Psi_0$  function can be defined either as an eigenfunction of the Hamiltonian in the reference P space or can include the relaxation of the reference space due to the inclusion of the correlation effects of the Q space. For the second choice, the correction  $\Psi_c$  would be just the Q space wave function  $\Psi_Q$ , while the other choice is to include the orthogonal complement of the P space in the correction

$$\Psi_c = \Psi_Q + \Psi'_P \quad (1.134)$$

Inclusion of the orthogonal complement in a full MRCC theory would cause problems but these do not present themselves in linear CEPA methods.

Several issues arise due to the multireference nature of the problem that are not present in the single reference theories. These include the problem that there is no unique P space configuration that produces a specific Q space configuration. Also, the idea of the cluster condition in a multireference method is not totally analogous to that in the single reference case and is not well defined. A related key issue involves the configuration spaces themselves. Using a multireference wavefunction, a product of Q space excitations does not necessarily produce an R space configuration, for example, a double excitation of a Q space configuration does not necessarily produce a quadruple

excitation from the P space (that would belong in the R space), this would depend on the configurations in the P space.

$$\hat{E}_j \hat{E}_i \Psi_0 \in Q \quad (1.135)$$

This causes over counting of some excitations, these terms are referred to as “redundancy terms”.

$$\langle \Phi_i^Q | \hat{C} | \Psi_0 \rangle \sum_{\substack{j \\ \hat{E}_j \hat{E}_i \Psi_0 \in Q}} \langle c_j \Psi_0 | \hat{H} - E | \Phi_j^Q \rangle = 0 \quad (1.136)$$

A full theory would subtract these redundancy terms.

Multireference methods differ in the definitions of  $\Psi_0$ ,  $E_0$ , the capture of the EPV terms (via definition of  $K_i$ ) and the accounting for redundancy terms.

### MRCI

The MRCISD method neglects higher order contributions and does not take account of EPV terms, i.e.  $K_i = 0$ . This introduces the extensivity error described earlier because unlinked term are not cancelled.

### MR-LCCM and MR-CEPA(0)

The single reference CEPA(0) and Linear Coupled Cluster are equivalent, but differ slightly for the MR case. Like the SR theory they both set  $K_i = E_c$  for the Q space, meaning the EPV terms are not subtracted so are retained in the overall theory. For the MR case, this is done for i in the Q space and  $K_i = 0$  otherwise. The difference between the multireference theories is due to the definition of  $\Psi_c$ . The MRLCCM method [19] [70] is based on the linearisation of the MRCC equations. This derivation does not include the orthogonal complement of the P space, as the full MRCC does not, so the correction only takes contributions from the Q space ( $\Psi_c = \psi_Q$ ). The MRCEPA(0) method [31] includes the orthogonal complement of the P space, and is therefore slightly different from MRLCCM despite using the same EPV capturing technique (i.e. that of not capturing them!).

$$\langle \Phi_i | \hat{H} - E + E_c | \Psi_0 + \Psi_c \rangle = \langle \Phi_i | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle = 0 \quad (1.137)$$

$$F[\Psi_c] = \langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle \quad (1.138)$$

## MR-ACPF

ACPF can be extended to the multireference case giving MRACPF[31], which acts as a size-extensivity/consistency correction to the MRCI method, and is very widely used in this form due to it decreasing the error at a trivial additional cost to the MRCI calculation. As in the single reference theory the EPV terms are accounted for in an averaged way. To account for the EPV terms,  $\frac{n}{2}$  electrons pairs are assumed to be non-interacting and all contribute equally to the correlation energy.

Defining  $K$  as

$$K_i = \frac{n-2}{n} E_c \quad i \in Q \quad K_i = 0 \quad i \in P \quad (1.139)$$

This gives the MRACPF equation,

$$\langle \Phi_i | \hat{H} - E | \Psi_0 + \Psi_c \rangle + \frac{n-2}{n} E_c \langle \Phi_i | \Psi_c \rangle = 0 \quad (1.140)$$

which can be formulated into a functional

$$F[\Psi_c] = \frac{\langle \Psi_0 + \Psi_c | \hat{H} - E | \Psi_0 + \Psi_c \rangle}{\langle \Psi_0 | \Psi_0 \rangle + \frac{2}{n} \langle \Psi_c | \Psi_c \rangle} \quad (1.141)$$

The functional may alternatively be written with the correlation correction wave function split into two a different mutually orthogonal parts,  $\Psi_a$  and  $\Psi_e$ .  $\Psi_a$  contains all determinants that have the same orbital occupation as the reference, outside the active space,  $\Psi_e$  contains all other determinants.

$$F[\Psi_c] = \frac{\langle \Psi_0 + \Psi_c | \hat{H} - E | \Psi_0 + \Psi_c \rangle}{\langle \Psi_0 | \Psi_0 \rangle + g_a \langle \Psi_a | \Psi_a \rangle + g_e \langle \Psi_e | \Psi_e \rangle} \quad (1.142)$$

The factor  $g_e$ , is set to  $2/n$  and the factor  $g_a$  is set to 1, giving the MRACPF defined earlier. The factor  $g_a$  is set to 1 with the argument that the reference is a CAS wavefunction then all combinations of double replacements are already included in the CI and hence cluster corrections are not needed for this class. Generally the reference may not be CAS but may be nearer enough for this approximation to hold, as a non-CAS choice of  $g_a$  may be difficult to define. The MRACPF method has several desirable attributes. It is invariant to orbital rotations within relevant orbital spaces, i.e. internal-internal and external-external rotations. ACPF is exact for non-interacting pairs of electrons and is size-consistent for identical subsystems using a single reference function. As shown MRACPF can also be formulated into an energy functional (the form shown

above) meaning that analytical energy derivatives are available. MRACPF can easily be related to other methods, for example, Multi-Reference Linearised Coupled Cluster (MR-LCCM) can be derived by setting  $g_e = g_a = 0$ . MRACPF is also closely related to MRAQCC (outlined below). The performance of this method has been tested substantially, and as a result has become popular as a choice of method that is approximately size extensive and consistent and of the same complexity of CI, directly utilising CI code in most cases. However, it has been shown that (MR)ACPF overestimates the effect of higher excitations[32] and also fails badly for small reference spaces. This second problem was known from the outset and was noted in the original MRACPF paper by Gdanitz & Ahlrichs([31], where it suggested using a large reference space.

### MRAQCC

Multireference Averaged Quadratic Coupled Cluster (MRAQCC)[32] builds on MRLCCM and MRACPF but is able to use a smaller reference space. Includes EPV terms arising in the quadratic part of the exact energy functional, but in an averaged way. It is simpler than a full quadratic coupled cluster, but builds towards that. As in the working for the single reference theory, a pair energy is defined where the correlation energy is equally distributed among  $\binom{2}{n}$  equivalent pairs rather than  $\frac{n}{2}$  as in (MR)ACPF.

Giving,

$$K_i = \binom{n-2}{2} \frac{E_c}{\binom{n}{2}} \quad (1.143)$$

$$\langle \Phi_i | \hat{H} - E | \Psi_0 + \Psi_c \rangle + G E_c \langle \Phi_i | \Psi_c \rangle = 0 \quad (1.144)$$

$$G = \frac{(n-2)(n-3)}{n(n-1)} \quad (1.145)$$

$$F[\Psi_c] = \frac{\langle \Psi_0 + \Psi_c | \hat{H} - E | \Psi_0 + \Psi_c \rangle}{\langle \Psi_0 | \Psi_0 \rangle + G \langle \Psi_c | \Psi_c \rangle} \quad (1.146)$$

Although (MR)AQCC is not exact for non-interacting pairs, an energy functional exists (Hellmann-Feynman theorem satisfied[71]), has analytical energy gradients, and can be applied to excited states. It can also calculate transition moments via a linear response extension MR-AQCC-LRT[72]. MR-AQCC is known to give similar results

to (MR)ACPF but performs better than MR-ACPF when a small reference space is used [32].

The multireference methods described so far are simple extensions of their single reference counterparts and have not made any attempt to account for the redundancy terms that arise as a result of the use of a multi-configurational reference function.

### MR-CEPA

The MR-CEPA method of Ruttink et al. [73] takes the idea of CEPA but tries to take account of the redundancy terms. This method splits the orbitals into 3 spaces, the inactive orbitals that doubly occupied in all references, the active whose occupation vary and the virtual that are unoccupied in all reference functions. Classes are created labelling excitations  $(k,l)$  where  $k$  is the number of holes in the inactive orbitals and  $l$  is the number of particles in the virtual orbitals. This classification is used to determine if the product of excitations belongs to the R space or the Q space, where products that remain in the Q space are the redundancy terms. The class of the product of two excitations  $E_i \in (k_1, l_1)$  and  $E_j \in (k_2, l_2)$  gives the class  $(k_1+k_2, l_1+l_2)$ . If  $k_1+k_2 \leq 2$  and  $l_1 + l_2 \leq 2$ , the term is a redundancy term. This method does not account for the redundancy terms exactly because it neglects single excitations and also retains some EPVs. Furthermore,  $k_i$  depends on  $c_k$ , meaning no functional exists, making gradients difficult to calculate.

### MR-SC2-CI [74, 75]

Accounts for EPVs and redundancy terms, so is rigorously size-extensive and consistent. Needs a lot of storage so makes it the most expensive correction (but additional cost is less than underlying MRCI calculation)[69]. Slightly more accurate than MR-AQCC and MR-ACPF [76][77]. It is an exact CEPA but not much more accurate than the approximate versions. Not associated with a functional hence no analytical derivatives. The additional accuracy is deemed to be not worth the considerable additional cost.

### *A posteriori* Corrections

The *a posteriori* size-consistency corrections are generally simple corrections to the correlation energy to account partially for quadruple excitations, applied after the con-

vergence of the MRCI procedure. As such these methods are cheap to calculate on top of underlying MRCI calculations, meaning they are easy to implement and readily used. Many of these corrections were developed from extension of single-reference corrections, but are introduced here as corrections to MRCI, as use of corrections for the single-reference case is outdated due to the use of Coupled Cluster theory.

The most commonly used *a posteriori* corrections are those devised and/or named for E. R. Davidson. The simplest Davidson Correction, by Langhoff and Davidson [78], aims to roughly negate the error via approximately including the effects of the quadruple excitations.

$$E_{DC} = (1 - c_0^2)(E_{CI} - E_0) \quad (1.147)$$

where  $c_0$  is the reference coefficient, and  $E_0$  the reference energy. The correction is therefore the magnitude of the correlation part of the wavefunction multiplied by the correlation energy. The multireference version of [79–81] can be derived from the single reference correction by the inserting for the  $n$ th state  $c_n^2 = \sum_{p \in P} c_p^2$  where  $c_p$  are the reference coefficients in the MRCI wavefunction and  $E_0$  the reference energy. Giving,

$$E_{MDC} = \left(1 - \sum_{p \in P} c_p^2\right) (E_{mrci} - E_0) \quad (1.148)$$

where  $E_0$  the reference energy. This simple correction is not normalised correctly, thus a later correction, known as the Renormalised Davidson Correction [82], whose single reference version takes the form,

$$E_{RDC} = \frac{1 - c_n^2}{c_n^2} (E_{CI}^n - E_0^n) \quad (1.149)$$

Extending this to the case for correction of MRCI can be done in several ways, two of which are discussed here. The  $c_n$  of the single reference correction can be replaced by the coefficient of the (fixed) reference function ( $c_{Rn}^0$ ) in the MRCI wave function, giving the MR fixed Renormalised Davidson Correction [83]

$$c_n = \langle \Psi_{ref}^n | \Psi_{MRCI}^n \rangle = \sum_R c_{Rn}^0 c_{Rn} \quad (1.150)$$

$$E_{RDCf}^n = \frac{1 - (\sum_R c_{Rn}^0 c_{Rn})^2}{\sum_R c_{Rn}^0 c_{Rn}} (E_{MRCI} - E_0) \quad (1.151)$$

Alternatively when the reference function may have a small overlap with the MRCI wave function, for example near avoided crossings, the relaxed wavefunction may be used [84, 85],

$$\Psi_{rlx}^n = \frac{\sum_R c_{Rn} \Phi_R}{\sqrt{\sum_R c_{Rn}^2}} \quad (1.152)$$

$$c_n = \sum_R c_{Rn}^2 \quad (1.153)$$

$$E_{RDCr}^n = \frac{1 - \sum_R c_{Rn}^2}{\sum_R c_{Rn}^2} (E_{MRCI}^n - E_0^n) \quad (1.154)$$

It is worth noting that the Davidson corrections are usually denoted as +Q, for example MRCI+Q, but care should be taken when running calculations as to which Davidson correction is implemented in the software being utilised, be it the Davidson or the Renormalised Davidson correction. The MOLPRO software package [86], used in the work presented in this thesis, uses the renormalised corrections, with a choice of fixed or relaxed coefficients.

Although there are many other corrections, such as the Meissner correction [87], the final *a posteriori* correction described here is that of the Pople correction [88] due to its use in MOLPRO it can therefore be compared against other corrections described above and against methods developed in this thesis. The single reference Pople correction can be written,

$$E_{PC} = \frac{\sqrt{n_c^2 + 2n_c \tan^2(2\theta)} - n_c}{2(\sec(2\theta) - 1)} E_{CI} - E_0 \quad (1.155)$$

where  $n_c$  is the number of correlated electrons. Unlike the Davidson corrections, this correction does disappear for the 2 electron case, when CISD is correct. Within MOLPRO, the extension to the multireference case can be made with either choice of fixed or relaxed coefficients, as outlined above.

### Performance of MR methods

The performance of the multireference methods described has been extensively studied, particularly in references [6, 26, 32, 51, 76]. These studies have shown that both MRACPF and MRAQCC out-perform MR-LCCM and MR-CEPA(0) in the prediction

of potential energy surfaces of small molecule examples. It has also been noted that MRACPF and MRAQCC perform better than many multireference Coupled Cluster methods that are more demanding [6, 53]. MRACPF and MRAQCC generally give very similar results for real systems[76], showing that the distribution of the correlation energy among all pairs instead of only over non-interacting pairs has little effect. Despite little difference between the two methods, MRACPF has been shown to fail for some examples[76], notably ozone[32], and in general under performs when a small active space is used[31, 76]. In examples with a larger active space, MRACPF tends to perform slightly better than MRAQCC but in general, has only small differences. The extension of these methods to include EPVs arising from virtual orbitals does not affect the energy greatly, and in several cases increases the consistency error[89]. A similar trend is seen for the -mc extensions, which were designed to account for the redundancy terms due to the MR nature [89]. In addition, functionals do not exist for those extensions, meaning analytic gradients can't be defined, and their applicability is limited.

In terms of the size-consistency error, MRACPF and MRAQCC reduce the size consistency error of MRCI by between 1-3 orders of magnitude [26, 89].

Of all of the multireference methods described here, the most widely used and cited are those of MRACPF and MRAQCC, due to their accuracy, existence of functionals and they can be implemented easily from existing MRCI programs and run at almost no extra cost (they have same complexity).

### **Performance of *a posteriori* Corrections**

The performance of the *a posteriori* corrections has been probed in references [87, 90, 91]. The study of Duch *et. al* [90] has shown that both the Meissner and Pople corrections outperform the Davidson corrections [87, 90, 91] for both the single-reference and multi-reference case in tests on size-consistency and size-extensivity. In accordance with this work, the Pople correction will be taken to be the benchmark for *a posteriori* corrections, and this will be compared against other multireference methods and the popular Davidson corrections.

### **Comparison**

Despite their popularity, MRACPF and MRAQCC are still only classed as “approx-

mately” extensive and consistent. The lack of size-consistency can be seen in the small molecule example of  $N_2$  and  $H_2O$ .

In this example the size-consistency error is measure by the energy difference between the energy calculated with separated molecules and the energy that is the sum of the energies of the two molecules in separate calculations.

|               | Error mH |
|---------------|----------|
| MRCISD        | 17.9696  |
| MRCISD+ $Q_f$ | 1.5887   |
| MRCISD+ $Q_r$ | 1.8323   |
| MRCISD+ $P_f$ | -0.1143  |
| MRCISD+ $P_r$ | 0.1547   |
| MRACPFSD      | -1.1183  |
| MRAQCCSD      | 0.0712   |

Table 1.1: Size-consistency error for  $N_2$  and  $H_2O$  using a cc-pVDZ basis

where “+Q” indicates the Renormalised Davidson correction and “+P” the Pople correction, with subscripts  $f$  and  $r$  indicating fixed or relaxed coefficients as described earlier. These energies were computed in MOLPRO using a cc-pVDZ basis, making use of the internally contracted MRCI method, from which the MRACPF and MRAQCC are modifications, hence all contain the contraction error. In agreement with the literature, the error in MRACPF is reduced by an order of magnitude and 2 for MRAQCC. The consistency error for MRAQCC is small ( $0.04 \text{ kcal mol}^{-1}$ ), below that of chemical accuracy ( $\sim 1 \text{ kcal mol}^{-1}$ ). The Pople correction is shown to perform much better than the renormalised Davidson corrections, in agreement with other studies.

The size-consistency error here is that when the molecules are measured at large separation, because the error can be measured easily and accurately. However, when the molecules are interacting, the error is likely to increase, and will certainly increase using larger systems. This has shown that MRAQCC has a small consistency error compared to the other methods, with the Pople correction performing well for its simple structure. Size-consistency is not the only property a method must have, it should also correctly give the shapes of potential energy curves and reproduce the FCI result as closely as possible, at least within chemical accuracy.

An appropriate example of the calculation of a potential energy curve is that for the insertion of Be into  $H_2$ . The full details of the calculation are described later in section 3.6.

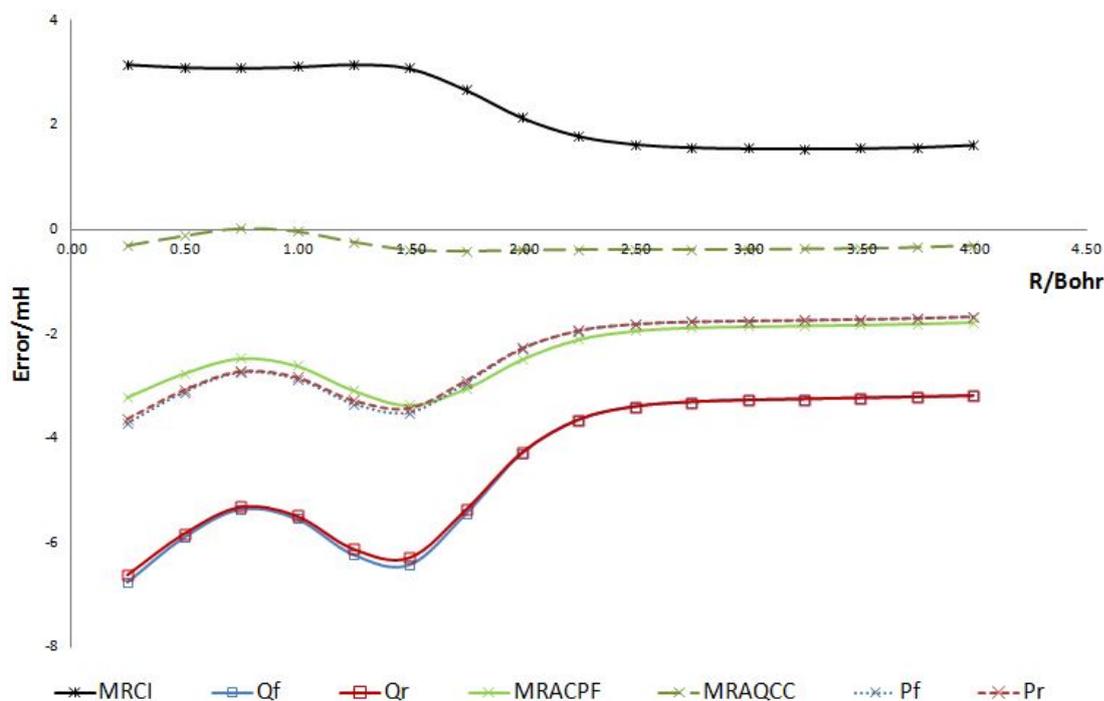


Figure 1.2: Potential energy curve errors to that of FCI for  $BeH_2$  dissociation calculated with a cc-pVDZ basis set. Qf (Qr) indicates the fixed (relaxed) renormalised Davidson correction and Pf (Pr) the fixed (relaxed) Pople correction

The mean error of each potential energy curve is given in table 1.2 (below), with a measure of how well the method reproduces the curve of the FCI calculation, this is the non-parallelism error (NPE). The NPE is calculation as the difference between the maximum error and minimum error. This measure gives some indication of how parallel the computed potential energy curve is to that of FCI, i.e. how precise it is, or conversely, how erratic the potential energy curve is.

These results agree with the stated previous studies that the Pople corrections perform better than the Davidson corrections, but MRACPF and MRAQCC out-perform *a posteriori* corrections by having a smaller mean error and NPE, better reproducing the FCI result. A recent comparative study from the Truhlar group [92] noted the good performance of MRACPF and MRAQCC but did show MRCI+Q to give quantitatively slightly more accurate results for the dissociation of  $F_2$ , however, it notes that this is due to a cancellation of errors, which cannot be relied upon in general.

|               | Mean Error/mH | NPE   |
|---------------|---------------|-------|
| MRCISD        | 2.275         | 1.617 |
| MRCISD+ $Q_f$ | -4.528        | 3.553 |
| MRCISD+ $Q_r$ | -4.489        | 3.419 |
| MRCISD+ $P_f$ | -2.417        | 2.056 |
| MRCISD+ $P_r$ | -2.389        | 1.959 |
| MRACPFSD      | -2.376        | 1.585 |
| MRAQCCSD      | -0.300        | 0.426 |

Table 1.2: Calculated BeH<sub>2</sub> potential energy curve mean and non-parallelism errors (in milli-Hartree) relative to FCI

The MRACPF energy is also below that of FCI and MRAQCC above it, as noted in previous studies. This can be explained by realising that MRACPF uses independent electron pairs to approximate the EPV terms, overestimating the contribution, while MRAQCC averages over all pairs, therefore including some interaction and reducing the error [76]. This example calculation highlights why MRAQCC is considered to be a relatively cheap alternative to the MRCI energy that performs better than MRCI. As stated earlier, MRAQCC also performs better than other corrective methods and other multireference methods. Evangelista & Gauss [53] reported results for this test system showing Mk-MRCCSD failed to reproduce the potential energy curve, breaking down at the point with most multireference character. Their results showed Mk-MRCCSD had a NPE of 2.24 mHartree and their icMRCC method an NPE of 1.04 mHartree. The same study also showed that MRACPFSD and MRAQCCSD had smaller errors than many multireference Coupled Cluster methods for the stretch of hydrogen fluoride, with only icMRCCSD performing slightly better. However, the error to FCI for the better performing multireference methods i.e. MRAQCC and MRACPF, remains chemically relevant for many systems and can only grow with increased system size. Accordingly, new rigorous corrective methods are required to give accurate and reliable results across a wider range of molecules. The work in this thesis aims to develop new methods that correct the errors outlined above, and must perform better than MRACPF, MRAQCC and the Pople and Davidson corrections, but simultaneously it goes beyond just the simple notion of correcting MRCI, to build towards a multireference Coupled Cluster theory.

Detailed analysis of the theories and performance of the methods outlined above has previously been carried out, the author specifically points to the review by Szalay[26] and also [44, 76, 89] for more information.

## Chapter 2

# A Multireference Approximate VCC Theory

To date no reliable MRCC exists, that has been proven to fulfil all criteria. As a result, MRCI based methods are still the most common form of calculation performed when dealing with multireference situations. In general MRACPF and MRAQCC has been shown to perform well, but they still are not strictly size-extensive and show size-consistency errors. As such, a method that improves upon the MRCI size consistency would be of use.

The previous chapter highlighted recent work on the LPF theories, which use a matrix transformation to introduce a linked denominator for a CEPA(0) like energy functional. For the 2 electron case, these methods are equivalent to CID. Additionally, the matrix inverse can be seen, via the binomial expansion, to introduce a series that approximates the linked form of variational coupled cluster theory to infinite order. Further developments of these methods has given increased robustness within a single reference framework in situations where single reference methods typically struggle or fail, such as at long bond length or when multiple bond breaking [11, 30].

The aim of the work presented in this chapter is to develop an LPF like matrix transformation of the pair amplitudes in MRCI that introduces a linked denominator.

Just as with single reference theory, the transformation is to be designed to introduce terms from VCC. As it would be done in a multireference framework, it would be an

approximation to a multireference variational coupled cluster.

This idea straddles different research areas, which explains the need for descriptions of many different methods in the previous chapter, as they bear some relation to, or their results can be compared to, the work presented here.

The author has not seen a multireference variational coupled cluster theory. Like with single reference, general implementations are based on projective methods. As such, the VCC equations must be analysed, but simplifications that arise when a single reference is used cannot be used here. Thus, the VCC expressions here will be of a general form.

In LPF theory it is the 3rd order VCC terms that are introduced via a transformation. Therefore the approach here will be to also analyse the 3rd order VCC energy and determine which terms are required for the 2 electron, 2 reference model case, and design a transformations to capture those terms. This will produce a method approximating terms in a 3rd order multireference variational Coupled Cluster theory, hence is referred to as MR3VCC.

## 2.1 Spin-Orbital Derivation of MR3VCC

Cluster operator:

$$\hat{T} = 2^{-2} T_{ab}^{ij} a^\dagger b^\dagger j i \quad (2.1)$$

Hamiltonian:

$$\hat{H} = \frac{1}{2} (pr|qs) p^\dagger s^\dagger q r \quad (2.2)$$

### 2.1.1 Isolation of Unlinked Parts of VCCD

#### Third order

As discussed in the previous chapter, it is the linked form of VCC that is infinite. An energy containing only linked diagrams, according to the linked diagram theorem

[7, 8], is extensive. Here, attempts are made to remove the unlinked terms from the 3rd order energy.

Relevant Hamiltonian fragment:

$$\hat{H} = \frac{1}{2} K_{ab}^{ij} a^\dagger b^\dagger j i \quad K_{ab}^{ij} = (a i | b j) \quad (2.3)$$

$$E = \frac{\mathcal{N}}{\mathcal{D}} \quad (2.4)$$

$$\mathcal{N} = 2\langle \hat{H}\hat{T} \rangle + \langle \hat{T}^\dagger \hat{H}\hat{T}^2 \rangle \quad (2.5)$$

$$= 2^{-2} \langle i^\dagger j^\dagger b a g^\dagger h^\dagger p o \rangle K_{ab}^{ij} T_{gh}^{op} + 2^{-7} \langle i^\dagger j^\dagger b a k^\dagger l^\dagger d c e^\dagger f^\dagger n m g^\dagger h^\dagger p o \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ef}^{mn} T_{gh}^{op} \quad (2.6)$$

$$= \frac{1}{2} \langle i^\dagger j^\dagger p o \rangle K_{ab}^{ij} T_{ab}^{op} + 2^{-5} \langle i^\dagger j^\dagger b a k^\dagger l^\dagger d f^\dagger n m g^\dagger h^\dagger p o \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cf}^{mn} T_{gh}^{op} \quad (2.7)$$

$$= \frac{1}{2} \langle i^\dagger j^\dagger p o \rangle K_{ab}^{ij} T_{ab}^{op} + 2^{-5} \langle i^\dagger j^\dagger b a k^\dagger l^\dagger n m g^\dagger h^\dagger p o \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{gh}^{op} + 2^{-4} \langle i^\dagger j^\dagger b a k^\dagger l^\dagger f^\dagger n m g^\dagger p o \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cf}^{mn} T_{gd}^{op} \quad (2.8)$$

$$= \frac{1}{2} \langle i^\dagger j^\dagger p o \rangle K_{ab}^{ij} T_{ab}^{op} + 2^{-4} \langle i^\dagger j^\dagger k^\dagger l^\dagger n m p o \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} \langle i^\dagger j^\dagger k^\dagger l^\dagger n m p o \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \quad (2.9)$$

The unlinked part of this is

$$2\langle \hat{T}^\dagger \hat{T} \rangle \langle \hat{H}\hat{T} \rangle = 2^{-6} \langle i^\dagger j^\dagger b a g^\dagger h^\dagger p o \rangle \langle k^\dagger l^\dagger d c e^\dagger f^\dagger n m \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ef}^{mn} T_{gh}^{op} \quad (2.10)$$

$$= 2^{-4} \langle i^\dagger j^\dagger p o \rangle \langle k^\dagger l^\dagger n m \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \quad (2.11)$$

In either the special case of a system containing only well-separated two-electron subsystems, or the case of a single Slater-determinant reference, this is equivalent to

$$2^{-4} \langle i^\dagger j^\dagger p o k^\dagger l^\dagger n m \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \quad (2.12)$$

We therefore seek to isolate and remove this term from the full numerator.

$$\begin{aligned} \mathcal{N} &= \frac{1}{2} \langle i^\dagger j^\dagger p o \rangle K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-4} \langle i^\dagger j^\dagger k^\dagger l^\dagger p o n m \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} \langle i^\dagger j^\dagger k^\dagger l^\dagger n m p o \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \end{aligned} \quad (2.13)$$

$$= \frac{1}{2} \langle i^\dagger j^\dagger p o \rangle K_{ab}^{ij} T_{ab}^{op}$$

$$\begin{aligned}
& + 2^{-4} \langle i^\dagger j^\dagger p k^\dagger l^\dagger onm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} \langle i^\dagger j^\dagger k^\dagger onm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{ol} \\
& + 2^{-3} \langle i^\dagger j^\dagger k^\dagger l^\dagger nmpo \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \tag{2.14}
\end{aligned}$$

$$\begin{aligned}
& = \frac{1}{2} \langle i^\dagger j^\dagger po \rangle K_{ab}^{ij} T_{ab}^{op} \\
& + 2^{-4} \langle i^\dagger j^\dagger pok^\dagger l^\dagger nm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} \langle i^\dagger j^\dagger p k^\dagger nm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{lp} \\
& + 2^{-3} \langle i^\dagger j^\dagger k^\dagger onm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{ol} \\
& + 2^{-3} \langle i^\dagger j^\dagger k^\dagger l^\dagger nmpo \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \tag{2.15}
\end{aligned}$$

$$\begin{aligned}
& = \frac{1}{2} \langle i^\dagger j^\dagger po \rangle K_{ab}^{ij} T_{ab}^{op} \\
& + 2^{-4} \langle i^\dagger j^\dagger pok^\dagger l^\dagger nm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} \langle i^\dagger j^\dagger k^\dagger onm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{ol} \\
& - 2^{-3} \langle i^\dagger j^\dagger nm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{kl} + 2^{-3} \langle i^\dagger j^\dagger k^\dagger onm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{ol} \\
& + 2^{-3} \langle i^\dagger j^\dagger k^\dagger l^\dagger nmpo \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \tag{2.16}
\end{aligned}$$

The unlinked term does not drop out, but must be contained within:

$$2^{-4} \langle i^\dagger j^\dagger pok^\dagger l^\dagger nm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}$$

The term  $2^{-3} \langle i^\dagger j^\dagger nm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{kl}$ , resembles the diagram -3C from VCC. A transformation could be constructed to capture this term, as it is known to be important in the single reference case. For 2 electrons it is the only term required in the 3rd order energy. However, its importance here is unknown. Terms that are also non-zero for the 2 electron model system will appear out of  $2^{-4} \langle i^\dagger j^\dagger pok^\dagger l^\dagger nm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}$ .

Therefore the unlinked terms must still be removed completely from the energy, to aid in determining the key energy contributions. To try to remove the unlinked term in a more systematic way, the density matrices are expanded in cumulants.

## 2.1.2 Development in Cumulants

The work of Mukherjee and Kutzelnigg, and their general normal ordering [93–96], has led to the widespread use of the cumulant expansion. This expansion allows the reduced density matrices (RDMs) to be expanded in terms of cumulants. The  $n$ -particle cumulant is the irreducible, connected part of the reduced density matrix, which cannot be described by lower order cumulants (or density matrices) or by disconnected products of cumulants. The remainder of the reduced density matrix is described by lower order cumulants and density matrices.

The reduced density matrices ( $\gamma_{1\dots n}^{1\dots n}$ ) can be expanded in cumulants ( $\lambda_{1\dots n}^{1\dots n}$ ) [97][96]

$$\gamma_t^p = \lambda_t^p \quad (2.17)$$

$$\gamma_{tu}^{pq} = \hat{\mathcal{A}} \left[ \left( \frac{1}{2!} \right)^2 \lambda_{tu}^{pq} + \frac{1}{2!} (\lambda_t^p \lambda_u^q) \right] = \lambda_{tu}^{pq} + \lambda_t^p \lambda_u^q - \lambda_u^p \lambda_t^q \quad (2.18)$$

$$\gamma_{tuv}^{pqr} = \hat{\mathcal{A}} \left[ \left( \frac{1}{3!} \right)^2 \lambda_{tuv}^{pqr} + \left( \frac{1}{2!} \right)^2 (\lambda_{tu}^{pq} \lambda_v^r) + \frac{1}{3!} (\lambda_t^p \lambda_u^q \lambda_v^r) \right] \quad (2.19)$$

$$\begin{aligned} \gamma_{tuvw}^{pqrs} = \hat{\mathcal{A}} & \left[ \left( \frac{1}{4!} \right)^2 \lambda_{tuvw}^{pqrs} + \left( \frac{1}{3!} \right)^2 (\lambda_{tuv}^{pqr} \lambda_w^s) + \left( \frac{1}{2!} \right)^5 (\lambda_{tu}^{pq} \lambda_{vw}^{rs}) \right. \\ & \left. + \left( \frac{1}{2!} \right)^3 (\lambda_t^p \lambda_u^q \lambda_v^r \lambda_w^s) + \frac{1}{4!} (\lambda_t^p \lambda_u^q \lambda_v^r \lambda_w^s) \right] \quad (2.20) \end{aligned}$$

The physical meaning of the  $n$ -particle cumulants is not fully understood. For a reference consisting of a single determinant, only the first order cumulant is non-zero, all higher order cumulants are zero. It is therefore logical that the higher order cumulants have been linked with describing electron correlation in multi-determinant theories. Kutzelnigg and Mukherjee [95] stated that the  $n$ -particle cumulant directly describes  $n$ -particle correlations.

Cumulants have been suggested to be the logical choice in density matrix based many body theories because they are extensive quantities, (unlike the reduced density matrices), in that they cannot be described by products of disconnected cumulants and hence truncation of the cumulant series at any point does not affect the extensivity of the method. The truncation of the series is desirable for computational reasons, hence work has been done to look at the affect of the truncation of the series on its accuracy in describing the RDM. Perturbation theory with a single reference shows that the higher order cumulants decrease in importance, meaning truncation of the cumu-

lant expansion of the RDM would have a negligible affect. Thus, since it has been shown that the 1 and 2 particle cumulants, in general, account for the majority of the higher order RDMs, it has become a common practice to neglect the third order cumulant and higher in the cumulant expansion of the RDMs. As mentioned, this still retains extensivity. Cumulant approximations of this nature are central to the Contracted Schrodinger (CSE) theories [98–102] and are used in the Canonical Transformation (CT) theory of Yanai and Chan [62–64], and both state-specific MRCC [103] and internally contracted MRCC [103–105]. Even though neglect of higher order cumulants does not affect extensivity, and whose affect on the accuracy of the RDM description has been deemed small when a single reference determinant is used, the question of the validity of the approximation in the multi-determinantal wave function case is still unanswered.

Recent work [106] has again questioned whether neglecting higher order cumulants has a significant affect by explicitly calculating the higher order cumulants in a few test cases and observing their magnitude. This work shows no decisive conclusion, because whether or not the cumulants decay at higher orders, and are therefore less important, seems to be system specific. Shamasundar[107] justifies the use of cumulant approximations by stating that in general states with multireference character, higher order cumulants will have less importance, unless extended electron delocalisation has occurred, in which case, higher order cumulants would be needed in the RDM decomposition. Based on perturbation theory arguments and the accuracy of methods involving them, approximate cumulant decompositions of RDMs are used here.

Make the approximation that the third and higher-order cumulants (as defined in [97]) are zero. In that case, the density matrices may be re-expressed as products of lower order density matrices,

$$\begin{aligned}
 \gamma_m^i &= \langle i^\dagger m \rangle \\
 \gamma_{mn}^{ij} &= \langle i^\dagger j^\dagger nm \rangle \\
 \gamma_{mno}^{ijk} &= \langle i^\dagger j^\dagger k^\dagger onm \rangle \\
 &= \mathcal{A}_{abc}^{ijk} \left( \frac{1}{4} \gamma_{mn}^{ij} \gamma_o^k - \frac{1}{3} \gamma_m^i \gamma_n^k \gamma_n^l \right) \\
 \gamma_{mnop}^{ijkl} &= \langle i^\dagger j^\dagger k^\dagger l^\dagger ponm \rangle
 \end{aligned} \tag{2.21}$$

$$= \hat{\mathcal{A}}_{mnop}^{ijkl} \left( \frac{1}{32} \gamma_{mn}^{ij} \gamma_{op}^{kl} - \frac{1}{8} \gamma_{mn}^{ij} \gamma_o^k \gamma_p^l + \frac{1}{8} \gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l + \frac{1}{8} \gamma_{mn}^{ij} \gamma_o^k \gamma_p^l + \frac{1}{24} \gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l \right) \quad (2.22)$$

$$= \hat{\mathcal{A}}_{mnop}^{ijkl} \left( \frac{1}{32} \gamma_{mn}^{ij} \gamma_{op}^{kl} + \frac{1}{6} \gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l \right) \quad (2.23)$$

Now starting from (2.9):

$$\begin{aligned} \mathcal{N} &= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} \left( \frac{1}{32} \gamma_{mn}^{ij} \gamma_{op}^{kl} + \frac{1}{6} \gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l \right) \\ &+ 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} \left( \frac{1}{32} \gamma_{mn}^{ij} \gamma_{op}^{kl} + \frac{1}{6} \gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l \right) \\ &= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-9} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (\gamma_{mn}^{ij} \gamma_{op}^{kl}) + \frac{1}{6} 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (\gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l) \\ &+ 2^{-8} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (\gamma_{mn}^{ij} \gamma_{op}^{kl}) + \frac{1}{6} 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (\gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l) \end{aligned} \quad (2.24)$$

$$\begin{aligned} &= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-9} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (4\gamma_{mn}^{ij} \gamma_{op}^{kl} - 16\gamma_{mn}^{il} \gamma_{op}^{kj} + 4\gamma_{mn}^{kl} \gamma_{op}^{ij}) \\ &+ \frac{1}{6} 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (4\gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l - 16\gamma_m^i \gamma_n^l \gamma_o^k \gamma_p^j + 4\gamma_m^k \gamma_n^l \gamma_o^i \gamma_p^j) \\ &+ 2^{-8} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (4\gamma_{mn}^{ij} \gamma_{op}^{kl} - 16\gamma_{mn}^{il} \gamma_{op}^{kj} + 4\gamma_{mn}^{kl} \gamma_{op}^{ij}) \\ &+ \frac{1}{6} 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \hat{\mathcal{A}}_{mnop}^{ijkl} (4\gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l - 16\gamma_m^i \gamma_n^l \gamma_o^k \gamma_p^j + 4\gamma_m^k \gamma_n^l \gamma_o^i \gamma_p^j) \end{aligned} \quad (2.25)$$

$$\begin{aligned} &= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-9} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (16\gamma_{mn}^{ij} \gamma_{op}^{kl} - 64\gamma_{mp}^{ij} \gamma_{on}^{kl} + 16\gamma_{op}^{ij} \gamma_{mn}^{kl} \\ &- 64\gamma_{mn}^{il} \gamma_{op}^{kj} + 256\gamma_{mp}^{il} \gamma_{on}^{kj} - 64\gamma_{op}^{il} \gamma_{mn}^{kj} \\ &+ 16\gamma_{mn}^{kl} \gamma_{op}^{ij} - 64\gamma_{mp}^{kl} \gamma_{on}^{ij} + 16\gamma_{op}^{kl} \gamma_{mn}^{ij}) \\ &+ \frac{1}{6} 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (16\gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l - 64\gamma_m^i \gamma_p^j \gamma_o^k \gamma_n^l + 16\gamma_o^i \gamma_p^j \gamma_m^k \gamma_n^l \\ &- 64\gamma_m^i \gamma_n^l \gamma_o^k \gamma_p^j + 256\gamma_m^i \gamma_p^l \gamma_o^k \gamma_n^j - 64\gamma_o^i \gamma_p^l \gamma_m^k \gamma_n^j \\ &+ 16\gamma_m^k \gamma_n^l \gamma_o^i \gamma_p^j - 64\gamma_m^k \gamma_p^l \gamma_o^i \gamma_n^j + 16\gamma_o^k \gamma_p^l \gamma_m^i \gamma_n^j) \\ &+ 2^{-8} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} (16\gamma_{mn}^{ij} \gamma_{op}^{kl} - 64\gamma_{mp}^{ij} \gamma_{on}^{kl} + 16\gamma_{op}^{ij} \gamma_{mn}^{kl} \\ &- 64\gamma_{mn}^{il} \gamma_{op}^{kj} + 256\gamma_{mp}^{il} \gamma_{on}^{kj} - 64\gamma_{op}^{il} \gamma_{mn}^{kj} \\ &+ 16\gamma_{mn}^{kl} \gamma_{op}^{ij} - 64\gamma_{mp}^{kl} \gamma_{on}^{ij} + 16\gamma_{op}^{kl} \gamma_{mn}^{ij}) \\ &+ \frac{1}{6} 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} (16\gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l - 64\gamma_m^i \gamma_p^j \gamma_o^k \gamma_n^l + 16\gamma_o^i \gamma_p^j \gamma_m^k \gamma_n^l \\ &- 64\gamma_m^i \gamma_n^l \gamma_o^k \gamma_p^j + 256\gamma_m^i \gamma_p^l \gamma_o^k \gamma_n^j - 64\gamma_o^i \gamma_p^l \gamma_m^k \gamma_n^j) \end{aligned}$$

$$\begin{aligned}
& + 16\gamma_m^k \gamma_n^l \gamma_o^i \gamma_p^j - 64\gamma_m^k \gamma_p^l \gamma_o^i \gamma_n^j + 16\gamma_o^k \gamma_p^l \gamma_m^i \gamma_n^j \quad (2.27) \\
& = \frac{1}{2}\gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (\gamma_{op}^{ij} \gamma_{mn}^{kl} \\
& + \gamma_{mn}^{ij} \gamma_{op}^{kl} - 4\gamma_{mp}^{ij} \gamma_{on}^{kl} \\
& - 2\gamma_{mn}^{il} \gamma_{op}^{kj} - 2\gamma_{op}^{il} \gamma_{mn}^{kj} + 8\gamma_{mp}^{il} \gamma_{on}^{kj}) \\
& + \frac{1}{3} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (\gamma_o^i \gamma_p^j \gamma_m^k \gamma_n^l + \gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l - 2\gamma_m^i \gamma_p^j \gamma_o^k \gamma_n^l \\
& - 2\gamma_m^i \gamma_n^l \gamma_o^k \gamma_p^j + 8\gamma_m^i \gamma_p^l \gamma_o^k \gamma_n^j - 2\gamma_o^i \gamma_p^l \gamma_m^k \gamma_n^j - 2\gamma_m^k \gamma_p^l \gamma_o^i \gamma_n^j) \\
& + 2^{-2} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} (\gamma_{mn}^{ij} \gamma_{op}^{kl} - \gamma_{mp}^{ij} \gamma_{on}^{kl} \\
& - \gamma_{mn}^{il} \gamma_{op}^{kj} + 4\gamma_{mp}^{il} \gamma_{on}^{kj} - \gamma_{op}^{il} \gamma_{mn}^{kj} - \gamma_{mp}^{kl} \gamma_{on}^{ij}) \\
& + \frac{1}{3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} (\gamma_m^i \gamma_n^j \gamma_o^k \gamma_p^l - 4\gamma_m^i \gamma_p^j \gamma_o^k \gamma_n^l + \gamma_o^i \gamma_p^j \gamma_m^k \gamma_n^l \\
& - 4\gamma_m^i \gamma_n^k \gamma_o^j \gamma_p^l + 16\gamma_m^i \gamma_p^k \gamma_o^j \gamma_n^l - 4\gamma_o^i \gamma_p^k \gamma_m^j \gamma_n^l \\
& + \gamma_m^k \gamma_n^l \gamma_o^i \gamma_p^j - 4\gamma_m^k \gamma_p^l \gamma_o^i \gamma_n^j + \gamma_o^k \gamma_p^l \gamma_m^i \gamma_n^j) \quad (2.28)
\end{aligned}$$

Unfortunately this isn't the best way to go since some of these terms mix not only terms 3B and 3C of VCC, but also the unlinked diagram. For example, the density matrix product  $\gamma_{mp}^{il} \gamma_{on}^{kj}$  multiplying  $K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}$ . To analyse what this term contains, we look at what it corresponds to in the single reference case, by assuming we only have one reference function and evaluating this term.

$$\begin{aligned}
& K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (\gamma_{mp}^{il} \gamma_{on}^{kj}) \\
& = K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (\delta_m^i \delta_p^l - \delta_p^i \delta_m^l) (\delta_o^k \delta_n^j - \delta_n^k \delta_o^j) \\
& = K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (\delta_m^i \delta_p^l \delta_o^k \delta_n^j + \delta_p^i \delta_m^l \delta_n^k \delta_o^j - \delta_p^i \delta_m^l \delta_o^k \delta_n^j - \delta_m^i \delta_p^l \delta_n^k \delta_o^j) \\
& = K_{ab}^{ij} \left( T_{cd}^{kl} T_{cd}^{ij} T_{ab}^{kl} + T_{cd}^{kl} T_{cd}^{lk} T_{ab}^{ji} - T_{cd}^{kl} T_{cd}^{lj} T_{ab}^{ki} - T_{cd}^{kl} T_{cd}^{ik} T_{ab}^{jl} \right) \\
& = \underbrace{K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{ij} T_{ab}^{kl}}_{\text{Term C}} + \underbrace{K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{lk} T_{ab}^{ji}}_{\text{Unlinked}} - \underbrace{K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{lj} T_{ab}^{ki}}_{\text{Term B}} - \underbrace{K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{ik} T_{ab}^{jl}}_{\text{Term B}}
\end{aligned}$$

Therefore the simple product of density matrices,  $\gamma_{mp}^{il} \gamma_{on}^{kj}$ , when combined with the integrals and amplitudes, contains terms corresponding to the third order terms B and C as well as an unlinked term. The unlinked term cannot be separated out from the product of density matrices, meaning that a transformation involving this product could not be constructed that would generate only linked diagrams, violating the rules for extensivity. If included in a transformation, this density matrix product term would also generate both 3B and 3C terms of 3rd order VCCD, however it has been shown

that generating 3B and 3C separately is advantageous [11].

It is better to delay the repackaging of the 2nd order and lower cumulants into density matrices, until the fully unlinked term appears. Assuming neglect of 3rd order cumulant  $\lambda_{mno}^{ijk}$  and higher, but keeping cumulants explicitly in the working,

$$\gamma_m^i = \langle i^\dagger m \rangle \quad (2.29)$$

$$\lambda_m^i = \gamma_m^i \quad (2.30)$$

$$\begin{aligned} \gamma_{mn}^{ij} &= \langle i^\dagger j^\dagger nm \rangle \\ &= \hat{\mathcal{A}}_{mn}^{ij} \left( \frac{1}{4} \lambda_{mn}^{ij} + \frac{1}{2} \lambda_m^i \lambda_n^j \right) \\ &= \lambda_{mn}^{ij} + \lambda_m^i \lambda_n^j - \lambda_n^i \lambda_m^j \end{aligned} \quad (2.31)$$

$$\lambda_{mn}^{ij} = \gamma_{mn}^{ij} - \gamma_m^i \gamma_n^j + \gamma_n^i \gamma_m^j \quad (2.32)$$

$$\begin{aligned} \gamma_{mno}^{ijk} &= \langle i^\dagger j^\dagger k^\dagger onm \rangle \\ &= \hat{\mathcal{A}}_{mno}^{ijk} \left( \frac{1}{4} \lambda_{mn}^{ij} \lambda_o^k + \frac{1}{2} \lambda_m^i \lambda_n^j \lambda_o^k \right) \end{aligned} \quad (2.33)$$

$$\begin{aligned} \gamma_{mnop}^{ijkl} &= \langle i^\dagger j^\dagger k^\dagger l^\dagger ponm \rangle \\ &= \hat{\mathcal{A}}_{mnop}^{ijkl} \left( \frac{1}{32} \lambda_{mn}^{ij} \lambda_{op}^{kl} + \frac{1}{8} \lambda_{mn}^{ij} \lambda_o^k \lambda_p^l + \frac{1}{24} \lambda_m^i \lambda_n^j \lambda_o^k \lambda_p^l \right) \end{aligned} \quad (2.34)$$

$$\begin{aligned} \mathcal{N} &= \frac{1}{2} \langle i^\dagger j^\dagger po \rangle K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-4} \langle i^\dagger j^\dagger k^\dagger l^\dagger ponm \rangle K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} \langle i^\dagger j^\dagger k^\dagger l^\dagger nmpo \rangle K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \end{aligned} \quad (2.35)$$

$$\begin{aligned} &= \frac{1}{2} \langle i^\dagger j^\dagger po \rangle K_{ab}^{ij} T_{ab}^{op} \\ &+ \left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right) \\ &\quad \hat{\mathcal{A}}_{mnop}^{ijkl} \left( \frac{1}{32} \lambda_{mn}^{ij} \lambda_{op}^{kl} + \frac{1}{8} \lambda_{mn}^{ij} \lambda_o^k \lambda_p^l + \frac{1}{24} \lambda_m^i \lambda_n^j \lambda_o^k \lambda_p^l \right) \end{aligned} \quad (2.36)$$

$$\begin{aligned} &= \frac{1}{2} \langle i^\dagger j^\dagger po \rangle K_{ab}^{ij} T_{ab}^{op} \\ &+ \left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right) \hat{\mathcal{A}}_{mnop} \left( \right. \\ &\quad \frac{1}{4} \lambda_{mn}^{ij} \lambda_{op}^{kl} + \frac{1}{2} \lambda_{mn}^{il} \lambda_{op}^{kj} \\ &\quad + \frac{1}{2} \lambda_{mn}^{ij} \lambda_o^k \lambda_p^l - 2 \lambda_{mn}^{il} \lambda_o^k \lambda_p^j + \frac{1}{2} \lambda_{mn}^{kl} \lambda_o^i \lambda_p^j \\ &\quad \left. + \lambda_m^i \lambda_n^j \lambda_o^k \lambda_p^l \right) \end{aligned} \quad (2.37)$$

$$\begin{aligned} &= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\ &+ \left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right) \left( \right. \end{aligned}$$

$$\begin{aligned}
& \lambda_{mn}^{ij} \lambda_{op}^{kl} - 4\lambda_{mp}^{ij} \lambda_{on}^{kl} + \lambda_{op}^{ij} \lambda_{mn}^{kl} \\
& - 2\lambda_{mn}^{il} \lambda_{op}^{kj} + 8\lambda_{mp}^{il} \lambda_{on}^{kj} - 2\lambda_{op}^{il} \lambda_{mn}^{kj} \\
& + 2\lambda_{mn}^{ij} \lambda_o^k \lambda_p^l - 8\lambda_{mp}^{ij} \lambda_o^k \lambda_n^l + 2\lambda_{op}^{ij} \lambda_m^k \lambda_n^l \\
& - 8\lambda_{mn}^{il} \lambda_o^k \lambda_p^j + 32\lambda_{mp}^{il} \lambda_o^k \lambda_n^j - 8\lambda_{op}^{il} \lambda_m^k \lambda_n^j \\
& + 2\lambda_{mn}^{kl} \lambda_o^i \lambda_p^j - 8\lambda_{mp}^{kl} \lambda_o^i \lambda_n^j + 2\lambda_{op}^{kl} \lambda_m^i \lambda_n^j \\
& + 4\lambda_m^i \lambda_n^j \lambda_o^k \lambda_p^l - 32\lambda_m^i \lambda_p^j \lambda_o^k \lambda_n^l + 4\lambda_o^i \lambda_p^j \lambda_m^k \lambda_n^l
\end{aligned} \tag{2.38}$$

$$\begin{aligned}
& = \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) ( \\
& \lambda_{mn}^{ij} \lambda_{op}^{kl} + 2\lambda_{mn}^{ij} \lambda_o^k \lambda_p^l + 2\lambda_{op}^{kl} \lambda_m^i \lambda_n^j + 4\lambda_m^i \lambda_n^j \lambda_o^k \lambda_p^l \\
& + \lambda_{op}^{ij} \lambda_{mn}^{kl} + 2\lambda_{mn}^{kl} \lambda_o^i \lambda_p^j + 2\lambda_{op}^{ij} \lambda_m^k \lambda_n^l + 4\lambda_o^i \lambda_p^j \lambda_m^k \lambda_n^l \\
& - 4\lambda_{mp}^{ij} \lambda_{on}^{kl} - 8\lambda_{mp}^{ij} \lambda_o^k \lambda_n^l - 32\lambda_m^i \lambda_p^j \lambda_o^k \lambda_n^l \\
& - 2\lambda_{mn}^{il} \lambda_{op}^{kj} - 8\lambda_{mn}^{il} \lambda_o^k \lambda_p^j \\
& + 8\lambda_{mp}^{il} \lambda_{on}^{kj} + 32\lambda_{mp}^{il} \lambda_o^k \lambda_n^j \\
& - 2\lambda_{op}^{il} \lambda_{mn}^{kj} - 8\lambda_{op}^{il} \lambda_m^k \lambda_n^j \\
& - 8\lambda_{mp}^{kl} \lambda_o^i \lambda_n^j
\end{aligned} \tag{2.39}$$

$$\begin{aligned}
& = \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) ( \\
& \gamma_{mn}^{ij} \gamma_{op}^{kl} \\
& + \gamma_{op}^{ij} \gamma_{mn}^{kl} \\
& - 4\gamma_{mp}^{ij} \gamma_{on}^{kl} \\
& + 4\gamma_{mn}^{kj} \gamma_{op}^{li} \\
& + 8\lambda_{mp}^{il} \lambda_{on}^{kj} + 32\lambda_{mp}^{il} \lambda_o^k \lambda_n^j
\end{aligned} \tag{2.40}$$

$$\begin{aligned}
\mathcal{N} & = \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) (\gamma_{op}^{ij} \gamma_{mn}^{kl} \\
& + \gamma_{mn}^{ij} \gamma_{op}^{kl} - 4\gamma_{mp}^{ij} \gamma_{on}^{kl} \\
& + 4(\lambda_{mn}^{kj} + 2\gamma_m^k \gamma_n^j)(\lambda_{op}^{li} + 2\gamma_o^l \gamma_p^i) \\
& + 8\lambda_{mp}^{il} \lambda_{on}^{kj} + 32\lambda_{mp}^{il} \lambda_o^k \lambda_n^j
\end{aligned} \tag{2.41}$$

$$\begin{aligned}
&= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
&+ (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) (\gamma_{op}^{ij} \gamma_{mn}^{kl} \\
&+ \gamma_{mn}^{ij} \gamma_{op}^{kl} - 4 \gamma_{mp}^{ij} \gamma_{on}^{kl} \\
&+ 16 \gamma_m^k \gamma_n^j \gamma_o^l \gamma_p^i + 4 \lambda_{mn}^{kj} \lambda_{op}^{li} + 8 \lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8 \lambda_{op}^{li} \gamma_m^k \gamma_n^j \\
&+ 8 \lambda_{mp}^{il} \lambda_{on}^{kj} + 32 \lambda_{mp}^{il} \gamma_o^k \gamma_n^j) \tag{2.42}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
&+ (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) (\gamma_{op}^{ij} \gamma_{mn}^{kl} \\
&+ \gamma_{mn}^{ij} \gamma_{op}^{kl} - 4 \gamma_{mp}^{ij} \gamma_{on}^{kl} \\
&+ 4(\gamma_{pn}^{ij} - \lambda_{pn}^{ij})(\gamma_{mo}^{kl} - \lambda_{mo}^{kl}) \\
&+ 4 \lambda_{mn}^{kj} \lambda_{op}^{li} + 8 \lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8 \lambda_{op}^{li} \gamma_m^k \gamma_n^j \\
&+ 8 \lambda_{mp}^{il} \lambda_{on}^{kj} + 32 \lambda_{mp}^{il} \gamma_o^k \gamma_n^j) \tag{2.43}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
&+ (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) (\gamma_{op}^{ij} \gamma_{mn}^{kl} \\
&+ \gamma_{mn}^{ij} \gamma_{op}^{kl} - 4 \gamma_{mp}^{ij} \gamma_{on}^{kl} \\
&+ 4 \gamma_{pn}^{ij} \gamma_{mo}^{kl} \\
&- 4 \lambda_{pn}^{ij} \gamma_{mo}^{kl} - 4 \gamma_{pn}^{ij} \lambda_{mo}^{kl} \\
&+ 4 \lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4 \lambda_{mn}^{kj} \lambda_{op}^{li} + 8 \lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8 \lambda_{op}^{li} \gamma_m^k \gamma_n^j \\
&+ 8 \lambda_{mp}^{il} \lambda_{on}^{kj} + 32 \lambda_{mp}^{il} \gamma_o^k \gamma_n^j) \tag{2.44}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
&+ (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) (\gamma_{op}^{ij} \gamma_{mn}^{kl} \\
&+ \gamma_{mn}^{ij} \gamma_{op}^{kl} - 4 \gamma_{mp}^{ij} \gamma_{on}^{kl} \\
&- 4 \gamma_{mp}^{ij} \gamma_{on}^{kl} \\
&- 4 \lambda_{pn}^{ij} \lambda_{mo}^{kl} + 8 \lambda_{pn}^{ij} \gamma_m^k \gamma_o^l - 4 \lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4 \gamma_p^i \gamma_n^j \lambda_{mo}^{kl} - 4 \gamma_n^i \gamma_p^j \lambda_{mo}^{kl} \\
&+ 4 \lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4 \lambda_{mn}^{kj} \lambda_{op}^{li} + 8 \lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8 \lambda_{op}^{li} \gamma_m^k \gamma_n^j \\
&+ 8 \lambda_{mp}^{il} \lambda_{on}^{kj} + 32 \lambda_{mp}^{il} \gamma_o^k \gamma_n^j) \tag{2.45}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
&+ (2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}) (\gamma_{op}^{ij} \gamma_{mn}^{kl} \\
&+ \gamma_{mn}^{ij} \gamma_{op}^{kl} - 8 \gamma_{mp}^{ij} \gamma_{on}^{kl} \\
&+ 8 \lambda_{pn}^{ij} \gamma_m^k \gamma_o^l + 4 \gamma_p^i \gamma_n^j \lambda_{mo}^{kl} - 4 \gamma_n^i \gamma_p^j \lambda_{mo}^{kl} + 8 \lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8 \lambda_{op}^{li} \gamma_m^k \gamma_n^j + 32 \lambda_{mp}^{il} \gamma_o^k \gamma_n^j)
\end{aligned}$$

$$-4\lambda_{pn}^{ij}\lambda_{mo}^{kl} + 4\lambda_{mn}^{kj}\lambda_{op}^{li} + 8\lambda_{mp}^{il}\lambda_{on}^{kj} \quad (2.46)$$

This is an approximate form of the 3rd order VCCD energy. No simplifications have been made by considering only a single reference function, this form assumes a reference of multiple functions but is completely general and hence the form of an approximate VCC method should aim to approximate this form as closely as possible.

### 2.1.3 Approximate MRVCC 3rd Order Energy Terms

The approximate MRVCC 3rd order energy terms are given below.

$$\begin{aligned} E_{3vcc} \approx & \left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right. \\ & \left( \gamma_{op}^{ij} \gamma_{mn}^{kl} + \gamma_{mn}^{ij} \gamma_{op}^{kl} - 8\gamma_{mp}^{ij} \gamma_{on}^{kl} + 8\lambda_{pn}^{ij} \gamma_m^k \gamma_o^l \right. \\ & + 4\gamma_p^i \gamma_n^j \lambda_{mo}^{kl} - 4\gamma_n^i \gamma_p^j \lambda_{mo}^{kl} + 8\lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8\lambda_{op}^{li} \gamma_m^k \gamma_n^j \\ & \left. \left. + 32\lambda_{mp}^{il} \gamma_o^k \gamma_n^j - 4\lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4\lambda_{mn}^{kj} \lambda_{op}^{li} + 8\lambda_{mp}^{il} \lambda_{on}^{kj} \right) \right) \quad (2.47) \end{aligned}$$

The approximate energy terms are referred to according to by the names in the following table, labelling the terms 1 to 12. When multiplying out the energy expression above, each ‘‘term’’, that is, each density matrix product or cumulant product, can be combined with  $2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}$  or  $2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}$ , which give different virtual pairings and will be referred to as pairing  $\alpha$  and pairing  $\beta$ . Term  $1\alpha$  therefore refers to density matrix product term 1, with  $\alpha$  virtual pairing  $2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}$ . The terms and their name, for ease of reference is given in table 2.1.

Many of the terms in this form can be easily seen to be generalisations of the simplified single reference case, which are the terms the single reference LPF methods are designed to capture, such as term 3C and 3B. To aid in determining the major contributors to the energy (i.e. the most important terms to capture in an approximation to the full form), a comparison with the single reference theory is needed.

#### Single Reference Limit

To aid in determining which terms give the most important contributions to the third order energy, correspondences with the single-reference theory must be sought and used as a guide.

| Name                   | Term  |
|------------------------|---|
| 1                      | $\gamma_{op}^{ij} \gamma_{mn}^{kl}$   |
| 2                      | $\gamma_{mn}^{ij} \gamma_{op}^{kl}$   |
| 3                      | $-8\gamma_{mp}^{ij} \gamma_{on}^{kl}$   |
| 4                      | $+8\lambda_{pn}^{ij} \gamma_m^k \gamma_o^l$   |
| 5                      | $+4\gamma_p^i \gamma_n^j \lambda_{mo}^{kl}$   |
| 6                      | $-4\gamma_n^i \gamma_p^j \lambda_{mo}^{kl}$   |
| 7                      | $+8\lambda_{mn}^{kj} \gamma_o^l \gamma_p^i$   |
| 8                      | $+8\lambda_{op}^{li} \gamma_m^k \gamma_n^j$   |
| 9                      | $+32\lambda_{mp}^{il} \gamma_o^k \gamma_n^j$  |
| 10                     | $-4\lambda_{pn}^{ij} \lambda_{mo}^{kl}$   |
| 11                     | $+4\lambda_{mn}^{kj} \lambda_{op}^{li}$   |
| 12                     | $+8\lambda_{mp}^{il} \lambda_{on}^{kj}$   |
| <b>Virtual pairing</b> |   |
| $1\alpha$              | $\gamma_{op}^{ij} \gamma_{mn}^{kl} \times 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}$ |
| $1\beta$               | $\gamma_{op}^{ij} \gamma_{mn}^{kl} \times 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}$ |
| ...                    | ...   |

Table 2.1: Approximate MRVCC 3rd Order Energy Terms

The approximate form of the energy assuming a reference containing multiple reference functions, as found above, is of the form:

$$\begin{aligned}
E = & \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + \left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right) \\
& \left( \gamma_{op}^{ij} \gamma_{mn}^{kl} + \gamma_{mn}^{ij} \gamma_{op}^{kl} - 8\gamma_{mp}^{ij} \gamma_{on}^{kl} + 8\lambda_{pn}^{ij} \gamma_m^k \gamma_o^l \right. \\
& + 4\gamma_p^i \gamma_n^j \lambda_{mo}^{kl} - 4\gamma_n^i \gamma_p^j \lambda_{mo}^{kl} + 8\lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8\lambda_{op}^{li} \gamma_m^k \gamma_n^j \\
& \left. + 32\lambda_{mp}^{il} \gamma_o^k \gamma_n^j - 4\lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4\lambda_{mn}^{kj} \lambda_{op}^{li} + 8\lambda_{mp}^{il} \lambda_{on}^{kj} \right) \quad (2.48)
\end{aligned}$$

The terms involving products of density matrices can be evaluated using a single reference determinant, shown below, and can be compared to single reference 3rd order VCC terms, shown in brackets.

$$\left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \right) \left( \gamma_{op}^{ij} \gamma_{mn}^{kl} \right) = 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{ij} (\text{Unlinked}) \quad (2.49)$$

$$(2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}) (\gamma_{mn}^{ij} \gamma_{op}^{kl}) = 2^{-3} K_{ab}^{ij} T_{cd}^{op} T_{cd}^{ij} T_{ab}^{op} (C) \quad (2.50)$$

$$(2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}) (-8\gamma_{mp}^{ij} \gamma_{on}^{kl}) = -K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{il} T_{ab}^{kj} (B) - K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{jk} T_{ab}^{li} (B) \quad (2.51)$$

|                                       | Virtual pairing   |   |
|---------------------------------------|---|---|
|                                       | $\times 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}$ | $\times 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op}$ |
| $\gamma_{op}^{ij} \gamma_{mn}^{kl}$   | Unlinked  | A   |
| $\gamma_{mn}^{ij} \gamma_{op}^{kl}$   | C   | A   |
| $-8\gamma_{mp}^{ij} \gamma_{on}^{kl}$ | B   | D   |
| All other terms                       | 0   | 0   |

Table 2.2: Single reference limit of the approximate 3rd order MRVCC terms

In the single reference theory, there exist special relationships for the 3rd order terms for the 2-electron case.

$$\begin{aligned} \langle \hat{T}^\dagger \hat{H} \hat{T} \hat{T} \rangle_L &= \langle \hat{T}^\dagger \hat{H} \hat{T} \hat{T} \rangle - \langle \hat{T}^\dagger \hat{T} \rangle \langle \hat{H} \hat{T} \rangle \\ \langle \hat{T}^\dagger \hat{H} \hat{T} \hat{T} \rangle_L + \langle \hat{T}^\dagger \hat{T} \rangle \langle \hat{H} \hat{T} \rangle &= \langle \hat{T}^\dagger \hat{H} \hat{T} \hat{T} \rangle \end{aligned} \quad (2.52)$$

For 2 electrons . . .

$$\begin{aligned} \langle \hat{T}^\dagger \hat{H} \hat{T} \hat{T} \rangle_L + \langle \hat{T}^\dagger \hat{T} \rangle \langle \hat{H} \hat{T} \rangle &= 0 \\ \langle \hat{T}^\dagger \hat{H} \hat{T} \hat{T} \rangle_L &= 3A + 3B + 3C + 3D = 3B + 3C = -3C = \frac{1}{2}3B \end{aligned} \quad (2.53)$$

In summary, the linked and unlinked terms cancel completely to give 0 as the third order energy. Terms in the linked energy cancel each other to leave an object that is the same as  $-3C$ , which is the same as  $\frac{1}{2}3B$  or  $3B + 3C$ . The different LPF methods differed in which of these terms they captured, all being equivalent for the 2 electron special case.

An equivalent analysis is carried out for the approximate 3rd order energy evaluated for the model system, both in the limit of a single reference and in the general multireference case.

## 2.1.4 Two Electron Case

$$|0\rangle = \frac{1}{2} C^{mn} |mn\rangle \equiv \frac{1}{2} C^{mn} m^\dagger n^\dagger |\text{vac}\rangle \quad (2.54)$$

$$\gamma_j^i = \langle 0 | i^\dagger j | 0 \rangle \quad (2.55)$$

$$= C^{il} C^{jl} \quad (2.56)$$

$$\gamma_{kl}^{ij} = \langle 0 | i^\dagger j^\dagger lk | 0 \rangle \quad (2.57)$$

$$= C^{ij} C^{kl} \quad (2.58)$$

$$\lambda_{kl}^{ij} = \gamma_{kl}^{ij} - \gamma_k^i \gamma_l^j + \gamma_l^i \gamma_k^j \quad (2.59)$$

$$= C^{ij} C^{kl} - C^{im} C^{km} C^{jn} C^{ln} + C^{im} C^{lm} C^{jn} C^{kn} \quad (2.60)$$

Assume, without loss of generality, that we have a natural orbital basis, ie that  $\gamma$  is diagonal:

$$\gamma_k^i = (c^{(i)})^2 \delta_{ik} \quad (2.61)$$

$$C^{ij} = c^{(i)} \delta_{j,[i]} \quad (2.62)$$

$$\gamma_{kl}^{ij} = c^{(i)} c^{(k)} \delta_{j,[i]} \delta_{l,[k]} \quad (2.63)$$

Restricting the excitation operators making up  $\hat{T}$  to exclude the null space,

$$T_{ab}^{ij} = c^{(i)} \delta_{j,[i]} T_{ab} \quad (2.64)$$

Normalisation:

$$1 = \sum_{ij} \frac{1}{2} (C^{ij})^2 = \frac{1}{2} \sum_i (c^{(i)})^2 \quad (2.65)$$

In the special case of just two orbitals, then  $c^{(1)} = c^{(2)} = 1$ , otherwise all coefficients are less than 1.

$$\begin{aligned} \mathcal{N}_C &= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (c^{(i)} c^{(o)} \delta_{j,[i]} \delta_{p,[o]} c^{(m)} c^{(k)} \delta_{n,[m]} \delta_{l,[k]} \\ &+ c^{(i)} c^{(m)} \delta_{j,[i]} \delta_{n,[m]} c^{(o)} c^{(k)} \delta_{p,[o]} \delta_{l,[k]} \\ &- 8 c^{(i)} c^{(m)} \delta_{j,[i]} \delta_{p,[m]} c^{(o)} c^{(k)} \delta_{n,[o]} \delta_{l,[k]} \\ &- 8 \lambda_{pn}^{ij} \gamma_m^k \gamma_o^l - 8 \gamma_p^i \gamma_n^j \lambda_{mo}^{kl} + 8 \lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8 \lambda_{op}^{li} \gamma_m^k \gamma_n^j + 32 \lambda_{mp}^{il} \gamma_o^k \gamma_n^j \\ &- 4 \lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4 \lambda_{mn}^{kj} \lambda_{op}^{li} + 8 \lambda_{mp}^{il} \lambda_{on}^{kj}) \\ &= \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\ &+ 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (c^{(i)} c^{(o)} \delta_{j,[i]} \delta_{p,[o]} c^{(m)} c^{(k)} \delta_{n,[m]} \delta_{l,[k]} \end{aligned} \quad (2.66)$$

$$\begin{aligned}
& +c^{(i)} c^{(m)} \delta_{j,[i]} \delta_{n,[m]} c^{(o)} c^{(k)} \delta_{p,[o]} \delta_{l,[k]} \\
& -8c^{(i)} c^{(m)} \delta_{j,[i]} \delta_{p,[m]} c^{(o)} c^{(k)} \delta_{n,[o]} \delta_{l,[k]} \\
& + \dots ) \tag{2.67}
\end{aligned}$$

$$\begin{aligned}
& = \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + 2^{-4} c^{(i)} c^{(k)} c^{(m)} c^{(o)} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} (\delta_{j,[i]} \delta_{p,[o]} \delta_{n,[m]} \delta_{l,[k]} \\
& + \delta_{j,[i]} \delta_{n,[m]} \delta_{p,[o]} \delta_{l,[k]} - 8\delta_{j,[i]} \delta_{p,[m]} \delta_{n,[o]} \delta_{l,[k]} + \dots) \tag{2.68}
\end{aligned}$$

$$\begin{aligned}
& = \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + 2^{-4} \sum_{ijklmnop} c^{(i)} c^{(k)} c^{(m)} c^{(o)} K_{ab}^{ij} c^{(k)} \delta_{l,[k]} T_{cd} c^{(m)} \delta_{n,[m]} T_{cd} c^{(o)} \delta_{p,[o]} T_{ab} \\
& ((1+1)\delta_{j,[i]} \delta_{n,[m]} \delta_{p,[o]} \delta_{l,[k]} \\
& - 8\delta_{j,[i]} \delta_{p,[m]} \delta_{n,[o]} \delta_{l,[k]}) + \dots \tag{2.69}
\end{aligned}$$

$$\begin{aligned}
& = \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + 2^{-4} \sum_{ikmo} K_{ab}^{i,[i]} c^{(i)} (c^{(k)})^2 (c^{(m)})^2 (c^{(o)})^2 ((1+1) - 8\delta_{[m],[o]}) + \dots \tag{2.70}
\end{aligned}$$

$$\begin{aligned}
& = \frac{1}{2} \gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} \\
& + 2^{-4} \sum_i K_{ab}^{i,[i]} c^{(i)} \sum_k (c^{(k)})^2 \left( \sum_{mo} (c^{(m)})^2 (c^{(o)})^2 (1+1) - 8 \sum_m (c^{(m)})^4 \right) + \dots \tag{2.71}
\end{aligned}$$

In the case of just two orbitals, then  $c^{(1)} = c^{(2)} = 1$ . The above is expected to give zero, because, as noted earlier, the linked 3rd order energy terms (B+C) are equal and opposite to the unlinked terms for the 2 electron case. For the multireference case, there are more than 2 orbitals. If the exact answer for 2 electrons is simply a sum of terms 3B and 3C, then the sum above still containing the unlinked term must be zero and would show that the neglected terms (products of cumulants) are not required to be correct for this example system.

Evaluation of the approximate 3rd order energy derived above using the cumulant expansion must be evaluated for the model system.

### Evaluation for the model system

The total 3rd order energy including all linked term and unlinked terms is 0 for 2 electrons, however, the linked form of the energy is non-zero. If the approximate form of the energy derived above is to be relied upon, then it must also show the same

behaviour for the 2-electron, 2-reference model system define in the Appendix.

It can be shown that the evaluation of the approximate 3rd order total energy for the model system does indeed give 0, as desired, and the linked energy is shown to be non-zero.

The linked energy is given below ( $\bar{1}$  indicates  $\beta$  spin and  $1$   $\alpha$  spin),

$$\begin{aligned}
 E_{3vcc} &= E_L + E_{UL} = 0 \\
 E_L &= \left( \begin{aligned}
 &- 8 K_{33}^{1\bar{1}} \cos^4 \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\
 &- 24 K_{33}^{1\bar{1}} \cos^3 \theta \sin \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\
 &- 24 K_{33}^{1\bar{1}} \cos^2 \theta \sin^2 \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\
 &- 8 K_{33}^{1\bar{1}} \cos \theta \sin^3 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\
 &- 8 K_{33}^{2\bar{2}} \sin^4 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} T_{23}^{2\bar{1}} \\
 &- 24 K_{33}^{2\bar{2}} \cos \theta \sin^3 \theta 2^{-3} T_{23}^{2\bar{2}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\
 &- 24 K_{33}^{2\bar{2}} \cos^2 \theta \sin^2 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\
 &- 8 K_{33}^{2\bar{2}} \cos^3 \theta \sin \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \end{aligned} \right) \quad (2.72)
 \end{aligned}$$

The linked energy contains the contributions from 23 terms. However, terms involving second order cumulants sum to zero (terms 4-12). As mentioned, in the single-reference theory for 2 electrons, the linked energy terms partially cancel to give a linked energy that is the same as term  $-3C$  or  $\frac{1}{2}3B$ . Similar cancellations occur here for the evaluation of the multi-reference model system, the linked energy in 2.72 can be seen to reduce to be the same as simple individual contributions. The evaluation of terms  $1\beta$  and  $2\alpha$  are give below,

$$\begin{aligned}
 &\text{Term } 1\beta \gamma_{op}^{ij} \gamma_{mn}^{kl} K_{ab}^{ij} \left( 2^{-4} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right) \quad (2.73) \\
 &= - 8 K_{33}^{1\bar{1}} \cos^4 \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\
 &\quad - 24 K_{33}^{1\bar{1}} \cos^3 \theta \sin \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\
 &\quad - 24 K_{33}^{1\bar{1}} \cos^2 \theta \sin^2 \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\
 &\quad - 8 K_{33}^{1\bar{1}} \cos \theta \sin^3 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\
 &\quad - 8 K_{33}^{2\bar{2}} \sin^4 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} T_{23}^{2\bar{1}} \\
 &\quad - 24 K_{33}^{2\bar{2}} \cos \theta \sin^3 \theta 2^{-3} T_{23}^{2\bar{2}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\
 &\quad - 24 K_{33}^{2\bar{2}} \cos^2 \theta \sin^2 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}}
 \end{aligned}$$

$$- 8 K_{33}^{2\bar{2}} \cos^3 \theta \sin \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \quad (2.74)$$

$$\begin{aligned} & \text{Term } 2\alpha \gamma_{mn}^{ij} \gamma_{op}^{kl} K_{ab}^{ij} \left( 2^{-4} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \right) \quad (2.75) \\ = & + 8 K_{33}^{1\bar{1}} \cos^4 \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\ & + 24 K_{33}^{1\bar{1}} \cos^3 \theta \sin \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\ & + 24 K_{33}^{1\bar{1}} \cos^2 \theta \sin^2 \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\ & + 8 K_{33}^{1\bar{1}} \cos \theta \sin^3 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} \\ & + 8 K_{33}^{2\bar{2}} \sin^4 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{2\bar{2}} T_{23}^{2\bar{1}} \\ & + 24 K_{33}^{2\bar{2}} \cos \theta \sin^3 \theta 2^{-3} T_{23}^{2\bar{2}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\ & + 24 K_{33}^{2\bar{2}} \cos^2 \theta \sin^2 \theta 2^{-3} T_{33}^{2\bar{2}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \\ & + 8 K_{33}^{2\bar{2}} \cos^3 \theta \sin \theta 2^{-3} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} T_{33}^{1\bar{1}} \quad (2.76) \end{aligned}$$

The results of the evaluation of these terms can be related to the linked energy (eq. 2.72) very simply. For the 2 electron, 2 reference model, the linked approximate 3rd order VCC energy can be written,

$$E_L = 1\beta = -2\alpha \quad (2.77)$$

It is shown above that in the single reference limit term  $2\alpha$  reduces to VCC term 3C, so it is in complete analogy that the linked energy for the model can be described completely by  $-2\alpha$ , just as in single reference the linked energy for 2 electrons can be written as  $-3C$ . The other term here that gives the full linked energy is term  $1\beta$ , which, in the single reference limit, reduces to diagram A. This is a slightly different trend to the single reference case.

The approximate 3rd order energy contains some terms involving 2nd order cumulants (terms 4-12), that sum to 0 for the model. They may be important for the general multireference case, but their importance is not probed here. One can look at their importance by re-analysing the 4th order density matrix and showing how good the approximations to it are when containing the different terms.

For the model case, term  $1\beta$  or  $-2\alpha$  are the only terms needed for the description of the linked energy. Accordingly attempts should be made at capturing these terms via a transformation of the pair amplitudes.

## The Transformation

As stated, it is apt to capture either term  $1\beta$  or the negative of term  $2\alpha$  via a transformation, as either of these terms give the correct linked energy for the model 2-electron multireference system. A pragmatic choice of which term to capture is to choose term  $-2\alpha$ , as the pairing of virtual indices is simpler than that of term  $1\beta$ . Also, another consideration is the comparison to single reference theory. Term  $-2\alpha$  is the multireference generalisation of the single reference 3rd order VCC diagram -3C. -3C is the term captured by LPF+1D, which has shown to perform well (despite being a simple approximation to VCC)[11, 30], and this term is known to be important in the general case. Without further analysis of the multireference terms, capture of  $-2\alpha$  is a logical choice.

To restate, term  $-2\alpha$  is chosen to be captured by a transformation because it is exact for the model system and reduces directly to LPF+1 when a single reference function is used. Possible transformations to capture further terms are outlined later but not considered in detail in this work.

The requirement is the transformed amplitude must capture the correct 1st order energy but also the negative 3rd order energy term  $2\alpha$ .

$$\mathcal{N} = \frac{1}{2}\gamma_{op}^{ij} K_{ab}^{ij} T_{ab}^{op} - 2^{-4}K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} \gamma_{mn}^{ij} \gamma_{op}^{kl} \quad (2.78)$$

$$= \frac{1}{2}K_{ab}^{ij} \gamma_{mn}^{ij} 2T_{ab}^{mn} \quad (2.79)$$

An appropriate transformation matrix,  $\mathbf{U}$ , can be defined,

$$\mathbf{U} = \mathbf{1} + \mathbf{\Delta} \quad (2.80)$$

$$U_{op}^{mn} = \delta_{op}^{mn} + \Delta_{op}^{mn} \quad (2.81)$$

$$\Delta_{op}^{mn} = \frac{1}{2}\eta_{kl}^{mn} \gamma_{op}^{kl} \quad (2.82)$$

$$\eta_{kl}^{mn} = \frac{1}{2}T_{cd}^{mn} T_{kl}^{cd} \gamma_{op}^{kl} \quad (2.83)$$

The definition of a general transformed amplitude follows;

$${}_q T_{ab}^{mn} = \frac{1}{2} \left( \delta_{op}^{mn} + \frac{1}{4} T_{cd}^{kl} T_{cd}^{mn} \gamma_{op}^{kl} \right)^{-\frac{q}{2}} T_{ab}^{op} \quad (2.84)$$

$${}_q T_{ab}^{mn} = \frac{1}{2} \left( \mathbf{U}^{-\frac{q}{2}} \right)_{op}^{mn} T_{ab}^{op} \quad (2.85)$$

The result of the transformation can be viewed using the binomial expansion  $(1 + x)^{-1} = 1 - x + \dots$

$$\frac{1}{2}K_{ab}^{ij}\gamma_{mn}^{ij}(U^{-1}T_{ab}^{op}) = \frac{1}{2}K_{ab}^{ij}\gamma_{mn}^{ij}T_{ab}^{mn} - 2^{-4}K_{ab}^{ij}\gamma_{mn}^{ij}T_{cd}^{mn}T_{cd}^{kl}\gamma_{op}^{kl}T_{ab}^{op} + \dots \quad (2.86)$$

This is the 1st order energy contribution plus the negative of term  $2\alpha$  as required, showing the transformation matrix does as required. This transformation can be constructed at a cost in complexity of no more than  $\sigma^4 v^2$ , by first computing the intermediate matrix  $\eta$  before contraction with the density matrix to make  $\Delta$  to be used in the transformation matrix,  $\mathbf{U}$ .

Computational scaling:

$$T_{cd}^{mn}T_{kl}^{cd}\gamma_{op}^{kl} \xrightarrow{\sigma^4 v^2} \eta_{kl}^{mn}\gamma_{op}^{kl} \xrightarrow{\sigma^6} \Delta_{op}^{mn}$$

$$U_{op}^{mn} = \delta_{op}^{mn} + \Delta_{op}^{mn}$$

The formation of this transformation matrix is no more complex than the limiting steps in both Coupled Cluster and Configuration Interaction, however, the number of internal orbitals is greatly increased in the multireference case, making the  $\mathbf{U}$  matrix more demanding to construct, and invert.

### 2.1.5 Positivity of the Transformation Matrix

As negative powers of the  $\mathbf{U}$  matrix are needed in the transformation, the  $\mathbf{U}$  matrix must proven to be positive definite, to be sure the matrix inverse exists and can be found. Without this property the theory cannot work.

In order to look at the positive definite-ness of  $\mathbf{U}$ , the component matrices must be analysed.

$$\mathbf{U} = \delta + \Delta \quad (2.87)$$

$$U_{op}^{mn} = \delta_{op}^{mn} + \frac{1}{4}T_{cd}^{mn}T_{kl}^{cd}\langle 0|k^\dagger l^\dagger p o|0\rangle \quad (2.88)$$

$$U_{op}^{mn} = \delta_{op}^{mn} + \frac{1}{4}T_{cd}^{mn}T_{kl}^{cd}\langle \phi_{-2}^{kl}|\phi_{op}^{-2}\rangle \quad (2.89)$$

Where  $\langle \phi_{-2}^{kl}|$  denotes an  $n - 2$  electron state, and  $kl$  indicates the electrons annihilated. The elements of the  $\Delta$  matrix are formed as inner products of vectors, showing  $\Delta$  is a

Grammian matrix and is therefore positive-semi-definite, meaning the lowest possible eigenvalue is 0 [108]. The addition of the unit matrix will shift the eigenvalues  $\mathbf{U}$  when compared to  $\Delta$  by 1. This results in the lowest possible eigenvalue now being positive and hence  $\mathbf{U}$  is positive definite, meaning it is non-singular and negative matrix powers can be found.

### Model Multireference 2-Electron Case

The multi-reference LPF theory must be formulated to be exact for 2 electrons. Here is the model two electron case, with 2 reference determinants (defined in Appendix). Consider a 2-electron system with 2 reference determinants. There are 2 occupied orbitals with different symmetry and one virtual orbital.

$$|1^2\rangle = 1\bar{1}$$

$$|2^2\rangle = 2\bar{2}$$

The reference wavefunction:

$$|0\rangle = \cos\theta|1^2\rangle + \sin\theta|2^2\rangle$$

Due to there being only 1 virtual orbital, there is only 1 excited configuration:

$$|3^2\rangle = 3\bar{3}$$

The total wavefunction,

$$|\Psi\rangle = \cos\theta|1^2\rangle + \sin\theta|2^2\rangle + \mu|3^2\rangle$$

$$|\Psi\rangle = |0\rangle + \mu|3^2\rangle$$

The excitation operators are also orthogonalised and are shown together with its action on the reference wavefunction.

$$\hat{T} = \left( \cos\theta\hat{E}_{31,31} + \sin\theta\hat{E}_{32,32} \right) \mu$$

$$\hat{T}|0\rangle = (\cos^2\theta + \sin^2\theta)\mu|3^2\rangle$$

$$T_{11}^{33} = \cos\theta\mu$$

$$T_{22}^{33} = \sin \theta \mu$$

The  $\gamma$  matrix needed for this method and its elements are defined using  $\gamma_{op}^{kl} = \langle k^\dagger l^\dagger | p o \rangle$

$$\gamma = \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix}$$

The  $\eta$  matrix must also be found.

$$\eta = \begin{pmatrix} \cos^2 \theta \mu^2 & \sin \theta \cos \theta \mu^2 \\ \sin \theta \cos \theta \mu^2 & \sin^2 \theta \mu^2 \end{pmatrix}$$

Evaluating all terms that contribute to the  $\Delta$  matrix,

$$\Delta = \eta \gamma$$

$$\Delta = \begin{pmatrix} \cos^2 \theta \mu^2 & \sin \theta \cos \theta \mu^2 \\ \sin \theta \cos \theta \mu^2 & \sin^2 \theta \mu^2 \end{pmatrix}$$

Forming the  $\mathbf{U}$  matrix,

$$U_{kl}^{ij} = \delta_{kl}^{ij} + \eta_{kl}^{ij} \quad (2.90)$$

$$\mathbf{U} = \begin{pmatrix} 1 + \cos^2 \theta \mu^2 & \sin \theta \cos \theta \mu^2 \\ \sin \theta \cos \theta \mu^2 & 1 + \sin^2 \theta \mu^2 \end{pmatrix}$$

This matrix needs inverting, the inverse matrix can be visualised via Kramer's closed form method for finding the inverse of a matrix.

$$\mathbf{U}^{-1} = \begin{pmatrix} \frac{1 + \sin^2 \theta \mu^2}{1 + \mu^2} & -\frac{\sin \theta \cos \theta \mu^2}{1 + \mu^2} \\ -\frac{\sin \theta \cos \theta \mu^2}{1 + \mu^2} & \frac{1 + \cos^2 \theta \mu^2}{1 + \mu^2} \end{pmatrix}$$

The transformation then follows,

$$\mathbf{U}^{-1} \mathbf{T} = \tilde{\mathbf{T}}$$

$$\begin{pmatrix} \frac{1 + \sin^2 \theta \mu^2}{1 + \mu^2} & -\frac{\sin \theta \cos \theta \mu^2}{1 + \mu^2} \\ -\frac{\sin \theta \cos \theta \mu^2}{1 + \mu^2} & \frac{1 + \cos^2 \theta \mu^2}{1 + \mu^2} \end{pmatrix} \begin{pmatrix} T_{11}^{33} \\ T_{22}^{33} \end{pmatrix} = \begin{pmatrix} \tilde{T}_{11}^{33} \\ \tilde{T}_{22}^{33} \end{pmatrix}$$

$$\begin{aligned}\frac{(1 + \sin^2 \theta \mu^2) T_{11}^{33} - (\sin \theta \cos \theta \mu^2) T_{22}^{33}}{1 + \mu^2} &= \tilde{T}_{11}^{33} \\ \frac{(-\sin \theta \cos \theta \mu^2) T_{11}^{33} + (1 + \cos^2 \theta \mu^2) T_{22}^{33}}{1 + \mu^2} &= \tilde{T}_{22}^{33}\end{aligned}$$

Remembering the definition of the amplitudes (above),

$$\begin{aligned}\tilde{T}_{11}^{33} &= \frac{\cos \theta \mu + \sin^2 \theta \mu^2 \cos \theta \mu - \sin \theta \cos \theta \mu^2 \sin \theta \mu}{1 + \mu^2} \\ \tilde{T}_{22}^{33} &= \frac{-\sin \theta \cos \theta \mu^2 \cos \theta \mu + \sin \theta \mu + \cos^2 \theta \mu^2 \sin \theta \mu}{1 + \mu^2} \\ \tilde{T}_{11}^{33} &= \frac{T_{11}^{33}}{1 + \mu^2} \\ \tilde{T}_{22}^{33} &= \frac{T_{22}^{33}}{1 + \mu^2}\end{aligned}$$

This is simply the standard amplitude with the introduction of the MRCI pair denominator,  $1 + \langle \hat{T}^\dagger \hat{T} \rangle$ , making it the same as MRCI. This shows the equivalence of a term involving the transformed amplitude to that of the term in MRCI in the 2-electron, 2-reference model system.

Alternatively, the inverse of the U transformation can be visualised via the binomial expansion of  $(1 + \Delta)^{-1}$ . Thus the transformation is as follows,

$$U^{-1}T = \tilde{T}$$

$$\left( \left( \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \cos^2 \theta \mu^2 & \sin \theta \cos \theta \mu^2 \\ \sin \theta \cos \theta \mu^2 & \sin^2 \theta \mu^2 \end{pmatrix} + \dots \right) \begin{pmatrix} T_{11}^{33} \\ T_{22}^{33} \end{pmatrix} \right) = \begin{pmatrix} \tilde{T}_{11}^{33} \\ \tilde{T}_{22}^{33} \end{pmatrix}$$

$$\begin{aligned}\cos \theta \mu - \cos^2 \theta \mu^2 \cos \theta \mu - \sin \theta \cos \theta \mu^2 \sin \theta \mu + \dots &= \tilde{T}_{11}^{33} \\ \sin \theta \mu - \sin \theta \cos \theta \mu^2 \cos \theta \mu - \sin^2 \theta \mu^2 \sin \theta \mu + \dots &= \tilde{T}_{22}^{33}\end{aligned}$$

$$\begin{aligned}T_{11}^{33} - T_{11}^{33} T_{11}^{33} T_{11}^{33} - T_{11}^{33} T_{22}^{33} T_{22}^{33} &= \tilde{T}_{11}^{33} \\ T_{22}^{33} - T_{22}^{33} T_{11}^{33} T_{11}^{33} - T_{22}^{33} T_{22}^{33} T_{22}^{33} &= \tilde{T}_{22}^{33}\end{aligned}$$

Showing the first order terms still being present, while capturing the terms corresponding to the -C 3rd order diagram, which ensures the desired behaviour for 2-electrons.

## Redundancy Removal

As a result of the multireference nature of the wavefunction, the excitation operators form a linearly dependent set. Thus, these redundancies must be removed. Due to the form of the equations, the method of Evangelista & Gauss [53] cannot be used, as it results in orthogonalisation of the operators, causing  $\mathbf{U}$  to be in a different basis to the rest of the work.

Therefore, the method used in the MOLPRO icMRCI [58] code is used, that of pair deletion.

$$\gamma = (\cos^2 \theta) \quad (2.91)$$

$$\mathbf{T} = T_{33}^{11,1} \quad (2.92)$$

Therefore the  $\mathbf{U}$  matrix can be easily found and inverted.

$$\begin{aligned} \mathbf{U} &= \mathbf{1} + \Delta \\ \mathbf{U} &= \mathbf{1} + (\gamma_{11}^{11} T_{33}^{111} T_{111}^{33}) \\ \mathbf{U} &= \mathbf{1} + \cos^2 \theta \mu^2 \cos^2 \theta \\ \mathbf{U}^{-1} &= \frac{1}{1 + \cos^2 \theta \mu^2 \cos^2 \theta} \end{aligned} \quad (2.93)$$

The  $\mathbf{U}$  matrix introduces the pair norm correctly:

$$\langle \hat{T}^\dagger \hat{T} \rangle = \langle \psi_{11}^{33} | \psi_{11}^{33} \rangle (T_{111}^{33})^2 \quad (2.94)$$

$$= \cos^2 \theta \mu^2 \cos^2 \theta \quad (2.95)$$

This shows using the method of pair deletion, the correct pair norm is still introduced, but the excitations to the external space are coming from 1 determinant only because of the redundancy.

## Further Manipulation

This section attempts to further manipulate the terms 4-12 in table 2.1, which are not needed for the model multireference case but may be important in the general case.

Looking at the (approximate) 3rd order energy contributions:

$$\begin{aligned}
E_{3vcc} = & \left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right) \\
& \left( \gamma_{op}^{ij} \gamma_{mn}^{kl} + \gamma_{mn}^{ij} \gamma_{op}^{kl} - 8\gamma_{mp}^{ij} \gamma_{on}^{kl} + 8\lambda_{pn}^{ij} \gamma_m^k \gamma_o^l \right. \\
& + 4\gamma_p^i \gamma_n^j \lambda_{mo}^{kl} - 4\gamma_n^i \gamma_p^j \lambda_{mo}^{kl} + 8\lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8\lambda_{op}^{li} \gamma_m^k \gamma_n^j \\
& \left. + 32\lambda_{mp}^{il} \gamma_o^k \gamma_n^j - 4\lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4\lambda_{mn}^{kj} \lambda_{op}^{li} + 8\lambda_{mp}^{il} \lambda_{on}^{kj} \right) \quad (2.96)
\end{aligned}$$

The terms involving 2nd order cumulants disappear in the single reference case. Although these terms were shown not to be important for the multireference model system, these terms may be important in the general multireference case and should be simplified into a computable form. The terms in question have been manipulated and rearranged below.

$$\begin{aligned}
& + 8\lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8\lambda_{op}^{li} \gamma_m^k \gamma_n^j + 4\lambda_{mn}^{kj} \lambda_{op}^{li} \\
& + 8\lambda_{mp}^{il} \lambda_{on}^{kj} + 32\lambda_{mp}^{il} \gamma_o^k \gamma_n^j \\
& - 8\lambda_{pn}^{ij} \gamma_m^k \gamma_o^l - 8\gamma_p^i \gamma_n^j \lambda_{mo}^{kl} - 4\lambda_{pn}^{ij} \lambda_{mo}^{kl}
\end{aligned}$$

Upon inspection the first line looks like the product  $\gamma_{mn}^{kj} \gamma_{op}^{li}$  and the 3rd line like  $-\gamma_{pn}^{ij} \gamma_{mo}^{kl}$ . This observation can be probed by expansion of the density matrix products and comparing to the terms in the energy expression.

$$\begin{aligned}
4\gamma_{mn}^{kj} \gamma_{op}^{li} & = 4 (\lambda_{mn}^{kj} + 2\gamma_m^k \gamma_n^j) (\lambda_{op}^{li} + 2\gamma_o^l \gamma_p^i) \\
& = 4 (\lambda_{mn}^{kj} \lambda_{op}^{li} + 4\gamma_m^k \gamma_n^j \gamma_o^l \gamma_p^i + 2\gamma_m^k \gamma_n^j \lambda_{op}^{li} + 2\gamma_o^l \gamma_p^i \lambda_{mn}^{kj}) \quad (2.97)
\end{aligned}$$

$$\begin{aligned}
-4\gamma_{pn}^{ij} \gamma_{mo}^{kl} & = 4 (\lambda_{pn}^{ij} + 2\gamma_p^i \gamma_n^j) (\lambda_{mo}^{kl} + 2\gamma_m^k \gamma_o^l) \\
& = 4 (\lambda_{pn}^{ij} \lambda_{mo}^{kl} + 4\gamma_p^i \gamma_n^j \gamma_m^k \gamma_o^l + 2\gamma_p^i \gamma_n^j \lambda_{mo}^{kl} + 2\gamma_o^l \gamma_m^k \lambda_{pn}^{ij}) \quad (2.98)
\end{aligned}$$

Thus

$$\begin{aligned}
4\gamma_{mn}^{kj} \gamma_{op}^{li} - 4\gamma_{pn}^{ij} \gamma_{mo}^{kl} & = +8\lambda_{mn}^{kj} \gamma_o^l \gamma_p^i + 8\lambda_{op}^{li} \gamma_m^k \gamma_n^j + 4\lambda_{mn}^{kj} \lambda_{op}^{li} \\
& - 8\lambda_{pn}^{ij} \gamma_m^k \gamma_o^l - 8\gamma_p^i \gamma_n^j \lambda_{mo}^{kl} - 4\lambda_{pn}^{ij} \lambda_{mo}^{kl} \quad (2.99)
\end{aligned}$$

which is the same set of terms that appear in the energy expression, Realising this relation, the terms in the energy expression involving 2nd order cumulants can now be simplified,

$$+ 4\gamma_{mn}^{kj} \gamma_{op}^{li} - 4\gamma_{pn}^{ij} \gamma_{mo}^{kl} + 8\lambda_{mp}^{il} \lambda_{on}^{kj} + 32\lambda_{mp}^{il} \gamma_o^k \gamma_n^j$$

Despite re-expressing the terms involving 2nd order cumulants as products of density matrices, it can be seen that they still sum to zero in the single reference limit and for the model MR case.

The energy expression now reads,

$$\begin{aligned}
 E_{3vcc} \approx & \left( 2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op} + 2^{-3} K_{ab}^{ij} T_{cd}^{kl} T_{ca}^{mn} T_{bd}^{op} \right) \\
 & \left( \gamma_{op}^{ij} \gamma_{mn}^{kl} + \gamma_{mn}^{ij} \gamma_{op}^{kl} - 8 \gamma_{mp}^{ij} \gamma_{on}^{kl} \right. \\
 & \left. + 4 \gamma_{mn}^{kj} \gamma_{op}^{li} - 4 \gamma_{pn}^{ij} \gamma_{mo}^{kl} + 8 \lambda_{mp}^{il} \lambda_{on}^{kj} + 32 \lambda_{mp}^{il} \gamma_o^k \gamma_n^j \right) \quad (2.100)
 \end{aligned}$$

### Further Transformations

Although not utilised in the current theory, these matrices are likely to be required in future implementations in order to generate more 3rd order energy terms to approximate VCC more closely and therefore produce a more complete theory.

The selection of which 3rd order terms to include in the theory and therefore which terms require production via a matrix transformation must be decided based on theoretical arguments based on approximation of the 4th order density matrix, and the numerical performance of pilot implementations and tests. It has already been indicated above that no extra terms are needed to be exact for the 2 electron, 2 reference model system and no further argument on which terms should be captured in a transformation is given here, but transformation matrices that would capture these terms are given, for completeness.

| Term  | Transformation Matrix  | Complexity          |
|---|--|---------------------|
| $(2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}) \gamma_{op}^{ij} \gamma_{mn}^{kl}$    | Unwanted   | -                   |
| $(2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}) \gamma_{mn}^{ij} \gamma_{op}^{kl}$    | $\Delta_{op}^{mn} = T_{cd}^{mn} T_{cd}^{kl} \gamma_{op}^{kl}$<br>$U_{op}^{mn} = \delta_{op}^{mn} + \Delta_{op}^{mn}$     | $o^4 v^2$ and $o^6$ |
| $-(2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}) 8 \gamma_{mp}^{ij} \gamma_{on}^{kl}$ | $\Delta_o^m = T_{cd}^{mn} T_{cd}^{kl} \gamma_{on}^{kl}$<br>$U_o^m = \delta_o^m + \Delta_o^m$                             | $o^4 v^2$ and $o^5$ |
| $(2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}) 4 \gamma_{mn}^{kj} \gamma_{op}^{li}$  | $\Delta^{mn,ki,op} = T_{cd}^{mn} T_{cd}^{kl} \gamma_{op}^{li}$<br>$U^{mn,ki,op} = \delta^{mn,ki,op} + \Delta^{mn,ki,op}$ | $o^4 v^2$ and $o^7$ |
| $-(2^{-4} K_{ab}^{ij} T_{cd}^{kl} T_{cd}^{mn} T_{ab}^{op}) 4 \gamma_{pn}^{ij} \gamma_{mo}^{kl}$ | $\Delta_o^n = T_{cd}^{mn} T_{cd}^{kl} \gamma_{mo}^{kl}$<br>$U_o^n = \delta_o^n + \Delta_o^n$                             | $o^4 v^2$ and $o^5$ |

Table 2.3: Transformations to capture all terms that are products of density matrices in the approximate 3rd order energy (2.100)

## 2.2 Spin Free Formulation

The partial numerator only containing the term  $2\alpha$  (corresponding to single-reference diagram 3C) has until now been expressed only in spin-orbital form.

$$\mathcal{N} = \sum_{ijklmnopabcd} \frac{1}{16} K_{ab}^{ij} T_{ab}^{op} \gamma_{mn}^{ij} T_{cd}^{mn} T_{cd}^{kl} \gamma_{op}^{kl} \quad \text{spin orbitals}$$

However, a spin-summed, purely spatial form is needed to be comparable to, and to be integrated into, the MRCI code within MOLPRO.

In spatial form, the numerator:

$$\mathcal{N} = \sum_{ijklmnopabcd} \frac{1}{16} K_{ab}^{ij} T_{ab}^{op} \gamma_{mn}^{ij} T_{cd}^{mn} T_{cd}^{kl} \gamma_{op}^{kl} \quad \text{spin orbitals}$$

$$\mathcal{N} = \sum_{ijklmnopabcd} \left( \frac{1}{16} K_{\bar{a}\bar{b}}^{\bar{i}\bar{j}} T_{\bar{a}\bar{b}}^{\bar{o}\bar{p}} \gamma_{\bar{m}\bar{n}}^{\bar{i}\bar{j}} T_{\bar{c}\bar{d}}^{\bar{m}\bar{n}} T_{\bar{c}\bar{d}}^{\bar{k}\bar{l}} \gamma_{\bar{o}\bar{p}}^{\bar{k}\bar{l}} \right. \\ \left. + 4 K_{\bar{a}\bar{b}}^{\bar{i}\bar{j}} T_{\bar{a}\bar{b}}^{\bar{o}\bar{p}} \gamma_{\bar{m}\bar{n}}^{\bar{i}\bar{j}} T_{\bar{c}\bar{d}}^{\bar{m}\bar{n}} T_{\bar{c}\bar{d}}^{\bar{k}\bar{l}} \gamma_{\bar{o}\bar{p}}^{\bar{k}\bar{l}} \right)$$

$$+ \frac{1}{16} K_{\bar{a}\bar{b}}^{\bar{i}\bar{j}} T_{\bar{a}\bar{b}}^{\bar{o}\bar{p}} \gamma_{\bar{m}\bar{n}}^{\bar{i}\bar{j}} T_{\bar{c}\bar{d}}^{\bar{r}\bar{s}} T_{\bar{c}\bar{d}}^{\bar{k}\bar{l}} \gamma_{\bar{o}\bar{p}}^{\bar{k}\bar{l}} \Big) \quad \text{spatial orbitals}$$

where  $\tilde{i}$  indicates an electron with  $\alpha$  spin and  $\bar{i}$  an electron with  $\beta$  spin. For non-singlet states, spin-rotation of the spin factor associated with the spin-orbital causes a change in the state. Accordingly, invariant states and objects are instead used in spin-adapted many electron methods in MOLPRO. Therefore, the term to be captured via a transformation (term  $2\alpha$ ) needs to be re-expressed using spin-free operators, which are invariant to spin rotations and can be written as invariant linear combinations of spin-orbital operators. The spin-free (singlet) excitation operator can be written as a linear combination of spin-orbital operators,

$$E_{pq} = \tilde{p}^\dagger \tilde{q} + \bar{p}^\dagger \bar{q} \quad (2.101)$$

Re-expressing using spin summed pure spatial operators,

$$\hat{T} = \sum_{ij} \sum_{ab} T_{ab}^{ij} \hat{E}_{ai} \hat{E}_{bj} \quad (2.102)$$

$$= \sum_{ij} \sum_{ab} T_{ab}^{ij} \left( \tilde{a}^\dagger \tilde{b}^\dagger \tilde{i} \tilde{j} + \bar{a}^\dagger \bar{b}^\dagger \bar{i} \bar{j} + a^\dagger \bar{b}^\dagger \bar{j} i + a^\dagger \tilde{b}^\dagger \tilde{i} j \right) \quad (2.103)$$

$$= \sum_{ij} \sum_{ab} \frac{1}{4} T_{\bar{a}\bar{b}}^{\bar{i}\bar{j}} \bar{a}^\dagger \bar{b}^\dagger \bar{i} \bar{j} + \frac{1}{4} T_{\bar{a}\bar{b}}^{\bar{i}\bar{j}} \bar{a}^\dagger \bar{b}^\dagger \bar{i} \bar{j} + T_{\bar{a}\bar{b}}^{\bar{i}\bar{j}} \tilde{a}^\dagger \tilde{b}^\dagger \tilde{i} \tilde{j} \quad (2.104)$$

$$\Gamma_{mn}^{ij} = \langle (\tilde{i}^\dagger \tilde{m} + \bar{i}^\dagger \bar{m}) (\tilde{j}^\dagger \tilde{n} + \bar{j}^\dagger \bar{n}) - \delta_m^j (\tilde{i}^\dagger \tilde{n} + \bar{i}^\dagger \bar{n}) \rangle \quad (2.105)$$

$$= \langle \tilde{i}^\dagger \tilde{j}^\dagger \tilde{n} \tilde{m} + \bar{i}^\dagger \bar{j}^\dagger \bar{n} \bar{m} + \tilde{i}^\dagger \bar{j}^\dagger \bar{n} \tilde{m} + \bar{i}^\dagger \tilde{j}^\dagger \tilde{n} \bar{m} - \delta_m^j (\tilde{i}^\dagger \tilde{n} + \bar{i}^\dagger \bar{n}) \rangle \quad (2.106)$$

$$= \gamma_{\tilde{m}\tilde{n}}^{\tilde{i}\tilde{j}} + \gamma_{\bar{m}\bar{n}}^{\bar{i}\bar{j}} + \gamma_{\tilde{m}\bar{n}}^{\tilde{i}\bar{j}} + \gamma_{\bar{m}\tilde{n}}^{\bar{i}\tilde{j}} \quad (2.107)$$

It becomes immediately clear that re-expression using these operators involves more than a simple ‘translation’ between the two representations. The spatial expression, above (2.107), involves products of density matrices, but does not contain products of density matrices of mixed spin, as a product of spin-summed density matrices would.

The product of 2 spin-summed 2nd-order reduced density matrices is shown.

$$\begin{aligned} \Gamma_{mn}^{ij} \Gamma_{op}^{kl} = & \underbrace{\gamma_{\tilde{m}\tilde{n}}^{\tilde{i}\tilde{j}} \gamma_{\tilde{o}\tilde{p}}^{\tilde{k}\tilde{l}}}_{\text{spin-}\alpha} + \underbrace{\gamma_{\bar{m}\bar{n}}^{\bar{i}\bar{j}} \gamma_{\bar{o}\bar{p}}^{\bar{k}\bar{l}}}_{\text{spin-}\beta} + \underbrace{\gamma_{\tilde{m}\bar{n}}^{\tilde{i}\bar{j}} \gamma_{\tilde{o}\tilde{p}}^{\tilde{k}\tilde{l}}}_{\text{mixed}} + \underbrace{\gamma_{\bar{m}\tilde{n}}^{\bar{i}\tilde{j}} \gamma_{\bar{o}\tilde{p}}^{\bar{k}\tilde{l}}}_{\text{mixed}} \\ & + \underbrace{\gamma_{\tilde{m}\tilde{n}}^{\tilde{i}\tilde{j}} \gamma_{\tilde{o}\tilde{p}}^{\tilde{k}\tilde{l}}}_{\text{spin-}\alpha} + \underbrace{\gamma_{\bar{m}\bar{n}}^{\bar{i}\bar{j}} \gamma_{\bar{o}\bar{p}}^{\bar{k}\bar{l}}}_{\text{spin-}\beta} + \underbrace{\gamma_{\tilde{m}\bar{n}}^{\tilde{i}\bar{j}} \gamma_{\tilde{o}\tilde{p}}^{\tilde{k}\tilde{l}}}_{\text{mixed}} + \underbrace{\gamma_{\bar{m}\tilde{n}}^{\bar{i}\tilde{j}} \gamma_{\bar{o}\tilde{p}}^{\bar{k}\tilde{l}}}_{\text{mixed}} \\ & + \underbrace{\gamma_{\tilde{m}\bar{n}}^{\tilde{i}\bar{j}} \gamma_{\tilde{o}\tilde{p}}^{\tilde{k}\tilde{l}}}_{\text{mixed}} + \underbrace{\gamma_{\bar{m}\tilde{n}}^{\bar{i}\tilde{j}} \gamma_{\bar{o}\tilde{p}}^{\bar{k}\tilde{l}}}_{\text{mixed}} + \underbrace{\gamma_{\tilde{m}\tilde{n}}^{\tilde{i}\tilde{j}} \gamma_{\tilde{o}\tilde{p}}^{\tilde{k}\tilde{l}}}_{\text{spin-}\alpha} + \underbrace{\gamma_{\bar{m}\bar{n}}^{\bar{i}\bar{j}} \gamma_{\bar{o}\bar{p}}^{\bar{k}\bar{l}}}_{\text{spin-}\beta} \end{aligned}$$

$$+ \gamma_{\underline{m}\underline{n}}^{\underline{i}\underline{j}} \gamma_{\underline{o}\underline{p}}^{\underline{k}\underline{l}} + \gamma_{\underline{m}\underline{n}}^{\underline{i}\underline{j}} \gamma_{\underline{o}\underline{p}}^{\underline{k}\underline{l}} + \gamma_{\underline{m}\underline{n}}^{\underline{i}\underline{j}} \gamma_{\underline{o}\underline{p}}^{\underline{k}\underline{l}} + \underline{\gamma_{\underline{m}\underline{n}}^{\underline{i}\underline{j}} \gamma_{\underline{o}\underline{p}}^{\underline{k}\underline{l}}} \quad (2.108)$$

The underlined terms are those that are present in the spatial expression above, but all other terms are not present. Therefore, the spatial expression does not contain a product of spin-summed density matrices. Shamasundar has noted that the spin-orbital reduced density matrices of non-singlet states cannot be re-expressed in terms of spin-free reduced density matrices alone [107]. It can be seen that this is the case for a 1st-order reduced density matrix evaluated for a set of  $(2S + 1)$  many-electron wavefunctions  $\{\Psi_{M_s}^S, M_s = -S, \dots, +S\}$

$$\gamma_{\underline{q}}^{\underline{p}} = \langle \Psi_{M_s}^S | \hat{p}^\dagger \underline{q} | \Psi_{M_s}^S \rangle \quad (2.109)$$

$$= \frac{1}{2} \Gamma_{\underline{q}}^{\underline{p}} + \zeta_{\underline{q}}^{\underline{p}} \quad (2.110)$$

where  $\zeta_{\underline{q}}^{\underline{p}}$  is a spin density matrix, which depends on both the total spin quantum numbers  $S$  and  $M_s$ . To get this far in the derivation of the spin-summed spin-free form, the spin-orbital form has been found, and ‘translated’ into spatial form and attempted to be written in a spin-free form via spin-summation (as Shamasundar notes it can’t be done). For a true spin-free form, the whole derivation of the theory, from the 3rd order VCC energy, must be re-done using spin-free operators.

Spin-free excitation operators are routinely used (and will be in this theory), as are spin-free density matrices which are expectation values of spin-free operators. Cumulants are not expectation values of spin-free operators, hence derivation of their spin-free form is more complicated. Spin-free cumulants, required for the cumulant approximation of the 4th order density matrix, have only recently become known and their use is still problematic.

Therefore, the most pragmatic option is not to re-derive the equations but to simply adjust factors to get the spin-free expression correct for the 2-electron singlet and triplet cases, the error for all other cases will be averaged out.

The 2 electron triplet case:

$$\Gamma_{mn}^{ij} = \gamma_{\underline{m}\underline{n}}^{\underline{i}\underline{j}} \quad ; \quad T_{ab}^{ij} = \frac{1}{4} T_{\underline{a}\underline{b}}^{\underline{i}\underline{j}} \quad (2.111)$$

$$E = 8 K_{ab}^{ij} T_{ab}^{op} T_{cd}^{mn} T_{cd}^{kl} \Gamma_{mn}^{ij} \Gamma_{op}^{kl} \quad (2.112)$$

The 2 electron singlet case:

$$\Gamma_{mn}^{ij} = \gamma_{\underline{m}\underline{n}}^{\underline{i}\underline{j}} + \gamma_{\underline{m}\underline{n}}^{\underline{j}\underline{i}} \quad ; \quad T_{ab}^{ij} = \frac{1}{2} T_{\underline{a}\underline{b}}^{\underline{i}\underline{j}} \quad (2.113)$$

$$E = 8 K_{ab}^{ij} T_{ab}^{op} T_{cd}^{mn} T_{cd}^{kl} \Gamma_{mn}^{ij} \Gamma_{op}^{kl} \quad (2.114)$$

### Symmetric and Antisymmetric

The spin-summed excitation operators are required to be re-expressed in terms of a symmetric and antisymmetric part, in the same way the MRCI code works as defined in ref [58]. That work uses the excitation operator  $\hat{C}$  while the work presented here so far uses  $\hat{T}$ .

$$\hat{C} = \frac{1}{2} \sum_{i \geq j} \sum_{ab} \sum_p C_{ab}^{ij,p} \left( \hat{E}_{ai,bj} + p \hat{E}_{aj,bi} \right) \quad (2.115)$$

$$\hat{T} = \sum_{ij} \sum_{ab} T_{ab}^{ij} \hat{E}_{ai,bj} \quad (2.116)$$

$$\begin{aligned} &= \sum_i T_{ab}^{ii} \hat{E}_{ai,bi} + \sum_{i>j} T_{ab}^{ij} \hat{E}_{ai,bj} + \sum_{i<j} T_{ab}^{ij} \hat{E}_{ai,bj} \\ &= \frac{1}{4} (T_{ab}^{ij} + T_{ba}^{ij}) \left( \hat{E}_{ai,bj} + \hat{E}_{aj,bi} \right) + \frac{1}{4} (T_{ab}^{ij} - T_{ba}^{ij}) \left( \hat{E}_{ai,bj} - \hat{E}_{aj,bi} \right) \end{aligned} \quad (2.117)$$

To allow re-expression of the T amplitudes in terms of a symmetric part and antisymmetric part, a comparison between the definition of  $\hat{C}$  (as used in ref. [58]) and  $\hat{T}$  (used here) must be made.

For the  $i=j$  case:

$$\hat{C} = \sum_{ab} C_{ab}^{ii,1} E_{ai,bi} \quad (2.118)$$

$$\hat{T} = \sum_{ab} T_{ab}^{ii} E_{ai,bi} \quad (2.119)$$

$$\text{Hence, } C_{ab}^{ii,1} = T_{ab}^{ii} \quad (2.120)$$

For the case  $i \neq j$ :

$$\hat{C} = \frac{1}{2} C_{ab}^{ij,1} \left( \hat{E}_{ai,bj} + \hat{E}_{aj,bi} \right) + \frac{1}{2} C_{ab}^{ij,-1} \left( \hat{E}_{ai,bj} - \hat{E}_{aj,bi} \right) \quad (2.121)$$

$$\begin{aligned} \hat{T} &= \frac{1}{4} (T_{ab}^{ij} + T_{ba}^{ij}) \left( \hat{E}_{ai,bj} + \hat{E}_{aj,bi} \right) + \frac{1}{4} (T_{ab}^{ij} - T_{ba}^{ij}) \left( \hat{E}_{ai,bj} - \hat{E}_{aj,bi} \right) \\ & \quad (2.122) \end{aligned}$$

$$C_{ab}^{ij,1} = \frac{1}{2} (T_{ab}^{ij} + T_{ba}^{ij}) \quad (2.123)$$

$$C_{ab}^{ij,-1} = \frac{1}{2} (T_{ab}^{ij} - T_{ba}^{ij}) \quad (2.124)$$

$$\text{Hence } T_{ab}^{ij} = C_{ab}^{ij,1} + C_{ab}^{ij,-1} \quad (2.125)$$

This also holds up for the case when  $i=j$ , because the antisymmetric amplitude will be zero by design.

Now that the amplitudes have been written in a combination of symmetric and anti-symmetric parts, the energy expression can be re-expressed using these amplitudes, together with the spin-free operators shown earlier.

Inserting the amplitude definitions into the unrestricted energy expression, it becomes:

$$8 K_{ab}^{ij} T_{cd}^{mn} T_{cd}^{kl} \Gamma_{op}^{kl} T_{ab}^{op} \Gamma_{mn}^{ij} \quad (2.126)$$

$$= 8 K_{ab}^{ij} (C_{cd}^{mn,1} + C_{cd}^{mn,-1}) (C_{cd}^{kl,1} + C_{cd}^{kl,-1}) \Gamma_{op}^{kl} (C_{ab}^{op,1} + C_{ab}^{op,-1}) \Gamma_{mn}^{ij} \quad (2.127)$$

$$= 8 K_{ab}^{ij} (C_{cd}^{mn,1} C_{cd}^{kl,1} + C_{cd}^{mn,-1} C_{cd}^{kl,-1}) \Gamma_{op}^{kl} (C_{ab}^{op,1} + C_{ab}^{op,-1}) \Gamma_{mn}^{ij} \quad (2.128)$$

Further manipulation must be carried out because the density matrices can be split up into symmetric and antisymmetric parts also.

$$\Gamma_{op}^{kl} = \langle 0 | \hat{E}_{ko,lp} | 0 \rangle \quad (2.129)$$

$$= \frac{1}{2} \langle 0 | \hat{E}_{ko,lp} + \hat{E}_{kp,lo} | 0 \rangle + \frac{1}{2} \langle 0 | \hat{E}_{ko,lp} - \hat{E}_{kp,lo} | 0 \rangle \quad (2.130)$$

Then using the definition of the overlap matrix (2nd order DM) used in the MRCI code ([58]),  $S_{op}^{kl,p} = \langle 0 | \hat{E}_{ko,lp} + p \hat{E}_{kp,lo} | 0 \rangle$ , then

$$\Gamma_{op}^{kl} = \frac{1}{2} S_{op}^{kl,1} + \frac{1}{2} S_{op}^{kl,-1} \quad (2.131)$$

The final spin-free form of the part of the energy required can be written fully,

$$= 4 K_{ab}^{ij} (C_{cd}^{mn,1} C_{cd}^{kl,1} + C_{cd}^{mn,-1} C_{cd}^{kl,-1}) (S_{op}^{kl,1} + S_{op}^{kl,-1}) (C_{ab}^{op,1} + C_{ab}^{op,-1}) \Gamma_{mn}^{ij} \quad (2.132)$$

$$= 4 K_{ab}^{ij} (C_{cd}^{mn,1} C_{cd}^{kl,1} S_{op}^{kl,1} C_{ab}^{op,1} + C_{cd}^{mn,-1} C_{cd}^{kl,-1} S_{op}^{kl,-1} C_{ab}^{op,-1}) \Gamma_{mn}^{ij} \quad (2.133)$$

$$= 2 K_{ab}^{ij} C_{cd}^{mn,1} C_{cd}^{kl,1} S_{op}^{kl,1} C_{ab}^{op,1} S_{mn}^{ij,1} + 2 K_{ab}^{ij} C_{cd}^{mn,-1} C_{cd}^{kl,-1} S_{op}^{kl,-1} C_{ab}^{op,-1} S_{mn}^{ij,-1} \quad (2.134)$$

$$= 2 K_{ab}^{ij} C_{cd}^{mn,P} C_{cd}^{kl,P} S_{op}^{kl,P} C_{ab}^{op,P} S_{mn}^{ij,P} \quad (2.135)$$

This is the part of the energy to capture via a transformation, and is now written using the same operators as the MRCI code in MOLPRO. The transformation can be re-defined in this form:

$$U_p^{-\frac{q}{2}} C_{ab}^{op,P} = \left( \delta_{op}^{mn} + C_{cd}^{mn,P} C_{cd}^{kl,P} S_{op}^{kl,P} \right)^{-\frac{q}{2}} C_{ab}^{op,P} = {}_q C_{ab}^{mn,P} \quad (2.136)$$

This definition of the energy and of the transformation cannot be correct for the general case, as the spin free formalism is not correct.

## 2.3 Redundancies and Conclusions

In multireference methods, complications arise due to redundancies in the excitation operator, due to the equivalence of pairs of electrons. The redundancies in the excitation operator show themselves as singularities in the pair overlap matrix  $S_{ij,kl}^{(p)}$ ,

$$S_{ij,kl}^{(p)} = \langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 0 \rangle \quad (2.137)$$

In the icMRCI code in MOLPRO, the pairs with small values in the overlap matrix (below a cutoff) are deleted until the overlap matrix is no longer singular, termed here “pair deletion”. However, recent research [53] shows that this pair deletion method cannot be used for a multireference Coupled Cluster if orbital invariance is to be retained, as all pair products are needed when considering powers of the amplitudes. This is not a consideration in MRCI because it does not contain powers of the operators/amplitudes. Instead the cited work suggested to use a singular value decomposition to remove redundancies, while retaining orbital invariance. This also has the effect of orthogonalising the excitation operators.

The theory described in this chapter uses explicit orbital labels not orthogonal pairs. Like the theory presented here, the single-reference LPF theories use explicit orbital labels. However, the single-reference situation is fundamentally different to the multireference case in that redundancies do not arise and the electron pairs are automatically orthogonal as a result of the reference consisting of a single Slater determinant. The orthogonality of the pairs can be shown simply in spin-orbital form,

$$\langle \psi_{kl}^{cd} | \psi_{ij}^{ab} \rangle = \delta_a^c \delta_b^d \delta_i^k \delta_j^l - \delta_a^c \delta_b^d \delta_j^k \delta_i^l - \delta_b^c \delta_a^d \delta_i^k \delta_j^l + \delta_b^c \delta_a^d \delta_j^k \delta_i^l \quad (2.138)$$

The single-reference theory can be written using explicit electron/orbital labels due to the orthogonality of the electron pairs. In the multireference case, electron pairs are not automatically orthogonal, and as stated above, orthogonalisation and redundancy removal must be carried out using a singular value decomposition. The result of the singular value decomposition is a set of non-redundant orthogonal excitation operators, generating orthogonal configurations, meaning they cannot be labelled by explicit electron/orbital labels. The use of orthogonal pairs cannot be integrated into the work presented in this chapter, which has been built using explicit electron labels. Without the use of a singular value decomposition (generating orthogonal pairs) and instead using pair deletion, the method presented here loses its property of the invariance to orbital rotations, which has been highlighted as a key property.

In addition to the lack of orbital invariance, it was noted earlier that the spin-free formalism does not correspond to the spin-orbital form, with its behaviour for more than 2 electrons unknown. With these errors in mind, better MRVCC approximations must be sought in a spin-free form that are orbital invariant, which can be ensured by using a singular value decomposition to eliminate redundancies rather than pair deletion, resulting in a theory based on orthogonal electron pairs.



# Chapter 3

## A Multireference LPF+1 Theory

### 3.1 Preliminaries

As mentioned in the introductory chapter, MRCI is not size consistent or extensive and there have been methods developed to try to correct this behaviour. The aim of this present chapter is to define a transformation of the pair amplitudes within MRCI akin to the LPF theories that introduces a linked denominator and via a binomial expansion can be seen to capture 3rd order VCC energy terms, hence approximating VCC, within a multireference framework.

Unlike the single reference theory where the orbital spaces are split into occupied and virtual, the multireference case has the virtual and internal orbital spaces. The internal orbital space consists of those orbitals occupied in any of the references. The internal orbitals are further split into active and core, core being orbitals doubly occupied in all configurations and active being the rest. Excitations of the kind in CI and Coupled Cluster can occur between active and virtual or between active orbitals.

The icMRCI energy can be written as follows,

$$E_{corr} = \frac{\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} + \frac{2 \langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle + \langle 0 | \hat{T}^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \quad (3.1)$$

where  $|\Psi\rangle = |0\rangle + \hat{c}|0\rangle + \hat{T}|0\rangle$

$$\begin{aligned}
|0\rangle &= \hat{c}_0 |\Psi_0\rangle = \frac{1}{n_{act}!} \sum_{uv\dots z} \hat{a}^{uv\dots z} |\Psi_0\rangle c_{uv\dots z} \\
\langle 0 | \hat{c} | 0 \rangle &= \mathbf{c}_0^\dagger \mathbf{c}' = 0
\end{aligned} \tag{3.2}$$

The  $|0\rangle$  wavefunction is the reference function, it is a CASSCF wavefunction generated from the single input configuration  $|\Psi_0\rangle$ .  $\hat{c}$ , here represents excitations amongst the internal configurations from the starting 0th order CASSCF wavefunction, i.e. it carries out the pure active excitations, those between active orbitals. It produces the orthogonal complement mentioned in the introductory chapter, whose coefficients of the internal configurations are written  $\mathbf{c}'$ . This operator ensures the result of its action is orthogonal to the reference state.

Initially development of an MRLPF method should focus on the double excitations to the virtual orbitals (external space), excitations that are often termed doubly-external or pair-external.

Therefore the energy of the MRLPF+1 would be of the form:

$$\begin{aligned}
E_{MRLPF}^{corr} &= \frac{\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \\
&\quad + 2 \langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle + \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle
\end{aligned} \tag{3.3}$$

where the terms involving double-external excitations have been separated from those without. The development of an MRLPF+1 method must introduce the linked denominator, and simultaneously capture terms from 3rd order VCC. This method must also be correct for 2 electrons, must be invariant to the choice of orbitals, reduce to the single reference LPF+1 theory and be size extensive. Initial development focuses on the transformation of the pair amplitudes, removing the extensivity error introduced by unlinked pair terms. This neglects the extensivity error of the first term in the energy above, where the complete pair norm, including unlinked terms is used.

As with the LPF theories the transformation of the amplitudes takes the form,

$${}_q\mathbf{T} = \mathbf{U}^{-\frac{q}{2}} \mathbf{T} \tag{3.4}$$

Where the transformation matrix will be of the form,

$$\mathbf{U} = \mathbf{1} + \eta \tag{3.5}$$

Therefore the development of a multireference LPF theory must focus on the development of an appropriate  $\eta$  matrix that satisfies the requirements stated above.

### 3.1.1 Extending LPF

The LPF+1  $\eta$  matrix in spin orbital form can be written,

$$\eta_{kl}^{ij} = \langle \Psi_0 | \hat{T}^\dagger k l j^\dagger i^\dagger \hat{T} | \Psi_0 \rangle \quad (3.6)$$

A basic multireference formalism would simply use the same  $\eta$  matrix but have multiple configurations in the 0th order wavefunction.

This formalism is fundamentally different to the single reference LPF theory. In LPF, the  $ji$  creation operators must match the occupied orbitals destroyed in  $\hat{T}$ , otherwise,  $ji$  are already occupied and the wavefunction is annihilated. In the multireference case,  $ji$  does not need to match  $\hat{T}$  in order for the contribution to be non-zero.  $ij$  may create in any orbitals that are not occupied in a specific reference, so this includes the orbitals/electrons annihilated in  $\hat{T}$  and any other reference determinant where  $ij$  are empty. This formalism would lead to amplitudes being included in the  $\Delta$  matrix that are nothing to do with  $ij$  and  $kl$ , and they would therefore not be linked to the amplitude being transformed.

An illustrative example can show how the amplitudes from  $\eta$  are not necessarily linked to the amplitude which is being transformed.

$$\begin{aligned} (\delta_{kl}^{ij} + \eta_{kl}^{ij})^{-1} T_{ab}^{kl} &= (\delta_{kl}^{ij} + \langle \hat{T}^\dagger k l j^\dagger i^\dagger \hat{T} \rangle)^{-1} T_{ab}^{kl} \\ &= \left( \delta_{kl}^{ij} + T_{cd}^{ij} T_{kl}^{cd} + T_{cd}^{uv} T_{wx}^{cd} + T_{\alpha\beta}^{yz} T_{st}^{\alpha\beta} + \dots \right)^{-1} T_{ab}^{kl} \end{aligned} \quad (3.7)$$

where  $uvwxyz$  are internal indices and  $\alpha$  and  $\beta$  indicate virtual orbitals. Expanding this using the binomial expansion gives terms such as:

$$T_{ab}^{ij} - T_{cd}^{ij} T_{kl}^{cd} T_{ab}^{kl} - T_{cd}^{uv} T_{wx}^{cd} T_{ab}^{kl} - T_{\alpha\beta}^{yz} T_{st}^{\alpha\beta} T_{ab}^{kl} \quad (3.8)$$

The 3rd and 4th terms in 3.8 can give unlinked contributions. This inclusion of unlinked diagrams is directly against the principle of LPF and violates the linked-diagram rules required for size-extensivity. For this reason, this direct generalisation of the LPF+1 theory to the case of multiple reference determinants cannot be used.

Applying this to the model case gives the correct denominator, however, it is conceptually different to the single reference LPF theory. In SR-LPF the hole density matrix creation indices must match the indices that are excited in the  $\hat{T}$  operator. If the indices do not match, then electrons are trying to be created in orbitals that are already occupied, therefore annihilating the wavefunction. This ensures that only linked diagrams are included because if the contribution to  $\eta$  does not link to the amplitude to be transformed, then it is annihilated. In the multi reference case the application of a creation operator indexed as internal does not necessarily need to match the indices of the  $\hat{T}$  operator, because the  $i^\dagger j^\dagger$  orbitals may not be occupied and therefore creating there does not annihilate the wavefunction. In summary, in SR-LPF the  $\eta$  creation operators,  $i^\dagger j^\dagger$ , must match the  $\hat{T}$  operator indices because those indices are the only space in which you can create that does not result in annihilation, this is not true in the multireference case where there are multiple references you could additionally create in.

#### Single Reference

$$o p j^\dagger i^\dagger \hat{T}_{kl}^{ab} |0\rangle = 0 \quad (ij \neq kl) \quad (3.9)$$

$$o p j^\dagger i^\dagger \hat{T}_{ij}^{ab} |0\rangle \neq 0 \quad (3.10)$$

#### Multi Reference

$$o p j^\dagger i^\dagger \hat{T}_{kl}^{ab} |0\rangle \neq 0 \quad (ij \neq kl) \quad (3.11)$$

### Second Possibility

The aforementioned difference in the multi-reference LPF method compared to the single reference theory must lead to a different formalism. The deficiency with the previous method above, i.e. the single reference LPF+1 transformation applied to the multi-reference case, was that the indices of the creation operators did not necessarily have to match the  $\hat{T}$  operator, as mentioned above. A simple modification to suit the multireference case is to change the definition of  $\eta$  so that the creation operators always match those annihilated by the  $\hat{T}$  operator, giving an effect analogous to the single reference theory. This can be accomplished via the use of a commutator, as shown below for the  $\eta$  matrix that generates term 3B (analogous to single reference

LPF0).

$$\eta_o^i = \langle \hat{T}^\dagger [oi^\dagger, \hat{T}] \rangle \quad (3.12)$$

$$(3.13)$$

The commutator is only non-zero when the creation operators  $i$  and  $j$  match the indices in the  $\hat{T}$  operator. In SR LPF  $ijj^\dagger i^\dagger \hat{T}|0\rangle = \hat{T}|0\rangle$ , this commutator approach can be proved to have the same property.

$$\begin{aligned} & [ii^\dagger, \hat{T}]|0\rangle \\ &= ii^\dagger \hat{T}|0\rangle - \hat{T}ii^\dagger|0\rangle \end{aligned} \quad (3.14)$$

$$= ii^\dagger |\Phi_{mn}^{ab}\rangle - \hat{T}ii^\dagger|0\rangle \quad (3.15)$$

If  $i \neq m$

$$j \text{ occupied} : 0 - 0 \quad (3.16)$$

$$j \text{ unoccupied} : |\Phi_{mn}^{ab}\rangle - |\Phi_{mn}^{ab}\rangle \quad (3.17)$$

$$\text{therefore } [ii^\dagger, \hat{T}]|0\rangle = 0 \quad (3.18)$$

If  $i = m$

$$|\Phi_{mn}^{ab}\rangle - 0 \quad (3.19)$$

$$\text{therefore } [ii^\dagger, \hat{T}]|0\rangle \neq 0 \quad (3.20)$$

This therefore may appear to be a satisfactory generalisation of LPF to the case of multiple reference functions, however, when applied to the model 2-electron 2-reference example, the method does not generate the CI denominator that we require for 2 electrons. Looking at the 1,1 element of  $\eta$ ,

$$\langle \hat{T}^\dagger [11^\dagger, \hat{T}] \rangle \quad (3.21)$$

$$\hat{T}^\dagger [(11^\dagger \hat{T}|0\rangle) - (\hat{T}11^\dagger)] \quad (3.22)$$

$$= \hat{T}^\dagger [(11^\dagger (\cos\theta T_{33}^{11}|3^2\rangle + \sin\theta T_{33}^{22}|3^2\rangle)) - (\hat{T} \sin\theta |2^2\rangle)] \quad (3.23)$$

$$= \hat{T}^\dagger [( \cos\theta T_{33}^{11}|3^2\rangle + \sin\theta T_{33}^{11}|3^2\rangle) - (\sin\theta T_{33}^{22}|3^2\rangle)] \quad (3.24)$$

$$= \hat{T}^\dagger [\cos\theta T_{33}^{11}|3^2\rangle] \quad (3.25)$$

$$= \cos\theta \cos\theta T_{33}^{11} \langle 3^2|3^2\rangle + \cos\theta \sin\theta T_{33}^{11} T_{33}^{22} \langle 3^2|3^2\rangle \quad (3.26)$$

$$= \cos^2\theta T_{33}^{11}T_{33}^{11} + \cos\theta\sin\theta T_{33}^{11}T_{33}^{22} \quad (3.27)$$

This is only part of the required denominator,  $(\cos\theta T_{33}^{11} + \sin\theta T_{33}^{22})^2$ . The 2,2 element contains the other half of the required denominator  $\sin^2\theta T_{33}^{22} + \cos\theta\sin\theta T_{33}^{11}T_{33}^{22}$ . The operators in the commutator can be normal ordered, but this does not give the required denominator.

$$\eta_j^i = \langle [\hat{T}^\dagger, [ij^\dagger, \hat{T}]] \rangle \quad (3.28)$$

$$ij^\dagger = \delta_{ij} - j^\dagger i \quad (3.29)$$

$$= [\delta_{ij} - j^\dagger i, \hat{T}] \quad (3.30)$$

$$= (\delta_{ij} - j^\dagger i)\hat{T} - \hat{T}(\delta_{ij} - j^\dagger i) \quad (3.31)$$

$$= -j^\dagger i \hat{T} + \hat{T} j^\dagger i \quad (3.32)$$

$$= -\langle [\hat{T}^\dagger, [j^\dagger i, \hat{T}]] \rangle \quad (3.33)$$

This method of building a direct analogue of LPF in a MR framework is not exact for 2 electrons.

The transformation must be linked to the amplitude it is transforming, giving only linked diagrams. However, in order to be exact for 2 electrons, terms that are not linked in the internal indices must be included. This is not a problem in single reference LPF because there are only 2 electrons, hence only 2 occupied indices, therefore all amplitudes in the transformation matrix are linked to the amplitude to be transformed, generating only linked diagrams.

In the MR case, there are more than 2 internal orbitals for the 2 electrons. This is because each reference has 2 internal orbitals, hence in total there are more than 2 internal orbitals. Therefore there is the opportunity for amplitudes to share or not share their internal indices, this does not happen in the SR-LPF for 2 electrons, it only happens with more electrons which defines the method by only including linked diagrams. So, not all internal indices match those of the untransformed amplitude (as they do in SR-LPF) and are therefore not included in the transformation because they would not produce linked diagrams. Hence these methods do not produce the full denominator for the 2 electron case because it would require unlinked contributions. It is therefore difficult to produce a theory that only contains linked diagrams and simultaneously is exact for 2 electrons, it can only be achieved in SR-LPF because

there are only 2 occupied orbitals, which is a special case.

### A way to progress....

A key reason why the previous attempts at a multireference LPF+1 theory do not give the desired results (exact for 2 electrons and reduce to single-reference theory), is the non-vanishing case of the  $\hat{T}$  excitation operator not matching indices on the  $\eta$  matrix. This leads to the presence of a density matrix evaluated across the reference. However, this difficulty can be overcome by realising that for a single reference theory with a Hartree-Fock determinant,  $\langle \Phi_{ab}^{ij} | \Phi_{kl}^{ab} \rangle$  is non-zero only if the internal orbital indices match, meaning that the electron pairs are orthogonal. In the multireference case, electron pairs can be made to be orthogonal via an orthogonalisation procedure, meaning that the indices on the  $\eta$  matrix must match the  $\hat{T}$  excitation operator. (This was the aim of the section above but yielded no good result). The multireference LPF+1 theory can then be formulated in terms of orthogonal pairs. A theory based on orthogonal pairs may also use the redundancy removal procedure of Evangelista & Gauss [53] (unlike the work in the previous chapter), which also orthogonalises the operators, which will preserve orbital invariance.

## 3.1.2 Operator Considerations

The work in chapter 2 attempted to approximate 3rd order variational coupled cluster theory. It started from a spin-orbital formulation and was later translated to a spin free formulation. This led to difficulties, so attempts here to form a multireference analogue of LPF+1 use the spin free operators from the outset, as these are the operators used in the MRCI program in MOLPRO.

A comparison of the operators used in different methods must be made. Below are 3 parameterisations of the excitation operator  $\hat{T}$ . Type 1 is in spin orbital notation, Type 2 is a spatial operator used in closed shell coupled cluster theory in MOLPRO and Type 3 is a spin summed operator used for open shell icMRCI code in MOLPRO (as used in the definition of the icMRCI wavefunction in 1.125).

$$\text{Spin Orbitals: } \hat{T} = \frac{1}{2} \sum_{ij} \sum_{ab} T_{ab}^{ij} a^\dagger b^\dagger j i \quad (3.34)$$

$$\text{Spatial Orbitals} \quad \hat{T} = \sum_{ij} \sum_{ab} T_{ab}^{ij} \hat{E}_{ai,bj} \quad (3.35)$$

$$\text{Spatial with parity} \quad \hat{C} = \frac{1}{2} \sum_{i \geq j} \sum_{ab} \sum_p C_{ab}^{ij,p} \left( \hat{E}_{ai,bj} + p \hat{E}_{aj,bi} \right) \quad (3.36)$$

Compare

$$\hat{T} = \sum_{ij} \sum_{ab} T_{ab}^{ij} \hat{E}_{ai,bj} \quad (3.37)$$

$$= \sum_i T_{ab}^{ii} \hat{E}_{ai,bi} + \sum_{i>j} T_{ab}^{ij} \hat{E}_{ai,bj} + \sum_{i<j} T_{ab}^{ij} \hat{E}_{ai,bj} \quad (3.38)$$

$$= \frac{1}{4} (T_{ab}^{ij} + T_{ba}^{ij}) \left( \hat{E}_{ai,bj} + \hat{E}_{aj,bi} \right) + \frac{1}{4} (T_{ab}^{ij} - T_{ba}^{ij}) \left( \hat{E}_{ai,bj} - \hat{E}_{aj,bi} \right) \quad (3.39)$$

To allow re-expression of the T amplitudes in terms of a symmetric part and antisymmetric part, a comparison between the definition of  $\hat{C}$  and  $\hat{T}$  must be made.

For the  $i=j$  case:

$$\hat{C} = \sum_{ab} C_{ab}^{ii,1} E_{ai,bi} \quad (3.40)$$

$$\hat{T} = \sum_{ab} T_{ab}^{ii} E_{ai,bi} \quad (3.41)$$

$$\text{Hence, } C_{ab}^{ii,1} = T_{ab}^{ii} \quad (3.42)$$

For the case  $i \neq j$ :

$$\hat{C} = \frac{1}{2} C_{ab}^{ij,1} \left( \hat{E}_{ai,bj} + \hat{E}_{aj,bi} \right) + \frac{1}{2} C_{ab}^{ij,-1} \left( \hat{E}_{ai,bj} - \hat{E}_{aj,bi} \right) \quad (3.43)$$

$$\hat{T} = \frac{1}{4} (T_{ab}^{ij} + T_{ba}^{ij}) \left( \hat{E}_{ai,bj} + \hat{E}_{aj,bi} \right) + \frac{1}{4} (T_{ab}^{ij} - T_{ba}^{ij}) \left( \hat{E}_{ai,bj} - \hat{E}_{aj,bi} \right) \quad (3.44)$$

$$C_{ab}^{ij,1} = \frac{1}{2} (T_{ab}^{ij} + T_{ba}^{ij}) \quad (3.45)$$

$$C_{ab}^{ij,-1} = \frac{1}{2} (T_{ab}^{ij} - T_{ba}^{ij}) \quad (3.46)$$

$$\text{Hence } T_{ab}^{ij} = C_{ab}^{ij,1} + C_{ab}^{ij,-1} \quad (3.47)$$

This also holds up for the case when  $i=j$ , because the antisymmetric amplitude will be zero by design.

The LPF+1D theory constructed from the  $\hat{T}$  using spatial orbitals only will not be the same as the theory constructed that includes the parity sum. In LPF+1D the amplitude transformation introduces the linked denominator

$$U_{kl}^{ij} = \delta_{kl}^{ij} + \frac{1}{2} T_{cd}^{ij} T_{kl}^{cd} \quad (3.48)$$

$${}_q T_{ab}^{ij} = \frac{1}{2} \left( \mathbf{U}^{-\frac{q}{2}} \right)_{kl}^{ij} T_{ab}^{kl} \quad (3.49)$$

However, using the parity summed operators, the transformation can only be constructed to involve amplitudes of the same parity, to transform amplitudes of the same parity

$$U_{klp}^{ijp} = \delta_{klq}^{ijp} + C_{cd}^{ijp} C_{klp}^{cd} \quad (3.50)$$

$${}_q C_{ab}^{ijp} = \left( \mathbf{U}^{-\frac{q}{2}} \right)_{kl}^{ij} C_{ab}^{klp} \quad (3.51)$$

Hence, the result is just different because using the parity sum operator does not introduce the full denominator required. The different partitioning causes a difference in the transformation, and hence the energy.

Despite the difference in the theories, they are equivalent in the limiting case of 2 electrons.

The Singlet Operator:

$$\frac{1}{2} \left( \hat{E}_{ai,bj} + \hat{E}_{aj,bi} \right) \quad (3.52)$$

$$= \frac{1}{2} (a^\dagger i + \bar{a}^\dagger \bar{i}) (b^\dagger j + \bar{b}^\dagger \bar{j}) + \frac{1}{2} (a^\dagger j + \bar{a}^\dagger \bar{j}) (b^\dagger i + \bar{b}^\dagger \bar{i}) \quad (3.53)$$

$$= \frac{1}{2} (a^\dagger b^\dagger ij + \bar{a}^\dagger \bar{b}^\dagger \bar{i} \bar{j} + \bar{a}^\dagger b^\dagger \bar{i} j + a^\dagger \bar{b}^\dagger i \bar{j} + a^\dagger b^\dagger ji + \bar{a}^\dagger \bar{b}^\dagger \bar{j} \bar{i} + \bar{a}^\dagger b^\dagger \bar{j} i + a^\dagger \bar{b}^\dagger j \bar{i}) \quad (3.54)$$

$$= \frac{1}{2} (a^\dagger b^\dagger ij + \bar{a}^\dagger \bar{b}^\dagger \bar{i} \bar{j} + \bar{a}^\dagger b^\dagger \bar{i} j + a^\dagger \bar{b}^\dagger i \bar{j} - a^\dagger b^\dagger ij - \bar{a}^\dagger \bar{b}^\dagger \bar{i} \bar{j} - \bar{a}^\dagger b^\dagger \bar{i} j - a^\dagger \bar{b}^\dagger i \bar{j}) \quad (3.55)$$

$$= \frac{1}{2} (\bar{a}^\dagger b^\dagger \bar{i} j - \bar{a}^\dagger b^\dagger i \bar{j} + a^\dagger \bar{b}^\dagger i \bar{j} - a^\dagger \bar{b}^\dagger \bar{i} j) \quad (3.56)$$

$$= \frac{1}{2} (\bar{a}^\dagger b^\dagger (\bar{i} j - i \bar{j}) + a^\dagger \bar{b}^\dagger (i \bar{j} - \bar{i} j)) \quad (3.57)$$

The Triplet Operator:

$$\frac{1}{2} \left( \hat{E}_{ai,bj} - \hat{E}_{aj,bi} \right) \quad (3.58)$$

$$= \frac{1}{2} (a^\dagger i + \bar{a}^\dagger \bar{i}) (b^\dagger j + \bar{b}^\dagger \bar{j}) - \frac{1}{2} (a^\dagger j + \bar{a}^\dagger \bar{j}) (b^\dagger i + \bar{b}^\dagger \bar{i}) \quad (3.59)$$

$$= \frac{1}{2} (a^\dagger b^\dagger ij + \bar{a}^\dagger \bar{b}^\dagger i\bar{j} + \bar{a}^\dagger b^\dagger i\bar{j} + a^\dagger \bar{b}^\dagger i\bar{j} - a^\dagger b^\dagger ji - \bar{a}^\dagger \bar{b}^\dagger j\bar{i} - \bar{a}^\dagger b^\dagger j\bar{i} - a^\dagger \bar{b}^\dagger j\bar{i}) \quad (3.60)$$

$$= \frac{1}{2} (a^\dagger b^\dagger ij + \bar{a}^\dagger \bar{b}^\dagger i\bar{j} + \bar{a}^\dagger b^\dagger i\bar{j} + a^\dagger \bar{b}^\dagger i\bar{j} + a^\dagger b^\dagger ij + \bar{a}^\dagger \bar{b}^\dagger i\bar{j} + \bar{a}^\dagger b^\dagger i\bar{j} + a^\dagger \bar{b}^\dagger i\bar{j}) \quad (3.61)$$

$$= a^\dagger b^\dagger ij + \frac{1}{2} (\bar{a}^\dagger b^\dagger + a^\dagger \bar{b}^\dagger) (i\bar{j} + i\bar{j}) + \bar{a}^\dagger \bar{b}^\dagger i\bar{j} \quad (3.62)$$

In the 2 electron singlet case,  $i=j$ , therefore the only excitation operator to survive in both representations is  $E_{a1,b1}$ , and the antisymmetric part of the parity sum is 0.

$$\begin{aligned} \text{Spatial Orbitals: } \hat{T} &= \sum_{ab} T_{ab}^{11} \hat{E}_{a1,b1} \\ &= \sum_{ab} T_{ab}^{11} (a^\dagger b^\dagger ij + \bar{a}^\dagger \bar{b}^\dagger i\bar{j} + \bar{a}^\dagger b^\dagger i\bar{j} + a^\dagger \bar{b}^\dagger i\bar{j}) \end{aligned} \quad (3.63)$$

$$= \sum_{ab} T_{ab}^{11} (\bar{a}^\dagger b^\dagger i\bar{j} + a^\dagger \bar{b}^\dagger i\bar{j}) \quad (3.64)$$

$$\text{Spatial with parity: } \hat{C} = \frac{1}{2} \sum_{ab} C_{ab}^{ij,+1} (\hat{E}_{a1,b1} + \hat{E}_{a1,b1}) \quad (3.65)$$

$$= \frac{1}{2} \sum_{ab} C_{ab}^{ij,+1} (\bar{a}^\dagger b^\dagger i\bar{j} + a^\dagger \bar{b}^\dagger i\bar{j}) \quad (3.66)$$

$$\hat{T} = \hat{C} \quad (3.67)$$

In the 2 electron triplet case,  $i \neq j$  and the spin of the electrons are the same,

$$\text{Spatial Orbitals: } \hat{T} = \sum_{ab} T_{ab}^{12} \hat{E}_{a1,b2} + T_{ab}^{21} \hat{E}_{a2,b1} \quad (3.68)$$

$$\begin{aligned} &= \sum_{ab} T_{ab}^{12} (a^\dagger b^\dagger ij + \bar{a}^\dagger \bar{b}^\dagger i\bar{j} + \bar{a}^\dagger b^\dagger i\bar{j} + a^\dagger \bar{b}^\dagger i\bar{j}) \\ &= \sum_{ab} (a^\dagger b^\dagger ij) T_{ab}^{12} \end{aligned} \quad (3.69)$$

$$\begin{aligned} \text{Spatial with parity: } \hat{C} &= \frac{1}{2} \sum_{ab} C_{ab}^{21,1} (\hat{E}_{a2,b1} + \hat{E}_{a1,b2}) \\ &\quad + C_{ab}^{21,-1} (\hat{E}_{a2,b1} - \hat{E}_{a1,b2}) \end{aligned} \quad (3.70)$$

$$= \frac{1}{2} \sum_{ab} C_{ab}^{21,1} (0) + C_{ab}^{21,-1} (a^\dagger b^\dagger ij) \quad (3.71)$$

$$\text{Therefore: } \hat{C} = \hat{T} \quad (3.72)$$

Therefore a theory constructed using the operators in the icMRCI code ( $\hat{E}$  operators and a parity sum) is not the same as the single reference LPF+1 theory (as published) in general, but is the same for both 2-electron cases, the singlet and triplet.

This can be seen in the simple examples in the table below showing the LPF+1 theory implemented with the parity summed operators is equivalent to the published full LPF+1D method for 2 electrons (helium), and a system of 2 separated 2 electron subsystems (He-He), but different otherwise (Be example).

|                                | CID          | LPF+1D (parity operators) | LPF+1D       |
|--------------------------------|--------------|---------------------------|--------------|
| He                             | -2.88759250  | -2.88759250               | -2.88759250  |
| He <sub>2</sub> (r=5.915 Bohr) | -5.77472763  | -5.77519127               | -5.77519127  |
| Be <sub>2</sub> (r=3.444)      | -29.18316881 | -29.19755451              | -29.19764021 |

Table 3.1: Comparison of LPF+1 implementations using different operators. Basis=cc-pVDZ. Energy in Hartree

A theory constructed using the spin free operators with the parity sum is theoretically different to the LPF+1D method because it doesn't introduce the full linked denominator, but still may be useful. Since the MRLPF+1D method is designed to be a correction to MRCI, it must necessarily be constructed using the same operators as the MRCI code in order to be implemented in MOLPRO. The theory presented here will continue to use the spin-free operators with a parity sum, but will change notation slightly from amplitudes denoted  $C_{ab}^{ijp}$  resembling the icMRCI paper of Werner and Knowles [58], to denoting the same amplitude as  $T_{ab}^{ijp}$ . This change is necessary to switch focus from the differing  $\hat{T}$  operators to highlighting that the amplitude is related to the pair-externals and should not be confused with the internal coefficients that are denoted  $c_I$ .

### 3.1.3 Redundancy Removal

The internally contracted nature of the reference function causes singularities to appear in the metric. For the two-determinant-reference example,

$$|\Psi_0\rangle = \hat{a}^{uv}|0\rangle_{c_{uv}} + \hat{a}^{wx}|0\rangle_{c_{wx}} \quad (3.73)$$

Excitations of electrons from these active orbitals into the virtual orbitals generate a linearly dependent set. The linearly dependent excitations can be removed without affecting the convergence of the theory to FCI. As this problem is due to the contracted nature of the reference function, internally contracted MRCI also encounters the same problem.

In the icMRCI formulation of Werner and Knowles[58], the matrix  $\mathbf{T}^{(p)}$ , is a square matrix that rotates a non-redundant subset of the individual excitation operators; there are multiple possibilities for the elimination of excitations to achieve this, and for CI, the choice does not matter. The icMRCI implementation simply removes redundant pairs based on their near zero values in the metric. Evangelista & Gauss[53] and Köhn & Hanauer[55], have recently shown that unlike icMRCI, the icMRCC solution depends on the method of removal of these redundancies to make the excitation basis, and invariance of the energy with respect to active-active rotations is also dependent on the method of redundancy removal. They provide a proof that orbital invariance can be maintained if the non-redundant excitation operators are formed by a transformation, of the redundant operator set, of the form,

$$\hat{\mathbf{T}}_{\text{in}} = \hat{\mathbf{T}} \mathbf{Z} \quad (3.74)$$

In their work they use a singular value decomposition (SVD), which will also be used in the present work, as the procedure serves to create a linearly independent excitation basis and canonically orthogonalises it; a requisite of the MRLPF+1D theory.

The procedure for removal of redundancies is now described,

$$S_{ij,kl}^{(p)} = \langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 0 \rangle \quad (3.75)$$

$$\mathbf{s}_{(p)} = \mathbf{V}^\dagger \mathbf{S}^{(p)} \mathbf{V} \quad (3.76)$$

where  $\mathbf{V}$  is a matrix that has  $M$  rows,  $M' \leq M$  columns, that is the left slice of an  $M \times M$  orthogonal matrix that brings  $\mathbf{S}$  to diagonal form including its null space.

$\mathbf{s}$  is thus an  $M' \times M'$  diagonal matrix. All singular values  $\mathbf{S}$  below a cutoff  $\epsilon$ , are neglected, leaving a set of linearly independent columns in the rectangular matrix  $\mathbf{s}^{-\frac{1}{2}\epsilon}$ , of dimension  $M \times M_{in}$ , whose elements are defined below,

$$\left(\mathbf{s}_\epsilon^{-\frac{1}{2}}\right)_{pq} = \begin{cases} \left(\mathbf{s}^{-\frac{1}{2}}\right)_{pp} & \text{if } p = q \text{ and } s_{pp} \geq \epsilon \\ 0 & \text{if } p \neq q \text{ or } s_{pp} < \epsilon \end{cases}$$

The transformation matrix  $\mathbf{Z}^{(p)}$  to transform between orthogonal and elementary generators is defined through

$$\mathbf{Z}^{(p)} = \mathbf{V} \mathbf{s}_\epsilon^{-\frac{1}{2}} \quad (3.77)$$

The transformation of the individual excitation operators and amplitudes in the cluster operator  $\hat{T}$  is defined previously (above), but is restated here.

$$\hat{\mathbf{T}}_{in} = \mathbf{Z}^{(p)} \hat{\mathbf{T}} \quad (3.78)$$

$$\hat{E}_{Dp}^{ab} = \sum_{i \geq j} \hat{E}_{ijp}^{ab} T_{D,ij}^{(p)} \quad (3.79)$$

$$C_{ab}^{ijp} = \sum_D T_{ij,D}^{(p)} T_{ab}^{Dp} \quad (3.80)$$

The cutoff value for the singular values/eigenvalues of  $\mathbf{S}$  used by Evangelista & Gauss[53] and Hanauer & Köhn [55] varies, typically being between  $10^{-4}$  or  $10^{-6}$ [59].

## 3.2 Construction of the Transformation

### 3.2.1 Form of $\eta$ using Orthogonal Pairs

As stated above, a transformation of the amplitudes is required of the form,

$${}_q T_{kl}^{ij} = \left( \mathbf{U}^{-\frac{q}{2}} \right) T_{ab}^{kl} \quad (3.81)$$

Where the transformation matrix can be written as,

$$U_{kl}^{ij} = \delta_{kl}^{ij} + \eta_{kl}^{ij} \quad (3.82)$$

The addition of the identity matrix gives the correct behaviour when considering the binomial expansion of the powered matrix, leaving the  $\eta$  matrix which is designed to contract with the amplitude to be transformed to generate terms present in 3rd order VCC. For limiting chosen cases, the matrix transformation must introduce the correct denominator for the amplitudes.

The challenge is to design an appropriate  $\eta$  matrix.

Like the LPF+1  $\eta$  matrix, an analogous matrix can be written in purely spatial operator form. Unlike the single reference case, the multireference  $\eta$  matrix does not simplify easily.

$$\eta_{ijp}^{kl\gamma} = \langle 0 | \hat{T}^\dagger \hat{E}_{kl\gamma}^{cd} \hat{E}_{ab}^{ijp} \hat{T} | 0 \rangle \quad (3.83)$$

$$= \langle 0 | \hat{E}_{ef}^{mnc} \hat{E}_{kl\gamma}^{cd} \hat{E}_{ab}^{ijp} \hat{E}_{rsq}^{gh} | 0 \rangle T_{rsq}^{gh} T_{ef}^{mnc} \quad (3.84)$$

Using the definition and commutation relations

$$\Psi_{\text{CI}} = (1 + \hat{T}) \Psi_0 \quad (3.85)$$

$$\hat{T} = \sum_{i \geq j} \sum_p \sum_{ab} T_{ab}^{ijp} \hat{E}_{ijp}^{ab} \quad \hat{E}_{ijp}^{ab} = \frac{1}{2} (\hat{E}_{ai,bj} + p \hat{E}_{bi,aj}) \quad (3.86)$$

$$= \sum_D \sum_p \sum_{ab} T_{ab}^{Dp} \hat{E}_{Dp}^{ab} \quad (3.87)$$

$$\hat{E}_{pq,kl} = \hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps} \quad (3.88)$$

$$[\hat{E}_{pq}, \hat{E}_{rs}] = \delta_{qr} \hat{E}_{ps} - \delta_{ps} \hat{E}_{rq} \quad (3.89)$$

Gives,

$$\eta_{ijp}^{kl\gamma} = \langle 0 | \frac{1}{2} (\hat{E}_{mk,nl} + \epsilon \hat{E}_{ml,nk}) \frac{1}{2} (\hat{E}_{ir,js} + p \hat{E}_{is,jr}) | 0 \rangle (T_{rsq}^{gh} T_{ef}^{mn\epsilon} (\delta_{ag}\delta_{bh} + \delta_{ah}\delta_{bg}) (\delta_{ce}\delta_{df} + \delta_{cf}\delta_{de})) \quad (3.90)$$

$$\eta_{ijp}^{kl\gamma} = \langle 0 | \hat{E}_{kle}^{mn} \hat{E}_{rsp}^{ij} | 0 \rangle T_{rsq}^{gh} T_{ef}^{mn\epsilon} (\delta_{ag}\delta_{bh} + \delta_{ah}\delta_{bg}) (\delta_{ce}\delta_{df} + \delta_{cf}\delta_{de}) \quad (3.91)$$

$$= \langle 0 | \hat{E}_{kle}^{mn} \hat{E}_{rsp}^{ij} | 0 \rangle T_{rsq}^{ab} T_{ab}^{mn\epsilon} \quad (3.92)$$

$$= \langle 0 | \hat{E}_{klq}^{mn} \hat{E}_{rsp}^{ij} | 0 \rangle T_{rsq}^{ab} T_{ab}^{mnq} \quad (3.93)$$

$$\eta_{ij}^{klp} = \langle 0 | \hat{E}_{klp}^{mn} \hat{E}_{rsp}^{ij} | 0 \rangle T_{rsp}^{ab} T_{ab}^{mnp} \quad (3.94)$$

This is a higher order density matrix that is costly to compute. To compute the density matrix efficiently we can insert a resolution of the identity as a sum over the complete space of orbital products (Something better), that is equivalent to the factorisation of the 4th order reduced density matrix into a product of two 2nd order transition density matrices.

$$\eta_{ij}^{klp} = \langle 0 | \hat{E}_{klp}^{mn} | I \rangle \langle I | \hat{E}_{rsp}^{ij} | 0 \rangle \frac{1}{2} T_{rsp}^{ab} T_{ab}^{mnp} \quad (3.95)$$

$$(3.96)$$

This can be re-expressed using orthogonal pairs, where the transformation between orthogonal and elementary generators is defined through

$$S_{ij,kl}^{(p)} = \langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 0 \rangle \quad (3.97)$$

$$\langle 0 | \hat{E}_{ab}^{ijp} \hat{E}_{klq}^{cd} | 0 \rangle = \frac{1}{2} \delta_{pq} (\delta_{ac}\delta_{bd} + p \delta_{ad}\delta_{bc}) S_{ij,kl}^{(p)} \quad (3.98)$$

$$\hat{E}_{Dp}^{ab} = \sum_{i \geq j} \hat{E}_{ijp}^{ab} Z_{D,ij}^{(p)} \quad (3.99)$$

$$T_{ab}^{ijp} = \sum_D Z_{ij,D}^{(p)} T_{ab}^{Dp} \quad (3.100)$$

The form of the matrix  $\mathbf{Z}^{(p)}$  used to transform between orthogonal and elementary generators is that given in (3.1.3)

The  $\eta$  matrix can thus be re-expressed using the canonical excitation operators.

$$\eta_E^{Dp} = \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.101)$$

This defines the  $\eta$  matrix used in the MRLPF+1 theory (efficient construction of the matrix is discussed later).

### 3.2.2 $\eta$ in the Limiting Cases

**Single Reference case** For the single reference case there will only be one internal reference configuration, hence the sum over I is limited to one configuration.

$$\eta_E^{Dp} = \frac{1}{2} \langle 0 | \hat{E}_{Ep}^F | 0 \rangle \langle 0 | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.102)$$

The transition density matrices can be shown to reduce to the identity matrix

$$\langle 0 | \hat{E}_D^{Gp} | 0 \rangle = Z_{G,kl}^{(p)} Z_{D,ij}^{(p)} \langle 0 | \hat{E}_{ik,jl} | 0 \rangle \quad (3.103)$$

$$= \delta_{DG} \quad (3.104)$$

$$\eta_E^{Dp} = \delta_{DG} \delta_{EF} \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.105)$$

$$= \frac{1}{2} T_{Fp}^{ab} T_{ab}^{Ep} \quad (3.106)$$

This is the required norm for the single reference theory and is the same as the  $\eta$  matrix used in the single reference LPF+1 theory.

In the special case when there are only 2 electrons there will only be 1 non-redundant pair, hence  $\eta$  is a 1-by-1 matrix.

$$\eta_D^{Dp} = \frac{1}{2} T_{Dp}^{ab} T_{ab}^{Dp} \quad (3.107)$$

Also the correct norm of the special case.

#### The multireference case 2 electrons (only 1 non-redundant pair)

As noted, for 2 electrons case there will only ever be 1 non-redundant excitation operator, therefore the U matrix will be only 1 element in size.

$$\eta_E^{Dp} = \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.108)$$

$$\eta_D^{Dp} = \langle 0 | I \rangle \langle I | 0 \rangle \frac{1}{2} T_{Dp}^{ab} T_{ab}^{Dp} \quad (3.109)$$

$$= c_{0I} c_{0I} \frac{1}{2} T_{Dp}^{ab} T_{ab}^{Dp} \quad (3.110)$$

This gives the norm required if the square of the reference coefficients sums to 1, i.e.  $\sum_I c_{0I}^2 = 1$ , (as it does in the single reference case or with any reference normalised to 1). If this condition is satisfied then the  $\eta$  matrix introduces the pair norm.

$$\eta = (T_{ab}^{Dp})^2 \quad (3.111)$$

which is the norm. This is also shown to give the norm for the 2-reference, 2-electron model system below (in the next subsection).

### The general multireference case of more than 2 electrons

$$\eta_E^{Dp} = \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.112)$$

$$\eta_E^{Dp} = \langle 0 | I_E^F \rangle \langle I | 0_G^D \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.113)$$

Or expressed slightly differently,

$$\eta_E^{Dp} = \langle 0 | I_E^F \rangle \langle I_D^G | 0 \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.114)$$

In this form it is easy to see that when only 1 reference function is used, or in the multireference case of 2 electrons, then the trace of the  $\eta$  matrix gives the norm. However it is also clear to see that for a general multireference case of more than 2 electrons then the matrix products scale the norm by some factor.

### 3.2.3 The Transformation Matrix

Like the single reference LPF theories, the transformation matrix is defined as the identity matrix plus the  $\eta$  matrix found above. The addition of the identity matrix ensures the correct behaviour when considering the binomial expansion of the powered  $\mathbf{U}$  matrix that is used to transform the amplitudes.

$$\mathbf{U} = \mathbf{1} + \eta$$

$$U_E^{D(p)} = \delta_E^D + \eta_E^{D(p)} \quad (3.115)$$

$${}_q T_{ab}^{Dp} = \left( \mathbf{U}^{-\frac{q}{2}} \right)_E^D T_{ab}^{Ep} \quad (3.116)$$

This matrix is required to give the norm of the wavefunction, local to a particular electron pair.

A look at the limiting cases can show the behaviour of the  $\eta$  matrix, is that which is required.

$$\begin{aligned} \langle \psi_{ijp}^{ab} | \psi_{klq}^{cd} \rangle &= \frac{1}{2} \delta_{pq} (\delta_{ac} \delta_{bd} + \delta_{ad} \delta_{bc}) \langle \hat{E}_{ik,jl} + p \hat{E}_{il,jk} \rangle \\ &= \frac{1}{2} \delta_{pq} (\delta_{ac} \delta_{bd} + \delta_{ad} \delta_{bc}) S_{ij,kl}^{(p)} \end{aligned} \quad (3.117)$$

### 3.2.4 Diagonals of $\mathbf{U}$

The transformation of the amplitudes does not give the correct normalisation for 2-electrons, it should include  $\langle 0|\hat{c}^\dagger\hat{c}|0\rangle$  in the normalisation, as this is contained in the MRCI denominator,

$$\langle 0|\hat{C}^\dagger\hat{C}|0\rangle + \langle 0|\hat{T}^\dagger\hat{T}|0\rangle \quad (3.118)$$

Therefore  $\langle 0|\hat{C}^\dagger\hat{C}|0\rangle$  must be introduced via the transformation. The diagonal elements of the  $\mathbf{U}$  matrix must therefore contain  $\mathbf{c}^\dagger\mathbf{c}$ , the dot product of the internals coefficient vectors.

$${}_qT_{ab}^{Dp} = \left(\mathbf{U}^{-\frac{q}{2}}\right)_E^{D(p)} T_{ab}^{Ep} \quad (3.119)$$

$$\mathbf{U}_E^{D(p)} = \delta_E^D + \eta_E^{D(p)} \quad (3.120)$$

$$\eta_E^{D(p)} = \frac{1}{2} \langle 0|\hat{E}_{Ep}^F|I\rangle \langle I|\hat{E}_{Gp}^D|0\rangle \frac{1}{2} T_{cd}^{D(p)} T_{E(p)}^{cd} \quad (3.121)$$

This gives the same normalisation as MRCI in the limiting 2-electron case. The terms involving the transformed amplitudes are now correct for the 2-electron MR case.

### 3.2.5 Positivity of the Transformation Matrix

The MRLPF+1D method involves introducing a local linked denominator for each amplitude via a matrix transformation of that amplitude, where the transformation matrix is the  $\mathbf{U}$  matrix defined above, powered to a either -1 or  $-\frac{1}{2}$ . In order for the inverse of a matrix to be guaranteed to exist, its determinant must not be 0 and its eigenvalues must all be positive. If the matrix is positive-definite then the matrix inverse exists, it is therefore very important that the  $\mathbf{U}$  matrix is positive-definite.

In order to look at the positive definite-ness of  $\mathbf{U}$ , the component matrices must be analysed.

$$\mathbf{U} = \mathbf{c}^\dagger\mathbf{c} \delta + \eta \quad (3.122)$$

$$U_{DE}^{(p)} = (\mathbf{c}^\dagger\mathbf{c} \delta_E^D) + \langle 0|\hat{E}_{Ep}^F|I\rangle \langle I|\hat{E}_{Gp}^D|0\rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.123)$$

$$U_{DE}^{(p)} = (\mathbf{c}^\dagger \mathbf{c} \delta_E^D) + \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | 0_G^D \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.124)$$

This shows that all elements of the matrices  $\langle 0 | \hat{E}_{Ep}^F | I \rangle$  and  $\langle I | \hat{E}_{Gp}^D | 0 \rangle$  are defined as inner products of a set of vectors. It follows that the  $\eta$  matrix is a Gramian matrix and is therefore positive-semi-definite, meaning the lowest possible eigenvalue is 0 [108]. The addition of the unit matrix will shift the eigenvalues  $\mathbf{U}$  when compared to  $\eta$ .

The dot product  $\mathbf{c}^\dagger \mathbf{c}$  is necessarily positive as it is the inner product of the two vectors. Addition of this to the diagonals of  $\Delta$  shifts the eigenvalues of  $\Delta$  by the scalar value of the dot product. In this work, intermediate normalisation is used, meaning the dot product of the internal coefficient vectors is greater than 1, therefore shifting the eigenvalues to be positive. This results in the lowest possible eigenvalue now being positive and hence  $\mathbf{U}$  is positive definite, meaning it is non-singular and negative matrix powers can be found.

The removal of linear dependencies either via a singular value decomposition (as described above), or by simply deleting electron pairs until there is no redundancy, is vital in ensuring that the  $\eta$  matrix is positive-definite, and that the matrix inverse exists. Without this property the determinant of the matrix is 0 and the method collapses because there is no denominator introduced and no binomial series approximation to the 3rd order VCC infinite series. In fact, the  $\eta$  matrix itself, which is the Gram matrix of a set of vectors is positive definite if the set of vectors are linearly independent. Hence,  $\eta$  and  $\mathbf{U}$  are positive-definite.

### 3.2.6 Complexity of Making $\mathbf{U}$

The difficulty in making  $\mathbf{U}$  lies in the difficulty of making  $\eta$ .

$$\mathbf{U} = (\mathbf{c}^\dagger \mathbf{c} \delta) + \eta \quad (3.125)$$

$$U_{DE}^{(p)} = (\mathbf{c}^\dagger \mathbf{c} \delta_E^D) + \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{Gp}^{ab} T_{ab}^{Fp} \quad (3.126)$$

$$\eta_E^{D(p)} = \mathbf{Z}_{F,mn}^{(p)} \mathbf{Z}_{E,kl}^{(p)} \left( \langle 0 | \hat{E}_{klp}^{mn} | I \rangle \right) \mathbf{Z}_{D,ij}^{(p)} \mathbf{Z}_{G,rs}^{(p)} \left( \langle I | \hat{E}_{rsp}^{ij} | 0 \rangle \right) \frac{1}{2} T_G^{ab} T_{ab}^F \quad (3.127)$$

$$\eta_E^{D(p)} = \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \Delta_F^{G(p)} \quad (3.128)$$

$$\eta_E^{D(p)} = \langle 0 | \hat{E}_{Ep}^F | I \rangle Y_{DF}^{(I)(p)} \quad (3.129)$$

$$\eta_E^{D(p)} = N_{FE}^{(I)(p)} Y_{DF}^{(I)(p)} \quad (3.130)$$

$$\eta_E^{D(p)} = M_{EF}^{(I)(p)} Y_{DF}^{(I)(p)} \quad (3.131)$$

Where the intermediates and their computational scaling are given,

$$M_{DG}^{(I)(p)} = \langle I | \hat{E}_{Gp}^D | 0 \rangle \quad (3.132)$$

$$N_{FE}^{(I)(p)} = \langle 0 | \hat{E}_{Ep}^F | I \rangle \quad (3.133)$$

$$\Delta_{GF}^{(p)} = \frac{1}{2} \sum_{ab} T_{Gp}^{ab} T_{ab}^{Fp} \quad \dots \mathcal{O}(n_p^2 \cdot v^2) \quad (3.134)$$

$$Y_{DF}^{(I)(p)} = M_{DG}^{(I)(p)} \Delta_{GF}^{(p)} \quad \dots \mathcal{O}(n_p^3 \cdot n_I) \quad (3.135)$$

$$\eta_E^{D(p)} = N_{FE}^{(I)(p)} Y_{DF}^{(I)(p)} \quad \dots \mathcal{O}(n_p^3 \cdot n_I) \quad (3.136)$$

Where  $n_p$  is the number of orthogonal pairs,  $n_I$  the number of internal configurations and  $v$ , the number of virtual orbitals. The intermediate  $M^{(I)(p)}$  (and by extensions its Hermitian conjugate  $N^{(I)(p)}$ ) are evaluated efficiently via the method of Knowles and Werner[58, 109]. The most demanding step in the construction of  $\mathbf{U}$  is no more demanding than steps involved in the construction of MRCI objects. Evaluating the 4th-order density matrix as a product of two 2nd-order transition density matrices is done efficiently using the method of Knowles and Werner in the existing icMRCI implementation in MOLPRO. This method is efficient as the formulism allows relatively small objects to be stored in high-speed memory, with the computational bottleneck being matrix multiplication, which on modern machines is done very efficiently. This method is considered quick enough to be used to recalculate objects each time they are needed instead of writing the large amount of data to a filesystem which furthers slows the process. The evaluation of the transition density matrices is an efficient process already carried out in the icMRCI theory, and their recalculation for inclusion in the  $\eta$  matrix causes no steep increase in computational cost.

### Matrix Inversion

The simplest matrix inversion routines which use Gauss-Jordan elimination scale as the cube of the side length, with more optimised routines reducing the scaling. The inversion of the  $\mathbf{U}$  matrix therefore has a maximum cost of the cube of the number of non-redundant pairs,  $\mathcal{O}(n_p^3)$ . However, the  $\mathbf{U}$  is blocked according to the irreducible representations of the symmetry point group, which significantly reduces the cost of inversion.

### 3.2.7 Transformation Model

The  $\eta$  matrix is evaluated here for a general 2-electron, 2-reference model, in order to show dependence on reference coefficients and pair amplitudes. The  $\eta$  matrix is then evaluated for the specific model case, using the reference coefficients and amplitudes previously defined.

#### General Model System

##### Elements of $S^{(p)}$

$$S_{ij,kl}^{(p)} = \langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 0 \rangle \quad (3.137)$$

$$\begin{aligned} S_{11,11}^{(1)} &= \langle 0 | \hat{E}_{11} \hat{E}_{11} | 0 \rangle - \delta_{11} \langle 0 | \hat{E}_{11} | 0 \rangle + p \langle 0 | \hat{E}_{11} \hat{E}_{11} | 0 \rangle - \delta_{11} p \langle 0 | \hat{E}_{11} | 0 \rangle \\ &= \langle 0 | \hat{E}_{11} \hat{E}_{11} | 1^2 \rangle c_{01} - \delta_{11} \langle 0 | \hat{E}_{11} | 1^2 \rangle c_{01} + \langle 0 | \hat{E}_{11} \hat{E}_{11} | 2^2 \rangle c_{02} \\ &\quad - \delta_{11} \langle 0 | \hat{E}_{11} | 2^2 \rangle c_{02} + p \langle 0 | \hat{E}_{11} \hat{E}_{11} | 1^2 \rangle c_{01} - \delta_{11} p \langle 0 | \hat{E}_{11} | 1^2 \rangle c_{01} \\ &\quad + p \langle 0 | \hat{E}_{11} \hat{E}_{11} | 2^2 \rangle c_{02} - \delta_{11} p \langle 0 | \hat{E}_{11} | 2^2 \rangle c_{02} \\ &= 4c_{01}^2 - 2c_{01}^2 + 4pc_{01}^2 - 2pc_{01}^2 = 4c_{01}^2 \end{aligned} \quad (3.138)$$

$$\begin{aligned} S_{22,22}^{(1)} &= \langle 0 | \hat{E}_{22} \hat{E}_{22} | 0 \rangle - \delta_{22} \langle 0 | \hat{E}_{22} | 0 \rangle + p \langle 0 | \hat{E}_{22} \hat{E}_{22} | 0 \rangle - \delta_{22} p \langle 0 | \hat{E}_{22} | 0 \rangle \\ &= 4c_{02}^2 - 2c_{02}^2 + 4pc_{02}^2 - 2pc_{02}^2 = 4c_{02}^2 \end{aligned} \quad (3.139)$$

$$\begin{aligned} S_{11,22}^{(1)} &= \langle 0 | \hat{E}_{12} \hat{E}_{12} | 0 \rangle + p \langle 0 | \hat{E}_{12} \hat{E}_{12} | 0 \rangle \\ &= \langle 0 | 1^2 \rangle c_{02} + p \langle 0 | 1^2 \rangle c_{02} \\ &= 2c_{01}c_{02} + 2pc_{01}c_{02} = 4c_{01}c_{02} \end{aligned} \quad (3.140)$$

$$S_{22,11}^{(1)} = 4c_{01}c_{02} \quad (3.141)$$

Giving the matrix:

$$S = \begin{pmatrix} 4c_{01}^2 & 4c_{01}c_{02} \\ 4c_{01}c_{02} & 4c_{02}^2 \end{pmatrix}$$

The S matrix has eigenpairs,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{c_{02}}{c_{01}} \\ 1 \end{pmatrix} \quad \text{and} \quad \left( 4(c_{01}^2 + c_{02}^2) \right) \begin{pmatrix} \frac{c_{01}}{c_{02}} \\ 1 \end{pmatrix} \quad (3.142)$$

This highlights the redundancies because one eigenvalue is 0.

Following the method to remove redundancies defined in a previous section, the transformation matrix to transform between orthogonal and non-orthogonal representations

$$\mathbf{Z}^{(p)} = \mathbf{V}_S^{-\frac{1}{2}} = \begin{pmatrix} \frac{c_{01}}{4(c_{01}^2 + c_{02}^2)} & \frac{c_{02}}{4(c_{01}^2 + c_{02}^2)} \\ 0 & 0 \end{pmatrix}$$

$$T_{ab}^{Dp} = \begin{pmatrix} 2\sqrt{c_{01}^2 + c_{02}^2} (c_{01}T_{ab}^{11} + c_{02}T_{ab}^{22}) \\ 0 \end{pmatrix}$$

The  $\eta$  matrix can be formed using the previously defined intermediates (see 3.2.6).

$$\Delta = \begin{pmatrix} 4(c_{01}^2 + c_{02}^2) \sum_{ab} (c_{01}T_{ab}^{11} + c_{02}T_{ab}^{22})^2 \\ 0 \end{pmatrix}$$

### Transition density matrix elements

$$\langle 0 | \hat{E}_{ik,jl} + p\hat{E}_{il,jk} | I \rangle \quad (3.143)$$

$$\langle 0 | \hat{E}_{ik,jl} + p\hat{E}_{il,jk} | 1^2 \rangle = \begin{pmatrix} 4c_{01} & 0 \\ 4c_{02} & 0 \end{pmatrix}$$

$$\langle 0 | \hat{E}_{ik,jl} + p\hat{E}_{il,jk} | 2^2 \rangle = \begin{pmatrix} 0 & 4c_{01} \\ 0 & 4c_{02} \end{pmatrix}$$

### Orthogonalisation of transition density matrices

$$M_{DE}^I = \mathbf{Z}^{(p)}_{E,kl} \mathbf{Z}^{(p)}_{D,ij} \langle 0 | \hat{E}_{ik,jl} + p\hat{E}_{il,jk} | I \rangle \quad (3.144)$$

$$M^1 = \begin{pmatrix} c_{01} \end{pmatrix}$$

$$M^2 = \begin{pmatrix} c_{02} \end{pmatrix}$$

### Product of orthogonal transition density matrix and $\Delta$

$$Y_{DF}^I = M_{DG}^I \Delta_{GF} \quad (3.145)$$

$$Y^1 = \left( c_{01} \Delta_1 \right)$$

$$Y^2 = \left( c_{02} \Delta_1 \right)$$

**Multiplication of orthogonal transition density matrix and Y**

$$\eta_{DE}^I = M_{EF}^I Y_{DF}^I \quad (3.146)$$

$$\eta^1 = \left( 4 c_{01}^2 \Delta_{11} \right)$$

$$\eta^2 = \left( 4 c_{02}^2 \Delta_{11} \right)$$

**Sum over internal configurations to form complete  $\eta$  matrix**

$$\eta = \left( 4 (c_{01}^2 + c_{02}^2) \Delta_{11} \right)$$

This is the same as the pair norm.

$$\langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle = 4 (c_{01}^2 + c_{02}^2) \Delta_{11} \quad (3.147)$$

### Specific Model System

Following the evaluation of the model system using general amplitudes and coefficients (above), the transformation is evaluation using the coefficients of the specific model system (see Appendix).

$$S = \begin{pmatrix} 4 \cos^2 \theta & 4 \cos \theta \sin \theta \\ 4 \cos \theta \sin \theta & 4 \sin^2 \theta \end{pmatrix}$$

Following the singular value decomposition (SVD)

$$T_{ab}^{Dp} = \left( 2 \sqrt{c_{01}^2 + c_{02}^2} (c_{01} T_{111}^{ab} + c_{02} T_{221}^{ab}) \right)$$

$$T_{ab}^{Dp} = \left( 2 \sqrt{\cos^2 \theta + \sin^2 \theta} (\cos^2 \theta \mu + \sin^2 \theta \mu) = 2\mu \right)$$

$$\Delta = \left( 4 (\cos^2 \theta + \sin^2 \theta) \sum_{ab} (\cos^2 \theta \mu + \sin^2 \theta \mu)^2 \right)$$

### Transition density matrix elements

$$\langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | I \rangle \quad (3.148)$$

$$\langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 1^2 \rangle = \begin{pmatrix} 4 \cos \theta & 0 \\ 4 \sin \theta & 0 \end{pmatrix}$$

$$\langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 2^2 \rangle = \begin{pmatrix} 0 & 4 \cos \theta \\ 0 & 4 \sin \theta \end{pmatrix}$$

### Orthogonalisation of transition density matrices

$$M_{DE}^I = \mathbf{Z}^{(p)}_{E,kl} \mathbf{Z}^{(p)}_{D,ij} \langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | I \rangle \quad (3.149)$$

$$M^1 = \begin{pmatrix} \cos \theta \end{pmatrix}$$

$$M^2 = \begin{pmatrix} \sin \theta \end{pmatrix}$$

### Product of orthogonal transition density matrices and $\Delta$

$$Y_{DF}^I = M_{DG}^I \Delta_{GF} \quad (3.150)$$

$$Y^1 = \left( 4 \mu^2 \cos \theta \right)$$

$$Y^2 = \left( 4 \mu^2 \sin \theta \right)$$

### Multiplication of tdm(orth).Y

$$\eta_{DE}^I = M_{EF}^I Y_{DF}^I \quad (3.151)$$

$$\eta^1 = \left( 4 \mu^2 \cos^2 \theta \right)$$

$$\eta^2 = \left( 4 \mu^2 \sin^2 \theta \right)$$

### Sum over I to form complete $\eta$ matrix

$$\eta = \left( 4 \mu^2 \right)$$

This is the same as the pair norm.

$$\langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle = 4 \mu^2 \quad (3.152)$$

The Transformation matrix is of side length of the number of non-redundant pairs, for the model this results in a  $1 \times 1$  matrix.

$$U = \mathbf{c}^\dagger \mathbf{c} + \mu^2$$

Using the definitions and relations for the model system (see Appendix),

$$|\Psi\rangle = |0\rangle + \sigma |1\rangle + \mu |3^2\rangle$$

$$|0\rangle = \cos \theta |1^2\rangle + \sin \theta |2^2\rangle$$

$$\mathbf{c}_0 = (\cos \theta \quad \sin \theta)$$

$$(1 + \hat{c})|0\rangle = |0\rangle + |1\rangle$$

$$|1\rangle = \sigma(-\sin \theta |1^2\rangle + \cos \theta |2^2\rangle)$$

$$\mathbf{c}' = \sigma(-\sin \theta \quad \cos \theta)$$

$$\mathbf{c} = \mathbf{c}_0 + \mathbf{c}'$$

$$\mathbf{c} = (\cos \theta - \sigma \sin \theta \quad \sin \theta + \sigma \cos \theta)$$

$$\mathbf{c}_0^\dagger \mathbf{c}_0 = 1$$

$$\mathbf{c}_0^\dagger \mathbf{c}' = (\cos \theta + \sin \theta)(1 - \sigma)$$

$$\mathbf{c}_0^\dagger \mathbf{c} = 1$$

$$\mathbf{c}^\dagger \mathbf{c} = 1 + \sigma^2$$

$$U = 1 + \sigma^2 + \mu^2 \quad (3.153)$$

$${}_q T_{ab}^{Dp} = \left( \mathbf{U}^{-\frac{\mathbf{q}}{2}} \right)_E^D T_{ab}^{Ep} \quad (3.154)$$

$${}_q T_{ab}^1 = \frac{2\mu}{(1 + \sigma^2 + \mu^2)^{\frac{\mathbf{q}}{2}}} \quad (3.155)$$

This shows that the transformation introduces the MRCI denominator for the simple 2-electron model system. The transformed amplitudes must be embedded in the energy terms.

### Inclusion in Energy

Looking at the term:  $2\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle$

Evaluating gives:

$$2\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) | 3^2 \rangle {}_2T_{33}^1 = \frac{2c_I \langle I | (\hat{H} - E_0) | 3^2 \rangle 2\mu}{1 + \sigma^2 + 4\mu^2} \quad (3.156)$$

Now the term:  $\langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle$

The transformation of the pair amplitude introduces the same division as MRCI.

$${}_1T_{33}^1 \langle 3^2 | (\hat{H} - E_0) | 3^2 \rangle {}_1T_{33}^1 = \frac{\langle 3^2 | (\hat{H} - E_0) | 3^2 \rangle 4\mu^2}{1 + \sigma^2 + 4\mu^2} \quad (3.157)$$

Which matches the terms in MRCI for this model case, and for any 2 electron system.

One condition for the method to work, which is implied by the SVD, is that the pair amplitude vector  $\mathbf{T}$  has to be an eigenvector of the pair overlap matrix  $\mathbf{S}$ .

### 3.2.8 As an Approximation to VCC

Apart from giving the correct norm in the limiting cases, the amplitude transformation is required to generate correctly the equivalent of the -3C 3rd order VCC energy

diagram.

Looking at the binomial expansion,

$$(\mathbf{U}_{(p)}^{-1})_E^D T_{cd}^{Ep} = T_{cd}^{Dp} - \eta_E^{Dp} T_{cd}^{Ep} + \dots \quad (3.158)$$

$$= T_{cd}^{Dp} - \left( \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{ab}^{Fp} T_{ab}^{Gp} \right) T_{cd}^{Qp} + \dots \quad (3.159)$$

The term introduced looks like a multireference version of diagram -3C, similar to that seen in the analysis in the previous chapter using explicit orbital labels. As such this can be seen as a simple approximation to an MRVCC. In the single reference limit, it was shown earlier that  $\mathbf{U}$  reduces to the single reference LPF+1  $\mathbf{U}$  matrix, and the transformation defined here gives the same as the LPF+1D theory (within limits of operators used).

However, the above binomial expansion is only true when the closed form expression is of the form,

$$\mathbf{U} = (1 + \eta)^{-1} \quad (3.160)$$

To agree with MRCI for the model system,  $c^\dagger c$  is used on the diagonals of the  $\mathbf{U}$  matrix. This is not in general 1, but  $1 + Y$ , so when considering the binomial expansion, the first and second terms above get scaled by  $\frac{1}{c^\dagger c}$  and  $\frac{1}{(c^\dagger c)^2}$  respectively.

This method is still correct for the model system, but its behaviour in the general case is to be determined. More theoretical work must be done to investigate this, including a rigorous re-expression of MRVCC terms in an orthogonal basis. The focus of the work in this chapter was to develop a transformation that acts to correct MRCI, as such, the above analysis is yet to be done, but is required. Subject to this analysis, the use of  $c^\dagger c$  on the diagonals may change, as it is included to mimic MRCI (for the model system), but may cause errors in the energy. Numerical results will also aid in testing whether or not the diagonals are defined as 1 or  $c^\dagger c$ .

### 3.3 Internals Energy

The icMRCI energy can be written as follows,

$$E_{MRCI} = E_0 + \frac{\langle 0 | (1 + \hat{c} + \hat{T})^\dagger (\hat{H} - E_0) (1 + \hat{c} + \hat{T}) | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \quad (3.161)$$

where  $|\Psi\rangle = |0\rangle + \hat{c}|0\rangle + \hat{T}|0\rangle$

The  $|0\rangle$  wavefunction is the starting 0th order CASSCF wavefunction. Expanding and separating terms,

$$E_{corr} = \frac{\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} + \frac{2\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle + \langle 0 | \hat{T}^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \quad (3.162)$$

This energy functional is a basis for comparison for the MRLPF method, as it is exact for the 2-electron case. Parts of the expression look familiar to MRLPF but terms involving  $\hat{T}$  do not use the full denominator, instead a local linked denominator is introduced via a matrix transformation. This matrix transformation also acts to approximate VCC at 3rd order and capture a subset of VCC terms to infinite order. Using the transformation of the doubles amplitudes to introduce the required norm for 2-electrons, the energy equation may be re-written.

$$E_{MRLPF} = E_0 + \frac{\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} + 2\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle + \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle \quad (3.163)$$

In the MRLPFD method, the active-active excitations are not actually included and in fact  $\hat{c}$  simply allows the internal coefficients to vary, which is equivalent to the presence of internal active-active excitations. The exclusion of active-active excitations in this way is also carried out in icMRCC theory [53, 55].

$$|0\rangle = \hat{c}_0 |\Psi_0\rangle = \frac{1}{n_{act}!} \sum_{uv..z} \hat{a}^{uv..z} |\Psi_0\rangle c_{uv..z} \quad (3.164)$$

$$= c_{0I}|I\rangle \quad (3.165)$$

$$\hat{c}|0\rangle = c_I|I\rangle \quad (3.166)$$

$$\langle 0|\hat{c}|0\rangle = 0 \quad (3.167)$$

If the MRLPF method is to be exact for the 2-electron case, then a comparison with MRCI (above, top of page) must be made, as MRCI is exact for that limiting case.

### 3.3.1 Exactness for 2-Electrons

The terms involving only internal excitations ( $\langle 0|(\hat{H} - E_0)\hat{c}|0\rangle$  and  $\langle \hat{c}^\dagger(\hat{H} - E_0)\hat{c}|0\rangle$ ) have a different denominator in MRLPF than in MRCI; the MRCI denominator contains  $\langle 0|\hat{T}^\dagger\hat{T}|0\rangle$  while MRLPF does not contain this contribution.

It may seem undesirable to introduce the pair norm to a part of the energy that contains no mention of the pair excitations, but in order to be exact for the 2-electron case, the method must be the same as MRCI. Hence the internal excitations must have a denominator containing  $\langle 0|\hat{T}^\dagger\hat{T}|0\rangle$ , but, as discussed at length earlier, this term leads to size-extensivity issues so it is not desirable to include it in the denominator (this was the initial motivation for the LPF and MRLPF theories).

To repeat, this is the same issue that terms involving the doubles amplitudes faced, resulting in the development of the LPF methods, and the transformation of the pair amplitudes defined earlier. Accordingly, efforts were made here to introduce a linked denominator via the matrix transformation applied to the pair amplitudes, so that the electrons in the pair amplitudes are related to the electrons occupied in the specific internal configuration (the fact they are related shows why part of the pair norm is needed). In order to use the same transformation matrix as that used by the pair amplitudes, internal configurations must be associated to a pair of electrons in order to have indices to transform around.

The definition of the internally contracted MRCI wavefunction shows that the internal configurations are generated from the multiconfigurational reference function by deletion and creation of 2 electrons. These internal states have previously been written

here as  $|I\rangle$ .

$$c_I|I\rangle = \sum_{tuvw} c^{tuvw} \hat{E}_{tu,vw}|0\rangle \quad (3.168)$$

The definition of the internal configurations in this way immediately associates the configurations to a pair of orbital indices. These can be used to link to the powered U transformation matrix. However, the U is indexed by orthogonal pairs, not explicit orbital labels, so cannot be linked to the internal coefficients.

Several further attempts were made to associate the internal configuration with an electron pair, one such attempt is given below.

$$\mathbf{c}' = \mathbf{X}\mathbf{c} \quad (3.169)$$

$$X_{IJ} = \langle I|\hat{E}_{Qp}^P|J\rangle \left( \mathbf{U}^{-\frac{1}{2}} \right)_{PQ} \quad (3.170)$$

Using orthogonal pairs means U is 1x1 for the model case and this generates the correct denominator for the 2 electron model.

$$X_{IJ} = \langle I|J\rangle \left( \mathbf{U}^{-\frac{1}{2}} \right)_{11} \quad (3.171)$$

$$X_{IJ} = \delta_{IJ} \left( \mathbf{U}^{-\frac{1}{2}} \right)_{11} \quad (3.172)$$

$$E c_I = \langle K|\hat{H}|J\rangle c_K c_J (\mathbf{U}^{-1}) \quad (3.173)$$

This is only correct for the 2-electron model and the behaviour in the general case is unknown. This is also impractical for computational implementation, as it requires all pair annihilations and pair creations on *each* internal configuration, whose number may be  $> 10^5$  and increases rapidly with an increase in internal orbitals. Although relatively efficient construction of the X matrix may be achieved by careful accumulation of the contributions, this is still computationally demanding and wasteful.

These efforts at developing a transformation of the internal coefficients does not give anything satisfactory for the general case, hence it is appropriate to look at how the icMRCC of Evangelista & Gauss [53] accounts for the internal excitations.

$$H_{\mu\nu}^{eff} c_\nu = \langle \mu|e^{-\hat{T}} \hat{H} e^{\hat{T}}|\nu\rangle c_\nu = E c_\mu \quad (3.174)$$

Analogously, we are guided by unitary coupled cluster, and define an effective Hamiltonian.

$$\langle I|e^{(\hat{T}^\dagger - \hat{T})} \hat{H} e^{(\hat{T} - \hat{T}^\dagger)}|J\rangle = E c_I \quad (3.175)$$

Making a linear approximation of this in a non-orthogonal spin orbital basis (illustrative),

$$\langle I | i^\dagger j^\dagger lk \hat{H} m^\dagger n^\dagger po | J \rangle = E_{c_I} \quad (3.176)$$

The normalisation can be introduced via the  $\mathbf{U}$  matrix transformation,

$$\langle I | i^\dagger j^\dagger lk \hat{H} m^\dagger n^\dagger po | J \rangle_{c_J} \left( \mathbf{U}^{-\frac{1}{2}} \right)_{ij,kl} \left( \mathbf{U}^{-\frac{1}{2}} \right)_{mn,op} = E_{c_I} \quad (3.177)$$

$$\langle I' | \hat{H} | J' \rangle_{c_J} = E_{c_I} \quad (3.178)$$

$$|I'\rangle = i^\dagger j^\dagger lk |I\rangle \left( \mathbf{U}^{-\frac{1}{2}} \right)_{ij,kl} \quad (3.179)$$

This associates an internal configuration with indices  $ij$  because  $ij$  are occupied in that configuration, therefore a linked denominator should link to  $ij$ . This idea can be reformulated using orthogonal pairs, with only 1 pair existing for a 2 electrons system, the correct norm can be achieved for the model case.

$$|I'\rangle = Q^\dagger P |I\rangle \left( \mathbf{U}^{-\frac{1}{2}} \right)_{QP} \quad (3.180)$$

$$|I'\rangle = 1^\dagger 1 |I\rangle \left( \mathbf{U}^{-\frac{1}{2}} \right)_{11} \quad (3.181)$$

$$E_{c_I} = \frac{\langle I | \hat{H} | J \rangle}{1 + \langle \hat{C}^\dagger \hat{C} \rangle + \langle \hat{T}^\dagger \hat{T} \rangle} \quad (3.182)$$

For the 2 electron case the introduced denominators multiply to give the effect of  $\mathbf{U}^{-1}$ , ensuring the correct denominator, but the behaviour for the general case is unknown.

However, again this method of generating a linked denominator is computationally demanding, requiring the generation of many configurations. Therefore the above efforts to include a linked denominator for the internal coefficients either is unsatisfactory in the general case or far too computationally demanding to account for the small increase in accuracy.

The simplest method of dealing with the size-extensivity problem of the denominator is to use the ACPF or AQCC factors, making the denominator almost intensive. Using the ACPF factors is computationally extremely simple to compute, and has the benefit of giving the exact denominator for the 2 electron case. It is therefore a pragmatic choice

to deal with unlinked nature of the denominator when a actual linked denominator cannot be introduced.

$$\begin{aligned}
 E = E_0 + & \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 1 + \hat{c}^\dagger | 1 + \hat{c} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \\
 & + 2 \langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle + \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle \quad (3.183)
 \end{aligned}$$

### Evaluation For MR 2 Electron Model

The term in the energy containing only internal configurations,

$$\frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 1 + \hat{c}^\dagger | 1 + \hat{c} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \quad (3.184)$$

$$= \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{1 + \sigma^2 + \frac{2}{2} \mu^2 \langle 3^2 | 3^2 \rangle} \quad (3.185)$$

$$= \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{1 + \sigma^2 + \mu^2} \quad (3.186)$$

This ensures the method is exact for 2 electrons and 2 references. The factor of  $\frac{2}{n}$  also ensures it is correct for two identical but separate 2-electron systems, such as separated Helium atoms.

### 3.3.2 Single Reference Limit

In the above equations there is an assumption, without which the theory would not reduce exactly to the single reference LPF theory or correspond theoretically to the CI case. The assumption is that the coefficient of the reference is 1, giving  $1 + c^2$  in the denominator and on the diagonals of the U matrix. This is intermediate normalisation, and must be imposed when equation solving. Both limits are only satisfied when the constraint,  $\mathbf{c}_0^\dagger \mathbf{c} = 1$ , is imposed, giving intermediate normalisation. To impose the constraint, the Lagrange multiplier method is used.

$$\begin{aligned}
E_{MRLPF} = E_0 &+ \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 1 + \hat{c}^\dagger | 1 + \hat{c} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \\
&+ 2 \langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle + \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle
\end{aligned} \tag{3.187}$$

$$\mathcal{L} = E_{MRLPF} + \lambda (\mathbf{c}_0^\dagger \mathbf{c} - 1) \tag{3.188}$$

### 3.4 Full Energy Expression

Using intermediate normalisation and an ACPF like denominator, for reasons given in the previous sections, the full expression for the MRLPF+1 energy can now be defined.

$$\begin{aligned}
 E = E_0 + & \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 1 + \hat{c}^\dagger | 1 + \hat{c} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \\
 & + 2 \langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle + \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle \quad (3.189)
 \end{aligned}$$

Where the transformed amplitudes are given by,

$${}_q T_{ab}^{Dp} = \left( \mathbf{U}^{-\frac{q}{2}} \right)_E^{D(p)} T_{ab}^{Ep} \quad (3.190)$$

$$\mathbf{U}_E^{D(p)} = \mathbf{c}^\dagger \mathbf{c} \delta_E^D + \eta_E^{D(p)} \quad (3.191)$$

$$\eta_E^{D(p)} = \frac{1}{2} \left( \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{cd}^{D(p)} T_{E(p)}^{cd} \right) \quad (3.192)$$

#### 3.4.1 Full Energy Evaluation for the Model System

Full evaluation of the energy equation for the two electrons system, to show equivalence to MRCI for this limiting case. This is done above, in part, for the transformation but is included here for completeness.

##### Model

$$S = \begin{pmatrix} 4 \cos^2 \theta & 4 \cos \theta \sin \theta \\ 4 \cos \theta \sin \theta & 4 \sin^2 \theta \end{pmatrix}$$

Following a singular value decomposition (SVD)

$$C_{ab}^{Dp} = \left( 2 \sqrt{c_1^2 + c_2^2} (c_1 \cos \theta \mu + c_2 \sin \theta \mu) \right)$$

$$\Delta = \left( 4 (c_1 c_2 + c_2^2) \sum_{ab} (c_1 \cos \theta \mu + c_2 \sin \theta \mu)^2 \right)$$

##### tdm Elements

$$\langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | I \rangle \quad (3.193)$$

$$\begin{aligned}\langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 1^2 \rangle &= \begin{pmatrix} 4 \cos \theta & 0 \\ 4 \sin \theta & 0 \end{pmatrix} \\ \langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | 2^2 \rangle &= \begin{pmatrix} 0 & 4 \cos \theta \\ 0 & 4 \sin \theta \end{pmatrix}\end{aligned}$$

**Orthogonalisation of tdm**

$$M_{DE}^I = \mathbf{Z}^{(p)}_{E,kl} \mathbf{Z}^{(p)}_{D,ij} \langle 0 | \hat{E}_{ik,jl} + p \hat{E}_{il,jk} | I \rangle \quad (3.194)$$

$$\begin{aligned}M^1 &= \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \\ M^2 &= \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix}\end{aligned}$$

**Product of orthogonal tdm and  $\Delta$** 

$$Y_{DF}^I = M_{DG}^I \Delta_{GF} \quad (3.195)$$

$$\begin{aligned}Y^1 &= \begin{pmatrix} 4 \mu^2 \cos \theta \\ 4 \mu^2 \sin \theta \end{pmatrix} \\ Y^2 &= \begin{pmatrix} 4 \mu^2 \sin \theta \\ 4 \mu^2 \cos \theta \end{pmatrix}\end{aligned}$$

**Multiplication of tdm(orth).Y**

$$\eta_{DE}^I = M_{EF}^I Y_{DF}^I \quad (3.196)$$

$$\begin{aligned}\eta^1 &= \begin{pmatrix} 4 \mu^2 \cos^2 \theta \\ 4 \mu^2 \sin^2 \theta \end{pmatrix} \\ \eta^2 &= \begin{pmatrix} 4 \mu^2 \sin^2 \theta \\ 4 \mu^2 \cos^2 \theta \end{pmatrix}\end{aligned}$$

**Sum over I to form complete  $\eta$  matrix**

$$\eta = \begin{pmatrix} 4 \mu^2 \\ 4 \mu^2 \end{pmatrix}$$

This the same as the pair norm.

$$\langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle = 4 \mu^2 \quad (3.197)$$

**MRLPF Energy**

$$E_{MRLPF} = E_0 + \frac{\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{1 + \sigma^2 + \frac{2}{2} \mu^2}$$

$$+ \frac{2\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle + \mu \langle 3^2 | (\hat{H} - E_0) | 3^2 \rangle}{1 + \sigma^2 + \mu^2} \quad (3.198)$$

### MRCI Energy

$$E_{corr} = \frac{\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{1 + \sigma^2 + \frac{2}{2}\mu^2} + \frac{2\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle + \langle 0 | \hat{T}^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \quad (3.199)$$

$$E_{corr} = \frac{\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} + \frac{2\langle 0 | (1 + \hat{c})^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle + \mu \langle 3^2 | (\hat{H} - E_0) | 3^2 \rangle}{1 + \sigma^2 + \mu^2} \quad (3.200)$$

These are the same, meaning the MRLPF+1D method performs as expected for the model 2 electron, 2 reference case.

### 3.5 Direct CI

The solutions of the CI equation are found by minimisation of the Rayleigh quotient,

$$\begin{aligned}\rho &= \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = E \\ \frac{\partial \rho}{\partial \mathbf{c}} &= 0\end{aligned}\quad (3.201)$$

Giving the eigenvalue equation

$$\mathbf{H} \mathbf{c}^k = E_k \mathbf{c}^k \quad (3.202)$$

Conventional-CI approaches explicitly construct the full Hamiltonian matrix. However this is computationally demanding and requires the storage of the full matrix whose dimensions may approach  $10^6 \times 10^6$ .

Alternatively the result of matrix-vector products are accumulated, and only they are stored. Differentiating the CI energy as above gives an eigenvalue equation to be solved, splitting this into the different spaces gives

$$\langle \psi_{Dp}^{ab} | \hat{H} - E | \Psi \rangle = 0 \quad (3.203)$$

$$\langle \psi_S^a | \hat{H} - E | \Psi \rangle = 0 \quad (3.204)$$

$$\langle \psi_I | \hat{H} - E | \Psi \rangle = 0 \quad (3.205)$$

This defines the residual vectors,

$$\langle \psi_{Dp}^{ab} | \hat{H} - E | \Psi \rangle = R_{Dp}^{ab} = \left( \frac{1}{2} (\mathbf{G}^{Dp} + p(\mathbf{G}^{Dp})^\dagger) - E \mathbf{T}^{Dp} \right)_{ab} \quad (3.206)$$

$$\langle \psi_S^a | \hat{H} - E | \Psi \rangle = R_S^a = [\mathbf{g}^S - E \mathbf{T}^S]_a \quad (3.207)$$

$$\langle \psi_I | \hat{H} - E | \Psi \rangle = R^I = g^I - E c^I \quad (3.208)$$

The contributions  $\mathbf{G}^{Dp}$ ,  $\mathbf{g}^S$  and  $g^I$  are calculated *directly* in operator form from the integrals. These contributions to the residual vectors are accumulated throughout the program, requiring only the storage of these residual vectors rather than the full size Hamiltonian matrix. This is a form of Direct-CI.

Within MOLPRO The contributions to the residual vectors are calculated in parts throughout the CI code. The evaluation of these residual vectors, along with the updates to the amplitudes occur in the macroiterations of the direct-CI code.

The use of microiterations is designed to speed up convergence, by using relatively cheap iterations to optimise coefficients, which will ultimately result in fewer expensive macroiterations being needed. The microiterations use the Davidson subspace method [110] and are described later.

### 3.5.1 Pair Residual

As stated above, differentiation of the CI energy with respect to the amplitudes,  $T$ , yields an eigenvalue equation.

$$R_{ab}^{Dp} = \langle \phi_{ab}^{Dp} | \hat{H} - E_{CID} | \psi_{CID} \rangle = 0 \quad (3.209)$$

These equations can be solved iteratively, each iteration computing the residual  $R$ .

$$R_{ab}^{Dp} = G_{ab}^{Dp} - \epsilon_{CID} T_{ab}^{Dp} \quad (3.210)$$

$$G_{ab}^{Dp} = \langle \phi_{ab}^D | \hat{H} - E_0 | \psi_{CID} \rangle \quad (3.211)$$

$\epsilon_{CID} = E_{CID} - E_0$  is the correlation energy introduced by CID.

The MRLPF+1 pair residual is more complicated than that of MRCI due to the use of transformed amplitudes. Care must be taken to account for the change in the transformed amplitudes as a result of a change in the underlying amplitudes. The methodology to do this is based on the corrected LPF+1 residual by Robinson & Knowles [33], that showed how to differentiate the transformed amplitudes with respect to the change in the underlying  $T$  amplitude. The multireference MRLPF+1 residual makes use of this differentiation to minimise the MR-Linked Pair Functional theory with respect to the doubles amplitudes.

### Preliminary Definitions

Energy expression:

$$E = E_0 + \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 1 + \hat{c}^\dagger | 1 + \hat{c} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} + 2 \langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle + \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle \quad (3.212)$$

using definitions:

$$\hat{T} = \sum_{i \geq j} \sum_p \sum_{ab} T_{ab}^{ijp} \hat{E}_{ijp}^{ab} \quad \hat{E}_{ijp}^{ab} = \frac{1}{2}(\hat{E}_{ai,bj} + p\hat{E}_{bi,aj}) \quad (3.213)$$

$$= \sum_D \sum_p \sum_{ab} T_{ab}^{Dp} \hat{E}_{Dp}^{ab} \quad (3.214)$$

$$\hat{E}_{Dp}^{ab} = \sum_{i \geq j} \hat{E}_{ijp}^{ab} \mathbf{Z}_{D,ij}^{(p)} \quad (3.215)$$

$$T_{ab}^{ijp} = \sum_D \mathbf{Z}_{D,ij}^{(p)} T_{ab}^{Dp} \quad (3.216)$$

$${}_q T_{ab}^{Dp} = \left( \mathbf{U}^{-\frac{q}{2}} \right)_{Ep}^{Dp} T_{ab}^{Ep} \quad (3.217)$$

$$\mathbf{U}_E^{D(p)} = \mathbf{c}^\dagger \mathbf{c} \delta_E^D + \eta_E^{D(p)} \quad (3.218)$$

$$\eta_E^{D(p)} = \frac{1}{2} \left( \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{cd}^{D(p)} T_{E(p)}^{cd} \right) \quad (3.219)$$

### Finding The Residual

The energy defined above can be written in simple terms of an internals and terms involving pair excitations

$$E = E_0 + E_I + E_P \quad (3.220)$$

$$E_P = 2 \langle 0 | (1 + \hat{C}^\dagger) (\hat{H} - E_0) | \Phi_{Dp}^{ab} \rangle {}_2 T_{ab}^{Dp} + \langle 0 | {}_1 \hat{T}^\dagger (\hat{H} - E_0) | \Phi_{Dp}^{ab} \rangle {}_1 T_{ab}^{Dp} \quad (3.221)$$

$$E_I = \frac{\langle 0 | (1 + \hat{C}^\dagger) (\hat{H} - E_0) (1 + \hat{C}) | 0 \rangle}{\langle 1 + \hat{C}^\dagger | 1 + \hat{C} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \quad (3.222)$$

This separation aids in finding the target  $\frac{1}{2} \frac{\partial E}{\partial T_{ab}^{Dp}}$ .

$$\frac{1}{2} dE = \frac{1}{2} \frac{\partial E_I}{\partial T_{ab}^{Dp}} T_{ab}^{Dp} + \frac{1}{2} \frac{\partial E_P}{\partial T_{ab}^{Dp}} T_{ab}^{Dp} \quad (3.223)$$

$$\frac{1}{2} dE = R_{Dp}^{ab} T_{ab}^{Dp} \quad (3.224)$$

$$\frac{1}{2} \frac{\partial E_I}{\partial T_{ab}^{Dp}} = \Omega_{ab}^{Dp} \quad (3.225)$$

The partial derivatives can be written,

$$\frac{1}{2} \frac{\partial E_P}{\partial T_{ab}^{Dp}} = 2 \langle 0 | (1 + \hat{C}^\dagger) (\hat{H} - E_0) | \Phi_{Dp}^{ab} \rangle d({}_2T_{ab}^{Dp}) + \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) | \Phi_{Dp}^{ab} \rangle d({}_1T_{ab}^{Dp}) \quad (3.226)$$

$$\frac{1}{2} \frac{\partial E_P}{\partial T_{ab}^{Dp}} = ({}_1G_{Dp}^{ab} + {}_2G_{Dp}^{ab}) \quad (3.227)$$

$$\frac{1}{2} \frac{\partial E}{\partial T_{ab}^{Dp}} = ({}_1G_{Dp}^{ab} + {}_2G_{Dp}^{ab}) + \Omega_{Dp}^{ab} \quad (3.228)$$

$$\frac{1}{2} \frac{\partial E}{\partial T_{ab}^{Dp}} = (G_{Dp}^{ab}) + \Omega_{Dp}^{ab} \quad (3.229)$$

$$\frac{1}{2} \frac{\partial E}{\partial T_{ab}^{Dp}} = R_{Dp}^{ab} \quad (3.230)$$

$$\frac{1}{2} dE = R_{Dp}^{ab} dT_{ab}^{Dp} \quad (3.231)$$

The problem is now to find  $G_{Dp}^{ab}$ , by performing the appropriate differentiations, and with the following definitions,

$$\begin{aligned} {}_2\mathcal{V}_{Dp}^{ab} &= \langle \Phi_0 | \hat{H} | \Phi_{Dp}^{ab} \rangle \\ {}_1\mathcal{V}_{Dp}^{ab} &= \langle \Phi_0 | {}_1\hat{T}^\dagger \hat{H} | \Phi_{Dp}^{ab} \rangle \end{aligned} \quad (3.232)$$

Unlinked terms must be removed from  ${}_1\mathcal{V}_{Dp}^{ab}$ .

$${}_1\mathcal{V}_{Dp}^{ab} = \langle \Phi_0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) | \Phi_{Dp}^{ab} \rangle \quad (3.233)$$

$${}_1\mathcal{V}_{Dp}^{ab} = \langle \Phi_0 | {}_1\hat{T}^\dagger \hat{H} | \Phi_{Dp}^{ab} \rangle - \langle \Phi_0 | {}_1\hat{T}^\dagger E_0 | \Phi_{Dp}^{ab} \rangle \quad (3.234)$$

$${}_1\mathcal{V}_{Dp}^{ab} = \langle \Phi_0 | {}_1\hat{T}^\dagger \hat{H} | \Phi_{Dp}^{ab} \rangle - T_{cd}^{Er} \langle \Phi_{cd}^{Er} | \Phi_{Dp}^{ab} \rangle E_0 \quad (3.235)$$

$${}_1\mathcal{V}_{Dp}^{ab} = \langle \Phi_0 | {}_1\hat{T}^\dagger \hat{H} | \Phi_{Dp}^{ab} \rangle - T_{ab}^{Er} \Gamma_{Er}^{Dp} E_0 \quad (3.236)$$

$${}_1\mathcal{V}_{Dp}^{ab} = \langle \Phi_0 | {}_1\hat{T}^\dagger \hat{H} | \Phi_{Dp}^{ab} \rangle - T_{ab}^{Ep} \Gamma_E^{D(p)} E_0 \quad (3.237)$$

$${}_1\mathcal{V}_{Dp}^{ab} = \langle \Phi_0 | {}_1\hat{T}^\dagger \hat{H} | \Phi_{Dp}^{ab} \rangle - T_{ab}^{Dp} E_0 \quad (3.238)$$

we may write the following.

$$\frac{1}{2} \frac{\partial E}{\partial T_{ab}^{Dp}} = \sum_{q=1}^2 {}_q\mathcal{V}_{Dp}^{ab} d({}_qT_{ab}^{Dp}) + \Omega_{ab}^{Dp} \quad (3.239)$$

$$\text{i.e.} = ({}_1G_{Dp}^{ab} + {}_2G_{Dp}^{ab}) + \Omega_{ab}^{Dp} \quad (3.240)$$

$$= R_{Dp}^{ab} \quad (3.241)$$

This involves the derivative of  $\mathbf{U}$  as given.

$$U_E^{D(p)} = (c^\dagger c) \delta_E^D + \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{cd}^{D(p)} T_{E(p)}^{cd} \quad (3.242)$$

$$\begin{aligned} dU_E^{D(p)} &= d \left( \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \frac{1}{2} T_{cd}^{D(p)} T_{E(p)}^{cd} \right) \\ &= \langle 0 | \hat{E}_{Ep}^F | I \rangle \langle I | \hat{E}_{Gp}^D | 0 \rangle \left( \frac{1}{2} dT_{cd}^{G(p)} T_{F(p)}^{cd} + \frac{1}{2} T_{cd}^{G(p)} dT_{F(p)}^{cd} \right) \end{aligned} \quad (3.243)$$

The problem of how to differentiate the transformed amplitudes was solved by Robinson and Knowles, and is shown below.

$$\begin{aligned} d({}_q T_{ab}^{Dp}) &= \left( \mathbf{U}_{(p)}^{-\frac{q}{2}} d\mathbf{T} + d \left[ \mathbf{U}_{(p)}^{-\frac{q}{2}} \right] \mathbf{T} \right)_{ab}^{Dp} \\ &= \left( \mathbf{U}_{(p)}^{-\frac{q}{2}} d\mathbf{T} \right)_{ab}^{Dp} + \left( d \left[ \mathbf{U}_{(p)}^{-\frac{q}{2}} \right] \mathbf{T} \right)_{ab}^{Dp} \end{aligned} \quad (3.244)$$

Inserting this into our expression for the derivative, gives:

$$\begin{aligned} \frac{1}{2} \frac{\partial E}{\partial T_{ab}^{Dp}} &= \sum_{q=1}^2 {}_q \mathcal{V}_{Dp}^{ab} \left( \mathbf{U}_{(p)}^{-\frac{q}{2}} d\mathbf{T} + d \left( \mathbf{U}_{(p)}^{-\frac{q}{2}} \right) \mathbf{T} \right)_{ab}^{Dp} + \Omega_{ab}^{Dp} dT_{ab}^{Dp} \\ &= \sum_{q=1}^2 {}_q A_{Dp}^{ab} dT_{ab}^{Dp} + \sum_p \sum_{q=1}^2 {}_q B_D^{E(p)} d \left( \mathbf{U}_{(p)}^{-\frac{q}{2}} \right)_E^D + \Omega_{ab}^{Dp} dT_{ab}^{Dp} \\ &= \sum_{q=1}^2 {}_q G_{Dp}^{ab} dT_{ab}^{Dp} + \Omega_{ab}^{Dp} dT_{ab}^{Dp} \\ &= R_{Dp}^{ab} \end{aligned} \quad (3.245)$$

The difficult step is finding  $G$ , similar to the LPF+1 residual. The working for the first term,  $A$ , is complete, so focus must lie on the second term  $B$ . The evaluation of  $A$  is simple, but the second term,  $B$ , requires the derivative of the transformed amplitudes. Intermediates can be defined to reduce computational complexity. The method of how to find the derivative of the transformed amplitudes was derived in [11], and casts the

problem into spectral form, where  $\mathbf{X}$  is the matrix of eigenvectors of the  $\mathbf{U}$  transformation matrix, and  $\epsilon$  the associated eigenvalues. For a detailed description of this derivation, the reader is directed to [11]. The method is used here, with only different intermediates defined due to the difference in the form of the  $\mathbf{U}$  matrix.

$$\begin{aligned}
& \sum_{q=1}^2 {}_q B_D^E d(\mathbf{U}^{-\frac{q}{2}})_E^D \\
&= \sum_{q=1}^2 {}_q B_{DE}^{(p)} d(\mathbf{U}_{(p)}^{-\frac{q}{2}})_{ED} \\
&= \sum_{q=1}^2 {}_q B_{DE}^{(p)} \sum_{F,G} dU_{FG}^{(p)} \left[ \sum_{\substack{k=1, l=1 \\ k \neq l, \epsilon_k \neq \epsilon_l}}^{n_p} \frac{\epsilon_k^{-\frac{q}{2}} - \epsilon_l^{-\frac{q}{2}}}{\epsilon_k - \epsilon_l} X_{El}^{(p)} X_{Dk}^{(p)} X_{Fl}^{(p)} X_{Gk}^{(p)} - \frac{q}{2} \sum_{k=1}^{n_p} \epsilon_k^{-\frac{q}{2}-1} X_{Ek}^{(p)} X_{Ek}^{(p)} X_{Fk}^{(p)} X_{Gk}^{(p)} \right] \\
&= \sum_{q=1}^2 \sum_{i,F,G} dU_{FG}^{(p)} \left[ \sum_{\substack{k=1, l=1 \\ k \neq l, \epsilon_k \neq \epsilon_l}}^{n_p} \frac{\epsilon_k^{-\frac{q}{2}} - \epsilon_l^{-\frac{q}{2}}}{\epsilon_k - \epsilon_l} X_{Ek}^{(p)} X_{Fl}^{(p)} X_{Gk}^{(p)} {}_q C_{Dl}^{(p)} - \frac{q}{2} \sum_{k=1}^{n_p} \epsilon_k^{-\frac{q}{2}-1} X_{Dk}^{(p)} X_{Fk}^{(p)} X_{Gk}^{(p)} {}_q C_{Dk}^{(p)} \right] \\
&= \sum_{q=1}^2 \sum_{F,G} dU_{FG}^{(p)} \left[ \sum_{\substack{k=1, l=1 \\ k \neq l, \epsilon_k \neq \epsilon_l}}^{n_p} \frac{\epsilon_k^{-\frac{q}{2}} - \epsilon_l^{-\frac{q}{2}}}{\epsilon_k - \epsilon_l} X_{Fl}^{(p)} X_{Gk}^{(p)} {}_q D_{lk}^{(p)} - \frac{q}{2} \sum_{k=1}^{n_p} \epsilon_k^{-\frac{q}{2}-1} X_{Fk}^{(p)} X_{Gk}^{(p)} {}_q D_{kk}^{(p)} \right] \\
&= \sum_{q=1}^2 \sum_{F,G} dU_{FG}^{(p)} \left[ \sum_{k=1}^{n_p} X_{Gk}^{(p)} {}_q F_{kF}^{(p)} - \frac{q}{2} {}_q E_{FG}^{(p)} \right] \\
&= \sum_{q=1}^2 \sum_{F,Gm,n} dU_{FG}^{(p)} [{}_q H_{FG}^{(p)} - \frac{q}{2} {}_q E_{FG}^{(p)}] \\
&= \sum_{q=1}^2 \sum_{F,G} dU_{FG}^{(p)} {}_q I_{FG}^{(p)} \\
&= \sum_{q=1}^2 dU_G^{F(p)} {}_q I_{FG} \\
&= \frac{1}{2} \sum_{q=1}^2 \left( \langle 0 | \hat{E}_{GP}^D | I \rangle \langle I | \hat{E}_{EP}^F | 0 \rangle \left( T_{ab}^{E(p)} dT_{D(p)}^{ab} + dT_{ab}^{E(p)} T_{D(p)}^{ab} \right) \right) {}_q I_{FG} \\
&= \frac{1}{2} \sum_{q=1}^2 \left( N_{DG}^{I(p)} M_{FE}^{I(p)} \left( T_{ab}^{E(p)} dT_{D(p)}^{ab} + dT_{ab}^{E(p)} T_{D(p)}^{ab} \right) \right) {}_q I_{FG} \\
&= \frac{1}{2} \sum_{q=1}^2 N_{DG}^{I(p)} O_{EG}^{I(p)} \left( T_{ab}^{E(p)} dT_{D(p)}^{ab} + dT_{ab}^{E(p)} T_{D(p)}^{ab} \right) \\
&= \frac{1}{2} \sum_{q=1}^2 M_{GD}^{I(p)} O_{EG}^{I(p)} \left( T_{ab}^{E(p)} dT_{D(p)}^{ab} + dT_{ab}^{E(p)} T_{D(p)}^{ab} \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \sum_{q=1}^2 q P_{DE} \left( T_{ab}^{E(p)} dT_{D(p)}^{ab} + dT_{ab}^{E(p)} T_{D(p)}^{ab} \right) \\
&= \frac{1}{2} \sum_{q=1}^2 \left( P + (P)^T \right)_E^D T_{D(p)}^{ab} dT_{ab}^{E(p)} \\
&= \frac{1}{2} \sum_{q=1}^2 q Q_{DE} T_{D(p)}^{ab} dT_{ab}^{E(p)} \\
&= \frac{1}{2} \sum_{q=1}^2 q K_{E(p)}^{ab} dT_{ab}^{E(p)}
\end{aligned} \tag{3.246}$$

Where  $\mathbf{X}$  is the matrix containing the eigenvectors of  $\mathbf{U}$  as its columns, and  $\epsilon$  is the diagonal matrix of eigenvalues corresponding to  $\mathbf{X}$ .

Combining this with the first component of  $G$ , we get the following.

$$\frac{\partial E}{\partial T_{ab}^{Dp}} = \sum_{q=1}^2 \left[ q A_{Dp}^{ab} + \frac{1}{2} q K_{Dp}^{ab} \right] + \Omega_{ab}^{Dp} \tag{3.247}$$

From this, we can read off the new, correct contribution to the residual.

$$G_{Dp}^{ab} = \sum_{q=1}^2 \left[ q A_{Dp}^{ab} + \frac{1}{2} q K_{Dp}^{ab} \right] \tag{3.248}$$

The definitions of the intermediates quantities used are as follows.

$$q A_{Dp}^{ab} = \left( \mathbf{U}_{(p)}^{-\frac{q}{2}} \right)_D^E q \mathcal{V}_{Ep}^{ab} \tag{3.249}$$

$$q B_E^{D(p)} = \frac{1}{2} q \mathcal{V}_{Dp}^{ab} T_{ab}^{E(p)} \quad (\text{not symmetric}) \tag{3.250}$$

$$q C_{DE}^{(p)} = \sum_{k=1}^{n_p} q B_{Dk}^{(p)} X_{kE}^{(p)} \tag{3.251}$$

$$q D_{DE}^{(P)} = \sum_{k=1}^{n_p} q C_{kD}^{(p)} X_{kE}^{(p)} \tag{3.252}$$

$$q E_{DE}^{(p)} = \sum_{k=1}^{n_p} \epsilon_k^{-\frac{q}{2}-1} X_{Dk}^{(p)} X_{Ek}^{(p)} q D_{kk}^{(p)} \tag{3.253}$$

$$q F_{DE}^{(P)} = \sum_{k=1, k \neq i, \epsilon_k \neq \epsilon_i}^{n_p} \frac{\epsilon_i^{-\frac{q}{2}} - \epsilon_k^{-\frac{q}{2}}}{\epsilon_i - \epsilon_k} X_{Ek}^{(p)} q D_{kD}^{(p)} \tag{3.254}$$

$$q H_{DE}^{(p)} = \sum_{k=1}^{n_p} X_{Ek}^{(p)} q F_{kD}^{(p)} \tag{3.255}$$

$${}_q I_{DE}^{(p)} = {}_q H_{DE}^{(p)} - \frac{q}{2} {}_q E_{DE}^{(p)} \quad (3.256)$$

$${}_q M_{DE}^{I(p)} = \langle I | \hat{E}_E^D | 0 \rangle \quad (3.257)$$

$${}_q N_{DE}^{I(p)} = \langle 0 | \hat{E}_E^D | I \rangle \quad (3.258)$$

$${}_q O_{DE(p)}^I = M_{DF}^{I(p)} {}_q I_{FE}^{(p)} \quad (3.259)$$

$${}_q P_{DE(p)} = M_{FE}^{I(p)} O_{GE(p)}^I \quad (3.260)$$

$${}_q Q_{DE(p)} = P_{DE} + P_{ED} \quad (3.261)$$

$${}_q K_{E(p)}^{ab} = Q_{DE(p)} T_{Dp}^{ab} \quad (3.262)$$

When there are four indices, each index belongs to the set of electrons or holes (for i,j,k,... and a,b,c,... respectively). When there are two indices, each index belongs to the set of electron or hole pairs (for i,j,k,... and a,b,c,... respectively).

It should be noted that the only difference between this residual and that of the original LPF+1 residual of Robinson & Knowles lies in the definition of  $\mathbf{U}$  and hence small changes in the intermediates needed.

The final contribution  $\Omega$  for the residual  $\mathbf{R}$  must be found.

$$E_I = \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\langle 1 + \hat{c}^\dagger | 1 + \hat{c} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} \quad (3.263)$$

$$\Omega_{ab}^{Dp} = \frac{1}{2} \frac{\partial E_I}{\partial T_{ab}^{Dp}} = - \frac{\langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) (1 + \hat{c}) | 0 \rangle}{\left( \langle 1 + \hat{c}^\dagger | 1 + \hat{c} \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle \right)^2} \frac{4}{n} T_{ab}^{Dp} \quad (3.264)$$

### Finding The Update

The pair amplitudes,  $T_{ab}^{Dp}$ , are updated using the residuals obtained above, and simple perturbation theory.

$$R_{ab}^{Dp} = G_{ab}^{Dp} + \Omega_{ab}^{Dp} \quad (3.265)$$

$$\Delta T_{ab}^{Dp} = \frac{-R_{ab}^{Dp}}{\langle \Phi_{Dp}^{ab} | \hat{H} - E_m | \Phi_{Dp}^{ab} \rangle} \quad (3.266)$$

### 3.5.2 Internals Residual

The derivative of the energy with respect to the change in the internal coefficients is required.

$$\begin{aligned} \frac{\partial E}{\partial c_I} = & \frac{2 \left( \langle 0 | \hat{H} | I \rangle + \langle I | \hat{H} | J \rangle c_J \right) \left( 1 + c_J^2 + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle \right)}{\left( 1 + c_J^2 + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle \right)^2} \\ & - \frac{\left( 4 \langle 0 | \hat{H} | J \rangle c_J + 2 c_K \langle K | \hat{H} | J \rangle c_J \right) c_J}{\left( 1 + c_J^2 + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle \right)^2} \\ & + \frac{\partial 2 c_I \langle I | \hat{H} | \hat{T} | 0 \rangle}{\partial c_I} + \frac{\partial \langle \hat{T} \hat{H} | \hat{T} \rangle}{\partial c_I} \end{aligned} \quad (3.267)$$

The derivatives of the terms involving transformed amplitudes now must be found. The transformed amplitudes themselves have dependence on  $c_I$ , due to the  $c^\dagger c$  on the diagonal of the  $\mathbf{U}$  matrix. The coefficients used in making the transformed amplitudes are fixed, thus cannot be changed when considering the change in the energy with respect to the change in internal coefficients. As a result, a linear approximation of the derivative must be made.

The  $\mathbf{U}$  matrix must therefore be differentiated with respect to the change of the internal coefficients. For the general case the matrix differentiation technique of Robinson & Knowles, previously used for differentiation with respect to the change in the  $T_{ij}^{ab}$  doubles amplitudes, can be used. However, the matrix  $d\mathbf{U}$  can be defined as the derivative of the  $\mathbf{U}$  matrix with respect to a change in the internal coefficients, which only affects the diagonal elements of the  $\mathbf{U}$  matrix. This means that  $\mathbf{U}$  and  $d\mathbf{U}$  commute, so the basic chain rule of differentiation can be used, without using the complicated method of Robinson & Knowles.

$$d(\mathbf{U}^x) = x d\mathbf{U} \mathbf{U}^{x-1} \quad (3.268)$$

$$d\left(\mathbf{U}^{-\frac{1}{2}}\right) = -\frac{1}{2} d\mathbf{U} \mathbf{U}^{-\frac{3}{2}} \quad (3.269)$$

$$d(\mathbf{U}^{-1}) = -d\mathbf{U} \mathbf{U}^{-2} \quad (3.270)$$

$$d\mathbf{U} = d(\mathbf{U})_{PQ} = d(\mathbf{U}_{PQ}) = \frac{\partial \mathbf{U}_{PQ}}{\partial c_I} = \delta_Q^P 2 c_I \quad (3.271)$$

Using these;

$$\frac{\partial \left( c_I \langle I | \hat{H} | \hat{T} | 0 \rangle \right)}{\partial c_I} \quad (3.272)$$

$$\begin{aligned}
&= \langle I | \hat{H} | \hat{T} | 0 \rangle + c_I \langle I | \hat{H} | \phi_{Qp}^{ab} \rangle \frac{\partial 2T_{ab}^{Qp}}{\partial c_I} \\
&= \left( -\langle I | \hat{H} | \phi_{Qp}^{ab} \rangle (\mathbf{U}^{-1})_P^Q (\mathbf{dU})_R^P (\mathbf{U}^{-1})_S^R T_{ab}^S \right) c_I + \langle I | \hat{H} | \hat{T} | 0 \rangle \\
&= \left( -\langle I | \hat{H} | \phi_{Qp}^{ab} \rangle (\mathbf{U}^{-1})_R^Q (\mathbf{U}^{-1})_S^R T_{ab}^S \right) 2c_I^2 + \langle I | \hat{H} | \hat{T} | 0 \rangle \quad (3.273)
\end{aligned}$$

$$\frac{\partial \left( \langle 0 | \hat{T} \hat{H} | \hat{T} | 0 \rangle \right)}{\partial c_I} \quad (3.274)$$

$$\begin{aligned}
&= 2 \langle 0 | \hat{T} \hat{H} | \phi_{Qp}^{ab} \rangle \frac{\partial (\mathbf{U}^{-\frac{1}{2}})_P^Q T_{ab}^{Pp}}{\partial c_I} \\
&= -\frac{1}{2} 2 c_J \langle 0 | \hat{T} \hat{H} | \phi_{Qp}^{ab} \rangle (\mathbf{U}^{-\frac{3}{2}})_P^Q T_{ab}^{Pp} \\
&= -2 c_J \langle 0 | \hat{T} \hat{H} | \phi_{Qp}^{ab} \rangle (\mathbf{U}^{-\frac{3}{2}})_P^Q T_{ab}^{Pp} \quad (3.275)
\end{aligned}$$

Inserting the definitions of the terms involving transformed amplitudes into the derivative expression, the full first derivative can be defined.

$$\begin{aligned}
\frac{\partial E}{\partial c_I} &= \frac{2 \left( \langle 0 | \hat{H} | I \rangle + \langle I | \hat{H} | J \rangle c_J \right) \left( 1 + c_J^2 + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle \right)}{\left( 1 + c_J^2 + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle \right)^2} \\
&\quad - \frac{\left( 4 \langle 0 | \hat{H} | J \rangle c_J + 2 c_K \langle K | \hat{H} | J \rangle c_J \right) c_J}{\left( 1 + c_J^2 + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle \right)^2} \\
&\quad + \left( -\langle I | \hat{H} | \phi_{Qp}^{ab} \rangle (\mathbf{U}^{-1})_R^Q (\mathbf{U}^{-1})_S^R T_{ab}^S \right) 2c_I^2 + \langle I | \hat{H} | \hat{T} | 0 \rangle \\
&\quad - 2 c_J \langle 0 | \hat{T} \hat{H} | \phi_{Qp}^{ab} \rangle (\mathbf{U}^{-\frac{3}{2}})_P^Q T_{ab}^{Pp} \quad (3.276)
\end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial c_I} = \frac{\partial E}{\partial c_I} + \lambda c_I^0 \quad (3.277)$$

$$\lambda = -\mathbf{c}_0^\dagger \sigma \quad (3.278)$$

The Lagrange multiplier is found via Schmidt Orthogonalisation.

### 3.5.3 Subspace Method

The equations above defining the macroiterations are enough to define the working equations of the MRLPF+1 method. However, as with CI convergence can be accelerated via use of the Davidson subspace method. The method is outlined here for the CI case before specific application to the MRLPF+1 equations.

In the Davidson subspace method[110], the eigenvectors are constructed as linear combinations of expansion vectors. The first expansion vector is the current optimised vector, with the others being previous guesses. These expansion vectors form the columns of a matrix,  $\mathbf{V}$ ,

$$c_I = \sum_i \alpha_i V_{Ii} \quad (3.279)$$

The optimal linear combination for CI is found via the subspace eigenvalue equation,

$$V_{Ii} H_{IJ} V_{Jj} \alpha_j = V_{Ii} V_{Ij} \alpha_j \epsilon \quad (3.280)$$

$$\bar{\mathbf{H}} \bar{\alpha} = \bar{\mathbf{S}} \bar{\alpha} \epsilon \quad (3.281)$$

In each macroiteration, the residuals,  $\mathbf{G}^{Dp}, \mathbf{g}^S$  and  $g^I$  are projected onto the space of expansion vectors, giving  $(\mathbf{G}^{Dp})^{(n)}, (\mathbf{g}^S)^{(n)}$  and  $(g^I)^{(n)}$  for each expansion vector(m).

$$\Psi_{ext}^n = \sum_S \sum_a (T^S)^n_a \Psi_S^a + \sum_{Dp} \sum_{ab} (T^{Dp})^n_{ab} \Psi_{Dp}^{ab} \quad (3.282)$$

Internal coefficients

$$\Psi = \sum_I c^I \Psi_I + \sum_n \alpha_n \Psi_{ext}^n \quad (3.283)$$

The microiterations optimise the  $\alpha_n$  coefficients and the coefficients of the internals. These coefficients are used to get better amplitudes  $T_{ab}^{Dp}, T_S^a$  and a new expansion vector is obtained.

The optimisation of the  $\alpha_n$  coefficients is done simultaneously with the optimisation of the internal coefficients. To do this the subspace eigenvalue equation is solved

(3.281), where the first  $n_{ext}$  elements of the vector  $\alpha$  are the coefficients of the external expansion vectors, and the remaining elements are the internal expansion coefficients, giving a full length  $\alpha$  vector of length  $n_{int} + n_{ext}$ . This matrix is typically less than  $10 \times 10$  in dimension which is significantly smaller than the full internal and external spaces.

However, the method described so far is that carried out for the MRCI problem, the MRLPF+1 method described in this chapter cannot be formed into an eigenvalue equation of the form above, so the equations must be solved slightly differently.

The MRLPF+1D energy cannot be written as an eigenvalue equation. Therefore, the expansion vectors do not contain any information regarding the pair external configurations, so in a doubles only theory, the external expansion vectors contain nothing, hence their Hamiltonian matrix elements in the space of expansion vectors are set to zero. If singles are included, the external expansion vectors contain only the singles information.

$$\Psi_{ext}^n = \sum_S \sum_a (T^S)_a^n \Psi_S^a \quad (3.284)$$

All objects are now written in this new basis, with orthogonalised expansion vectors.

Using the definitions:

$$c_I = \sum_i \alpha_i V_{Ii} \quad (3.285)$$

$$h_{ij} = V_{Ii} H_{IJ} V_{Jj} \quad (3.286)$$

$$S_{ij} = V_{Ii} V_{Ij} \quad (3.287)$$

$${}_q T_{ab}^Q = \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \Delta)^{-\frac{q}{2}} \right)_{PQ} T_{ab}^P \quad (3.288)$$

The energy can be defined as,

$$E_{MRLPF} = E_0 + \frac{\alpha_i V_{Ii} H_{IJ} V_{Jj} \alpha_j}{\alpha_i V_{Ii} V_{Ij} \alpha_j + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} + 2\alpha_i V_{Ii} \langle I | \hat{H} \hat{T} | 0 \rangle + \langle 0 | \hat{T} \hat{H} \hat{T} | 0 \rangle \quad (3.289)$$

As mentioned,  $\langle 0 | \hat{T} \hat{H} \hat{T} | 0 \rangle$  is not covered by the expansion vectors, and the internal coefficient vector  $\mathbf{c}$  has been written as a linear combination of internal expansion vectors. The vector corresponding to the “best” linear combination of expansion vectors,

$\alpha$ , now appears in place of  $c$ , including on the diagonals of the U matrix involved in the pair amplitude transformation.

Implementing the constraint on the internal coefficients, the Lagrangian to be minimised is defined,

$$\mathcal{L} = E_{MRLPF} + \lambda(\alpha^\dagger \alpha_0 - 1) \quad (3.290)$$

As seen above in 3.289, unlike the MRCI situation, the MRLPF+1D energy equation cannot be written as an eigenvalue equation, it has a different form, and must be solved differently.

To optimise the coefficients in the  $\alpha$  vector, i.e. to get “best” linear combination of the expansion vectors, the derivative of the Lagrangian must be zero with respect to the change in the coefficients. Accordingly, the derivative must be found.

### Derivative Of The Energy With Respect To $\alpha$

Now look at the derivative of the terms contributing to the energy that involve transformed amplitudes. Firstly, the energy contribution,  $E_c$ , containing  ${}_2\hat{T}$ .

$$\begin{aligned} E_c &= \alpha_i V_{Ii} \langle I | (\hat{H} - E_0) {}_2\hat{T} | 0 \rangle \\ &= \alpha_i V_{Ii} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-1} \right)_{PQ} T_{ab}^P \\ &= \alpha_i V_{Ii} G_I \end{aligned} \quad (3.291)$$

$$\frac{\partial E_c}{\partial \alpha_i} = V_{Ii} G_I + \alpha_k V_{Ik} \frac{\partial G_I}{\partial \alpha_i}$$

$$\frac{\partial G_I}{\partial \alpha_i} = -2 \alpha_j S_{ij} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-2} \right)_{PQ} T_{ab}^P$$

$$\begin{aligned} \frac{\partial E_c}{\partial \alpha_i} &= V_{Ii} G_I \\ &+ \alpha_k V_{Ik} \cdot \left( -2 \alpha_j S_{ij} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-2} \right)_{PQ} T_{ab}^P \right) \end{aligned}$$

$$\frac{\partial E_c}{\partial \alpha_i} = V_{Ii} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-1} \right)_{PQ} T_{ab}^P$$

$$+ \alpha_k V_{Ik} \left( -2 \alpha_j S_{ij} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-2} \right)_{PQ} T_{ab}^P \right) \quad (3.292)$$

$$= V_{Ii} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle {}_2T_{ab}^Q - (2 \alpha_j S_{ij}) \alpha_k V_{Ik} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle {}_4T_{ab}^Q \quad (3.293)$$

Secondly, the term involving  ${}_1\hat{T}$ .

$$\langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle \quad (3.294)$$

$$= \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-\frac{1}{2}} \right)_{SR} T_{cd}^S \langle R_{cd} | (\hat{H} - E_0) | Q_{ab} \rangle \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-\frac{1}{2}} \right)_{PQ} T_{ab}^P$$

$$\begin{aligned} & \frac{\partial \langle 0 | {}_1\hat{T}^\dagger (\hat{H} - E_0) {}_1\hat{T} | 0 \rangle}{\partial \alpha_i} \\ &= \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-\frac{1}{2}} \right)_{SR} T_{cd}^S \langle R_{cd} | (\hat{H} - E_0) | Q_{ab} \rangle \frac{\partial \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-\frac{1}{2}} \right)_{PQ} T_{ab}^P}{\partial \alpha_i} \\ &= \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-\frac{1}{2}} \right)_{SR} T_{cd}^S \langle R_{cd} | (\hat{H} - E_0) | Q_{ab} \rangle \left( (\alpha_i S_{ij} \alpha_j \mathbf{1} + \mathbf{\Delta})^{-\frac{3}{2}} \right)_{PQ} T_{ab}^P (-\alpha_j S_{ij}) \\ &= 2 {}_1T_{cd}^R \langle R_{cd} | (\hat{H} - E_0) | Q_{ab} \rangle {}_3T_{ab}^Q \cdot (-\alpha_j S_{ij}) \end{aligned} \quad (3.295)$$

Therefore the total derivative of the energy with respect to a change in the  $\alpha$  coefficients is as follows,

$$\begin{aligned} \frac{\partial E_{MRLPF}}{\partial \alpha_i} &= \frac{h_{ij} \alpha_j}{\alpha_k S_{kl} \alpha_l + \frac{2}{n} \langle \hat{T}^\dagger \hat{T} \rangle} - \frac{\alpha_k h_{kl} \alpha_l}{\alpha_k S_{kl} \alpha_l + \frac{2}{n} \langle \hat{T}^\dagger \hat{T} \rangle} \left( \frac{S_{ij} \alpha_j}{\alpha_k S_{kl} \alpha_l + \frac{2}{n} \langle \hat{T}^\dagger \hat{T} \rangle} \right) \\ &+ V_{Ii} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle {}_2T_{ab}^Q \\ &- (2 \alpha_j S_{ij}) \alpha_k V_{Ik} \langle I | (\hat{H} - E_0) | Q_{ab} \rangle {}_4T_{ab}^Q \\ &+ 2 {}_1T_{cd}^R \langle R_{cd} | (\hat{H} - E_0) | Q_{ab} \rangle {}_3T_{ab}^Q \cdot (-\alpha_j S_{ij}) \end{aligned} \quad (3.296)$$

$$\frac{\partial \mathcal{L}}{\partial \alpha_i} = \frac{\partial E_{MRLPF}}{\partial \alpha_i} + \lambda \alpha_{0i} \quad (3.297)$$

$$\text{Where } \lambda = -\frac{\partial \mathcal{L}}{\partial \alpha_i} \alpha_{0i} \quad (3.298)$$

The constraint on the coefficients is implemented via Schmidt orthogonalisation.

The coefficients  $\alpha$  are then found iteratively using a simple Newton-Raphson procedure or by the Augmented Hessian method of Lengsfeld[111].

### 3.5.4 Singles

single and double excitations. As stated in the introduction, it is common for a method to include single and double excitations, as is done in CISD and CCSD. This extends to a multireference wavefunction where singly and doubly-external configurations are generated, for example in icMRCISD [58], MRACPFSD [32], MRAQCCSD [31], icMRCCSD [53, 55] and Mk-MRCCSD[49]. In order for the MRLPFD method to be comparable to other multireference methods and therefore be of use, the method must allow for the inclusion of single excitations. The inclusion of singles amplitudes into the single reference LPF transformation has proved problematic as it destroys the desirable properties possessed by the transformation when only double amplitudes are present [11]. To avoid this issue, the extensions of LPF (i.e. AVCC, QVCC) use orbital optimisation to account for single excitations [11, 33, 39]. Due to these difficulties, in this work singles may be included in the MRLPFD formalism in the same way they are included in MRCI; in a linear fashion, without use of a transformation.

$$E_{MRLPFSD} = E_0 + \frac{\langle 0 | (1 + \hat{c} + \hat{s})^\dagger (\hat{H} - E_0) (1 + \hat{c} + \hat{s}) | 0 \rangle}{\langle 0 | 0 \rangle + \langle 0 | \hat{c}^\dagger \hat{c} | 0 \rangle + \langle 0 | \hat{s}^\dagger \hat{s} | 0 \rangle + \frac{2}{n} \langle 0 | \hat{T}^\dagger \hat{T} | 0 \rangle} + 2 \langle 0 | (1 + \hat{c}^\dagger) (\hat{H} - E_0) \hat{T} | 0 \rangle + \langle 0 | \hat{T}^\dagger (\hat{H} - E_0) \hat{T} | 0 \rangle \quad (3.299)$$

$$\mathbf{U} = (\mathbf{c}^\dagger \mathbf{c} + \mathbf{s}^\dagger \mathbf{s}) \mathbf{1} + \eta \quad (3.300)$$

This offers no handling of the size extensivity/consistency error introduced by the singles, but does offer a practical approach to improve the correlation energy.

## 3.6 Proposed Tests

The method presented in this chapter, MRLPFD, must be tested to show both its theoretically predicted attributes and its performance in comparison to other multireference methods currently in use. In addition, these numerical tests on the example systems can aid in the further development of the theory by showing the effects of changes to the method (for example, changing the normalisation condition  $c_0^\dagger c = 1$ ). An outline of a proposed set of tests designed to ascertain the performance of MRLPFD is now given. This list is not exhaustive but has been compiled to include the primary key tests.

### 3.6.1 Size Consistency and Extensivity

The theory presented here uses an LPF like transformation of the pair amplitudes in order to introduce a linked denominator, and additionally approximate VCC. As the method is implemented within an MRCI code, the emphasis is on the degree to which the size extensivity and consistency error is improved. In MRCI it is the pair norm which introduces the major part of the error, but in this work there is a linked denominator, so is extensive. The following proposed numerical tests are required to see the change in the extensivity/consistency error.

**N<sub>2</sub> and H<sub>2</sub>O** – This small molecule example, mentioned earlier (1.4.3), has been widely used to look at the consistency of methods, with comparisons for all methods outlined in the introduction[89].

**CH<sub>2</sub> ··· nHe** – The CH<sub>2</sub> ··· nHe model system of Cave and Davidson [112] is a general test of size-consistency. This example takes CH<sub>2</sub> and places a He atom 10Å away, along the principal symmetry axis. The calculation is then repeated, each time adding more helium atoms placed along the symmetry axis at the same regular (10Å) intervals. This adds a 2 electron system, for which the new MRLPF method should be exact, to the existing system. Without size-consistency, the description of CH<sub>2</sub> gets worse with every helium atom added, as can be seen via the difference in energy to FCI. Methods described earlier such as MRACPF, MRAQCC and MRCEPA(0) have been shown to reduce the MRCI error by an order of magnitude [26] and are termed

approximately-consistent. The method developed in this thesis must perform at least as well as the widely used methods of MRACPF and MRAQCC, so numerical testing and analysis of its size-consistency error is needed.

### 3.6.2 Potential Energy Curves

As stated, multireference methods are utilised when static correlation is considerable, which often occurs at long bond lengths. Accordingly, the ability of MRLPFD to reproduce the correct potential energy curve must be tested. For potential energy curves it is appropriate not only to look at the standard error compared to FCI, but also the non-parallelism error (NPE), the difference between the maximum error and minimum error. This measure gives some indication of how parallel the computed potential energy curve is to that of FCI, i.e. how precise it is, or conversely, how erratic the PES is. Aside from the qualitative behaviour, it is desirable for the method to be more accurate than competing multireference methods, recovering more correlation energy. This accuracy must also be tested.

**HF Dissociation** – The dissociation of hydrogen fluoride is another widely used model to test the performance of multireference methods. As such, there is a wide array of literature to compare against. Interestingly a recent review[50] has shown that for this system that MRACPF and MRAQCC give smaller NPE errors than most new MRCC methods. New icMRCC methods have given smaller absolute errors and smaller NPE errors [53]. The performance of MRACPF and MRAQCC makes it an interesting model for comparison for the MRLPFD method, where the MRLPFD method would need to be compared against MRACPF/MRAQCC results from MOLPRO, to remove differences due to contraction error.

**N<sub>2</sub> Potential Energy Surface** – The stretching of N<sub>2</sub> is often used as a test of new methods because it is a challenging problem, due to the static correlation involved. It is therefore an appropriate example with which to test the accuracy of MRLPFD. Around the equilibrium bond distance, the zeroth order wavefunction is dominated by a single configuration that has doubly occupied bonding  $\pi$  orbitals, and can be described by a single reference method. However, on stretching the bond the excited configurations become more important and a multireference method must be used.

**BeH<sub>2</sub> Potential Energy Curve** – The insertion of beryllium into H<sub>2</sub> is a model system introduced by Purvis *et al.*[113], where the potential energy curve of the <sup>1</sup>A<sub>1</sub> ground state is computed. This system has become a standard test for new multireference methods [31, 49, 53] as it is a symmetry forbidden reaction where the dominant configuration changes as the beryllium is moved.

The system is set with beryllium placed in the two-dimensional coordinate system at (0, 0). The hydrogen atoms are then placed at (x, y) and (x, -y), where x is the distance to the beryllium atom, and y is found via  $y(x) = 2.54 - 0.46x$ . The energy is then computed for a range of x, while constrained to a *c*<sub>2v</sub> symmetry.

For molecular BeH<sub>2</sub>, the dominant configuration in the zeroth-order wavefunction is,

$$|\psi_1\rangle = |(1a_1)^2(2a_1)^2(1b_2)^2\rangle \quad (3.301)$$

At large x, when the system exists as Be + H<sub>2</sub> (after dissociation), the dominant configuration is,

$$|\psi_2\rangle = |(1a_1)^2(2a_1)^2(3a_1)^2\rangle \quad (3.302)$$

An NPE can then be computed for the potential energy surface. This has been a popular test for multireference methods, and as such, a wide array of comparisons can be made with other multireference methods,

**Symmetric Water Stretch** – The symmetric stretch of water is an apt test for MRLPFD as it requires a large reference space. It has been used many times as a test of both single- and multi-reference methods [44, 53]. The ground state wavefunction around the equilibrium geometry is dominated by one configuration and can be described well by a Hartree-Fock determinant,

$$\Psi_{HF}\rangle = |(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2(1b_1)^2\rangle \quad (3.303)$$

However, upon stretching the hydrogen-oxygen bonds, a CASSCF wavefunction is required to qualitatively describe the potential energy surface correctly, with 6 electrons distributed in 5 orbitals, giving 28 reference configurations [53]. It has been shown that accuracy of JM based MRCC methods decreases with an increased active space [50]. This offers the opportunity to compare MRLPFD against the more costly JM based MRCC methods.

### 3.7 Comments and Future Development

The MRLPF+1D method defined in this chapter aims to eliminate the size extensivity error of MRCI. MRLPF+1D introduces a linked denominator via a transformation of the pair amplitudes that is similar in nature to the LPF+1 theory[29], and follows the spirit of coupled pair functional. The transformation also introduces an infinite series to approximate a multireference variational Coupled Cluster theory.

The pair norm is not the only source of error in MRCI, but it is the largest source of error, with the pair external excitations being more numerous than any other. Therefore, it is known that the method is indeed another “approximately extensive” method, but relative to others such as MRACPF and MRAQCC, the error will, theoretically, be much less. Numerical tests, such as those mentioned above, are needed to investigate this.

As an approximation to VCC, the transformation defined here gives a very simple first approximation to VCC, which, using the framework developed here, can be extended to capture more VCC terms and give a better approximation. Which other terms to be captured, and the form of their transformation is work yet to be done.

MRCI itself is invariant to orbital rotations in the separate orbital spaces[52]. The difference in this work to MRCI do not affect the orbital invariance. The transformation of the amplitudes is written in tensor notation, and is invariant to orbital rotations. The use of a linearly independent set of excitation operators is key to ensuring the orbital invariance of the method, and lack of such a set was one of the key reason why the “Multireference approximate VCC theory” development was halted. The use of the MRACPF style denominator for the internals (and singles if included), also does not affect orbital invariance, as MRACPF is already known to be invariant[31]. A more detailed theoretical argument may be discussed in the future, but would hinge on the same comments made here. It is also possible to prove orbital invariance through numerical calculations within MOLPRO. this has not been done here as, at present, the code implementation contains bugs.

There also exist two issues within the theory that are open to change based on numerical calculations (those described in section 3.6), these are, firstly, the use of  $\langle 0|\hat{T}^\dagger\hat{T}0\rangle$

in the internal denominator, and secondly, the use of  $c^\dagger c$  on the diagonals of the  $U$  matrix.

The rationale behind the use of the pair norm in the denominator of the part of the energy involving only the internal is based on what MRCI does. The author is aware that it may seem illogical, but the development of the method here has been based on still being equivalent to MRCI for the model system, which requires the use of the MRACPF style denominator.

The use of  $c^\dagger c$  on the diagonals of the  $U$  matrix may also be unwanted, based on the arguments in 3.2.8. Again, its inclusion is based on the equivalence to MRCI for the model case. Whether it should be used or not will be guided by the numerical results of test calculations such as those outlined earlier.

After taking into account the results of those numerical calculations, the MRLPF+1D method that is of most use in the general case, can be definitively defined.

## Chapter 4

# Conclusions and Comments

The aim of the work presented in this thesis was to develop new methods to approximately solve the non-relativistic electronic Schrödinger equation, under the Born-Oppenheimer approximation; the solution of which can provide valuable insights into many areas of chemistry, physics and materials science.

The first chapter introduces the subject of electronic structure theory and some of the most common methods employed in approximate solution of the Schrödinger equation. We then focus on the need for a multiconfigurational reference wavefunction and describe some common multireference methods. Of particular importance are MRCC, MRCI and MRCI based methods, that are relevant to the work developed here.

Chapter 2 shows the development of “MR3VCC”. This is a direct approximation to a multireference variational Coupled Cluster theory. This approach started with the 3rd order VCC energy and decomposed the 4th order density matrices to allow removal of unlinked terms. The importance of the remaining terms was determined and an LPF-like transformation was formed to capture the desired 3rd order energy term, and a subset of VCC terms to infinite order. It was noted that the spin-free form was not equivalent to the spin-orbital form, only being the same for the two electron case. Development and implementation of this theory was halted as it was noted during the research that the method of redundancy removal was key to retaining orbital invariance; requiring a singular value decomposition and use of orthogonal electron pairs. Accordingly, focus switched to an LPF-like method formed using orthogonal pairs.

The MRLPF+1 theory of chapter 3 is a direct extension of the LPF+1 theory. It uses the same basic form of the transformation matrix, with significant complications arising due to the multiconfigurational reference wavefunction. Importantly this theory uses a singular value decomposition to remove redundancies and therefore uses orthogonal pairs, in order to retain orbital invariance. Evaluation of the 2-electron 2-reference model system is given, as well as limiting case of a single reference function. This chapter also details how to embed the transformation of the pair amplitudes into the MRCI energy functional, including inclusion within the efficient Davidson subspace method for solving the MRCI equations.

This method captures static correlation and uses an approximate Coupled Cluster approach to describe dynamic correlation. Therefore this method is capable of describing the many body effects seen in chemistry, such as van de Waals and dispersion forces, that other cheaper electronic structure methods (like Hartree-Fock) struggle to model. MRLPF+1 is a multireference method, and can therefore be applied to study bond breaking, cases of near degenerate states and other multireference situations that single reference methods are incapable of describing with chemical accuracy, even the single reference LPF+1 method.

Numerical tests of the MRLPF+1 method (see 3.6) are required to show how well the method corrects the MRCI size-extensivity and consistency errors and its general performance on capturing correlation energy. The properties of extensivity and consistency are most obvious when calculating bond dissociation energies and separating fragments. As such, the MRLPF+1 method is likely to show strong performance compared to MRCI for such examples. The application of this theory is not limited to those cases, but is likely to improve upon the performance of MRCI in general, as will be highlighted by the proposed tests. The method should perform better than existing MRCI corrections, such as the Davidson and Pople corrections, but must also perform better than other multireference methods such as MRACPF and MRAQCC. Aside from correcting MRCI, this method approximates variational Coupled Cluster theory. In this regard it is a simple approximation, but extensions can be made to better approximate VCC. As it is an approximate Coupled Cluster method, the performance of MRLPF+1 should be compared with that of modern MRCC approaches, such as Mk-MRCC and the recently developed icMRCC theories of the groups of Gauss and Köhn (outline

and references in introduction), whose computational scaling is greater than that of MRLPF+1.

The consideration of cost is important when developing *ab initio* electronic structure theories, as methods must be capable of providing relevant information in reasonable timescales. As noted earlier (see section 3.2.6), the MRLPF+1 method is slightly more costly than MRCI but doesn't have an increase in complexity. In addition, the work presented here is embedded in the internally contracted MRCI theory and implemented within as such within MOLPRO. This greatly reduces the computational demand as the number of amplitudes is not dependent on the number of configurations, just the number of correlated orbitals. Accordingly internally contracted MRCI, although it contains a small contraction error, is far more applicable than the standard uncontracted MRCI, and as such has been widely used. Embedding MRLPF+1 within the icMRCI theory should allow the MRLPF+1 method to be used in all situations that icMRCI is currently used in, examples include the small molecule studies in [114] and [115]. Work in the Werner groups has recently provided a more efficient implementation in MOLPRO using an integrated tensor framework [116], including an explicitly correlation version [117]. These methods have been used to treat dioxygen-copper complexes with different ligands [116] and study of the conical intersections of LiF [118] extending further the applicability of icMRCI. The MRLPF+1 could be implemented within this new icMRCI framework, which would further reduce computational cost and extend its applicability.

The MRLPF+1 offers an approximately size-extensive alternative to MRCI, MRACPF and MRAQCC, but also offers an opportunity to approximate a multireference Coupled Cluster theory (a multireference variational Coupled Cluster theory) that can be implemented as an extension of existing, efficient, icMRCI software.



# Appendix

## 2-Electron Multi-Reference Model System

A very simple model system is used in this thesis, to help guide the development of the methods and to compare between methods.

The multireference approximate variational coupled cluster theories must be formulated to be exact for this 2 electron case.

The model consists of two electrons, with 2 reference determinants. There are 2 occupied orbitals with different symmetry and one virtual orbital.

$$|1^2\rangle = 1\bar{1} \quad (4.1)$$

$$|2^2\rangle = 2\bar{2} \quad (4.2)$$

Due to there being only 1 virtual orbital, there is only 1 excited configuration:

$$|3^2\rangle = 3\bar{3} \quad (4.3)$$

and 2 amplitudes,  $T_{3\bar{3}}^{1\bar{1}}$  and  $T_{3\bar{3}}^{2\bar{2}}$

The reference wavefunction:

$$|0\rangle = \cos \theta |1^2\rangle + \sin \theta |2^2\rangle \mathbf{c}_0 = (\cos \theta \quad \sin \theta) \quad (4.4)$$

The MRCI wavefunction (without internal excitations):

$$|\Psi\rangle = |0\rangle + \mu |3^2\rangle \quad (4.5)$$

### Internal Excitations

For a full energy expression, the internal energy must be considered. The only Internal excitations,

$$\begin{aligned}
 (1 + \hat{c})|0\rangle &= |0\rangle + |1\rangle \\
 |\Psi_{MRCI}\rangle &= |0\rangle + \sigma|1\rangle + \mu|3^2\rangle \\
 |1\rangle &= \sigma(-\sin\theta|1^2\rangle + \cos\theta|2^2\rangle) \\
 \mathbf{c}' &= \sigma \begin{pmatrix} -\sin\theta & \cos\theta \end{pmatrix} \\
 \\ 
 \mathbf{c} &= \mathbf{c}_0 + \mathbf{c}' \\
 \mathbf{c} &= (\cos\theta - \sigma\sin\theta \quad \sin\theta + \sigma\cos\theta)
 \end{aligned}$$

The choice of  $|1\rangle$  ensures normalisation and orthogonality with the reference:

$$\begin{aligned}
 (1 + \hat{c})|0\rangle &= |0\rangle + \sigma|1\rangle \\
 &= (\cos\theta - \sigma\sin\theta)|1^2\rangle + (\sigma\cos\theta - \sin\theta)|2^2\rangle \quad (4.6)
 \end{aligned}$$

$$\begin{aligned}
 \langle(1 + \hat{c}^\dagger)|1 + \hat{c}\rangle &= (\cos^2\theta + \sigma^2\sin^2\theta)\langle 1^2|1^2\rangle + (\sigma^2\cos^2\theta + \sin^2\theta)\langle 2^2|2^2\rangle \\
 &= \cos^2\theta + \sigma^2\sin^2\theta + \sigma^2\cos^2\theta + \sin^2\theta \\
 &= 1 + \sigma^2 \quad [n.b. \quad \sin^2\theta + \cos^2\theta = 1] \quad (4.7)
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{c}_0^\dagger \mathbf{c}_0 &= 1 \\
 \mathbf{c}_0^\dagger \mathbf{c}' &= (\cos\theta + \sin\theta)(1 - \sigma) \\
 \mathbf{c}_0^\dagger \mathbf{c} &= 1 \\
 \mathbf{c}^\dagger \mathbf{c} &= 1 + \sigma^2
 \end{aligned}$$

## List of Acronyms

|          |  |
|----------|--|
| ACPF     | Average Coupled-Pair Functional                                  |
| AQCC     | Average Quadratic Coupled Cluster                                |
| AVCC     | Approximate Variational Coupled Cluster                          |
| CBH      | Campbell-Baker-Hausdorff (expansion)                             |
| cc-pVXZ  | correlation consistent polarised valence $X$ -tuple zeta (basis) |
| CCSD     | Coupled Cluster Singles and Doubles                              |
| CISD     | Configuration Interaction Singles and Doubles                    |
| CEPA     | Coupled Electron-Pair Approximation                              |
| CPF      | Coupled-Pair Functional  |
| CSE      | Contracted Schrödinger Equation                                  |
| CT       | Canonical Transformation   |
| EPV      | Exclusion Principle Violating                                    |
| FCI      | Full Configuration Interaction                                   |
| HOMO     | Highest Occupied Molecular Orbital                               |
| icMRCC   | Internally Contracted MRCC                                       |
| icMRCI   | Internally Contracted MRCI                                       |
| LCCM     | Linearized Coupled Cluster Method                                |
| LPF      | Linked (electron) Pair Functional                                |
| LPF $_x$ | Linked Pair Functional variant $\lambda = x$                     |
| LUMO     | Lowest Unoccupied Molecular Orbital                              |
| MCSCF    | Multi-Configurational Self-Consistent Field                      |
| Mk-MRCC  | Mukherjee's state specific MRCC                                  |
| MR       | Multi-Reference  |
| MRCC     | MultiReference Coupled Cluster                                   |
| MRCI     | MultiReference Configuration Interaction                         |
| MRVCC    | MultiReference Variational Coupled Cluster                       |
| MR3VCC   | 3rd Order Approximate MRVCC                                      |
| MR-ACPF  | MultiReference ACPF  |

|            |  |
|------------|--|
| MR-AQCC    | MultiReference AQCC  |
| MR-CEPA    | MultiReference CEPA  |
| MR-CEPA(0) | MultiReference CEPA(0)   |
| MR-LCCM    | MultiReference LCCM  |
| MR-SC2-CI  | MultiReference Self-Consistent Size-Consistent Configuration Interaction |
| NPE        | Non-Parallelism Error  |
| PES        | Potential Energy Surface   |
| QCC        | Quadratic Coupled Cluster (method)                                       |
| QVCC       | Quasi-Variational Coupled Cluster  |
| RDM        | Reduced Density Matrix. nRDM: n-body RDM.                                |
| RHF        | Restricted Hartree-Fock  |
| SCF        | Self-Consistent Field  |
| SR         | Single-Reference   |
| TCC        | Traditional Coupled Cluster  |
| tdm        | transition density matrix  |
| VCC        | Variational Coupled Cluster  |

# Bibliography

- [1] D. R. Hartree, Proc. Cambridge Phil. Soc. **24**, 89 (1928).
- [2] V. Fock, Z. Phys. **61**, 126 (1930).
- [3] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, 1996.
- [4] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
- [5] G. G. Hall, Proc. R. Soc. London. Ser. A. Math. Phys. Sci. **205**, 541 (1951).
- [6] D. I. Lyakh, M. Musia, V. F. Lotrich, and R. J. Bartlett, Chem. Rev. **112**, 182 (2012).
- [7] K. A. Brueckner, Phys. Rev. **100**, 36 (1955).
- [8] J. Goldstone, Proc. R. Soc. A Math. Phys. Eng. Sci. **239**, 267 (1957).
- [9] T. D. Crawford and H. F. Schaefer, *An Introduction to Coupled Cluster Theory for Computational Chemists*, pages 33–136, John Wiley and Sons, Inc., 2007.
- [10] E. Merzbacher, *Quantum Mechanics*, Wiley, New York, 2nd edition, 1970.
- [11] J. Robinson, *Approximate Variational Coupled Cluster Theories*, PhD thesis, School of Chemistry, Cardiff University, 2012.
- [12] T. V. Voorhis and M. Head-Gordon, J. Chem. Phys. **113**, 8873 (2000).
- [13] J. A. Pople, M. Head-Gordon, and K. Raghavachari, J. Chem. Phys. **87**, 5968 (1987).
- [14] J. D. Watts, M. Urban, and R. J. Bartlett, Theor. Chim. Acta **90**, 341 (1995).

- [15] J. Paldus, J. Cizek, and B. Jeziorski, *J. Chem. Phys.* **90**, 4356 (1989).
- [16] S. Koch and W. Kutzelnigg, *Theor. Chim. Acta* **59**, 387 (1981).
- [17] R. J. Barlett and G. D. Purvis III, *Int. J. Quantum Chem.* **14**, 561 (1978).
- [18] J. Čížek, *Advan. Chem. Phys.* **14**, 35 (1969).
- [19] W. D. Laidig and R. J. Bartlett, *Chem. Phys. Lett.* **104**, 424 (1984).
- [20] R. Ahlrichs, *J. Chem. Phys.* **48**, 1819 (1968).
- [21] W. Meyer, *J. Chem. Phys.* **58**, 1017 (1973).
- [22] C. Kollmar and F. Neese, *Molecular Physics* **108**, 2449 (2010).
- [23] W. Kutzelnigg, Unconventional Aspects of Coupled-Cluster Theory, in *Recent Progress in Coupled Cluster Methods*, edited by P. Čársky, J. Paldus, and J. Pittner, pages 1–61, Springer, 2010.
- [24] W. Kutzelnigg and P. von Herigonte, *Electron Correlation at the Dawn of The 21st century*, volume 36 of *Advances in Quantum Chemistry*, Academic Press, 1999.
- [25] S. Koch and W. Kutzelnigg, *Theor. Chim. Acta* **59**, 387 (1981).
- [26] P. Szalay, Towards state-specific formulation of multireference coupled-cluster theory: Coupled electron pair approximations (cepa) leading to multireference configuration interaction (mr-ci) type equations, in *Recent Advances in Coupled-Cluster Methods*, edited by R. J. Bartlett, page 81, World Scientific, Singapore, 1997.
- [27] R. Ahlrichs, P. Scharf, and C. Ehrhardt, *J. Chem. Phys.* **82**, 890 (1985).
- [28] D. P. Chong and S. R. Langhoff, *J. Chem. Phys.* **84**, 5606 (1986).
- [29] P. J. Knowles and B. Cooper, *J. Chem. Phys.* **133**, 224106 (2010).
- [30] B. Cooper, *Benchmark Calculations of Projection and Variational Coupled Cluster Methods*, PhD thesis, School of Chemistry, Cardiff University, 2010.
- [31] R. J. Gdanitz and R. Ahlrichs, *Chem. Phys. Lett.* **143**, 413 (1988).

- [32] P. G. Szalay and R. J. Bartlett, *Chem. Phys. Lett.* **214**, 481 (1993).
- [33] J. B. Robinson and P. J. Knowles, *J. Chem. Phys.* **135**, 044113 (2011).
- [34] U. S. Mahapatra, B. Datta, and D. Mukherjee, *J. Chem. Phys.* **110**, 6171 (1999).
- [35] F. A. Evangelista, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **127**, 024102 (2007).
- [36] P. Pulay and S. Sæbø, *Chem. Phys. Lett.* **117**, 37 (1985).
- [37] T. Van Voorhis and M. Head-Gordon, *Chem. Phys. Lett.* **330**, 585 (2000).
- [38] P. G. Szalay, M. Nooijen, and R. J. Bartlett, *J. Chem. Phys.* **103**, 281 (1995).
- [39] J. B. Robinson and P. J. Knowles, *J. Chem. Phys.* **136**, 054114 (2012).
- [40] T. J. Lee and P. R. Taylor, *Int. J. Quantum Chem.* **36**, 199 (1989).
- [41] K. Boguslawski, P. Tecmer, O. Legeza, and M. Reiher, *J. Phys. Chem. Lett.* **3**, 3129 (2012).
- [42] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory*, John Wiley & Sons, LTD, Chichester, 2000.
- [43] H.-J. Werner, Matrix-Formulated Direct Multiconfiguration Self-Consistent Field and Multiconfiguration Reference Configuration-Interaction Methods, in *Adv. Chem. Phys.*, pages 1–62, John Wiley & Sons, Inc., 1987.
- [44] P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, *Chem. Rev.* **112**, 108 (2012).
- [45] K. Andersson, P. k. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96** (1992).
- [46] D. Mukherjee, R. K. Moitra, and A. Mukhopadhyay, *Mol. Phys.* **33**, 955 (1977).
- [47] B. Jeziorski and H. J. Monkhorst, *Phys. Rev. A* **24**, 1668 (1981).
- [48] J. Masik, I. Hubac, and P. Mach, *J. Chem. Phys.* **108**, 6571 (1998).
- [49] U. S. Mahapatra, B. Datta, and D. Mukherjee, *J. Chem. Phys.* **110**, 6171 (1999).
- [50] S. Das, D. Mukherjee, and M. Kállay, *J. Chem. Phys.* **132**, 074103 (2010).

- [51] S. Chattopadhyay, D. Pahari, D. Mukherjee, and U. S. Mahapatra, *J. Chem. Phys.* **120**, 5968 (2004).
- [52] L. Kong, *Int. J. Quantum Chem.* **110**, 2603 (2010).
- [53] F. A. Evangelista and J. Gauss, *J. Chem. Phys.* **134**, 114102 (2011).
- [54] F. A. Evangelista, M. Hanauer, A. Köhn, and J. Gauss, *J. Chem. Phys.* **136**, 204108 (2012).
- [55] M. Hanauer and A. Köhn, *J. Chem. Phys.* **134**, 204111 (2011).
- [56] M. Hanauer and A. Köhn, *J. Chem. Phys.* **136**, 204107 (2012).
- [57] M. Hanauer and A. Kohn, *J. Chem. Phys.* **137**, 131103 (2012).
- [58] H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **89**, 5803 (1988).
- [59] A. Köhn, M. Hanauer, L. A. Mück, T.-C. Jagau, and J. Gauss, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **00**, 1 (2012).
- [60] P. Carsky, J. Paldus, and J. Pittner, *Recent Progress in Coupled Cluster Methods: Theory and Applications*, Challenges and Advances in Computational Chemistry and Physics, Springer, 2010.
- [61] R. Bartlett, *Recent Advances in Coupled-Cluster Methods*, Recent Advances in Computational Chemistry, World Scientific, 1997.
- [62] T. Yanai and G. K.-L. Chan, *J. Chem. Phys.* **124**, 194106 (2006).
- [63] T. Yanai and G. K.-L. Chan, *J. Chem. Phys.* **127**, 104107 (2007).
- [64] E. Neuscamman, T. Yanai, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 124102 (2009).
- [65] E. Neuscamman, T. Yanai, and G. K.-L. Chan, *Int. Rev. Phys. Chem.* **29**, 231 (2010).
- [66] V. V. Ivanov and L. Adamowicz, *J. Chem. Phys.* **112**, 9258 (2000).
- [67] D. I. Lyakh, V. V. Ivanov, and L. Adamowicz, *J. Chem. Phys.* **128**, 074101 (2008).

- [68] H.-J. Werner and E. Reinsch, *J. Chem. Phys.* **76** (1982).
- [69] P. G. Szalay and R. J. Bartlett, *J. Chem. Phys.* **103** (1995).
- [70] W. D. Laidig, P. Saxe, and R. J. Bartlett, *J. Chem. Phys.* **86**, 887 (1987).
- [71] P. G. Szalay, M. Nooijen, and R. J. Bartlett, *J. Chem. Phys.* **103**, 281 (1995).
- [72] P. Szalay, T. Muller, and H. Lischka, *Phys. Chem. Chem. Phys.* **2**, 2067 (2000).
- [73] P. J. A. Ruttink, J. H. van Lenthe, R. Zwaans, and G. C. Groenenboom, *J. Chem. Phys.* **94**, 7212 (1991).
- [74] J. Meller, J. Heully, and J. Malrieu, *Chem. Phys. Lett.* **218**, 276 (1994).
- [75] J.-P. Malrieu, J.-P. Daudey, and R. Caballol, *J. Chem. Phys.* **101**, 8908 (1994).
- [76] P. G. Szalay and R. J. Bartlett, *J. Chem. Phys.* **103**, 3600 (1995).
- [77] J. Meller, J. Malrieu, and J. Heully, *Mol. Phys.* **101**, 2029 (2003).
- [78] S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.* **8**, 61 (1974).
- [79] W. Butscher, S.-k. Shih, R. J. Buenker, and S. D. Peyerimhoff, *Chem. Phys. Lett.* **52**, 457 (1977).
- [80] G. Hirsch and P. Bruna, *Chem. Phys. Lett.* **52**, 442 (1977).
- [81] P. Bruna, S. Peyerimhoff, and R. Buenker, *Chem. Phys. Lett.* **72**, 278 (1980).
- [82] W. L. Luken, *Chem. Phys. Lett.* **58**, 421 (1978).
- [83] M. R. A. Blomberg, *J. Chem. Phys.* **78**, 5682 (1983).
- [84] S. Prime, C. Rees, and M. A. Robb, *Mol. Phys.* **44**, 173 (1981).
- [85] J. Simons, *J. Phys. Chem.* **93**, 626 (1989).
- [86] H.-J. Werner et al., MOLPRO, version 2012.1, a package of ab initio programs, 2012.
- [87] L. Meissner, *Chem. Phys. Lett.* **146**, 204 (1988).
- [88] J. A. Pople, R. Seeger, and R. Krishnan, *Int. J. Quantum Chem.* **12**, 149 (1977).

- [89] L. Füsti-Molnár and P. G. Szalay, *J. Phys. Chem.* **100**, 6288 (1996).
- [90] W. Duch and G. H. F. Diercksen, *J. Chem. Phys.* **101**, 3018 (1994).
- [91] P. G. Szalay, Configuration Interaction: Corrections for Size-Consistency, in *Encycl. Comput. Chem.*, edited by P. v. R. Schleyer et al., chapter Configurat, John Wiley & Sons, Ltd, Online, 2005.
- [92] K. R. Yang, A. Jalan, W. H. Green, and D. G. Truhlar, *J. Chem. Theory Comput.* **9**, 418 (2013).
- [93] D. Mukherjee, *Chem. Phys. Lett.* **274**, 561 (1997).
- [94] W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **107**, 432 (1997).
- [95] W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **110**, 2800 (1999).
- [96] W. Kutzelnigg and D. Mukherjee, *Chem. Phys. Lett.* **317**, 567 (2000).
- [97] M. Nooijen, M. Wladyslawski, and A. Hazra, *J. Chem. Phys.* **118**, 4832 (2003).
- [98] D. A. Mazziotti, *Acc. Chem. Res.* **39**, 207 (2006).
- [99] D. A. Mazziotti, *Phys. Rev. A* **81**, 062515 (2010).
- [100] D. A. Mazziotti, *Chem. Rev.* **112**, 244 (2012).
- [101] D. Mukherjee and W. Kutzelnigg, *J. Chem. Phys.* **114**, 2047 (2001).
- [102] W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **116**, 4787 (2002).
- [103] L. Kong, K. R. Shamasundar, O. Demel, and M. Nooijen, *J. Chem. Phys.* **130**, 114101 (2009).
- [104] D. Zgid, D. Ghosh, E. Neuscammann, and G. K.-L. Chan, *J. Chem. Phys.* **130**, 194107 (2009).
- [105] D. Datta, L. Kong, and M. Nooijen, *J. Chem. Phys.* **134**, 214116 (2011).
- [106] M. Hanauer and A. Köhn, *Chem. Phys.* **401**, 50 (2012).
- [107] K. R. Shamasundar, *J. Chem. Phys.* **131**, 174109 (2009).

- [108] G. E. Shilov, *Linear Algebra*, Dover, New York, 1st edition, 1977.
- [109] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **145**, 514 (1988).
- [110] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).
- [111] B. H. Lengsfeld, J. Chem. Phys. **73**, 382 (1980).
- [112] R. J. Cave and E. R. Davidson, J. Chem. Phys. **89**, 6798 (1988).
- [113] G. D. Purvis, R. Shepard, F. B. Brown, and R. J. Bartlett, Int. J. Quantum Chem. **23**, 835 (1983).
- [114] J. Czernek and O. Živný, J. Chem. Phys. **129**, 194305 (2008).
- [115] K. Stark and H.-J. Werner, J. Chem. Phys. **104**, 6515 (1996).
- [116] K. R. Shamasundar, G. Knizia, and H.-J. Werner, J. Chem. Phys. **135**, 054101 (2011).
- [117] T. Shiozaki, G. Knizia, and H.-J. Werner, J. Chem. Phys. **134**, 034113 (2011).
- [118] T. Shiozaki and H.-J. Werner, J. Chem. Phys. **134**, 184104 (2011).