RICE UNIVERSITY

Describing strong correlations with mean-field approximations

by

Takashi Tsuchimochi

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree
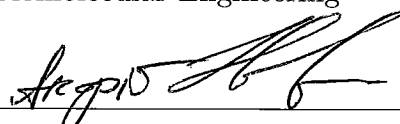
Doctor of Philosophy

Approved, Thesis Committee:

Gustavo E. Scuseria, Chair
Robert A. Welch Professor of Chemistry
and Physics and Astronomy

Cecilia Clementi
Wiess Career Development Chair,
Professor of Chemistry and Chemical and
Biomolecular Engineering

Andriy Nevidomskyy
Assistant Professor of Physics and
Astronomy

Houston, Texas

April, 2012

ABSTRACT

Describing strong correlations with mean-field approximations

by

Takashi Tsuchimochi

Strong electron correlations in electronic structure theory are purely quantum effects arising as a result of degeneracies in molecules and materials, and exhibit significantly different yet interesting characters than do weak correlations. Although weak correlations have recently been able to be described very efficiently and accurately within single particle pictures, less known are good prescriptions for treating strong correlations efficiently. Brute-force calculations of strong correlations in wave function theories tend to be very computationally-intensive, and are usually limited to small molecules for applications.

Breaking symmetry in a mean-field approximation is an efficient alternative to acquire strong correlations with, in many cases, qualitatively accurate results. The symmetry broken in quantum chemistry has been traditionally of spin, in so-called unrestricted methods, which typically break spatial symmetry as a consequence, and *vice versa*, in most situations. In this work, we present a novel approach to accurately describing strong correlations with a mean-field cost by means of Hartree-Fock-Bogoliubov (HFB) theory. We are inspired by the number-symmetry-breaking in HFB, which, with an attractive particle interaction, accounts for strong correlations, while maintaining spin and spatial symmetry. We show that this attractive interaction must be restricted to the chemically-relevant orbitals in an active space

to obtain physically meaningful results. With such constraints, our constrained pairing mean-field theory (CPMFT) can accurately describe potential energy curves of various strongly-correlated molecular systems, by cleanly separating strong and weak correlations. To achieve the correct dissociation limits in hetero-atomic molecules, we have modified our CPMFT functional by adding asymptotic constraints. We also include weak correlations by combining CPMFT with density functional theory for chemically accurate results, and reveal the connection between CPMFT and traditional unrestricted methods.

The similarity between CPMFT and unrestricted methods leads us to the idea of constrained active space unrestricted mean-field approaches. Motivated by CPMFT, we *partially* retrieve spin-symmetry that has been fully broken in unrestricted methods. We allow symmetry breaking *only* in an active space. This constrained unrestricted Hartree-Fock (CUHF) is an interpolation between two extrema: the fully broken-symmetry solution and the symmetry preserved solution. This thesis defines the theory behind and reports the results of CUHF. We first show that, if an active space is chosen to include only open-shell electrons, CUHF reduces to restricted open-shell Hartree-Fock (ROHF), and such CUHF proves in many ways significantly better than the traditional ROHF scheme. We then develop perturbation theory with CUHF as the zeroth order, and apply the methods to calculating singlet-triplet splitting energies, where a balanced description of correlation effects between singlet and triplet states is crucial.

Acknowledgments

Contents

4 Constrained Unrestricted Hartree-Fock: ROHF theory made simple 81

5 Constrained Active Space Unrestricted Mean-Field Theory 96

Illustrations

Tables

Preface

This thesis is based on parts of my research conducted with Professor Gustavo E. Scuseria at Rice University. Most of the results shown have been published previously in a series of papers:

- Takashi Tsuchimochi and Gustavo E. Scuseria,
  Strong correlations via constrained-pairing mean-field theory,
  Journal of Chemical Physics **131** 121102 (2009).

- Gustavo E. Scuseria and Takashi Tsuchimochi,
  Constrained-pairing mean-field theory. II. Exact treatment of dissociations to nondegenerate orbitals,
  Journal of Chemical Physics **131** 164119 (2009).

- Takashi Tsuchimochi, Gustavo E. Scuseria, and Andreas Savin,
  Constrained-pairing mean-field theory. III. Inclusion of density functional exchange and correlation effects via alternative densities,
  Journal of Chemical Physics **132** 024111 (2010).

- Takashi Tsuchimochi, Thomas M. Henderson, Gustavo E. Scuseria, and Andreas Savin,
  Constrained-pairing mean-field theory. IV. Inclusion of corresponding pair constraints and connection to unrestricted Hartree-Fock theory,
  Journal of Chemical Physics **133** 134108 (2010).

- Takashi Tsuchimochi and Gustavo E. Scuseria,
  Communication: ROHF theory made simple,

Journal of Chemical Physics **133** 141102 (2010).

- Takashi Tsuchimochi and Gustavo E. Scuseria,
  Constrained active space unrestricted mean-field methods for controlling spin contamination,
  Journal of Chemical Physics **134** 064101 (2011).

- Jason K. Ellis, Carlos A. Jiménez-Hoyos, Thomas M. Henderson, Takashi Tsuchimochi, and Gustavo E. Scuseria,
  Constrained-pairing mean-field theory. V. Triplet pairing formalism,
  Journal of Chemical Physics **135**, 034112 (2011).

Chapter 1

Introduction

The last few decades have witnessed the great success of computational chemistry as a tool for estimating chemical properties of a variety of molecules and solids. With methodologies based on single reference wave functions, one can calculate many useful quantities very accurately, such as heats of formation, barrier heights of chemical reactions, ionization energies, and excitation spectra. The success in single reference wave function theories, which are usually based on Hartree-Fock (HF) theory [1] as a starting point, however, have been mostly limited to *weakly correlated* (non-degenerate) systems. Such methods include perturbation theory, coupled-cluster theory, and density functional theory. For *strongly correlated* cases, *e.g.*, dissociated molecules and degenerate systems, a ground state wave function can only be correctly described with a linear combination of more than one Slater determinants. Therefore, methods that rely on a single reference determinant fail to account for effects that originate from degeneracies [2–6].

In molecular systems, strong correlations exist in different "flavors" and forms [3, 7]. The first one is referred to as "angular correlation" and is ubiquitous in atoms, appearing, for example, in the four-electron Be series. The $2s$ and $2p$ orbitals in the Be series are nearly-degenerate in energy, requiring a multi-reference wave function to correctly describe the behavior of the change in total energies with respect to the atomic number. Another type of strong correlations is "left-right correlation." The simplest case of left-right strong correlation is the dissociation of a closed-shell

electron pair to open-shell fragments, such as $H_2 \rightarrow H + H$. In this case, the bonding orbital and the anti-bonding orbital become exactly degenerate to each other at an infinite separation of two H atoms, and the correct electronic structure requires two Slater determinants in the same footing for the wave function.

A single reference method can achieve the correct energetic limits in some cases by breaking the spatial and spin-symmetry in its wave function. For example, the total energy of unrestricted HF (UHF) at the dissociation of H_2 is twice the energy of each fragment. It should be pointed out, however, that there are cases where UHF is not energetically correct at the dissociation limit, as in O_2 and CO_2 [8]. In addition, the electronic structure of UHF is typically broken-symmetry, *e.g.*, the α and β electron densities do not belong to the symmetry of the target system, and therefore such description is physically incorrect. The correct description of the dissociated H_2 is such that each atom has 0.5 α and 0.5 β electrons.

Complete active space self-consistent field (CASSCF) [9] is one of the multi-reference methods that can correctly describe such strong correlations, by spanning its wave function as a linear combination of all possible determinants within the chosen active space. If an active space is appropriately chosen, CASSCF offers qualitatively good descriptions on strong correlations. However, the computational cost of CASSCF grows exponentially with the increasing number of electrons and orbitals, limiting its application to small or medium size systems. Density matrix renormalization group (DMRG) recently has become a popular method for properly taking into account the multi-reference character of strong correlations with a computational cost that scales polynomially, as opposed to CASSCF. However, DMRG in its current formalism can handle only 1-dimensional (or 2-dimensional at most) systems, while 3-dimensional systems are, unfortunately, out of reach.

Therefore, the efficient yet accurate description of strong correlations in electronic structure theory remains an elusive goal. Despite its importance in many physical and chemical processes, a first-principles, accurate, black-box, and computationally inexpensive scheme for strong correlations remains unknown. One of the main purposes of this thesis is to provide a novel scheme to describe strong correlations at a mean-field cost ($\mathcal{O}(N^3)$). We will present a method inspired by electron number fluctuations and pairings in the mean-field Hartree-Fock-Bogoliubov (HFB) theory [10, 11], but these effects are constrained to an active space. This method, called constrained pairing mean-field theory (CPMFT), is a density matrix functional which is considered a mixture of HF and HFB with an attractive pairing interaction. CPMFT is built upon the previous work of Staroverov and Scuseria [12], who showed that the HFB scheme with an effective, scaled pairing interaction (ζ-HFB) can accurately describe the strong correlations occurring in the Be series and in molecular dissociations done over the correct spatial and spin-symmetry restricted HF (RHF) reference. Using the full attractive interaction, $-1/r_{12}$, for the pairing interaction along with an active space, CPMFT can accurately dissociate *any* closed-shell molecule to its fragments with the correct spatial and spin-symmetries, while UHF *cannot* in some cases, as explained above. Obviously, that CPMFT offers reasonable potential energy surfaces indicates it also predicts reasonable heats of formation and barrier heights. Yet, CPMFT is a mean-field approximation. This is a remarkable achievement because CPMFT, due to its efficient computational cost, can handle very large systems which for CASSCF and DMRG calculations are intractable. Since our results indicate that CPMFT can only include strong correlations, in order to achieve "chemical accuracy," we will combine it with density functional theory (DFT) in an attempt to include weak correlations arising due to electrons avoiding collisions with each other.

We will also investigate the deep connection between CPMFT and UHF.

With the connection between CPMFT and UHF we develop, we will also take advantage of the concept of CPMFT to generalize UHF, in order to control spin-contamination, which offers a measure of the qualitative error in a UHF wave function. This constrained UHF (CUHF) has been shown to be useful for obtaining restricted open-shell HF (ROHF) in a robust manner with physically meaningful orbitals and orbital energies, which traditional ROHF schemes do *not* possess. As a consequence, CUHF is able to offer accurate ionization potentials and excitation energies in open-shell systems, via Koopmans' theorem and time-dependent HF. Furthermore, by introducing an active space as in CPMFT, CUHF can control its spin-contamination. We use second-order peturbation theory guided by Møller and Plesset (MP2) for CUHF to include weak correlation, and Löwdin's spin-projection operator to include strong correlation. To show the advantage of active space CUHF, we will benchmark singlet-triplet splitting energies, one of the most important properties for small conjugated organic dyes such as those used in solar cells [13–15] and single-molecule magnets [16–19]. It will be demonstrated that CUHF gives very accurate results, while UHF without controlling spin-contamination typically fails.

The discussion of this work will proceed as follows. Chapter 2 reviews the basic features of HF and HFB as well as electron correlations, and also discusses the connection between Corrected HF (CHF) of Csányi and Arias [20] and ζ-HFB of Staroverov and Scuseria [12]. Chapter 3 provides the main ideas behind CPMFT and its development, including the results obtained in this work. Chapter 4 applies the concept of CPMFT to UHF to define CUHF, which is an alternative formulation of ROHF. In Chapter 5, we generalize this idea and describe second-order perturbation theory and spin-projection on top of CUHF. We also present the results of CUHF ap-

plied to singlet-triplet splitting energies. Finally, Chapter 6 includes some concluding remarks, regarding the successes and prospects of these methods.

Chapter 2

Basic Theories

2.1 Hartree-Fock

All of the problems in non-relativistic quantum chemistry under the Born-Oppenheimer approximation can be exactly solved by obtaining a wave function $|\Psi\rangle$ via the Schrödinger equation,

$$\hat{\mathcal{H}}|\Psi\rangle = E|\Psi\rangle, \tag{2.1}$$

where E is the total energy, and $\hat{\mathcal{H}}$ is the electronic Hamiltonian of the system,

$$\hat{\mathcal{H}} = \sum_{pq} h_{pq} c_p^\dagger c_q + \frac{1}{4} \sum_{pqrs} (\langle pq|rs \rangle - \langle pq|sr \rangle) c_p^\dagger c_q^\dagger c_s c_r. \tag{2.2}$$

Here, $c_p^\dagger$ and c_p are the creation and annihilation operators of an electron on a molecular spin orbital $|\phi_p\rangle$, and

$$h_{pq} = \langle \phi_p | \left(\frac{1}{2} \nabla^2 - \sum_A \frac{Z_A}{r_A} \right) | \phi_q \rangle,$$

$$\langle pq|rs \rangle = \langle \phi_p \phi_q | \frac{1}{r_{12}} | \phi_r \phi_s \rangle. \tag{2.3a}$$

Throughout this work, we will use $p, q, \dots$ for spin orbitals ϕ, $i, j, \dots$ for spatial orbitals ψ, and $\mu, \nu, \dots$ for atomic orbitals (AO) χ. In other words,

$$\phi(\mathbf{x}) = \begin{cases} \psi^\alpha(\mathbf{r})\alpha(\omega) \\ \quad \text{or} \\ \psi^\beta(\mathbf{r})\beta(\omega) \end{cases} \tag{2.4}$$

where $\mathbf{x} = (\mathbf{r}, \omega)$, and

$$\psi_i^\sigma(\mathbf{r}) = \sum_\mu C_{\mu i}^\sigma \chi_\mu(\mathbf{r}), \tag{2.5}$$

where σ is the spin index (α or β), and $\mathbf{C}^\sigma$ is the molecular orbital (MO) coefficients matrix of spin σ. The last equation is known as the linear-combination-of-atomic-orbitals (LCAO) approximation.

The essential difficulty of dealing with Eqs.(2.1, 2.2) is that they can be solved neither analytically, nor numerically in most situations. Therefore, one has to introduce an approximation to the equation.

The Hartree-Fock (HF) approximation is one of the most basic theories in the electronic structure field, and is often taken as a starting point of more accurate wave function calculations, like perturbation theory. Its wave function $|\Phi_{\mathrm{HF}}\rangle$ is a Slater determinant given by

$$|\Phi_{\mathrm{HF}}\rangle = \prod_p^{N_e} c_p^\dagger |-\rangle, \tag{2.6}$$

where N_e is the number of electrons, and $|-\rangle$ a physical vacuum. The HF energy can be evaluated as

$$\begin{aligned} E_{\mathrm{HF}} &= \frac{\langle \Phi_{\mathrm{HF}} | \hat{\mathcal{H}} | \Phi_{\mathrm{HF}} \rangle}{\langle \Phi_{\mathrm{HF}} | \Phi_{\mathrm{HF}} \rangle} \\ &= \sum_p h_{pp} + \frac{1}{2} \sum_{pq} \left(\langle pq|pq \rangle - \langle pq|qp \rangle \right), \end{aligned} \tag{2.7}$$

or more generally

$$E_{\mathrm{HF}} = \sum_{pq} h_{pq} \gamma_{qp} + \frac{1}{2} \sum_{pqrs} \left(\langle pr|qs \rangle - \langle pr|sq \rangle \right) \gamma_{pq} \gamma_{rs}, \tag{2.8}$$

where

$$\gamma_{pq} = \langle \Phi_{\mathrm{HF}} | c_q^\dagger c_p | \Phi_{\mathrm{HF}} \rangle \tag{2.9}$$

is the one-particle density matrix. Note that, in HF, $\boldsymbol{\gamma}$ is idempotent and Hermitian, *i.e.*, $\boldsymbol{\gamma}^2 = \boldsymbol{\gamma}$ and $\boldsymbol{\gamma}^\dagger = \boldsymbol{\gamma}$. In general $\boldsymbol{\gamma}$ can be partitioned into each spin-block as follows:

$$\boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\gamma}^\alpha & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\gamma}^\beta \end{pmatrix}, \tag{2.10}$$

where $\boldsymbol{\gamma}^\sigma$ ($\sigma = \alpha, \beta$) means the density matrix of σ spin. In *restricted* HF (RHF), the spatial parts of the α and β orbitals ϕ are identical and thus $\boldsymbol{\gamma}^\alpha = \boldsymbol{\gamma}^\beta$, while in *unrestricted* HF (UHF), they can be different, $\boldsymbol{\gamma}^\alpha \neq \boldsymbol{\gamma}^\beta$; thus, UHF has more variational freedom.

To minimize $E_{\mathrm{HF}}[\{\phi_i\}]$ subject to the orthonormal condition $\langle \phi_p | \phi_q \rangle = \delta_{pq}$, one constructs the Lagrangian $\mathcal{L}_{\mathrm{HF}}$,

$$\mathcal{L}_{\mathrm{HF}} = E_{\mathrm{HF}} + \sum_{pq} \epsilon_{qp} (\langle \phi_p | \phi_q \rangle - \delta_{pq}), \tag{2.11}$$

where ϵ_{qp} is the Lagrange multiplier. The variation in $\mathcal{L}_{\mathrm{HF}}$ gives

$$\begin{aligned} \delta\mathcal{L}_{\mathrm{HF}} &= \sum_p \langle \delta\phi_p | h | \phi_p \rangle + \sum_{pq} (\langle \delta\phi_p \phi_q | \phi_p \phi_q \rangle - \langle \delta\phi_p \phi_q | \phi_q \phi_p \rangle) - \sum_{pq} \epsilon_{qp} \langle \delta\phi_p | \phi_q \rangle + c.c. \\ &= 0, \end{aligned} \tag{2.12}$$

which reduces to

$$F|\phi_p\rangle = \sum_{q=1} \epsilon_{qp} |\phi_q\rangle, \tag{2.13}$$

with the HF Fock matrix

$$F_{pq} = \langle \phi_p | F | \phi_q \rangle = h_{pq} + \frac{1}{2} \sum_{rs} (\langle pr|qs \rangle - \langle pr|sq \rangle) \gamma_{rs}. \tag{2.14}$$

By unitary-rotating $\{|\phi_p\rangle\}$ to $\{|\phi'_p\rangle\}$, one can get the canonical (diagonal) form:

$$F|\phi'_p\rangle = \epsilon'_p |\phi'_p\rangle. \tag{2.15}$$

Thus, each HF orbital is associated with its one-particle (orbital) energy. In the canonical basis $\boldsymbol{\gamma}$ is also diagonal, and we obtain

$$[\mathbf{F}, \boldsymbol{\gamma}] = \mathbf{0}. \tag{2.16}$$

Note that, since E_{HF} can be considered a density matrix functional, $\mathbf{F}$ is also obtained as the derivative of E_{HF} with respect to the density matrix,

$$F_{pq} = \frac{\partial E_{\text{HF}}}{\partial \gamma_{qp}}. \tag{2.17}$$

In practice, we generally work in the LCAO approximation. In the AO basis, the HF equation to solve is given by the Roothaan equation,

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon}, \tag{2.18}$$

where $S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle$ is the overlap matrix between atomic orbitals, and $\boldsymbol{\epsilon}$ is a diagonal matrix of the eigenvalues (orbital energies).

2.2 Correlation energy

Here we discuss the basic idea of electron correlation within electronic structure theory.

In the HF approximation, the potential that each electron feels is the average due to the other electrons (a mean-field approximation), and therefore HF does not address most electron correlations, especially between opposite spins. Hence, the HF energy is always an upper bound for the exact solution of Eq.(2.1), which is known as full configuration interaction (FCI). The terminology "correlation energy" is defined as the difference in energy between HF and the exact solution. Historically, correlation energy has been divided to two types. The first is *weak* (dynamical)

correlation, which is a consequence of the "dance" of electrons, *i.e.*, electrons trying to avoid each other. For example, HF does not prohibit two opposite spin electrons from occupying the same position, and therefore weak correlation is not taken into account in HF. The second type of correlation is known as *strong* (non-dynamical) correlation, which arises due to degeneracies in a system. If there are determinants that are energetically close to the HF determinant, the qualitatively correct wave function is a linear combination of them, and HF is qualitatively incorrect. This is always observed in dissociation of closed-shell molecules to open-shell fragments, as already mentioned in the Introduction.

2.3 Hartree Fock Bogoliubov

The Hartree Fock Bogoliubov theory is a generalization of HF and the Bardeen-Cooper-Schrieffer (BCS) model to describe pairing correlations in nuclei and superconductivity in solids. Its wave function is still a Slater determinant, but with *quasiparticles* $\beta^\dagger, \beta$, which are a linear combination of particle (electron) creation and annihilation operators $c^\dagger, c$:

$$\beta_p^\dagger = \sum_q \left(U_{qp} c_q^\dagger + V_{qp} c_q \right),\tag{2.19a}$$

$$\beta_p = \sum_q \left(U_{qp}^* c_q + V_{qp}^* c_q^\dagger \right),\tag{2.19b}$$

or

$$\begin{pmatrix} \beta \\ \beta^\dagger \end{pmatrix} = \begin{pmatrix} \mathbf{U}^\dagger & \mathbf{V}^\dagger \\ \mathbf{V}^\mathrm{T} & \mathbf{U}^\mathrm{T} \end{pmatrix} \begin{pmatrix} c \\ c^\dagger \end{pmatrix} = \mathbf{W}^\dagger \begin{pmatrix} c \\ c^\dagger \end{pmatrix}.\tag{2.20}$$

Since we want quasiparticles to obey the conventional fermion commutation rules, $\mathbf{W}$ must be a unitary matrix that transforms between particles and quasiparticles. $\mathbf{U}$

and $\mathbf{V}$ can be decomposed by the Bloch-Messiah theorem as

$$
\begin{pmatrix} \mathbf{U} & \mathbf{V}^* \\ \mathbf{V} & \mathbf{U}^* \end{pmatrix} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^* \end{pmatrix} \begin{pmatrix} \bar{\mathbf{U}} & \bar{\mathbf{V}} \\ \bar{\mathbf{V}} & \bar{\mathbf{U}} \end{pmatrix} \begin{pmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^* \end{pmatrix}, \tag{2.21}
$$

where $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ have special diagonal forms:

$$
\bar{\mathbf{U}} = \begin{pmatrix} \bar{\mathbf{U}}_1 & & & \\ & \bar{\mathbf{U}}_2 & & \\ & & \ddots & \\ & & & \bar{\mathbf{U}}_{N_{\mathrm{orbs}}} \end{pmatrix}, \tag{2.22}
$$

$$
\bar{\mathbf{V}} = \begin{pmatrix} \bar{\mathbf{V}}_1 & & & \\ & \bar{\mathbf{V}}_2 & & \\ & & \ddots & \\ & & & \bar{\mathbf{V}}_{N_{\mathrm{orbs}}} \end{pmatrix}, \tag{2.23}
$$

where N_{orbs} is the number of spin orbitals and

$$
\bar{\mathbf{U}}_p = \begin{pmatrix} u_p & 0 \\ 0 & u_p \end{pmatrix}, \tag{2.24}
$$

$$
\bar{\mathbf{V}}_p = \begin{pmatrix} 0 & v_p \\ -v_p & 0 \end{pmatrix}, \tag{2.25}
$$

with $0 \leq u_p \leq 1$ and $0 \leq v_p \leq 1$.

We define the HFB wave function as a Slater determinant built with quasiparticles,

$$
|\Phi_{\mathrm{HFB}}\rangle = \prod_p \beta_p |-\rangle, \tag{2.26}
$$

so that

$$
\beta_k |\Phi_{\mathrm{HFB}}\rangle = 0, \tag{2.27}
$$

for all k. With the Bloch-Messiah decomposition, the wave function can be written as

$$
|\Phi_{\mathrm{HFB}}\rangle = \prod_p (u_p + v_p a_p^\dagger a_{\bar{p}}^\dagger) |-\rangle, \tag{2.28}
$$

where $\bar{p}$ indicates the "conjugate state" of p and is defined by the 2×2 matrix of Eqs.(2.24,2.25), and $\{a^\dagger, a\}$ are the electron creation and annihilation operators in the natural orbital (NO) basis where the HFB density matrix

$$\gamma_{pq} = \langle \Phi_{\text{HFB}} | c_q^\dagger c_p | \Phi_{\text{HFB}} \rangle \tag{2.29}$$

is diagonal. $\{a^\dagger\}$ and $\{c^\dagger\}$ are related by a unitary transformation:

$$a_p^\dagger = \sum_q D_{qp} c_q^\dagger. \tag{2.30}$$

In addition, we also define the *anomalous density* or *pairing matrix*,

$$\kappa_{pq} = \langle \Phi_{\text{HFB}} | c_q c_p | \Phi_{\text{HFB}} \rangle, \tag{2.31}$$

and $\boldsymbol{\gamma}$ and $\boldsymbol{\kappa}$ can be expressed in terms of $\mathbf{U}$ and $\mathbf{V}$ as

$$\boldsymbol{\gamma} = \mathbf{V}^* \mathbf{V}^{\text{T}}, \tag{2.32a}$$

$$\boldsymbol{\kappa} = \mathbf{V}^* \mathbf{U}^{\text{T}}. \tag{2.32b}$$

$\boldsymbol{\kappa}$ is anti-symmetric by its definition of Eq.(2.31). Note that an HFB wave function (Eq.(2.28)) violates number symmetry. In other words, $|\Phi_{\text{HFB}}\rangle$ is clearly not an eigenfunction of the number operator $\hat{N} = \sum_p c_p^\dagger c_p = \sum_p a_p^\dagger a_p$:

$$\hat{N} | \Phi_{\text{HFB}} \rangle \neq N_e | \Phi_{\text{HFB}} \rangle, \tag{2.33}$$

where N_e is the total number of electrons in the system. This is a natural consequence of the Bogoliubov transformation to quasiparticles (Eqs.(2.19, 2.20)). In order for a wave function to be meaningful, therefore, we require the average number of electrons in each determinant to be N_e, *i.e.*,

$$\langle \Phi_{\text{HFB}} | \hat{N} | \Phi_{\text{HFB}} \rangle = N_e. \tag{2.34}$$

Furthermore, $\mathbf{W}$ being unitary gives the following relations:

$$\boldsymbol{\gamma} - \boldsymbol{\gamma}^2 = -\boldsymbol{\kappa}\boldsymbol{\kappa}^\dagger, \tag{2.35a}$$

$$\boldsymbol{\gamma}\boldsymbol{\kappa} = \boldsymbol{\kappa}\boldsymbol{\gamma}^*. \tag{2.35b}$$

When $\boldsymbol{\kappa}$ is non-zero, $\boldsymbol{\gamma}$ is clearly non-idempotent. Information about pairing correlations is carried by the anomalous density matrix $\boldsymbol{\kappa}$. It is also useful to introduce the generalized density matrix or quasi-particle density matrix $\mathbf{R}$,

$$\mathbf{R} = \begin{pmatrix} \boldsymbol{\gamma} & \boldsymbol{\kappa} \\ -\boldsymbol{\kappa}^* & \mathbf{I} - \boldsymbol{\gamma}^* \end{pmatrix}, \tag{2.36}$$

which is associated with a quasiparticle single determinant wave function, and therefore $\mathbf{R}^2 = \mathbf{R}$ and $\mathbf{R}^\dagger = \mathbf{R}$, as in the HF density matrix.

The energy expression for HFB is given by

$$\begin{aligned} E_{\text{HFB}} &= \langle \Phi_{\text{HFB}} | \hat{\mathcal{H}} | \Phi_{\text{HFB}} \rangle \\ &= \sum_{pq} h_{pq} \gamma_{qp} + \frac{1}{2} \left(\langle pr | qs \rangle - \langle pr | sq \rangle \right) \gamma_{pq} \gamma_{rs} \\ &\quad + \frac{1}{4} \left(\langle pq | rs \rangle - \langle pq | sr \rangle \right) \kappa_{pq}^* \kappa_{rs}, \end{aligned} \tag{2.37}$$

where the first two terms correspond to the HF energy (Eq.(2.8)), and the last term, called the *pairing energy*, represents pairing correlation. One minimizes Eq.(2.37) subject to the idempotency condition of $\mathbf{R}$ and Eq.(2.34). The resulting HFB Lagrangian $\mathcal{L}_{\text{HFB}}$ becomes

$$\mathcal{L}_{\text{HFB}} = E_{\text{HFB}} + \mu \left(\sum_p \gamma_{pp} - N_e \right), \tag{2.38}$$

where the last term is introduced to fulfill the requirement Eq.(2.34), with μ being the chemical potential as a Lagrange multiplier. We should mention that Eq.(2.37) yields a lower energy than E_{HF} if the electron-electron interaction is attractive, as

in superconductors. With the electronic structure Hamiltonian, which has a repulsive electron-electron interaction, the pairing energy is always positive, and therefore Eq.(2.37) will reduce to the HF energy by the variational principle [21].

In the same way as in HF, one can derive the quasiparticle Hamiltonian $\mathbf{H}$ by

$$H_{pq}^{\text{HFB}} = \frac{\partial \mathcal{L}_{\text{HFB}}}{\partial R_{qp}}, \tag{2.39}$$

which is equivalent to

$$\mathbf{H}^{\text{HFB}} = \begin{pmatrix} \mathbf{F}^{\text{HFB}} & \mathbf{\Delta}^{\text{HFB}} \\ -\mathbf{\Delta}^{\text{HFB}*} & \mathbf{F}^{\text{HFB}*} \end{pmatrix}, \tag{2.40}$$

where

$$F_{pq}^{\text{HFB}} = \frac{\partial \mathcal{L}_{\text{HFB}}}{\partial \gamma_{qp}} = h_{pq} + \frac{1}{2}\sum_{rs}(\langle pr|qs\rangle - \langle pr|sq\rangle)\gamma_{rs} + \mu\delta_{pq}, \tag{2.41a}$$

$$\Delta_{pq}^{\text{HFB}} = \frac{\partial \mathcal{L}_{\text{HFB}}}{\partial \kappa_{qp}^*} = \frac{1}{4}\sum_{rs}(\langle pq|rs\rangle - \langle pq|sr\rangle)\kappa_{rs}. \tag{2.41b}$$

Note that $\mathbf{F}^{\text{HFB}} = \mathbf{F} + \mu\mathbf{I}$ is the HF Fock matrix (Eq.(2.14)) with the chemical potential μ in its diagonal. For the singlet pairing, in which the pairing correlation occurs only between opposite electron pairs, Eq.(2.40) becomes

$$\mathbf{H}^{\text{HFB}} = \begin{pmatrix} \mathbf{F}^{\alpha} + \mu\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{\Delta}^{\alpha\beta} \\ \mathbf{0} & \mathbf{F}^{\beta} + \mu\mathbf{I} & \mathbf{\Delta}^{\beta\alpha} & \mathbf{0} \\ \mathbf{0} & -\mathbf{\Delta}^{\alpha\beta*} & -\mathbf{F}^{\alpha*} - \mu\mathbf{I} & \mathbf{0} \\ -\mathbf{\Delta}^{\beta\alpha*} & \mathbf{0} & \mathbf{0} & -\mathbf{F}^{\beta*} - \mu\mathbf{I} \end{pmatrix}. \tag{2.42}$$

The HFB equation to solve is

$$\mathbf{H}^{\text{HFB}} \begin{pmatrix} \mathbf{U} & \mathbf{V}^* \\ \mathbf{V} & \mathbf{U}^* \end{pmatrix} = \begin{pmatrix} \mathbf{U} & \mathbf{V}^* \\ \mathbf{V} & \mathbf{U}^* \end{pmatrix} \begin{pmatrix} \boldsymbol{\epsilon} & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\epsilon} \end{pmatrix}, \tag{2.43}$$

where $\boldsymbol{\epsilon}$ is positive definite. Therefore, the eigenvalues of the HFB Hamiltonian come out as *conjugate-pairs*, ϵ_i and $-\epsilon_i$. In the quasiparticle basis β_i, both $\mathbf{H}$ and $\mathbf{R}$ are diagonal, and thus we have

$$[\mathbf{H}^{\text{HFB}}, \mathbf{R}] = \mathbf{0}, \tag{2.44}$$

which corresponds to Eq.(2.16) of HF.

HFB has been successfully used in nuclear and superconductor physics, with attractive interactions between particles. However, as already mentioned above, the electronic structure Hamiltonian contains a repulsive interaction between electrons. Therefore, when applied to molecular systems, HFB immediately reduces to HF.

In what follows, we limit our discussion to the real-orbital and closed-shell case (restricted HFB, or RHFB), in which case we have

$$\boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\gamma}^\alpha & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\gamma}^\beta \end{pmatrix} = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix}, \tag{2.45a}$$

$$\boldsymbol{\kappa} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\kappa}^{\alpha\beta} \\ \boldsymbol{\kappa}^{\beta\alpha} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{K} \\ -\mathbf{K} & \mathbf{0} \end{pmatrix}, \tag{2.45b}$$

for each spin-block. Here we have defined $\mathbf{P} = (\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta)/2$ since $\boldsymbol{\gamma}^\alpha = \boldsymbol{\gamma}^\beta$. Also, Eqs.(2.35) indicate

$$\mathbf{PK} - \mathbf{KP} = \mathbf{0}, \tag{2.46a}$$

$$\mathbf{P} - \mathbf{P}^2 = \mathbf{K}^2. \tag{2.46b}$$

$\mathbf{P}$ is *not* idempotent unless one ends up with an HF solution, *i.e.*, $\mathbf{K} = 0$. The RHFB energy expression can be given by

$$
\begin{aligned}
E_{\text{HFB}} &= 2\sum_{ij} h_{ij} P_{ij} + \sum_{ijkl} \left(2\langle ik|jl\rangle - \langle ik|lj\rangle\right) P_{ji} P_{lk} + \sum_{ijkl} \langle ij|kl\rangle K_{ij} K_{kl} \\
&+ \mu\left(N_e - 2\sum_i P_{ii}\right),
\end{aligned} \tag{2.47}
$$

in the spatial orbital basis. The RHFB Hamiltonian reduces to

$$\mathbf{H}^{\text{RHFB}} = \begin{pmatrix} \mathbf{F}^{\text{cs}} + \mu\mathbf{I} & \boldsymbol{\Delta} \\ \boldsymbol{\Delta} & -\mathbf{F}^{\text{cs}} - \mu\mathbf{I} \end{pmatrix}, \tag{2.48}$$

where

$$F_{ij}^{\text{cs}} = h_{ij} + \sum_{kl} (2\langle ik|jl \rangle - \langle ik|lj \rangle) P_{kl}, \tag{2.49a}$$

$$\Delta_{ij} = \Delta_{ij}^{\alpha\beta} \equiv \sum_{kl} \langle ij|kl \rangle K_{kl}. \tag{2.49b}$$

For more details about HFB, the reader is referred to Ref. [10].

2.4 Natural orbitals and natural occupations

In this section, we will briefly discuss the *natural orbitals* (NO). The NOs are the basis where the (half) *charge* density matrix $\mathbf{P} = \frac{1}{2}\left(\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta\right)$ is diagonal:

$$\mathbf{P}\mathbf{C}_i^{\text{NO}} = n_i \mathbf{C}_i^{\text{NO}}, \tag{2.50}$$

where the NO coefficient matrix $\mathbf{C}^{\text{NO}}$ is a unitary matrix that rotates orbitals to the NOs,

$$|\psi_i^{\text{NO}}\rangle = \sum_j C_{ji}^{\text{NO}} |\psi_j\rangle, \tag{2.51}$$

and n_i are the natural occupations of each $|\psi_i^{\text{NO}}\rangle$, and obey $0 \leq n_i \leq 1$.

In RHF, $\boldsymbol{\gamma}^\alpha = \boldsymbol{\gamma}^\beta$ and both spin density matrices are idempotent. This immediately means the RHF charge density is also idempotent, $\mathbf{P}^2 = \mathbf{P}$ and therefore n_i have to be integer, *i.e.*, either 1 or 0. On the other hand, in UHF, $\boldsymbol{\gamma}^\alpha \neq \boldsymbol{\gamma}^\beta$ and thus n_i can be fractional, but one can express the non-idempotency of $\mathbf{P}$ as

$$\mathbf{P} - \mathbf{P}^2 = \mathbf{M}^2, \tag{2.52}$$

where $\mathbf{M}$ is the spin magnetization density matrix defined as

$$\mathbf{M} = \frac{1}{2}(\boldsymbol{\gamma}^\alpha - \boldsymbol{\gamma}^\beta). \tag{2.53}$$

Because in UHF we write $\mathbf{P}$ as the sum of two idempotent matrices, its eigenvalues, *i.e.*, natural occupations, can be 0, 1, $\frac{1}{2}$, or appear in so-called "corresponding pairs" $(n, 1-n)$ [22].

In RHFB, $\mathbf{P}$ is also non-idempotent; hence, it has fractional occupations, but they are not necessarily corresponding pairs. However, $\mathbf{P}$ for RHFB holds a similar relation as Eq.(2.52), as given by Eq.(2.46b). From Eq.(2.46a), in the NO basis, $\mathbf{K}$ is also a diagonal matrix, and Eq.(2.46b) indicates that its eigenvalues are $\kappa_i = \pm\sqrt{n_i(1-n_i)}$.

2.5 Corrected HF and ζ-HFB

In 2000, Csányi and Arias proposed a natural orbital functional,[20] called corrected HF (CHF), whose energy expression is given by:

$$
\begin{aligned}
E_{\text{CHF}} = {} & 2\sum_i n_i h_{ii} + \sum_{ij} n_i n_j \left(2\langle ij|ij\rangle - \langle ij|ji\rangle\right) \\
& - \sum_{ij} \sqrt{n_i(1-n_i)n_j(1-n_j)}\langle ij|ji\rangle.
\end{aligned}
\tag{2.54}
$$

They arrived at Eq.(2.54) by expanding the two-particle density matrix (2PDM) as a tensor product of the one-particle density matrices. Therefore, the 2PDM of CHF is not guaranteed to be N-representable, meaning that there is no wave function associated with CHF. As can be seen, E_{CHF} is a functional of the NOs and natural occupations, *i.e.*, the density matrix, and thus it falls into a natural orbital functional or density matrix functional.

Later, Staroverov and Scuseria noticed that this functional is very similar to the RHFB energy.[12] In the NO basis, the RHFB energy is,

$$
\begin{aligned}
E_{\text{HFB}} = {} & 2\sum_i n_i h_{ii} + \sum_{ij} n_i n_j \left(2\langle ij|ij\rangle - \langle ij|ji\rangle\right) \\
& + \sum_{ij} \sqrt{n_i(1-n_i)n_j(1-n_j)}\langle ij|ji\rangle,
\end{aligned}
\tag{2.55}
$$

where we have taken the signs of K_{ii} to be positive and the term that involves μ is neglected because it is zero at convergence. Clearly, the difference between these two functionals is the sign of the pairing energy (the last term). This analogy led Staroverov and Scuseria to define a more general functional, called ζ-HFB (or equivalently, ζ-CHF), by setting a scaling factor ζ in front of the pairing energy:

$$
\begin{aligned}
E_{\zeta-\mathrm{HFB}} &= 2\sum_i n_i h_{ii} + \sum_{ij} n_i n_j \left(2\langle ij|ij\rangle - \langle ij|ji\rangle\right) \\
&\quad -\zeta \sum_{ij} \sqrt{n_i(1-n_i)n_j(1-n_j)}\langle ij|ji\rangle.
\end{aligned} \tag{2.56}
$$

Hence, when $\zeta = -1$, Eq.(2.56) reduces to HFB, while if $\zeta = 1$ it becomes identical to CHF if $\zeta = 1$, which we will refer to as 1HFB, hereafter. We remind the reader that if n_i is either 1 or 0, Eq.(2.56) also immediately reduces to the HF energy expression, Eq.(2.8).

The ability of the ζ-HFB scheme to estimate strong correlations in atomic and molecular systems has been well demonstrated [12]. For example, with $\zeta = 1.12$, ζ-HFB significantly improved the correlation energies in the Be series compared to CHF ($\zeta = 1$). In dissociation curves of a few diatomic molecules (FH, F_2, and N_2), ζ-HFB with $\zeta = 0.6 \sim 0.8$ predicted a certain amount of strong correlation as the internuculear distance R increases, while it reproduced the RHF energy near equilibrium, where only the weak correlation effect is important and strong correlation should not be expected at all. On the other hand, CHF almost always yields unbound dissociation curves, as it captures too much strong correlation at large R. This result clearly indicates ζ-HFB can correctly pick up only strong correlation, but not weak correlation.

An interesting fact about ζ-HFB and CHF is that both methods do all of this over the correct spatial- and spin-symmetry surfaces. In other words, orbitals do

not break the spatial symmetry of the system, and the α and β electron densities are equivalent each other for singlet systems, as they should be. On the other hand, as we have mentioned, this feature is not observed in UHF. UHF always breaks spatial- and spin-symmetry in the dissociation limit. However, while ζ-HFB showed a significant improvement over CHF, a major concern is that ζ is system-dependent. This means that ζ-HFB uses different attractive interactions for different systems in order to obtain the best results. Hence, our motivation provides a corrected scheme which utilizes a system-independent attractive potential, *i.e.*, with a constant ζ.

Chapter 3

Constrained Pairing Mean-Field Theory

3.1 Analysis of 1HFB

Here we develop a new method based on ζ-HFB. Our intent is to break the electron number symmetry, in order to include pair correlation with an attractive potential, which we expect to contain a certain amount of strong correlation [2–6]. HFB is connected to the N-electron multireference antisymmetrized geminal power wave function (AGP) through the number symmetry projection [12, 23], and hence, the appearance of strong correlation should not be unexpected. As we will see for the simple case of H_2, ζ-HFB does not include correlations for most molecules near equilibrium, yielding the non-correlated RHF solution. As the system dissociates, 1HFB (CHF) adds strong (left-right) correlations, but significantly overcorrelates at large internuclear separations in almost all cases. This behavior was considered a serious drawback [24], and the model lost attention. To address this problem, ζ-HFB has the optimal attractive interaction $-\zeta/r$, but the strength of the interaction varies by system, which prohibits its wide-spread use in electronic structure calculations. In this chapter, however, we show that by incorporating a few simple constraints, a mixture of HF and 1HFB methods for inactive and active sets of orbitals, respectively, can overcome the overcorrelation problem and yield an accurate mean-field approximation for strong correlations. To further analyze ζ-HFB, below we studied it from various perspectives.

Assuming that each dissociated fragment is uncorrelated and requires a method that includes only strong correlation between fragments, the desired energy at a dissociation limit is the sum of the restricted open-shell HF (ROHF) fragments, since ROHF preserves both spatial- and spin-symmetries while not containing any correlation energy. Thus, the optimal value, ζ_{opt}, may be defined for each molecular system, where ζ_{opt}-HFB reproduces the ROHF fragments at a dissociation limit. This ζ_{opt} is found to be always less than 1 for all the tested molecular systems, with the exception of H_2 with a minimal basis [12, 25].

The H_2 molecule is a special example because it can be dissociated *exactly* with $\zeta = 1$ if a minimal basis set is used *i.e.*, 2-electron-2-orbital. This is demonstrated in Figure 3.1, where we used an STO-6G basis. This is the *only* case where 1HFB reproduces the correct strong correlation at the dissociation limit of a molecule. For example, when a larger 6-31G basis is used, 1HFB overcorrelates already at 3Å. ζ_{opt} in this case is 0.9785. Therefore, we can conclude ζ_{opt} is *basis set dependent*.
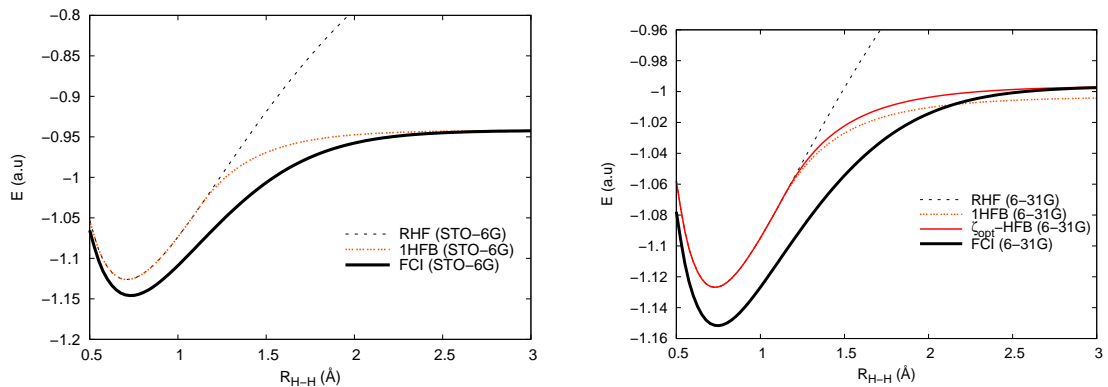
Figure 3.1 : Dissociation curves of the H_2 molecule with an STO-6G basis (Left) and a 6-31G basis (Right).

Furthermore, since the correct dissociation limit of the H_2 molecule is the ROHF fragments of two H atoms, which are equivalent to the FCI fragments in this case,

Table 3.1 : Natural occupations at the dissociation limit of H_2.

Occupations	1HFB (STO-6G)	1HFB (6-31G)	ζ_{opt}-HFB (6-31G)	FCI (6-31G)
$n_{\sigma_{1s}}$	0.500000	0.497934	0.498055	0.500000
$n_{\sigma_{1s}^*}$	0.500000	0.497934	0.498055	0.500000
$n_{\sigma_{2s}}$	-	0.002066	0.001945	0.000000
$n_{\sigma_{2s}^*}$	-	0.002066	0.001945	0.000000

the natural occupations should be 0.5 in the bonding and anti-bonding orbitals for both α and β spins, while other high lying orbitals should have exactly 0 occupations. 1HFB with a minimum STO-6G basis (2 orbitals) actually captures this feature. On the other hand, 1HFB with a 6-31G basis (4 orbitals) gives different occupations, as can be seen in Table 3.1. Although ζ_{opt}-HFB reproduces the correct energy of FCI at dissociation, it is immediately evident that the resulting electron density is incorrect because the eigenvalues of the ζ_{opt}-HFB density matrix (natural occupations) are different from the correct values of FCI.

These results indicate that the parametrized ζ_{opt}-HFB model should not be pursued even though it yields the correct limit energetically. Instead, we consider the 1HFB energy expression Eq.(2.56) with $\zeta = 1$ at the dissociation of H_2. Substituting the correct natural occupations and orbitals produced by FCI (or ROHF for each fragment), Eq.(2.56) can indeed reproduce the correct energy:

$$E_{1\text{HFB}} = h_{\sigma_{1s}\sigma_{1s}} + h_{\sigma_{1s}^*\sigma_{1s}^*} + J_{\sigma_{1s}\sigma_{1s}^*} - K_{\sigma_{1s}\sigma_{1s}^*} \tag{3.1}$$

where $J_{\sigma_{1s}\sigma_{1s}^*} = \langle \psi_{\sigma_{1s}}\psi_{\sigma_{1s}^*} | \psi_{\sigma_{1s}}\psi_{\sigma_{1s}^*} \rangle$ and $K_{\sigma_{1s}\sigma_{1s}^*} = \langle \psi_{\sigma_{1s}}\psi_{\sigma_{1s}^*} | \psi_{\sigma_{1s}^*}\psi_{\sigma_{1s}} \rangle$ are Coulomb and exchange integrals. Given that the last two terms cancel out each other at dissociation, $J_{\sigma_{1s}\sigma_{1s}^*} = K_{\sigma_{1s}\sigma_{1s}^*}$, this expression correctly reproduces the energy of two

isolated hydrogen atoms. Importantly, this result holds for any basis set.

This simple observation led us to seek a theoretical model that can remove small occupations occurring in irrelevant orbitals (σ_{2s}, σ_{2s}^*) in the 1HFB functional to alleviate its overcorrelation problem. Constraining its natural occupations exactly corresponds to removing unnecessary pairing interactions between certain "inactive" orbitals. Such a method is, hereafter, referred to as constrained pairing mean-field theory (CPMFT) [2–6], and will be explained in the next section.

3.2 General Theory

3.2.1 The CPMFT energy functional

In CPMFT, we partition the orbitals into core, active, and virtual spaces. Only the orbitals in the active space are subject to a Bogoliubov transformation with an attractive pairing interaction; the core and virtual orbitals are treated by HF. In this way, we "constrain" pairing interactions exclusively to the active space. To achieve this goal, we first introduce additional constraints to divide orbital spaces into fully occupied (core, $n = 1$), fractionally occupied (active), and virtual ($n = 0$). These constraints can be expressed as

$$\sum_{c}^{M_c} n_c = N_C, \tag{3.2}$$

$$\sum_{c}^{M_c} (n_c^2 - n_c) = 0, \tag{3.3}$$

$$\sum_{a}^{M_a} n_a = N_A, \tag{3.4}$$

$$\sum_{v}^{M_v} n_v = 0, \tag{3.5}$$

$$\sum_v^{M_v} (n_v^2 - n_v) = 0, \tag{3.6}$$

where c, a, and v run over the core, active, and virtual orbitals, respectively, and $2N_C$ and $2N_A$ are the number of electrons in the core and active spaces (recall that we will work in the closed-shell, spatial orbital basis). M_c, M_a, and M_v are the number of core, active, and virtual orbitals. Because $M_c = N_c$ always, the only possible solution to Eq.(3.3) is that all $n_c = 1$, thus Eq.(3.4) is automatically satisfied and redundant (core orbitals are singly occupied in $\mathbf{P}$). Similarly, Eq.(3.5) implies that all $n_v = 0$, thus Eq.(3.6) is redundant, too. On the other hand, $M_a > N_A$, in general. In summary, the resulting Lagrangian including all constraints becomes

$$\mathcal{L}^{\text{CPMFT}} = E^{\text{CPMFT}} + \mu_C \left(\sum_c n_c - N_C \right) + \mu_V \sum_v n_v + \mu_A \left(\sum_a n_a - N_A \right), \tag{3.7}$$

where the different μ's are the chemical potentials (Lagrange multipliers) constraining the electron numbers in each space. E^{CPMFT} above is nothing but the 1HFB energy expression,

$$E^{\text{CPMFT}} = 2 \sum_{ij} h_{ij} P_{ij} + \sum_{ijkl} \left(2\langle ik|jl \rangle - \langle ik|lj \rangle \right) P_{ij} P_{kl} - \sum_{ijkl} \langle ij|kl \rangle K_{ij} K_{kl}. \tag{3.8}$$

The mathematical solution to the CPMFT constrained problem consists in setting $\mu_C = -\infty$ and $\mu_V = \infty$. The physical meaning of this is that μ_C (μ_V) pulls down (pushes up) the energies of core (virtual) orbitals so that these levels get fully occupied (empty). In practice, we may set the chemical potential of the inactive spaces $\mu_V = -\mu_C$, and gradually increase μ_V smoothly to convergence. A slow μ increase allows for efficient mixing between spaces until orbital optimization is achieved, as dictated by the variational procedure. In Figure 3.2, we present a pictorial description of the CPMFT orbitals and optimization process. The core and virtual spaces are treated by RHF, whereas the active space is treated by 1HFB.
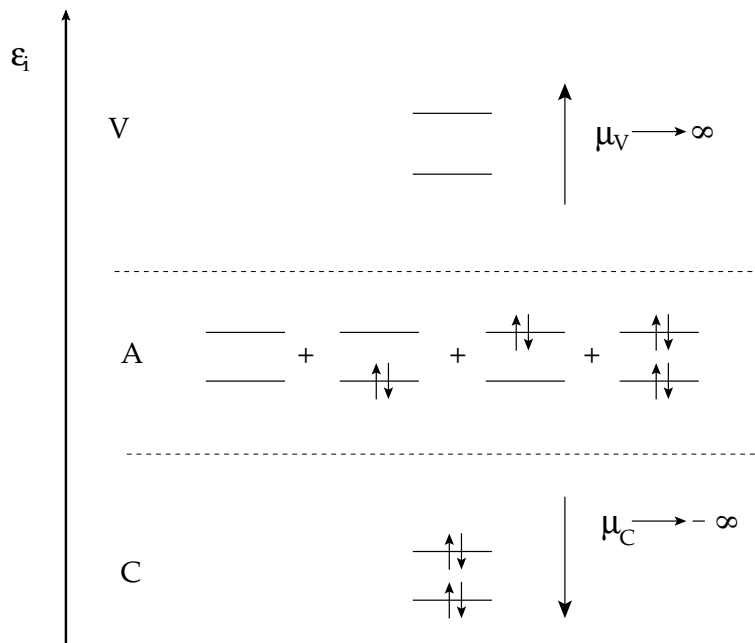
Figure 3.2 : Description of each space in CPMFT.

3.2.2 Properties of CPMFT

We note that Eq.(3.8) just flips the sign of the last term, the pairing energy, in HFB (Eq.(2.47)). Changing the sign of the pairing term changes the sign of $\boldsymbol{\Delta}$ so that the CPMFT Hamiltonian is

$$\mathcal{H}_{\text{CPMFT}} = \begin{pmatrix} \mathbf{F}^{\text{cs}} + \boldsymbol{\mu} & -\boldsymbol{\Delta} \\ -\boldsymbol{\Delta} & -\mathbf{F}^{\text{cs}} - \boldsymbol{\mu} \end{pmatrix}, \tag{3.9}$$

where each matrix is the same as in RHFB, Eqs.(2.49a, 2.49b),

$$F_{ij}^{\text{cs}} = h_{ij} + \sum_{kl}(2\langle ik|jl\rangle - \langle ik|lj\rangle)P_{kl} = f_{ij}, \tag{3.10a}$$

$$\Delta_{ij} = \sum_{kl}\langle ij|kl\rangle K_{kl}. \tag{3.10b}$$

Here, cs denotes "closed-shell" and $\mathbf{f}$ is the standard Fock matrix appearing in RHF. $\boldsymbol{\mu}$ is in the NO basis diagonal and consists of μ_C, μ_A, and μ_V. Except for the negative

sign in front of $\boldsymbol{\Delta}$ in Eq.(3.9), CPMFT follows the same procedure as in HFB, noting that

$$\mathbf{PK} - \mathbf{KP} = \mathbf{0} \tag{3.11a}$$

$$\mathbf{P} - \mathbf{P}^2 = \mathbf{K}^2, \tag{3.11b}$$

which arises from the fact that the generalized density matrix $\mathbf{R}$ is idempotent. However, changing the sign of the pairing energy and the pairing matrix severs the connection between the HFB wave function $|\Phi_{\mathrm{HFB}}\rangle$ and the CPMFT energy.

We can, indeed, view the CPMFT energy as the expectation value of a model Hamiltonian with respect to a particle-number violating determinant:

$$H_0|\Phi\rangle = E_{\mathrm{CPMFT}}|\Phi\rangle, \tag{3.12a}$$

$$H_0 = \sum_{ij}\left\{(h_{ij} + F_{ij}^{\mathrm{cs}})a_i^\dagger a_j - \frac{1}{2}\Delta_{ij}a_i^\dagger a_j^\dagger - \frac{1}{2}\Delta_{ij}^\star a_i a_j\right\}. \tag{3.12b}$$

This quadratic model Hamiltonian, however, is *not* the mean-field of the physical Hamiltonian with respect to a quasiparticle determinant, as we have introduced an effective attractive particle interaction. We can also interpret the CPMFT energy as a hybrid of HF and HFB where HF uses $2/r_{12}$ as the electron-electron repulsion operator and HFB uses $-1/r_{12}$. In order to see this, consider the Coulomb operator split into two parts:

$$\frac{1}{r_{12}} = \frac{1+\zeta}{r_{12}} - \frac{\zeta}{r_{12}}. \tag{3.13}$$

Since the RHF energy expression is given by

$$E^{\mathrm{RHF}} = 2\sum_{ij}h_{ij}P_{ij} + \sum_{ijkl}\left(2\langle ik|jl\rangle - \langle ik|lj\rangle\right)P_{ij}P_{kl}, \tag{3.14}$$

applying the first and second terms of Eq.(3.13) to E^{RHF} and E^{HFB} (Eq.(2.37)), respectively, gives

$$E^{\mathrm{CPMFT}} = (1+\zeta)E^{\mathrm{RHF}} - \zeta E^{\mathrm{HFB}} \tag{3.15}$$

with $\zeta = 1$. This argument of course holds for 1HFB—although in 1HFB all the orbitals are subject to the attractive pairing interaction while in CPMFT this effect is limited to the chosen active space.

Nevertheless, we have a fruitful alternative viewpoint, which is to envision the CPMFT energy expression of Eq.(3.8) as defining a model two-particle density matrix $\boldsymbol{\Gamma}$ such that the energy in the spin-orbital basis is

$$E_{\text{CPMFT}} = \text{Tr}(\mathbf{h}\,\boldsymbol{\gamma}) + \text{Tr}(\mathbf{v}\,\boldsymbol{\Gamma}_{\text{CPMFT}}) \tag{3.16}$$

where $\mathbf{v}$ is the two-particle part of the Hamiltonian, and $\mathbf{h}$ is the one-particle part. In terms of spin-orbitals, we have

$$\Gamma_{pq}^{rs} = \frac{1}{2}\left(\gamma_p^r\gamma_q^s - \gamma_p^s\gamma_q^r - \kappa_{pq}\kappa^{rs}\right), \tag{3.17}$$

with lower (upper) indices corresponding to bra (ket) indices. As we have already seen, the first two terms in this model 2PDM correspond to HF,hole-hole, whereas the last term introduces strong correlation via $\mathbf{K}$, which is a measure of nonidempotency for $\mathbf{P}$. This last term is an important quantity in the cumulant decomposition of density matrices [26], but in our work appears naturally from the idempotency of the quasi-particle density matrix $\mathbf{R}$. As can be seen, Eq.(3.17) includes particle-particle and hole-hole but not particle-hole correlation terms. As shown below, this simple ansatz describes strong correlations very accurately. We suggest that the cumulant term offers a compelling definition of what we intuitively understand as strong correlation.

If we use this model two-particle density matrix to *define* expectation values of two-particle operators, then, as derived in Appendix A, it becomes apparent that CPMFT has no particle number fluctuations, but does have spin contamination. In making this choice, we are inevitably working with a density matrix functional and are effectively utilizing some form of a statistical ensemble theory.

3.2.3 Alternative constraints

In practice, the CPMFT scheme presented above works reasonably well, but requires numerous SCF iterations with small μ_C and μ_V to allow for sufficient relaxation of orbitals between the active and inactive spaces. However, we identified an alternative procedure which constrains $\kappa_i = \sqrt{n_i - n_i^2}$ instead of working with n_i. In other words, we constrain κ to be zero for the core and virtual spaces. This allows us to write the following Lagrangian:

$$\mathcal{L}^{\mathrm{CPMFT}} = E^{\mathrm{CPMFT}} + 2 \sum_{ij} \tilde{\mu}_{ij} K_{ij} + \mu_A \left(\sum_a n_a - N_A \right), \qquad (3.18)$$

where i and j run over the core and virtual orbitals, and for the active space the number of electrons are yet to be constrained. The factor of 2 arises because the equation is written in the spatial orbital basis. By transforming Eq.(3.7) to Eq.(3.18), we have removed μ_C and μ_V which approach infinity, which is not computationally favorable. Meanwhile, we have gained in Eq.(3.18) a Lagrange multiplier matrix $\tilde{\boldsymbol{\mu}}$ which is to be determined. Since now we have more Lagrange multipliers in $\tilde{\boldsymbol{\mu}}$ than Eq.(3.7) (μ_C and μ_V), it might appear that we have just made the equations more complicated. However, in Appendix B, we derive the exact condition that $\tilde{\boldsymbol{\mu}}$ has to satisfy *at convergence* of the CPMFT equation, which is simply

$$\tilde{\mu}_{ij} = \begin{cases} \Delta_{ij} & \text{for } i, j \in c \text{ or } i, j \in v \\ 0 & \text{otherwise.} \end{cases} \qquad (3.19)$$

This means $\tilde{\boldsymbol{\Delta}}$ defined as the derivative of $\mathcal{L}$ with respect to K at each SCF iteration is given by

$$\tilde{\boldsymbol{\Delta}} = -\boldsymbol{\Delta} + \tilde{\boldsymbol{\mu}} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\Delta}^{ca} & \boldsymbol{\Delta}^{cv} \\ \boldsymbol{\Delta}^{ac} & \boldsymbol{\Delta}^{aa} & \boldsymbol{\Delta}^{av} \\ \boldsymbol{\Delta}^{vc} & \boldsymbol{\Delta}^{va} & \mathbf{0} \end{pmatrix}. \qquad (3.20)$$

We should note that the ca, ac, va, and av blocks of $\tilde{\boldsymbol{\Delta}}$ contribute to the orbital rotation between the active and the inactive space, which is also done in Eq.(3.7)

Table 3.2 : The CPMFT results on dissociated H_2 with different constraining schemes.

Schemes	Energy	Cycles
Eq.(3.7)	-0.999 884	52
Eq.(3.18)	-0.999 891	19

via small μ_C and μ_V. The aa block is required for the orbital rotation among the active space. On the other hand, $\boldsymbol{\Delta}^{cv}$ and $\boldsymbol{\Delta}^{vc}$ do not play a role: they are related to core-virtual relaxation, but this can be done via the Fock part only, since it is the HF space. This exactly corresponds to μ_C and μ_V being infinity at convergence. Hence, $\boldsymbol{\Delta}^{cv}$ and $\boldsymbol{\Delta}^{vc}$ are found to be *arbitrary*, *i.e.*, they can be set to zero.

The above formulation of the CPMFT constraints allows faster and numerically more robust convergence in CPMFT than Eq.(3.7) because we do not deal with large μ_C and μ_V in actual calculations. Both constraint schemes should not change the final results, *e.g.*, the energy and NOs, if the orbital rotation is appropriately achieved. Table 3.2 shows the results of the H_2 dissociation limit, computed with both schemes. With Eq.(3.7), it takes more SCF cycles than with the new formula, Eq.(3.18). Furthermore, its energy is slightly higher, because the active orbitals and the inactive orbitals are not efficiently mixed during the calculation: the efficacy of an orbital relaxation with Eq.(3.7) depends on the increase rate of μ_C (and μ_V), and in order to relax orbitals completely, one would need a quasi-static process (infinite numbers of cycles).

3.2.4   Results on some dissociation curves

As a paradigmatic example of strong correlation, we present results for the dissociation potential of the H_2 and N_2 in Figure 3.3. The latter is an especially challenging case because it requires up to six-electron excitations to yield the correct curve. In both cases, RHF includes no correlations, failing miserably as R becomes larger. On the other hand, 1HFB acquires too much correlation, as mentioned earlier, producing a physically incorrect description. Unrestricted HF improves over RHF, but the density is symmetry-broken: for example, at the dissociated N_2, three α electrons localize on one N atom, and three β electrons on the other, with a similar phenomenon occurring in the H_2 case. Complete active space self-consistent field with six-electron-six-orbitals (CASSCF(6,6)) calculations, which include all excitations for six electrons within six active orbitals, also yield a correct potential curve for N_2 including a part of weak and strong correlations near equilibrium and only strong correlations toward dissociation. CPMFT(6,6), meaning six-electron-six-orbitals in the active space in a similar way as CASSCF, gives a very accurate description of the dissociation of the N_2 molecule (so does CPMFT(2,2) for the H_2 molecule). It accomplishes this feat while preserving spatial and spin symmetries. Note that the behavior of the CPMFT curves differs from that of UHF, and we will discuss the connection between these two methods later on.

Another interesting example is the potential energy curve for the ethylene torsion, as presented in Figure 3.4. At a dihedral angle of 90°, there is an exact degeneracy between the π^2 and π^{*2} configurations that CPMFT(2,2) handles very well. It is well known that single reference methods such as RHF have a cusp at 90°, yielding an unphysically large barrier height. This is similar to H_2 dissociation, but on a different coordinate.
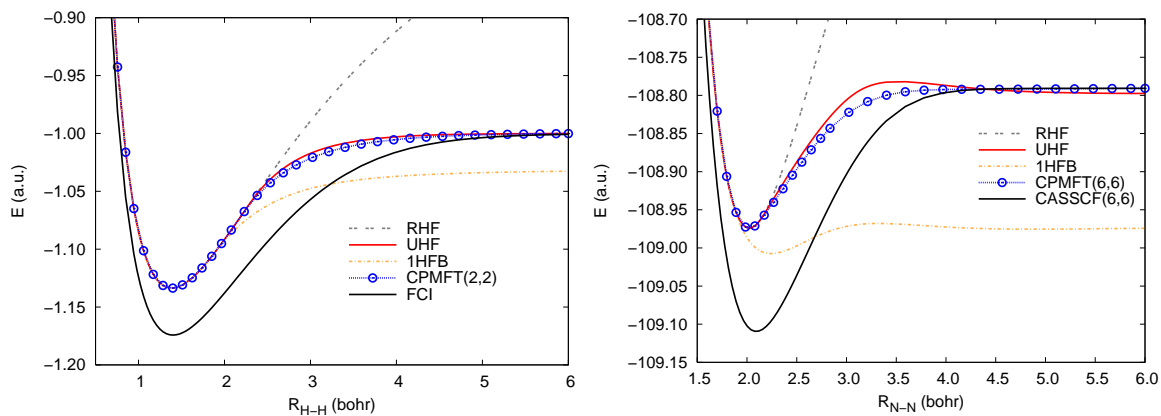
Figure 3.3 : Potential energy curves of H_2 and N_2 calculated with Gaussian cc-pV5Z and 6-311++G** basis sets, respectively.
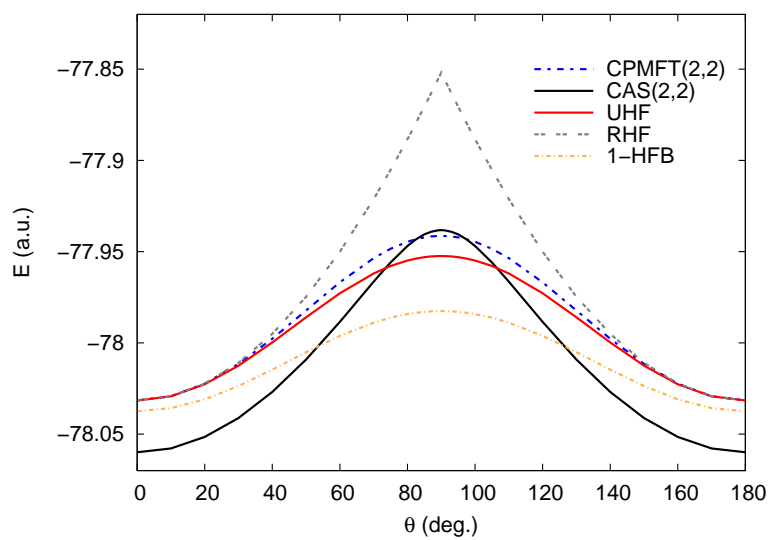
Figure 3.4 : Potential energy curves of the distorted ethylene molecule calculated with a 6-31G* basis. θ is the dihedral angle between the CH_2 planes.

3.3 Generalization to nondegenerate cases

3.3.1 A two level model system

To better understand how CPMFT accomplishes its feats, we will discuss in detail a simple two level model system: a closed-shell electron pair in two orbitals, $e.g.$, H_2 in a minimum basis. For this case, using our symmetry adapted formalism, we have only two occupations $n_1 + n_2 = 1$. It follows then that $\kappa_1^2 = n_1(1 - n_1) = n_2(1 - n_2) = \kappa_2^2$. Defining

$$W = 2J_{11} + 2J_{22} + 4K_{12} - 4J_{12} \tag{3.21}$$

with $J_{ij} = \langle \psi_i \psi_j | \psi_i \psi_j \rangle$ and $K_{ij} = \langle \psi_i \psi_j | \psi_j \psi_i \rangle$ and the diagonal elements of the full CI Hamiltonian:

$$E_1 = 2h_{11} + J_{11} \tag{3.22}$$

$$E_2 = 2h_{22} + J_{22}, \tag{3.23}$$

the energy expression Eq.(2.37) becomes:

$$E(n_1) = E_2 + n_1(E_1 - E_2 - W) + n_1^2 W, \tag{3.24}$$

or equivalently

$$E(n_2) = E_1 + n_2(E_2 - E_1 - W) + n_2^2 W. \tag{3.25}$$

For $n_1 = 1$ and $n_2 = 0$ (RHF case), either one of these two expressions yields

$$E^{\text{RHF}} = E_1. \tag{3.26}$$

For clarity, we emphasize again that $n_1 = 1$ here means $n_{1\alpha} = n_{1\beta} = 1$, so that there are a total of two electrons being considered.

We can calculate $\frac{dE}{dn_1} = 0$ to determine the critical n_0 where this energy expression is minimum:

$$n_0 = \frac{1}{2} + \frac{E_2 - E_1}{2W}. \tag{3.27}$$

The CPMFT solution appears when $n_0 < 1$. This condition reads

$$\frac{E_2 - E_1}{W} < 1. \tag{3.28}$$

Using that in RHF, $\varepsilon_{\sigma_g} = h_{11} + J_{11}$ and $\varepsilon_{\sigma_u} = h_{22} + 2J_{12} - K_{12}$, we can rewrite this expression in terms of the usual molecular orbital energies as

$$\varepsilon_{\sigma_u} - \varepsilon_{\sigma_g} - \frac{J_{11} + J_{22}}{2} - K_{12} < 0, \tag{3.29}$$

which can be compared with the RHF triplet instability condition [27, 28]:

$$\varepsilon_{\sigma_u} - \varepsilon_{\sigma_g} - J_{12} - K_{12} < 0. \tag{3.30}$$

The CPMFT solution appears when the energy gap $E_2 - E_1$ closes and becomes smaller than W. Depending on how J_{12} compares to $(J_{11} + J_{22})/2$, the CPMFT solution may appear before or after the UHF solution (Coulson-Fischer point).

It is interesting to note that for the regular HFB case (repulsive pairing interaction of $1/r$ and positive sign in pairing energy), the energy expressions Eq.(3.24) and Eq.(3.25) are still valid with $W = -4J_{12}$. This means, however, that $E_1 - E_2 - W$ is always positive, thus the RHF solution always has a lower energy than HFB, *i.e.*, $n_2 = 0$ in Eq.(3.25). Also note that $\frac{d^2 E}{dn_1^2} = 2W$, so the critical point n_0 is guaranteed to be a minimum only for positive W, as in CPMFT.

At dissociation of a strongly correlated pair (also referred to as "fully entangled"), we want $n_1 = n_2 = \frac{1}{2}$. From Eq.(3.27), $n_0 = \frac{1}{2}$ if $E_1 = E_2$, which yields

$$2h_{11} + J_{11} - 2h_{22} - J_{22} = 0. \tag{3.31}$$

Using that for $n_1 = n_2 = \frac{1}{2}$,

$$f_{11} = h_{11} + \frac{J_{11}}{2} + J_{12} - \frac{K_{12}}{2} \tag{3.32}$$

$$f_{22} = h_{22} + \frac{J_{22}}{2} + J_{12} - \frac{K_{12}}{2} \tag{3.33}$$

we obtain that the $n_0 = \frac{1}{2}$ condition is fulfilled if

$$f_{11} - f_{22} = 0, \tag{3.34}$$

which is always true for dissociation to degenerate orbitals. From Eq.(3.24), we get that the energy is

$$E(n = \frac{1}{2}) = h_{11} + h_{22} + J_{12} - K_{12}. \tag{3.35}$$

In the dissociation limit $J_{12} = K_{12}$, the orbitals become degenerate, and $h_{11} = h_{22}$ is the energy of an isolated hydrogen atom. Thus, CPMFT can dissociate this electron pair to the correct energy and occupations of $\frac{1}{2}$. In a large AO basis, the variational procedure in CPMFT rotates the orbitals between active and inactive spaces yielding the correct NOs for describing dissociation in such basis.

3.3.2 Dissociation to non-degenerate orbitals

The dissociation of an electron pair to non-degenerate orbitals, as for example in XH (X=Li, B, or F), is more challenging than the degenerate case, $e.g.$, H_2. In the latter case, we showed above that the predicted CPMFT occupations are exactly $\frac{1}{2}$. In the non-degenerate case, however, the resulting occupations at dissociation are different from $\frac{1}{2}$, an effect that we will refer to as "spilling." In essence, the variational principle drives electron density to the lower energy orbital. This is evident from Eq.(3.34) when f_{11} is different from f_{22}. As an example, we present in Figure 3.5 the plot of the

CPMFT(2,2) energy as a function of occupations for the BH molecule with a 6-31G basis set. The resulting occupations at dissociation are $n_1 = 0.56$ and $n_2 = 0.44$, instead of $\frac{1}{2}$ and $\frac{1}{2}$. In other words, the minimum of the CPMFT functional does not occur at the desired value of $\frac{1}{2}$ unless the dissociated orbitals become degenerate, i.e., $f_{11} = f_{22}$. The spilling of occupations results in a small overshooting of the correct dissociation energy, as shown in Figure 3.5 where CPMFT(2,2) is compared to CASSCF(2,2), 1HFB, and RHF. Note, however, the substantial improvement from 1HFB to CPMFT(2,2), despite spilling.
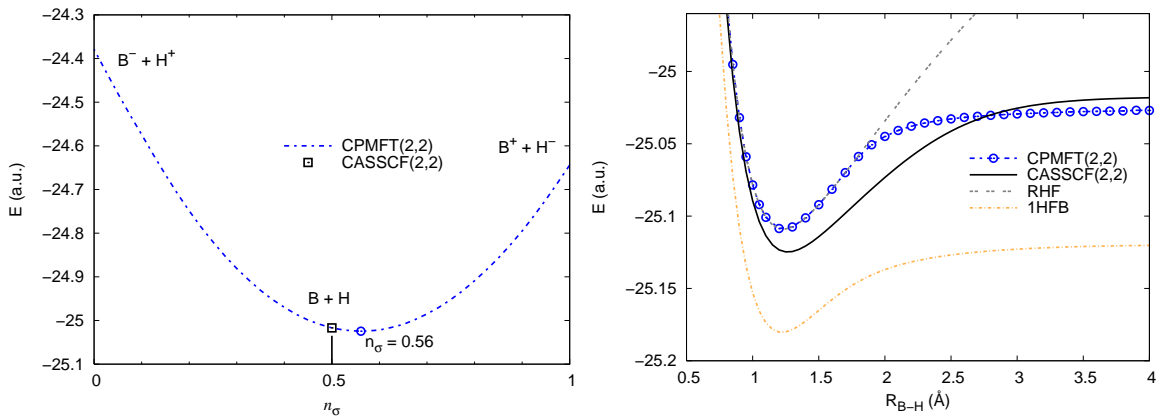
Figure 3.5 : CPMFT(2,2) energy as a function of occupation numbers (Left) and dissociation curves (Right) of the BH molecule calculated with a 6-31G basis and different methods.

Before proposing a solution to the non-degenerate dissociation problem, a small digression is in order. First, we need to more rigorously define the meaning of *dissociation* within the context of our model. We define the dissociation state of a molecule as a collection of atoms (or molecular fragments) in specific spin states with orbital occupations limited to 0, 1, and $\frac{1}{2}$, corresponding to empty, fully-occupied, and half-occupied (fully-entangled) orbitals. The choice of these states determines the space and spin symmetry of the molecular state to be studied. Most importantly,

we want to emphasize that no other occupations (except for 0, $\frac{1}{2}$, and 1) are allowed in CPMFT at dissociation. Natural orbitals with $\frac{1}{2}$ occupation at dissociation are strongly correlated and form and break chemical bonds near equilibrium. We also note that if the fragments at dissociation are chosen as closed-shell (*i.e.*, no fully entangled electrons) and are treated with RHF, then the active space is null, and CPMFT reduces to RHF, which is size consistent in this case.[*] If the fragments are closed-shell and treated with 1HFB, there are no inter-fragment, strongly correlated electrons (there may be some in the fragment itself) and in this case, 1HFB is size consistent too.

Strictly speaking, NOs have no energy associated with them. At dissociation, however, NOs in CPMFT become strongly localized on fragments, which is the correct behavior also observed in CASSCF. Furthermore, when written in the NO basis, the active block of the Fock matrix becomes diagonal as the molecule approaches dissociation. In the degenerate case, the off-diagonal terms are zero because of symmetry; in the non-degenerate case, the NOs localize and they do not overlap, also yielding zero off-diagonal active terms. At CPMFT convergence, with the infinite inactive chemical potentials introduced in Section 3.2.1, the Fock matrix is *effectively* diagonal because the inactive/active blocks become negligible and can be set to zero as the diagonal of the inactive block approaches $\pm\infty$. Thus, NOs effectively become ROHF orbitals, and we can associate their energies with those of fragment ROHF orbitals. The Fock matrix at dissociation is depicted in Figure 3.6.

A remarkable feature of the CPMFT energy functional is that it yields the correct dissociation energy if fed the correct occupation numbers ($\frac{1}{2}$) for the strongly

---

[*]A method is size consistent if the total energy of fragments infinitely separated from each other is equivalent to the sum of the energies of each fragment.

$$\begin{pmatrix}
f_{11}^{cc} & & 0 & f_{11}^{ca} & \cdots & f_{m1}^{ca} & & & \\
& \ddots & & \vdots & \ddots & \vdots & & 0 & \\
0 & & f_{mm}^{cc} & f_{1m}^{ca} & \cdots & f_{mm}^{ca} & & & \\
\hline
f_{11}^{ca} & \cdots & f_{1m}^{ca} & f_{11}^{aa} & & 0 & f_{11}^{av} & \cdots & f_{m1}^{av} \\
\vdots & \ddots & \vdots & & \ddots & & \vdots & \ddots & \vdots \\
f_{m1}^{ca} & \cdots & f_{mm}^{ca} & 0 & & f_{mm}^{aa} & f_{1m}^{av} & \cdots & f_{mm}^{av} \\
\hline
& & & f_{11}^{av} & \cdots & f_{1m}^{av} & f_{11}^{vv} & & 0 \\
& 0 & & \vdots & \ddots & \vdots & & \ddots & \\
& & & f_{m1}^{av} & \cdots & f_{mm}^{av} & 0 & & f_{mm}^{vv}
\end{pmatrix}$$

Figure 3.6 : Fock matrix of CPMFT at dissociation with superscripts c, a, and v representing core, active, and virtual orbitals, respectively. The inactive (core and virtual) blocks can always be diagonalized by an orbital rotation, and the inactive/active blocks become negligible at convergence. Note that the active block is diagonal at dissociation (but, in general, non-diagonal outside dissociation).

correlated orbitals. In the two-orbital case, this is easily seen from Eq.(3.35), which yields the correct energy at dissociation for $n = \frac{1}{2}$ even if orbitals 1 and 2 are not degenerate. However, as already pointed out, the minimum does not occur at $n = \frac{1}{2}$ unless orbitals 1 and 2 are degenerate. The variational principle in CPMFT produces spilling in the non-degenerate case. This is incorrect and will be addressed in the following subsection by adding constraints to CPMFT that will be referred to as "asymptotic" constraints for reasons that should become clear below.

3.3.3 Dissociation of polyatomic molecules

Even for multiply bonded diatomic molecules, CPMFT dissociates exactly to ROHF atoms, as shown for N_2 in Section 3.2.4. Here, we present a simple argument showing that CPMFT can dissociate *any* polyatomic molecule exactly to ROHF fragments or atoms. To achieve this goal, the active space of CPMFT should match the open-shell

orbital spaces of the ROHF fragments. We begin by recognizing that at dissociation, the variational principle drives the CPMFT NOs to become identical with ROHF orbitals, as is the case for CASSCF. The core orbitals of the two methods are thus the same. The core orbitals are doubly occupied in the ROHF fragments and are inactive in CPMFT, thus their energy contribution (given by an RHF expression) is identical. If there were total energy differences between CPMFT and the sum of ROHF fragments, they would originate in the active space. At dissociation, our model guarantees that all active CPMFT occupations are $n_i = \kappa_i = \frac{1}{2}$ (for the non-degenerate case; spilling is addressed in the next section). Using that $J_{ii} = K_{ii}$, the CPMFT energy expression, Eq.(3.8), becomes:

$$E^{\text{CPMFT}} = \sum_i^{\text{act}} h_{ii} + \sum_{i>j}^{\text{act}} (J_{ij} - K_{ij}). \tag{3.36}$$

From the ROHF perspective, this expression is identical to the contribution of high-spin open-shells to the total energy. This proves that the CPMFT energy is identical to the sum of fragment ROHF energies provided that (1) the active space in CPMFT corresponds with the fragment ROHF open-shell orbitals, (2) the CPMFT NOs converge to fragment ROHF MOs, and (3) the occupation of fully-entangled orbitals at dissociation is exactly $\frac{1}{2}$. The latter two conditions are intrinsic and built into our model (*vide infra*). In summary, CPMFT will always dissociate any polyatomic molecule exactly to ROHF fragments if the active space is chosen adequately.

3.3.4 Asymptotic constraints

In order to solve the spilling problem in dissociations to non-degenerate orbitals, we modify the CPMFT energy functional and introduce a constraint that, *at dissociation*, does not change the energy but does change the effective Hamiltonian such that

the new minimum of the CPMFT functional occurs at occupations of exactly $\frac{1}{2}$ for strongly correlated electrons. In other words, we mathematically reformulate the problem such that the non-degenerate orbitals look degenerate to the functional. To achieve this, we add to the CPMFT energy expression the following constraint:

$$\phi = 2 \operatorname{Tr}_a \left[\mathbf{U}\mathbf{G}(\mathbf{P})\right], \tag{3.37}$$

where $\mathbf{U}$ is a matrix of Lagrange multipliers, $\mathbf{G}(\mathbf{P})$ is a polynomial of the density matrix designed to satisfy a number of conditions explained below, and the a index in the trace means that we restrict the sum to active orbitals, as $\mathbf{U}$ is defined on the active space only. This constraint, via Eq.(3.10a), contributes the following terms to the Fock matrix:

$$
\begin{aligned}
u_{ij} &= \frac{1}{2}\frac{\partial \phi}{\partial P_{ji}} = \sum_{kl} U_{kl}\frac{\partial G_{kl}}{\partial P_{ji}} \\
&= U_{ij}G'_{ij}.
\end{aligned} \tag{3.38}
$$

Note that there is no sum in the last expression defining G'_{ij}. Mathematically, the conditions on $\mathbf{G}$ and $\mathbf{G}'$ are simple to express in the NO basis. They guarantee that ϕ is zero at dissociation $(n = \frac{1}{2})$ and in the RHF region $(n = 0, 1)$, and they modify the equation determining the n_0 minimum. The constraints are:

$$G_{ij} = 0 \quad \text{if} \quad \{n_i, n_j\} = \{0, \frac{1}{2}, 1\}, \tag{3.39}$$

$$G'_{ij} = 0 \quad \text{if} \quad \{n_i, n_j\} = \{0, 1\}, \tag{3.40}$$

$$G'_{ij} = 1 \quad \text{if} \quad n_i = n_j = \frac{1}{2}. \tag{3.41}$$

The first condition leaves the energy unchanged both in the RHF region and at dissociation; the second condition leaves the Fock matrix unchanged in the RHF region, while the last one changes it at dissociation. To satisfy all these constraints,

we assume a minimalist approach of the lowest-order polynomial in $\mathbf{P}$, which leads to a unique 5th degree solution:

$$\mathbf{G}(\mathbf{P}) = 16\mathbf{P}^2(\mathbf{P} - \mathbf{I})^2(\mathbf{P} - \frac{\mathbf{I}}{2}). \tag{3.42}$$

Note, however, that alternative expressions of higher-degree polynomials are possible. The expression for G'_{ij} is presented below. With ϕ added to the CPMFT energy expression, the $n_0 = \frac{1}{2}$ condition in the NO basis now reads:

$$f_{11} + U_{11} - f_{22} - U_{22} = 0. \tag{3.43}$$

We simply need to choose U_{11} and U_{22} to satisfy this equation. There are, however, an infinite number of solutions to this problem. We thus propose the following physically-motivated choice. We first add and subtract the chemical potential μ and write the above equation as $(f_{11}+U_{11}+\mu_A)-(f_{22}+U_{22}+\mu_A) = 0$. By choosing $U_{ii} = -(f_{ii}+\mu_A)$, a clear physical interpretation emerges: U_{11} and U_{22} are chemical potential shifts defining *orbital* chemical potentials ($\mu_i = \mu_A + U_{ii} = -f_{ii}$) such that for the purpose of our model, the non-degenerate orbitals 1 and 2 effectively become degenerate at zero energy. This procedure is described pictorially in Figure 3.7. The $\mathbf{U}$ potential can be easily obtained from a converged CPMFT calculation at dissociation. For the degenerate case, $f_{11} = f_{22}$, so trivially $\mu_1 = \mu_2 = -f_{11} = -f_{22}$.

In summary, we have introduced ϕ, $\mathbf{U}$, and $\mathbf{G}(\mathbf{P})$ in CPMFT such that:

- The energy expression is unchanged at dissociation.

- The energy minimum occurs always at $n = \frac{1}{2}$.

- The model reduces to RHF when $\mathbf{P}$ is idempotent.

In order to distinguish this extended model from CPMFT, we will refer to it as CPMFT+ϕ in the rest of this section. Note that the ϕ condition is non-zero in the
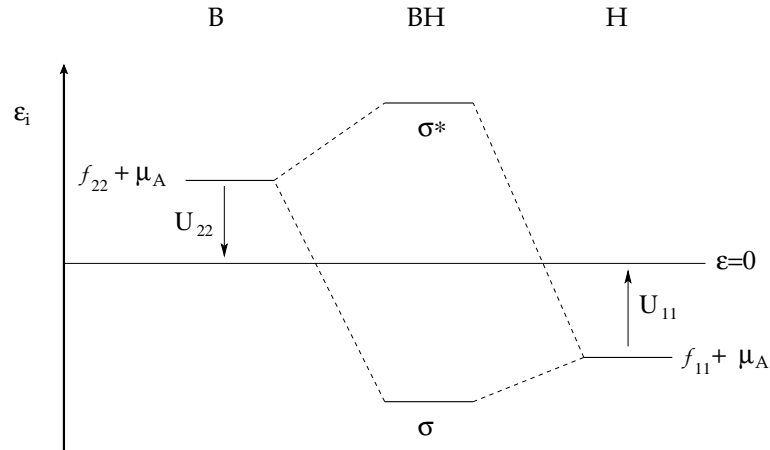
Figure 3.7 : Dissociations to non-degenerate orbitals are characterized by orbitals with different energies localized on different fragments. Introducing chemical potential shifts, degeneracy can be effectively restored and the correct dissociation energy obtained with CPMFT.

region between the RHF solution (idempotent density matrix) and dissociation (fully entangled pair). In between these two regions, ϕ can contribute to the energy, and is assumed and verified in multiple test cases to remain small. $\mathbf{G}(\mathbf{P})$ changes the one-particle Hamiltonian, but not the 2PDM ansatz of Eq.(3.17). In this sense, ϕ is an exact condition (a Lagrange constraint adding to zero in a mathematical sense) only at dissociation, thus the name "asymptotic constraint." Our suggestion is that the static part of the correlation energy (strong correlation) can be obtained from this smooth interpolation between the strongly correlated (dissociation) and uncorrelated (RHF) limits. This is a very reasonable objective for a mean-field model.

3.3.5   Examples of non-degenerate dissociations

To test our CPMFT+ϕ, we first discuss BH with a cc-pVTZ basis set. Energy curves and occupation numbers as the eigenvalue of the charge density $\mathbf{P} = (\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta)/2$ are
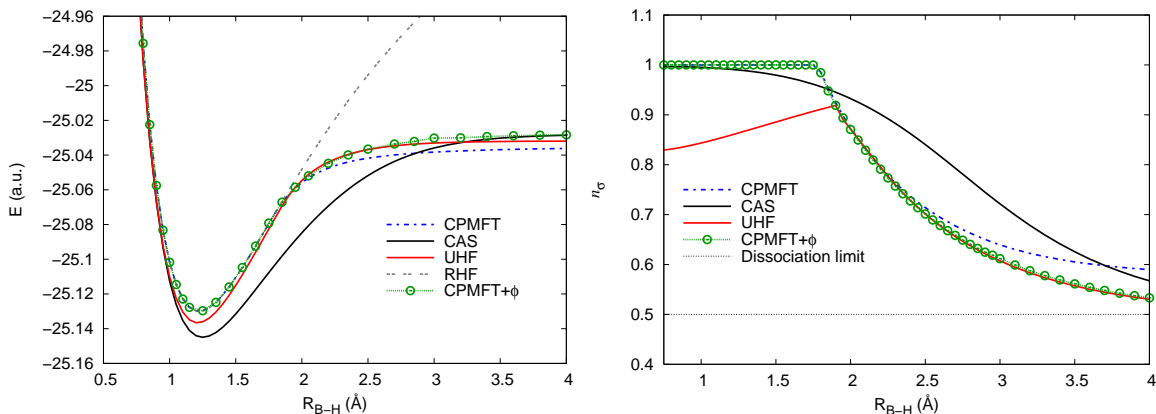
Figure 3.8 : Potential energy curves (Left) and change in occupation number in the σ orbital (Right) for the BH molecule calculated with a cc-pVTZ basis set.

presented in Figure 3.8. CPMFT+ϕ fixes the spilling problem and yields ROHF energies at dissociation. For CPMFT, the occupation number in the bonding σ orbital, n_σ, initially follows a horizontal line ($n = 1$) until pairing occurs. The CASSCF onset of natural occupations before CPMFT is due to weak correlation included in the CASSCF. Note, however, the bizarre behavior of the UHF occupation before joining the CPMFT curve around 2 Å. Also, we note that the CPMFT occupations are not necessarily exact; the model targets the energy, not the occupations. In Figure 3.9 we present results for LiH, HCN, CO, and C$_2$H$_4$. We used a cc-pVTZ basis set except for HCN which was computed with a 6-31G** basis set. Note how in the LiH case the appearance of the CPMFT(2,2) solution is much earlier than the Coulson-Fischer point. Also, CPMFT+ϕ has much less strong correlation than CPMFT near the equilibrium point; the ϕ constraint raises the energy to address spilling and overcorrelation at dissociation. Another example presented here is HCN with all bonds dissociating simultaneously. Note that in this case, the active space is (6,6), so CPMFT+ϕ is by no means limited to (2,2) active spaces. The last two examples involve breaking double bonds in CO and C$_2$H$_4$, which are all correctly dissociated with CPMFT+ϕ.
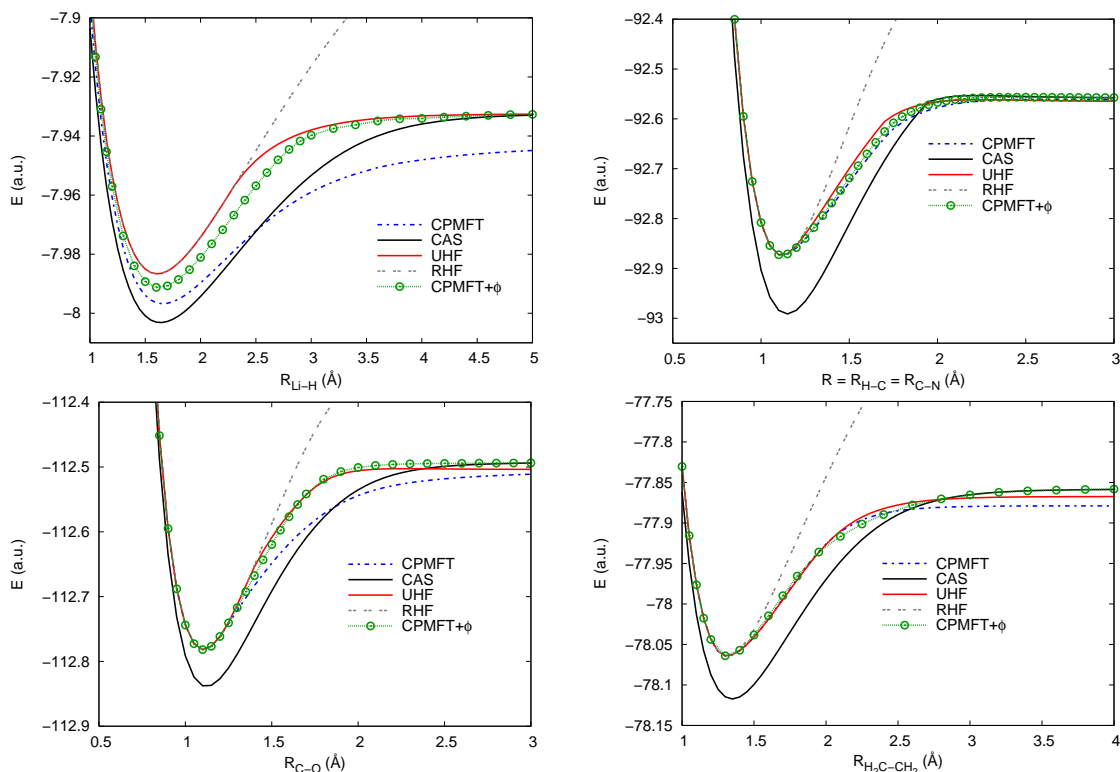
Figure 3.9 : Potential energy curves for the LiH [(2,2)], HCN [(6,6)], CO [(4,4)], and C_2H_4 [(4,4)] molecules. The numbers in brackets are active spaces.

One final important consideration is the observation that the B, C, N, and O atoms are ROHF unstable, *i.e.*, they yield lower atomic energies (and spin contamination) when calculated with UHF. This explains the differences between UHF and CPMFT+ϕ (and CAS) at dissociations involving these atoms.

In summary, we have extended the original CPMFT model, which exactly dissociates electron pairs into degenerate orbitals, to deal with dissociations to nondegenerate orbitals. To achieve this goal, we have introduced the concept of asymptotic constraints. These conditions are satisfied as Lagrange constraints only under special circumstances—molecular dissociation in our case. They modify the energy everywhere else, but are designed to exert no effect in the absence of strong correla-

tion (RHF solution). The extended CPMFT+ϕ model developed in this section can smoothly dissociate *any* polyatomic molecule exactly to restricted open-shell atoms or fragments. The above benchmarks demonstrate how CPMFT accounts for strong correlation in an accurate, yet computationally inexpensive manner. From next section, we will denote CPMFT+ϕ by just CPMFT, unless otherwise stated.

3.4   Inclusion of weak correlations

Until now, CPMFT($+\phi$) includes only strong correlations, while neglecting weak, dynamical correlations. Here we extend the model to account for weak correlations. Natural candidates to do this are exchange-correlation functionals from density functional theory (DFT), since (1) DFT is also computationally very efficient, (2) a DFT correlation functional includes mostly weak correlation, and (3) we can take advantage of the large body of work developed over the past several decades. In an attempt to blend DFT with CPMFT, however, we unfortunately face a fundamental obstacle: in strongly-correlated systems the vast majority of DFT work assumes that strong correlation is captured by spatial and spin-symmetry breaking and weak correlation via a correlation functional which depends on broken symmetry densities (orbitals and densities are spin-polarized in cases where spin polarization should be zero) [29, 30]. CPMFT orbitals and densities have the correct space and spin symmetries, and if fed into standard DFT exchange-correlation subroutines, they would, in general, return poor quality results in such situations. However, the traditional dogma of working with symmetry broken densities in DFT is a choice, not necessarily an imposition of the theory. We believe that the preference of working with unrestricted orbitals stems from the need to describe left-right correlations, which are essentially nonlocal in a symmetry adapted formalism, whereas DFT, in its traditional formulation, prefers to

be local or semilocal (*i.e.*, dependent upon quantities such as the density and orbitals evaluated at a reference point).

To achieve this goal, we here build upon previous work [31–34] demonstrating that the total density and the on-top density are viable alternatives to the standard α and β densities used in DFT. In what follows, we mix CPMFT with HF and DFT in an attempt to take advantage of the best that each model has to offer. We mix DFT exchange with CPMFT in regular hybrid [35] and range-separated schemes [36–44]. We also add weak correlation to CPMFT using DFT functionals via alternative densities. The justification for these approaches is well founded as will be described in detail below, and in Appendix C.

3.4.1 Alternative densities in Kohn-Sham theory

First, we briefly discuss the approach of defining alternative densities based on the on-top pair density $P_2(\mathbf{r})$ and the total density, $\rho(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r})$, as independent variables in Kohn-Sham (KS) theory [45]. $P_2(\mathbf{r})$ is the diagonal part of the pair density, *i.e.*, $P_2(\mathbf{r}) = P_2(\mathbf{r}, \mathbf{r}' = \mathbf{r})$. Using Löwdin's normalization, the pair density is defined as

$$P_2(\mathbf{r}, \mathbf{r}') = \frac{N_e(N_e - 1)}{2} \int \Psi^*(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2, \cdots, \mathbf{r}_N, \sigma_N)$$
$$\times \Psi(\mathbf{r}'_1, \sigma_1, \mathbf{r}'_2, \sigma_2, \cdots, \mathbf{r}_{N_e}, \sigma_{N_e})$$
$$d^3\mathbf{r}_3 \cdots d^3\mathbf{r}_N d\sigma_1 d\sigma_2 \cdots d\sigma_{N_e}\big|_{\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}, \mathbf{r}'_1 = \mathbf{r}'_2 = \mathbf{r}'} \tag{3.44}$$

where N is the number of electrons, Ψ is the wave function of the system under consideration, and $\mathbf{r}$ and σ are spatial and spin coordinates. Eq.(3.44) is equivalent to

$$P_2(\mathbf{r}, \mathbf{r}') = \sum_{pqrs} \Gamma_{pq}^{rs} \, \phi_p(\mathbf{r}) \, \phi_q(\mathbf{r}') \, \phi_r^*(\mathbf{r}) \, \phi_s^*(\mathbf{r}') \tag{3.45}$$

where we recall that $\mathbf{\Gamma}$ is the 2PDM and ϕ_p are orthonormal spin-orbitals. Recall, also, that we use $p, q, \cdots$ indices for denoting spin-orbitals, and $i, j, \cdots$ for spin-integrated spatial-orbitals.

KS calculations with approximations have shown that good results can be obtained only when the energy functional is chosen to depend on α and β electron densities individually, or equivalently, on the total density, $\rho(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r})$, and the spin polarization density $m(\mathbf{r}) = \rho_\alpha(\mathbf{r}) - \rho_\beta(\mathbf{r})$. The latter has proven to be particularly useful for dissociating closed-shell molecules. KS calculations traditionally use symmetry breaking, which in many situations ensures size-consistency. In such cases, however, the significance of a non-zero m is not clear, as it should be zero when the total spin $S = 0$.

A way out of the dilemma of having to choose between the right spin- and spatial-symmetries versus having size-consistency is to imagine that the spin polarization appearing in unrestricted DFT calculations can in fact represent another quantity: an alternative spin polarization density χ_m. In order to take advantage of existing functionals of ρ and m, χ_m can be generated from the density and α and β densities of the KS determinant. In this way, one can convert any functional of ρ and m (or ρ_α and ρ_β) into a functional of ρ and χ_m. Using this prescription, existing functionals can be used as previously, but now m has a new interpretation, *viz.*, generating an alternative, physically relevant, spin polarization.

In the current CPMFT model, only closed-shell situations have been considered, so $m(\mathbf{r}) = 0$. When dissociating closed-shell molecules, fragments show up for which (when treated individually) $m(\mathbf{r}) \neq 0$. For each individual fragment, the CPMFT energy expression becomes that obtained with an ROHF determinant. In this limiting situation, we would like to take the complement of the CPMFT calculations with

correlation density functionals designed for KS calculations. Since $m(\mathbf{r}) = 0$, we can achieve this by using the alternative spin polarization density χ_m.

There are many ways to generate such an alternative spin polarization density. Following the work of Yamaguchi and coworkers [46, 47], one can generate alternative densities from the pair density of spin-unrestricted Slater determinants,

$$P_2(\mathbf{r}, \mathbf{r}') = \frac{1}{2} \left[\rho(\mathbf{r})\rho(\mathbf{r}') - |\gamma^\alpha(\mathbf{r}, \mathbf{r}')|^2 - |\gamma^\beta(\mathbf{r}, \mathbf{r}')|^2 \right] \tag{3.46}$$

where

$$\gamma^\sigma(\mathbf{r}, \mathbf{r}') = \sum_{ij} \gamma_{ij}^\sigma \, \psi_i^\sigma(\mathbf{r}) \, \psi_j^\sigma(\mathbf{r}') \tag{3.47}$$

and $\rho_\sigma(\mathbf{r}) = \gamma^\sigma(\mathbf{r}, \mathbf{r})$. Recall that $\boldsymbol{\gamma}^\sigma$ and $\psi_i^\sigma(\mathbf{r})$ are the 1PDM and spatial orbitals $\psi_i(\mathbf{r})$ of σ spin. Eq.(3.46) yields the so-called on-top pair density when $\mathbf{r}' = \mathbf{r}$,

$$P_2(\mathbf{r}) = \rho_\alpha(\mathbf{r})\rho_\beta(\mathbf{r}). \tag{3.48}$$

In other words, one can replace m in any functional of ρ and m with [31–34]

$$\chi_m(\mathbf{r}) = \sqrt{\rho(\mathbf{r})^2 - 4P_2(\mathbf{r})}. \tag{3.49}$$

3.4.2 Alternative densities in CPMFT

We now elaborate on the extension of this scheme for CPMFT. Recall that the CPMFT 2PDM in the spin-orbital basis is

$$\Gamma_{pq}^{rs} = \frac{1}{2} \left(\gamma_p^r \gamma_q^s - \gamma_p^s \gamma_q^r - \kappa_{pq} \kappa^{rs} \right) \tag{3.50}$$

where we recall $\boldsymbol{\gamma}$ and $\boldsymbol{\kappa}$ are the 1PDM and pairing matrix in the spin-orbital basis, and $-\boldsymbol{\kappa}^2 = \boldsymbol{\gamma} - \boldsymbol{\gamma}^2$ is positive definite. To avoid any confusion, we remind the reader that $\boldsymbol{\gamma}$ is *not* the charge density matrix $\mathbf{P} = \frac{1}{2} \left(\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta \right)$, but is a block-matrix with

respect to spin blocks (Eq.(2.45)). Also remember that one can always diagonalize $\boldsymbol{\gamma}^\sigma$ to obtain the natural spin occupation numbers n_i^σ as its eigenvalues, and the NOs, ψ^{NO}, in which the pairing matrix is also diagonal, $\kappa_i^{\sigma\sigma'} = \sqrt{n_i^\sigma - (n_i^\sigma)^2} = \sqrt{n_i^{\sigma'} - (n_i^{\sigma'})^2}$ ($\sigma \neq \sigma'$) in general. Note that $0 \leq n_i^\sigma \leq 1$. The absence of strong correlation is characterized by $\boldsymbol{\kappa} = \mathbf{0}$, which yields an idempotent 1PDM, *i.e.*, regular RHF.

Substituting Eq.(3.50) into Eq.(3.45) and then setting $\mathbf{r}' = \mathbf{r}$ yields the on-top pair density of CPMFT,

$$P_2(\mathbf{r}) = \rho_\alpha(\mathbf{r})\rho_\beta(\mathbf{r}) + \frac{1}{2}\left[\kappa_{\alpha\beta}(\mathbf{r})^2 + \kappa_{\beta\alpha}(\mathbf{r})^2\right]. \tag{3.51}$$

In Eq.(3.51), $\kappa_{\sigma\sigma'}(\mathbf{r})$ is the pairing matrix in real space. In analogy with Eq.(3.47),

$$\kappa_{\sigma\sigma'}(\mathbf{r}, \mathbf{r}') = \sum_{ij} \kappa_{ij}^{\sigma\sigma'} \psi_i^\sigma(\mathbf{r}) \psi_j^{\sigma'}(\mathbf{r}'), \tag{3.52}$$

and $\kappa_{\sigma\sigma'}(\mathbf{r}) \equiv \kappa_{\sigma\sigma'}(\mathbf{r}, \mathbf{r}' = \mathbf{r})$. Although we restrict ourselves to closed-shell systems, in this chapter, we will retain the spin labels α and β for the sake of clarity. Therefore, $\boldsymbol{\gamma}^\alpha = \boldsymbol{\gamma}^\beta$, $\boldsymbol{\kappa}^{\alpha\beta} = -\boldsymbol{\kappa}^{\beta\alpha}$, $\psi_i(\mathbf{r}) = \psi_i^\alpha(\mathbf{r}) = \psi_i^\beta(\mathbf{r})$, and thus $\rho_\alpha(\mathbf{r}) = \rho_\beta(\mathbf{r})$ and $\kappa_{\alpha\beta}(\mathbf{r}) = -\kappa_{\beta\alpha}(\mathbf{r})$.

Note that Eq.(3.51) is normalized to the correct number of electron pairs as well as Eq.(3.50), which is correctly normalized to $N_e(N_e - 1)/2$. However, the 2PDM of CPMFT is not N-representable as there is no wave function associated with it.

Now for practical purposes, we introduce the alternative densities $\chi_\alpha(\mathbf{r})$ and $\chi_\beta(\mathbf{r})$ satisfying

$$\rho(\mathbf{r}) = \chi_\alpha(\mathbf{r}) + \chi_\beta(\mathbf{r}), \tag{3.53}$$

and

$$\chi_m(\mathbf{r}) = \chi_\alpha(\mathbf{r}) - \chi_\beta(\mathbf{r}). \tag{3.54}$$

We put α and β indices on χ for convenience, but they do not represent actual electron densities unless $\kappa_{\alpha\beta}(\mathbf{r})$ is zero. Our first mapping model for defining χ_α and χ_β is based on Eq.(3.49). Using Eq.(3.49) and Eq.(3.51), one can easily derive

$$\chi_\alpha(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \kappa_{\alpha\beta}(\mathbf{r}) \tag{3.55}$$

$$\chi_\beta(\mathbf{r}) = \rho_\alpha(\mathbf{r}) - \kappa_{\alpha\beta}(\mathbf{r}) \tag{3.56}$$

where we have used $\rho_\alpha(\mathbf{r}) = \rho_\beta(\mathbf{r})$ and $\kappa_{\alpha\beta}(\mathbf{r}) = -\kappa_{\beta\alpha}(\mathbf{r})$. In a practical sense, we replace $\rho_\sigma(\mathbf{r})$ by $\chi_\sigma(\mathbf{r})$ in a DFT exchange-correlation functional E_{xc}. Therefore, for the exchange-correlation energy (but not its derivatives), the implementation of this scheme is simple and straightforward: one substitutes $\chi_\sigma(\mathbf{r})$ into existing DFT subroutines. It should be mentioned that in cases where E_{xc} depends on density gradients and/or kinetic energy densities, one needs to use the corresponding quantities in terms of $\chi_\sigma(\mathbf{r})$. For example, for a generalized gradient approximation (GGA) like Perdew-Burke-Ernzerhof functional (PBE) [48], $E_{xc}^{\mathrm{GGA}}[\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta]$ should be replaced by $E_{xc}^{\mathrm{GGA}}[\chi_\alpha, \chi_\beta, \nabla\chi_\alpha, \nabla\chi_\beta]$.

Note, however, the possibility of $\chi_\beta(\mathbf{r})$ in the above definition Eq.(3.56) becoming negative. The natural occupations, x_i^σ, of $\chi_\sigma(\mathbf{r})$ are

$$x_i^\alpha = n_i + \sqrt{n_i - n_i^2} \tag{3.57}$$

$$x_i^\beta = n_i - \sqrt{n_i - n_i^2}, \tag{3.58}$$

where we have used $n_i = n_i^\alpha = n_i^\beta$. As is illustrated in Figure 3.10, x_i^β is negative when $0 < n_i < 0.5$ and, therefore, so can be $\chi_\beta(\mathbf{r})$. This result is not surprising, as our CPMFT 2PDM ansatz in Eq.(3.50) is not N-representable. While the existence of a transformation to alternative densities $\chi_\sigma(\mathbf{r})$ is well founded, a negative $\chi_\beta(\mathbf{r})$ would be quite problematic for all present exchange-correlation functionals. One could
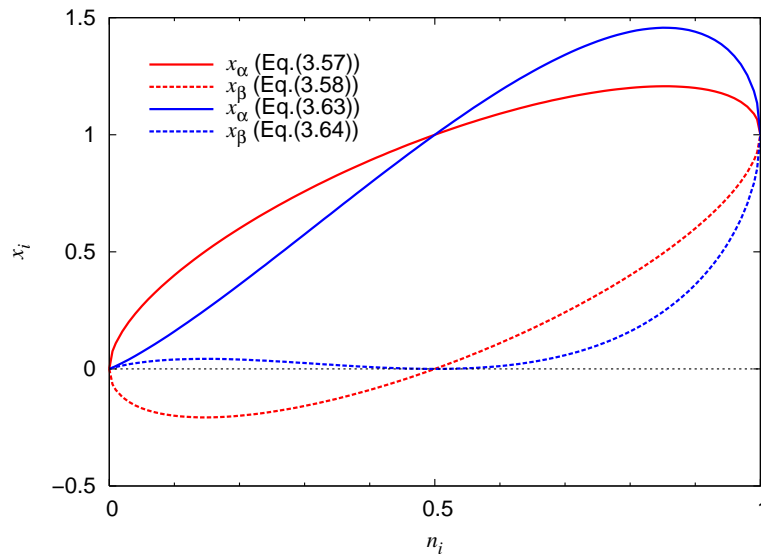
Figure 3.10 : Plots of x_i^σ in Eqs.(3.57, 3.58) (in red) and Eqs.(3.63, 3.64) (Transformation A, in blue) as functions of n_i. Note that x_i^β of Eq.(3.58) is negative when $0 < n_i < 0.5$ while that of Eq.(3.64) is positive everywhere.

extend exchange-correlation functionals to treat negative densities, but the physical meaning of the result would not be clear.

Fortunately, there are many ways to construct $E_{xc}^{\text{on-top}}[\rho(\mathbf{r}), P_2(\mathbf{r})]$, *e.g.*, choosing alternative densities $\chi_\sigma(\mathbf{r})$ for E_{xc}. When fed into exchange-correlation functionals, we want $\chi_\sigma(\mathbf{r})$ to behave as if they were *real* densities. In other words, they have to be positive, continuous, and differentiable. We will also require that

1. $x_i^\alpha = x_i^\beta = 0$ if $n_i = 0$

2. $x_i^\alpha = x_i^\beta = 1$ if $n_i = 1$

3. $x_i^\alpha = 1, x_i^\beta = 0$ if $n_i = 0.5$

4. $x_i^\alpha + x_i^\beta = n_i^\alpha + n_i^\beta = 2n_i$.

The first and second conditions guarantee that $E_{xc}[\chi_\alpha(\mathbf{r}), \chi_\beta(\mathbf{r})]$ produces exactly

$E_{xc}[\rho_\alpha(\mathbf{r}), \rho_\beta(\mathbf{r})]$ when there is no strong correlation (*i.e.*, our model reduces to a standard restricted KS solution due to the idempotency of the 1PDM). They also enforce $\chi_\sigma^{\text{core}}(\mathbf{r}) = \rho_\sigma^{\text{core}}(\mathbf{r})$ for the core orbitals because their occupation numbers are always 1 for both α and β spins. The third condition yields the correct exchange-correlation energy for strongly correlated systems (*e.g.*, dissociation of a molecule). The fourth condition ensures Eq.(3.53). We stress that this only applies to the active orbitals of CPMFT, where n_i can be fractional. All of the above conditions are satisfied in Eqs.(3.55, 3.56), except for the positivity condition for $\chi_\beta(\mathbf{r})$. Based on these requirements, we propose the following alternative densities

$$\chi_\alpha(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \int \left[\rho_\alpha(\mathbf{r}, \mathbf{r}')\kappa_{\alpha\beta}(\mathbf{r}', \mathbf{r}) + \kappa_{\alpha\beta}(\mathbf{r}, \mathbf{r}')\rho_\alpha(\mathbf{r}', \mathbf{r})\right] d\mathbf{r}' \qquad (3.59)$$

$$\chi_\beta(\mathbf{r}) = \rho_\alpha(\mathbf{r}) - \int \left[\rho_\alpha(\mathbf{r}, \mathbf{r}')\kappa_{\alpha\beta}(\mathbf{r}', \mathbf{r}) + \kappa_{\alpha\beta}(\mathbf{r}, \mathbf{r}')\rho_\alpha(\mathbf{r}', \mathbf{r})\right] d\mathbf{r}', \qquad (3.60)$$

which are simple extensions of Eqs.(3.55, 3.56). We will refer to this transformation as *Transformation A*. In matrix form, they are equivalent to

$$\boldsymbol{\chi}^\alpha = \boldsymbol{\gamma}^\alpha + \left(\boldsymbol{\gamma}^\alpha \boldsymbol{\kappa}^{\alpha\beta} + \boldsymbol{\kappa}^{\alpha\beta} \boldsymbol{\gamma}^\alpha\right) = \mathbf{P} + (\mathbf{PK} + \mathbf{KP}) \qquad (3.61)$$

$$\boldsymbol{\chi}^\beta = \boldsymbol{\gamma}^\alpha - \left(\boldsymbol{\gamma}^\alpha \boldsymbol{\kappa}^{\alpha\beta} + \boldsymbol{\kappa}^{\alpha\beta} \boldsymbol{\gamma}^\alpha\right) = \mathbf{P} - (\mathbf{PK} + \mathbf{KP}). \qquad (3.62)$$

The eigenvalues of the above $\boldsymbol{\chi}^\sigma$ are

$$x_i^\alpha = n_i + 2n_i\sqrt{n_i - n_i^2} \qquad (3.63)$$

$$x_i^\beta = n_i - 2n_i\sqrt{n_i - n_i^2}. \qquad (3.64)$$

Eqs.(3.63, 3.64) are also shown in Figure 3.10. With this transformation, it is guaranteed that $\chi_\beta(\mathbf{r})$ is always positive.

Given the requirements above, a family of plausible candidates for $\boldsymbol{\chi}_\sigma$ in matrix
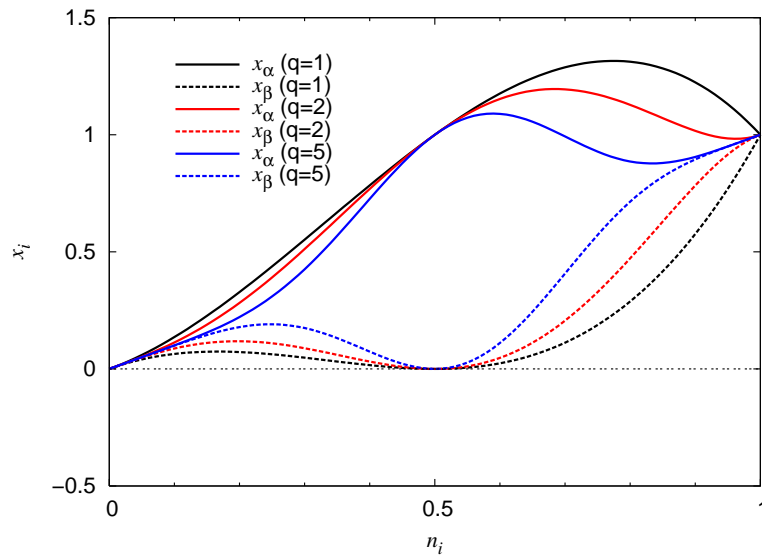
Figure 3.11 : Plots of x_i^σ with Transformation B for several q as a function of n_i.

form are

$$\boldsymbol{\chi}^\alpha = \frac{1}{2}\left(\boldsymbol{\gamma}^\alpha + 2^{2q}(\boldsymbol{\kappa}^{\alpha\beta})^q\boldsymbol{\gamma}^\alpha(\boldsymbol{\kappa}^{\alpha\beta})^q\right) \tag{3.65}$$

$$\boldsymbol{\chi}^\beta = \frac{1}{2}\left(\boldsymbol{\gamma}^\alpha - 2^{2q}(\boldsymbol{\kappa}^{\alpha\beta})^q\boldsymbol{\gamma}^\alpha(\boldsymbol{\kappa}^{\alpha\beta})^q\right) \tag{3.66}$$

where q is a positive number, and we call them Transformation B. Although these $\boldsymbol{\chi}^\sigma$ do not have explicit analytical expressions in $\rho(\mathbf{r})$ and $P_2(\mathbf{r})$, one could expand them in terms of these two variables in real space. Figure 3.11 shows the behavior of natural occupations of this family for different q. Note that these alternative densities also satisfy Conditions 1-4. Compared to Transformation A, they substantially differ in the way they approach the restricted KS densities when occupations are close to 1 or 0, especially when q is large. Also, when $q = \frac{1}{2}$, Transformation B is equivalent to Transformation A, because $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\kappa}^{\alpha\beta}$ commute.

3.4.3 Constrained-pairing generalized Kohn-Sham

As discussed previously, the 2PDM of CPMFT is postulated to include only strong correlation. Thus, the addition of standard DFT correlation functionals to CPMFT seems like a viable way of adding weak correlation effects to the model. In other words, we assume that the pairing energy in CPMFT and regular DFT correlation functionals handle strong and weak correlation, respectively. However, it turns out that a combination of 100% HF-exchange E_x^{HF} plus 100% pairing E_p (defined below) plus 100% DFT correlation E_c^{DFT} would not be fully satisfactory since in many molecules at equilibrium (where the 1PDM is idempotent) this combination reduces to RHF $+ E_c^{\mathrm{DFT}}$, which is well known to be a less-than-ideal approximation in terms of accuracy. Therefore, inclusion of a portion of DFT exchange (E_x^{DFT}) along with E_c^{DFT} seems to be an interesting alternative worth exploring.

It is normally assumed that, as opposed to E_c^{DFT}, E_x^{DFT} introduces a portion of strong correlation via its localization (and approximation) of the exchange hole. In doing so, self-interaction error is also introduced. In order to avoid double-counting of strong correlation from the CPMFT pairing energy E_p and E_x^{DFT}, we resort to adiabatic connection arguments for mixing them.

Here we propose a novel scheme to blend CPMFT with regular DFT. Global hybrid functionals [35] contain a portion of nonlocal HF-type exchange potential,

$$E_{xc}^{\mathrm{hyb}} = aE_x^{\mathrm{HF}} + (1-a)E_x^{\mathrm{DFT}} + E_c^{\mathrm{DFT}}, \tag{3.67}$$

where

$$E_x^{\mathrm{HF}} = -\frac{1}{2}\sum_{ijkl}\langle ik|lj\rangle(\gamma_{ij}^{\alpha}\gamma_{kl}^{\alpha} + \gamma_{ij}^{\beta}\gamma_{kl}^{\beta}) = -\sum_{ijkl}\langle ik|lj\rangle P_{ij}P_{kl}, \tag{3.68}$$

and a is a parameter which can be fit to the experimental values of chemical properties such as heats of formation, or can be determined theoretically without experimental

fits. We note that E_x^{HF} is not an explicit functional of electron density but a functional of KS orbitals, and therefore in most cases hybrid functionals are treated as the generalized KS (GKS) scheme, in which the derivative with respect to orbitals is used as the KS potential. We add to Eq.(3.67) the pairing energy E_p defined as

$$E_p = -\sum_{ijkl}\langle ij|kl\rangle\kappa_{ij}^{\alpha\beta}\kappa_{kl}^{\alpha\beta} = -\sum_{ijkl}\langle ij|kl\rangle K_{ij}K_{kl}. \qquad (3.69)$$

so that our exchange-correlation-pairing (xcp) energy is given by

$$\begin{aligned} E_{xcp}^{\text{CPGKS}} &= a(E_x^{\text{HF}}[\mathbf{P}] + E_p[\mathbf{K}]) \\ &\quad + (1-a)E_x^{\text{DFT}}[\chi_\alpha, \chi_\beta] + E_c^{\text{DFT}}[\chi_\alpha, \chi_\beta]. \qquad (3.70) \end{aligned}$$

We call this scheme Constrained-Pairing GKS (CPGKS). Note that $E_p \leq 0$. The coefficient of E_p must be the same as that of E_x^{HF} to obtain correct energies at dissociation. Otherwise, strong correlation may be double-counted. A simple rationalization of the CPGKS energy expression arises from considering a mixture of the HF, HFB, and DFT energy expressions with electron-electron interactions given by $\frac{2a}{r_{12}}$, $-\frac{a}{r_{12}}$, and $\frac{1-a}{r_{12}}$, respectively. If all three are added, the CPGKS energy expression is obtained. Note that HF with $\frac{2a}{r_{12}}$ has excess exchange energy, but it is exactly compensated with the exchange energy from HFB with $-\frac{a}{r_{12}}$ (recall that HFB itself has an HF-type exchange term). Most importantly, note that E_x^{DFT} and E_c^{DFT} in Eq.(3.70) are now explicit functions of $\chi_\alpha(\mathbf{r})$ and $\chi_\beta(\mathbf{r})$ introduced in the last section, whereas E_x^{HF} and E_p are dependent on the 1PDM and pairing matrix, respectively. In other words, CPGKS can be considered a *hybrid* scheme where HF, HFB, and semilocal DFT are blended. We emphasize that for this mixture to yield meaningful results, the alternative densities need to be chosen as described earlier and the fractions of E_p and E_x^{DFT} have to add up to one.

The above argument can be extended to a range-separated hybrid scheme [36–44], *e.g.*, with the electron-electron interaction split into

$$\frac{1}{r_{12}} = \frac{2\text{erf}(\omega r_{12})}{r_{12}} - \frac{\text{erf}(\omega r_{12})}{r_{12}} + \frac{\text{erfc}(\omega r_{12})}{r_{12}}, \tag{3.71}$$

where erf and erfc are the error function and complementary error function, and ω is a parameter determining the range separation. The first term is used for HF, the second for HFB (both exchange and pairing), and the third for DFT-exchange. Again, E_x^{HF} with an interaction of $2\text{erf}(\omega r_{12})/r_{12}$ from HF part partially cancels out with the $-\text{erf}(\omega r_{12})/r_{12}$ interaction of HFB, yielding $\text{erf}(\omega r_{12})/r_{12}$ as the total HF-type exchange interaction. The resulting energy expression is

$$E_{xcp} = E_x^{lr-\text{HF}} + E_x^{sr-\text{DFT}} + E_c^{\text{DFT}} + E_p^{lr}, \tag{3.72}$$

where superscripts lr and sr stand for the long-range and short-range interactions, respectively. The first term uses a long-range interaction $\text{erf}(\omega r_{12})/r_{12}$, the second a short-range $\text{erfc}(\omega r_{12})/r_{12}$, the third full $1/r_{12}$, and the last a long-range *attractive* $-\text{erf}(\omega r_{12})/r_{12}$ potential. If one chooses $\omega = 0.4$ bohr^{-1} and PBE as the DFT functional, then Eq.(3.72) becomes LC-ωPBE [42] when the last term is zero. In other words, in the absence of strong correlation, Eq.(3.72) yields the LC-ωPBE result. Hereafter, we refer to this functional as Constrained-Pairing LC-ωPBE (CPLC-ωPBE).

3.4.4 Exchange-correlation potentials in CPGKS

In order to achieve self-consistency in a CPGKS calculation, we need the derivatives of the exchange-correlation energy with respect to physical densities $\rho_\sigma(\mathbf{r})$ and

anomalous density $\kappa_{\alpha\beta}(\mathbf{r})$. Using the chain rule, the potentials are

$$v_\sigma^\rho(\mathbf{r}) = \sum_{\sigma'} \int \frac{\partial E_{xc}}{\partial \chi_{\sigma'}(\mathbf{r}')} \frac{\partial \chi_{\sigma'}(\mathbf{r}')}{\partial \rho_\sigma(\mathbf{r})} d\mathbf{r}' \tag{3.73}$$

$$v_{\alpha\beta}^\kappa(\mathbf{r}) = \sum_{\sigma'} \int \frac{\partial E_{xc}}{\partial \chi_{\sigma'}(\mathbf{r}')} \frac{\partial \chi_{\sigma'}(\mathbf{r}')}{\partial \kappa_{\alpha\beta}(\mathbf{r})} d\mathbf{r}'. \tag{3.74}$$

Note that v^ρ and v^κ yield contributions to the Fock Hamiltonian ($\mathbf{F}^{\mathrm{cs}}$) and pairing field ($\boldsymbol{\Delta}$), respectively. In this section, we derive the explicit forms of these contributions for Transformation A. After some simple algebra, we find that the exchange-correlation potential matrices are

$$\mathbf{V}_\alpha^\rho = \frac{1}{2}\frac{\partial E_{xc}}{\partial \boldsymbol{\gamma}^\alpha} = \frac{1}{2}\left(\mathbf{W} + \tilde{\mathbf{W}}\boldsymbol{\kappa}^{\alpha\beta} + \boldsymbol{\kappa}^{\alpha\beta}\tilde{\mathbf{W}}\right) = \mathbf{V}_\beta^\rho \tag{3.75}$$

$$\mathbf{V}_{\alpha\beta}^\kappa = \frac{1}{2}\frac{\partial E_{xc}}{\partial \boldsymbol{\kappa}^{\alpha\beta}} = \frac{1}{2}\left(\tilde{\mathbf{W}}\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\alpha\tilde{\mathbf{W}}\right) \tag{3.76}$$

where

$$\mathbf{W} = \left(\frac{\partial}{\partial \boldsymbol{\chi}_\alpha} + \frac{\partial}{\partial \boldsymbol{\chi}_\beta}\right) E_{xc} \tag{3.77}$$

$$\tilde{\mathbf{W}} = \left(\frac{\partial}{\partial \boldsymbol{\chi}_\alpha} - \frac{\partial}{\partial \boldsymbol{\chi}_\beta}\right) E_{xc}. \tag{3.78}$$

Note that $\mathbf{V}_{\alpha\beta}^\kappa = -(\mathbf{V}_{\beta\alpha}^\kappa)^{\mathrm{T}}$ in the same way that $\boldsymbol{\Delta}_{\alpha\beta} = -(\boldsymbol{\Delta}_{\beta\alpha})^{\mathrm{T}}$, *i.e.*, they are anti-symmetric.

Similarly, it is straightforward to obtain the corresponding potentials for Transformation B. As explained above, we separate the NOs between core, active, and virtual spaces introducing the chemical potential for the active space and the Lagrange multiplier matrix $\boldsymbol{\lambda}$ for the inactive space. Therefore, once the desired $\boldsymbol{\chi}_\sigma$ are obtained, Eqs.(3.75, 3.76) are not restricted to any particular orbital space: they are used for constructing the entire Fock Hamiltonian and pairing field.
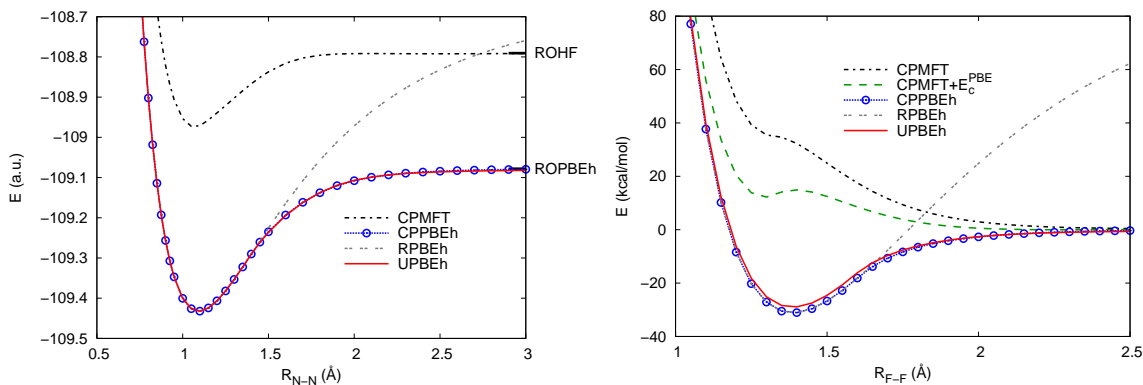
Figure 3.12 : Potential energy curves of the N_2 and F_2 molecules with a 6-311++G** basis set.

3.4.5 Results

Diatomic molecules

We begin by presenting benchmark calculations for the dissociation curves of N_2 and F_2 with the 6-311++G** basis set, using the constrained-pairing scheme for the PBEh functional [49–52] (CPPBEh), which uses PBE exchange and correlation functionals [48] with the global mixing constant $a = 0.25$ in Eq.(3.70). In these calculations, we have used Transformation A for $\chi_\sigma(\mathbf{r})$. Our calculations are carried out in D_{2h} symmetry, but in the correct $D_{\infty h}$ symmetry, the sum of occupation numbers in σ_g-σ_u, $\pi_{x,u}$-$\pi_{x,g}$, and $\pi_{y,u}$-$\pi_{y,g}$ active orbital pairs are all fixed to two each (one for each spin). This is referred to as the "corresponding pair" constraint and will be discussed later. This restriction is easily applied by using different chemical potentials (as a Lagrange multiplier) that control the occupation numbers of each pair. In Figure 3.12, we compare the CPPBEh dissociation curve with that of unrestricted PBEh (UPBEh) for the N_2 molecule. Interestingly, we find little difference in energies: both energies are almost identical one another. In Table 3.3, we list the total energies for the N_2 molecule at several R_{N-N}. The good agreement between CPPBEh and UPBEh

indicates that our formulae for alternative densities (Transformation A) are quite reasonable.

Table 3.3 : Total energy of the N_2 molecules calculated with a 6-311++G** basis set (in Hartree).

R_{N-N} (Å)	RPBEh	UPBEh	CPPBEh[a]
0.8	-108.902 12	-108.902 12	-108.902 12
1.2	-109.406 38	-109.406 38	-109.406 38
1.6	-109.164 22	-109.194 61	-109.193 06
2.0	-108.971 29	-109.106 95	-109.107 90
2.4	-108.853 10	-109.087 58	-109.086 26
2.8	-108.782 66	-109.083 10	-109.080 64
3.2	-108.740 75	-109.081 70	-109.078 85
3.6	-108.715 51	-109.081 07	-109.078 24
4.0	-108.699 89	-109.080 86	-109.078 01

[a] Transformation A is used. The sum of the restricted open-shell PBE energies of two isolated N atoms is -109.077 81 Hartree.

Transformation B presented in Eqs.(3.65, 3.66) behaves differently. For $q \geq 1$, the critical bond length (R_c) where the CPPBEh solution appears is longer compared to UPBEh and CPPBEh with Transformation A. For example, in Figure 3.12, while R_c for the latter two methods are both 1.4 Å, its value for Transformation B with $q = 1$ is around 1.6 Å (not shown in Figure 3.12). This happens because Transformation B approaches the restricted PBE solution much faster when the occupation numbers of the active orbitals are close to zero or one. For the remainder of this section, we report CPGKS results only with Transformation A.

Also, we emphasize that in these examples, while UPBEh breaks the spin- and spatial-symmetries as it approaches dissociation, CPPBEh correctly preserves both. It also shows significant improvement over CPMFT and RPBEh because it has both

weak and strong correlation, and yields a much more accurate dissociation energy D_0, 218.9 kcal/mol, than does CPMFT, which gives 111.2 kcal/mol; the experimental value is 225.1 kcal/mol [53].

F_2 is a dramatic example where a balance between E_x^{DFT} and E_c^{DFT} is needed in order for the molecule to be bound (Figure 3.12). Pure CPMFT, which includes no weak correlation, predicts a repulsive F_2 potential energy curve, implying that substantial weak correlation is missing near equilibrium. Merely adding E_c^{PBE} does not fix this problem as can be seen in the CPMFT+E_c^{PBE} curve: it produces a (local) minimum around 1.297 Å, which is too short compared to the experimental bond length of $R_e = 1.412$ Å [54]. On the other hand, CPPBEh gives a bound curve with $R_e = 1.387$ Å, similar to the good performance of PBEh for this system. Again, UPBEh gives a reasonable curve but does so by breaking spin- and spatial-symmetries.
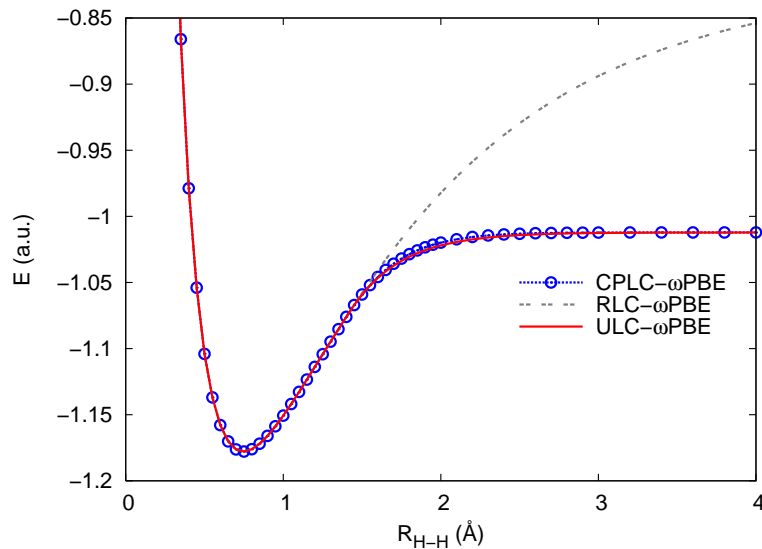
Figure 3.13 : Potential energy curves of the H_2 molecule with range-separated hybrid functionals.

Figure 3.13 shows the H_2 dissociation curve of CPLC-ωPBE along with ULC-
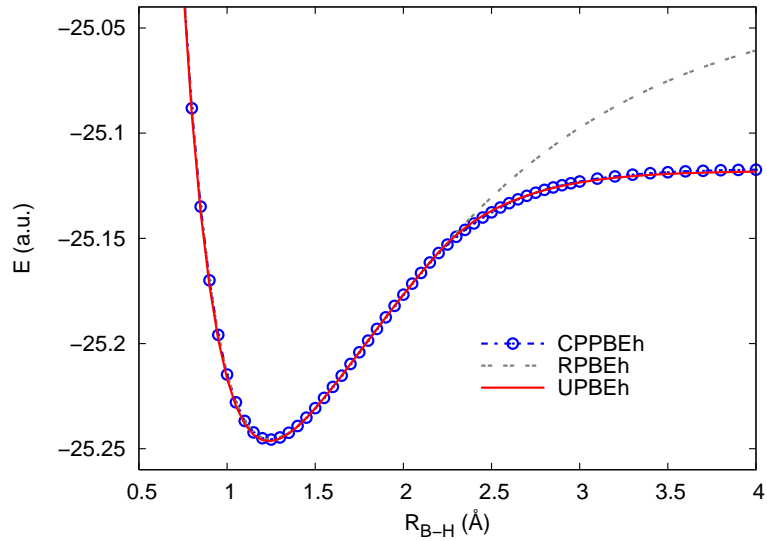
Figure 3.14 : Potential energy curves of the BH molecule.

ωPBE. Considering that ULC-ωPBE is very accurate for this system, the mixture of CPMFT and DFT with a range-separated interaction proves also successful.

Our approach is by no means limited to homo-nuclear systems, but can be used for hetero-nuclear systems using the previously discussed CPMFT+ϕ scheme. We recall that, in CPMFT+ϕ, we equilibrate the energies of the non-degenerate (active) orbitals of the dissociated fragments using a Lagrangian matrix $\mathbf{U}$ which plays the role of an orbital chemical potential shift and fixes the occupation numbers of dissociated orbitals to the correct $\frac{1}{2}$ value. Then, along a dissociation curve, one uses $\mathbf{U}$ in combination with a polynomial of the 1PDM that imposes constraints on both the dissociation and weak (RHF) correlation limits. It is remarkable that for the BH molecule using CPMFT+ϕ and alternative densities (see Figure 3.14), CPPBEh gives very accurate dissociation curves compared to UPBEh.

Although all the hybrid examples presented thus far show significant improvement

over both CPMFT and restricted GKS, our CPGKS yields negligible differences in dissociation curves compared to unrestricted GKS (UGKS). This is to be expected, because UPBEh and ULC-ωPBE give very good results for these systems. Therefore, the fact that CPGKS produces dissociation curves of the same quality as UGKS indicates that our transformation Eqs.(3.59, 3.60) are a reasonable starting point for further development of functionals better suited to CPGKS.

In order to test whether CPGKS yields significant differences compared to UGKS, we consider a more challenging example for hybrid UGKS, one where strong correlation is ubiquitous: the chromium dimer, Cr_2. The ground state of Cr_2 is $^1\Sigma_g^+$, and dissociates to two septet Cr atoms. The twelve valence electrons have a strong multiconfigurational electronic structure at equilibrium, which makes calculations extremely difficult. It has been reported that unrestricted hybrid functionals predict the bond length to be too long, and the dissociation energy D_0 to be too small [55]. Edgecombe and Becke showed that a hybrid functional was able to describe both bond length and D_0 accurately after a simple spin-projection to $S = 0$ [56].

We performed CPPBEh(12,12) and CPLC-ωPBE(12,12) calculations using the CPMFT+ϕ scheme. The Cr $4s$ and $3d$ orbitals are not degenerate at dissociation; thus, a small chemical potential is needed to equilibrate them. In these calculations, we also constrain the number of electrons in each irreducible representation of the point group symmetry, as we did in the N_2 calculations: the sum of occupation numbers in σ_g-σ_u, π_g-π_u, and δ_g-δ_u active orbitals are all fixed to four each. The basis set used in the Cr_2 calculations is the contracted 14s11p6d2f/10s8p3d2f Gaussian set of Scuseria and Schaefer [57].

In Figure 3.15, we plot the potential energy curves of the Cr_2 molecule obtained with different methods. None of these approaches predict the shoulder at long R_e ob-
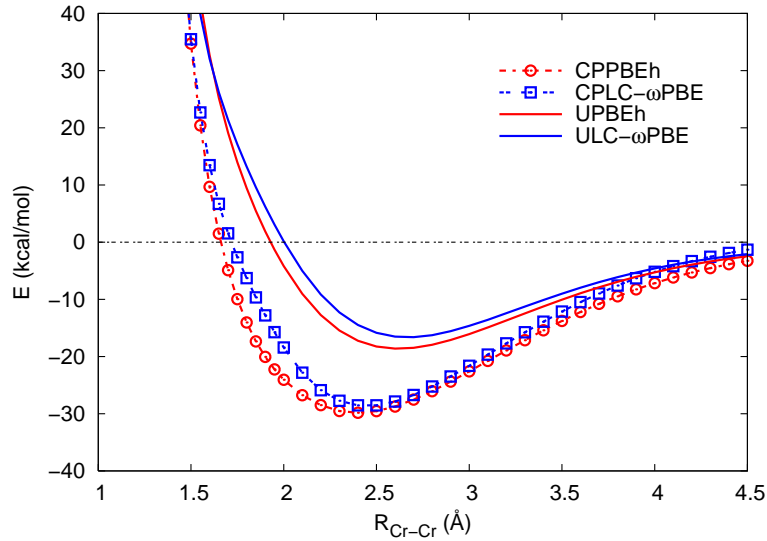
Figure 3.15 : Potential energy curves of the Cr_2 molecule.

tained with other multireference wave functions methods [58]. CPPBEh and CPLC-ωPBE predict an extended R_e of 2.400 Å and 2.448 Å, respectively, relative to the experimental value of 1.679 Å [59]. However, they both yield shorter R_e compared to their parent functional (Table 3.4). The dissociation energy of Cr_2 (D_e) obtained with CPPBEh and CPLC-ωPBE (Table 3.4) is much more reasonable than those of hybrid UGKS.

Table 3.4 : Bond length (Å) and dissociation energy (kcal/mol) for Cr_2.

	CPPBEh	CPLC-ωPBE	UPBEh	ULC-ωPBE	Exptl.
R_e	2.400	2.448	2.617	2.670	1.679[a]
D_e	29.8	28.6	18.6	16.6	34.0[b]

[a] Ref. [59].
[b] Estimated from the experimental vibrational frequency of 480.6 cm^{-1} (Ref. [60]) and D_0 of 33.3 kcal/mol (Ref. [61]).
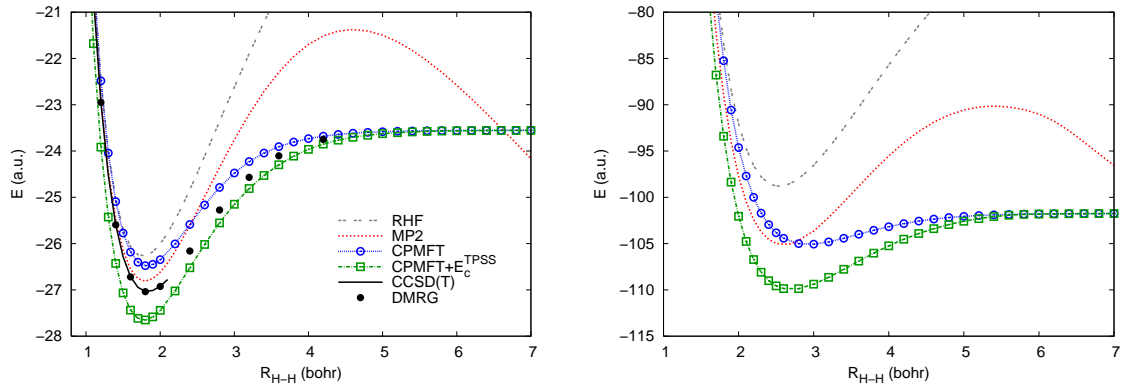
Figure 3.16 : Potential energy curves for the symmetric dissociations of a H_{50} chain (Left) and a $6 \times 6 \times 6$ hydrogen cube (Right).

Hydrogen network

As a last example, we discuss our results on hydrogen networks, namely, a one dimensional H_{50} chain and a three dimensional (3D) $6 \times 6 \times 6$ hydrogen cube, with a uniform internuclear distance R_{H-H} between adjacent atoms. In the limit of $R_{H-H} \to \infty$, these model systems dissociate to 50 and 216 isolated hydrogen atoms, respectively. As R_{H-H} increases, they display a metal-insulator transition. Both systems are paradigmatic models for strongly-correlated Mott insulators [62] and cannot be addressed by CASSCF due to the staggering number of active configurations, 10^{27} and 10^{123} for the 1D and 3D cases, respectively. In the linear case, the DMRG approach [63], a very accurate multi-reference wave function method, is available [62]. However, for the 3D case, DMRG is not (yet) applicable. As shown below, accurate CPMFT treatment of these model systems only requires diagonalizing Hamiltonians of moderate dimensions: 100×100 and 432×432, respectively.

Figure 3.16 presents potential energy curves for each system computed with several methods using an STO-6G basis. This choice of minimal basis is made to allow comparison with results in the literature [62]. All electrons and orbitals are explicitly

treated in our calculations. The active spaces of CPMFT are (50,50) and (216,216) for the 1-D and 3-D systems, respectively. In both cases, RHF and second-order Møller-Plesset perturbation theory (MP2) yield unreliable curves, while coupled cluster singles, doubles, and perturbative triples (CCSD(T)) has convergence difficulties at long internuclear distances. The RHF method misses a considerable amount of strong correlation, as we have already discussed. MP2 may both miss or exaggerate strong correlation, as seen from Figure 3.16. CCSD(T), despite its single reference character, is quite good at covering correlation near equilibrium of the 1D system, giving quantitatively similar results to the reference DMRG. This is probably because CCSD(T) includes infinite-order excitation terms in a balanced way, and most correlation in the vicinity of equilibrium happens to be rather weak correlation.

For our CPGKS cases of these systems, we used CPMFT with the TPSS correlation energy [64]. Although at first glance CPMFT+E_c^{TPSS} appears to overcorrelate the 1D chain near equilibrium, this is partially a basis set effect: CCSD(T) and DMRG are essentially underestimating correlation effects due to the small basis used in the calculation (a necessity to make the CCSD(T) and DMRG calculations affordable).

Table 3.5 : Correlation energy (in Hartree) of a H_{50} chain at R_e.

Basis set	MP2	CCSD(T)	CPMFT	CPMFT+E_c^{TPSS}
STO-6G[a]	-0.545 55	-0.765 47	-0.211 12	-1.405 62
cc-pVDZ	-0.976 45	-1.213 41	-0.203 42	-1.443 22
cc-pVTZ	-1.123 26	-1.347 9[b]	-0.205 46	-1.430 83
cc-pVQZ	-1.162 74	-1.374[b]	-0.205 81	-1.430 66

[a] The DMRG correlation energy with the STO-6G basis is -0.772 67 (Ref. [62]).
[b] Estimated value by extrapolation of a H_n chain to $n = 50$.

To further investigate the accuracy of CPMFT, we performed calculations of the

1-D hydrogen chain near equilibrium (R_e = 1.8 bohr) with very large cc-pVnZ basis sets, where n =D, T, Q. We compare our correlation energies with CCSD(T), which yields results close to DMRG (Figure 3.16) with the STO-6G basis, and is thus considered very accurate for the 1D chain near R_e. For the cc-pVTZ and cc-pVQZ bases, we extrapolated the CCSD(T) correlation energy of H_{50} from shorter chains, as it was not feasible to compute CCSD(T) for H_{50} within a reasonable time-frame. Results are presented in Table 3.5. While convergence of the CCSD(T) correlation energy with basis set size is slow, the correlation energies of both CPMFT and CPMFT+E_c^{TPSS} converge reasonably well. Strong correlation is expected to be fairly basis set independent, a property reproduced by CPMFT.
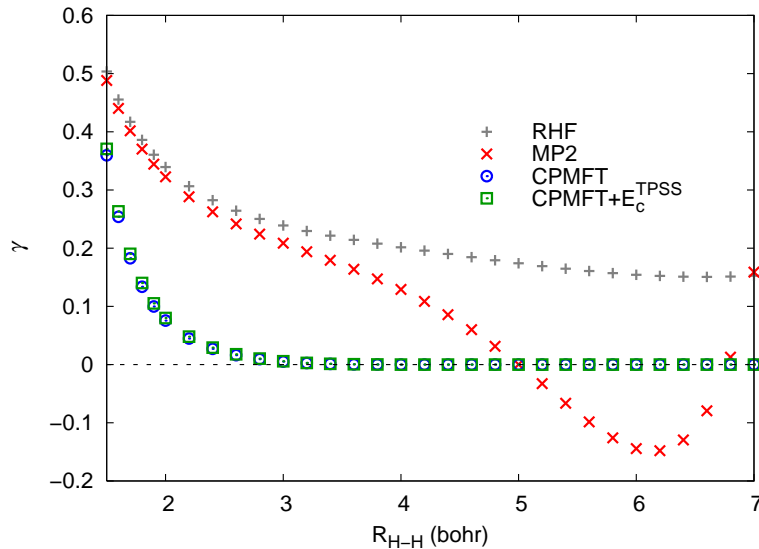
Figure 3.17 : Decay of off-diagonal density matrix term (γ_{12}) for two hydrogen atoms at diagonal vertices in a $4 \times 4 \times 4$ hydrogen cube.

Finally, in order to visualize the metal-insulator transition explicitly, we have plotted the off-diagonal density matrix term γ_{12} between two hydrogen atoms at diagonal vertices in a $4 \times 4 \times 4$ cube (Figure 3.17). Both RHF and MP2 off-diagonal

terms remain substantially above zero as R_{H-H} increases, implying that the electrons are delocalized. The CPMFT and CPMFT+E_c^{TPSS} off-diagonal γ_{12}, on the other hand, rapidly decay toward zero. Evidently, only CPMFT reveals the gradual metal-insulator transition in this hydrogen cube.

3.5 Corresponding pair constraints and connection to UHF

3.5.1 CPMFT and UHF

To this point, we have demonstrated the success of CPMFT in describing strong left-right correlations. Here, we investigate the close relation between CPMFT and UHF, to provide greater insight into how to improve CPMFT—or even UHF.

Consider the UHF treatment of a system where the number of spin-up and spin-down electrons is the same. The spin-up and spin-down density matrices $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$ are both idempotent:

$$(\boldsymbol{\gamma}^\alpha)^2 - \boldsymbol{\gamma}^\alpha = (\boldsymbol{\gamma}^\beta)^2 - \boldsymbol{\gamma}^\beta = \mathbf{0}. \tag{3.79}$$

The charge density and spin magnetization (or polarization) density matrices are

$$\mathbf{P} = \frac{1}{2}\big(\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta\big), \tag{3.80a}$$

$$\mathbf{M} = \frac{1}{2}\big(\boldsymbol{\gamma}^\alpha - \boldsymbol{\gamma}^\beta\big). \tag{3.80b}$$

Traditionally, the UHF energy [65] is expressed in terms of the $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$ density matrices:

$$E_{\text{UHF}} = h_{ij}(\gamma_{ij}^\alpha + \gamma_{ij}^\beta) + \frac{1}{2}\langle ij|kl\rangle(\gamma_{ik}^\alpha + \gamma_{ik}^\beta)(\gamma_{jl}^\alpha + \gamma_{jl}^\beta)$$
$$- \frac{1}{2}\langle ij|kl\rangle(\gamma_{il}^\alpha\gamma_{jk}^\alpha + \gamma_{il}^\beta\gamma_{jk}^\beta), \tag{3.81}$$

where we have put $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$ in the same basis (say, the atomic orbital basis). Although it is almost never presented in this way, we can also write the UHF energy

as a functional of $\mathbf{P}$ and $\mathbf{M}$, which yields

$$E_{\mathrm{UHF}}[\mathbf{P}, \mathbf{M}] = E_{\mathrm{cs}}[\mathbf{P}] + E_c[\mathbf{M}], \tag{3.82a}$$

$$E_{\mathrm{cs}}[\mathbf{P}] = 2h_{ij}P_{ij} + (2\langle ij|kl\rangle - \langle ij|lk\rangle)P_{ik}P_{jl} \tag{3.82b}$$

$$E_c[\mathbf{M}] = -\langle ij|kl\rangle M_{il}M_{jk}. \tag{3.82c}$$

Here, E_{cs} indicates the usual RHF energy expression given in terms of the charge density matrix $\mathbf{P}$, while E_c carries the "correlation" energy in terms of the spin magnetization density matrix $\mathbf{M}$. An utterly unexpected result is that the closed-shell CPMFT energy expression

$$E^{\mathrm{CPMFT}} = 2\sum_{ij} h_{ij}P_{ij} + \sum_{ijkl}(2\langle ik|jl\rangle - \langle ik|lj\rangle) - \sum_{ijkl}\langle ij|kl\rangle K_{ij}K_{kl}, \tag{3.83}$$

is identical to the UHF energy expression of Eq.(3.82), except that the spin density matrix $\mathbf{M}$ is replaced by the anomalous density matrix $\mathbf{K}$.[†] In cases in which UHF predicts strong correlation by breaking symmetry (*i.e.*, non-zero spin contamination) [66], $\mathbf{P}$ is not idempotent. Instead, it satisfies

$$\mathbf{P} - \mathbf{P}^2 = \frac{1}{2}(\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta) - \frac{1}{4}(\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta)^2 \tag{3.84a}$$

$$= \frac{1}{4}(\boldsymbol{\gamma}^\alpha - \boldsymbol{\gamma}^\beta)^2 = \mathbf{M}^2. \tag{3.84b}$$

This is one consequence of the idempotency of $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$. The second is

$$\mathbf{P}\mathbf{M} + \mathbf{M}\mathbf{P} = \mathbf{M}. \tag{3.85}$$

Note that the condition of Eq.(3.84) is the same as the CPMFT condition of Eq.(3.11b), again with $\mathbf{M}$ taking the role of $\mathbf{K}$. Both the magnetization density matrix $\mathbf{M}$ and the anomalous density matrix $\mathbf{K}$ are Hermitian.

[†]Note that UHF subtracts $\langle ij|kl\rangle M_{il}M_{jk}$ from the closed-shell energy, while CPMFT subtracts $\langle ij|kl\rangle K_{ij}K_{kl}$. However, these are the same if the basis functions and the anomalous density matrix are real, as is generally the case.

While CPMFT and UHF use the same energy expression (one with $\mathbf{K}$ and the other with $\mathbf{M}$), $\mathbf{K}$ and $\mathbf{M}$ are not identical even though with the same density matrix $\mathbf{P}$, we have $\mathbf{K}^2 = \mathbf{M}^2$. There are other important differences: both UHF and CPMFT impose an additional condition on these two matrices, which in UHF is given in Eq.(3.85), while in CPMFT is instead given in Eq.(3.11a). Additionally, $\mathbf{K}$ is positive semi-definite, while $\mathbf{M}$ is traceless (and thus has both positive and negative eigenvalues). Finally, because in UHF we write $\mathbf{P}$ as the half-sum of two idempotent matrices, its eigenvalues occur in what is known as "corresponding pairs" n_i and $1-n_i$ [67, 68].

That UHF has the corresponding pairs property has little to do with UHF *per se*. It originates simply from the fact [22] that the eigenvalues of a matrix that is the half-sum of two idempotent matrices are 0, 1, $\frac{1}{2}$, or a corresponding pair $(n, 1-n)$,[†] which can be easily shown as in Appendix D. Similarly, the eigenvalues of a matrix written as the half-difference of two idempotent matrices are 0, $\pm\frac{1}{2}$, or a corresponding pair $(-n, n)$. For example, $\mathbf{M}$ has eigenvalues adding to 0 in pairs, while $\mathbf{P}$ has eigenvalues adding to 1 in pairs. Quite generally, any non-integer eigenvalues of the charge density matrix from a single determinant method will be either $\frac{1}{2}$ or occur in a corresponding pair. Eigenvalues of $\frac{1}{2}$ could be part of a corresponding pair (for entangled electrons) or may occur singly for open shells. We should be clear that while matrices written as the sum of two idempotent matrices exhibit the corresponding pairs property, the converse is not necessarily true; a matrix whose eigenvalues come in corresponding pairs may or may not be the sum of two idempotents.

Unlike UHF, the eigenvalues of $\mathbf{P}$ in double Hamiltonian (DH) CPMFT, Eq.(3.9),

---

[†]More precisely, the eigenvalues of $\alpha\mathbf{A}+\beta\mathbf{B}$ for idempotent $\mathbf{A}$ and $\mathbf{B}$ are $0, \alpha, \beta$, or a corresponding pair, $(n, \alpha + \beta - n)$.

do not occur in corresponding pairs (except when the active space consists of two spatial orbitals). That said, the corresponding pairs property has some attractive features for CPMFT. Most importantly, it eliminates overcorrelation between orbital pairs in different symmetries. This is, ubiquitous, for example, in N_2 where the variational principle drives occupancy into orbitals at low energies and one must introduce multiple chemical potentials to retain the correct total number of σ and π electrons. A corresponding pair constraint controls this unphysical "spilling" and has the inherent attractive feature of limiting strong correlations to be an affair between orbital pairs.

Previously, we introduced the corresponding pairs feature within the DH-CPMFT framework using different chemical potentials (Lagrange multipliers) for different irreducible representations of the system (Section 3.4.5). However, in the general case where no spatial symmetry is present, imposition of this constraint leads to one Lagrange multiplier per orbital pair and a rather complicated, nonlinear optimization problem. A more satisfactory and much simpler approach is to write the CPMFT density matrix as

$$\mathbf{P} = \frac{1}{2}\left(\mathbf{A} + \mathbf{B}\right) \tag{3.86}$$

where $\mathbf{A}$ and $\mathbf{B}$ are auxiliary density matrices, individually idempotent and Hermitian ($\mathbf{A}^2 = \mathbf{A} = \mathbf{A}^\dagger$ and similarly for $\mathbf{B}$). As with UHF, this decomposition enforces the corresponding pairs condition automatically, obviating the need to enforce this condition via Lagrange multipliers. Eigenvalues of 0 or 1 in $\mathbf{P}$ correspond to virtual or core orbitals, respectively, while paired eigenvalues correspond to active orbitals. Further, by choosing $\mathbf{A}$ and $\mathbf{B}$ to trace to half the number of electrons, we guarantee that $\mathbf{P}$ does likewise, and we thus have no need of any chemical potential. By making this decomposition, in other words, we can avoid the Lagrange multipliers μ of the

double-Hamiltonian approach entirely, and thus simplify the computation. Note that once we have converged solutions for $\mathbf{A}$ and $\mathbf{B}$ (and thus $\mathbf{P}$ and $\mathbf{K}$), we could, if desired, extract the Lagrange multipliers of the DH-CPMFT approach.

The critical mathematical difference between UHF and CPMFT as formulated in this manner is that in UHF, we obtain $\mathbf{M}$ from the spin-up and spin-down density matrices, while in CPMFT, we acquire $\mathbf{K}$ from the total density matrix alone (since $\mathbf{K}$ satisfies the condition of Eq.(3.11b), commutes with $\mathbf{P}$, and is positive semi-definite). To rephrase, CPMFT with corresponding pairs defines $\mathbf{P}$ from $\mathbf{A}$ and $\mathbf{B}$ as in Eq.(3.86), but differs from UHF in constructing

$$\mathbf{K} = \sqrt{\mathbf{P} - \mathbf{P^2}} = \frac{1}{2}\sqrt{(\mathbf{A} - \mathbf{B})^2} = \frac{1}{2}|\mathbf{A} - \mathbf{B}|. \tag{3.87}$$

from auxiliary density matrices $\mathbf{A}$ and $\mathbf{B}$ while UHF builds $\mathbf{P}$ and $\mathbf{M}$ from $\boldsymbol{\gamma}^{\alpha}$ and $\boldsymbol{\gamma}^{\beta}$, as shown in Eq.(3.80). Note in Eq.(3.87), our definition of the absolute value of a matrix from the square root of the square. In practice, to calculate the absolute value of a matrix one needs to diagonalize it, flip the sign of the negative eigenvalues and transform back to the original basis. Both the square root and absolute value of a matrix are positive definite matrices and both have a convergent polynomial series expansion if the matrix is positive definite with eigenvalues between 0 and 1, as is the case here.

To make the comparison between CPMFT and UHF more concrete, consider the case where $\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices and let $\mathbf{M} = \frac{1}{2}(\mathbf{A} - \mathbf{B})$. Idempotency of $\mathbf{A}$

and $\mathbf{B}$ requires that in the natural orbital basis we have

$$\mathbf{A} = \begin{pmatrix} n & k \\ k & 1-n \end{pmatrix} \tag{3.88a}$$

$$\mathbf{B} = \begin{pmatrix} n & -k \\ -k & 1-n \end{pmatrix} \tag{3.88b}$$

$$\mathbf{P} = \begin{pmatrix} n & 0 \\ 0 & 1-n \end{pmatrix} \tag{3.88c}$$

$$\mathbf{M} = \begin{pmatrix} 0 & k \\ k & 0 \end{pmatrix} \tag{3.88d}$$

$$\mathbf{K} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \tag{3.88e}$$

$$k = \sqrt{n(1-n)}. \tag{3.88f}$$

When $\mathbf{A}$ and $\mathbf{B}$ are of larger dimension, then in the natural orbital basis they are block diagonal with 2×2 blocks of the form given above. This is essentially a consequence of Eq.(3.85), which in the natural basis becomes

$$(n_i + n_j)M_{ij} = M_{ij} = \frac{1}{2}(A_{ij} - B_{ij}), \tag{3.89}$$

the solutions to which are $M_{ij} = 0$ and $n_i + n_j = 1$. Because we also have $A_{ij} + B_{ij} = 2\, n_i\, \delta_{ij}$ from $\mathbf{A} + \mathbf{B} = \mathbf{P}$, which is diagonal, we conclude that for $i \neq j$, we must either have $A_{ij} = B_{ij} = 0$ or $n_i + n_j = 1$ (in other words, the two eigenvalues form a corresponding pair). When the occupation numbers are degenerate, the natural orbitals are not uniquely defined and we can thus choose them such that $\mathbf{A}$ and $\mathbf{B}$ still have this structure. In the core (virtual) space, $\mathbf{A} = \mathbf{B} = \mathbf{1}$ ($\mathbf{A} = \mathbf{B} = \mathbf{0}$).

Before we continue to the working equations for CPMFT in this UHF-like framework, let us pause to clarify that CPMFT and UHF are different methods. While we have expressed the UHF energy as a density matrix functional, we could also write it

as an expectation value

$$E_{\text{UHF}} = \langle \Phi_{\text{UHF}} | H | \Phi_{\text{UHF}} \rangle, \tag{3.90}$$

with $|\Phi_{\text{UHF}}\rangle$ constrained to be a single determinant. This is *not* true of the CPMFT energy expression, and in fact there appears to be no wave function associated with CPMFT. This may seem somewhat surprising, in light of the intimate connection between CPMFT and HFB theory, in which there certainly *is* a wave function, albeit one which violates particle number conservation. As we have said, we lose the HFB wave function because we have by fiat changed the sign of the pairing energy. Additionally, unlike UHF, the spin-density is zero everywhere for closed shells, even in the presence of strong correlation.

Is CPMFT is equivalent to projected UHF (PUHF) [69]? No. Projecting the UHF determinant onto a spin eigenfunction, one finds that the charge density matrix of the UHF determinant and the spin-projected state have the same eigenfunctions [68]. Spin projection, in other words, changes only the occupation numbers of the charge density matrix, but not the natural orbitals. The fact that the UHF and CPMFT natural orbitals are different should lay to rest any concerns that CPMFT is merely a projected UHF.

Another fundamental difference between CPMFT and UHF is the onset of the appearance of the solution with an energy lower than RHF. As shown in Section 3.3.1, the CPMFT solution for a two-level model system appears when the RHF orbital energy gap reduces to

$$\varepsilon_2 - \varepsilon_1 < \frac{J_{11} + J_{22}}{2} + K_{12}, \tag{3.91}$$

whereas the UHF Coulson-Fischer instability point is determined by

$$\varepsilon_2 - \varepsilon_1 < J_{12} + K_{12}. \tag{3.92}$$

Because all two-electron integrals in the equations above are positive, the CPMFT solution appears inevitably when the orbital gap closes, and strong correlation becomes manifest, such as along a dissociation curve.

3.5.2 Working equations

Let us now return to the solution of the CPMFT equations in this UHF-like framework. For convenience, we repeat the energy expression here:

$$E_{\mathrm{CPMFT}} = E_{\mathrm{cs}} - \langle ij|kl\rangle K_{ij}K_{kl}. \tag{3.93}$$

We then minimize the energy with respect to (idempotent) $\mathbf{A}$ and $\mathbf{B}$ matrices. The derivatives of E_{cs} in Eq.(3.93) with respect to $\mathbf{A}$ and $\mathbf{B}$ give the usual closed-shell Fock matrix obtained from $\mathbf{P}$. That is

$$\frac{\partial E_{\mathrm{cs}}}{\partial A_{ij}} = \frac{\partial E_{\mathrm{cs}}}{\partial B_{ij}} = \frac{1}{2}\frac{\partial E_{\mathrm{cs}}}{\partial P_{ij}} = F_{ij}^{\mathrm{cs}}. \tag{3.94}$$

The differences with UHF arise from differentiating the last term of the CPMFT energy, E_p. Taking derivatives with respect to $\mathbf{A}$ leads to an effective potential $\tilde{\mathbf{\Delta}}$, given by

$$\tilde{\Delta}_{ij} = \frac{\partial E_p}{\partial A_{ij}} = \frac{\partial E_p}{\partial K_{kl}}\frac{\partial K_{kl}}{\partial A_{ij}} = -2\Delta_{kl}\frac{\partial K_{kl}}{\partial A_{ij}}. \tag{3.95}$$

This is essentially the same result that we get from differentiating E_c^{UHF} of Eq.(3.82):

$$\frac{\partial E_c^{\mathrm{UHF}}}{\partial \gamma_{ij}^{\alpha}} = \frac{\partial E_c^{\mathrm{UHF}}}{\partial M_{kl}}\frac{\partial M_{kl}}{\partial \gamma_{ij}^{\alpha}} = -2\Delta_{kl}^{\mathrm{UHF}}\frac{\partial M_{kl}}{\partial \gamma_{ij}^{\alpha}}, \tag{3.96}$$

where

$$\Delta_{kl}^{\mathrm{UHF}} = \langle km|nl\rangle M_{mn} = \langle kl|mn\rangle M_{mn} \tag{3.97}$$

resembles $\mathbf{\Delta}$ except that we replace $\mathbf{K}$ with $\mathbf{M}$. In UHF, however, we simply have

$$\frac{\partial M_{kl}}{\partial \gamma_{ij}^{\alpha}} = \frac{1}{2}\delta_{ik}\delta_{jl}, \tag{3.98}$$

while in CPMFT, the derivative of $\mathbf{K}$ with respect to $\mathbf{A}$ is obtained by differentiating both sides of $\mathbf{K}^2 = \frac{1}{4}(\mathbf{A} - \mathbf{B})^2$. This gives

$$\frac{\partial K_{km}}{\partial A_{ij}} K_{ml} + K_{km} \frac{\partial K_{ml}}{\partial A_{ij}} = \frac{1}{2} \left(M_{jl}\delta_{ki} + M_{ki}\delta_{jl} \right). \tag{3.99}$$

In the natural orbital basis, where $\mathbf{K}$ is diagonal with eigenvalues κ_i, we have

$$\frac{\partial K_{kl}}{\partial A_{ij}} = \frac{1}{2} \frac{M_{jl}\delta_{ki} + M_{ki}\delta_{jl}}{\kappa_k + \kappa_l}. \tag{3.100}$$

Thus, in the natural orbital basis the effective potential $\tilde{\boldsymbol{\Delta}}$ is

$$\tilde{\Delta}_{ij} = -\frac{\Delta_{il}M_{jl}}{\kappa_i + \kappa_l} - \frac{\Delta_{kj}M_{ki}}{\kappa_k + \kappa_j}. \tag{3.101}$$

Since

$$\frac{\partial \mathbf{K}}{\partial \mathbf{A}} = -\frac{\partial \mathbf{K}}{\partial \mathbf{B}}, \tag{3.102}$$

the equations we ultimately solve are $[\mathbf{F}^{\mathrm{A}}, \mathbf{A}] = \mathbf{0}$ and $[\mathbf{F}^{\mathrm{B}}, \mathbf{B}] = \mathbf{0}$, where $\mathbf{F}^{\mathrm{A}}$ and $\mathbf{F}^{\mathrm{B}}$ are effective Fock matrices given by

$$\mathbf{F}^{\mathrm{A}} = \mathbf{F}^{\mathrm{cs}} + \tilde{\boldsymbol{\Delta}}, \tag{3.103a}$$

$$\mathbf{F}^{\mathrm{B}} = \mathbf{F}^{\mathrm{cs}} - \tilde{\boldsymbol{\Delta}}. \tag{3.103b}$$

At first glance, the right-hand-side of Eq.(3.101) might appear to be divergent unless all κ_i are non-zero. However, since forcing $\Delta_{ij} = 0$ actually gives the condition $K_{ij} = 0$, we simply set $\Delta_{ij} = 0$ for the inactive-inactive (core and virtual) block where $\mathbf{K}$ must be zero (see Section 3.2.3). Therefore, in Eq.(3.101), such divergent terms due to inactive orbitals are simply removed from the sum.

3.5.3 Results

We have implemented this version of CPMFT in the `Gaussian` suite of programs [70]. Each calculation requires the specification of the number N_a of active natural

orbitals. Due to the corresponding pairs constraint, the number of active electrons is always equal to N_a – in other words, we always work at half-filling. In order to obtain an appropriate initial guess for $\mathbf{A}$ and $\mathbf{B}$, we mix the coefficients of the N_a orbitals closest to the Fermi level, just as one would do to break spatial symmetries in UHF. The natural orbital pairs closest to the Fermi energy correspond to those whose occupations are closest to half and half.

In single bond systems where we normally choose the active space to be two electrons in two orbitals, the corresponding pair constraint is automatically satisfied, and no difference is observed between the results using this approach and those using our previous double-Hamiltonian approach (that is, diagonalization of the double-Hamiltonian constructed from $\mathbf{F}$ and $\boldsymbol{\Delta}$). However, in DH-CPMFT, one must adjust the chemical potential μ at every iteration of the SCF procedure to control the number of electrons in the active space. Because we must adjust the chemical potential, we must diagonalize the double Hamiltonian of Eq.(3.9) several times in each SCF cycle, until the resulting density matrix has the proper trace. In contrast, the current approach requires no chemical potential, since we have $\mathrm{Tr}(\mathbf{P}) = 1/2\,\mathrm{Tr}(\mathbf{A} + \mathbf{B})$. Because both $\mathbf{A}$ and $\mathbf{B}$ trace to the correct number of electrons, so too does $\mathbf{P}$. This is a significant operational advantage of the present implementation.

For systems with larger active spaces, the present approach differs from DH-CPMFT, although as mentioned above, we can impose the corresponding pairs constraint in DH-CPMFT in some special cases by including different chemical potentials for different irreducible representations. We illustrate this with the case of N_2. Table 3.6 shows the total energy of N_2 at 2.0 Å. We use the cc-pVTZ basis set and choose six active orbitals and six active electrons. The current scheme yields a slightly higher energy than does DH-CPMFT with only one chemical potential, as one would expect

Table 3.6 : CPMFT energies of N_2 at R = 2.0 Å. Also included are the number of diagonalization steps required, N_{diag}, and the number of SCF cycles required for convergence.

Scheme	Energy (Hartree)	N_{diag}	SCF cycles
DH-CPMFT(6,6)[a]	-108.79901762	118	32
DH-CPMFT(6,6)[b]	-108.79715442	121	34
CPMFT(6)	-108.79715442	12	12

[a]Single chemical potential.
[b]Corresponding pairs enforced by multiple chemical potentials.

since we have imposed an additional constraint on the system. Also, as one would expect, it produces the same results as does DH-CPMFT with the corresponding pairs constraint enforced by additional Lagrange multipliers. However, removing the chemical potentials results in considerable computational savings. In Figure 3.18, we show the N_2 dissociation curves from CPMFT in the double-Hamiltonian approach and in the corresponding pairs framework. In this case, the corresponding pairs constraint has only a minor effect on the energy.

We have also performed a CPMFT calculation of the C_2 molecule with the 6-31G basis set. Near equilibrium, C_2 has significant strong correlation due to near-degeneracy between the RHF occupied σ_{2s}^* and unoccupied σ_{2p_z} orbitals. As the molecule is stretched, however, the π_x, π_y, π_x^*, and π_y^* orbitals become degenerate, while the σ_{2s}^*–σ_{2p_z} interaction becomes weak. We have, therefore, chosen our active space to be six electrons in six orbitals for this system. In Figure 3.19, we show the total energy of C_2 as a function of bond length. The CASSCF energy includes all strong correlation that results from these orbital interactions (plus some weak correlation). Without the corresponding pairs constraint, DH-CPMFT strongly over-correlates nearly everywhere. Adding the corresponding pairs constraint significantly
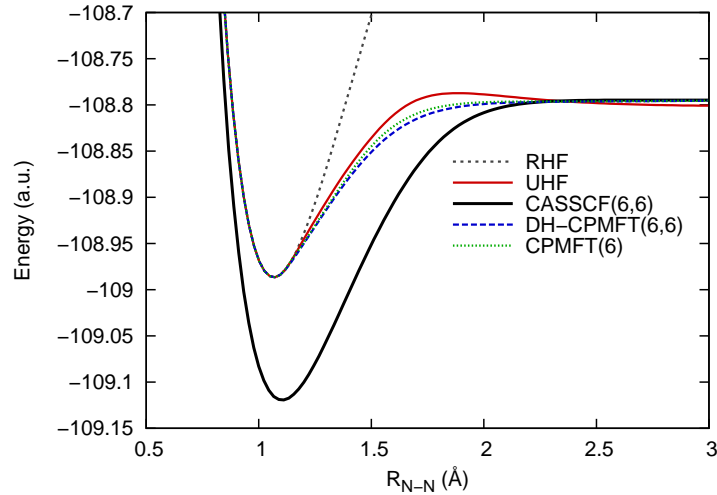
Figure 3.18 : Potential energy curves of N_2 calculated with the cc-pVTZ basis set.

reduces this overcorrelation. Near equilibrium, it gives results between UHF and CASSCF. Unfortunately, it still overcorrelates as the molecule dissociates. This is due to electron "spilling" between $\sigma_{2s}^{\star}$ and σ_{2p_z} orbitals. As R$\rightarrow \infty$, only the π orbitals should be strongly correlated; including these σ orbitals in the active space at large internuclear separation allows them to correlate and lower the energy unphysically. If we remove two orbitals from the active space, we produce the curve marked CPMFT(4). This approaches the correct dissociation limit, but undercorrelates at equilibrium where the active space should be larger. The correct solution for this molecule involves introducing renormalized one-body potentials in CPMFT(6) that eliminate the spilling at dissociation. While going to the right dissociation limit is important, it is perhaps less critical than getting the correct behavior near equilibrium. Note that CPMFT(4) dissociates correctly to two ROHF carbon atoms, while UHF instead dissociates to two spin-contaminated UHF carbon atoms and CASSCF(6,6)
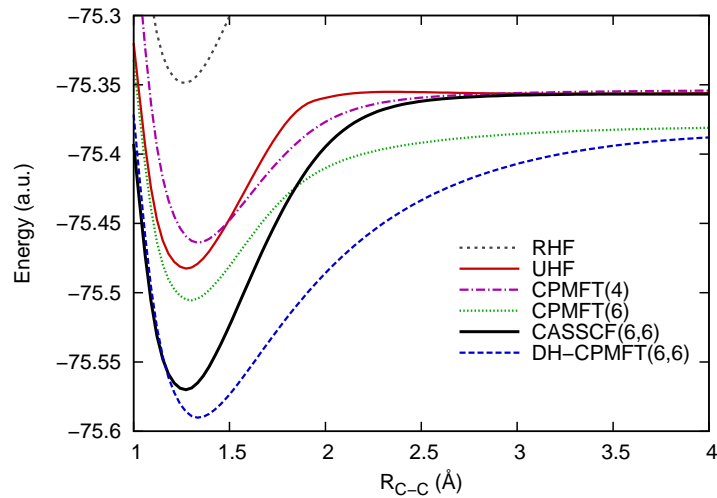
Figure 3.19 : Potential energy curves of C_2 calculated with the 6-31G basis set.

has some weak correlation at dissociation.

Finally, we stress the differences between UHF and CPMFT by analyzing the dissociation of the CO_2 molecule. The ground state of CO_2 near equilibrium is a closed-shell singlet with no expected strong correlation. Indeed both UHF and CPMFT reduce to the RHF solution near R_e. However, when the molecule is symmetrically stretched and the two oxygen atoms are simultaneously separated from the carbon atom, the correct dissociation limit corresponds to all three atoms in their triplet ground state. This situation cannot be handled by UHF. In CO_2 near R_e, there are six electrons associated with bond formation, three with spin-up and three with spin-down. At dissociation, UHF might assign two spin-up electrons to one oxygen atom and two spin-down electrons on the other, which puts both oxygen atoms in their triplet ground state. However, with only one electron of each spin remaining, the best UHF can do is to assign a singlet state to the carbon atom, which is clearly
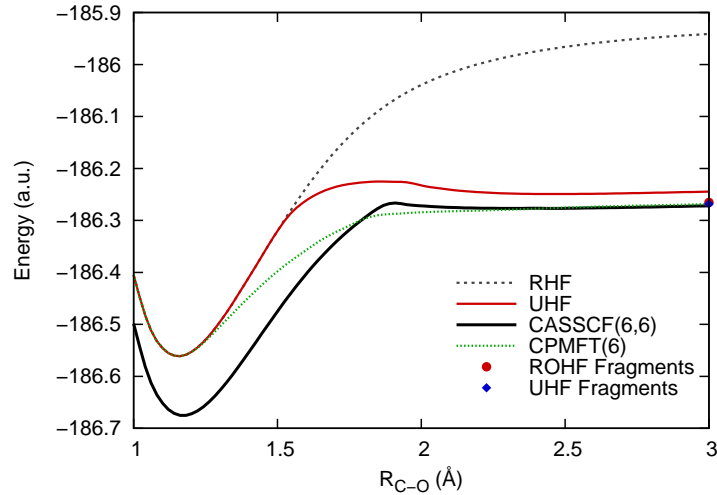
Figure 3.20 : Potential energy curves for the double dissociation of CO_2 calculated with the 3-21G basis set.

incorrect and not the lowest energy state. In simple words, UHF runs out of broken symmetry degrees of freedom (it has only two) to model the dissociation of CO_2 (Figure 3.20) and misses the correct dissociation limit by ~ 20 mHartree. Resolving this problem requires us to break further a symmetry, $\hat{S}_z$, which is done in generalized HF [8]. The bumps in the dissociation curves correspond to crossings of different solutions to the respective SCF equations and we have plotted the lowest energy state at each R. Because spin states are treated in CPMFT through an "ensemble" representation, one that yields zero spin magnetization density everywhere, the CPMFT solution for this dissociation has two half spins up and two half spins down on each of the three atoms, leading to the correct energy corresponding to the sum of ROHF atomic energies. Note that CPMFT(6) in Figure 3.20 contains a one-body potential arising from an asymptotic constraint as explained in Section 3.3, $i.e.$, CPMFT+ϕ.

3.6 Discussion

In this chapter, we have presented our novel mean-field method to describe strong correlations which are ubiquitous in (near-)degenerate systems. With an attractive pairing potential of HFB constrained to an active space, CPMFT successfully reduces the overcorrelation effects occurring in 1HFB. CPMFT core and virtual orbitals have the characteristics of symmetry preserved RHF orbitals. CPMFT has been shown to cleanly separate strong and weak correlations, and accurately dissociates *any* molecules to the ROHF fragments by introducing asymptotic constraints as a one-body Hamiltonian. To achieve chemical accuracy in our method, weak correlations were incorporated into CPMFT via DFT exchange-correlation functionals, by transforming densities ρ to alternative densities χ. Finally, we have investigated the close relation between CPMFT and UHF, and the corresponding pair constraints were considered in CPMFT to further avoid overcorrelation that might occur in some situations such as C_2. The connection between CPMFT and UHF will turn out to be useful and can be exploited to control spin-contamination in UHF. We will discuss this scheme in the following chapters.

Chapter 4

Constrained Unrestricted Hartree-Fock: ROHF theory made simple

4.1   Introduction

In the last chapter, we have developed a novel theory called CPMFT for treating strong correlations within an independent quasiparticle picture. We have mentioned that the UHF energy can be written as a functional of the charge density matrix $\mathbf{P} = (\boldsymbol{\gamma}^{\alpha} + \boldsymbol{\gamma}^{\beta})/2$ and the spin density matrix $\mathbf{M} = (\boldsymbol{\gamma}^{\alpha} - \boldsymbol{\gamma}^{\beta})/2$, in the same way as in CPMFT but with $\mathbf{M}$ instead of $\mathbf{K} = (\mathbf{P} - \mathbf{P}^2)^{1/2}$. This connection between CPMFT and UHF turns out to be enlightening for reformulating ROHF as a constrained UHF (CUHF) theory.

ROHF theory was formulated by Roothaan some 50 years ago [71]. A major drawback of this model is the lack of a unique effective Fock operator [72]. Even though the ROHF wave function and total energy obtained from different coupling schemes are the same, the resulting orbitals and orbital energies are different and lead to post-ROHF results that generally depend on them. The interpretation and physical picture emerging from Roothaan's open-shell theory have always been somewhat blurry. Attempts to resolve these ambiguities, as well as many paradoxes resulting from them, are well documented in the literature [73–76].

On the other hand, the physical picture of unrestricted HF (UHF) is clear [65]. It is a single-determinant wave function with well-defined α and β orbital energies

obeying Koopmans' theorem [1]. It is straightforward to use it in post-UHF calculations by simply treating the α and β orbitals explicitly and separately. The notorious problem in UHF, however, is spin-contamination: the wave function is not an eigenfunction of $\hat{S}^2$. This weakness is ubiquitous and a serious detriment when bonds are stretched. If the UHF wave function suffers from severe spin-contamination, as is the case when strong static correlation is present, then UHF is no longer a good starting reference point for post-UHF treatments of correlation or excited states. Once lost, good quantum numbers are difficult to recover, so when possible, it is preferable to use ROHF as a starting point, despite the ambiguities regarding its associated Fock operator.

Here, we introduce our CUHF scheme which leads to well-defined α and β Fock operators with straightforward interpretation and no spin-contamination. The ROHF wave function, energy, and charge and spin densities remain the same; only the ROHF Fock operator is replaced by two UHF-like counterparts. As shown in benchmarks below, the meaning of the resulting orbitals and orbital energies is much more physical than in Roothaan's approach and provide a base for treatments of electron correlation and excited states.

The familiar energy expression in ROHF is

$$E_{\text{ROHF}} = 2 \sum_i f_i h_{ii} + \sum_{ij} f_i f_j (2a_i^j \langle ij|ij \rangle - b_i^j \langle ij|ji \rangle), \tag{4.1}$$

where h_{ij} are one-electron integrals, $\langle ij|kl \rangle$ are two-electron integrals in Dirac's notation, a_i^j and b_i^j are the coupling coefficients, and f_i are the orbital occupations: 1 for core (doubly-occupied, c) and 0 for virtual (unoccupied, v) orbitals. In the case of high-spin open-shell systems under consideration, $a = 1$, $b = 2$, and $f = 1/2$ for open-shells orbitals (singly-occupied, o). Roothaan's effective Fock operator is defined

as

$$\mathbf{F}_{\mathrm{ROHF}} = \begin{pmatrix} \mathbf{R}_{cc} & \mathbf{F}_{co}^{\beta} & \mathbf{F}_{cv}^{\mathrm{cs}} \\ \mathbf{F}_{oc}^{\beta} & \mathbf{R}_{oo} & \mathbf{F}_{ov}^{\alpha} \\ \mathbf{F}_{vc}^{\mathrm{cs}} & \mathbf{F}_{vo}^{\alpha} & \mathbf{R}_{vv} \end{pmatrix} \begin{array}{l} \text{core } (c) \\ \text{open } (o) \\ \text{virtual } (v) \end{array} \qquad (4.2)$$

where $\mathbf{F}^{\alpha}$ and $\mathbf{F}^{\beta}$ are UHF α and β Fock matrices, and $\mathbf{F}^{\mathrm{cs}} = (\mathbf{F}^{\alpha} + \mathbf{F}^{\beta})/2$. At SCF convergence, all off-diagonal $\mathbf{F}_{\mathrm{ROHF}}$ terms become zero. The choice of the diagonal elements in Eq.(4.2) is completely arbitrary within a set of A and B coupling parameters:

$$\mathbf{R}_{cc} = A_{cc}\mathbf{F}_{cc}^{\alpha} + B_{cc}\mathbf{F}_{cc}^{\beta} \qquad (4.3a)$$

$$\mathbf{R}_{oo} = A_{oo}\mathbf{F}_{oo}^{\alpha} + B_{oo}\mathbf{F}_{oo}^{\beta} \qquad (4.3b)$$

$$\mathbf{R}_{vv} = A_{vv}\mathbf{F}_{vv}^{\alpha} + B_{vv}\mathbf{F}_{vv}^{\beta}. \qquad (4.3c)$$

Different values for these parameters have been suggested in the literature [72]. Although they do not affect the ROHF wave function and energy, they affect orbital energies whose physical meaning is obscured because of this dependence. Choices guided to determine "canonical" sets that satisfy Koopmans' theorem may result in violations of the aufbau principle [74, 76]. In the next section, these problems are resolved by abandoning the use of a single Fock operator. We will obtain the ROHF wave function by *projecting* the UHF wave function self-consistently.

4.2 Theory

We start from spin-contamination in UHF which is given by [5, 77–79]

$$\delta_s = \langle S^2 \rangle - S_z(S_z + 1) = N_\beta - \mathrm{Tr}(\boldsymbol{\gamma}^{\alpha}\boldsymbol{\gamma}^{\beta}) \qquad (4.4)$$

where $S_z = (N_\alpha - N_\beta)/2$ and N_σ ($\sigma = \alpha, \beta$) is the number of σ electrons in the system. It should be pointed out that if δ_s is made zero then this UHF wave function not

only gives the correct $\langle S^2 \rangle$ but is also necessarily a spin eigenstate. This is because the UHF wave function consists of determinants with $S \geq M$ and it cannot have determinants with $S < M$. Furthermore, since the wave function has to be a single Slater determinant, in the limit of zero δ_s the resulting state is the corresponding high-spin ROHF wave function. The previously proposed spin-constrained UHF (SUHF) approach [79] introduces a Lagrange multiplier λ in UHF to achieve this goal, $\delta_s = 0$. However, this is exact only in the limit of $\lambda \to \infty$. In this limit, the effective SUHF Fock matrices remain in the form of Eq.(4.2) [80]. Recently, Glushkov has suggested a similar approach [81]. We here propose an alternative method based on restricting natural occupations and spin density eigenvalues via finite Lagrange multipliers.

Before we proceed, we would like to recall that, in UHF, the natural occupations n are eigenvalues of $\mathbf{P}$; they can be 0, 1, $\frac{1}{2}$, or appear in "corresponding pairs" $(n, 1-n)$ [68]. This is a rigorous mathematical result following from $\mathbf{P}$ being the half sum of two idempotent density matrices [22]. In high-spin systems, the number of $\frac{1}{2}$ occupations is $N_\alpha - N_\beta = N_s$ (we assume $N_\alpha > N_\beta$ always). Note that $\mathrm{Tr}\mathbf{P} = (N_\alpha + N_\beta)/2 = N_e/2$, where N_e is the number of electrons. For clarity, we discuss below only the case where the number of orbitals N_{orbs} is greater than N_e, but our results hold for $N_{\mathrm{orbs}} \leq N_e$ too. The UHF $\boldsymbol{\gamma}^\sigma$ are block-diagonal in the NO basis:

$$\boldsymbol{\gamma}^\alpha = \begin{pmatrix} \boldsymbol{\gamma}_1^\alpha & & & & \\ & \ddots & & & \\ & & \boldsymbol{\gamma}_{N_{cp}}^\alpha & & \\ & & & \mathbf{I} & \\ & & & & \mathbf{0} \end{pmatrix}, \; \boldsymbol{\gamma}^\beta = \begin{pmatrix} \boldsymbol{\gamma}_1^\beta & & & \\ & \ddots & & \\ & & \boldsymbol{\gamma}_{N_{cp}}^\beta & \\ & & & \mathbf{0} \\ & & & & \mathbf{0} \end{pmatrix} \tag{4.5}$$

where N_{cp} is the number of corresponding pairs and is N_β for $N_{\mathrm{orbs}} > N_e$, and

$$\boldsymbol{\gamma}_i^\alpha = \begin{pmatrix} n_i & +m_i \\ +m_i & 1 - n_i \end{pmatrix}, \; \boldsymbol{\gamma}_i^\beta = \begin{pmatrix} n_i & -m_i \\ -m_i & 1 - n_i \end{pmatrix} \tag{4.6}$$

with $m_i = \sqrt{n_i - n_i^2}$. The identity matrix in $\boldsymbol{\gamma}^\alpha$ accounts for unpaired electrons, traces to N_s, and is substituted by a corresponding zero matrix in $\boldsymbol{\gamma}^\beta$. The other zero matrix has dimension $N_v = N_{\text{orbs}} - N_s - 2N_{cp}$ and corresponds to virtual $(n = 0)$ unpaired orbitals. In the NO basis, $\mathbf{M}$ is

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 & & & & \\ & \ddots & & & \\ & & \mathbf{M}_{N_{cp}} & & \\ & & & \frac{1}{2}\mathbf{I} & \\ & & & & \mathbf{0} \end{pmatrix}, \tag{4.7}$$

where, from Eq.(4.6), $\mathbf{M}_i = (\boldsymbol{\gamma}_i^\alpha - \boldsymbol{\gamma}_i^\beta)/2$ is

$$\mathbf{M}_i = \begin{pmatrix} 0 & m_i \\ m_i & 0 \end{pmatrix} \tag{4.8}$$

which is traceless with eigenvalues $\pm m_i$. The full spectrum of $\mathbf{M}$ also includes $\frac{1}{2}$ and 0 eigenvalues, thus tracing to $N_s/2$. Using the idempotency of $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$, we get

$$\text{Tr}(\boldsymbol{\gamma}^\alpha \boldsymbol{\gamma}^\beta) = \frac{N_e}{2} - 2\,\text{Tr}\mathbf{M}^2. \tag{4.9}$$

Considering Eqs.(4.7) and (4.8), it is evident that

$$\text{Tr}\mathbf{M}^2 = 2\sum_i^{N_{cp}} m_i^2 + \frac{N_s}{4}, \tag{4.10}$$

and hence,

$$\delta_s = N_\beta - \text{Tr}(\boldsymbol{\gamma}^\alpha \boldsymbol{\gamma}^\beta) = 4\sum_i^{N_{cp}} m_i^2. \tag{4.11}$$

From Eq.(4.11), it is immediately clear that to eliminate spin-contamination in UHF all m_i should be zero. Therefore, we propose to formulate ROHF as a constrained UHF scheme that enforces this condition. We stress that this is a very similar procedure to what we have done for CPMFT, *i.e.*, forcing κ_i to be zero for desired orbital spaces (see Section 3.2.3 and Appendix B). From Eq.(4.6), $m_i = 0$ implies

that corresponding pair occupations become constrained to values of 1 and 0, thus effectively creating core (c) and virtual (v) orbital blocks with multiple degeneracies within themselves. To enforce these constraints, we introduce Lagrange multipliers λ_{ij} and then write

$$\mathcal{L}_{\text{CUHF}} = E_{\text{UHF}} + 2\sum_{ij}{}' \lambda_{ij} M_{ij}, \tag{4.12}$$

where the factor of 2 is introduced for algebraic convenience and does not change the final result because the last term is always zero at convergence. The prime on the summation restricts it to cv and vc blocks. $\mathbf{M}$ is unconstrained in the oo block and zero in other blocks. We next derive equations for λ_{ij}.

The UHF energy is normally written as a functional of $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$. In Section 3.5, we have shown that the UHF energy expression can be written alternatively as $E_{\text{UHF}} = E_{\text{cs}}[\mathbf{P}] - E_{\text{c}}[\mathbf{M}]$, a functional of $\mathbf{P}$ and $\mathbf{M}$ (Eq.(3.82)). The derivatives of E_{cs} with respect to $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$ yield the usual closed-shell Fock matrix

$$\frac{\partial E_{\text{cs}}}{\partial \gamma_{ji}^\alpha} = \frac{\partial E_{\text{cs}}}{\partial \gamma_{ji}^\beta} = \frac{1}{2}\frac{\partial E_{\text{cs}}}{\partial P_{ji}} \equiv F_{ij}^{\text{cs}} = \frac{1}{2}(F_{ij}^\alpha + F_{ij}^\beta). \tag{4.13}$$

On the other hand, the derivatives of E_{c} are

$$\frac{\partial E_{\text{c}}}{\partial \gamma_{ji}^\alpha} = -\frac{\partial E_{\text{c}}}{\partial \gamma_{ji}^\beta} = -\sum_{kl}\langle ik|lj\rangle M_{kl} \equiv \Delta_{ij}^{\text{UHF}} = \frac{1}{2}(F_{ij}^\alpha - F_{ij}^\beta). \tag{4.14}$$

Hence,

$$\mathbf{F}^\alpha = \mathbf{F}^{\text{cs}} + \boldsymbol{\Delta}^{\text{UHF}} \tag{4.15a}$$

$$\mathbf{F}^\beta = \mathbf{F}^{\text{cs}} - \boldsymbol{\Delta}^{\text{UHF}}, \tag{4.15b}$$

which are the usual UHF Fock matrices. Now, the CUHF Fock matrices additionally require the derivatives of the constraints in Eq.(4.12) with respect to γ^α and γ^β, which

are trivially λ_{ij} and $-\lambda_{ij}$, respectively. Defining $\mathbf{\Delta}^{\mathrm{CUHF}}$ as

$$
\Delta_{ij}^{\mathrm{CUHF}} \equiv
\begin{cases}
\Delta_{ij}^{\mathrm{UHF}} + \lambda_{ij} & \text{if } \{i \in c \wedge j \in v\}, \\
 & \text{or } \{i \in v \wedge j \in c\} \\
\\
\Delta_{ij}^{\mathrm{UHF}} & \text{otherwise}
\end{cases}
\tag{4.16}
$$

yields the CUHF α and β Fock matrices,

$$
\tilde{\mathbf{F}}^\alpha = \mathbf{F}^{\mathrm{cs}} + \mathbf{\Delta}^{\mathrm{CUHF}},
\tag{4.17a}
$$

$$
\tilde{\mathbf{F}}^\beta = \mathbf{F}^{\mathrm{cs}} - \mathbf{\Delta}^{\mathrm{CUHF}}.
\tag{4.17b}
$$

At first glance, it might seem necessary to perform a $c \times v$ dimensional search for finding all λ_{ij}. They are, however, analytically determined by the CUHF equations, namely, $[\tilde{\mathbf{F}}^\alpha, \boldsymbol{\gamma}^\alpha] = 0$ and $[\tilde{\mathbf{F}}^\beta, \boldsymbol{\gamma}^\beta] = 0$. If these two equations are subtracted and divided by 2, after some elementary algebra one arrives at

$$
[\mathbf{F}^{\mathrm{cs}}, \mathbf{M}] + [\mathbf{\Delta}^{\mathrm{CUHF}}, \mathbf{P}] = 0.
\tag{4.18}
$$

Partitioning these matrices into core, open, and virtual blocks gives,

$$
\mathbf{F}_{co}^{\mathrm{cs}} + \mathbf{\Delta}_{co}^{\mathrm{CUHF}} = \tilde{\mathbf{F}}_{co}^\beta = 0,
\tag{4.19a}
$$

$$
\mathbf{F}_{vo}^{\mathrm{cs}} - \mathbf{\Delta}_{vo}^{\mathrm{CUHF}} = \tilde{\mathbf{F}}_{vo}^\alpha = 0,
\tag{4.19b}
$$

$$
\mathbf{\Delta}_{cv}^{\mathrm{CUHF}} = 0,
\tag{4.19c}
$$

where we have used $\mathbf{P}_{cc} = \mathbf{I}$, $\mathbf{P}_{vv} = \mathbf{P}_{cv} = \mathbf{P}_{co} = \mathbf{P}_{vo} = 0$, and $\mathbf{P}_{oo} = \mathbf{M}_{oo} = \frac{1}{2}\mathbf{I}$. Together with Eq.(4.16), Eq.(4.19c) implies that $\boldsymbol{\lambda}_{cv} = -\mathbf{\Delta}_{cv}^{\mathrm{UHF}}$ at convergence. It must be noted that during the iterative procedure, $\boldsymbol{\lambda}$ can have different values, simply because Eq.(4.19c) is satisfied only at convergence. Therefore, at each SCF cycle, this choice does not guarantee $\delta_s = 0$. We will demonstrate, however, that with this choice of $\boldsymbol{\lambda}$ at every iteration, the CUHF calculations converge rapidly. Therefore, we choose

this value for $\boldsymbol{\lambda}_{cv}$. Also note that Eqs.(4.19) yield the SCF conditions for Roothaan's ROHF. Finally, our CUHF α and β Fock matrices are

$$\tilde{\mathbf{F}}^{\alpha} = \begin{pmatrix} \mathbf{F}_{cc}^{\alpha} & \mathbf{F}_{co}^{\alpha} & \mathbf{F}_{cv}^{cs} \\ \mathbf{F}_{oc}^{\alpha} & \mathbf{F}_{oo}^{\alpha} & \mathbf{F}_{ov}^{\alpha} \\ \mathbf{F}_{vc}^{cs} & \mathbf{F}_{vo}^{\alpha} & \mathbf{F}_{vv}^{\alpha} \end{pmatrix}, \qquad \tilde{\mathbf{F}}^{\beta} = \begin{pmatrix} \mathbf{F}_{cc}^{\beta} & \mathbf{F}_{co}^{\beta} & \mathbf{F}_{cv}^{cs} \\ \mathbf{F}_{oc}^{\beta} & \mathbf{F}_{oo}^{\beta} & \mathbf{F}_{ov}^{\beta} \\ \mathbf{F}_{vc}^{cs} & \mathbf{F}_{vo}^{\beta} & \mathbf{F}_{vv}^{\beta} \end{pmatrix}. \qquad (4.20)$$

These CUHF Fock matrices differ from the UHF ones only in the cv and vc blocks, and are different from Roothaan's effective Fock matrix of Eq.(4.2). Our CUHF procedure yielding ROHF is surprisingly straightforward: one simply performs UHF with Fock matrices replaced by Eqs.(4.20). A more detailed algorithm is described in Appendix E. These Fock matrices eliminate the ambiguities arising in ROHF theory and produce a more physical UHF-like picture. In open-shell molecules, α and β electrons feel different potentials; our $\tilde{\mathbf{F}}^{\alpha}$ and $\tilde{\mathbf{F}}^{\beta}$ operators are not identical to each other and yield α orbitals different from β orbitals that are true "canonical orbitals" obtained by diagonalization. However, unlike UHF, they have no spin-contamination, which is removed by Lagrangian constraints. Their eigenvalues ε_i^{σ} are physical orbital energies in the sense that they are associated with individual α and β orbitals, satisfy Koopmans' theorem, and the aufbau principle [82], as opposed to many ROHF canonicalizations of Eq.(4.2) [76]. Our orbitals have been previously proposed in the literature as semi-canonical orbitals for Møller-Plesset perturbation theory [83] and used in an *ad hoc* fashion with ROHF [84, 85]. Our present work shows that the Fock matrices for which these orbitals are eigenfunctions appear from a constrained UHF optimization that eliminates spin-contamination.

4.3   Results

4.3.1   Convergence

We have implemented CUHF in the `Gaussian` suite of programs [70] and verified that our procedure converges to the ROHF energy. Table 4.1 shows the number of SCF cycles that it took to converge calculations for difficult open-shell cases. Unlike many ROHF schemes, CUHF has no issues with SCF convergence and its behavior is similar to UHF. This is undoubtedly related to the satisfaction of the aufbau principle in our method.

4.3.2   Koopmans' theorem for ionization potentials

In HF, an N_e-electron Slater determinant $|\Psi^{N_e}\rangle$ has occupied and virtual spin orbitals with orbital energies ε_p and ε_v. Koopmans' theorem states that $-\varepsilon_p$ is identical to the ionization potential computed as the difference in energy between $|\Psi^{N_e}\rangle$ and the ionized state $|\Psi_p^{N_e-1}\rangle$ by removing an electron from ϕ_p. Similarly, $-\varepsilon_v$ corresponds to the electron affinity as the difference in energy between $|\Psi^{N_e}\rangle$ and the state $|\Psi_v^{N_e+1}\rangle$ by adding an electron to ϕ_v. This is a well-known approximation both in RHF and UHF, but regular ROHF schemes do *not* obey this theorem. Note that Koopmans' theorem is valid only for the highest-occupied-molecular-orbital (HOMO) in KS-DFT, but generally orbitals do not have any physical meaning.

Since Koopmans' theorem is valid for CUHF [86], orbital energies approximate ionization potentials (IP) and electron affinities. In Table 4.2 we summarize first IPs estimated via HOMO energies, $\varepsilon_{\mathrm{HOMO}}$, for 24 open-shell compounds selected from the G2 set [87], as well as the mean (ME) and mean absolute errors (MAE). Molecular geometries are optimized with B3LYP/6-31G(2df,p) [35]. CUHF results with a 6-

311++G(3df,3pd) basis are compared to UHF and the default ROHF implementation in `Gaussian` (parameters of McWeeny and Diercksen [88] denoted as MD). In all systems, the CUHF $\varepsilon_{\mathrm{HOMO}}$ captures the correct physics, yielding IPs comparable to those of UHF, yet preserving the correct $\langle S^2 \rangle$ expectation value.

We have compared our CUHF orbital energies with those obtained by Eq.(4.2) with the parameters recently suggested by Plakhutin, Gorelik, and Breslavskaya (PGB) [72] and Davidson and Plakhutin (DP) [75]. These parameterizations are chosen to obey Koopmans' theorem. However, both schemes usually yield poor SCF convergence (Table 4.1). Therefore, as a simple remedy to obtain PGB and DP results in this work, we have used the converged ROHF wave function and then diagonalized Eq.(4.2) with DP parameters in a single shot. The eigenvalues thus obtained are identical to those from the self-consistent PGB and DP schemes. In Table 4.2, we also list the ME and MAE of the PGB scheme. For most systems, PGB gives first IPs very similar to CUHF. In Table 4.3, we present valence orbital energies for a model high spin transition metal complex $MnCl_2(H_2O)_2$ in C_{2v} symmetry. This is an example where ionization from closed shells is easier than ionization from ROHF open shells [89]. Again, CUHF poses no convergence problems and the orbital energies are in fair agreement with those of PGB and DP.

Why are the results obtained from our scheme and the PGB canonicalization of ROHF [72] similar? The ROHF equations require that at convergence $\mathbf{F}_{co}^{\beta}$, $\mathbf{F}_{ov}^{\alpha}$, and $\mathbf{F}_{cv}^{\mathrm{cs}}$ are all $\mathbf{0}$ [71]. Then the PGB Fock matrix becomes [72]

$$\mathbf{F}_{\mathrm{PGB}} = \begin{pmatrix} \mathbf{F}_{cc}^{\beta} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{oo}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_{vv}^{\alpha} \end{pmatrix}, \tag{4.21}$$

Table 4.1 : Number of SCF cycles to convergence using default criteria in the Gaussian 09 program (guess=Harris, no damping, no level-shift, DIIS) for representative open-shell molecules. Convergence threshold is set to Tight. NC means "no convergence" within 128 SCF cycles.

Method	$O_2{}^a$	$NO_2{}^b$	$MnCl_2(H_2O)_2{}^c$	LiH^{-d}	$C_6H_5{}^e$	Fe^f	Mn^f	Co^f
MD^g	11	19	22	NC	14	NC	10	NC
Roothaan^h	10	15	20	NC	14	NC	12	NC
Davidson^i	8	14	NC	NC	NC	NC	NC	NC
Guest et al.^{j}	10	16	21	NC	14	NC	10	NC
Binkley et al.^{k}	8	16	NC	NC	NC	NC	NC	NC
Faegri et al.^{l}	8	14	50	NC	NC	NC	NC	NC
GAMESS GVB^m	8	15	NC	NC	14	11	15	12
PGB^n	8	NC	NC	NC	NC	NC	NC	NC
Davidson et al.^{o}	NC	29	NC	NC	NC	NC	NC	NC
UHF	10	16	14	30	NC	10	10	11
CUHF (this work)	9	16	13	22	14	10	10	10

^{a}R$_{O-O}$=1.20752Å. aug-cc-pVTZ.
^{b}R$_{N-O}$=1.1934Å, $\angle_{O-N-O}$= 134.1°. aug-cc-pVTZ.
^{c}High-spin sextet, C_{2v} symmetry. 6-31G(d,p) (5d,7f). Geometry optimized with
 UB3LYP/6-31G(d,p).
^{d}R$_{Li-H}$=10Å. 3-21G.
^{e}6-31G(d). Geometry optimized with UB3LYP/6-31G(d).
^{f}6-31G(d). The ground state electronic configuration.
^{g}Ref. [88].
^{h}Ref. [71].
^{i}E. R. Davidson, Chem. Phys. Lett. **21**, 565 (1973).
^{j}M. F. Guest and V. R. Saunders, Mol. Phys. **28**, 819 (1974).
^{k}J. S. Binkley et al., Mol. Phys. **28**, 1423 (1974).
^{l}K. Faegri and R. Manne, Mol. Phys. **31**, 1037 (1976).
^{m}For details, see Ref.[72]
^{n}Ref. [72].
^{o}Ref. [75].

Table 4.2 : $\varepsilon_{\text{HOMO}}$ of open-shell systems (in eV).

System	ROHF		CUHF	UHF	Exptl.[a]
	MD[b]	PGB[c]			IP
H	-3.40	-13.60	-13.60	-13.60	13.60
Li	-1.44	-5.34	-5.34	-5.34	5.39
B	-1.57	-8.44	-8.44	-8.67	8.30
C	-2.38	-11.80	-11.80	-11.95	11.26
N	-3.29	-15.46	-15.46	-15.55	14.54
O	-4.87	-14.37	-14.37	-14.21	13.61
F	-6.55	-18.62	-18.62	-18.54	17.42
Na	-1.35	-4.95	-4.95	-4.95	5.14
Al	-1.22	-5.72	-5.72	-5.94	5.98
Si	-1.98	-8.09	-8.09	-8.20	8.15
P	-2.85	-10.66	-10.66	-10.67	10.49
S	-4.07	-10.11	-10.11	-10.30	10.36
Cl	-5.40	-13.00	-13.00	-13.09	12.97
OH	-4.48	-14.13	-14.13	-13.98	13.01
PH_2	-2.89	-9.94	-9.94	-10.25	9.82
SH	-4.00	-10.31	-10.31	-10.35	10.37
NH	-3.25	-13.79	-13.79	-13.82	13.49
O_2	-3.86	-14.52	-14.52	-15.25	12.07
S_2	-3.34	-10.05	-10.05	-10.46	9.36
CH_3	-2.01	-10.18	-10.18	-10.46	9.84
C_2H_5	-1.65	-9.54	-9.25	-9.51	8.12
CN	-6.21	-13.68	-13.68	-14.17	13.60
HCO	-2.60	-10.88	-10.40	-10.73	8.14
CH_3O	-3.93	-12.29	-12.29	-12.16	10.73
ME	-7.38	0.57	0.54	0.68	
MAE	7.38	0.64	0.61	0.71	

[a]Ref. [87].
[b]Ref. [88].
[c]Ref. [72].

Table 4.3 : Orbital energies of $MnCl_2(H_2O)_2$ (in eV).

MO	PGB+DP		CUHF	
	α	β	α	β
b_2	-11.310	-11.079	-10.807	-11.079
a_2	-11.415	-11.159	-10.909	-11.159
a_1	-11.362	-11.168	-10.910	-11.168
b_1	-11.327	-11.162	-11.195	-11.162
b_2	-12.083	-11.676	-11.741	-11.676
a_1	-13.112	-12.405	-11.684	-12.405
$3d_{x^2-y^2}^{Mn}$	-15.748		-15.742	
b_1	-16.470	-16.319	-16.635	-16.319
$3d_{yz}^{Mn}$	-16.552		-17.262	
a_1	-16.613	-16.417	-17.728	-16.417
$3d_{zx}^{Mn}$	-16.658		-15.811	
$3d_{xy}^{Mn}$	-16.701		-17.129	
$3d_{z^2}^{Mn}$	-16.924		-17.095	
a_1	-18.942	-18.665	-19.257	-18.665

while the α and β CUHF Fock matrices are

$$\tilde{\mathbf{F}}^\alpha = \begin{pmatrix} \mathbf{F}_{cc}^\alpha & \mathbf{F}_{co}^\alpha & \mathbf{0} \\ \mathbf{F}_{oc}^\alpha & \mathbf{F}_{oo}^\alpha & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_{vv}^\alpha \end{pmatrix}, \tag{4.22a}$$

$$\tilde{\mathbf{F}}^\beta = \begin{pmatrix} \mathbf{F}_{cc}^\beta & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{oo}^\beta & \mathbf{F}_{ov}^\beta \\ \mathbf{0} & \mathbf{F}_{vo}^\beta & \mathbf{F}_{vv}^\beta \end{pmatrix}. \tag{4.22b}$$

At convergence, both methods diagonalize the same blocks of the UHF Fock matrices, *i.e.*, $\mathbf{F}_{cc}^\beta$ and $\mathbf{F}_{vv}^\alpha$. Thus, the β occupied and α virtual orbitals in CUHF correspond to the PGB core (closed-shell) and virtual orbitals, respectively. For the open-shell space, however, the methods do not necessarily yield the same orbitals. In other words, while PGB diagonalizes only the $\mathbf{F}_{oo}^\alpha$ sub-block, CUHF treats the whole occupied α space

in Eq.(4.22a), where the off-diagonal blocks $\mathbf{F}_{co}^{\alpha}$ and $\mathbf{F}_{oc}^{\alpha}$ are non-zero in general. As a result, the canonical α CUHF orbitals (in which $\tilde{\mathbf{F}}^{\alpha}$ is diagonal) may not possess the "open-shell" orbitals explicitly, but they are represented as linear combinations of the core and open-shell α natural orbitals.

4.3.3   Excitation energies

Last, we present excitation energies of five small open-shell molecules calculated with time-dependent HF (TDHF) based on UHF and CUHF with a 6-311++G(3df,3pd) basis. The bond lengths for BeF and CO^{+} (not included in the G2 set) are 1.355 and 1.078 Å, respectively. For TD-CUHF, we have used CUHF orbitals and orbital energies in the TD-UHF procedure. Although this TD-CUHF scheme is not rigorous (one should calculate the linear response of $\tilde{\mathbf{F}}$ instead of $\mathbf{F}$), this simple approximation turns out to be quite reasonable, as shown in Table 4.4. When the UHF spin-contamination (δ_s) is small, TD-UHF and TD-CUHF give very similar results. As δ_s becomes larger, however, TD-UHF greatly overestimates the excitation energies. On the other hand, by retaining a spin projected reference ($\delta_s = 0$), TD-CUHF gives more reasonable excitation energies, outperforming TD-UHF in spin-contaminated situations.

Table 4.4 : TDHF valence (V) and Rydberg (R) excitation energies (in eV) of open-shell molecules. Numbers in parentheses are the UHF spin-contamination δ_s.

System	State	CUHF	UHF	Exptl.[a]
BeF	V $^2\Pi$	4.19	4.20	4.14
(0.001)	R $^2\Sigma^+$	6.33	6.34	6.16
	R $^2\Sigma^+$	6.54	6.54	6.27
BeH	V $^2\Pi$	2.64	2.69	2.48
(0.002)	R $^2\Pi$	6.25	6.26	6.32
CH$_3$	R ^{2}A$_1'$	6.23	6.54	5.73
(0.012)	R ^{2}A$_2''$	7.34	7.73	7.44
CO$^+$	V $^2\Pi$	4.84	6.93	3.26
(0.141)	V $^2\Sigma^+$	9.81	11.10	5.82
CN	V $^2\Pi$	0.85	4.11	1.32
(0.397)	V $^2\Sigma^+$	1.62	5.41	3.22
ME		0.41	1.43	
MAE		0.81	1.44	

[a]Taken from Ref. [90].

Chapter 5

Constrained Active Space Unrestricted Mean-Field Theory

5.1 Introduction

In this chapter, we extend CUHF and use constraints to allow or remove spin-contamination depending on different orbital spaces. The spirit here is in many ways similar to CPMFT where we allow an attractive pairing electron-electron interaction (and the concomitant particle number symmetry breaking) in an active space. The premise here is to benefit from what symmetry breaking has to offer while limiting its negative effects. In this sense, our specific objective is to allow spin-contamination in a limited active orbital space while eliminating it from the remainder. As shown below, this approach can be used in myriad ways with very promising results in terms of the quality of the resulting wave functions and energies. If the active space is properly chosen, our generalized CUHF method benefits from a controlled broken-symmetry effect while it avoids the negative consequences of massive spin-contamination characteristic of UHF in so many cases.

We test the validity of this model on singlet-triplet splittings of several molecules. While the triplet state can usually be well approximated by a single Slater determinant (high-spin ROHF), the diradical singlet state requires a multi-determinant wave function. To tackle this problem, we apply two different approximations: Noodleman's approach with broken-symmetry solutions [91, 92] and Löwdin's projection

operator scheme [67, 93, 94]. To include weak correlation effects, we also present a simple extension of our model to second-order Møller-Plesset perturbation theory (MP2) [83].

In Section 5.2.1, we present our generalized CUHF theory by following the previous derivations in Chapter 4. In Section 5.2.2, we discuss the character of the unpaired orbitals in the broken-symmetry solutions and introduce a useful concept called spin-deviation. We then consider in Section 5.2.3 Noodleman's approach and Löwdin's spin-projection scheme for calculating singlet-triplet energy splittings, and extend MP2 for CUHF in Section 5.2.4. In Section 5.2.5, the computational details of our benchmark calculations are given. Section 5.3 investigates the size-consistency problem in projected UHF (PUHF) and discusses several benchmarks to compare different schemes resulting from projecting CUHF. In Appendix E, we also present a detailed CUHF algorithm for easy computational implementation.

5.2 Theory: CUHF with an active space

5.2.1 Generalization of CUHF

Spin-contamination in single determinant wave functions has many pros and cons. On the one hand, it offers a description of radicals, for which broken-symmetry (BS) UHF solutions have been widely used [47, 95]. For clarity, it should be noted that in this chapter we use the term "broken-symmetry" exclusively for the low-spin solutions, $M = 0$. The broken-symmetry solutions tend to localize strongly-correlated electrons and can thus treat most bond-breaking processes. On the other hand, the broken-symmetry wave function can sacrifice many other properties because it is not a pure spin-state. As mentioned in Section 5.1, we here propose a method to selectively

control spin-contamination in UHF. To achieve this goal, we open an active space in CUHF, *i.e.*, we allow non-zero δ_s only in a limited orbital space. By doing this, we can avoid undesired consequences of spin-contamination while benefiting from a controlled broken-symmetry effect. Accordingly, we divide the orbital space into core (c), virtual (v), and active (a), where the latter always includes all open-shell orbitals. We shall call this method CUHF(N_a), where N_a is the number of active orbitals and satisfies $N_s \leq N_a \leq N_e$. It is evident that CUHF(N_s) is the previously developed CUHF for ROHF. Note that due to the corresponding pairs property in CUHF, the number of active electrons is always equal to the number of active orbitals in a manner similar to corresponding pair CPMFT in Section 3.5. A pictorial description of CUHF(2) is presented in Figure 5.1.

We again consider Eq.(4.4) for spin-contamination δ_s in UHF. Now, in the NO basis, δ_s can be divided into two contributions,

$$\delta_s = \delta_s^{cv} + \delta_s^a, \tag{5.1}$$

where the first and second terms are spin-contamination in the core-virtual and active spaces, respectively. The idea in CUHF(N_a) is to impose $\delta_s^{cv} = 0$,

$$\delta_s^{cv} = N_c - \sum_{ij}^{N_c} |\langle \phi_i^\alpha \mid \phi_j^\beta \rangle|^2 \tag{5.2a}$$

$$= 4 \sum_i^{N_c} m_i^2 \to 0, \tag{5.2b}$$

while leaving δ_s^a unconstrained. Here N_c is the number of core orbitals (doubly-occupied, see Figure 5.1), and therefore $2N_c$ is the number of core electrons. The CUHF conditions can be satisfied by introducing Lagrange multipliers λ_{ij}. The resulting energy expression is essentially the same as Eq.(4.12),

$$\mathcal{L}_{\text{CUHF}(N_a)} = E_{\text{UHF}} + 2 \sum_{ij}' \lambda_{ij} M_{ij}, \tag{5.3}$$

where, similarly to Eq.(4.12), the prime on the summation in Eq.(5.3) indicates that $i \in c \wedge j \in v$ or $i \in v \wedge j \in c$. Therefore, the Lagrange multiplier matrix $\boldsymbol{\lambda}$ has the form

$$
\boldsymbol{\lambda} = \begin{pmatrix} \overset{c}{\mathbf{0}} & \overset{a}{\mathbf{0}} & \overset{v}{\boldsymbol{\lambda}_{cv}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\lambda}_{vc} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{matrix} c \\ a \\ v \end{matrix} .
\tag{5.4}
$$

$\boldsymbol{\lambda}$ is Hermitian, and hence $\boldsymbol{\lambda}_{cv} = \boldsymbol{\lambda}_{vc}^{\dagger}$. The CUHF Fock matrices $\tilde{\mathbf{F}}^{\sigma}$ are the derivatives of Eq.(5.3), so recalling that $\mathbf{M} = (\boldsymbol{\gamma}^{\alpha} - \boldsymbol{\gamma}^{\beta})/\mathbf{2}$,

$$
\tilde{\mathbf{F}}^{\alpha} = \mathbf{F}^{\alpha} + \boldsymbol{\lambda},
\tag{5.5a}
$$

$$
\tilde{\mathbf{F}}^{\beta} = \mathbf{F}^{\beta} - \boldsymbol{\lambda},
\tag{5.5b}
$$

where $F_{ij}^{\sigma} = \partial E_{\mathrm{UHF}}/\partial \gamma_{ij}^{\sigma}$ are the UHF Fock matrices.

The CUHF SCF equations are still given by $[\tilde{\mathbf{F}}^{\alpha}, \boldsymbol{\gamma}^{\alpha}] = \mathbf{0}$ and $[\tilde{\mathbf{F}}^{\beta}, \boldsymbol{\gamma}^{\beta}] = \mathbf{0}$, and therefore by Eq.(4.18). Due to Eq.(5.2b), we obtain the following forms for $\mathbf{P}$ and $\mathbf{M}$ in the NO basis:

$$
\mathbf{P} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{aa} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},
\tag{5.6a}
$$

$$
\mathbf{M} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{aa} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}.
\tag{5.6b}
$$

Substituting Eqs.(5.6a) and (5.6b) into Eq.(4.18) yields

$$
\boldsymbol{\Delta}_{cv}^{\mathrm{UHF}} + \boldsymbol{\lambda}_{cv} = \mathbf{0},
\tag{5.7a}
$$

$$
\mathbf{F}_{ca}^{\mathrm{cs}} \mathbf{M}_{aa} + \boldsymbol{\Delta}_{ca}^{\mathrm{UHF}} (\mathbf{I} - \mathbf{P}_{aa}) = \mathbf{0},
\tag{5.7b}
$$

$$
[\mathbf{F}_{aa}^{\mathrm{cs}}, \mathbf{M}_{aa}] + [\boldsymbol{\Delta}_{aa}^{\mathrm{UHF}}, \mathbf{P}_{aa}] = \mathbf{0}.
\tag{5.7c}
$$
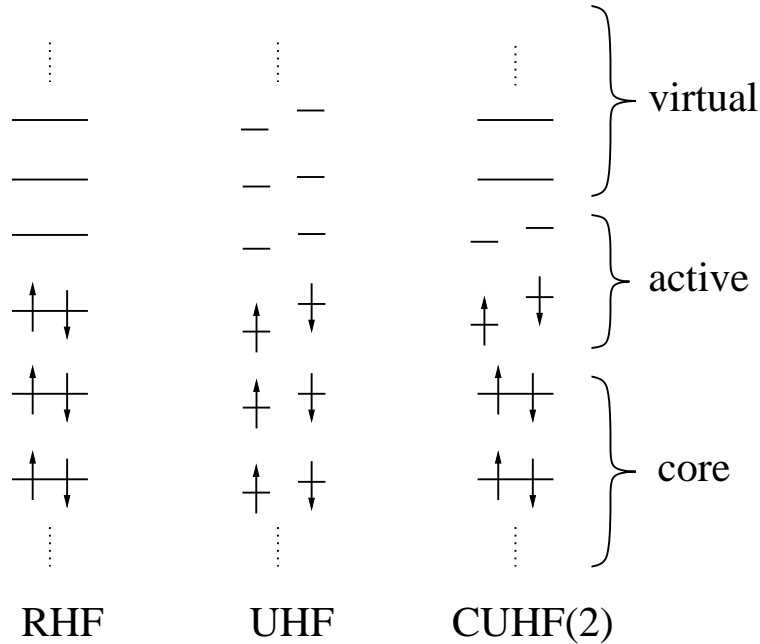
Figure 5.1 : Spin-symmetry preserved orbitals and broken-symmetry orbitals in each method. $CUHF(N_a)$ is a mixture of RHF and UHF, having RHF-type symmetry-preserved orbitals in the core and virtual spaces and UHF-type broken-symmetry orbitals in the active space. This figure represents a case of $N_a = 2$.

Together with Eq.(5.4), Eq.(5.7a) gives the necessary and sufficient conditions for determining $\boldsymbol{\lambda}$ at convergence: it consists of the cv and vc parts of $-\boldsymbol{\Delta}^{\text{UHF}}$.

$$\tilde{\mathbf{F}}^\alpha = \begin{pmatrix} \mathbf{F}^\alpha_{cc} & \mathbf{F}^\alpha_{ca} & \mathbf{F}^{\text{cs}}_{cv} \\ \mathbf{F}^\alpha_{ac} & \mathbf{F}^\alpha_{aa} & \mathbf{F}^\alpha_{av} \\ \mathbf{F}^{\text{cs}}_{vc} & \mathbf{F}^\alpha_{va} & \mathbf{F}^\alpha_{vv} \end{pmatrix}, \tag{5.8a}$$

$$\tilde{\mathbf{F}}^\beta = \begin{pmatrix} \mathbf{F}^\beta_{cc} & \mathbf{F}^\beta_{ca} & \mathbf{F}^{\text{cs}}_{cv} \\ \mathbf{F}^\beta_{ac} & \mathbf{F}^\beta_{aa} & \mathbf{F}^\beta_{av} \\ \mathbf{F}^{\text{cs}}_{vc} & \mathbf{F}^\beta_{va} & \mathbf{F}^\beta_{vv} \end{pmatrix}. \tag{5.8b}$$

If $N_a = N_e$, our active-space CUHF scheme is equivalent to full UHF because there is no c space (all electrons are active), so $\boldsymbol{\lambda} = \mathbf{0}$ and consequently $\tilde{\mathbf{F}}^\sigma = \mathbf{F}^\sigma$. If $N_a = N_s$, *i.e.*, if one makes the active space identical with the open-shell space, then the method is exactly the same as our previously proposed CUHF, shown to converge
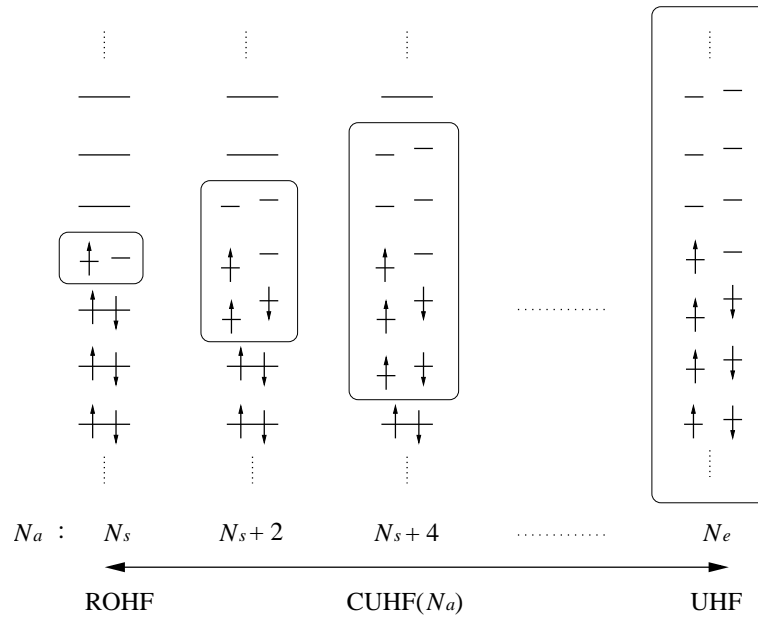
Figure 5.2 : CUHF(N_a) as a function of N_a. The rectangles depict the active spaces. CUHF(N_a) can be ROHF, UHF, or an intermediate of them, depending on the choice of the active space.

to the high-spin ROHF solution in a robust fashion. Therefore, CUHF(N_a) may be considered an interpolation between ROHF and UHF (Figure 5.2).

Generally, one should choose the active space to include the most important orbitals, such as unpaired orbitals. Thus, we choose the N_a α and β spin-orbitals near the Fermi energy as active orbitals and allow them to break symmetry while other orbitals (categorized as core and virtual) retain it. In this way, one can eliminate undesired spin-contamination that may arise from irrelevant orbitals far from the Fermi energy. Of course, it is obvious from Eq.(4.11) that the most significant contribution to δ_s in UHF is made by the NOs (except open-shell orbitals) whose n_i significantly deviate from 1 and 0. In most cases, these orbitals are required to describe the physics of the process under consideration (like bond-breaking, diradical character, etc.). The

aim of CUHF is to limit symmetry breaking to only these physically-relevant orbitals. In summary, we will consider cases where the active space is limited to this subset of significant NOs while making the rest inactive (c or v).

Finally, the CUHF concepts discussed above are not limited to HF and they are applicable to KS-DFT [45], provided that $\langle \hat{S}^2 \rangle$ in KS-DFT is computed in the same way as in HF and δ_s is also given by Eq.(4.4).

5.2.2 Measure of singly-occupied character and spin-deviation

Suppose that we have obtained a UHF broken-symmetry solution for a diradical compound. If the UHF wave function characterizes the diradical well, then all UHF occupations are nearly 1 and 0 except for one corresponding-pair with occupations n_k and $1 - n_k$, where n_k should be around 0.5. The kth sub-block density matrix $\boldsymbol{\gamma}_k^\alpha$ in Eq.(4.5) can be diagonalized to give α occupied and virtual orbitals,

$$\mathbf{V}^\dagger \boldsymbol{\gamma}_k^\alpha \mathbf{V} = \mathbf{V}^\dagger \begin{pmatrix} n_k & m_k \\ m_k & 1 - n_k \end{pmatrix} \mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.9}$$

Let the first column in the right-hand side of Eq.(5.9) be orbital ϕ_k^α and the second column $\phi_{k'}^\alpha$. This immediately means that the k (k') orbital is occupied by one (zero) α electron. Similarly, we can also diagonalize the β part, γ_k^β, to obtain occupied ϕ_k^β and unoccupied $\phi_{k'}^\beta$ orbitals. The overlap between the ϕ_k^α and ϕ_k^β orbitals, $S_{\alpha\beta}$, measures the diradical character of the system, *i.e.*, how localized the α and β unpaired electrons are [96],

$$S_{\alpha\beta} = \langle \phi_k^\alpha | \phi_k^\beta \rangle = 2n_k - 1. \tag{5.10}$$

For example, if $S_{\alpha\beta} = 0$, there is zero spatial overlap, so the system is purely diradical. If $S_{\alpha\beta} = 1$, the orbitals are completely delocalized, there are no radicals associated with ϕ_k^α and ϕ_k^β, and the solution corresponds to R(O)HF. It is easy to show that

for CUHF(2), $S_{\alpha\beta}$ is directly related to spin-contamination in the broken-symmetry singlet CUHF solution,

$$\delta_s = 1 - S_{\alpha\beta}^2 = 4(n_k - n_k^2). \tag{5.11}$$

The above equation is closely related to the concepts of odd electron distribution of Yamaguchi and coworkers [47], and the effectively unpaired electrons of Staroverov and Davidson [97]. Eq.(5.11) also has a connection with the anomalous density $\kappa_k = \sqrt{n_k - n_k^2}$ in CPMFT. It can be assumed from Eq.(5.11) that in the broken-symmetry CUHF(2) wave functions, only the ϕ_k^α and ϕ_k^β orbitals carry spin-contamination and are sufficient for characterizing the diradical character of the system. In UHF, not only these orbitals but also other orbitals bring additional spin-contamination. If this additional contamination is non-negligible, then the UHF wave function no longer represents a diradical, but it contains more unpaired electrons than needed. Based on this fact, it is useful to introduce the concept of *spin-deviation* σ_s, which is a measure of how different the CUHF and UHF wave functions are for a given system. Spin-deviation may be defined by

$$\sigma_s = \langle S^2 \rangle_{\text{UHF}} - \langle S^2 \rangle_{\text{CUHF}}, \tag{5.12}$$

which is exactly the additional spin-contamination in UHF irrelevant to the diradical character. It is worth noting that if σ_s is zero, the UHF and CUHF(N_a) wave functions are equivalent.

5.2.3 Singlet-triplet splittings with CUHF(2)

Singlet-triplet energy differences (gaps) are the most common multiplet-splittings, and are a fundamental property in organic chemistry. They represent important excited states in organic diradical compounds, and are also connected to magnetic

coupling constants in the Heisenberg Hamiltonian. Here, we discuss two approximations to evaluate singlet-triplet gaps.

Noodleman's approximation

The Heisenberg Hamiltonian $\hat{\mathcal{H}}_H$ between fragments A and B is

$$\hat{\mathcal{H}}_H = -J\hat{S}_A \cdot \hat{S}_B, \tag{5.13}$$

where J is the coupling constant. For the case of two interacting spins, J is exactly the energy difference between the singlet and triplet states of the system, ΔE_{ST}. Noodleman proposed to use the high-spin and broken-symmetry solutions to approximate J by [91, 92, 98–100]

$$J = \frac{2(E(\text{BS}) - E(\text{HS}))}{\langle S^2 \rangle_{\text{HS}} - \langle S^2 \rangle_{\text{BS}}}, \tag{5.14}$$

which has been widely used within KS-DFT. Using Eq.(5.14), one can approximate ΔE_{ST} by

$$\Delta E_{\text{ST}}^{\text{N}} = \frac{2(E(\text{BS}) - E(\text{T}))}{\langle S^2 \rangle_{\text{T}} - \langle S^2 \rangle_{\text{BS}}}, \tag{5.15}$$

which can be considered a generalization of Ziegler's approximation [101, 102],

$$\Delta E_{\text{ST}}^{\text{Z}} = 2\left(E(\text{BS}) - E(\text{T})\right). \tag{5.16}$$

While Ziegler's approximation Eq.(5.16) requires the broken-symmetry wave function to be a 50-50% mixture of singlet and triplet, Noodleman's approximation does not.

Spin Projection

Consider the following two-determinant singlet wave function,

$$|\Psi_{\text{S}}\rangle = c_1|\psi_r^\alpha \psi_r^\beta\rangle + c_2|\psi_s^\alpha \psi_s^\beta\rangle, \tag{5.17}$$

where c_1 and c_2 are variationally optimized coefficients. We can approximate $|\Psi_S\rangle$ by projecting a single-determinant broken-symmetry wave function $|\Phi_{BS}\rangle$. As a matter of fact, $|\Phi_{BS}\rangle$ does contain $|\Psi_S\rangle$. It is known that the $M = 0$ UHF wave function can be expanded as [68, 78, 103]

$$|\Phi_{UHF}\rangle = d_1|\psi_r^\alpha \psi_r^\beta\rangle + d_2|\psi_s^\alpha \psi_s^\beta\rangle + d_3(|\psi_r^\alpha \psi_s^\beta\rangle + |\psi_r^\beta \psi_s^\alpha\rangle) + \cdots, \tag{5.18}$$

where for simplicity we omit closed-shells below ψ_r and ψ_s. Spin-contamination terms with $S + 1, S + 2, \cdots$ correspond to excited determinants. The true singlet state is a linear combination of determinants which include excitations from the closed-shells too, but this is weak correlation and it will be accounted for using MP2. The important strong correlation is captured by Eq.(5.18). Similarly, it can be shown that the $M = 0$ CUHF(2) wave function can be expressed as

$$|\Phi_{CUHF(2)}\rangle = d_1'|\psi_r^\alpha \psi_r^\beta\rangle + d_2'|\psi_s^\alpha \psi_s^\beta\rangle + d_3'(|\psi_r^\alpha \psi_s^\beta\rangle + |\psi_r^\beta \psi_s^\alpha\rangle), \tag{5.19}$$

without any extra determinants. Note that the first two terms correspond to Eq.(5.17), while the last term is the low-spin triplet. The first two terms can be extracted using Löwdin's projection operator $\hat{P}_s$,

$$\hat{P}_s = \sum_{l \neq s} \hat{A}_{s,l}, \tag{5.20}$$

where

$$\hat{A}_{s,l} = \frac{\hat{S}^2 - l(l+1)}{s(s+1) - l(l+1)}. \tag{5.21}$$

Upon spin-projection, the resulting projected UHF (PUHF) wave function is

$$|\Phi_{PUHF}\rangle = \hat{P}_s|\Phi_{UHF}\rangle = |\Phi_{UHF}\rangle + |\tilde{\Phi}_{UHF}\rangle, \tag{5.22}$$

with

$$|\tilde{\Phi}_{UHF}\rangle = \sum_i \frac{|\Psi_i\rangle\langle\Psi_i|\hat{P}_s|\Phi_{UHF}\rangle}{\langle\Phi_{UHF}|\hat{P}_s|\Phi_{UHF}\rangle}, \tag{5.23}$$

where $|\Psi_i\rangle$ are excited determinants. The PUHF energy can be evaluated as [104–107],

$$E_{\text{PUHF}} = \langle\Phi_{\text{UHF}}|\hat{P}_s\hat{H}\hat{P}_s|\Phi_{\text{UHF}}\rangle \qquad (5.24\text{a})$$

$$= \langle\Phi_{\text{UHF}}|\hat{H}\hat{P}_s|\Phi_{\text{UHF}}\rangle \qquad (5.24\text{b})$$

$$= E_{\text{UHF}} + \Delta E_{\text{PUHF}}, \qquad (5.24\text{c})$$

where ΔE_{PUHF} is the correction to the UHF energy,

$$\Delta E_{\text{PUHF}} = \langle\Phi_{\text{UHF}}|\hat{H}|\tilde{\Phi}_{\text{UHF}}\rangle. \qquad (5.25)$$

Note that, in deriving above equations, we have used $\hat{P}_s^2 = \hat{P}_s = \hat{P}_s^\dagger$ and $[\hat{P}_s, \hat{H}] = 0$.

It is common to approximate $\hat{P}_s$ by just one single annihilator $A_{s,l}$ to eliminate only the next higher spin-contaminant of the UHF wave function. Hence, unless otherwise noted, we apply only $A_{0,1}$ to the broken-symmetry UHF (and CUHF) wave functions, which usually removes the most dominant contamination for $M = 0$. Therefore, $E(\text{S})$ may now be approximated with the broken-symmetry UHF wave function by

$$E(\text{S}) \approx E(\text{PBS}) = \langle\Phi_{\text{UHF}}^{\text{BS}}|\hat{H}\hat{A}_{0,1}|\Phi_{\text{UHF}}^{\text{BS}}\rangle, \qquad (5.26\text{a})$$

where PBS stands for the projected broken-symmetry solution. Note that in this expression we make the approximation that $\hat{A}_{0,1}$ commutes with $\hat{H}$ and is idempotent.

Furthermore, the UHF high-spin triplet state also contains spin-contamination from $S = 2, 3, \cdots$, although it is usually small compared to the broken-symmetry $M = 0$ state. Therefore, we also apply $\hat{A}_{1,2}$ to the triplet UHF wave function to approximate $E(\text{T})$,

$$E(\text{T}) \approx E(\text{PT}) = \langle\Phi_{\text{UHF}}^{\text{T}}|\hat{H}\hat{A}_{1,2}|\Phi_{\text{UHF}}^{\text{T}}\rangle. \qquad (5.27)$$

On the other hand, with CUHF(2), $\hat{A}_{0,1}$ completely removes all the spin-contamination in the broken symmetry solution of $M = 0$ because the only spin-contaminant is the

third term ($S = 1$) in Eq.(5.19). In other words, in CUHF(2), $\hat{A}_{0,1}$ is equivalent to the "full" projection operator. Hence, it commutes with $\hat{H}$, and is idempotent. Accordingly,

$$E(\mathrm{S}) = \langle \Phi^{\mathrm{BS}}_{\mathrm{CUHF}(2)} | \hat{H} \hat{A}_{0,1} | \Phi^{\mathrm{BS}}_{\mathrm{CUHF}(2)} \rangle. \tag{5.28}$$

Note that the projection is not needed for the triplet CUHF(2) wave function since it is already spin-adapted (ROHF).

Finally the singlet-triplet splitting with the projection scheme is given by

$$\Delta E^{\mathrm{P}}_{\mathrm{ST}} = \begin{cases} E(\mathrm{PBS}) - E(\mathrm{PT}) & \mathrm{PUHF} \\ E(\mathrm{S}) - E(\mathrm{T}) & \mathrm{PCUHF}(2) \end{cases}. \tag{5.29}$$

As we will show, the PUHF approximation in Eq.(5.29) is not adequate in many cases, while it proves very useful with PCUHF, especially when weak dynamical correlation via MP2 is properly introduced.

This approach can be easily generalized for other multiplet splittings within a larger CUHF active space. For example, if there are four unpaired electrons, it is desirable to use CUHF(4). Computing the singlet state with CUHF(4) would require annihilation of two spin-components, $l = 1, 2$.

Furthermore, one should note that PUHF and PCUHF presented here are in the so-called projection-after-variation scheme, in which one first obtains converged UHF and CUHF wave functions and then projects them. This is different from variation-after-projection which variationally minimizes a projected wave function [69].

5.2.4 Constrained MP2 based on CUHF

An accurate description of electronic states requires the inclusion of weak electron correlation. Here we consider a second-order perturbation theory approximation (MP2) [83] to the CUHF energy, which we call constrained unrestricted MP2 (CUMP2). We

start from our CUHF Fock operators as the zeroth order Hamiltonian $\hat{H}_0$,

$$\hat{H}_0 = \sum_i \hat{\tilde{F}}(i). \tag{5.30}$$

The perturbation $\hat{V}$ that we apply is

$$\hat{V} = \hat{H} - \hat{H}_0, \tag{5.31}$$

where $\hat{H}$ is the physical Hamiltonian. The CUHF wave function $|\Phi_0\rangle$ satisfies

$$\hat{H}_0|\Phi_0\rangle = \sum_i \varepsilon_i|\Phi_0\rangle. \tag{5.32}$$

The first-order wave function $|\Phi_1\rangle$ is in general

$$|\Phi_1\rangle = -\sum_{ia} \frac{\langle\Phi_i^a|\hat{V}|\Phi_0\rangle}{\varepsilon_a - \varepsilon_i}|\Phi_i^a\rangle - \frac{1}{4}\sum_{ijab} \frac{\langle\Phi_{ij}^{ab}|\hat{V}|\Phi_0\rangle}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}|\Phi_{ij}^{ab}\rangle, \tag{5.33}$$

where i, j and a, b run over occupied and virtual orbitals, respectively. Given

$$\langle\Phi_i^a|\hat{V}|\Phi_0\rangle = \langle\Phi_i^a|\hat{H}|\Phi_0\rangle = F_{ia}, \tag{5.34a}$$

$$\langle\Phi_{ij}^{ab}|\hat{V}|\Phi_0\rangle = \langle\Phi_{ij}^{ab}|\hat{H}|\Phi_0\rangle = \langle ij||ab\rangle, \tag{5.34b}$$

the CUMP2 correlation energy E_2 is

$$\begin{aligned} E_2 &= \langle\Phi_0|\hat{V}|\Phi_1\rangle \\ &= -\sum_{ia} \frac{|F_{ia}|^2}{\varepsilon_a - \varepsilon_i} - \frac{1}{4}\sum_{ijab} \frac{|\langle ij||ab\rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - i - \varepsilon_j}. \end{aligned} \tag{5.35}$$

Not surprisingly, if the reference is CUHF(N_s), Eq.(5.35) corresponds to the RMP2 approximation of Knowles et al. [86]. One desirable feature of RMP2 is that even though the first order correction to the wave function is spin-contaminated, the second-order energy correction is not [86]. Eq.(5.35) yields standard UMP2 for the CUHF(N_e) reference. Thus, CUMP2 may also be regarded as an interpolation between RMP2 and UMP2.

Generalizations of MP2 to PUHF have been discussed independently by Schlegel [104, 105], and Knowles and Handy [106, 107]. In this work, we use the energy expression for projected UMP2 (PUMP2) derived by Schlegel,

$$E_{\text{PUMP2}} = E_{\text{UMP2}} + \Delta E_{\text{PUHF}} \left(1 - \frac{\langle \Phi_1 | \tilde{\Phi}_0 \rangle}{\langle \tilde{\Phi}_0 | \tilde{\Phi}_0 \rangle} \right), \qquad (5.36)$$

which is already implemented in the `Gaussian` program as PMP2 [70]. In the above equation, $|\tilde{\Phi}_0\rangle$ is defined in Eq.(5.23). A similar equation is obtained for projected CUMP2 (PCUMP2).

5.2.5 Computational details

We have implemented the methods described above in the `Gaussian` suite of programs [70]. Calculations with CUHF(2) and CUMP2(2) confirm that, for high-spin triplet states, our codes yield the ROHF and RMP2 solutions, respectively. It is convenient to start the CUHF iterative procedure from UHF orbitals after sorting the NOs in decreasing order of occupations n_i, and taking the first N_c to be core, the next N_a to be active, and the rest to be virtual. This selection is always possible because of the UHF corresponding-pair property, and has the additional advantage of avoiding getting trapped in local minima. Recall that the core, active, and virtual blocks are well-defined only in the NO basis. The $\langle S^2 \rangle$ values in Eq.(5.15) needed for UMP2 and CUMP2 in broken-symmetry situations are taken from UHF and CUHF, respectively. For projected schemes, Eq.(5.29) is used. In the next section, we will first discuss size-consistency in PUHF and PCUHF, and then present results of benchmark calculations including (i) the small molecules NH, OH$^+$, O$_2$, and NF, (ii) CH$_2$, (iii) trimethylenemethane (TMM), and (iv) o-, m-, and p-benzynes.

5.3 Results and discussion

5.3.1 The size-consistency problem in PUHF

Size-consistency is an important and desirable feature. A method is deemed size-consistent when the sum of fragment energies E_A and E_B is identical to the E_{AB} fragments as the fragments become non-interacting,

$$E_{AB} \xrightarrow[R_{AB}\to\infty]{} E_A + E_B. \tag{5.37}$$

In bond-breaking situations, UHF is usually size-consistent, while RHF typically is not.

Davidson pointed out that PUHF is *not* size-consistent even if the full projection operator $\hat{P}_s$ in Eq.(5.20) is employed [100]. In this section, we examine the size-consistency of PCUHF with two examples: the CN and O_2 dimers at infinite separation. We carry out CUHF(N_a) calculations for monomer A with $N_a = N_A, N_A - 2, \cdots, N_s + 2, N_s$, where N_A and N_s are the total and unpaired number of electrons in A, respectively. We then obtain PCUHF(N_a) after *full* spin-projection so that the PCUHF wave function has the correct S value of $N_s/2$. For the separated dimer AA, we use an active space of twice the size of the monomer, $2N_a$, and test the high-spin $M = 2N_s$ and $M = 0$ cases. Note that, since in these calculations we remove all the spin-contaminants with $S > M$, the resulting PCUHF($2N_a$) wave functions for dimers are eigenfunctions of $\hat{S}^2$ with $S = M$. Also note that CUHF(N_A) and CUHF($2N_A$) correspond to UHF for the monomer and dimer, respectively, and therefore yield PUHF after full spin projection. In Table 5.1, we report E_A as well as the energy differences $\Delta E = E_{AA} - 2E_A$. For a method to be size-consistent, ΔE must be zero.

Calculations in Table 5.1 are carried out with the cc-pVTZ basis set at geometries

of $R_{CN} = 1.16945\text{Å}$ and $R_{O_2} = 1.20639\text{Å}$ for the monomers, and $R_{AA} = 10^5\text{Å}$ for

Table 5.1 : Size-consistency in PUHF and PCUHF. The PCUHF(N_a) energies of monomer A (E_A) and energy difference ΔE are presented in Hartree and kcal mol^{-1}, respectively.

	CN ($N_A = 13$)						
PCUHF(N_a)	$(13)^a$	(11)	(9)	(7)	(5)	(3)	$(1)^b$
$E_A(S = 1/2)$	-92.2359	-92.2359	-92.2359	-92.2350	-92.2305	-92.2077	-92.1964
$\Delta E(S = 1)^c$	8.75	8.74	8.74	8.53	7.38	2.18	0.00
$\Delta E(S = 0)^c$	2.08	2.08	2.08	1.99	1.53	0.14	0.00

	O$_2$ ($N_A = 16$)							
PCUHF(N_a)	$(16)^a$	(14)	(12)	(10)	(8)	(6)	(4)	$(2)^b$
$E_A(S = 1)$	-149.6376	-149.6376	-149.6376	-149.6374	-149.6340	-149.6294	-149.6186	-149.6083
$\Delta E(S = 2)^c$	4.09	4.09	4.09	4.05	3.53	2.86	1.38	0.00
$\Delta E(S = 0)^c$	0.06	0.06	0.06	0.06	0.05	0.03	0.01	0.00

[a]Corresponds to PUHF.
[b]Corresponds to ROHF for the monomer.
[c]The active space is $2N_a$ for the dimer and N_a for the monomer.

the separated dimers. The full projected results can be obtained from the `Gaussian` output by requesting an MP4 calculation. We list ΔE for both cases.

In the CN monomer, the seven α and six β electrons yield $N_s = 1$. Therefore, the $S = 1/2$ CUHF(N_a) wave function is obtained after full projection, and its energy E_A (Hartree) is presented in Table 5.1. In the CN dimer, the $M = 1$ and $M = 0$ CUHF wave functions generally have spin-contamination from $S = 2, 3, \dots$ and $S = 1, 2, \dots$ components, respectively. In PCUHF(N_a), these contaminants are all removed and we obtain the $S = 1$ and $S = 0$ spin-adapted states. In both cases, ΔE is generally nonzero, $i.e.$, PCUHF(N_a) and PUHF are not size-consistent. It is worth noting, however, that PCUHF is size-consistent for $N_a = N_s$. The reason for this is the following: for the monomer minimum active space (N_s), there is only one $S = N_s/2$ configuration state function (CSF) in PCUHF, which is exactly the high-spin ROHF determinant. The $S = 1$ PCUHF($2N_s$) wave function is also ROHF, and thus $\Delta E(S = 1)$ is zero. On the other hand, the $S = 0$ PCUHF($2N_s$) wave function contains two $S = 0$ CSFs, $|\sigma_g^2\rangle$ and $|\sigma_u^2\rangle$, with the same $1/\sqrt{2}$ weight, which give the lowest energy in this CI space. Therefore, the energy of PCUHF($2N_s$) is twice E_A also for the $S = 0$ case, $i.e.$, $\Delta E = 0$.

The argument above fails when $N_a \neq N_s$. For the monomer, the PCUHF(N_a) wave function contains more than one $S = 1/2$ CSF with CI coefficients that are not variationally optimized. The PCUHF($2N_a$) wave function for the dimer also contains more than one $S = 0$ CSF with unoptimized CI coefficients. Since neither set of CI coefficients yields the lowest energy state in each CI space, the result is in general not size-consistent. This size-inconsistency problem is especially pronounced in PUHF and is unavoidable unless one uses $N_a = N_s$.

Similar results are obtained in the O_2 case, whose dimer has four unpaired elec-

trons. While PCUHF with $N_a = N_s = 2$ yields size-consistent results, other choices of N_a result in small errors. As expected, the error increases with N_a. The maximum error is obtained when PCUHF corresponds to PUHF.

We believe that these size-consistency results are important because, even if one fully projects UHF, there are intrinsic errors due to size-inconsistency, and the quality of results obtained with full-PUHF is inherently questionable. However, these errors can be avoided by using a minimal PCUHF(N_s) space. This is the motivation for our focus on CUHF(2) for singlet-triplet splittings.

5.3.2   NH, OH$^+$, O$_2$, and NF

We here discuss UHF and CUHF applications to singlet-triplet ($^1\Delta - {}^3\Sigma^-$) energy splittings in NH, OH$^+$, O$_2$, and NF. Geometries were optimized with UMP2/TZ2P (taken from Ref. [108]) and single point calculations used the cc-pVTZ basis set.

Due to the exact degeneracy between π orbitals, all of these systems exhibit pure unpaired orbitals in the broken-symmetry solutions. Therefore, the ideal $\langle S^2 \rangle$ values for $M = 1$ and $M = 0$ are 2 and 1, respectively. This is exactly what CUHF(2) and CUMP2(2) achieve. On the other hand, in UHF and PUHF, $\langle S^2 \rangle$ deviates from these values as shown in Table 5.2 and hence spin-deviation σ_s is non-zero. Nevertheless, there is balance in UHF where σ_s in the $M = 1$ and $M = 0$ cases are moderate and comparable in size. Therefore, UHF and UMP2 can be expected to give reasonably accurate results for the energy splittings as they indeed do. Table 5.3 summarizes the $^1\Delta - {}^3\Sigma^-$ energy splittings calculated by different methods, as well as ME and MAE. For comparison purposes, MP2 data obtained from open-shell singlet ROHF (OSS-MP2) [108] are also included in the table. The MAEs of UHF and UMP2 are surprisingly small, and are of quality similar to the CUHF and CUMP2 results.

Table 5.2 : $\langle S^2 \rangle$ in UHF and PUHF.

System	$\langle S^2 \rangle_{\mathrm{UHF}}$		$\langle S^2 \rangle_{\mathrm{PUHF}}$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
NH	2.0151	1.0106	2.0001	0.0839
OH^+	2.0137	1.0079	2.0000	0.0625
O_2	2.0420	1.0251	2.0009	0.2012
NF	2.0215	1.0136	2.0001	0.1079

Table 5.3 : Singlet-Triplet splitting energies (kcal mol^{-1}) for small diradicals predicted with different approximations. The active space in CUHF and CUMP2 is chosen to be (2) for all systems. The cc-pVTZ basis is used except for OSS-MP2, which uses TZ2P.

Method	NH	OH^+	O_2	NF	ME	MAE
UHF	38.69	51.41	37.09	39.03	4.93	5.08
UMP2	38.43	51.22	21.47	34.79	-0.15	0.72
PUHF	30.69	42.63	18.02	29.31	-6.47	6.47
PUMP2	29.87	41.85	7.25	26.22	-10.33	10.33
CUHF	41.98	54.73	32.13	41.28	5.90	5.90
CUMP2	37.69	50.38	24.54	34.12	0.05	0.90
PCUHF	41.47	54.14	30.63	40.84	5.14	5.14
PCUUMP2	39.83	52.66	27.12	37.74	2.71	2.71
OSS-MP2[a]	38.52	52.14	24.27	34.23	0.66	0.95
Exptl.[b]	39	50.55	22.64	34.32		

[a]OPT2 values taken from Ref. [108].
[b]Taken from Ref. [109].

In PUHF, however, spin-contamination becomes "unbalanced," *i.e.*, the $M = 1$ wave function has little spin-contamination and $\langle S^2 \rangle \approx 2$, while the $M = 0$ state is heavily contaminated. This unbalance may explain the large errors in the MAEs of projected methods based on UHF, especially PUMP2, which underestimates E(PBS) for $M = 0$. On the other hand, the MAEs of PCUHF and PCUMP2 remain small since both are spin-adapted wave functions. Although PCUMP2 has a slightly larger error than CUMP2, it is certainly more accurate than PUMP2.

5.3.3 CH$_2$

The next benchmark example is methylene, CH$_2$. The ground state is ^{3}B$_1$ and the low lying singlet excited states are ^{1}A$_1$ and ^{1}B$_1$. We have used the T2ZP full CI geometries [110], and a cc-pVTZ basis for our single-point energy calculations. The zero point energy (ZPE) is estimated from configuration interaction singles and doubles (CISD) frequency calculations [111] and the difference in ZPE between states, ΔZPE, is added when making direct comparison between the computed ΔE_{ST} and the experimental values.

In Table 5.4, we present ΔE_{ST} for different approximations as well as $\langle S^2 \rangle$ for UHF and CUHF. Again, spin-deviation is moderate in UHF, and therefore the UHF and CUHF results are similar—although CUHF is slightly better. When adding dynamical correlation, UMP2 and CUMP2 both yield very accurate ΔE_{ST} for ^{1}A$_1$ with Noodleman's approximation, but they underestimate the FCI value for ^{1}B$_1$. On the other hand, projected schemes address this issue by explicitly treating the multireference character of ^{1}B$_1$ correctly. Both PCUHF and PCUMP2 are much more accurate than PUHF and PUMP2.

Table 5.4 : Total energies (Hartree) and ΔE_{ST} (kcal mol^{-1}) of CH_2. UHF and CUHF $\langle S^2 \rangle$ values are also listed.

Method	$^3B_1$	$^1A_1{}^a$	$^1B_1{}^b$
UHF	-38.9377	13.0	17.7
($\langle S^2 \rangle$)	(2.0162)	(0.7180)	(0.9631)
UMP2	-39.0555	9.8	14.7
PUHF	-38.9406	17.5	28.5
PUMP2	-39.0574	12.7	24.9
CUHF	-38.9322	11.6	17.8
($\langle S^2 \rangle$)	(2.0000)	(0.6727)	(0.9423)
CUMP2	-39.0554	9.3	14.2
PCUHF	-38.9322	15.7	32.7
PCUMP2	-39.0554	12.7	29.3
FCI(T2ZP)^c	-39.0667	11.1	35.6
Exp.^{d}		9.0	32.9
Exp. $- \Delta$ZPE^e		9.4	33.0

^{a}All fully unrestricted methods, CUHF and CUMP2 yield
 a symmetry-broken electronic state, a mixture of mostly
 $^1A_1$ and $^3B_1$ states.
^{b}All fully unrestricted methods, CUHF and CUMP2 yield
 a mixture of mostly $^1B_1$ and $^3B_1$ states.
^{c}Ref. [110].
^{d}Ref. [112].
^{e}CISD estimate, Ref. [111].

5.3.4 TMM

Trimethylenemethane is an example of a non-Kekulé molecule that cannot be represented with resonance structures (Figure 5.3). In TMM, the ground state is $^3A_2'$ with two unpaired electrons occupying a_2 and $2b_1$ π orbitals. These π orbitals are energetically degenerate and the two electrons thus also couple into a $^1A_1$ singlet

state. The experimental ΔE_{ST} between these two states is 16.1 kcal mol^{-1} [113]. Table 5.5 summarizes the total energies for $^3\text{A}_2'$, ΔE_{ST}, and $\langle S^2 \rangle$ values for UHF and CUHF. In this system, the UHF spin-deviation turns out to be large, more than 0.1 for both states. Hence, the energy difference between UHF and CUHF is also rather large (about 35 mHartree in $^3\text{A}_2'$). Consequently, UHF predicts a very poor ΔE_{ST}, 25 kcal mol^{-1} larger than the target value of 17.7 kcal mol^{-1} after adding the ZPE correction [114] to the experimental result. PUHF unexpectedly gives a reasonable ΔE_{ST} of 11.2 kcal mol^{-1}, but we believe that this originates from adventitious cancellation of errors, as indicated by the remaining large spin-contamination in the $^1\text{A}_1$ state. Also note how poor the PUMP2 result is at 0.5 kcal mol^{-1}.
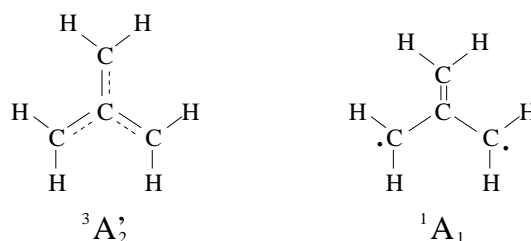
Figure 5.3 : The TMM molecule.

The PCUHF and PCUMP2 wave functions are pure spin states for TMM, and hence both give reasonably accurate results of 10.5 and 20.8 kcal mol^{-1}, respectively. It should be pointed out that the PCUHF energy of $^1\text{A}_1$ is very similar to that of CASSCF(2,2) (for $^3\text{A}_2'$ they are equivalent: ROHF). This indicates that the PCUHF orbitals and CI coefficients mimic those of CASSCF(2,2). The latter optimizes both at an additional computational expense, resulting in an energy that is slightly lower than that of PCUHF.

We have also tested the performance of our constrained unrestricted scheme on

Table 5.5 : Total energies (Hartree) and ΔE_{ST} (kcal mol^{-1}) for TMM calculated by several methods with cc-pVTZ. Geometries optimized with CASSCF(10,10)/cc-pVDZ were taken from Ref. [114].

| | $^3\mathrm{A}_2'$ | | $^1\mathrm{A}_1$ | |
Method	E	$\langle S^2 \rangle$	$\langle S^2 \rangle$	ΔE_{ST}
UHF	-154.9529	2.2166	1.1139	42.9
UMP2	-155.5420			26.6
PUHF	-154.9698	2.0098	0.8681	11.2
PUMP2	-155.5558			0.5
CUHF	-154.9172	2.0000	0.9998	14.9
CUMP2	-155.5572			35.5
PCUHF	-154.9172	2.0000	0.0000	10.5
PCUMP2	-155.5572			20.8
CASSCF(2,2)	-154.9172	2.0000	0.0000	9.8
UB3LYP[a]	-155.9888	2.0327	1.0055	22.0
UHSE[b]	-155.7986	2.0444	1.0071	24.6
ULC-ωPBE[c]	-155.8703	2.0752	1.0080	29.6
CUB3LYP[a]	-155.9823	2.0000	1.0000	17.5
CUHSE[b]	-155.7900	2.0000	1.0000	18.8
CULC-ωPBE[c]	-155.8568	2.0000	0.9999	19.0
Exp.[d]				16.1
Exp. $-\Delta$ZPE[e]				17.7

[a]Geometries optimized with UB3LYP/cc-pVTZ.
[b]Geometries optimized with UHSE/cc-pVTZ.
[c]Geometries optimized with ULC-ωPBE/cc-pVTZ.
[d]Ref. [113].
[e]CASSCF(4,4)/cc-pVDZ value taken from Ref. [114].

generalized unrestricted Kohn-Sham (CUGKS) hybrid and range-separated functionals, namely, B3LYP [35], HSE [40, 41, 115, 116] and LCωPBE [42–44]. Recall that "generalized" in GKS is meant to indicate nonlocal KS potentials from derivatives of the energy with respect to the one-particle density matrix as opposed to the electron density. Geometries were optimized with each functional and the cc-pVTZ basis set. Since KS/GKS determinants are not strictly wave functions, $\langle S^2 \rangle$ is *not* given by Eq.(4.4) [117]. Nevertheless, it is known that Eq.(4.4) is a reasonable approximation to KS $\langle S^2 \rangle$ values [118, 119], and we have therefore used them here. The functionals listed above contain a large amount of HF exchange and are more prone to triplet instabilities than is pure KS without HF exchange. Therefore, as shown in Table 5.5, the $\langle S^2 \rangle$ values for all functionals are unbalanced: the $^3A_2'$ state is very spin-contaminated (spin-deviation σ_s of $\sim$0.03-0.07), especially for ULCωPBE, while the $^1A_1$ state has near zero σ_s. This unbalance yields inaccurate ΔE_{ST} predictions of 22.0, 24.6, and 29.6 kcal mol^{-1} for UB3LYP, UHSE, and ULCωPBE, respectively. It is noteworthy, however, that by imposing active space constraints, all CUGKS functionals dramatically improve results over GKS, strongly suggesting that our constrained approach may be a very useful alternative to regular GKS in heavily contaminated cases.

5.3.5   *o*-, *m*-, and *p*-benzynes

Last, we discuss benchmarks on a series of benzynes. All calculations used UB3LYP geometries optimized with a 6-31G(d) basis followed by single-point energy calculations with the cc-pVTZ basis set. For these compounds, we have used the same functionals mentioned in the previous section. The ground state is singlet $^1A_1$ for *o*- and *m*-benzynes, and ^1A_g for *p*-benzyne. Therefore, ΔE_{ST} is negative.

For these molecules, the UHF $\langle S^2 \rangle$ is significantly different from that of CUHF,

i.e., σ_s is very large (Table 5.6). Consequently, UHF and UMP2 have substantial errors in ΔE_{ST} when compared to experiment. As reported by Davidson and Clark [100], since the $M = 0$ UHF wave function has not only triplet but also quintet and higher spin contaminants with significant weights, PUHF increases δ_s and σ_s for *p*-benzyne, as well as for *o*-benzyne, making the computed ΔE_{ST} completely unreliable (Table 5.7). Similar poor quality results are obtained for PUMP2. This is perhaps one of the reasons why projection schemes are not very popular among quantum chemists.

Table 5.6 : $\langle S^2 \rangle$ for *o*-, *m*-, and *p*-benzyne molecules calculated with cc-pVTZ.

	o		*m*		*p*	
	$^3\mathrm{B}_2$	$^1\mathrm{A}_1$	$^3\mathrm{B}_2$	$^1\mathrm{A}_1$	$^3\mathrm{B}_{1u}$	$^1\mathrm{A}_g$
UHF	2.4154	1.3482	2.7574	0.0000	2.4076	1.7583
PUHF	2.0874	3.5537	2.3530	0.0000	2.0869	5.2197
CUHF	2.0000	0.6201	2.0000	0.0000	2.0000	0.9925
PCUHF	2.0000	0.0000	2.0000	0.0000	2.0000	0.0000
UB3LYP	2.0077	0.0000	2.0205	0.0000	2.0067	0.9442
UHSE	2.0099	0.0000	2.0328	0.0000	2.0088	0.9745
ULCωPBE	2.0102	0.1743	2.0606	0.0000	2.0090	1.0324
CUB3LYP	2.0000	0.0000	2.0000	0.0000	2.0000	0.9177
CUHSE	2.0000	0.0000	2.0000	0.0000	2.0000	0.9373
CULCωPBE	2.0000	0.0534	2.0000	0.0000	2.0000	0.9732

Interestingly, all methods (including UHF) predict *m*-benzyne to be closed-shell in the singlet state. This is perhaps correct, as we observe a rather large CI coefficient (0.989) for the closed-shell HF configuration in the CASSCF(4,4) wave function. It thus appears that in *m*-benzyne, the addition of dynamical correlation is important.

Table 5.7 : Total energies (Hartree) and ΔE_{ST} (kcal mol^{-1}) for o-, m-, and p-benzyne molecules calculated with cc-pVTZ. Geometries optimized with UB3LYP/6-31G(d).

| | o | | m | | p | | |
	$^3B_2$	ΔE_{ST}	$^3B_2$	ΔE_{ST}	$^3B_{1u}$	ΔE_{ST}	MAE
UHF	-229.4711	-29.6	-229.4917	20.5	-229.4780	-31.1	25.7
UMP2	-230.3121	-24.4	-230.3048	-38.2	-230.3213	25.6	20.1
PUHF	-229.4879	-127.5	-229.5258	49.7	-229.4951	-292.9	149.7
PUMP2	-230.3274	-123.5	-230.3362	-30.0	-230.3368	-274.2	121.9
CUHF	-229.4627	-14.2	-229.4669	12.8	-229.4691	-0.2	20.3
CUMP2	-230.3413	-34.4	-230.3469	-26.3	-230.3499	-2.1	3.6
PCUHF	-229.4627	-24.5	-229.4669	12.8	-229.4691	-0.5	16.6
PCUMP2	-230.3413	-37.8	-230.3469	-26.3	-230.3499	-1.6	2.6
UB3LYP	-230.9432	-31.7	-230.9432	-13.0	-230.9522	-4.9	5.1
UHSE	-230.6794	-27.8	-230.6794	-15.6	-230.6885	-4.4	5.4
ULCωPBE	-230.7539	-27.8	-230.7539	-15.4	-230.7636	-3.4	5.2
CUB3LYP	-230.9409	-33.3	-230.9409	-16.3	-230.9498	-4.3	3.3
CUHSE	-230.6765	-29.7	-230.6765	-18.3	-230.6853	-3.5	3.5
CULCωPBE	-230.7510	-27.7	-230.7510	-18.7	-230.7604	-1.6	4.7
Exp.^{a}		-37.5		-21.0		-3.8	
Exp. $- \Delta$ZPE^b		-38.0		-20.6		-3.5	

^{a}Ref. [120].
^{b}ZPEs computed with UB3LYP/cc-pVTZ.

UHF, PUHF, CUHF, and PCUHF all fail to predict $^1A_1$ to be the lowest state. Although CUHF and PCUHF eliminate much spin-contamination, their results deviate by 20.3 and 16.6 kcal mol^{-1}. Adding MP2 correlation to both (CUMP2 and PCUMP2 schemes) substantially improves the MAE to 3.6 and 2.6 kcal mol^{-1}, respectively.

For these benzynes, DFT performs reasonably well, yielding MAEs of ~ 5 kcal mol^{-1}. Because the σ_s values are small, the CUGKS results for ΔE_{ST} are similar to the regular GKS results. In most cases, however, the constrained approach yields a small improvement.

Chapter 6

Conclusions

Strong correlation is ubiquitous in electronic structure theory. In many situations, degeneracies play an important role in determining the properties of molecules and materials, necessitating a multi-determinant wave function to properly describe such effects. Despite its importance, there have been rather few theoretical approaches to capture strong correlations appropriately, while there has been significant progress and development of methods that deal with weak correlations in the last few decades. Most multi-determinant methods are simply computationally intractable due to the exponential scaling of their cost. The aim of this thesis has been to develop a novel theoretical scheme that achieves the correct description of strong correlations within a mean-field cost. To accomplish this goal, we have considered a mixture of HF and HFB with an active space, named Constrained Pairing Mean-Field Theory (CPMFT). We have shown that CPMFT is capable of cleanly separating strong and weak correlations, yet only accounts for the former [2]. We have generalized CPMFT to describe any molecular dissociations through an asymptotic one-body potential [3], as well as to include weak correlations via the wisdom of density functional theory (DFT) [4]. We have also understood the connection between CPMFT and broken-symmetry unrestricted Hartree-Fock (UHF), and successfully merged the corresponding pairs property of UHF into CPMFT, in order to reduce computational costs and surmount the overcorrection problem in some cases with promising results [5]. This study also elucidated a direct connection between CPMFT and UHF. A fundamental difference

is that in CPMFT, the symmetry restoration step is done on the two-particle density matrix instead of the wave function. This leads to some form of an ensemble model that is not N-representable, yet contains *no* particle-number fluctuations.

Furthermore, the concept behind CPMFT was applied to UHF in a reverse manner. As separate work [77], we have developed a constrained unrestricted Hartree-Fock method (CUHF) for obtaining restricted open-shell HF (ROHF) directly from UHF by projecting the wave function self-consistently. The results obtained from this CUHF method proved promising in terms of the convergence behavior and the quality of semi-canonical orbitals. We have further studied this model for allowing possible spin-contamination δ_s in an active space while removing it from the rest [121]. This scheme was applied to both HF and Kohn-Sham DFT. Our constrained active space unrestricted method, CUHF(N_a), was regarded as an interpolation between the fully restricted (ROHF) and fully unrestricted approaches (UHF). Second-order Møller-Plesset perturbation theory (MP2) on CUHF (CUMP2) was also derived, another interpolation between restricted MP2 and unrestricted MP2.

With CUHF(N_a), we have presented benchmark calculations on singlet-triplet energy splittings of various systems using Noodleman's approximation. We have also revisited and utilized Löwdin's projection operator method for restoring spin quantum numbers. The latter approach acquires strong correlations by breaking symmetry in CUHF then restoring it via spin-projection. Based on this projection scheme, we have proposed the projected CUHF (PCUHF) and projected CUMP2 (PCUMP2) methods. Both schemes do have wave functions associated with them, unlike CPMFT. We have demonstrated that constraining spin-contamination to an active space has beneficial effects, yielding approximations that are much more accurate than regular UHF and projected UHF, particularly when the latter methods suffer from large spin-

contamination. In cases where DFT exchange functionals contain large amounts of HF exchange and a tendency to break spin-symmetry, our constrained scheme also exerts a positive effect. The calculated singlet-triplet splittings for trimethylenemethane and three benzynes show significant improvement when B3LYP, HSE, and LC-ωPBE are allowed to become unrestricted spin density functionals only in an active space.

The main drawback of the current formalism of CPMFT and CUHF is that the active space has to be chosen by chemical intuition. In CPMFT, it is usually preferred to take the number of valence orbitals or the number of open-shell electrons at dissociation of the system. However, as we have seen in the example of C_2, this choice might neglect the so-called "intruder states" that are not important for the dissociation limit but play a fundamental role in the vicinity of equilibrium. Another open question in CPMFT is how it handles open-shell systems. The open-shell treatment in CPMFT requires a procedure that involves "blocking" open-shell orbitals out of the Bogoliubov transformation. This is a formidable task even in HFB, and its generalization to CPMFT has not been clarified. Finally, describing angular correlation that comes from near-degeneracies in atoms such as $2s$-$2p$ in Be may require improvement on the CPMFT functional. Overall, however, CPMFT offers an accurate description for strong correlations if the active space is properly chosen, and is an interesting model for further development, which we are the first to explore.

The ideas presented in this work can be applied in myriad ways to both HF and DFT, as well as to many properties other than those benchmarked here: heats of formation and barrier heights. The computational implementation of our constrained methods is rather straightforward and simple. Indeed, our work stimulated and revived further work on symmetry breaking and restoration [69, 122], ideas that were actively pursued many years ago but seem to have been abandoned. As shown both

in our CPMFT and CUHF work, strong correlations can be accurately modeled with broken-symmetry mean-field approaches if adequate strategies for symmetry restoration are developed.

Appendix A

Properties of the CPMFT model two-Particle density matrix

The CPMFT model two-particle density matrix is

$$(\mathbf{\Gamma}_{\text{CPMFT}})_{ij}^{kl} = \frac{1}{2}\gamma_i^k\gamma_j^l - \frac{1}{2}\gamma_i^l\gamma_j^k - \frac{1}{2}\kappa_{ij}\kappa^{kl}. \tag{A.1}$$

where i, j, k, and l are spin-orbitals and $\boldsymbol{\gamma}$ and $\boldsymbol{\kappa}$ are the density matrix and anomalous density matrix in the spin-orbital basis (*i.e*, they are of dimension $2N \times 2N$, where N is the size of the atomic orbital basis). In general, $\boldsymbol{\gamma}$ is Hermitian and $\boldsymbol{\kappa}$ is antisymmetric. When everything is real (which we take for simplicity; this does not affect our conclusions), the idempotent HFB quasiparticle density matrix is

$$\mathbf{R} = \begin{pmatrix} \boldsymbol{\gamma} & \boldsymbol{\kappa} \\ -\boldsymbol{\kappa} & 1 - \boldsymbol{\gamma} \end{pmatrix}. \tag{A.2}$$

Idempotency tells us that

$$\boldsymbol{\gamma}\,\boldsymbol{\kappa} - \boldsymbol{\kappa}\,\boldsymbol{\gamma} = \mathbf{0}, \tag{A.3a}$$

$$\boldsymbol{\gamma}^2 - \boldsymbol{\kappa}^2 = \boldsymbol{\gamma}. \tag{A.3b}$$

We recall that for closed shells [12],

$$\boldsymbol{\gamma} = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix}, \tag{A.4a}$$

$$\boldsymbol{\kappa} = \begin{pmatrix} \mathbf{0} & \mathbf{K} \\ -\mathbf{K} & \mathbf{0} \end{pmatrix}, \tag{A.4b}$$

$$\mathbf{0} = \mathbf{P}\,\mathbf{K} - \mathbf{K}\,\mathbf{P}, \tag{A.4c}$$

$$\mathbf{P} = \mathbf{P}^2 + \mathbf{K}^2. \tag{A.4d}$$

We can define an analogous model two-particle density matrix for HFB, for which all the conditions on $\boldsymbol{\kappa}$, $\boldsymbol{\gamma}$, $\mathbf{K}$, and $\mathbf{P}$ are the same, but where

$$(\boldsymbol{\Gamma}_{\text{HFB}})_{ij}^{kl} = \frac{1}{2}\gamma_i^k\gamma_j^l - \frac{1}{2}\gamma_i^l\gamma_j^k + \frac{1}{2}\kappa_{ij}\kappa^{kl}. \tag{A.5}$$

Finally, the UHF two-particle density matrix is

$$(\boldsymbol{\Gamma}_{\text{UHF}})_{ij}^{kl} = \frac{1}{2}\gamma_i^k\gamma_j^l - \frac{1}{2}\gamma_i^l\gamma_j^k \tag{A.6}$$

where $\boldsymbol{\gamma}$ is idempotent. We have

$$\boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\gamma}^{\alpha\alpha} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\gamma}^{\beta\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{P}+\mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}-\mathbf{M} \end{pmatrix}, \tag{A.7a}$$

$$\mathbf{P} = \mathbf{P}^2 + \mathbf{M}^2, \tag{A.7b}$$

$$\mathbf{M} = \mathbf{P}\,\mathbf{M} + \mathbf{M}\,\mathbf{P}. \tag{A.7c}$$

A.1   Partial trace of the two-particle density matrix

An important condition on the two-particle density matrix is that it traces to the one-particle density matrix. That is, we must have

$$\Gamma_{ij}^{il} = \frac{N-1}{2}\gamma_j^l. \tag{A.8}$$

Here, repeated indices are to be summed to simplify notation.

The partial trace condition is satisfied by the UHF two-matrix and the CPMFT

model two-matrix, but not by the HFB model two-matrix:

$$\Gamma_{ij}^{il} = \frac{1}{2} \left(\gamma_i^i \gamma_j^l - \gamma_i^l \gamma_j^i \mp \kappa_{ij} \kappa^{il} \right) \tag{A.9a}$$

$$= \frac{1}{2} \left[N\gamma_j^l - (\boldsymbol{\gamma}^2)_j^l \pm (\boldsymbol{\kappa}^2)_j^l \right] \tag{A.9b}$$

$$= \frac{1}{2} \left[N\gamma_j^l - (\boldsymbol{\gamma} + \boldsymbol{\kappa}^2)_j^l \pm (\boldsymbol{\kappa}^2)_j^l \right] \tag{A.9c}$$

$$= \frac{N-1}{2} \gamma_j^l - \frac{1}{2} \left[(\boldsymbol{\kappa}^2)_j^l \mp (\boldsymbol{\kappa}^2)_j^l \right]. \tag{A.9d}$$

Here, the top (bottom) sign in $\pm$ and $\mp$ corresponds to CPMFT (HFB), and we have used antisymmetry of $\boldsymbol{\kappa}$. Explicitly, we have

$$(\boldsymbol{\Gamma}_{\text{CPMFT}})_{ij}^{il} = \frac{N-1}{2} \gamma_j^l. \tag{A.10a}$$

$$(\boldsymbol{\Gamma}_{\text{HFB}})_{ij}^{il} = \frac{N-1}{2} \gamma_j^l - (\boldsymbol{\kappa}^2)_j^l. \tag{A.10b}$$

Note that by N we mean the trace of the one-particle density matrix $\boldsymbol{\gamma}$, which should be the number of particles in the system.

A.2 Particle number fluctuations

In order to work out particle number fluctuations, we need the expectation values of $\hat{N}$ and $\hat{N}^2$, with $\hat{N}$ the number operator, given as

$$\hat{N} = \delta_{pq} a_p^\dagger a_q. \tag{A.11}$$

We have already noted that the expectation value of $\hat{N}$ is just $\text{Tr}(\boldsymbol{\gamma})$. The expectation value of $\hat{N}^2$ requires the two-particle density matrix:

$$\langle \hat{N}^2 \rangle = \delta_{pq} \, \delta_{rs} \, \langle a_p^\dagger a_q a_r^\dagger a_s \rangle \tag{A.12a}$$

$$= \delta_{pq} \, \delta_{rs} \left(-\langle a_p^\dagger a_r^\dagger a_q a_s \rangle + \delta_{qr} \langle a_p^\dagger a_s \rangle \right) \tag{A.12b}$$

$$= \delta_{pq} \, \delta_{rs} \left(2\,\Gamma_{pr}^{qs} + \delta_{qr} \gamma_p^s \right) \tag{A.12c}$$

$$= 2\,\Gamma_{pr}^{pr} + \gamma_p^p. \tag{A.12d}$$

If the two-particle density matrix obeys the partial trace condition, the particle number fluctuations are automatically zero. This is thus true of UHF and of CPMFT. However, HFB has particle number fluctuations:

$$
\begin{aligned}
\langle \hat{N}^2 \rangle_{\text{HFB}} &= (N-1)\gamma_j^j - 2(\boldsymbol{\kappa}^2)_j^j + \gamma_j^j \\
&= N^2 - 2\operatorname{Tr}(\boldsymbol{\kappa})^2
\end{aligned}
\tag{A.13}
$$

implying that

$$
\sigma_N^2 = \langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2 = -2\operatorname{Tr}(\boldsymbol{\kappa}^2).
\tag{A.14}
$$

Note that this is positive, as it should be, since $-\boldsymbol{\kappa}^2 = \boldsymbol{\gamma} - \boldsymbol{\gamma}^2$ and occupation numbers are between 0 and 1, inclusive. In the closed-shell case, we have $\sigma_N^2 = 4\operatorname{Tr}(\mathbf{K}^2)$.

A.3 Spin contamination

Evaluating spin contamination is more complicated than evaluating particle number fluctuations, not least because we need an expression for $\langle \hat{S}^2 \rangle$ for a general two-particle density matrix $\boldsymbol{\Gamma}$. We begin by noting that

$$
\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2
\tag{A.15a}
$$

$$
= \hat{S}_z + \hat{S}_z^2 + \hat{S}_-\hat{S}_+,
\tag{A.15b}
$$

where $\hat{S}_\pm$ is the spin raising/lowering operator. We are interested here in the closed-shell case (*i.e.* $N_\alpha = N_\beta$ with a block-diagonal $\boldsymbol{\gamma}$).

In the closed-shell case, the contribution to $\langle \hat{S}^2 \rangle$ from the first term is zero. We must evaluate the contribution from the next piece using our model two-particle density matrix. We have

$$
\hat{S}_z^2 = \underbrace{\sum_i \hat{s}_z(i)^2}_{\hat{X}_z} + \underbrace{\sum_{i \neq j} \hat{s}_z(i)\hat{s}_z(j)}_{\hat{Y}_z}.
\tag{A.16}
$$

The first (second) term is a one-particle (two-particle) operator. Note that we could also write

$$\hat{Y}_z = 2 \sum_{i>j} \hat{s}_z(i)\hat{s}_z(j) \tag{A.17}$$

which explains the factor of 2 that might otherwise appear to be missing below.

Evaluating the contribution to $\langle \hat{S}_z^2 \rangle$ from $\hat{X}_z$ is straightforward, and we get just

$$\langle \hat{X}_z \rangle = \frac{1}{4}(N_\alpha + N_\beta) = \frac{1}{2}\mathrm{Tr}(\mathbf{P}). \tag{A.18}$$

The nonzero matrix elements of $\hat{Y}_z$ are

$$(Y_z)_{k_\alpha l_\alpha}^{i_\alpha j_\alpha} = \langle i_\alpha j_\alpha | \hat{Y}_z | k_\alpha l_\alpha \rangle = \frac{1}{2}\delta_k^i \delta_l^j, \tag{A.19a}$$

$$(Y_z)_{k_\alpha l_\beta}^{i_\alpha j_\beta} = \langle i_\alpha j_\beta | \hat{Y}_z | k_\alpha l_\beta \rangle = -\frac{1}{2}\delta_k^i \delta_l^j, \tag{A.19b}$$

$$(Y_z)_{k_\beta l_\alpha}^{i_\beta j_\alpha} = \langle i_\beta j_\alpha | \hat{Y}_z | k_\beta l_\alpha \rangle = -\frac{1}{2}\delta_k^i \delta_l^j, \tag{A.19c}$$

$$(Y_z)_{k_\beta l_\beta}^{i_\beta j_\beta} = \langle i_\beta j_\beta | \hat{Y}_z | k_\beta l_\beta \rangle = \frac{1}{2}\delta_k^i \delta_l^j. \tag{A.19d}$$

Here, we are working in an orthornomal basis set.

The relevant components of the CPMFT and HFB two-particle density matrices are

$$\Gamma_{i_\alpha j_\alpha}^{k_\alpha l_\alpha} = \frac{1}{2}\left(\gamma_{i_\alpha}^{k_\alpha}\gamma_{j_\alpha}^{l_\alpha} - \gamma_{i_\alpha}^{l_\alpha}\gamma_{j_\alpha}^{k_\alpha}\right), \tag{A.20a}$$

$$\Gamma_{i_\alpha j_\beta}^{k_\alpha l_\beta} = \frac{1}{2}\left(\gamma_{i_\alpha}^{k_\alpha}\gamma_{j_\beta}^{l_\beta} \mp \kappa_{i_\alpha j_\beta}\kappa^{k_\alpha l_\beta}\right), \tag{A.20b}$$

$$\Gamma_{i_\beta j_\alpha}^{k_\beta l_\alpha} = \frac{1}{2}\left(\gamma_{i_\beta}^{k_\beta}\gamma_{j_\alpha}^{l_\alpha} \mp \kappa_{i_\beta j_\alpha}\kappa^{k_\beta l_\alpha}\right), \tag{A.20c}$$

$$\Gamma_{i_\beta j_\beta}^{k_\beta l_\beta} = \frac{1}{2}\left(\gamma_{i_\beta}^{k_\beta}\gamma_{j_\beta}^{l_\beta} - \gamma_{i_\beta}^{l_\beta}\gamma_{j_\beta}^{k_\beta}\right), \tag{A.20d}$$

where the top (bottom) sign corresponds to CPMFT (HFB).

Contracting the density matrices with the matrix elements, we obtain

$$\langle \hat{Y}_z \rangle = \left(\frac{N_\alpha - N_\beta}{2}\right)^2 - \frac{1}{4}\mathrm{Tr}(\boldsymbol{\gamma}_{\alpha\alpha}^2 + \boldsymbol{\gamma}_{\beta\beta}^2 \mp \boldsymbol{\kappa}_{\alpha\beta}^2 \mp \boldsymbol{\kappa}_{\beta\alpha}^2) \tag{A.21}$$

where we have used antisymmetry of $\boldsymbol{\kappa}$. Working in our closed-shell case, this reduces to

$$\langle \hat{Y}_z \rangle = -\frac{1}{2}\mathrm{Tr}(\mathbf{P}^2 \mp \mathbf{K}^2). \tag{A.22}$$

In total, then, we find that $\langle \hat{S}_z^2 \rangle$ in CPMFT and HFB is given by

$$\langle \hat{S}_z^2 \rangle = \frac{1}{2}\mathrm{Tr}(\mathbf{P} - \mathbf{P}^2 \pm \mathbf{K}^2) \tag{A.23a}$$

$$= \frac{1}{2}\mathrm{Tr}(\mathbf{K}^2 \pm \mathbf{K}^2). \tag{A.23b}$$

Thus, we end up with

$$\langle \hat{S}_z^2 \rangle_{\mathrm{HFB}} = 0, \tag{A.24a}$$

$$\langle \hat{S}_z^2 \rangle_{\mathrm{CPMFT}} = \mathrm{Tr}(\mathbf{K}^2). \tag{A.24b}$$

The contribution to $\langle \hat{S}^2 \rangle$ from $\hat{S}_-\hat{S}_+$ must also be evaluated using the model two-particle density matrix. Expanding this operator in terms of contributions from individual electrons, we have

$$\hat{S}_-\,\hat{S}_+ = \underbrace{\sum_i \hat{s}_-(i)\hat{s}_+(i)}_{\hat{X}} + \underbrace{\sum_{i\neq j} \hat{s}_-(i)\hat{s}_+(j)}_{\hat{Y}}. \tag{A.25}$$

The first term is the one-particle operator $\hat{X}$, and the second is the two-particle operator $\hat{Y}$.

Since $\hat{X}$ does nothing to spin-down electrons but annihilates spin-up electrons, we clearly have

$$\langle \hat{X} \rangle = N_\beta = \mathrm{Tr}(\mathbf{P}). \tag{A.26}$$

To take the expectation value of $\hat{Y}$, it proves useful to symmetrize it so that it acts the same on the two electrons. Since operators acting on different electrons commute,

we have

$$\hat{Y} = \sum_{i \neq j} \hat{s}_-(i)\hat{s}_+(j) \tag{A.27a}$$

$$= \frac{1}{2} \sum_{i \neq j} \left(\hat{s}_-(i)\hat{s}_+(j) + \hat{s}_+(i)\hat{s}_-(j) \right) \tag{A.27b}$$

$$= \sum_{i > j} \left(\hat{s}_-(i)\hat{s}_+(j) + \hat{s}_+(i)\hat{s}_-(j) \right). \tag{A.27c}$$

The only nonzero matrix elements of $\hat{Y}$ are

$$Y_{k_\alpha l_\beta}^{i_\beta j_\alpha} = \langle i_\beta j_\alpha | \hat{Y} | k_\alpha l_\beta \rangle = \delta_k^i \delta_l^j, \tag{A.28a}$$

$$Y_{k_\beta l_\alpha}^{i_\alpha j_\beta} = \langle i_\alpha j_\beta | \hat{Y} | k_\beta l_\alpha \rangle = \delta_k^i \delta_l^j. \tag{A.28b}$$

The relevant spin components of the CPMFT and HFB model two-particle density matrix are

$$\Gamma_{i_\alpha j_\beta}^{k_\beta l_\alpha} = \frac{1}{2} \left(-\gamma_{i_\alpha}^{l_\alpha} \gamma_{j_\beta}^{k_\beta} \mp \kappa_{i_\alpha j_\beta} \kappa^{k_\beta l_\alpha} \right), \tag{A.29a}$$

$$\Gamma_{i_\beta j_\alpha}^{k_\alpha l_\beta} = \frac{1}{2} \left(-\gamma_{i_\beta}^{l_\beta} \gamma_{j_\alpha}^{k_\alpha} \mp \kappa_{i_\beta j_\alpha} \kappa^{k_\alpha l_\beta} \right), \tag{A.29b}$$

where again CPMFT (HFB) corresponds to the top (bottom) sign.

Contracting the two-particle density matrix with the matrix elements gives us

$$\langle \hat{Y} \rangle = -\mathrm{Tr}(\boldsymbol{\gamma}_{\alpha\alpha}\boldsymbol{\gamma}_{\beta\beta} \mp \boldsymbol{\kappa}_{\alpha\beta}\boldsymbol{\kappa}_{\alpha\beta}). \tag{A.30}$$

In the closed-shell case, using the results in Eq.(A.4), this becomes

$$\langle \hat{Y} \rangle = -\mathrm{Tr}(\mathbf{P}^2 \mp \mathbf{K}^2). \tag{A.31}$$

Then the expectation value of $\hat{S}_-\hat{S}_+$ is given by

$$\langle \hat{S}_-\hat{S}_+ \rangle = \mathrm{Tr}(\mathbf{P} - \mathbf{P}^2 \pm \mathbf{K}^2) \tag{A.32a}$$

$$= \mathrm{Tr}(\mathbf{K}^2 \pm \mathbf{K}^2). \tag{A.32b}$$

We therefore have

$$\langle \hat{S}_- \hat{S}_+ \rangle_{\mathrm{HFB}} = 0, \tag{A.33a}$$

$$\langle \hat{S}_- \hat{S}_+ \rangle_{\mathrm{CPMFT}} = 2\mathrm{Tr}(\mathbf{K}^2). \tag{A.33b}$$

Combining Eqs.(A.24) and (A.33) gives us the total spin contamination in HFB and in CPMFT:

$$\langle \hat{S}^2 \rangle_{\mathrm{HFB}} = 0, \tag{A.34a}$$

$$\langle \hat{S}^2 \rangle_{\mathrm{CPMFT}} = 3\mathrm{Tr}(\mathbf{K}^2). \tag{A.34b}$$

For UHF in cases in which there is strong correlation, we have the familiar formula

$$\langle \hat{S}^2 \rangle = s(s+1) + N_\beta - \mathrm{Tr}(\boldsymbol{\gamma}_{\alpha\alpha}\,\boldsymbol{\gamma}_{\beta\beta}). \tag{A.35}$$

For the closed-shell case, using the results in Eq.(A.7), we have

$$\langle \hat{S}^2 \rangle = \mathrm{Tr}[\mathbf{P} - (\mathbf{P} + \mathbf{M})(\mathbf{P} - \mathbf{M})] \tag{A.36}$$

$$= \mathrm{Tr}(\mathbf{P} - \mathbf{P}^2 + \mathbf{M}^2)$$

$$= 2\,\mathrm{Tr}(\mathbf{M}^2).$$

Appendix B

Exact constraints for the inactive space in CPMFT

We here derive the alternative constraints Eq.(3.19) that are imposed to the CPMFT Hamiltonian, starting from Eq.(3.18). First let us recall that $\tilde{\boldsymbol{\mu}}$ is non-zero only for the c and v orbitals in the NO basis,

$$
\tilde{\boldsymbol{\mu}} \;=\; \begin{pmatrix} \tilde{\boldsymbol{\mu}}_{cc} & \mathbf{0} & \tilde{\boldsymbol{\mu}}_{cv} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \tilde{\boldsymbol{\mu}}_{vc} & \mathbf{0} & \tilde{\boldsymbol{\mu}}_{vv} \end{pmatrix}. \tag{B.1}
$$

Now it is evident that the derivative of $\mathcal{L}^{\mathrm{CPMFT}}$ with respect to $\mathbf{K}$ includes not only $\boldsymbol{\Delta}$ but also $\tilde{\boldsymbol{\mu}}$, meaning the Hamiltonian in this scheme is given by

$$
\mathbf{H} \;=\; \begin{pmatrix} \mathbf{F} & -\boldsymbol{\Delta} + \tilde{\boldsymbol{\mu}} \\ -\boldsymbol{\Delta} + \tilde{\boldsymbol{\mu}} & -\mathbf{F} \end{pmatrix}, \tag{B.2}
$$

where $\mathbf{F} = \mathbf{F}^{\mathrm{cs}} + \boldsymbol{\mu}$. Consider the CPMFT equation (HFB equation) $[\mathbf{H}, \mathbf{R}] = \mathbf{0}$ in terms of $\mathbf{F}$, $\boldsymbol{\Delta}$, $\tilde{\boldsymbol{\mu}}$, $\mathbf{P}$, and $\mathbf{K}$:

$$
[\mathbf{F}, \mathbf{P}] - [\boldsymbol{\Delta} - \tilde{\boldsymbol{\mu}}, \mathbf{K}] = \mathbf{0}, \tag{B.3}
$$

$$
\{\mathbf{F}, \mathbf{K}\} + \{\boldsymbol{\Delta} - \tilde{\boldsymbol{\mu}}, \mathbf{P}\} - \boldsymbol{\Delta} = \mathbf{0}. \tag{B.4}
$$

At convergence of CPMFT, in the NO basis, we require that $\mathbf{P}$ and $\mathbf{K}$ be

$$
\mathbf{P} \;=\; \begin{pmatrix} \mathbf{I}_{cc} & \mathbf{0}_{ca} & \mathbf{0}_{cv} \\ \mathbf{0}_{ac} & \mathbf{P}_{aa} & \mathbf{0}_{av} \\ \mathbf{0}_{vc} & \mathbf{0}_{va} & \mathbf{0}_{vv} \end{pmatrix}, \tag{B.5}
$$

$$
\mathbf{K} \;=\; \begin{pmatrix} \mathbf{0}_{cc} & \mathbf{0}_{ca} & \mathbf{0}_{cv} \\ \mathbf{0}_{ac} & \mathbf{K}_{aa} & \mathbf{0}_{av} \\ \mathbf{0}_{vc} & \mathbf{0}_{va} & \mathbf{0}_{vv} \end{pmatrix}, \tag{B.6}
$$

where we recall that superscripts c, a, and v stand for the core, active, and virtual spaces. Substituting Eqs.(B.5) and (B.6) in the CPMFT equations yields the following set of equations: from Eq.(B.3)

$$\mathbf{F}_{ca}\left(\mathbf{P}_{aa} - \mathbf{I}\right) - \boldsymbol{\Delta}_{ca}\mathbf{K}_{aa} = \mathbf{0}, \tag{B.7}$$

$$\mathbf{F}_{va}\mathbf{P}_{aa} - \boldsymbol{\Delta}_{va}\mathbf{K}_{aa} = \mathbf{0}, \tag{B.8}$$

$$[\mathbf{F}_{aa}, \mathbf{P}_{aa}] - [\boldsymbol{\Delta}_{aa}, \mathbf{K}_{aa}] = \mathbf{0}, \tag{B.9}$$

$$\mathbf{F}_{cv} = \mathbf{0}, \tag{B.10}$$

and from Eq.(B.4)

$$\mathbf{F}_{ca}\mathbf{K}_{aa} + \boldsymbol{\Delta}_{ca}\mathbf{P}_{aa} = \mathbf{0}, \tag{B.11}$$

$$\mathbf{F}_{va}\mathbf{K}_{aa} + \boldsymbol{\Delta}_{va}\left(\mathbf{P}_{aa} - \mathbf{I}\right) = \mathbf{0}, \tag{B.12}$$

$$\{\mathbf{F}_{aa}, \mathbf{K}_{aa}\} + \{\boldsymbol{\Delta}_{aa}, \mathbf{P}_{aa}\} - \boldsymbol{\Delta}_{aa} = \mathbf{0}, \tag{B.13}$$

$$\boldsymbol{\Delta}_{cc} - \tilde{\boldsymbol{\mu}}_{cc} = \mathbf{0}, \tag{B.14}$$

$$\boldsymbol{\Delta}_{vv} - \tilde{\boldsymbol{\mu}}_{vv} = \mathbf{0}. \tag{B.15}$$

Note that Eqs.(B.9) and (B.13) correspond to the CPMFT equations of Eqs.(B.3) and (B.4) for the active space, and Eq.(B.10) to the Brillouin theorem of HF. The last two equations give the conditions for $\tilde{\boldsymbol{\mu}}_{cc}$ and $\tilde{\boldsymbol{\mu}}_{vv}$ that must be held at convergence. On the other hand, the above set of equations indicates that $\tilde{\boldsymbol{\mu}}_{cv}$ and $\tilde{\boldsymbol{\mu}}_{vc}$ are arbitrary.

Note that Eqs.(B.14) and (B.15) do not need to be satisfied at each SCF cycle unless a CPMFT calculation is converged. However, since there is no other simple prescription for $\tilde{\boldsymbol{\mu}}$ during the SCF cycles and Eqs.(B.14) and (B.15) are a sufficient approximation to the constraints, we *define* $\tilde{\boldsymbol{\mu}}$ via these equations and use it in each iteration.

Appendix C

Rationalization for CPMFT

We offer here a rationalization for the CPMFT model based on Legendre transforms [123]. Alternatively, one could use the Levy constrained-search formalism [124]. The universal density functional is

$$\mathcal{F}[\rho] = \sup_v \left(E(v) - \int v(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \right) \tag{C.1}$$

where

$$E(v) = \inf_\Psi \langle \Psi | T + V_{\mathrm{ee}} + \sum_{i=1}^{N_e} v(\mathbf{r}_i) | \Psi \rangle. \tag{C.2}$$

Here, T is the kinetic energy operator, V_{ee} the electron-electron interaction, and v the external potential. $\mathcal{F}[\rho]$ is split into a functional that is known ($\mathcal{F}_0[\rho]$) and a complement that needs to be approximated ($\bar{\mathcal{F}}_0[\rho]$),

$$\mathcal{F}[\rho] = \mathcal{F}_0[\rho] + \bar{\mathcal{F}}_0[\rho]. \tag{C.3}$$

A common choice for $\mathcal{F}_0$ is obtained by restricting E to a Hartree form [45],

$$\begin{aligned} \mathcal{F}_0[\rho] &\rightarrow \sup_v \left(E_{\mathrm{H}}(v, \rho) - \int v(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \right) \\ &= T_s[\rho] + U[\rho] \end{aligned} \tag{C.4}$$

where

$$E_{\mathrm{H}}(v, \rho) = \inf_\Psi \langle \Psi | T + \sum_{i=1}^{N_e} v(\mathbf{r}_i) | \Psi \rangle + U[\rho], \tag{C.5}$$

T_s is the non-interacting kinetic energy, and

$$U[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_1 d^3\mathbf{r}_2. \tag{C.6}$$

In this case, $\bar{\mathcal{F}}_0[\rho] \to E_{xc}[\rho]$. Another choice for $\mathcal{F}_0[\rho]$ is obtained by restricting E to a HF form [45]:

$$\mathcal{F}_0[\rho] \to \sup_v \left(E_{\mathrm{HF}}(v) - \int v(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \right) \tag{C.7}$$

where

$$E_{\mathrm{HF}}(v, \rho) = \inf_\Phi \langle \Phi | T + V_{\mathrm{ee}} + \sum_{i=1}^{N_e} v(\mathbf{r}_i) | \Phi \rangle. \tag{C.8}$$

Φ is restricted to Slater determinants. In this case, $\bar{\mathcal{F}}_0[\rho] \to E_c[\rho]$. In the same spirit, we can choose for $\mathcal{F}_0$ an HFB form

$$\mathcal{F}_0[\rho] \to \sup_v \left(E_{\mathrm{1HFB}}(v) - \int v(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \right) \tag{C.9}$$

where $E_{\mathrm{1HFB}}(v)$ is the HFB energy of a system in the external potential $v(\mathbf{r})$ and pairing interaction $-1/|\mathbf{r} - \mathbf{r}'|$. In this case, a new functional is obtained, $\bar{\mathcal{F}}_0[\rho] \to E_c^{\mathrm{1HFB}}[\rho]$. CPMFT is a mixture of HF and 1HFB. In all the cases above, there remains a density functional to be approximated (E_{xc}, E_c, E_c^{HFB}, E_c^{CPMFT}). The basic approximation for them could be a local density approximation or the use of alternative densities with existing exchange and correlation functionals described in Section 3.4.

Appendix D

Corresponding pairs property

Consider two idempotent matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$. The half sum of $\mathbf{T} = \frac{1}{2}(\mathbf{Q}_1 + \mathbf{Q}_2)$ has an eigenvector $\mathbf{e}$ with an eigenvalue t,

$$\mathbf{T}\mathbf{e} = t\mathbf{e}. \tag{D.1}$$

Then, using the fact that

$$
\begin{aligned}
\frac{1}{2}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{T}\mathbf{e} &= \frac{1}{4}(\mathbf{Q}_1 - \mathbf{Q}_2 + \mathbf{Q}_1\mathbf{Q}_2 - \mathbf{Q}_2\mathbf{Q}_1)\mathbf{e} \\
&= \frac{t}{2}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{e}
\end{aligned}
\tag{D.2}
$$

and thus

$$\frac{1}{4}(\mathbf{Q}_2\mathbf{Q}_1 - \mathbf{Q}_1\mathbf{Q}_2)\mathbf{e} = (\frac{1}{2} - t)\frac{1}{2}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{e}, \tag{D.3}$$

we arrive at

$$
\begin{aligned}
\mathbf{T}\frac{1}{2}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{e} &= \frac{1}{4}(\mathbf{Q}_1 - \mathbf{Q}_2 - \mathbf{Q}_1\mathbf{Q}_2 + \mathbf{Q}_2\mathbf{Q}_1)\mathbf{e} \\
&= \frac{1}{4}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{e} + (\frac{1}{2} - t)\frac{1}{2}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{e} \\
&= (1 - t)\frac{1}{2}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{e}.
\end{aligned}
\tag{D.4}
$$

Hence, $\tilde{\mathbf{e}} = \frac{1}{2}(\mathbf{Q}_1 - \mathbf{Q}_2)\mathbf{e}$ is also an eigenvector of $\mathbf{T}$ with an eigenvalue of $1 - t$, unless $\tilde{\mathbf{e}}$ vanishes for $t \in \{0, \frac{1}{2}, 1\}$.

Appendix E

CUHF algorithm

The detailed CUHF algorithm is as follows.

1. Choose the dimension of the active space N_a of the target system. It is suggested that N_a be the number of unpaired electrons. Then N_c is automatically determined $(N_c + N_a = N_\alpha)$.

2. Form $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$ as in UHF:

$$\gamma_{\mu\nu}^\sigma = \sum_i^{N_\sigma} C_{\mu i}^\sigma C_{\nu i}^{\sigma *} \tag{E.1}$$

   where $\mathbf{C}^\sigma$ is the CUHF σ canonical MO coefficients matrix.

3. Compute UHF $\mathbf{F}^\alpha$ and $\mathbf{F}^\beta$ from $\boldsymbol{\gamma}^\alpha$ and $\boldsymbol{\gamma}^\beta$.

4. Form $\mathbf{P}$, then orthonormalize and diagonalize it to obtain the NO coefficients matrix $\mathbf{C}^{\mathrm{NO}}$ and occupation numbers n_i.

5. Define the c, a, and v spaces in the NO basis. Open-shell orbitals must always be included in the active space. We are interested in the NOs whose occupations significantly deviate from 0 and 1. Thus, by sorting NOs in decreasing order of n_i, the first N_c orbitals are considered to be c, the next N_a orbitals are a, and the rest are all v.

6. Construct $\boldsymbol{\Delta}^{\mathrm{UHF}} = (\mathbf{F}^\alpha - \mathbf{F}^\beta)/2$, and transform it to the NO basis by using $\mathbf{C}^{\mathrm{NO}}$.

7. Build $\boldsymbol{\lambda}$ by flipping the sign of the cv and vc blocks of $\boldsymbol{\Delta}^{\mathrm{UHF}}$. Back transform $\boldsymbol{\lambda}$ to the atomic orbital (AO) basis.

8. Form $\tilde{\mathbf{F}}^{\alpha} = \mathbf{F}^{\alpha} + \boldsymbol{\lambda}$ and $\tilde{\mathbf{F}}^{\beta} = \mathbf{F}^{\beta} - \boldsymbol{\lambda}$.

9. Solve $\tilde{\mathbf{F}}^{\alpha}\mathbf{C}^{\alpha} = \mathbf{S}\mathbf{C}^{\alpha}\boldsymbol{\varepsilon}^{\alpha}$ and $\tilde{\mathbf{F}}^{\beta}\mathbf{C}^{\beta} = \mathbf{S}\mathbf{C}^{\beta}\boldsymbol{\varepsilon}^{\beta}$, where $\mathbf{S}$ and $\boldsymbol{\varepsilon}^{\sigma}$ are the AO overlap matrix and the CUHF σ orbital energies, respectively.

10. Go to 2.

The CUHF algorithm is very similar to UHF except for the additional steps 4-8. Methods designed for accelerating UHF convergence can be used straightforwardly in CUHF. Note that the above procedure is also very similar to CPMFT [2–5], except for the definition of $\boldsymbol{\Delta}^{\mathrm{UHF}}$.

Appendix F

Symmetry in CUHF and PCUHF wave functions

Here we discuss the symmetry of the $M = 0$ wave function in CUHF and PCUHF. Broken-symmetry solutions were obtained for CUHF in most calculations carried out in Chapter 5.

In our definition, core orbitals have $(\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta)/2$ eigenvalues of 1, *i.e.*, all the core orbitals are fully occupied. We remind the reader that the CUHF canonical orbitals do not always have a well-defined core space, while natural orbitals do. These natural core orbitals have the correct symmetry, as opposed to canonical orbitals which often break spatial symmetries, if one starts calculations from an initial guess with appropriate symmetry, *i.e.*, CUHF does not break symmetry in the core space. Thus, to consider the structure of a CUHF wave function, it is useful to define the first N_c CUHF α and β orbitals as the CUHF natural core orbitals, $\psi_c^\alpha = \psi_c^\beta = \psi_c^{\mathrm{NO}}$ where $c = 1, \cdots, N_c$, since a CUHF wave function is invariant with respect to unitary transformation in the core space. In the following, let us write this part of the wave function as $[CORE]$, which is symmetry adapted.

The active space is different. Once symmetry is broken, α and β orbitals do not remain equivalent. As explained in Section 5.2.1, the σ spin density matrix $\boldsymbol{\gamma}^\sigma$ in the NO basis can be written as a direct sum of 2×2 matrices as shown in Eq.(4.6). In particular, there is only one such matrix $\boldsymbol{\gamma}_k^\sigma$ with non-zero m_k for the CUHF(2) case. The rest of $\boldsymbol{\gamma}^\sigma$ are an identity matrix (core) and a zero matrix (virtual). In Section 5.2.2, we have defined localized α and β occupied spatial orbitals, ψ_k^α and ψ_k^β, by

diagonalizing $\boldsymbol{\gamma}_k^\alpha$ and $\boldsymbol{\gamma}_k^\beta$. For simplicity, let $\psi_k^\alpha = a$ and $\psi_k^\beta = b$. Then, the CUHF(2) wave function consists of

$$|\Phi_{\text{CUHF}}\rangle = [CORE]|a(\alpha)b(\beta)\rangle. \tag{F.1}$$

Therefore, to determine the symmetry of a CUHF(2) wave function, we only need to focus on the two-electron, two-orbital, broken-symmetry part of the wave function, $|a(\alpha)b(\beta)\rangle$. It can be said that if $\langle a|b\rangle = 1$, *i.e.*, $a = b$, the above wave function is a pure singlet, corresponding to the symmetry-adapted RHF solution. On the other hand, if $\langle a|b\rangle = 0$, this is a 50-50% mixture of a singlet and a triplet state. It is easy to show [68, 78] that upon spin-projection of this wave function, PCUHF(2), the following singlet state is obtained:

$$|\Phi_{\text{PCUHF}}\rangle = \frac{1}{\sqrt{2}}[CORE]\left\{|a(\alpha)b(\beta)\rangle + |b(\alpha)a(\beta)\rangle\right\}. \tag{F.2}$$

Note that the orbitals do not change during spin-projection (projection after variation). Now if a and b are symmetry adapted, then the PCUHF(2) wave function has definite symmetry. This is the case of CH_2 for the $^1B_1$ and $^3B_1$ mixture in Section 5.3.3, where a and b are of b_1 and a_1 symmetry. Therefore, after projection as in Eq.(F.2), the mixture reduces to $^1B_1$.

In many cases, a and b do not belong to any particular irreducible representation. For example, consider the UHF dissociation of H_2 where the a and b orbitals localize on the atoms. Note that CUHF(2) is equivalent to UHF in this case because there is no core space. A transformation is needed to obtain the correct symmetry orbitals:

$$\begin{pmatrix} \sigma_g \\ \sigma_u \end{pmatrix} = \begin{pmatrix} \frac{1}{2\sqrt{n_k}} & \frac{1}{2\sqrt{n_k}} \\ \frac{1}{2\sqrt{1-n_k}} & -\frac{1}{2\sqrt{1-n_k}} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \tag{F.3}$$

After some simple algebra, it can be shown that

$$|a(\alpha)b(\beta)\rangle + |b(\alpha)a(\beta)\rangle = c_1|\sigma_g(\alpha)\sigma_g(\beta)\rangle - c_2|\sigma_u(\alpha)\sigma_u(\beta)\rangle \tag{F.4}$$

where c_1 and c_2 are the so-called Sanibel coefficients as a function of n_k. Such a wave function turns out to be the correct Σ_g^+ state. This applies, for example, to the O_2 case in Section 5.3.2 where the a and b orbitals do not have the correct spatial symmetry but some linear combination of them yields both with π_g symmetry and the overall wave function Eq.(F.2) acquires $^1\Delta$ symmetry. As mentioned earlier, the singlet PCUHF wave function for O_2 we obtained is the $\pi_{x,g}(\alpha)\pi_{x,g}(\beta) - \pi_{y,g}(\alpha)\pi_{y,g}(\beta)$ component of $^1\Delta$, which is degenerate with the other component, $\pi_{x,g}(\alpha)\pi_{y,g}(\beta) - \pi_{y,g}(\alpha)\pi_{x,g}(\beta)$.

For all the systems tested in Section 5.3, we have confirmed that each PCUHF(2) wave function has the appropriate symmetry allowing direct comparison of ΔE_{ST} to the experimental values.

Appendix G

Glossary

The definition of each acronym and frequently used symbols are listed in Table G.1 and Table G.2.

Table G.1 : Definitions of acronyms used in this work.

Acronym	Definition
HF	Hartree-Fock
RHF	Restricted HF
UHF	Unrestricted HF
ROHF	Restricted Open-shell HF
HFB	Hartree-Fock-Bogoliubov
1HFB	ζ-HFB with $\zeta = -1$
CHF	Corrected HF
NO	Natural orbital
DFT	Density functional theory
CPMFT	Constrained pairing mean-field theory
CASSCF	Complete active space self-consistent field
CPGKS	Constrained pairing generalized Kohn-Sham DFT
UMP2	Unrestricted Second-Order Møller-Plesset Perturbation Theory
PUHF	Projected UHF
PUMP2	Projected UMP2
CUHF	Constrained UHF
CUMP2	Constrained UMP2
PCUHF	Projected CUHF
PCUMP2	Projected CUMP2
CUGKS	Constrained Unrestricted Generalized Kohn-Sham DFT

Table G.2 : Frequently used symbols.

Symbol	Definition
N_σ	Number of σ electrons ($\sigma = \alpha, \beta$)
N_e	$N_\alpha + N_\beta$
M	$(N_\alpha - N_\beta)/2$
N_s	$N_\alpha - N_\beta$
$\boldsymbol{\gamma}^\sigma$	One-particle density matrix of σ spin
$\boldsymbol{\Gamma}$	Two-particle density matrix
$\mathbf{P}$	Half charge density matrix, $(\boldsymbol{\gamma}^\alpha + \boldsymbol{\gamma}^\beta)/2$
$\mathbf{M}$	Spin density matrix, $(\boldsymbol{\gamma}^\alpha - \boldsymbol{\gamma}^\beta)/2$
$\mathbf{K}$	Anomalous density matrix, also known as pairing matrix, $\mathbf{K} = \sqrt{\mathbf{P} - \mathbf{P}^2}$
n_i	Occupation number of ith natural orbital
m_i	Eigenvalues of $\mathbf{M}$ as $\pm m_i$, $\sqrt{n_i(1-n_i)}$
κ_i	Eigenvalues of $\mathbf{K}$, $\sqrt{n_i(1-n_i)}$
$\rho(\mathbf{r})$	Electron density, $\rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r})$
$m(\mathbf{r})$	Spin polarization density, $\rho_\alpha(\mathbf{r}) - \rho_\beta(\mathbf{r})$
$P_2(\mathbf{r})$	On-top pair density
$\chi_m(\mathbf{r})$	Alternative spin polarization density
N_{cp}	Number of corresponding pairs in the natural orbital basis
N_c	Number of core orbitals in CUHF
N_a	Number of active orbitals in CUHF
$\mathbf{F}^\sigma$	UHF Fock matrix of σ spin
$\tilde{\mathbf{F}}^\sigma$	CUHF Fock matrix of σ spin
$\mathbf{F}^{\mathrm{cs}}$	$(\mathbf{F}^\alpha + \mathbf{F}^\beta)/2$ and $\mathbf{f}$
$\boldsymbol{\Delta}^{\mathrm{UHF}}$	$(\mathbf{F}^\alpha - \mathbf{F}^\beta)/2$
$\boldsymbol{\Delta}$	Pairing field in RHFB and CPMFT defined by Eqs.(2.49b, 3.10b)
μ	Chemical potential
δ_s	Spin-contamination
σ_s	Spin-deviation

Bibliography

[1] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Mineola, New York: Dover Publishing, 1996).

[2] T. Tsuchimochi and G. E. Scuseria, J. Chem. Phys. **131**, 121102 (2009).

[3] G. E. Scuseria and T. Tsuchimochi, J. Chem. Phys. **131**, 164119 (2009).

[4] T. Tsuchimochi, G. E. Scuseria, and A. Savin, J. Chem. Phys. **132**, 024111 (2010).

[5] T. Tsuchimochi, T. M. Henderson, G. E. Scuseria, and A. Savin, J. Chem. Phys. **133**, 134108 (2010).

[6] J. K. Ellis, C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, J. Chem. Phys. **135**, 034112 (2011).

[7] J. W. Hollett and P. M. W. Gill, J. Chem. Phys. **134**, 114111 (2011).

[8] C. A. Jiménez-Hoyos, T. M. Henderson, and G. E. Scuseria, J. Chem. Theory Comput. **7**, 2667 (2011).

[9] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).

[10] P. Ring and P. Schuck, *The Nuclear Many-Body Problem* (Springer-Verlag, 1980).

[11] J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems* (The MIT Press, 1986).

[12] V. N. Staroverov and G. E. Scuseria, J. Chem. Phys. **117**, 11107 (2002).

[13] A. Hagfeldt and M. Grätzel, Acc. Chem. Res. **33**, 269 (2000).

[14] N. Robertson, Angew. Chem., Int. Ed. **45**, 2338 (2006).

[15] B. O'Regan and M. Grätzelätzel, Nature (London) **353**, 737 (1991).

[16] R. Jain, K. Kabir, J. B. Gilroy, K. A. R. Mitchell, K. chung Wong, and R. G. Hicks, Nature **445**, 291 (2007).

[17] O. Kahn, *Molecular Magnetism* (VCH, New York, 1993).

[18] J. S. Miller, Inorg. Chem **39**, 4392 (2000).

[19] S. Ferlay, T. Mallah, R. Ouahes, P. Veillet, and M. Verdaguer, Nature **378**, 701 (1995).

[20] G. Csányi and T. A. Arias, Phys. Rev. B **61**, 7348 (2000).

[21] V. Bach, E. H. Lieb, and J. P. Solovej, J. Stat. Phys. **76**, 3 (1994).

[22] V. Rabanovich, Linear Algebra Appl. **390**, 137 (2004).

[23] A. J. Coleman, J. Math. Phys. **6**, 1425 (1965).

[24] N. N. Lathiotakis and M. A. L. Marques, J. Chem. Phys. **128**, 184103 (2008).

[25] V. N. Staroverov and G. E. Scuseria, J. Chem. Phys. **117**, 2489 (2002).

[26] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **110**, 2800 (1999).

[27] J. Čížek and J. Paldus, J. Chem. Phys. **47**, 3976 (1967).

[28] J. M. Bofill and P. Pulay, J. Chem. Phys. **90**, 3637 (1989).

[29] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).

[30] P. Mori-Śnchez, A. J. Cohen, and W. Yang, Phys. Rev. Lett. **102**, 066403 (2009).

[31] F. Moscardó and E. San-Fabian, Phys. Rev. A **44**, 1549 (1991).

[32] A. D. Becke, A. Savin, and H. Stoll, Theor. Chim. Acta. **91**, 147 (1995).

[33] J. P. Perdew, A. Savin, and K. Burke, Phys. Rev. A **51**, 4531 (1995).

[34] J. P. Perdew, M. Ernzerhof, K. Burke, and A. Savin, Int. J Quantum Chem. **61**, 197 (1997).

[35] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[36] A. Savin, "Recent developments and applications of modern density functional theory," (Elsevier, Amsterdam, 1996) pp. 327–357.

[37] T. Leininger, H. Stoll, H.-J. Werner, and A. Savin, Chem. Phys. Lett. **275**, 151 (1997).

[38] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, J. Chem. Phys. **115**, 3540 (2001).

[39] Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, and K. Hirao, J. Chem. Phys. **120**, 8425 (2004).

[40] J. Heyd, G. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).

[41] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **124**, 219906 (2006).

[42] O. A. Vydrov and G. E. Scuseria, J. Chem. Phys. **125**, 234109 (2006).

[43] O. A. Vydrov, J. Heyd, A. Krukau, and G. E. Scuseria, J. Chem. Phys. **125**, 074106 (2006).

[44] A. Vydrov, G. E. Scuseria, and J. P. Perdew, J. Chem. Phys. **126**, 154109 (2007).

[45] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[46] K. Yamaguchi and T. Fueno, Chem. Phys. **19**, 35 (1977).

[47] K. Takatsuka, T. Fueno, and K. Yamaguchi, Theoret. Chim. Acta **48**, 175 (1978).

[48] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. L **77**, 3865 (1996).

[49] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).

[50] M. Ernzerhof, J. P. Perdew, and K. Burke, Int. J. Quantum Chem. **64**, 285 (1997).

[51] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **110**, 5029 (1999).

[52] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

[53] G. Herzberg, *Molecular Spectra and Molecular Structure* (D. Van Norstrand Company, Inc., New Jersey, 1950).

[54] D. R. Lide, ed., *CRC Handbook of Chemistry and Physics, 89th ed.* (CRC Press/Taylor and Francis, Boca Raton, 2009).

[55] N. E. S. Y. Zhao and D. G. Truhlar, J. Phys. Chem. A **109**, 4388 (2005).

[56] K. E. Edgecombe and A. D. Becke, Chem. Phys. Lett. **244**, 427 (1995).

[57] G. E. Scuseria and H. F. S. III, Chem. Phys. Lett. **174**, 501 (1990).

[58] K. Andersson, B. O. R. P.-A. Malmqvist, and P.-O. Widmark, Chem. Phys. Lett. **230**, 391 (1994).

[59] V. E. Bondybey and J. H. English, Chem. Phys. Lett. **94**, 443 (1983).

[60] S. M. Casey and D. G. Leopold, J. Phys. Chem. **97**, 816 (1993).

[61] K. Hilpert and K. Ruthardt, Ber. Busenges. Phys. Chem. **91**, 724 (1987).

[62] J. Hachmann, W. Cardoen, and G. K. L. Chan, J. Chem. Phys. **125**, 144101 (2006).

[63] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[64] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

[65] J. A. Pople and R. K. Nesbet, J. Chem. Phys. **22**, 571 (1954).

[66] H. Fukutome, Int. J. Quantum Chem. **20**, 955 (1981).

[67] A. T. Amos and G. G. Hall, Proc. Roy. Soc. (London) **A263**, 483 (1961).

[68] J. E. Harriman, J. Chem. Phys. **40**, 2827 (1964).

[69] C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, Submitted .

[70] M. J. Frisch, G. W. Trucks, and H. B. S. *et al.* (GAUSSIAN, Revision G.01, 2007).

[71] C. C. Roothaan, Rev. Mod. Phys. **32**, 179 (1960).

[72] B. N. Plakhutin, E. V. Gorelik, and N. N. Breslavskaya, J. Chem. Phys. **125**, 204110 (2006).

[73] B. N. Plakhutin and E. R. Davidson, J. Math. Chem. **45**, 859 (2009).

[74] B. N. Plakhutin and E. R. Davidson, J. Phys. Chem. A **113**, 12386 (2009).

[75] E. R. Davidson and B. N. Plakhutin, J. Chem. Phys. **132**, 184110 (2010).

[76] K. R. Glaesemann and M. W. Schmidt, J. Phys. Chem. A **114**, 8772 (2010).

[77] T. Tsuchimochi and G. E. Scuseria, J. Chem. Phys. **133**, 141102 (2010).

[78] I. Zilberberg and S. P. Ruzankin, Chem. Phys. Lett. **394**, 165 (2004).

[79] J. S. Andrews, D. Jayatilaka, R. G. A. Bone, N. C. Handy, and R. D. Amos, Chem. Phys. Lett. **183**, 423 (1991).

[80] R. D. Amos, J. S. Andrews, N. C. Handy, and P. J. Knowles, Chem. Phys. Lett. **185**, 256 (1991).

[81] V. N. Glushkov, Int. J. Quantum Chem. **99**, 236 (2004).

[82] V. Bach, E. H. Lieb, M. Loss, and J. P. Solovej, Phys. Rev. Lett. **72**, 2981 (1994).

[83] C. Møller and M. Plesset, Phys. Rev. **46**, 618 (1934).

[84] N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, Chem. Phys. Lett. **164**, 185 (1989).

[85] W. J. Lauderdale, J. F. Stanton, J. Gauss, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **187**, 21 (1991).

[86] P. J. Knowles, J. S. Andrews, R. D. Amos, N. C. Handy, and J. A. Pople, Chem. Phys. Lett. **186**, 130 (1991).

[87] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **109**, 42 (1998).

[88] R. McWeeny and G. Diercksen, J. Chem. Phys. **49**, 4852 (1968).

[89] E. R. Davidson and A. E. Clark, Phys. Chem. Chem. Phys. **9**, 1881 (2007).

[90] S. Hirata and M. Head-Gordon, Chem. Phys. Lett. **314**, 291 (1999).

[91] L. Noodleman, J. Chem. Phys. **74**, 5737 (1981).

[92] L. Noodleman and E. R. Davidson, Chem. Phys. **109**, 131 (1987).

[93] P. O. Löwdin, Phys. Rev. **97**, 1509 (1955).

[94] P. O. Löwdin, Rev. Mod. Phys. **32**, 328 (1960).

[95] K. Yamaguchi, H. Fukui, and T. Fueno, Chem. Lett. **4**, 625 (1986).

[96] R. Caballol, O. Castell, F. Illas, I. de P. R. Moreira, and J. P. Malrieu, J. Phys. Chem. A **101**, 7860 (1997).

[97] V. N. Staroverov and E. R. Davidson, Chem. Phys. Lett. **330**, 161 (2000).

[98] S. Yamanaka, T. Kawakami, H. Nagao, and K. Yamaguchi, Chem. Phys. Lett. **231**, 25 (1991).

[99] A. E. Clark and E. R. Davidson, J. Chem. Phys. **115**, 7382 (2001).

[100] E. R. Davidson and A. E. Clark, Int. J. Quantum Chem. **103**, 1 (2005).

[101] T. Ziegler, A. Rauk, and E. J. Baerends, Theor. Chim. Acta. **43**, 261 (1977).

[102] C. J. Cramer, F. J. Dulles, D. J. Giesen, and J. Almlöf, Chem. Phys. Lett. **245**, 165 (1995).

[103] K. Yamaguchi, Chem. Phys. Lett. **33**, 330 (1975).

[104] H. B. Schlegel, J. Chem. Phys. **84**, 4530 (1986).

[105] H. B. Schlegel, J. Chem. Phys. **92**, 3075 (1988).

[106] P. J. Knowles and N. C. Handy, J. Phys. Chem. **92**, 3097 (1988).

[107] P. J. Knowles and N. C. Handy, J. Chem. Phys. **88**, 6991 (1988).

[108] J. S. Andrews, C. W. Murray, and N. C. Handy, Chem. Phys. Lett. **201**, 458 (1993).

[109] K. P. Huber and G. Herzberg, *Constants of diatomic molecules* (Van Nostrand Reinhold, New York, 1979).

[110] C. D. Sherrill, M. L. Leininger, T. J. Van, Huis, and H. F. Schaefer, III, J. Chem. Phys. **108**, 1040 (1998).

[111] Y. Yamaguchi, C. D. Sherrill, and H. F. S. III, J. Phys. Chem. **100**, 7911 (1996).

[112] P. Jensen and P. R. Bunker, J. Chem. Phys. **89**, 1327 (1998).

[113] P. G. Wenthold, J. Hu, R. R. Squires, and W. C. Lineberger, J. Am. Chem. Soc. **118**, 475 (1996).

[114] C. J. Cramer and B. A. Smith, J. Phys. Chem. **100**, 9664 (1996).

[115] A. F. Izmaylov, G. Scuseria, and M. J. Frisch, J. Chem. Phys. **125**, 104103 (2006).

[116] T. M. Henderson, A. F. Izmaylov, G. Scalmani, and G. E. Scuseria, J. Chem. Phys. **131**, 044108 (2009).

[117] J. Wang, A. D. Becke, and V. H. Smith, J. Chem. Phys. **102**, 3477 (1995).

[118] J. Gräfenstein and D. Cremer, Mol. Phys. **99**, 981 (2001).

[119] A. J. Cohen, D. J. Tozer, and N. C. Handy, J. Chem. Phys. **126**, 214104 (2007).

[120] P. G. Wenthold, R. R. Squires, and W. C. Lineberger, J. Am. Chem. Soc. **120**, 5279 (1998).

[121] T. Tsuchimochi and G. E. Scuseria, J. Chem. Phys. **134**, 064101 (2011).

[122] G. E. Scuseria, C. A. Jiménez-Hoyos, T. M. Henderson, K. Samanta, and J. K. Ellis, J. Chem. Phys. **135**, 124108 (2011).

[123] E. H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).

[124] M. Levy, Proc. Natl. Acad. Sci. U.S.A. **76**, 6062 (1979).