Fast electronic structure methods for studying highly-correlated molecular systems.

by

Gregory John Olaf Beran

B.S. (University of California, San Diego) 2000

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Chemistry

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Martin P. Head-Gordon, Chair
Professor William A. Lester, Jr.
Professor Alexis T. Bell

Fall 2005

The dissertation of Gregory John Olaf Beran is approved:

_____

Chair   Date

_____

Date

_____

Date

University of California, Berkeley

Fall 2005

Fast electronic structure methods for studying highly-correlated molecular systems.

Copyright 2005

by

Gregory John Olaf Beran

Abstract

Fast electronic structure methods for studying highly-correlated molecular systems.

by

Gregory John Olaf Beran

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Martin P. Head-Gordon, Chair

Though it is often very successful, the mean-field, or Hartree-Fock (HF), approximation for predicting molecular electronic structure in quantum chemistry cannot reliably describe systems with strong electronic correlation effects, such as transition states, bond-breaking, radicals, diradicals, and other interesting systems, and these difficulties can magnify when HF is used as a reference wave function for higher-level correlation methods. In a case study on the prediction of harmonic vibrational frequencies in diatomic radicals, we demonstrate the perils of using a poor HF reference with normally very accurate correlation methods. We also find that replacing the HF reference with one that is determined from a correlated wave function drastically reduces the pathologies in many cases. Unfortunately, these alternative methods are too computationally expensive to be applied to systems beyond a handful of atoms.

With the goal of developing efficient methods that can properly describe highly-correlated systems, we explore a two-step approach to computing the correlation energy. In this model, the leading correlation terms involving the chemically valence electrons are treated in a self-consistent coupled cluster fashion. The remaining correlations are approximated perturbatively from the improved reference.

We first demonstrate that such a procedure can viably approximate more expensive coupled cluster methods at reduced cost. Next, we explore the spatial character of these strongest correlation effects and find that they are often very localized in space. Even the simplest local model, Perfect Pairing (PP), improves upon HF in

many cases. We introduce an extremely efficient algorithm for PP and generalize the model to an unrestricted formalism to enable it to properly dissociate molecular bonds to their atomic limits and to treat systems with unpaired electrons.

Then a second-order perturbative correction to account for the numerous missing correlation terms is applied to the PP wave function. Overall, this method, PP(2), costs only up to a factor of a few times more than standard second-order perturbation theory based on the HF wave function, but it substantially improves predictions in many cases with strong correlations. We demonstrate the strengths and weaknesses of PP(2) in various interesting systems.

Professor Martin P. Head-Gordon
Dissertation Committee Chair

Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my wife, Suzanne, for her enduring love and support, particularly in the final months of my thesis. I would also like to thank my family, in whose eyes I am now fully viable.

My advisor Martin made graduate school a pleasure. I am extremely thankful for what he has taught me both as a researcher and as a friend. His guidance made this work possible. Equally important in my learning have been all of the members of the Head-Gordon group during my time there. Andreas Dreuw, Tony Dutoi, Steve Gwaltney, Yousung Jung, Alex Sodt, Joe Subotnik, and Yihan Shao have been particularly valuable friends and collaborators. I would also like to thank Geoff Galitz, who saved both my data and my sanity on a number of occasions.

Chapter 1

Introduction

1.1 Elementary quantum chemistry

Quantum mechanics provides the mathematical framework for describing a chemical reaction from the reactants, through any intermediate reactive species, and finally to the products. This information may be obtained by solving the Schrödinger equation,

$$\hat{H}_{tot}\Psi_{tot}(r_1, r_2, \ldots, r_n, R_1, R_2, \ldots, R_N) = E_{tot}\Psi_{tot}(r_1, r_2 \ldots, r_n, R_1, R_2, \ldots, R_N),$$

(1.1)

for the total wave function Ψ_{tot}, where the molecular Hamiltonian $\hat{H}_{tot}$ expressed in atomic units takes the form:

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \frac{1}{2}\sum_A \nabla_A^2 - \sum_{iA}\frac{Z_A}{|R_A - R_i|} + \sum_{ij}\frac{1}{|r_i - r_j|} + \sum_{AB}\frac{Z_A Z_B}{|R_A - R_B|}. \quad (1.2)$$

Here, r_i and r_j refer to electronic coordinates, R_A and R_B refer to nuclear coordinates, and Z_A and Z_B refer to the nuclear charges. The first two terms in $\hat{H}_{tot}$ correspond to the electronic and nuclear kinetic energies, followed by the nuclear-electronic attraction, the electron-electron and nuclear-nuclear repulsions. Thus, the accurate prediction of chemical processes requires the solution of this many-dimensional, second-order partial differential equation. In practice, this equation cannot be solved exactly for all but the most trivial of systems. It is therefore necessary to make various approximations that make the problem more tractable.

Given the large difference in masses of the electrons and nuclei (the mass of a hydrogen nucleus is roughly one thousand times that of an electron), we can perform an approximate separation of variables, known as the Born-Oppenheimer approximation. In this approximation, we solve the Schrödinger equation corresponding to the electronic Hamiltonian,

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{iA}\frac{Z_A}{|R_A - R_i|} + \sum_{ij}\frac{1}{|r_i - r_j|}, \tag{1.3}$$

for fixed positions of the nuclei $\mathbf{R}$ to obtain the electronic energy,

$$\hat{H}\Psi(r_1, r_2, \ldots, r_n; \mathbf{R}) = E\Psi(r_1, r_2, \ldots, r_n; \mathbf{R}). \tag{1.4}$$

This resulting solutions to the electronic Schrödinger equation for varying values of the nuclear position parameters $\mathbf{R}$ maps out the potential energy surface (PES). From this surface, one may study molecular dynamics, identify stable (minima) and transition state species (saddle points) on the PES, or compute reaction activation barriers, reaction kinetics, and reaction thermodynamics. Likewise, one may predict various spectroscopic properties such as infrared (IR), Raman, or nuclear magnetic resonance (NMR) frequencies by differentiating the PES with respect to various applied fields.

Therefore, the goal of present-day electronic structure theory is to accurately and uniformly predict these potential energy surfaces for small- to moderate-sized systems of chemical interest. Of course, to be useful, such predictions must be computationally affordable and should ideally be made available as tools that are easily used by practicing chemists. Furthermore, any approximations utilized should satisfy the desired criteria for a theoretical model chemistry.[6] (1) They should involve no system-specific, user-defined parameters. (2) Potential energy surfaces generated by the approximate methods should vary smoothly with changes in nuclear geometry and always remain differentiable. (3) The model should be size extensive, scaling appropriately with the number of electrons. This third criterion will be discussed in more detail below. From this point forward, we will consider only the solution of the electronic Schrödinger equation, which corresponds to the branch of quantum chemistry known as electronic structure theory.

Various strategies for solving the electronic Schrödinger equation exist. One of these, known as density functional theory (DFT)[7] takes advantage of various theorems proved by Hohenburg, Kohn, and Sham in the 1960s that demonstrate that the energy depends only on the three-dimensional density, rather than the full many-dimensional wave function $\Psi(r_1, r_2, \ldots, r_n)$. Unfortunately, although these theorems prove the existence of a functional that maps the density to the exact energy, they provide inadequate insight into what form this functional should take. Although many highly successful approximate density functionals have been developed, particularly in the last quarter century,[8, 9, 10, 11] there remain fundamental chemical interactions that current DFT functionals cannot describe, such as dispersion effects or highly correlated systems. Therefore, we shall focus primarily on the other strategy for solving the electronic Schrödinger equation: the wave-function based methods.

1.2 Wave-function-based methods

Virtually all, *ab initio* wave function methods begin with the mean-field Hartree-Fock (HF) approximation. In this model, the wave function is approximated in the form of a single, n-electron, antisymmetrized determinant of one-electron orbitals $|\phi_i\rangle$,

$$|\Phi_0\rangle = |\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\cdots\phi_n(\mathbf{r}_n)|. \tag{1.5}$$

These one-electron orbitals are, in turn, written as linear combinations of a finite number of atom-centered Gaussian basis functions (or predetermined contractions thereof) $|\chi_\mu\rangle$,

$$|\phi_i\rangle = \sum_\mu C_{\mu i}|\chi_\mu\rangle. \tag{1.6}$$

In the limit of an infinite basis, this latter expansion will be exact. Forcing the electrons to reside in separate, one-electron orbitals in this manner leads to a mean-field description of the electronic structure. The HF method solves for these molecular orbital coefficients $C_{\mu i}$ variationally. The resulting wave function provides a single configuration arranging the electrons in the lowest energy orbitals consistent with the Pauli antisymmetry principle.

Despite its simplicity, the HF approximation typically recovers roughly 99% of the total energy of the system. Unfortunately, chemistry depends on relative energies between related species, and this remaining one percent of the total energy, termed the correlation energy, is important for making quantitative chemical predictions. Formally, the correlation energy is defined as the difference between the exact energy and the mean-field (HF) energy in a given, finite basis:

$$E_{corr} = E_{exact} - E_{HF} \tag{1.7}$$

In order to recover the correlation energy, one resorts to another basis set expansion, this time in terms of n-electron determinants or "configurations", to find the exact wave function. Because the HF approximation typically captures the qualitative nature of the system, we often write this basis set expansion in terms of the degree of "excitation" of the other n-electron configurations from the HF one:

$$\begin{aligned} |\Psi_{exact}\rangle &= (1 + \hat{C}_1 + \hat{C}_2 + \hat{C}_3 + \cdots)|\Phi_0\rangle \tag{1.8} \\ &= |\Phi_0\rangle + \overset{singles}{\sum_{ia} c_i^a |\Phi_i^a\rangle} + \overset{doubles}{\sum_{ijab} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle} + \overset{triples}{\sum_{ijkabc} c_{ijk}^{abc} |\Phi_{ijk}^{abc}\rangle} + \cdots. \tag{1.9} \end{aligned}$$

The excitations are generated by n-body excitation operators $\hat{C}_n$. For example, in second quantization,

$$\hat{C}_2 = \sum_{ijab} c_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_i \hat{a}_b^\dagger \hat{a}_j. \tag{1.10}$$

In this notation, $i, j, k, \cdots$ refer to orbitals which are occupied in the HF reference determinant, and $a, b, c, \cdots$ refer to orbitals which are unoccupied (or virtual) in the HF reference. The operators $\hat{a}_i$ and $\hat{a}_a^\dagger$ annihilate an electron in orbital i and create one in orbital a, respectively. The doubles terms given in Eq. 1.10, for example, are determinants in which two electrons in the HF reference determinant are removed from orbitals i and j and placed into empty orbitals a and b. As in the HF approximation, the expansion coefficients $\{c_{ijk\cdots}^{abc\cdots}\}$ are solved for variationally. This method is called configuration interaction. If all possible configurations are included in the expansion, it is termed Full Configuration Interaction (FCI), and the method recovers the exact correlation energy within our chosen atomic orbital basis set. Unfortunately,

the number of configurations grows factorially with the number of orbitals/electrons, making the approach intractable for all but the smallest systems.

However, given that the HF approximation is generally fairly good, one might suspect that the most important terms in this expansion correspond to those configurations that are fairly similar to the HF determinant. Indeed, fairly good approximations are made by solving for the coefficients in a truncated configuration interaction expansion involving only the singles and doubles terms (CISD). If necessary, triples or quadruples could also be included, though, as will be discussed in more detail below, the computational effort grows rapidly with the inclusion of higher order terms in the CI expansion. Furthermore, size extensivity is lost in truncated CI expansions.

Alternatively, given the generally good quality of the HF wave function, we can approximate the exact energy and wave function using perturbation theory. In this case, we partition the Hamiltonian into the effective one-body operator, $\hat{F}$ which arises in the mean-field HF procedure (the zeroth order piece), and the remaining two-body interactions, $\hat{V}$ (the first order piece):

$$\hat{H} = \hat{F} + \hat{V}. \tag{1.11}$$

Working out the perturbative expansion, the first new contribution arises at second order and is termed second-order Møller-Plesset perturbation theory, or MP2.[12] Of course, one may continue to higher orders of the perturbation expansion. Empirical evidence suggests that this series is often not convergent and if higher accuracy is required, one should resort to alternative strategies.

Finally, the other major category of wave-function-based method are the coupled cluster (CC) methods.[13] Truncated CI expansions fail to satisfy size-consistency—the energy of two non-interacting subunits is not the same as the sum of the individual subunit energies. The CC Ansatz assumes a wave function of the form,

$$\begin{align} |\Psi\rangle &= e^{\hat{T}}|\Phi_0\rangle \tag{1.12} \\ &= (1 + \hat{T} + \frac{1}{2!}\hat{T}^2 + \frac{1}{3!}\hat{T}^3 + \cdots)|\Phi_0\rangle, \tag{1.13} \end{align}$$

where $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots$. The $\hat{T}_n$ operators in Eq. 1.13 are defined analogously

to those for the CI wave function in Eq. 1.10. For example,

$$\hat{T}_2 = \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_i \hat{a}_b^\dagger \hat{a}_j, \tag{1.14}$$

defines the double excitations.

In the limit of all possible excitations being included, the CC and CI Ansätze are equivalent, and both reproduce the FCI result. Truncated CC expansions, however, are vastly superior to their CI counterparts. Truncated CC expansions are properly size-consistent/size-extensive, and they tend to converge to the exact wave function more quickly with a given excitation order.[14, 15] Both of these observations stem from the exponential form of the wave function. Consider a simple model that includes only doubles (usually the leading correction to HF). In this case, the CC wave function has the form:

$$|\Psi_{CCD}\rangle = e^{\hat{T}_2}|\Phi_0\rangle = (1 + \hat{T}_2 + \frac{1}{2!}\hat{T}_2^2 + \frac{1}{3!}\hat{T}_2^3 + \cdots)|\Phi_0\rangle. \tag{1.15}$$

In addition to including all of the doubles, this wave function estimates the quadruples as products of two separate double excitations, the triples as products of three separate double excitations, and so forth. It is exactly these additional terms that make the wave function size extensive and provide a more balanced description of the correlations.

The CC equations could be solved variationally, like the HF and CI equations. However, it turns out that this is factorially complicated even for truncated CC models due to the exponential Ansatz.[16] Instead, the CC equations are usually solved projectively. Starting from the Schrödinger equation in the CC Ansatz,

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{1.16}$$

$$\hat{H}e^{\hat{T}}|\Phi_0\rangle = Ee^{\hat{T}}|\Phi_0\rangle, \tag{1.17}$$

we can left-multiply by $e^{-\hat{T}}$ to obtain a modified equation in terms of the similarity-transformed Hamiltonian, $\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$,

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E|\Phi_0\rangle. \tag{1.18}$$

In order to obtain the energy, then, we simply left-project both sides on to the reference determinant, $|\Phi_0\rangle$,

$$\langle \Phi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle = E \langle \Phi_0 | \Phi_0 \rangle = E. \tag{1.19}$$

This energy expression depends on the amplitudes $\{t\}$, and in order to solve for these we left-project with excited determinants. For example, to solve for CCSD, we would also left-project with all possible singly- and doubly-excited determinants $|\Phi_i^a\rangle$ and $|\Phi_{ij}^{ab}\rangle$,

$$\langle \Phi_i^a | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle \quad = E \langle \Phi_i^a | \Phi_0 \rangle \quad = 0 \tag{1.20}$$

$$\langle \Phi_{ij}^{ab} | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle \quad = E \langle \Phi_{ij}^{ab} | \Phi_0 \rangle \quad = 0. \tag{1.21}$$

These projections produce a set of coupled, non-linear equations for the amplitudes which must be solved iteratively. Unlike the variational solution of the CC equations, this approach leads to equations for the unknown amplitudes that naturally terminate. In the case of a doubles only method, the exponential expansion terminates with the estimated quadruples, meaning that the amplitude equations are only quadratic in nature (products of two doubles amplitudes leading to an estimated quadruple).

Currently, coupled cluster singles and doubles (CCSD),[17] or sometimes CCSD with a perturbative estimate of the triples, CCSD(T),[18, 19] are widely used for highly accurate calculations on small molecules (typically those with only a handful of non-hydrogen atoms).

1.3 Computational considerations and local correlation

Thus far we have outlined a hierarchy of successful approaches for computing the wave function and energies of molecules, beginning with the HF approach, which is typically only qualitatively accurate, to MP2, which behaves well for many standard, primarily-organic molecules near their equilibrium geometries, and finally to CCSD

and CCSD(T) for more accurate predictions across a wider range of species. Unfortunately, commensurate with the improvements in the description of the electron-electron correlations is a dramatic increase in computational cost. HF scales with the fourth power of the system size naively, and developments in the last decade or so have brought this down to quadratic or even and effective linear scaling.[20, 21, 22, 23] Canonical MP2 scales with the fifth power of the system size, CCSD with the sixth power, and CCSD(T) with the seventh power! This means that a doubling the size of the system to be studied increases the computational effort required by a factor of 32, 64, or even 128 times, respectively. Clearly, these steep scalings severely constrain the realm of applicability of the various correlation methods.

Fortunately, a little over twenty years ago Pulay demonstrated that these unphysically steep scalings could be significantly reduced by taking account of the localized nature of electron-electron correlations.[24] The molecular orbitals $|\phi_i\rangle$ obtained from a HF calculation, usually termed the canonical molecular orbitals, are in fact delocalized over the entire molecule. This means that a complete treatment of electron correlation must interact all possible electron configurations because all of the orbitals interact non-trivially.

However, the HF energy depends only on the space spanned by the occupied orbitals, rather than the representation of the orbitals spanning this space. Therefore, one may transform the basis in any manner that preserves the partitioning between the occupied and virtual spaces. Given this freedom, various researchers have proposed transformations that localize the canonical molecular orbitals into ones that look like the familiar bonding orbitals used in qualitative molecular orbital theory. Two such schemes, the Boys[25] and Pipek-Mezey[26] procedures, are widely used to localize the canonical orbitals. The Boys scheme localizes the orbitals via a functional that minimizes their spatial variances, while the Pipek-Mezey scheme minimizes the number of atomic centers on which each molecular orbital has an appreciable presence.

Using localization schemes like these in combination with numeric cutoffs based on the spatial distance between orbitals for which electrons to correlate, Saebø and Pulay demonstrated substantial computational savings could be obtained with minimal loss in accuracy.[24, 27, 28] Since their landmark work, a variety of local approximations

have been pursued,[29, 30, 31, 32, 33] with much of the work being done on improvements to the basic Saebø and Pulay local correlation model. In particular, it has been extended to coupled cluster methods.[34, 35, 36, 37] These Pulay-type models suffer from one significant flaw: the use of the numeric cutoff leads to discontinuous potential energy surfaces as orbitals move across the cutoff and suddenly become or cease to be correlated.[38]

In the late 1990s, Lee and co-workers proposed an alternative local correlation model that does not utilize user-defined numerical cutoffs or suffer from discontinuous energy surfaces.[39, 40] Instead, they partitioned the molecular correlations into spatial classes using atom-centered localized orbitals. Local truncations are made by including a subset of the one-atom, two-atom, three-atom, and four-atom pair excitations in MP2. In general, the more centers involved in the correlation, the less local it is. They found that for MP2, the triatomics-in-molecules, or TRIM, model recovered roughly 99% of the non-local MP2 correlation energy at a significantly reduced cost. Though these schemes do include some non-local terms, they recover the vast majority of the correlation energy with many fewer amplitudes than their non-local counterparts, scaling no worse than with the fourth power of system size. They also do not exhibit the discontinuities of the Saebø-Pulay models.

1.4 Breakdowns of the Hartree-Fock approximation

In addition to the steep computational scaling of these correlation methods, they also are all based on the fundamental assumption that the HF approximation provides a good, qualitative description of the system. When this is true, perturbation methods or low-order excitation CC methods (CCSD, for example) are able to approximate the correlation energy well. This situation corresponds to the case when the various $\{c\}$ or $\{t\}$ coefficients in Eqs. 1.9 and 1.13 are significantly smaller than 1, the coefficient of the HF determinant.

Consider now the case where one of these excited determinants, $|\Pi\rangle$ is also very

important and has a coefficient approaching the weight of the HF determinant. Now the simplest qualitatively-correct wave function takes the form,

$$|\Psi_{uncorr}\rangle = |\Phi_0\rangle + c_\pi|\Pi\rangle, \tag{1.22}$$

where $0 << c_\pi < 1$. We would say that this wave function captures the static correlation, meaning that it goes beyond the single-determinantal mean-field description in order to be flexible enough to describe the system qualitatively. Supposing now we applied an MP2-like correlation for each of these determinants, which amounts to an approximate treatment of double excitations from each configuration, this wave function would take the form,

$$|\Psi\rangle = (|\Phi_0\rangle + \sum_{ijab} t_{ij}^{ab}|\Phi_{ij}^{ab}\rangle) + c_\pi(|\Pi\rangle + \sum_{ijab} u_{ij}^{ab}|\Pi_{ij}^{ab}\rangle) \tag{1.23}$$

This type of correlation is termed dynamical correlation, and it is characterized by the short-range electron-electron scattering events. Of course, the determinant $|\Pi\rangle$ and its related configurations can be rewritten as excited versions of the HF reference $|\Phi_0\rangle$. Suppose that $|\Pi\rangle$ is simply a doubly-excited determinant relative to $|\Phi_0\rangle$,

$$c_\pi|\Pi\rangle = t_{kl}^{cd}|\Phi_{kl}^{cd}\rangle, \tag{1.24}$$

then the correlated wave function in Eq. 1.23 becomes,

$$|\Psi\rangle = |\Phi_0\rangle + \sum_{ijab} t_{ij}^{ab}|\Phi_{ij}^{ab}\rangle + t_{kl}^{cd}|\Phi_{kl}^{cd}\rangle + \sum_{ijab} t_{ijkl}^{abcd}|\Phi_{ijkl}^{abcd}\rangle. \tag{1.25}$$

Now, in terms of the HF reference, this wave function includes all doubly-excited configurations and a fraction of quadruply excited configurations! Clearly, a method like CCSD lacks the ability to accurately describe this wave function if we start from only the HF reference. A robust HF-based description would require CCSDTQ, which scales computationally with the ninth power of system size.

This example wave function is not simply an arbitrary construction, either. Such wave functions arise as chemical bonds are stretched. In fact, any case with more than one important configuration, which includes some radicals, diradicals, transition states, etc., requires a more careful treatment of the electronic correlation.

The standard approach for handling such strongly-correlated systems is to construct a multi-configurational reference function. This is accomplished either by the user choosing the important configurations to construct a useful multi-configurational self-consistent field (MCSCF) wave function, or by choosing a small number of "active" electrons and orbitals and then including all possible configurations involving those orbitals and electrons. This procedure, known as the complete active space self-consistent field (CASSCF) model,[41] is basically equivalent to FCI in a subspace of the orbitals. The only difference is that the active orbitals are variationally optimized to minimize the energy, therefore making sure the best possible orbitals are included in the typically small active space. Because such calculations are equivalent to FCI, they exhibit exponential computational scaling that limits their utility in practice to no more than fourteen active electrons and orbitals. These reference wave functions are then typically corrected perturbatively with various multi-reference perturbation theories.[42, 43, 44, 45]

In the hands of a skillful practitioner, these aforementioned methods are extremely powerful and can provide very good results. Unfortunately, given their computational expense, it is often difficult to assess whether, for example, the results are converged with respect to active space size in CASSCF calculations or whether an unexpected configuration is important for MCSCF calculations.

With the goal of making calculations more affordable and less prone to user error, Krylov *et al.* developed the valence orbital-optimized coupled cluster doubles (VOD) approximation to CASSCF.[46] Much like CCSD approximates FCI through the inclusion of singles and double excitations, VOD approximates CASSCF by performing a CCD calculation within an active space. The orbitals are simultaneously variationally optimized to find the best set of core occupied and active occupied and virtual orbitals. Single excitations are not included because in CCSD they are associated with orbital relaxation.[46] Since the orbitals are being relaxed via the orbital optimization, singles are neglected.

VOD scales with the sixth power of system size, analogously to CCSD, and enables the inclusion of many more active orbitals and electrons than standard CASSCF. Calculations on the order of 50 active orbitals/electrons are quite reasonable and

enable the user to simply select all valence orbitals as active, rather than requiring careful selection. Though not as robust as CASSCF (just as CCSD is less robust than FCI), VOD mimics CASSCF fairly well across a variety of difficult systems.

The formalism for VOD differs moderately from that for standard CCSD, and because it pertains to topics discussed in subsequent chapters, it merits further discussion. The standard CC equations, Eqs. 1.20 and 1.21 are projective rather than variational. Differentiating with respect to changes in the orbitals leads to a response term for the amplitudes. Alternatively, one can construct a Lagrangian functional,

$$L = \langle \Phi_0 | (1 + \hat{\Lambda}) e^{-\hat{T}} H e^{\hat{T}} | \Phi_0 \rangle, \tag{1.26}$$

with Lagrange multipliers given by the operator,

$$\hat{\Lambda} = \sum_{ijab} \lambda_{ab}^{ij} \hat{a}_i^\dagger \hat{a}_a \hat{a}_j^\dagger \hat{a}_b. \tag{1.27}$$

This functional satisfies the property that making the functional stationary with respect to the λ amplitudes,

$$\frac{\partial L_{CCSD}}{\partial \hat{\Lambda}} = 0, \tag{1.28}$$

produces the original equations for the t amplitudes given in Eq.1.21 and the correct CC energy. Likewise, making the functional stationary with respect to the t amplitudes,

$$\frac{\partial L_{CCSD}}{\partial \hat{T}} = 0 \tag{1.29}$$

produces equations for the λ amplitudes. When both sets of amplitudes have been solved for, orbital (or nuclear) derivatives can be obtained without explicit response terms. This formalism is simply another way to write the coupled cluster equations.

Orbital-optimized coupled cluster doubles (OD)[47] utilizes this Lagrangian formulation to go beyond standard coupled cluster doubles. In addition making the Lagrangian stable with respect to the amplitudes, it is made stationary with respect to the reference orbitals. In other words, the orbitals are variationally optimized to minimize the coupled cluster doubles. This enables the reference determinant in the coupled cluster wave function to find more reasonable orbitals (e.g. less symmetry-broken or spin-contaminated) than the HF ones and can substantially improve predictions, as will be seen in the next chapter.

VOD works basically the same way as OD, except instead of partitioning the reference orbitals into occupied and virtual subspaces, the orbitals are partitioned into four subspaces: inactive (or core) occupied, active (or valence) occupied, active virtuals, and inactive virtuals. VOD solves the coupled cluster equations only in the active occupied and virtual spaces. The orbital optimization procedure then requires mixing between all four subspaces: occupied and virtual, active and inactive.

VOD and OD have been demonstrated to reduce symmetry-breaking effects in molecules and remove other pathological effects. A second order perturbative correction, termed (2),[48, 5, 49, 3] has also been developed to correct VOD for the beyond-valence correlations and also some triple and quadruple excitations, which add dynamical correlation via the additional configurations included in the wave function.

1.5   Research directions

In the next chapter, we will explore the failure of many correlation methods to treat small radicals accurately and demonstrate how these difficulties arise from the inadequacy of the HF model.[50] More advanced correlation like OD(T) that do not rely on the HF reference are more robust and perform admirably for these particular systems. We also propose an inexpensive alternative to these advanced correlation methods that replaces the HF reference with one from density functional theory and then performs a standard CCSD(T) calculation. For these systems, the differences between this hybrid approach and the OD(T) approaches are minimal, while the computational savings are sizable. More recently, we applied similar techniques to a study on the cleavage of the N_2 triple bond under relatively mild reaction conditions using a molybdenum complex.[51] In that particular system, the mixture of high- and low-spin open-shell electronic states made reliable predictions very difficult with standard DFT or wave function approaches. Unfortunately, none of these novel coupled cluster techniques can be applied readily to systems with more than about ten non-hydrogen atoms. We desire more affordable methods that can perform reasonably well in these types of systems.

In Chapter 3, a two-part approach to computing the correlation energy is proposed and implemented.[52] A CCSD calculation is approximated using a CCSD calculation within a relatively small valence active space to obtain the most significant correlations robustly, and then a perturbative correction is used to obtain the remaining singles and doubles correlations. This basic model demonstrates the plausibility of such a two part approach for the efficient calculation of correlation energies in difficult systems.

Chapter 4 explores the spatial nature of valence space correlations, which typically comprise 30-40% of the total correlation energy.[53, 6] In particular, we decompose the correlations in VOD according to the number of spatial sites involved in the pair excitations and learn that same site correlations contribute some 50-70% of the valence correlation energy. Additionally including two-center valence correlations brings this figure to 85-95% of the valence correlation energy. The rest of the valence correlation energy is fairly evenly distributed among the many remaining three- and four-center correlations.

The one center model discussed in Chapter 4 already exists in the literature, and is termed the Generalized Valence Bond Perfect-Pairing (PP) approximation.[54, 55, 56] Though it was not originally written using the CC formalism, it can be and is then simply a strong local approximation to VOD.[57, 58, 53] In Chapter 5 we generalize the PP model to unrestricted and open-shell systems and develop and implement a highly efficient algorithm for PP that makes it very affordable and a possible replacement for HF in strongly-correlated systems.[59, 60]

Chapter 6 digresses to consider a technique for analyzing the important physical correlations from complex, many-body wave functions.[61] Inspired by the simplicity of the physical interpretation of PP wave functions, we utilize the singular value decomposition on the doubles piece of CCD-like wave functions to rewrite the doubles as a single pair excitation for each possible pair of electrons in the molecule. The pair excitations with the largest amplitudes are the most chemically significant, and we visualize these correlations in terms of the two-electron orbitals from which the pair originates and to which it is excited. The technique is completely general and may be applied to an arbitrary level of correlations in the wave function or to any MPn,

CC-, or CI-type wave function.

Finally, in Chapter 7 we discuss the derivation, implementation and reliability of a second-order perturbative correction built on top of PP, termed PP(2).[62] This efficient, two-part approach scales with the fifth power of the system size (like MP2, albeit with a larger prefactor) but obtains accuracy much closer to CCSD(T) than MP2 across a wider range of the potential energy surface. We conclude by demonstrating the utility of this new approach in various interesting chemical systems.

Chapter 2

Overcoming symmetry-breaking effects in small radicals

2.1 Introduction

Free radicals play a critical role in diverse areas of chemistry, including combustion, atmospheric chemistry, interstellar reactions, and biochemistry. Because of their high reactivity, studying them in the laboratory requires using special techniques like rare-gas matrices to trap and stabilize the radicals. Such experimental difficulties increase the need for accurate theoretical approaches to complement the experimental ones. A strong argument for using wave-function-based *ab initio* methods for the prediction of molecular structure and properties is their general reliability and capacity for systematic improvement. The hierarchy of Hartree-Fock (HF) theory, Second-Order Møller-Plesset Perturbation (MP2) theory, Coupled Cluster Single and Double excitations (CCSD)[17], and CCSD with a perturbative account of triple excitations (CCSD(T))[18] is typically used to systematically improve the quality of theoretical results. Extensive benchmarking has provided estimates of the accuracy of these different methods. For example, one study of 19 closed-shell molecules[63] found a mean absolute error in predicted bond lengths versus experiment of 0.57 pm for CCSD and 0.20 pm for CCSD(T) in the cc-pVTZ basis[64]. A different study of 10 small closed-shell molecules and 33 vibrational modes found a mean absolute

error for predicted harmonic frequencies of 3.7% and 2.2% for CCSD and CCSD(T), respectively, in the TZ(2df,2pd) basis[65].

However, both the magnitude of the errors and even the notion of systematic improvement along this hierarchy do not necessarily apply for open-shell molecules. In a recent study comparing bond lengths and frequencies of 33 small radical molecules against experiment using both unrestricted and restricted open-shell wave functions, Byrd *et al.*[66] found HF and MP2 theories to be too erratic for general use. CCSD performed much better, though at a somewhat lower statistical accuracy than is generally found for closed-shell molecules[63, 65]. Most surprising, however, was the fact that CCSD(T) results showed *little or no overall statistical improvement* over CCSD across a range of small to medium-sized basis sets, both for the restricted open-shell and the unrestricted cases. For the unrestricted case in the cc-pVTZ basis, mean absolute errors in bond lengths actually worsened, going from 0.56 to 1.25 pm for CCSD and CCSD(T) respectively, while harmonic frequencies demonstrated negligible improvement from mean absolute errors of 69 cm^{-1} for CCSD to 64 cm^{-1} for CCSD(T). The study also found that unrestricted B3LYP, which generated mean absolute errors of 0.75 pm and 71 cm^{-1}, performed similarly to unrestricted CCSD at a much reduced cost.

Although density functional theory (DFT) is a useful alternative to wave-function-based methods for radicals, in some cases one desires a higher level of accuracy than it or CCSD can provide. In this letter, we explore the cause of this failure of the (T) correction and propose several approaches which achieve better agreement with experiment. Perturbative methods and truncated coupled cluster (CC) theories cannot obtain accurate results when the HF reference is a poor one, as is often the case for radicals. Two obvious solutions exist to the problem of an inadequate reference: either improve the reference or utilize perturbative corrections that do not depend so strongly on the exact nature of the reference. In the former case, optimal orbital methods in the CC formalism[17, 67] such as Brueckner Coupled Cluster Doubles (BD)[68, 69] and Orbital-Optimized Coupled Cluster Doubles (OD)[47] may help to find an appropriate reference for each case. Alternatively, Kohn-Sham (KS) orbitals, which form a correlated single-determinantal description of the system, may provide a

reasonable and computationally inexpensive reference for CC methods as well. This approach is not entirely new: KS orbitals have been used as an alternative to HF ones successfully in Configuration Interaction (CI) excited-state calculations[70] and as a guiding function in Quantum Monte Carlo calculations[71]. Moreover, the similarity between Kohn-Sham and Brueckner orbitals has been suggested by several researchers[72, 73], and recently a more formal connection between the two has been proposed[74].

For the second case, we believe that the recently-developed (2) correction[48, 49] to coupled cluster wave functions may be less sensitive to the choice of reference and therefore supply a reasonable alternative to the (T) correction[18]. In the following sections we briefly describe the theory behind these methods and then investigate their behavior for a set of 12 first-row diatomic radicals.

2.2   Theory

Orbital-optimized coupled cluster methods have been described elsewhere[47, 75]. Here we will only briefly summarize them. OD[47], the most straightforward model, is based on the coupled cluster doubles (CCD) Ansatz, which can be written as a Lagrangian

$$L_{CCD} = \langle \Phi_0 | (1 + \hat{\Lambda}_2) e^{-\hat{T}_2} \hat{H} e^{\hat{T}_2} | \Phi_0 \rangle \tag{2.1}$$

where $|\Phi_0\rangle$ is the HF reference, $\hat{T}_2$ is the double excitation operator, and $\hat{\Lambda}_2$ is a Lagrange multiplier commonly used in CC gradient theory that acts as a double de-excitation operator. In standard CCD, the energy is obtained by making the Lagrangian stationary with respect to all excitation and de-excitation amplitudes in the sets $\{t_{ij}^{ab}\} \in \hat{T}_2$ and $\{\lambda_{ab}^{ij}\} \in \hat{\Lambda}_2$, which may be symbolically written as:

$$\frac{\partial L_{CC}}{\partial \hat{\Lambda}_2} = 0, \qquad \frac{\partial L_{CC}}{\partial \hat{T}_2} = 0. \tag{2.2}$$

In OD we solve the same equations, except we also make the wave function stationary with respect to variations in the orbitals by performing rotations (with individual angles θ_p^q) between the occupied and virtual subspaces until the minimum energy set

of orbitals is found:

$$\frac{\partial L_{CC}}{\partial \hat{\theta}} = 0. \tag{2.3}$$

This approach produces orbitals similar to the so-called Brueckner orbitals in the Brueckner doubles method[68], as is described in Ref. [47]. This optimized set of orbitals forms the "best" single reference for the given system/geometry as defined by the minimum CCD energy.

Beyond OD and CCSD, we utilize the non-iterative (2) correction[48, 49], which is a second-order perturbative correction based on the expansion of the similarity-transformed CC Hamiltonian, $e^{-\hat{T}}\hat{H}e^{\hat{T}}$. This correction is readily applicable to the CC methods described above and has been demonstrated to be more robust than the more common (T) correction although at a somewhat higher cost. For closed-shell molecules near equilibrium, (2) results are very similar to those from (T). However, in difficult diradical and bond-breaking cases, among others, it has been shown to be much more faithful to full CI. Formally, (2) scales as the seventh power of the system size—the same as (T)—but with a larger prefactor. In practice, a (2) calculation costs roughly 3-5 times more than a (T) one in our implementation. For more detail, the reader is referred to the original references.

Unfortunately, orbital optimization in the context of CC theories is computationally demanding at present. Ideally, one could efficiently find an adequately optimized or well-behaved set of orbitals from which further correlations could be included. In active space coupled cluster methods, one good choice might be the Imperfect Pairing[76] orbitals which we will show to approximate optimized orbitals at a reduced cost in Chapter 3. However, orbitals that are both cheaper and optimized for the full space are desirable. Unrestricted B3LYP has been shown to perform quite well for radicals.[66] More importantly, just like Brueckner methods, DFT approaches demonstrate improved stability against symmetry-breaking over Hartree-Fock theory[77]. This is advantageous from two perspectives. First, we want orbitals that vary smoothly and continuously with geometry changes, which does not occur near a symmetry-breaking point. Second, unlike symmetry-constrained Hartree-Fock wave functions, DFT or optimized/Brueckner orbitals require no constraints to re-

main symmetric. One way of viewing symmetry-breaking is as being driven by the asymptotic behavior of the system[78]: for example, the symmetric Hartree-Fock wave function for F_2^+ must break symmetry in order to approach the correct energetics of the two isolated atomic fragments, since the symmetric minimum is too high in energy. For correlated references (like Kohn-Sham or Brueckner), the potential well is deeper and the potential energy surface can smoothly reach the asymptotic limits without breaking symmetry. Thus, just as one may include correlations based on excitations from the HF orbitals, we explore the possibility of using BLYP Kohn-Sham orbitals[9, 10] (effectively, they form the natural orbitals for the BLYP density) as the reference from which to perform CC excitations. More specifically, the Kohn-Sham orbitals from the DFT calculation are used to form a single-determinantal reference, and from there the calculation proceeds like any other non-Hartree-Fock coupled cluster calculation[79, 19, 5]. The correlation energy is thus defined as the difference between the true energy and the energy expectation value of this KS determinant,

$$E_{corr} = E_{true} - \tilde{E}_0 = E_{true} - \langle \Phi_{KS} | \hat{H} | \Phi_{KS} \rangle. \tag{2.4}$$

If successful, this approach would provide a robust replacement for the HF reference with virtually no additional computational expense in high level coupled-cluster calculations.

2.3   Results and discussion

In this study, we compare the harmonic vibrational frequencies for twelve diatomic radicals (OF, F_2^+, CF, BO, NO, O_2^+, CN, CO^+, N_2^+, CH, FH^+, and OH) against experiment using the Dunning cc-pVTZ and cc-pVQZ basis sets[64] as calculated with CCSD, CCSD(T), CCSD(2), OD, OD(T), and OD(2). For simplicity, the notation OD(x) and CCSD(x) is sometimes used to refer to either the (T) or (2) theories. All correlated calculations use an unrestricted wave function, and only non-core orbitals were correlated. Experimental data were obtained from Refs [80, 81, 82, 83, 84]. All other single point calculations were performed using Q-Chem[85]. The harmonic

Table 2.1: Total energies (in hartrees) at the experimental geometry in the cc-pVTZ basis.

	R_e	CCSD	OD	KS-CCSD
CH	1.1199^a	-38.40709569	-38.40718259	-38.40565468
OH	0.9697^a	-75.63263205	-75.63260502	-75.63116680
FH$^+$	1.001^a	-99.75342511	-99.75341291	-99.75210724
BO	1.205^a	-99.85559766	-99.85479632	-99.85106296
CN	1.1718^a	-92.54722376	-92.54599638	-92.54143226
CO$^+$	1.1283^a	-112.62735434	-112.62608244	-112.62162841
N$_2^+$	1.11642^a	-108.78355436	-108.78199470	-108.77759668
CF	1.272^a	-137.59154043	-137.59088691	-137.58711350
NO	1.1508^a	-129.69838704	-129.69757854	-129.69381772
O$_2^+$	1.1164^a	-149.67115243	-149.67036505	-149.66639928
OF	1.354^b	-174.65391617	-174.65251797	-174.64811065
F$_2^+$	1.305^c	-198.70532070	-198.70425505	-198.70054976

^{a} Reference [81]. ^{b} Reference [83]. ^{c} Reference [82].

frequencies were determined from five-point fits of single-point calculations spaced 0.5 picometer apart around the experimental equilibrium geometry using PSI[86].

Since, to our knowledge, no results combining Kohn-Sham orbitals with coupled-cluster theory have been reported, in Table 2.1 we present the total energies at the experimental geometries for the various doubles-level methods and molecules in the cc-pVTZ basis. Notice that the differences in total energies between the different references are only a few millihartree. This fits well with the widely-held view that coupled-cluster methods have limited reference dependence energetically.

In Table 2.2, we present the root-mean-square (rms) absolute and percentage errors for each method for the entire test set described above. The individual frequencies are collected in Table 2.3. Figures 1 and 2 plot the errors of the calculated frequencies relative to experiment by method.

Looking at Table 2.2 and Figures 1 and 2, we see that CCSD(T) substantially improves upon CCSD for the radicals in general, though cases like NO and CO$^+$ still have errors of 216 and 106 cm^{-1}, respectively, in the cc-pVTZ basis (see Table 2.3). In the cc-pVQZ basis set, many of the errors decrease further for CCSD(T), but several outliers remain (CN, CO$^+$, and NO). These results suggest that although CCSD(T) performs very well for many radicals, the large scatter arising from certain

Table 2.2: Root-mean-square absolute and percent errors relative to experiment in the cc-pVTZ and cc-pVQZ basis sets for harmonic frequencies.

	CCSD	OD	CCSD(T)	OD(T)	CCSD(2)	OD(2)
Absolute rms errors (cm^{-1})						
cc-pVTZ	82	86	74	22	36	23
cc-pVQZ	92	105	75	21	42	26
Percent rms errors (%)						
cc-pVTZ	5.0	5.3	3.8	1.2	1.9	1.3
cc-pVQZ	5.5	5.8	3.9	1.2	2.2	1.6

cases makes it generally unreliable.

Two effects could be important in a proper description of radicals beyond the CCSD level: a better treatment of higher order excitations or a better initial reference than the HF one. Consider first the effects of approximately including triple and quadruple excitations through the (2) correction. CCSD(2) represents a significant improvement over both CCSD and CCSD(T) for both basis sets, reducing the rms percent errors by a factor of 2-3. In particular, CCSD(2) is superior to CCSD(T) in the outlying cases like NO, CN, and CO^+. This is not surprising, since previous studies have suggested that CCSD(2) depends less heavily on the quality of the reference wave function than CCSD(T)[49]. Despite the improvement with CCSD(2), the rms errors of 2-3% percent are still larger than is desirable. More importantly, the statistical errors are slightly misleading, since particular molecules like CO^+ and NO have errors much larger than the rest.

The other alternative is to modify the reference. Optimizing the orbitals as is done in OD should improve the reference (and therefore the correlation methods' accuracies). In fact, though we see little benefit upon optimizing the wave function at the doubles level (compare CCSD and OD rms errors—both methods behave similarly on most molecules), the difference between CCSD(T) and OD(T) is substantial. The rms errors decrease by almost a factor of four with the improved reference. In the cc-pVQZ basis, the error in the calculated NO frequency drops from 213 cm^{-1} to 4 cm^{-1} upon orbital optimization! The largest single OD(T) error is 37 cm^{-1}, as compared with 213 cm^{-1} for CCSD(T). Even this largest OD(T) error for CO^+ is a factor of three improvement over the same CCSD(T) frequency. This evidence
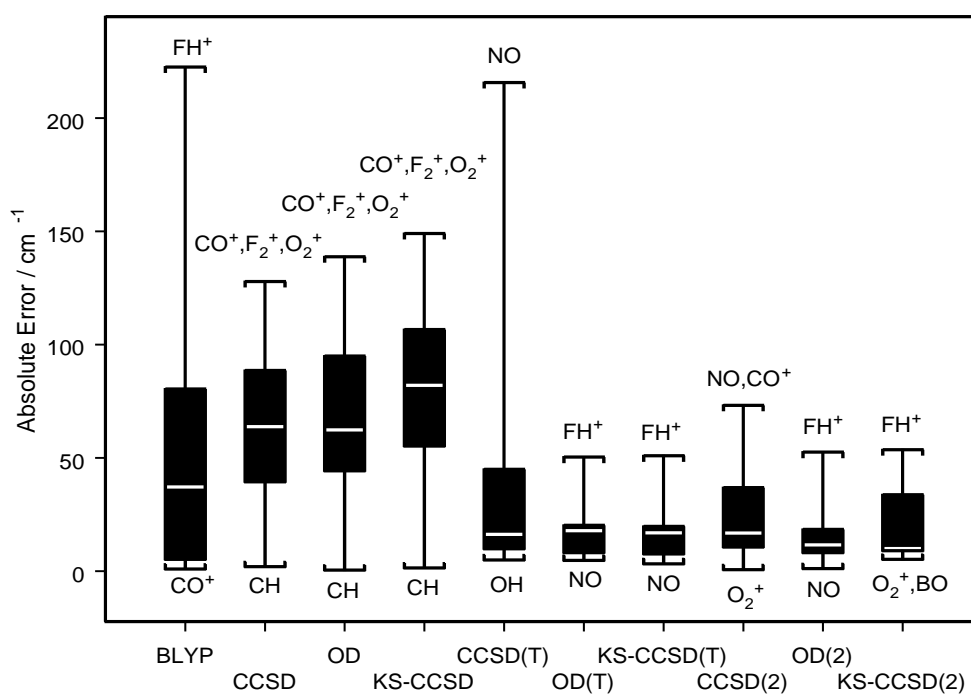
Figure 2.1: Errors (in cm^{-1}) versus experiment in calculated harmonic frequencies in the cc-pVTZ basis. The white line marks the median error, and the surrounding boxes delineate the middle 50 percent of results. The whiskers mark the extent of the data, and the adjacent molecules are responsible for those largest/smallest errors.
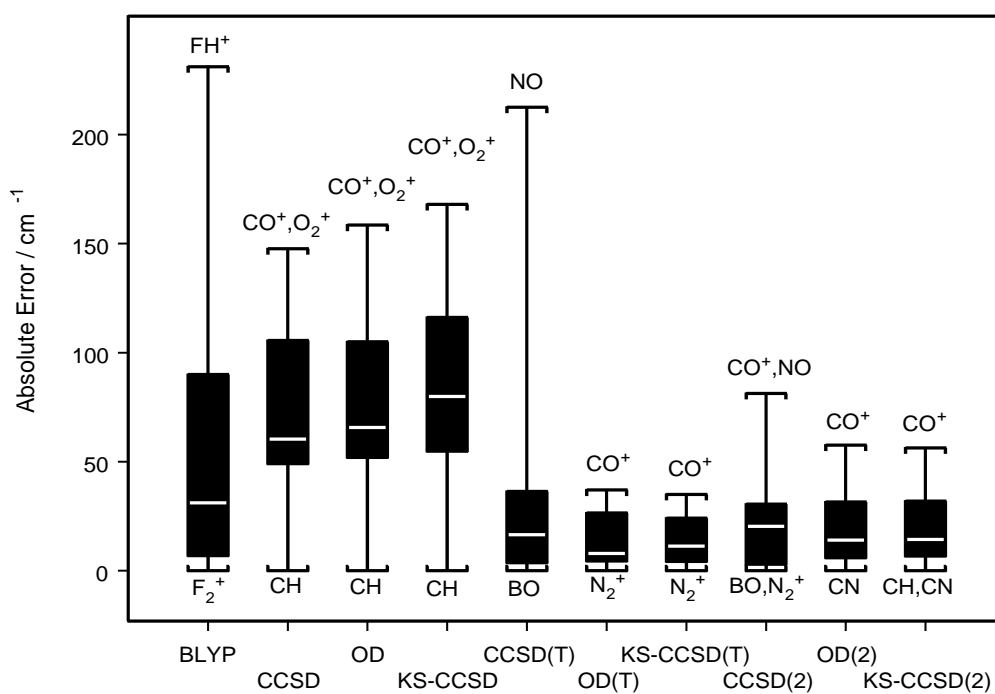
Figure 2.2: Errors (in cm^{-1}) versus experiment in calculated harmonic frequencies in the cc-pVQZ basis. The white line marks the median error, and the surrounding boxes the middle 50 percent of results. The whiskers mark the extent of the data, and the adjacent molecules are responsible for those largest/smallest errors.

Table 2.3: Errors (in cm^{-1}) relative to experimental harmonic vibrational frequencies in the cc-pVTZ and cc-pVQZ basis sets.

	CCSD	OD	CCSD(T)	OD(T)	CCSD(2)	OD(2)	Expt.
cc-pVTZ basis							
CH	-5.9	0.5	-18.5	-19.3	-16.9	-17.1	2858.5[c]
OH	39.5	42.0	7.0	6.9	7.8	8.2	3737.8[b]
FH$^+$	76	78	50	50	52	53	3090[a]
BO	44	50	-9	-20	-11	-7	1886[a]
CN	88.6	62.4	45.0	-19.2	36.9	-11.6	2068.6[b]
CO$^+$	124.2	134.4	106.2	17.9	64.0	38.4	2169.8[b]
N$_2^+$	64	77	-13	-16	-14	-9	2207[a]
CF	39	44	13	13	14	19	1308[a]
NO	85.9	95.0	215.7	-4.7	73.2	-1.1	1904.2[b]
O$_2^+$	127.9	138.8	9.9	8.2	0.7	4.8	1904.7[b]
OF	52	60	37	25	27	18	1053[a]
F$_2^+$	123	132	16	23	28	33	1104[c]
cc-pVQZ basis							
CH	14.6	15.5	-4.8	-5.7	-2.9	-3.2	2858.5[c]
OH	49.1	52.0	11.9	11.7	13.4	14.0	3737.8[b]
FH$^+$	58	61	28	29	31	32	3090[a]
BO	56	62	2	-8	2	6	1886[a]
CN	105.6	77.8	57.6	-6.1	58.8	2.1	2068.6[b]
CO$^+$	145.3	153.8	122.1	37.0	81.3	57.6	2169.8[b]
N$_2^+$	80	93	3	0	2	7	2207[a]
CF	33	38	4	3	5	10	1308[a]
NO	98.6	105.0	212.6	4.4	79.0	16.7	1904.2[b]
O$_2^+$	147.7	158.5	28.1	26.5	20.4	24.7	1904.7[b]
OF	60	66	36	29	28	32	1053[a]
F$_2^+$	128	137	17	23	29	34	1104[c]

[a] Reference [80].   [b] Reference [81].   [c] Reference [84].

Table 2.4: Root-mean-square absolute and percent errors relative to experiment for CCSD-based methods using UBLYP Kohn-Sham orbitals to approximate Brueckner orbitals.

	BLYP	KS-CCSD	KS-CCSD(T)	KS-CCSD(2)
Absolute rms errors (cm^{-1})				
cc-pVTZ	100	95	21	25
cc-pVQZ	110	114	20	28
Percent rms errors (%)				
cc-pVTZ	3.7	5.9	1.1	1.6
cc-pVQZ	3.8	6.3	1.0	1.7

strongly suggests that the failure of CCSD(T) for some of the radicals is the poor HF reference. OD(2) also improves significantly upon CCSD(2). The difference between the two theories is not nearly so large as for the (T) case, however, reflecting the superior performance of (2) over (T) with HF orbitals.

We notice that OD(T) follows experiment more faithfully on the whole than does OD(2) in these basis sets. It seems that given a good reference, the (T) theory is more accurate for these molecules. Such agreement must in a sense be fortuitous, however, since the (2) correction contains additional terms of quadruple excitation character. Of course, in the absence of a stable reference, the (2) correction behaves more robustly. The better orbitals also restore the systematic improvement from CCSD to CCSD(T). The OD(T) rms errors of 22 and 21 cm^{-1} compare favorably with the closed-shell CCSD(T) results for 13 molecules with 36 harmonic frequencies of Martin who found rms errors of 19 and 13 cm^{-1} for the cc-pVTZ and cc-pVQZ basis sets, respectively[87]. The neglected core correlation may also account for a sizable portion of the residual errors: Pawłowski *et al* recently attributed errors of up to 11 cm^{-1} in closed-shell diatomic molecules to it[88].

Because of the high computational expense of orbital-optimized methods, we also studied the effects of using unrestricted BLYP Kohn-Sham (KS) orbitals to approximate Brueckner orbitals for CCSD, CCSD(T), and CCSD(2) calculations. The data and statistics for these hybrid approaches and BLYP are presented in Tables 2.4 and 2.5. BLYP provides a reasonable starting point, with median errors significantly below those for CCSD. There is actually significant scatter for the hydrogen-

containing radicals with BLYP. Nevertheless, for both CCSD(T) and CCSD(2) with KS orbitals, we see substantial improvement over the calculations using HF orbitals, even in the cases where BLYP behaves poorly. In fact, these orbitals seem to be similar in quality to the OD orbitals—in most cases *KS-CCSD(x) frequencies are within a few wavenumbers of OD(x) values* at a drastically reduced computational expense. For CCSD, the Kohn-Sham orbitals make only a small difference (just like the OD orbitals) for the most part, except for CN (one of the molecules where orbital optimizations were particularly important at the doubles-level theories).

What is it about these alternative references that differentiates them from HF? Significant failures of unrestricted HF-based (UHF) methods for calculations of properties of open-shell molecules like doublet radicals are commonly blamed on spin contamination which can unphysically deform the UHF potential energy surface (PES), and therefore it provides a poor starting point for further correlations[89]. Although including correlation effects generally reduces spin contamination, the effect is not always sufficient to restore the PES. For example, the rate of convergence of the unrestricted MPn series is drastically reduced[90, 91, 92]. Optimized orbitals are known to reduce spin contamination significantly[93], as does density functional theory[94]. CN and CO$^+$ in particular have substantial spin contamination at the HF level, with $\langle \hat{S}^2 \rangle$ values of 1.00 and 1.16 (compared to 0.75 for a doublet spin eigenstate), respectively. Improving the reference using OD or BLYP virtually eliminates spin contamination, reducing $\langle \hat{S}^2 \rangle$ to 0.75-0.77. Figure 3 plots the ratio $\langle \hat{S}^2 \rangle_{ODref} / \langle \hat{S}^2 \rangle_{HF}$ and the OD(T) and CCSD(T) frequency errors for the cc-pVTZ basis, and Figure 4 plots the same for BLYP-CCSD(T) and CCSD(T). Small values (less than 1.0) of the ratio of $\langle \hat{S}^2 \rangle$ values indicate a marked decrease in spin contamination. Both of these heavily spin-contaminated molecules have some of the largest errors with CCSD(T), which is consistent with difficulties observed in the MPn series, and the frequencies dramatically improve with the better reference (OD or BLYP) and its decreased spin contamination. Similar results hold for the cc-pVQZ basis. The actual OD or CCSD wave functions decrease $\langle \hat{S}^2 \rangle$ even further, but this improvement does not help the (T) correction, which depends more directly on the HF reference (unlike (2), which is based off of the coupled-cluster wave function). However, spin contamination cannot

Table 2.5: Errors (in cm^{-1}) relative to experimental harmonic vibrational frequencies using BLYP Kohn-Sham orbitals as a reference for CC correlations

	BLYP	KS-CCSD	KS-CCSD(T)	KS-CCSD(2)	Experiment
cc-pVTZ basis					
CH	-143.1	1.4	-19.8	-17.7	2858.5
OH	-189.6	45.2	7.6	9.1	3737.8
FH$^+$	-223	82	51	54	3090
BO	-49	60	-23	-5	1886
CN	5.3	65.2	-17.8	-9.7	2068.6
CO$^+$	2.6	148.1	16.9	37.1	2169.8
N$_2^+$	27	84	-18	-10	2207
CF	-80	55	6	21	1308
NO	-55.7	106.7	-3.3	7.3	1904.2
O$_2^+$	-37.2	149.1	9.6	5.7	1904.7
OF	-14	84	12	35	1053
F$_2^+$	4	137	25	33	1104
cc-pVQZ basis					
CH	-135.5	16.5	-6.0	-3.8	2858.5
OH	-182.0	54.8	11.9	14.3	3737.8
FH$^+$	-231	64	28	32	3090
BO	-45	71	-11	7	1886
CN	5.2	79.9	-4.2	4.2	2068.6
CO$^+$	6.9	167.5	35.0	56.3	2169.8
N$_2^+$	26	100	-2	7	2207
CF	-90	48	-4	12	1308
NO	-55.7	116.2	5.6	17.1	1904.2
O$_2^+$	-31.1	168.0	27.4	25.2	1904.7
OF	-11	90	12	36	1053
F$_2^+$	0	142	24	35	1104

account for all of the errors. For example, NO displays relatively little spin contamination at the HF level.

More recently, Crawford *et al*[95] analyzed the effects of HF orbital instabilities ("instability volcanos"), or regions where the wave function is likely to break spatial symmetry, for both HF and correlated methods. Many of the molecules in this test set are either notorious for symmetry breaking or are isoelectronic with such molecules (NO for example is isoelectronic to O_2^+) . They found that the instability region on the PES is significantly larger for perturbative methods such as MPn and CCSD(T) than it is for a method like CCSD. Thus, the erratic behavior of CCSD(T) as compared to CCSD actually comes as little surprise. The improved performance of KS-CCSD(x) and OD(x) is also not too surprising, since both methods have been shown to be much more resistant (though not impervious[96, 77, 97, 98]) to symmetry-breaking[77, 97, 99, 100, 47]. Thus, it is most likely that rapid changes in the Hartree-Fock orbitals as a function of nuclear displacement are responsible for the poor properties predictions. Perhaps the success of KS-CCSD(x) lies simply in the fact that the Kohn-Sham reference changes more smoothly with nuclear displacement than the Hartree-Fock one, and therefore provides a better starting point for the perturbative corrections.

In order to analyze these spatial wave function instabilities in an "orbital-free" manner (*i.e.* in a way that is not confused by the various transformations among the orbitals), we look at how the reference one-particle density matrix changes with respect to nuclear displacements. In particular, the derivative with respect to internuclear distance $\mathbf{P^x}$ at the experimental equilibrium geometry is constructed by finite difference (step size $\delta = 0.0001$ Angstroms) in an orthogonalized atomic orbital basis:

$$\mathbf{P^x} = \frac{\mathbf{S(r_A} + \delta)^{\frac{1}{2}}\mathbf{P(r_A} + \delta)\mathbf{S(r_A} + \delta)^{\frac{1}{2}} - \mathbf{S(r_A)}^{\frac{1}{2}}\mathbf{P(r_A)}\mathbf{S(r_A)}^{\frac{1}{2}}}{\delta}$$

This quantity is the transition density from the experimental geometry at $\{\mathbf{r_A}\}$ to a new geometry at $\{\mathbf{r_A} + \delta\}$ and is calculated separately in the alpha and beta orbital spaces. The norm of this derivative density matrix gives a measure the maximum rate of change of the wave function with changes in the nuclear geometry.

A quantitative correlation between the absolute norm of $\mathbf{P^x}$ and the error in CCSD(T) was not observed (the decomposition of $\mathbf{P^x}$ into its different contributions
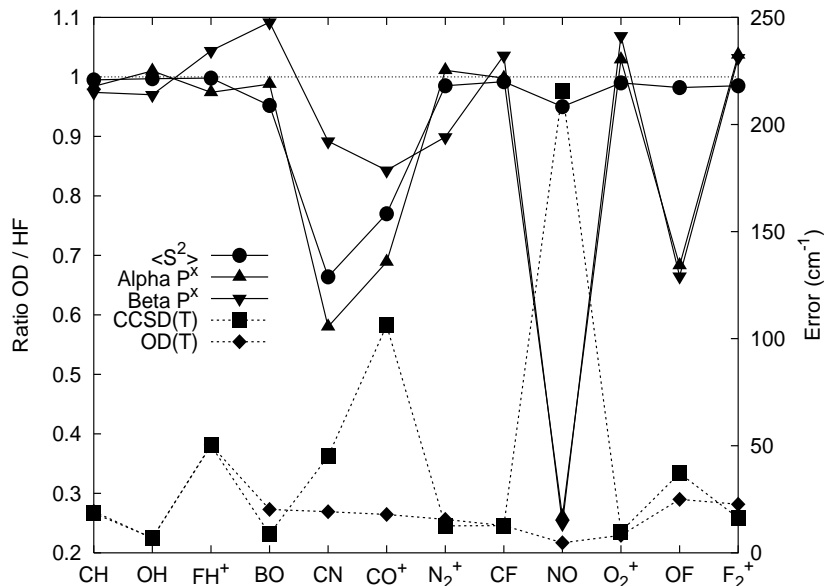
Figure 2.3: Ratio of the $\langle S^2 \rangle$ values and $\|\mathbf{P^x}\|$ for the OD and HF references (left axis), and the errors in calculated frequencies for OD(T) and CCSD(T) (right axis) in the cc-pVTZ basis. Smaller ratios (less than 1.0) indicate a reduction in the error or a more smoothly changing wave function upon improving the reference from HF to OD. Values greater than 1.0 indicate that the OD reference is changing more rapidly than the HF one.

is underway to see if such a correlation does exist), but a very good correlation exists between the ratio of the norms between OD (or BLYP) references and HF (*e.g.* $\|\mathbf{P^x_{ODref}}\|/\|\mathbf{P^x_{HF}}\|$), as is plotted in Figures 3 and 4 for the cc-pVTZ basis. Notice in particular the large improvement in the rate of change in the density going from HF to the OD reference for CN, CO^+, and NO. Each of these molecules exhibits dramatic improvement in the predicted frequency as the reference improves. For OF, the CCSD(T) results are reasonable, and they improve further with the more stable reference for OD(T)/BLYP-CCSD(T).

For FH^+, OD(T) or BLYP-CCSD(T) improve negligably upon CCSD(T) and the errors remain fairly large, so the spin contamination and $\|\mathbf{P^x}\|$ ratios near 1.0 come as no surprise. FH^+ apparently requires an even larger basis set (the error in the cc-pV5Z basis shrinks by a third) for more accurate predictions. In a few cases the rate
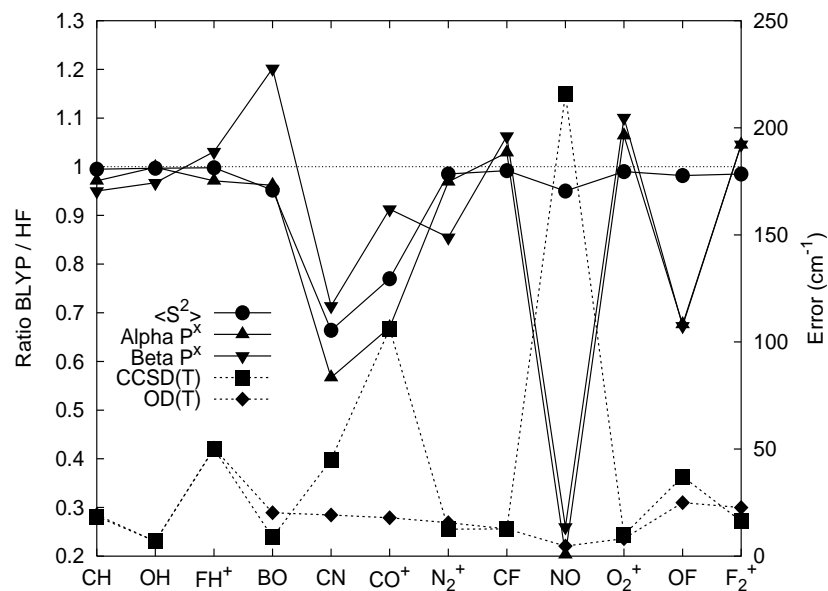
Figure 2.4: Ratio of the $\langle S^2 \rangle$ values and $\|\mathbf{P^x}\|$ for the BLYP and HF references (left axis), and the errors in calculated frequencies for BLYP-CCSD(T) and CCSD(T) (right axis) in the cc-pVTZ basis. Smaller ratios (less than 1.0) indicate a reduction in the error or a more smoothly changing wave function upon improving the reference from HF to BLYP. Values greater than 1.0 indicate that the BLYP reference is changing more rapidly than the HF one.

of change in the orbitals is slightly faster for OD than for HF (as indicated by a ratio > 1.0), but it does not seem to be sufficiently fast to cause serious problems in the predicted frequencies. In the worst case, the beta space ratio $\|\mathbf{P^x_{ODref}}\|/\|\mathbf{P^x_{HF}}\|$ for BO approaches 1.1, and we see that OD(T) is actually slightly worse than CCSD(T), though not horribly so. Overall, a similar correlation exists between reduction in errors and more smoothly changing reference wave functions in the cc-pVQZ basis set as well.

Finally, at first glance, it may seem surprising that a pathological symmetry-breaking molecule like F_2^+ does not also have a rapidly changing density. However, this simply means that in this particular basis, the symmetry-breaking point is sufficiently far away from the equilibrium geometry that the wave function is well-behaved in the relevant region. This also explains the relatively small errors observed for F_2^+ and O_2^+ frequencies in this study. Near the symmetry-breaking points, similar rapid changes in the density are observed at the HF level for these molecules (and they are reduced using the correlated determinants).

Thus, rapidly changing spin and spatial wave functions seem to create a poor starting point for perturbative methods. References like OD or DFT that resist symmetry-breaking (both spin and spatial) provide a superior starting point and can minimize unphysical property predictions. Of course, in extreme cases where even these methods break symmetry, their advantages may well disappear.

2.4   Conclusions

We have demonstrated that near-closed-shell CCSD(T)-type accuracy may be obtained when studying open-shell molecules like radicals using the OD(2) or OD(T) approaches, with rms percent errors for harmonic frequencies hovering around 1.2-1.6% for the cc-pVTZ and cc-pVQZ basis sets. The failure of CCSD(T) for open-shell molecules can likely be attributed to orbital instabilities (or near instabilities) in the Hartree-Fock solutions. We have analyzed these instabilities in terms of the extent of spin-contamination and the rate of change of the electron density in the reference wave functions, and demonstrated that correlated determinants like OD and DFT

tend to reduce their magnitudes and thus provide a better starting point for including correlations.

Though formally scaling the same as CCSD, the OD iterations involve additional repeated steps that scale to the *sixth power* with the number of orbitals. In practice, a large-basis-set OD(x) calculation is much more expensive than a similar CCSD(x) one and quickly becomes impractical at present. Thus, less costly alternatives are desirable. We recommend two candidates:

1. CCSD(2). Since the (2) correction appears much less sensitive to the choice of reference, this approach may make for a reasonable alternative to the optimized-orbital methods. However, there are clearly cases like NO and CO^+ where HF orbitals are inadequate.

2. CCSD(x) with Kohn-Sham orbitals. This approach is clearly the least expensive of the methods explored here, with KS-CCSD(x) costing only trivially more than a standard CCSD(x) calculation.

Of course, studies on both approaches with a wider set of test molecules is desirable, particularly for the KS-CCSD(x) methods. Nevertheless, it appears that radicals can be treated with accuracy approaching that of closed-shell molecules using "practical" methods.

Unfortunately, the definition of affordable depends drastically on what size system one wants to study. Methods like CCSD or CCSD(T) are really only practical up to 10 or so non-hydrogen atoms. For this reason, we would like to develop efficient methods to study difficult systems like these radicals that can be applied to tens of non-hydrogen atoms. In the next chapter, we explore a two-part approach to approximating the CCSD energy. The valence correlation, which typically corresponds to 30-40% of the total correlation energy and is typically the most challenging to treat correctly, will be treated at the CCSD level. The beyond-valence, or dynamical correlation, will be treated perturbatively, since it is exactly that type of correlation for which MP2 excels.

Chapter 3

A two-part approach to computing the correlation energy

3.1 Introduction

Despite recent advances in computer hardware, accurate *ab initio* electronic structure calculations on large molecules remain difficult due to the high formal scaling of the computational cost of standard wave-function-based correlation methods. The widely-used second-order Møller-Plesset perturbation theory (MP2) scales formally as N^5, and the generally more accurate Coupled Cluster Singles Doubles (CCSD) method scales as N^6, where N is the number of basis functions. However, these formal scalings are unphysical, [101] and one would like to reduce this formal scaling using appropriate physical and numerical approximations.

One such approach to reducing this high cost has been through the so-called local correlation methods pioneered by Saebø and Pulay.[27, 28] In methods of this type, molecular orbitals are localized in some fashion and excitation amplitudes (whether in the context of MP2 or coupled cluster theories) are explicitly treated only between orbitals within some predefined spatial distance. Though highly efficient and often quite accurate,[32] these methods contain a degree of arbitrariness in how these cut-offs are defined and also can produce discontinuous potential energy surfaces as molecular geometry changes.[28] These problems can be avoided at additional computational cost

by utilizing different criteria for selecting significant excitation amplitudes. One such model implemented for MP2 is the TRIM model,[40] in which only double excitations with one occupied and one virtual orbital on a common atom are retained. Another alternative is to use atomic orbital basis cutoffs directly.[31, 102]

The purpose of this chapter is to explore whether efficient approximations to CCSD can be formulated in a two-stage approach, similar to that employed in methods which first treat static correlation in an active space, followed by dynamical correlations associated with orbitals outside the active space. This is the strategy used in the CASSCF/CASPT2 pair of methods, for example.[42] However, instead of the factorially expensive CASSCF treatment of the active electrons,[103] we will use CCSD itself in the space of all valence orbitals. A subsequent second order correction outside the valence space will be performed using our recently developed similarity-transformed perturbation theory.[48, 5, 49, 3]

The motivation for doing this is both physical and algorithmic. We expect that beyond-valence correlations are weaker and will thus be more satisfactorily treated by perturbation theory than valence correlations. Furthermore, within the valence space, approximations to CCSD can be stronger than in the full space—for example the Perfect Pairing Generalized Valence Bond (PP-GVB) approximation [104, 54, 55, 56] retains only a linear number of pair excitations. Generalizations to restricted pairing [76, 105, 106] retain only a small quadratic number of amplitudes. Thus one can potentially combine the ability of these methods to treat very large active spaces with local MP2-like models for the beyond valence correlation.

This chapter takes the first step towards this general objective by addressing the extent to which an approximation of this type can be faithful to full CCSD. We do this by first obtaining an appropriate set of valence orbitals (in the 1:1 active space where each valence occupied orbital is associated with a correlating anti-bonding orbital, as in perfect pairing). The reference CCSD calculation is performed in this limited valence space, with no local approximations. Subsequently the second order perturbative correction is performed with only single and doubles terms as needed to mimic full CCSD. This hybrid theory is defined in the following section, and its computational cost in the absence of local approximations is discussed. It is then

tested against full CCSD for a range of structural and energetic properties.

3.2 Theory

Our target level of theory is the standard CCSD method, which involves a wave function of the following form:

$$|\Psi_{CC}\rangle = e^{\hat{T}}|\Phi_0\rangle \tag{3.1}$$

$$\hat{T} \equiv \hat{T}_1 + \hat{T}_2 = \sum_{i,a} t_i^a \{\hat{a}_a^\dagger \hat{a}_i\} + \frac{1}{4} \sum_{i,j,a,b} t_{ij}^{ab} \{\hat{a}_a^\dagger \hat{a}_i \hat{a}_b^\dagger \hat{a}_j\}, \tag{3.2}$$

Here $i, j, k, \ldots$ refer to occupied orbitals, and $a, b, c, \ldots$ refer to virtual orbitals. The energy and unknown cluster amplitudes are obtained by minimizing the functional:

$$L_{CCSD} = \langle \Phi_0 | (1 + \hat{\Lambda}) e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle \tag{3.3}$$

where

$$\hat{\Lambda} \equiv \hat{\Lambda}_1 + \hat{\Lambda}_2 = \sum_{i,a} \lambda_a^i \{\hat{a}_i^\dagger \hat{a}_a\} + \frac{1}{4} \sum_{i,j,a,b} \lambda_{ab}^{ij} \{\hat{a}_i^\dagger \hat{a}_a \hat{a}_j^\dagger \hat{a}_b\}, \tag{3.4}$$

by forcing the derivatives of the functional with respect to the λ and t amplitudes to be stationary. Symbolically, this may be written as:

$$\frac{\partial L_{CCSD}}{\partial \hat{\Lambda}} = 0, \qquad \frac{\partial L_{CCSD}}{\partial \hat{T}} = 0. \tag{3.5}$$

In CCSD, the occupied orbitals may be restricted to only valence orbitals (*e.g.* a frozen core approximation), but no restrictions are usually imposed on the virtual orbitals.

We approximate the full solution of the CCSD equations in two stages. The first step is to solve the CCSD problem in just a valence space. In other words, we restrict the t and λ amplitudes such that excitations occur only within an active space of valence orbitals. The second step is the addition of a perturbative second order correction for beyond-valence correlation.

We consider that active spaces appropriate for general chemical applicability are ones which correlate all valence electrons. Two principal alternatives are the full valence (V) active space and the perfect-pairing or (1:1) active space (PP). Disregarding

spatial symmetry, the full valence space defines the active orbitals to be equal to the number of valence atomic orbitals. The PP space (for closed shell molecules) defines the number of active orbitals to be twice the number of occupied valence orbitals, where there is (at least in principle) one correlating orbital for each occupied orbital.

To apply CCSD within a valence active space means that the active orbitals must have been optimized by another method. After all, one should not simply choose the lowest virtual orbitals from a Hartree-Fock calculation to be the active orbitals. As the basis set size grows, these lowest lying virtual orbitals become low-energy, diffuse Rydberg orbitals which cannot properly correlate occupied valence bonding orbitals. Correspondingly, the magnitude of the correlation energy from such a calculation limits toward zero as the basis set approaches completeness—clearly an unphysical result!

Several reasonable choices for valence optimized orbitals are possible. First, one could choose orbitals from a valence optimized orbital CCD (VOD)[46] or QCCD[75] calculation. However these calculations are still relatively expensive. In particular, they are far more expensive than performing CCSD in the valence space presuming the orbitals are available. It is therefore expedient to use a strong local correlation approximation, which will permit (relatively) inexpensive optimization of the valence active space. The local correlation approximation need not be quantitatively accurate; it must simply yield a reasonably faithful set of valence orbitals.

One such possibility is perfect pairing CCD, which yields approximate PP orbitals, based on retaining only a linear number of amplitudes.[76, 57, 58] A related possibility is to use imperfect pairing (IP) CCD, which also determines approximate PP orbitals, but based on retaining a quadratic number of double excitations.[76] Because of the increased flexibility in the IP functional, the orbitals obtained have been found to be closer to the VOD-optimized orbitals, and we therefore adopt it as our source of valence optimized orbitals. We therefore also adopt the PP active space. In summary, VCCSD calculations involve optimization of the PP active space within the IP Ansatz followed by CCSD with no local correlation approximations in this valence space.

As the VCCSD calculation is performed only within the PP valence space, it recovers only a modest fraction of the total correlation energy, perhaps in the vicinity

of 20% to 40% and diminishing with larger basis sets. Valence correlation can be viewed as a definition of static or non-dynamical correlation energy. To approach quantitatively accurate results, it is likely that a correction must be applied subsequent to VCCSD to account for the beyond-valence or dynamical correlations. We apply a perturbative, second-order correction based on our recently introduced similarity transformed perturbation theory[5, 49]

The resulting equations have the following general form:

$$E^{[0]} = E_{VCCSD} \tag{3.6}$$

$$E^{[1]} = 0 \tag{3.7}$$

$$E^{[2]} = \langle 0 | (1 + \Lambda) \bar{H}^{[1]} (E^{[0]} - \bar{H}^{[0]})^{-1} \bar{H}^{[1]} | 0 \rangle. \tag{3.8}$$

This correction contains terms involving single, double, triple, and quadruple excitations, as discussed in detail elsewhere.[5]

Since our goal is to approximate CCSD efficiently, we truncate this correction after the doubles terms, ignoring the computationally demanding triples and quadruples terms which are not present in CCSD. We term this procedure VCCSD(SD). Evaluation of the (SD) correction dominates the computational time. Overall, the method scales as $o^2 v^2 V^2$, where o is the number of active occupied orbitals, v is the number of active virtual orbitals, and V is the total number of virtual orbitals. In separate future work we will include the triples and quadruples terms to assess their ability to directly approximate CCSD(2)[49] (which is equivalent or superior to CCSD(T)[18]) from a VCCSD starting point.

3.3   Results and discussion

The VCCSD(SD) method has been implemented in a developmental version of Q-Chem. [85] A variety of tests have been performed in order to determine the accuracy with which VCCSD(SD) approximates standard CCSD results. We also make some

very preliminary tests of computational efficiency. Each of these is described below. In some cases, it is desirable to compare the results obtained with VCCSD(SD) with those obtained from a method where the active orbitals are fully optimized (and therefore represent the best choice of orbitals for such a method). For this reason, we also introduce VOD(SD), which is equivalent to VOD(2)[5] except that the perturbative correction is truncated as described above for VCCSD(SD).

3.3.1 The VCCSD energy and the choice of active space

The performance of any active space method depends critically upon the choice of the orbitals within that space. Here we address this issue in the context of VCCSD. Figure 3.1 plots the VCCSD and VOD energy (without the perturbative correction) versus the active space size for water in a cc-pVDZ basis.[64, 107] In this example, the water molecule has been symmetrically stretched to $1.75R_e$, thereby increasing the significance of the static correlation. The active space used is of the perfect-pairing (1:1) type until ten orbitals are active. Beyond this point, all occupied orbitals are active, so only the active virtual space grows. With no active orbitals we obtain the Hartree-Fock result, and with all 24 orbitals active we have the CCSD result.

Focusing first on the VOD curve in Figure 3.1, we see a dramatic gain in correlation energy at four active orbitals, corresponding to the inclusion of a correlating virtual for each pair of bonding electrons. Doubling the active space size to eight orbitals (the PP active space) adds in correlating orbitals for the lone pairs and obtains significantly more correlation energy. Although no rigorous partitioning of the static and dynamical correlation exists, this chemically-motivated active space translates into a suitable choice for obtaining a sizable fraction of the correlation energy roughly corresponding to the static correlation at minimal cost. Alternatively, once might consider something like a 'double-PP' active space with 16 active orbitals instead of eight, thereby providing three correlating orbitals to each occupied. In this case, a double-PP active space would recover almost all of the CCSD correlation energy. However, exploration of such extended active spaces will be the subject of future studies. For now, the remaining (dynamical) correlation will be obtained with

Figure 3.1: Energies as a function of active space size for H_2O, symmetrically stretched to $1.75R_e$, in a cc-pVDZ basis as calculated with VCCSD using IP orbitals, VCCSD using canonical orbitals, and VOD. The active space contains equal numbers of occupied and virtual orbitals until all five occupied orbitals are active, after which only virtual orbitals are added to the space. The vertical dashed line marks the PP active space used in a standard VCCSD calculation.
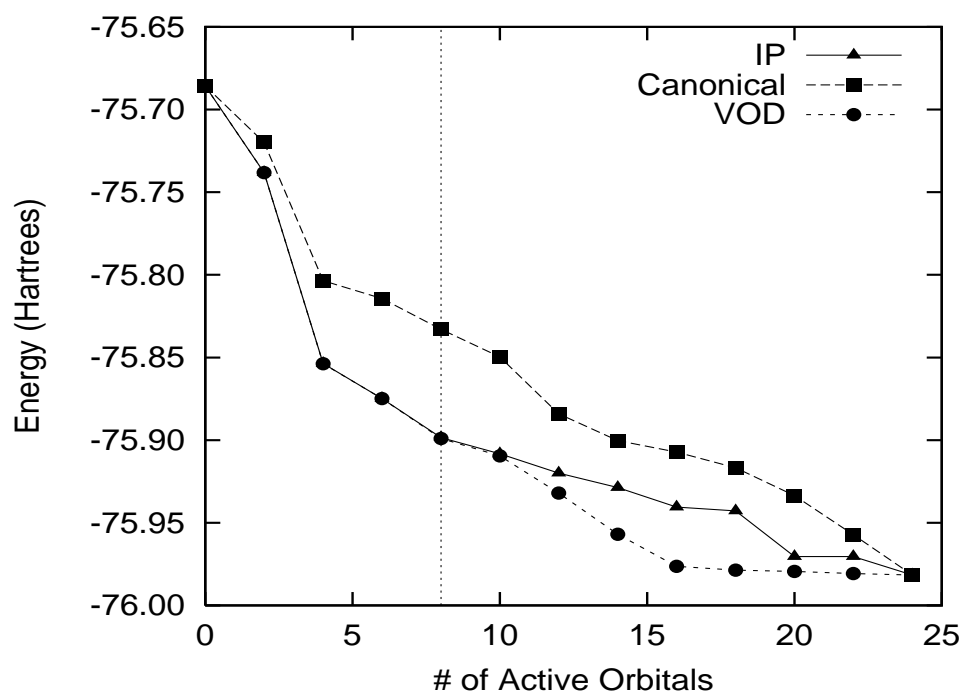
perturbation theory.

VCCSD with the IP-optimized orbitals behaves qualitatively the same as VOD up through ten active orbitals (i.e. in the regime in which IP is well-defined). Beyond ten active orbitals, no more occupied orbitals remain to pair with additional virtual orbitals, so subsequent active virtuals have little chemical significance in the IP model. Correspondingly, the rate of correlation energy recovery drops dramatically. Thus, the approximately-optimized orbitals (from an IP calculation) behave similarly to the fully optimized orbitals within the PP active space. As for VOD, this choice of active space balances efficiency and recovery of the correlation energy. In contrast, if we use the canonical orbitals instead, the active space must be roughly twice as large to recover the same portion of the correlation obtained with IP orbitals in the PP space, underscoring the importance of an optimized active space.

3.3.2 Potential energy surfaces and spectroscopic constants of diatomic molecules

Next we focus on the ability of VCCSD and VCCSD(SD) to reproduce CCSD potential energy surfaces and properties. Consider, for example, the surface produced by stretching the F_2 bond, as shown in Figure 3.2. Qualitatively, VCCSD(SD) reproduces the shape of the CCSD potential, though the absolute energy is somewhat in error. However, this absolute error in energy is, for the most part, unimportant in observable relative energies, as demonstrated below. First, we focus on the ability of VCCSD and VCCSD(SD) to correctly mimic the CCSD potentials in calculating the spectroscopic constants R_e, ω_e, B_e, D_e, α_e, and $\omega_e x_e$ for the singlet ground-state first- and second-row diatomic molecules.

Five single-point calculations about the experimental R_e values[81] spaced 10 picometers apart were used for the fit, which was performed with PSI.[86] In cases where the points were not well-centered about the experimental minimum, a shifted set of points was used to obtain better results. Unfortunately, the target cc-pCVTZ basis[64, 108] is unavailable for many of the elements in this study. Therefore, the cc-pV(T+d)Z basis[109] and the 6-311G(2df,2pd) basis[110, 111] were used as well.

Figure 3.2: The F_2 potential produced by stretching the F-F bond in a cc-pVTZ basis.
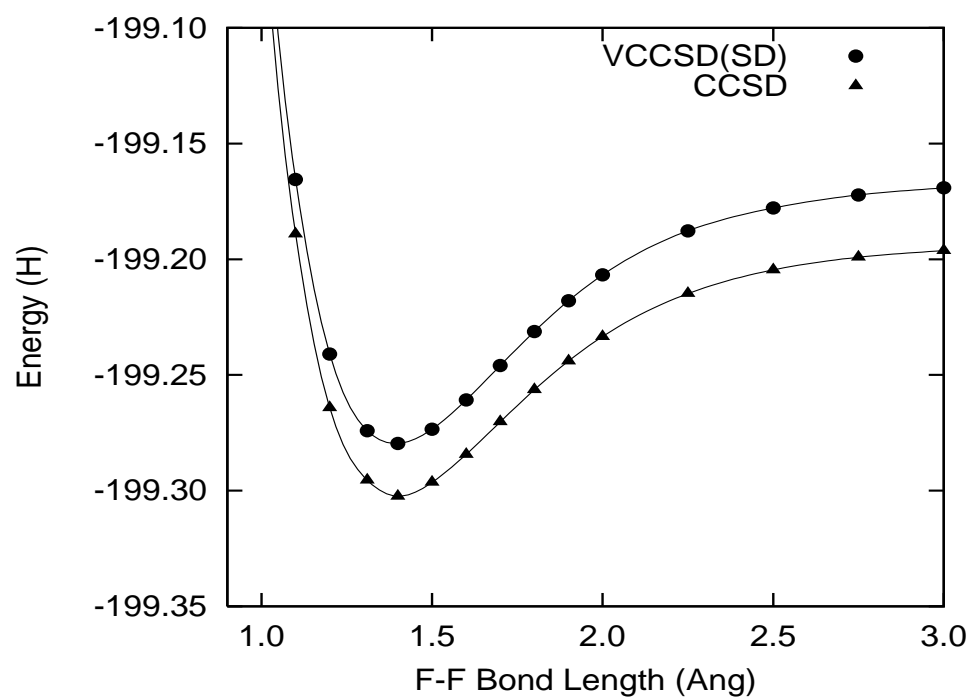
Table 3.1: Root-mean-square errors with respect to CCSD for VCCSD(SD) and MP2 for first- and second-row diatomic spectroscopic constants. R_e is in Å, while the other properties are in cm^{-1}.

	R_e	ω_e	B_e	D_e	α_e	$\omega_e x_e$
MP2[a]	0.0083	89.2	0.077	13.2×10^{-6}	0.0139	2.55
VCCSD[a]	0.0086	43.1	0.143	10.4×10^{-6}	0.0067	2.83
VCCSD(SD)[a]	0.0040	21.8	0.020	2.89×10^{-6}	0.0032	0.72
MP2[b]	0.0145	71.1	0.019	2.59×10^{-6}	0.0034	1.02
VCCSD[b]	0.0193	37.2	0.0830	4.53×10^{-6}	0.0076	2.81
VCCSD(SD)[b]	0.0085	13.7	0.005	1.49×10^{-6}	0.0032	1.17
MP2[c]	0.0242	107.9	0.0321	6.64×10^{-6}	0.0059	2.38
VCCSD[c]	0.0425	51.9	0.1379	6.67×10^{-6}	0.0032	9.61
VCCSD(SD)[c]	0.0175	22.9	0.0190	2.12×10^{-6}	0.0022	2.62

[a] Includes: BF, BH, C$_2$, CO, F$_2$, HF, N$_2$, NF, and NH, in the cc-pCVTZ basis.

[b] Includes: AlCl, AlF, AlH, Cl$_2$, CS, FCl, HCl, NP, P$_2$, SiO, and SiS, in the cc-pV(T+d)Z basis (Al-Cl) and the cc-pCVTZ basis (H-F).

[c] Includes: BeO, BeS, Li$_2$, LiH, MgO, MgS, Na$_2$, NaF, NaH in the 6-311G(2df,2pd) basis.

For the purpose of calculating statistics, the diatomics are divided into three groups based on the basis sets used, as explained in Table 3.1. The first group of diatomics contains BF, BH, C$_2$, CO, F$_2$, HF, N$_2$, NF, and NH, the second group contains AlCl, AlF, AlH, Cl$_2$, CS, FCl, HCl, NP, P$_2$, SiO, and SiS, and the third group contains BeO, BeS, Li$_2$, LiH, MgO, MgS, Na$_2$, NaF, and NaH. We compare VCCSD, VCCSD(SD), and MP2 results against CCSD results, since our aim is to test the faithfulness of VCCSD(SD) to CCSD rather than to concern ourselves with choosing appropriate basis sets for predicting experimental results. Mg$_2$ is excluded from the study because it was found that each theory predicted widely different bond lengths ranging from 4.5 to 7.5Å, none of which were near the experimental value of 3.89Å. Further study of this molecule is beyond the scope of this chapter.

Table 3.1 contains the root-mean-square (RMS) errors for MP2, VCCSD, and VCCSD(SD) as compared to the CCSD results for each group of diatomic molecules and each spectroscopic constant. For almost every constant, VCCSD(SD) results are a factor of two to four better overall than their corresponding MP2 results. Note

particularly the success of VCCSD(SD) in predicting harmonic frequencies, rotational constants, and centrifugal distortion constants as compared to MP2. The anharmonic and coupling constants prove slightly more difficult for VCCSD(SD). The second and third groups of anharmonic constants are the only categories in which MP2 has a lower RMS error than VCCSD(SD), primarily due to AlH for which VCCSD(SD) is 3.4 cm^{-1} (15.5%) off and Li$_2$, for which it is 6.1 cm^{-1} in error. However, the other VCCSD(SD)-calculated properties of AlH are in fairly good agreement with CCSD, so this case is not too significant. Moreover, the error in Li$_2$ is large but qualitatively correct. In contrast, MP2 predicts the wrong sign for this constant. The largest errors for VCCSD(SD) occur in the third group, which contains molecules with alkali and alkaline earth metals. This is not surprising, since the PP active spaces for these elements are quite small. Nevertheless, VCCSD(SD) generally outperforms MP2 in even these cases.

One molecule, AlCl, deserves closer examination. VCCSD(SD) has particular difficulty with AlCl, as is shown by the results in Table 3.2. Due to the rather large shift in R_e, the fit for AlCl was performed on points every ten picometers from 2.100113Å to 2.140113Å, rather than the standard five single points centered about the equilibrium bond length (2.130113Å). In any case, VCCSD(SD) does not correctly reproduce the potential surface. For comparison, VOD(SD) was run to determine whether the problem stems from the partitioning of the correlation energy or whether it is due to the difference between IP orbitals and fully optimized orbitals (as in VOD). We see that while VOD(SD) does fix the problem somewhat, especially in ω_e and $\omega_e x_e$, the other properties still differ substantially from CCSD. Apparently, AlCl does not lend itself well to a partitioned description of its correlation energy, at least in the perfect-pairing active space.

In contrast to the fairly systematic improvement of the results going from MP2 to VCCSD(SD), the uncorrected VCCSD behaves quite inconsistently. Although in a few cases it performs reasonably well, it generally fluctuates significantly, leading to rather large RMS errors as shown in Table 3.1. For example, VCCSD improves upon the MP2 description of harmonic frequencies, but it errs grossly in the rotational constants. The anharmonic constants are also poor in most cases. The valence

Table 3.2: Predicted spectroscopic constants for AlCl in the cc-pV(T+d)Z basis. R_e is in Å, while the other properties are in cm^{-1}.

	R_e	ω_e	B_e	D_e	α_e	$\omega_e x_e$
MP2	2.136411	489.9	0.24251	0.238×10^{-6}	0.00150	1.87
VCCSD(SD)	2.113053	517.5	0.24790	0.228×10^{-6}	0.00193	0.78
VOD(SD)	2.115792	492.7	0.24726	0.249×10^{-6}	0.00154	1.86
CCSD	2.137953	490.1	0.24216	0.236×10^{-6}	0.00149	1.99
Expt.[a]	2.130113	481.3	0.24393	0.250×10^{-6}	0.001611	1.95

[a] Values taken from Ref. [81]

correlation alone is insufficient to quantitatively describe potential energy surfaces and their associated spectroscopic constants (although it is of course better than no correlation at all).

We also present Tables 3.3 and 3.4 containing the actual computed values of ω_e and $\omega_e x_e$, which are representative of the best and worst cases for VCCSD(SD). For harmonic frequencies, the largest VCCSD(SD) errors are roughly 50 cm^{-1} for C_2 and MgO. In contrast, MP2 has errors of over 100 cm^{-1} or more for several molecules, including CO, N_2, NP, and MgO. In the last case, the VCCSD(SD) error is one fifth the size of the MP2 one. For uncorrected VCCSD, the errors are smaller than those for MP2, but larger errors on the order of 50 cm^{-1} or more are quite common, including most of the hydrides, N_2, P_2, and MgS. While MP2 and VCCSD(SD) tend to give fairly good results for a range of molecules with intermittent exceptions, VCCSD gives mediocre or poor results for a large number of cases. Looking at the anharmonic constants in Table 3.4, we again see that for most molecules the methods are fairly good, though each has its exceptions. Notably, VCCSD predicts the wrong sign of $\omega_e x_e$ for AlCl, MgS, and Na_2. Actually, in the case of Na_2, CCSD, MP2, and VCCSD all predict a negative $\omega_e x_e$. The experimental value, however, is 0.725 cm^{-1}. Only VCCSD(SD) gets the appropriate sign, even if it is about twice as large as the experimental value. The largest MP2 error is MgO, at 4.0 cm^{-1}, while the largest VCCSD error is 26.5 cm^{-1} for MgS, and the largest VCCSD(SD) error is 6.1 cm^{-1} for Li_2. Overall, while the difference between VCCSD(SD) and MP2 is less clear here, VCCSD(SD) performs as well as or slightly better than MP2 on the whole.

Table 3.3: Calculated harmonic frequencies ω_e using MP2, VCCSD, VCCSD(SD), and CCSD. CCSD value is the calculated one, while other columns are errors relative to CCSD. All values are in cm^{-1}. The horizontal lines mark the separations between the different groups of molecules described in the text. Basis sets are described in Table I.

	MP2	VCCSD	VCCSD(SD)	CCSD
BF	-3.47	-25.91	15.00	1427.36
BH	67.46	-90.10	-6.56	2367.86
C_2	-2.83	-9.62	46.54	1891.93
CO	-101.77	-6.38	16.79	2233.30
F2	0.51	-57.95	13.10	1020.64
HF	-34.26	-14.04	6.12	4214.75
N_2	-226.73	-37.73	6.62	2434.33
NF	22.01	20.26	34.50	1243.21
NH	60.36	-49.40	-11.30	3357.84
AlCl	-0.22	-8.38	27.37	490.10
AlF	-13.61	-9.73	0.56	801.94
AlH	28.61	-69.63	-15.67	1672.66
Cl_2	13.88	-63.20	18.13	572.21
CS	-27.71	-8.14	4.42	1359.65
FCl	-3.25	-1.54	-24.43	813.55
HCl	32.39	-41.28	-6.48	3027.38
NP	-196.82	-35.27	3.86	1400.05
P_2	-78.26	-53.37	6.68	821.09
SiO	-86.14	-5.37	5.82	1290.01
SiS	-19.18	-18.74	-0.94	778.90
BeO	-146.68	-20.05	9.70	1583.17
BeS	-18.01	-23.44	7.64	1020.36
Li_2	-12.13	0.67	-3.81	353.09
LiH	30.07	-112.09	-17.38	1416.22
MgO	282.72	8.77	55.46	746.49
MgS	30.78	-47.99	29.37	527.46
Na_2	-2.23	-1.20	8.55	159.28
NaF	-3.02	-9.95	-2.74	578.41
NaH	30.37	-90.84	-15.05	1158.69

In summary, VCCSD(SD) successfully reproduces the full CCSD spectroscopic properties of first- and second-row diatomic molecules. VCCSD(SD) is clearly superior to MP2 for bond lengths, harmonic frequencies, rotational constants, and centrifugal distortion constants, and as good or better for the anharmonic and vibration-rotation coupling constants.

3.3.3   Relative energies

To further test the faithfulness of VCCSD(SD) to standard CCSD, reaction energies for a series of 13 isogyric reactions[112] were calculated with MP2, VCCSD, VCCSD(SD), VOD(SD), and CCSD using accurate equilibrium geometries,[63] the results of which are shown in Table 3.5. Once again, we see a substantial improvement in accuracy relative to CCSD in going from MP2 to VCCSD(SD) and VOD(SD) in terms of the mean absolute error and the root-mean-square deviations. In fact, VOD(SD) and VCCSD(SD) behave rather similarly, though VOD(SD) does slightly better on the whole, as would be expected. More importantly, VCCSD without the perturbative correction is clearly unable to accurately reproduce the relative energies of these reactions. The first reaction energy even has the wrong sign. By itself, the valence space calculation is an insufficient model for the correlation energy. For four of the thirteen reactions, MP2 actually gets closer to the CCSD results than VCCSD(SD), but in most cases the difference between the two methods is rather small. Overall, VCCSD(SD) seems to satisfactorily reproduce CCSD reaction energies.

A more sensitive test of the accuracy of any method is its ability to accurately predict small relative energies such as torsional barriers. As an example, we focus on the torsional barrier of ethane rotating from the *anti* to *gauche* configurations. Table 3.6 compares the results for the torsional barrier of ethane for MP2, VCCSD(SD), VOD(SD), and CCSD with multiple basis sets. Once again, VCCSD(SD) (and VOD(SD)) performs much better than MP2 relative to the CCSD results, with errors of roughly half a percent versus 2-3% for MP2. Although in this case the absolute errors are rather small, the faithfulness of VCCSD(SD) to CCSD is encouraging. In

Table 3.4: Calculated anharmonic constants $\omega_e x_e$ using various methods. CCSD value is the calculated one, while other columns are errors relative to CCSD. All values are in cm^{-1}. The horizontal lines mark the separations between the different groups of molecules described in the text. Basis sets are described in Table I.

	MP2	VCCSD	VCCSD(SD)	CCSD
BF	0.14	-0.88	0.30	11.12
BH	-2.17	-6.67	0.66	48.30
C_2	-1.14	-0.97	-0.27	12.42
CO	0.45	-0.96	0.81	12.11
F_2	-0.58	2.35	-0.91	9.81
HF	-3.80	-3.74	1.23	89.85
N_2	4.45	-0.38	-0.09	13.27
NF	-0.13	-0.20	0.75	9.30
NH	-4.19	-2.31	0.67	75.54
AlCl	-0.12	-3.11	-1.21	1.99
AlF	-0.44	-0.27	0.16	4.06
AlH	0.25	8.57	3.37	21.64
Cl_2	0.05	0.14	-0.14	2.33
CS	-0.20	0.20	0.13	6.02
FCl	-0.25	0.73	-0.51	4.56
HCl	-1.45	-0.03	-0.20	52.23
NP	2.85	1.68	1.24	6.20
P_2	0.86	-0.25	-0.19	2.54
SiO	0.06	-0.32	0.08	5.35
SiS	0.02	0.24	0.51	2.28
BeO	1.69	-0.89	0.26	11.07
BeS	-1.98	-0.96	-0.53	6.44
Li_2	-3.21	0.91	6.10	1.79
LiH	-2.37	-3.62	-2.03	25.14
MgO	4.05	3.35	1.36	0.66
MgS	-1.49	-26.49	1.48	2.76
Na_2	-0.52	-1.61	2.55	-0.69
NaF	-0.14	-0.44	0.06	3.36
NaH	-3.06	-10.00	3.10	14.66

Table 3.5: Reaction energies for isogyric reactions[a] in the aug-cc-pVDZ basis. Errors are relative to CCSD results. Energies are in kcal/mol.

Reaction	MP2	VCCSD	VCCSD(SD)	VOD(SD)	CCSD
$CO+H_2 \rightarrow CH_2O$	-3.1	1.6	-3.8	-3.3	-2.4
$N_2+3H_2 \rightarrow 2NH_3$	-39.2	-26.3	-40.4	-40.8	-40.3
$C_2H_2+H_2 \rightarrow C_2H_4$	-50.9	-40.2	-51.0	-51.4	-53.8
$CO_2+4H_2 \rightarrow CH_4+2H_2O$	-60.7	-51.6	-65.4	-64.1	-66.2
$CH_2O+2H_2 \rightarrow CH_4+H_2O$	-63.9	-56.1	-63.9	-63.7	-63.7
$CO+3H_2 \rightarrow CH_4+H_2O$	-67.0	-54.5	-67.7	-67.0	-66.1
$HCN+3H_2 \rightarrow CH_4+NH_3$	-76.5	-60.8	-77.2	-77.2	-79.3
$H_2O_2+H_2 \rightarrow 2H_2O$	-137.7	-142.8	-139.4	-139.2	-136.0
$HNO+2H_2 \rightarrow H_2O+NH_3$	-181.8	-178.8	-180.9	-180.3	-178.3
$C_2H_2+3H_2 \rightarrow 2CH_4$	-110.6	-91.6	-109.3	-109.3	-113.6
$CH_2(^1A_1)+H_2 \rightarrow CH_4$	-133.4	-118.9	-126.8	-126.5	-125.2
$F_2+H_2 \rightarrow 2HF$	-143.6	-148.9	-141.7	-141.5	-138.3
$2CH_2(^1A_1) \rightarrow C_2H_4$	-207.2	-186.4	-195.3	-195.1	-190.6
Mean Absolute Error	4.0	10.3	2.2	2.1	
RMS Error	5.9	11.9	2.6	2.5	

[a] Reactions taken from Ref. [112] and the geometries from Ref. [63]

contrast, VCCSD seems to perform rather erratically, sometimes doing very well and sometimes behaving more like MP2, demonstrating once again the importance of the perturbative correction.

3.3.4 Benzene and symmetry-breaking

Using a strong local correlation approximation for the valence correlation energy is known to sometimes introduce artifacts in computed potential energy surfaces. One classic example is that both IP and PP-GVB predict broken symmetry in benzene—that is, they predict alternating longer and shorter bonds as a stable structure as a result of the localization procedure and the restrictions placed on the cluster amplitudes.[76] This deformation can be represented as a deviation of the bond angles from 60 degrees to alternating angles of $60 + \phi$ and $60 - \phi$. What impact if any does the use of these IP orbitals have on VCCSD(SD)?

The results of such calculations are shown in Figure 3.3, where the energy of

Table 3.6: Torsional barrier of ethane under *anti* to *gauche* rotation with increasing basis set size. Energies are in kcal/mol. Percent errors relative to CCSD are in parentheses.

	cc-pVDZ[a]		aug-cc-pVDZ[a]		cc-pVTZ[a]	
MP2	3.74	(3.1)	3.57	(2.2)	3.45	(2.0)
VCCSD	3.60	(-0.7)	3.59	(1.0)	3.37	(-0.2)
VCCSD(SD)	3.61	(-0.6)	3.49	(-0.2)	3.36	(-0.5)
VOD(SD)	3.60	(-0.9)	3.47	(-0.5)	3.35	(-0.9)
CCSD	3.63		3.49		3.38	

[a] Ref. [64]

benzene is calculated versus the deformation angle ϕ. The VCCSD(SD) results for the deformation of benzene correctly restore symmetry to benzene. This is because excitations from one localized bonding orbital are allowed to occur to localized anti-bonding orbitals on other sites in the VCCSD calculation, thereby returning to the picture of delocalized bonding. Thus we see no serious consequences from using the IP orbitals for VCCSD(SD) in this case where the IP potential surface itself is incorrect.

3.3.5 The Cope rearrangement

Having demonstrated the performance of VCCSD(SD) on simple systems, we shift our attention to a more difficult example: the Cope Rearrangement. This reaction has been the focus of many studies at different theoretical levels including semi-empirical, DFT, and multi-reference methods.[113, 114, 115, 116, 117] Unfortunately, the results of many of these studies contradicted one another both qualitatively and quantitatively, and only recently has the controversy mostly been resolved.[118] The difficulty arises because of the very flat energy landscape along a C_{2h} cut in the potential energy surface that connects a diyl intermediate at $R = 1.64$ Å and an aromatic transition state at $R = 2.19$ Å.[114, 116] One way of exploring the success of various levels of theory on this reaction is to follow the C_{2h} cut of the surface between these two structures. Both RHF and UHF are inadequate since RHF cannot properly describe the diradicaloid species involved, and UHF overestimates the stability of such species.[118] Studies also suggest that correlated methods based on HF such as CCSD

Figure 3.3: Deformation of benzene and symmetry breaking. Calculations performed in a 6-31G* basis [1, 2] and with radial distances fixed at the MP2/6-31G* level. The symmetry-breaking angle is described in the text.
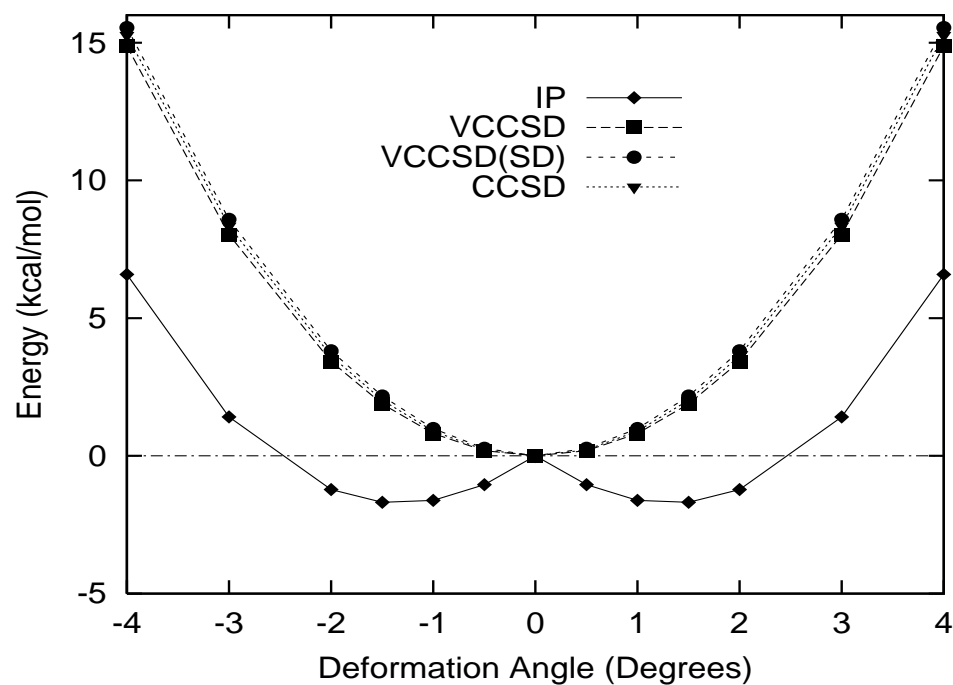
converge slowly with classes of excitation included.[119] Therefore, multi-reference methods have become standard in most of the accepted studies. As shown in the works cited above, a six-electron, six-orbital CASSCF calculation incorrectly places a small barrier between these two structures along the C_{2h} cut. The active space used includes the two π bonds and one σ bond explicitly involved in the rearrangement. However, they have shown that the addition of dynamical correlation through multi-reference perturbation theory eliminates this barrier, and predicts the true transition state for the Cope rearrangement (the minimum on this C_{2h} cut of the surface) to occur near 1.85 Å with a transition state barrier height of roughly 31 kcal/mol above hexadiene (the diyl and aromatic species lie at 33 and 35 kcal/mol, respectively).[116] This drastic change in behavior has been attributed to the dynamical correction allowing the sigma-bonding framework to relax during the reaction.[119] In a sense, this can be thought of as a triple excitation involving two active π electrons and an inactive σ one.

In this chapter, we apply VCCSD and VCCSD(SD) to this reaction and compare against CCSD and the multi-reference methods. All calculations are performed at CASSCF(6,6)-optimized geometries using the 6-31G* basis at 11 geometries with the interallylic distance ranging from 1.64 to 2.19 Å along the C_{2h} potential energy surface. In each case, restricted orbitals are used. VCCSD(SD) and related methods all use the PP active space, which contains 34 electrons in 34 orbitals for this system.

Figure 3.4 plots the C_{2h} potential energy surface as generated using CASSCF(6,6) and MRMP2, IP, VCCSD, and VCCSD(SD). In the figure, the MRMP2 minimum and end points were estimated from Ref. [116] and a smooth function interpolated between. The energies given are relative to the energy of hexadiene for each method. As discussed previously, CASSCF(6,6) places an unphysical barrier between the two intermediates. Not too surprisingly, IP performs poorly on this surface. We see two distinct states in this region: one roughly corresponding to the diyl and the other to the aromatic intermediate. Presumably, the occurrence of the second, lower-energy state in the aromatic region is a problem due to broken symmetry in the orbitals for the aromatic species, just as in benzene. The interesting question becomes, can VCCSD overcome this broken symmetry as in the benzene example? If we simply

Figure 3.4: C_{2h} cut along the Cope Rearrangement potential energy surface. The CASSCF(6,6) and MRMP2 results are compared against IP, VCCSD, VCCSD(SD), and CCSD in the 6-31G* basis using the CASSCF optimized geometries. Except for CASSCF, all the active space methods use the PP active space. The bond length R is the distance (in Å) between the two allyl groups in the chair structure, and energies are relative to hexadiene in kcal/mol. At short R is the diyl intermediate, and at long R is the aromatic species. The MRMP2 data were estimated from Ref. 40.
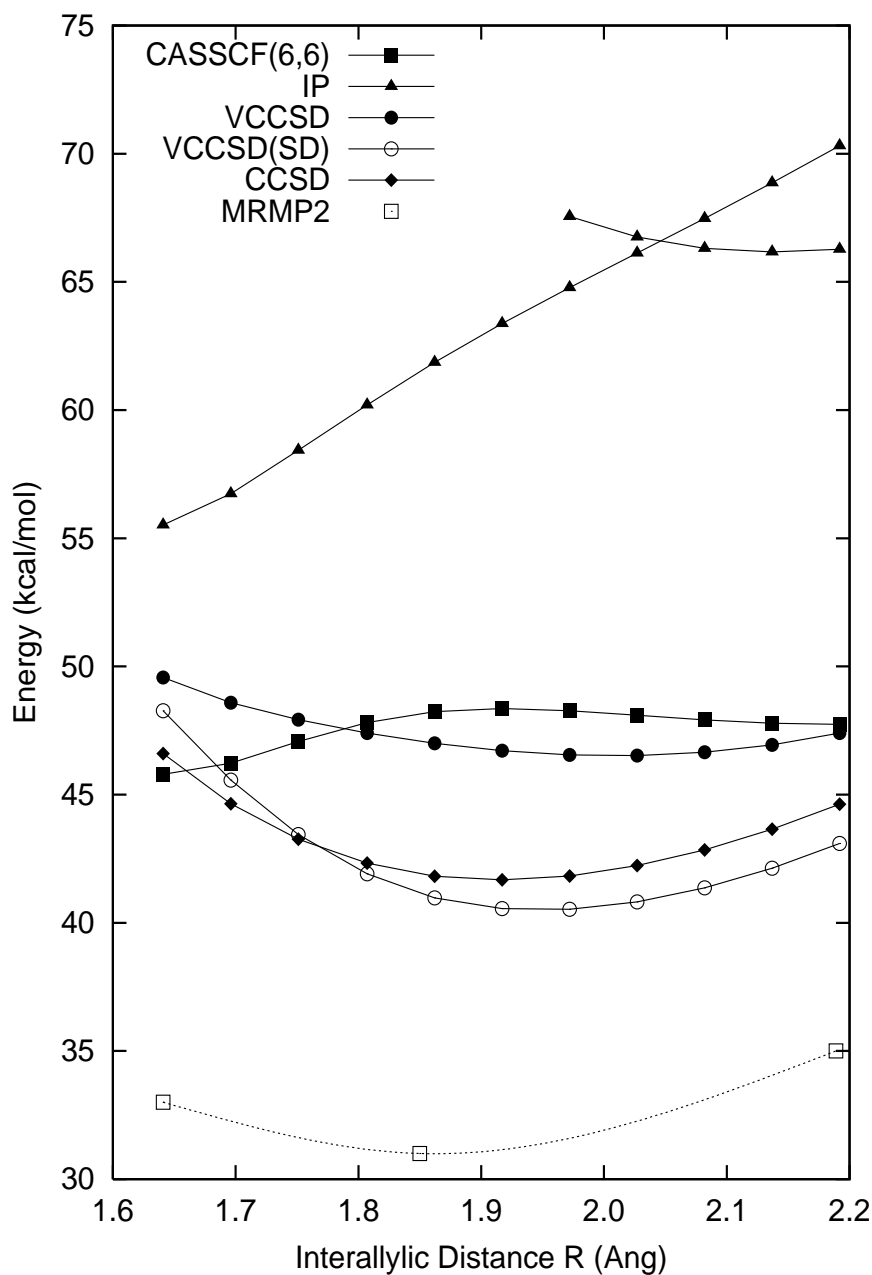
follow the lowest state at all times, the answer is no. VCCSD as calculated from these orbitals, though smoother than the IP surface, still demonstrates a kink where the orbitals break symmetry. In contrast, if we follow the symmetric (diyl) solution, VCCSD produces the smooth curve shown in Figure 3.4. This curve is qualitatively similar to the MRMP2 surface,[116] except that the minimum on these two curves is rather long (occurring at just over 2 Å instead of 1.85 Å), and the relative energies of the endpoint intermediates are flipped (though the actual differences are only a few kcal/mol—well within the model errors). Despite the failure of the IP model, the orbitals so-obtained are sufficient for VCCSD to obtain approximately the correct shape and relative energies of the intermediates as compared to the minimum as long as one remains on the symmetry-preserving IP surface. Unfortunately, the energy of these states relative to hexadiene is much too high; VCCSD places the reaction transition state at roughly 46 kcal/mol (the minimum of the C_{2h} surface), in sharp disagreement with the approximately 31 kcal/mol reported by Kozlowski *et al.*[116]

Presumably, a better description could be obtained by including dynamical correlation. In the present context, we examine the effects of a partial treatment of dynamical correlation in the form of VCCSD(SD), the results of which are plotted in Figure 3.4. Indeed, we see that the transition state barrier is lowered to about 41 kcal/mol (still high), but the aromatic species is much lower than the diyl one. Moreover, although the minimum moves to shorter interallylic distances, it is still longer than MRMP2. Although the perturbative singles and doubles terms are largely sufficient to describe the aromatic species, the more difficult diyl species presumably requires higher order excitations in order to be accurately described. These results agree fairly well with those from CCSD, though CCSD shows a smaller difference between the two endpoints and a slightly shorter transition state. Thus, although the VCCSD(SD) surface is lacking as compared to MRMP2, we see that it is reasonably similar to CCSD, even in this difficult case.

Once again we have demonstrated that VCCSD(SD) is rather faithful to CCSD, even in a very sensitive system. Unfortunately, in this example, both methods are lacking compared to higher level calculations. Intriguingly, we notice that in many senses, the pure valence space contribution of VCCSD gives a better description of

the surface than the full VCCSD(SD). This behavior could arise from either of two key effects: the increased size of the active space relative to the CASSCF calculations, or the incomplete treatment of dynamical correlation in the truncated perturbative correction (SD). We hope to report on a more extensive analysis of these issues in the near future.

3.3.6 Timings

As discussed in the Introduction, the purpose of this chapter is to assess the capability of a valence-based method to approximate CCSD. Future work will more fully exploit the potential efficiency of this approach. For the moment, we content ourselves with a simple preliminary comparison of the efficiency of such a method against standard methods. VCCSD(SD) and CCSD calculations on a growing linear alkane chain in the cc-pVDZ basis were performed and the timings noted, as is plotted in Figure 3.5. In all molecules in this set, the VCCSD(SD) calculation (including the time required to obtain the IP orbitals) is faster than the full CCSD calculation. The savings become more pronounced at larger system sizes, as expected. By comparing the timings of VCCSD(SD) with and without the IP calculation, one can see that obtaining the IP orbitals does not correspond to too large a fraction of of the total computational time, which was one of the major goals of this study. Of course, the cost of an MP2 calculation is still far below that of the VCCSD(SD) calculation. The next step, which we hope to report on in the future, is to apply local correlation approximations to both the valence part of the problem and the second order correction to generate a method well-suited to treating large systems.

3.4 Conclusions

This study has explored the validity of partitioning the correlation energy into static and dynamical contributions to approximate a correlation method (CCSD) which makes no such distinction. Such an approximation has the potential to be significantly more efficient if it is sufficiently accurate. Our model uses CCSD in

Figure 3.5: Comparison of timings on growing linear alkane chains for various methods in the cc-pVDZ basis. In this figure, IP-VCCSD(SD) refers to the time of both the IP calculation and the VCCSD(SD) calculation, while VCCSD(SD) includes only the actual VCCSD(SD) steps.
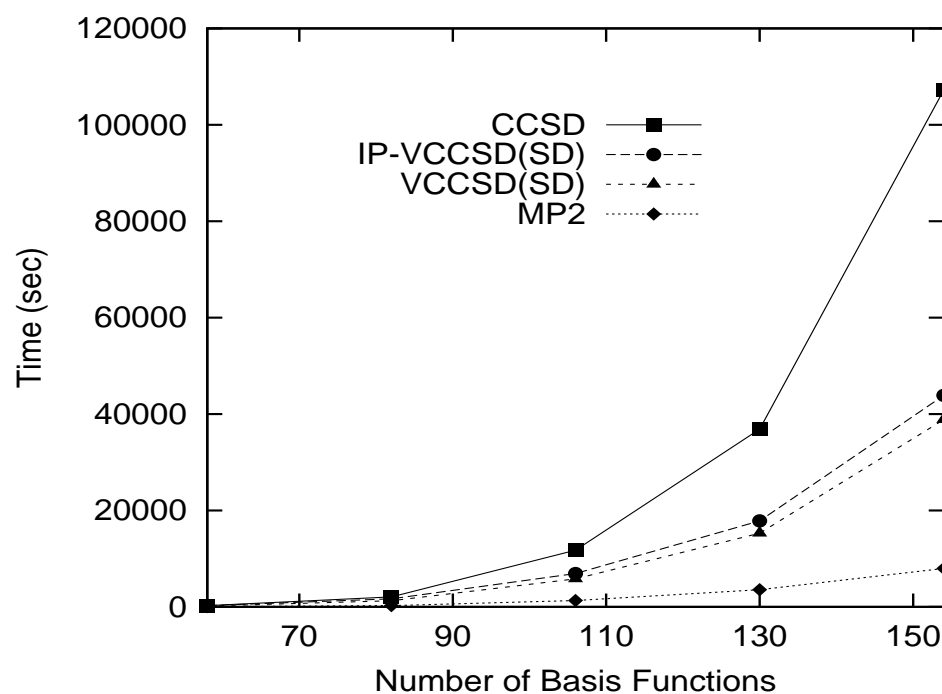
the perfect pairing active space of all valence electrons, plus a truncated perturbative (2) correction for non-valence single and double substitutions. If appropriately optimized orbitals are employed, this VCCSD(SD) model approximates full CCSD fairly satisfactorily for spectroscopic constants, structural properties, relative energies, and in the Cope Rearrangement. Residual deviations are likely to be largely the result of neglected effects of coupling between the correlation effects in the valence and non-valence space. Having demonstrated that such an approach can be successful, future work will be directed towards efficient local correlation treatments based on this framework. With this in mind, in the next chapter we will explore the localizability of the valence space correlation.

Chapter 4

The localizability of valence space electron-electron correlations in pair-based coupled cluster models

4.1 Introduction

The vast majority of wave-function-based electronic structure calculations begin with the Hartree-Fock (HF) approximation and correct for electron-electron correlations using perturbative or coupled cluster methods. Unfortunately, in cases such as transition states, bond-breaking, radicals, diradicals, and other highly-correlated systems, HF often fails to provide an even qualitatively correct potential energy surface (PES). For this reason, Complete Active Space Self-Consistent Field (CASSCF),[41] which includes all possible excitations within an active space is the method of choice for studying diverse regions of a PES, such as when following a reaction coordinate.

Unfortunately, the factorial computational scaling of CASSCF means that the active space must contain fewer than roughly twenty active orbitals and electrons, making it applicable only to either small molecules or very small active spaces in larger systems. In the latter case, one wonders whether the active space is sufficiently large to correctly reproduce the qualitative behavior of the potential energy surface. One might try to use more selective multi-configurational approaches to treat these larger

systems, but then the choice of effective active space becomes even more arbitrary. Clearly, more efficient active space methods to treat such systems are desirable.

Furthermore, efficient active space methods, together with beyond-active space corrections, may provide a means for obtaining highly accurate results at reduced computational expense. Such a possibility was explored with moderate success in the non-local VCCSD(SD) approach in Chapter 3, which used a CCSD calculation in the valence space and a perturbative treatment of the beyond-valence singles and doubles to approximate a full CCSD calculation. For example, one would like to predict accurate molecular geometries at low computational cost. Non-local active space coupled-cluster methods have shown promise in their predicted structures[120], and local models could drastically expand the range of applicability of high-level methods to structural prediction.

Non-local active space coupled cluster methods, such as Valence Orbital Optimized Coupled Cluster Doubles (VOO-CCD or simply VOD)[46, 47] and Valence Orbital-Optimized Quadratic Coupled Cluster Doubles (VQCCD)[75, 120], were developed as robust, less expensive replacements for CASSCF. These methods, often in concert with a second-order perturbative correction,[5, 48, 49, 3] have proven themselves to be much more reliable than more standard approaches across a range of systems.[120, 5, 3].

Though promising and far less computationally expensive than CASSCF, these non-local active space coupled cluster doubles methods nevertheless demonstrate a steep N^6 scaling with system size, making them practical only in the region of up to roughly 50 active orbitals. In contrast, the extremely affordable Generalized Valence Bond Perfect Pairing (GVB-PP) approximation can be calculated extremely efficiently for hundreds of active orbitals, and it can provide a qualitatively correct description for many problems in which static correlations are strong.[60] It also provides fairly good geometries as compared with both experiment and VQCCD[121, 120]. GVB-PP can be rewritten as a local coupled cluster doubles problem within an active space in which each pair of bonding electrons correlates with a single, spatially localized anti-bonding orbital, and all such pairs of orbitals are mutually orthogonal[57, 58]. These correlated electron pairs then interact in a mean field way

with each other. We classify this Ansatz as a one-center model, since electrons correlate solely in the spatial region from which they originated.

How does one progress beyond the GVB-PP model in order to obtain an accurate potential energy surface efficiently? The most obvious way is of course to relax the strong-orthogonality or perfect-pairing constraints in the GVB-PP wave function–i.e. compute the full GVB wave function.[122] Unfortunately, such expansions become computationally expensive and difficult to converge. Another possible approach is to use the GVB-PP natural orbitals in a configuration interaction (CI) expansion, such as GVB-RCI and related approaches.[123, 124, 125, 126, 127, 128] Like CASSCF, Complete RCI expansions also exhibit factorial computational cost, so in practice they are truncated. Typical truncated GVB-RCI expansions include open-shell singlet and triplet configurations formed within each natural orbital pair in addition to the GVB-PP configurations. These compact wave functions can accurately describe bond dissociation and other highly-correlated systems. Unfortunately, these wave functions, like other truncated CI wave functions, are generally not size consistent. Alternatively, perturbative corrections based on generalizations of Møller-Plesset perturbation theory to multi-reference wave functions have also been used with success to correct GVB-PP wave functions.[43, 44, 129, 130, 131] Such treatments typically estimate both the remaining valence and beyond-valence correlations excluded in the GVB-PP treatment.

Taking a different perspective, we note that a VOD wave function can be rewritten in terms of orthogonal, localized orbitals with no approximation. Including all possible double excitations in the cluster operator requires double excitations involving up to four spatial centers. Given the success of the one-center GVB-PP model for many problems, perhaps less restrictive local approximations for VOD could recover the vast majority of the correlation energy at a much reduced cost. Recently, Van Voorhis *et al* investigated what they termed the Imperfect Pairing (IP) approximation, which incorporated certain two-center terms into the wave function with moderate success[76]. In fact, this IP wave function is very similar to the GVB-RCI wave function mentioned above.[106] This chapter will investigate the localizability of electron-electron correlations in the valence space more generally, and, in partic-

ular, explore the success of two-center local correlation models in terms of (1) their recovery of the non-local correlation energy and (2) their ability to reproduce aspects of potential energy surfaces in terms of both energetics and equilibrium structures. These simplified correlation models will also be compared against HF to see what benefits can be obtained by a limited treatment of valence correlation for various systems.

4.2 The models

As mentioned above, these local correlation models may be viewed as approximate versions of the standard coupled cluster doubles (CCD) Ansatz in which some of the cluster amplitudes have been constrained to be zero. We assume that the wave function takes the form

$$|\Psi_{CCD}\rangle = e^{\hat{T}_2}|\Phi_0\rangle, \tag{4.1}$$

where $|\Phi_0\rangle$ is some single-determinantal reference that can be either the Hartree-Fock (HF) state or perhaps a reference obtained from orbital optimization at a higher level of theory, and $\hat{T}_2$ is a double excitation operator,

$$\hat{T}_2 = \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_i \hat{a}_b^\dagger \hat{a}_j. \tag{4.2}$$

Here, i, j, k refer to occupied orbitals, a, b, c, to virtual orbitals, μ, ν, σ to atomic orbitals, and i^*, j^*, k^* to virtual orbitals in same same region of space and paired with their respective occupied orbitals. As written above, the spins are over spin orbitals. The analogous double de-excitation operator $\hat{\Lambda}_2$ can also be defined as

$$\hat{\Lambda}_2 = \sum_{ijab} \lambda_{ij}^{ab} \hat{a}_i^\dagger \hat{a}_a \hat{a}_j^\dagger \hat{a}_b. \tag{4.3}$$

The energy can be calculated as the minimum of the Lagrangian,

$$L_{CCD} = \langle\Phi_0|(1 + \hat{\Lambda}_2)e^{-\hat{T}_2}\hat{H}e^{\hat{T}_2}|\Phi_0\rangle \tag{4.4}$$

with respect to the t and λ amplitudes.

In valence orbital-optimized coupled cluster doubles (VOD)[46], two additional aspects are necessary. First, the orbitals are varied in order to minimize the energy. This yields the optimal single-determinantal reference at the CCD level instead of the optimal mean-field orbitals obtained in the HF procedure. Second, it is necessary to define an active space. We believe that the active space should in general be large enough to contain the significant static correlations and also be well-defined from system to system. For these reasons, we typically choose all valence orbitals as active. Moreover, in terms of perfect pairing, each active occupied orbital requires a corresponding virtual orbital. Together, this is called the perfect-pairing active space. The coupled cluster equations are solved only within this space; that is, only excitations from active occupied orbitals to active virtual ones are included.

We can now begin to consider local approximations to VOD based on the class of excitations included. The primary change is the definition of the cluster operator, though the orbital optimization problem also changes slightly as will be discussed below. In the simplest case, the perfect-pairing approximation, we can write the cluster operator as a sum over active electron pairs:[58]

$$\hat{T}_{PP} = \sum_i t_{ii}^{i^*i^*} \hat{a}_{i^*}^\dagger \hat{a}_i \hat{a}_{\bar{i}^*}^\dagger \hat{a}_{\bar{i}}. \tag{4.5}$$

It includes only correlations of each electron pair in orbital i with a corresponding anti-bonding orbital i^* localized in the same region of space (thus, it is a one-center model).

In order to generalize slightly, one could include some or all two-center correlation terms. These come in two types: "covalent" and "ionic" terms. The imperfect pairing approximation (IP)[76], loosens the definition of the electron pair, and any two electrons are correlated together, but the excitations force the electrons to remain in their respective regions of space (or to undergo fermionic exchange), as shown in Figure 4.1. These additional terms are covalent in nature. The cluster operator takes the form

$$
\begin{aligned}
\hat{T}_{IP} &= \sum_{ij} t_{ij}^{i^*j^*} (\hat{a}_{i^*}^\dagger \hat{a}_i + \hat{a}_{\bar{i}^*}^\dagger \hat{a}_{\bar{i}})(\hat{a}_{j^*}^\dagger \hat{a}_j + \hat{a}_{\bar{j}^*}^\dagger \hat{a}_{\bar{j}}) \\
&\quad + t_{ij}^{j^*i^*} (\hat{a}_{j^*}^\dagger \hat{a}_i + \hat{a}_{\bar{j}^*}^\dagger \hat{a}_{\bar{i}})(\hat{a}_{i^*}^\dagger \hat{a}_j + \hat{a}_{\bar{i}^*}^\dagger \hat{a}_{\bar{j}}).
\end{aligned} \tag{4.6}
$$
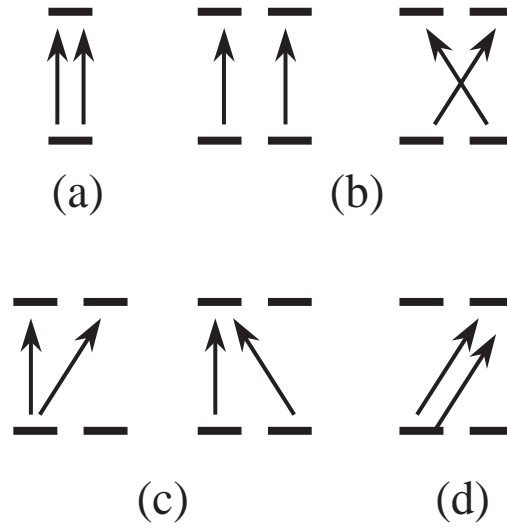
Figure 4.1: Different classes of one and two-center local excitations and their corresponding models. Each successive model includes all terms from the previous models in addition to the one(s) shown. (a) Perfect Pairing, (b) Imperfect Pairing, (c) Singly-Ionic Pairing, and (d) Doubly-Ionic Pairing.

One could go further, and include singly-ionic (SIP) or doubly-ionic (DIP) terms which allow the electrons to move around the molecule as they correlate. DIP is aesthetically pleasing in that it contains all two-center terms, and in that respect is a logical successor to PP. That is to say, whereas PP is exact (as compared to both VOD and CASSCF) for a single electron pair, DIP is fully equivalent to VOD for two pairs. Of course, one could also include some or all three-center terms (one can break them down into several subclasses) and four-center terms. With all terms included, we recover VOD with no approximation.

In general, the HF and non-local correlation energies are invariant to unitary transformations among the occupied or among the virtual orbitals. Only the space spanned by the two sets of orbitals determines the energy. Therefore, when performing a VOD calculation, one must mix orbitals only between the occupied and virtual subspaces, and between the active and inactive subspaces.[46] However, once local approximations are made such that each occupied correlates only with a certain subset of virtuals, it becomes necessary to also optimize the orbitals among the active occupieds and among the active virtuals. This adds some complexity to the orbital optimization problem, but the extra difficulty is outweighed by the computational savings elsewhere in the methods[132].

Investigation of these various models using a pilot code offers insight into the importance of different types of terms. In Figures 2-4, we decompose the VOD cluster amplitudes into these aforementioned classes for a set of fixed PP orbitals (*i.e.* orbitals optimized at the perfect-pairing level of theory) and plot the magnitude of each cluster amplitude t_{ij}^{ab}. These orbitals were chosen because they are well-localized, and they provide minimal bias among the different two-center models. Of course, the orbitals should be variationally optimized for each different approximation, but the approximate orbitals suffice for the investigation here. In effect, instead of performing a true VOD calculation, we are computing the valence space CCD energy using the PP orbitals.

We present three cases: propane, N_2, and allyl anion in the 6-31 or 6-31G* basis sets.[1, 2] In all three, the one-center amplitudes are unquestionably dominant and are up five times larger than the next largest amplitudes. The covalent two-center
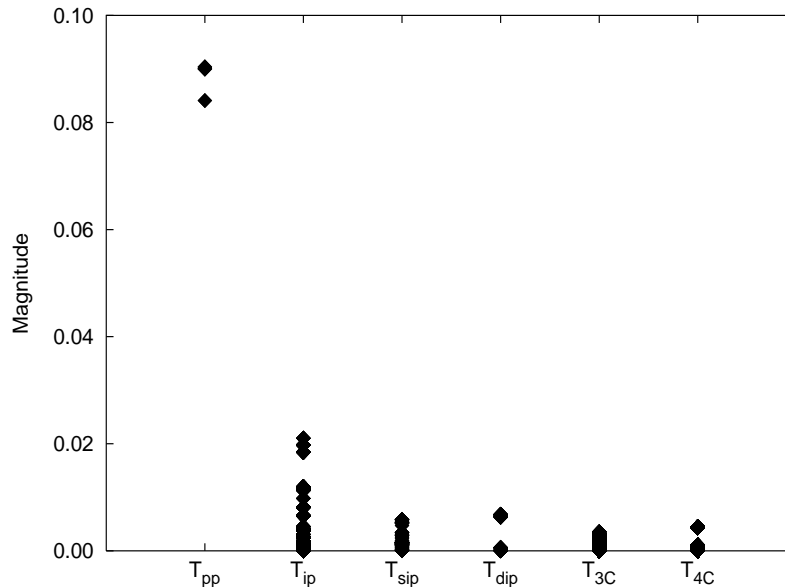
Figure 4.2: Decomposition of the full coupled-cluster T_2 vector into spatial classes for propane using localized perfect-pairing orbitals formed from the 6-31G basis.

amplitudes included in the IP model clearly offer the leading correction to PP, since they are also up to several times larger than the remaining amplitudes. For propane, the singly and doubly ionic amplitudes are the next most important (and roughly equal in magnitude), and they are somewhat larger than the multitude of three- and four-center amplitudes. For nitrogen, a similar picture exists, but now the doubly-ionic terms are substantially more important than singly-ionic ones.

At first it may seem surprising that doubly-ionic amplitudes are at least as important as the singly-ionic ones, since it is tempting to think of doubly-ionic terms as being associated with basis-set-superposition error (BSSE). Many of them do behave as two electrons in one region of space correlating with an orbital elsewhere in space. In fact, in studies on local, atom-based MP2 models doubly-ionic terms were relatively insignificant[40]. If we were using an atomic orbital framework in which the atomic centers defined the local truncations, this would certainly be true. However, in these models each orbital is paired with only one other orbital. Therefore, the dominant doubly-ionic amplitudes observed here are not associated with BSSE;
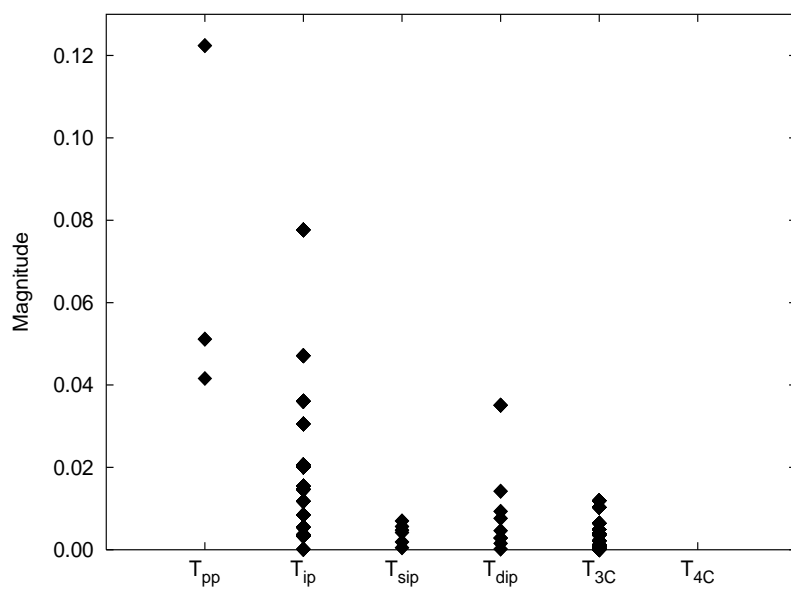
Figure 4.3: Decomposition of the full coupled-cluster T_2 vector into spatial classes for N_2 using localized perfect-pairing orbitals formed from the 6-31G* basis.
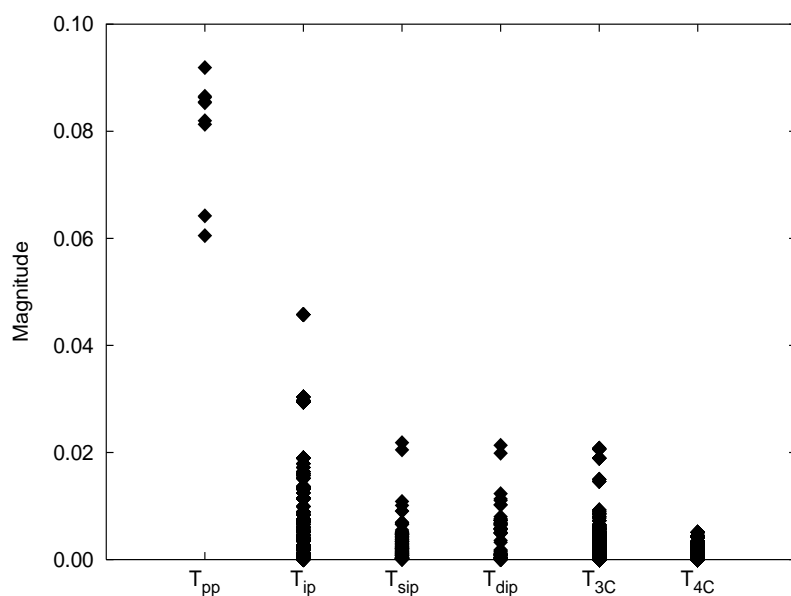
Figure 4.4: Decomposition of the full coupled-cluster T_2 vector into spatial classes for the allyl anion using localized perfect-pairing orbitals formed from the 6-31G basis.

rather they are correlations of one pair of electrons with other virtual orbitals in the *same* region of space. For example, the large DIP amplitudes in the nitrogen molecule are $(\sigma)^2 \rightarrow (\pi^*)^2$, $(\pi_x)^2 \rightarrow (\pi_y^*)^2$, $(n)^2 \rightarrow (\pi^*)^2$, and other similar *local* correlations. Thus, in a two-center model, one ought not ignore doubly-ionic excitations, particularly in molecules with multiple-bonding or lone pairs. In fact, including these local doubly-ionic amplitudes helps compensate for the difficulty in defining a correlating orbital for the lone pairs. The lone pair will correlate with all orbitals in the same region of space (and elsewhere).

This issue also reveals a weakness of these models. The focus on the electron pair using orthogonal orbitals rather than the atom means that a quadratic number of terms is necessary in a pair-based model just to describe all doubles-level atomic correlations. In a non-orthogonal atom-based local scheme,[40, 6] one could obtain the same atomic correlations with only a linear number of excitations (*i.e.* all excitations on the atom). Of course, such a model would still neglect interactions between atoms/bonds that are included in the two-center pair models. On the other hand, the use of non-orthogonal functions required by atom-centered models adds a different set of complexities to the theory.

Finally, the three- (T_{3C}) and four-center (T_{4C}) terms tend to be similar to one another in magnitude and are generally slightly smaller than the two-center ones, with the exception of molecules exhibiting resonance valence-bond structures. The allyl anion has four correlated, delocalized π-electrons, and in this case we find that the three-center terms are essentially of equal importance with the two-center ones. Decomposition of the three-center terms into different classes reveals no obvious truncations, and in many cases they are all of similar magnitude with four-center amplitudes. The only minor exception is that amplitudes corresponding to excitations where two electrons begin in the same occupied orbital and separate to two different virtuals (whether in a two- or three-center way) tend to be slightly smaller than the corresponding coalescence terms starting from different occupieds orbitals.

Thus, though it is clear that a two-center model alone cannot account for all correlations in molecules exhibiting significant electronic delocalization or resonance effects, DIP is likely to be a reasonable improvement over IP at relatively little ad-

ditional cost. In the rest of this chapter, we focus on the SIP and DIP models in comparison to PP, IP, and VOD for their treatment of valence correlations. The cluster operator for DIP takes the following form:

$$
\begin{aligned}
\hat{T}_{DIP} &= \hat{T}_{IP} + \sum_{ij}^{i \neq j} \Big[t_{ii}^{i*j*} (\hat{a}_{i*}^{\dagger} \hat{a}_i \hat{a}_{\bar{j}*}^{\dagger} \hat{a}_{\bar{i}} + \hat{a}_{j*}^{\dagger} \hat{a}_i \hat{a}_{\bar{i}*}^{\dagger} \hat{a}_{\bar{i}}) \\
&\quad + t_{ij}^{i*i*} (\hat{a}_{i*}^{\dagger} \hat{a}_i \hat{a}_{\bar{i}*}^{\dagger} \hat{a}_{\bar{j}} + \hat{a}_{i*}^{\dagger} \hat{a}_j \hat{a}_{\bar{i}*}^{\dagger} \hat{a}_{\bar{i}}) \\
&\quad + t_{ii}^{j*j*} \hat{a}_{j*}^{\dagger} \hat{a}_i \hat{a}_{\bar{j}*}^{\dagger} \hat{a}_{\bar{i}} \Big].
\end{aligned}
\tag{4.7}
$$

The SIP cluster operator simply eliminates the last term in Eq. 4.7. The simplest way to derive the equations for the DIP amplitudes is to eliminate the three- and four-center amplitudes in the standard CCD equations. At present, we have done so from a closed-shell, spin-restricted formulation[133], though one could readily generalize to open-shell restricted or unrestricted versions. The complete equations for the t amplitudes are presented in Appendix A.

The computational effort required for these models is dramatically reduced compared to their non-local counterparts. Whereas a VOD calculation scales with the sixth-power of system size, these calculations scale only cubicly once sparsity of the integrals between the localized molecular orbitals is properly accounted for. Thus, the effort is only the square root of that for VOD. Even if their performance is not entirely faithful to VOD, they can provide improved reference wave functions over Hartree-Fock for further calculations (such as corrections for dynamical correlation) at very low cost.

4.3 Implementation

The doubly-ionic pairing (DIP) model is implemented along the lines described in Ref. [132], and SIP is calculated by constraining all doubly-ionic excitations to have zero amplitude. Briefly, just as for IP, we have only a quadratic number of cluster amplitudes (as compared to a quartic number in VOD) but a cubic number of integrals (instead of quadratic) of the forms $\langle ij|ik \rangle$, $\langle ij|i^*k \rangle$, $\langle i^*j|i^*k \rangle$, and their exchange partners. This means that given a set of orbitals/integrals, the DIP amplitude equations

are trivially cheap to solve. The integrals are formed by building half-transformed coulomb and exchange matrices of the form $\langle i\mu|i\nu\rangle$ and then contracting over the atomic orbital μ and ν indices with molecular orbital coefficients.

In VOD or VQCCD, solving the amplitude equations is so expensive that the orbital optimization problem and the amplitude equations are solved simultaneously[47, 120]. This means that the Hellman-Feynman theorem which enables us to easily find the derivatives of the energy with respect to orbital rotations is not satisfied until we approach convergence. In the local case, since solving the cluster equations is so inexpensive, we can afford to solve them to convergence at each orbital rotation step. Thus, our general algorithm runs as follows[132]:

1. Obtain a guess for the localized orbital pairs

2. Build and transform the integrals

3. Solve for the t and λ amplitudes

4. Calculate the orbital rotation gradient and diagonal Hessian

5. Take an orbital step, and return to step 2.

For the purposes of this study, we have not implemented the orbital rotations of step 4. Instead, we simply use optimized orbitals from an IP or PP calculation as an approximate, fixed set of orbitals from which to perform the DIP calculation.

Actually, it turns out that one needs all N^4 integrals in solving for the SIP or DIP amplitudes, but these four-index integrals always occur in contraction with the singly- and doubly-ionic amplitudes, such as

$$t_2 \leftarrow \sum_m \langle ij^*|km^*\rangle t_{ii}^{i^*m^*}. \tag{4.8}$$

Since there are relatively so few of these amplitudes, we are able to define new intermediates that act as modified molecular orbital coefficients,

$$D_\mu^{\tilde{i}^*} = \sum_m C_\mu^m t_{ii}^{i^*m^*} \tag{4.9}$$

that enable the replacement of N^4 integrals with N^3 ones like $\langle ij^*|k\tilde{i}^*\rangle$ at a cost of only a larger prefactor.

4.4   Results and discussion

We evaluate the performance of the SIP and DIP models in regards to their ability to recover the correlation energy of the non-local VOD, and their ability to reproduce potential energy surfaces qualitatively and quantitatively through a variety of examples. The models are also compared to their computationally less-expensive relatives HF, PP, and IP. Unless otherwise indicated, PP orbitals are used for all local and non-local calculations to minimize the bias among the local orbitals. Furthermore, both HF and the pairing models use spin-restricted orbitals. The ionic models have been implemented in a developmental version of the Q-CHEM program package[85], and PP and IP are available in current release versions.

4.4.1   Recovery of the valence correlation energy

To begin, we assess what fraction of the non-local VOD correlation energy these local models recover given an identical set of PP orbitals. Consider first the three examples from above for which the cluster amplitudes were decomposed: propane, molecular nitrogen, and the allyl anion. The percent recoveries of the correlation energy are listed in Table 4.1. A few sample total energies are also reported in Table 4.2 to facilitate the reproduction of our work. For simple propane, none of the methods do particularly poorly. Even the minimal PP model recovers a sizeable 70% of the full correlation energy, and the simplest two-center model (IP) brings the recovery rate above 90%. Including singly- and doubly-ionic terms each provide a little extra correlation energy for a total recovery around 95%. Thus, for fairly simple, large-gap organic molecules, these two-center local models recover the vast majority of the correlation energy at a drastically reduced cost.

In the nitrogen case, it was observed that the doubly-ionic terms are important due to inter-pair correlations present in molecules with multiple bonding. This observation manifests itself in the very large improvement in going from IP/SIP (90-91% recovery) to DIP (97% recovery), demonstrating the benefit of including all two-center terms. Finally, for the allyl anion case, which has two delocalized electron pairs, the

Table 4.1: Percent correlation energy recovered as compared to non-local VOD with fixed perfect-pairing orbitals. Unless otherwise noted, calculations are in a cc-pVDZ basis.

Molecule	PP	IP	SIP	DIP
C_3H_8 [a]	70.5	90.4	93.0	94.8
N_2 [b]	51.3	90.0	91.3	96.9
$C_3H_5^-$ [a]	56.6	81.8	85.5	87.9
CH_4	77.8	95.7	97.4	98.6
C_2H_6	73.7	93.9	95.9	97.1
C_4H_{10}	70.8	92.1	94.2	95.5
C_6H_{14}	70.1	92.0	94.1	95.4
C_8H_{18}	69.6	91.7	93.9	95.2
C_2H_4	69.7	92.8	94.3	95.9
C_8H_{10}	62.4	88.0	90.6	92.3

[a] 6-31G basis [b] 6-31G* basis

correlations are less local (as indicated by larger three- and four-center cluster amplitudes observed in Figure 4.4). In that case, the PP recovery rate drops below 60%. IP provides the dominant correction, restoring up to 82% of the correlation energy, but even the full two-center model cannot quite recover 90%.

Finally, we briefly look at the long-range behavior of the models by examining the correlation energy recoveries for alkane chains C_nH_{2n+2} for n=1,2,4,6,8 and alkene chains C_nH_{n+2} for n=2,8 in the cc-pVDZ basis (See Table 4.1)[64]. Asymptotically (*i.e.* $n = 8$), PP recovers just under 70% of the VOD correlation energy, whereas IP brings that value up to almost 92%. Adding singly- and doubly- ionic terms each provide about 1.5% additional percentage points, for a total of around 95% for DIP—similar to what was observed for propane in the 6-31G basis set as described previously. Note also that DIP recovers roughly half of the correlation energy missing in IP. For the alkenes, which combine multiple bonding and electron delocalization effects, PP obtains much less of the correlation energy (62% for octatetraene). The other methods degrade as well, though to a lesser extent—DIP asymptotes to a value around 92% of the correlation energy. Though not as good as for the alkanes, this value still represents a reasonable improvement over IP (recovering a third of the missing correlation energy) and it recovers the vast majority of the correlations

Table 4.2: Absolute energies for some sample species using PP converged orbitals (except for HF). The alkanes here use $r_{CH} = 1.10$ Å, $r_{CC} = 1.54$ Å, $\langle HCH = 109.5°$. The geometry for CH_3F was obtained from Ref [4]

	PP	IP	SIP	DIP
CH_4 [a]	-40.25915262	-40.27311516	-40.274395613	-40.27532274
C_2H_6 [a]	-79.33963037	-79.36876008	-79.371567337	-79.37333716
F^- [b]	-99.38795411	-99.46835242	-99.46835242	-99.47272782
CH_3F [b]	-139.13678606	-139.20727304	-139.21055877	-139.21347173

[a] cc-pVDZ basis [b] 6-31G* basis

extremely efficiently.

Based on wider testing, these examples span the range of possibilities. PP does a moderately good job (particularly given its simplicity) of recovering the complete correlation energy within the active space. Imperfect Pairing is the leading correction and it recovers a much larger fraction of the correlation energy (from 20 to 50 percentage points more). However, if one wishes to attempt quantitative accuracy, at least all two-center terms are necessary. The DIP model recovers roughly 4-10 percentage points more correlation energy than IP. For molecules with intuitively local electronic structure, a complete two-center model recovers in the mid-ninety percent range. However, as the degree of electron-delocalization increases (or as multi-pair correlations become stronger), all of these models degrade. In that case, the full two-center models typically recover a little less than 90% of the correlation energy. The remaining correlation energy seems to be spread out almost evenly among the three- and four-center terms, making less restrictive approximations more difficult to define (and much more costly computationally). On the one hand, such a recovery is still very respectable, particularly given the economical computational cost of the models. On the other hand, a true test is how well the model does at predicting chemistry, which depends on the reliability of relative energies. We will assess both their reproduction of potential energy surfaces and reaction barriers, and their abilities to inexpensively predict molecular geometries.

Table 4.3: Barrier heights in kcal/mol for two simple reactions and percent recovery of the VOD correlation energy for the reactant complex and the transition state using PP orbitals.

	HF	PP	IP	SIP	DIP	VOD
$F^- + CH_3F \rightarrow CH_3F + F^-$						
Barrier (kcal/mol)	16.9	18.9	17.2	15.7	15.8	13.7
Complex (%)	0.0	47.9	91.5	92.9	94.4	
TS (%)	0.0	46.1	89.9	92.1	93.4	
$HOCl + HCl \rightarrow H_2O + Cl_2$						
Barrier (kcal/mol)	76.8	72.1	77.4	77.4	77.8	74.9
Complex (%)	0.0	60.5	86.4	90.0	93.4	
TS (%)	0.0	63.2	84.7	88.2	91.3	

4.4.2 Simple chemical reactions

Shifting the focus now to chemical reaction barriers and potential energy surfaces, we examine a variety of reactions from the simple S_N2 reaction to the very difficult Cope rearrangement. Starting with the S_N2 identity reaction of CH_3F and F^- to form the same species. Using CCSD(T)/cc-pVQZ+1 optimized geometries[4], the 6-31G* basis reaction barriers from the pre-reactive complex to the transition state are presented along with the percent correlation energy recoveries for both species in Table 4.3. Though the overall differences are only a couple of kcal/mol, each more elaborate correlation model progressively approaches the non-local barrier height of just under 14 kcal/mol. In particular, an improvement occurs from PP to IP, and from IP to the ionic pairing models. The residual error of 2 kcal/mol is a non-parallelarity error: the local models recover slightly more of the pre-reactive complex correlation energy than they do for the transition state, which is not surprising since the transition state contains a more delocalized electronic structure (two partial bonds). The improvement from one-center to two-center models is attributable to a smaller non-parallelarity error as the more elaborate models provide a better and more uniform description of the overall potential energy surface.

Next the activation barrier for the reaction,

$$HOCl + HCl \rightarrow H_2O + Cl_2 \tag{4.10}$$
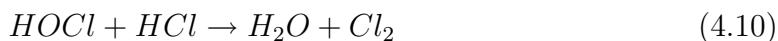
is calculated based on the B3LYP/6-31+G* basis set stationary points.[134] The results show a similar pattern to the S$_N$2 reaction. In this case, the methods all have a similar error with respect to VOD, but whereas PP underestimates the barrier, the others overestimate it. In this case, moving from IP to DIP does significantly improve the amount of correlation energy recovered (from 85 to 91% for the transition state), but the non-parallelarity error remains roughly constant.

In systems where only a single electron pair is strongly correlated, it is likely that even PP can provide a reasonable description. The effects of the two-center models, especially SIP and DIP, are not especially noticeable. In fact, neither of these systems provided a convincing need to go beyond HF, since it also performed reasonably on both activation barriers. Two-center correlated models should become important in cases where at least two pairs of electrons are strongly correlated. With this in mind, we focus on two more difficult reactions: the simultaneous removal of two hydrogen atoms from methane and the Cope rearrangement.

4.4.3   Reactions with multiple strongly-correlated electrons

Table 4.4 and Figure 4.5 compare how well the different local approximations conform to the VOD PES (using PP orbitals) for simultaneously removing two hydrogens from tetrahedral methane ($r_{CH} = 1.091, \angle HCH = 109.4°$) to form ^{1}CH$_2$ and ^{1}H$_2$. The slice of the PES is such that the H-H distance of the removed pair and all other degrees of freedom except the distance between the two hydrogens and the carbon stay fixed. At equilibrium, the distance from the midpoint of the hydrogen pair to the carbon is 0.63 Å, and 2.43 Å corresponds to an C-H distance of 2.59 Å. Table 4.4 and Figure 4.5 set the energy of equilibrium methane for each method to zero and demonstrate the non-parallelarity energy. Restricted HF is completely unreasonable for this stretch, since it cannot describe the dissociated products correctly. PP also rises in energy too quickly and introduces a sizeable error. IP behaves slightly better, but still introduces errors of several kcal/mol. Only SIP and DIP track VOD with a less than 1 kcal/mol error across most of this slice. DIP in particular remains parallel to VOD to within 0.6 kcal/mol throughout this range.
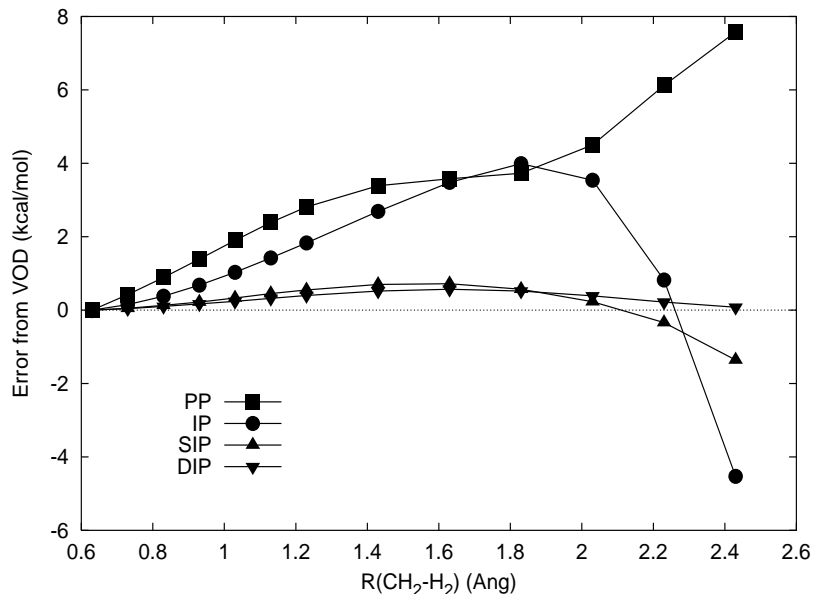
Figure 4.5: Error relative to VOD for the removal of two hydrogens from methane in the 6-31G* basis. All methods use perfect-pairing orbitals.

Table 4.4: VOD/6-31G* potential energy surface for removing to hydrogens from methane with a constant H-H distance. The distance coordinate measures from the midpoint between the two abstracted hydrogens and the carbon, and all other degrees of freedom are kept constant. Values are in kcal/mol.

$R(CH_2 - H_2)$	HF	PP	IP	SIP	DIP	VOD
0.63	0.00	0.00	0.00	0.00	0.00	0.00
0.83	5.40	0.89	0.38	0.13	0.10	10.23
1.03	12.78	1.90	1.03	0.33	0.24	39.39
1.23	22.34	2.81	1.83	0.55	0.40	76.53
1.43	34.25	3.39	2.69	0.70	0.52	113.72
1.63	48.56	3.58	3.48	0.72	0.57	146.40
1.83	64.83	3.73	3.99	0.57	0.52	172.30
2.03	81.20	4.50	3.54	0.23	0.39	190.55
2.23	55.20	6.12	0.82	-0.34	0.22	201.81
2.43	40.36	7.58	-4.53	-1.36	0.08	208.50

This system exemplifies the importance of two-center terms when multiple pairs of electrons are highly correlated. PP allows only a mean-field interaction between pairs, whereas DIP explicitly (albeit incompletely) correlates pairs together. The Cope Rearrangement will provide an even more difficult example, with three highly-correlated electron pairs

Before examining the Cope Rearrangement, however, it is instructive to study the behavior of these models on benzene. Considering only the (6,6) space corresponding to the π electrons and their bonding and anti-bonding orbitals, one can construct three orthogonal, localized orbitals. However, since benzene has six equivalent bonds, these three orbitals treat three of the bonds differently than the other three, and this results in a symmetry-breaking that deforms the D_{6h} structure to one with alternating short and long bonds (D_{3h}).

In a non-local formulation using IP orbitals, including all possible excitations restores the equivalency and favors the D_{6h} structure, as we saw in Chapter 3. However, once truncations are made, this equivalency is no longer necessarily true, as is pictured in Figure 4.6. For instance, PP prefers the D_{3h} geometry by about 3 kcal/mol for a 2° deformation. It was observed that IP does much to eliminate this symmetry-breaking, reducing it to only about 0.5 kcal/mol and 1°[76]. Unfortunately, neither SIP nor DIP is able to remedy this situation completely. SIP provides only a trivial improvement over IP, but DIP cuts the IP symmetry-breaking in half, to only 0.25 kcal/mol and 0.5°. Presumably this is due to the inclusion of correlations of each pair of π electrons equivalently with all of the π^* orbitals. On one hand, the symmetry breaking in DIP is energetically small and would not affect reaction barriers significantly. On the other hand, vibrational frequency or structural predictions would be very poor given such a potential energy surface. However, as has been demonstrated, the artifacts of symmetry-breaking in cases like this can be eliminated using a non-orthogonal formulation[105], to generate six equivalent, redundant occupied (and six virtual) orbitals, one occupied and virtual for each atom. This strategy is likely a better solution to the problem than simply including more non-local excitations.

Another way to view the benzene symmetry-breaking is as symptomatic of a severe non-parallelarity error. PP recovers only 59% of the VOD correlation energy for the
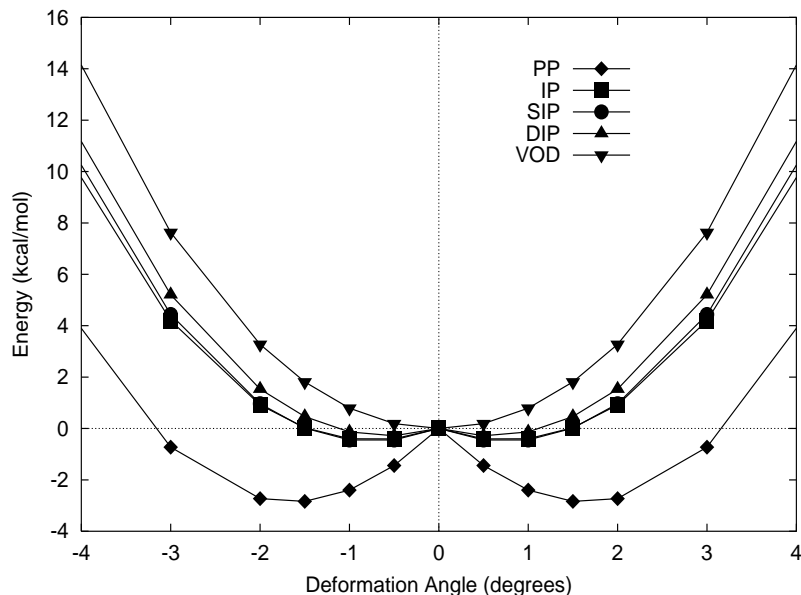
Figure 4.6: Deformation of benzene and symmetry breaking of local models in the 6-31G* basis using IP orbitals (except for PP, which uses its own orbitals) for the π-space.

D_{6h} structure and 69% at a deformation of 1.5°, for a 10 percentage-point difference in the recovery rate! For the same geometries, IP recovers 79% and 83%, respectively, reducing the gap to 4 percentage points. SIP increases the recovery to 86% and 90%, but does not alter the gap. But DIP reaches 89% and 92%—a non-parallelarity error of only 3 percentage points, which explains the improvement.

With some understanding of the behavior of these models on benzene, we can apply them to the Cope Rearrangement[115, 117, 116, 118]. This reaction is infamous in electronic structure theory for its difficulty: the transition region of the reaction involves two π bonds and one σ bond rearranging (see Figure 4.7) and requires an accurate treatment of static correlation not found in a typical single-reference method. However, the electronic structure of the "non-reactive" σ-bonding framework also couples strongly to the bonds being directly broken and formed, meaning that these other correlations must be treated accurately for an even qualitatively correct potential energy surface. For a nice review on the subject, see Ref [118] and references
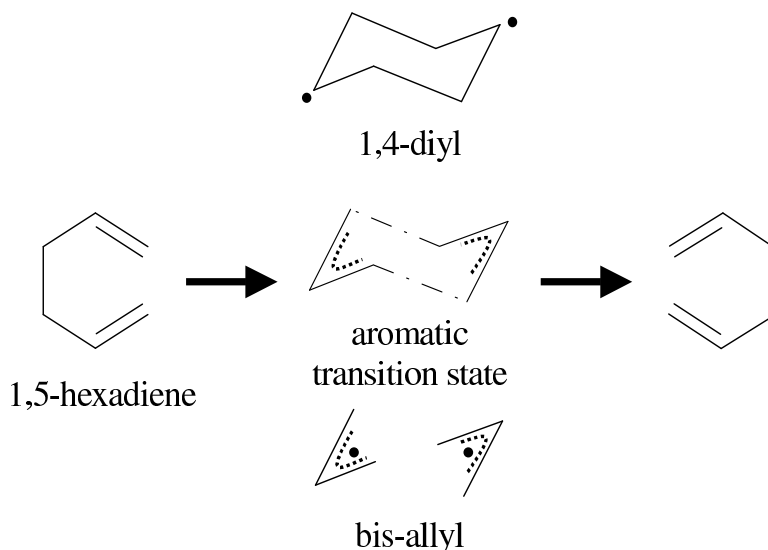
Figure 4.7: The Cope rearrangement. 1,5-hexadiene rearranges as shown by crossing through a C_{2h} chair species. The distance between the two allylic groups at the transition state/intermediate determines to what extent it is diradicaloid or aromatic.

therein. Here, we focus only on the ability for these various models to reproduce CASSCF and VOD in a minimal correlation space. The reaction involves a chair-like C_{2h} transition state/intermediate, as shown in Figure 4.7. The relevant question is how and where the reaction crosses this C_{2h} slice of the PES. Essentially, this slice defines a coordinate which varies the separation between two allyl fragments. In the six-orbital/six-electron, or (6,6), active space (i.e. the aforementioned two π and one σ orbitals, plus their corresponding π^* and σ^* orbitals), CASSCF predicts two essentially degenerate routes separated by a barrier on the C_{2h} surface slice. At 1.64 Å separation between the adjacent carbons on each allyl-fragment in the C_{2h} structure, CASSCF finds a diradicaloid intermediate (the 1,4-diyl in Figure 4.7). At 2.19 Å it finds a aromatic transition state. These two stationary points are separated by a maximum along the C_{2h} slice.

Using CASSCF geometries, VOD(6,6) mimics CASSCF fairly well, particularly for the aromatic species. VOD(6,6) does predict the diradicaloid intermediate to
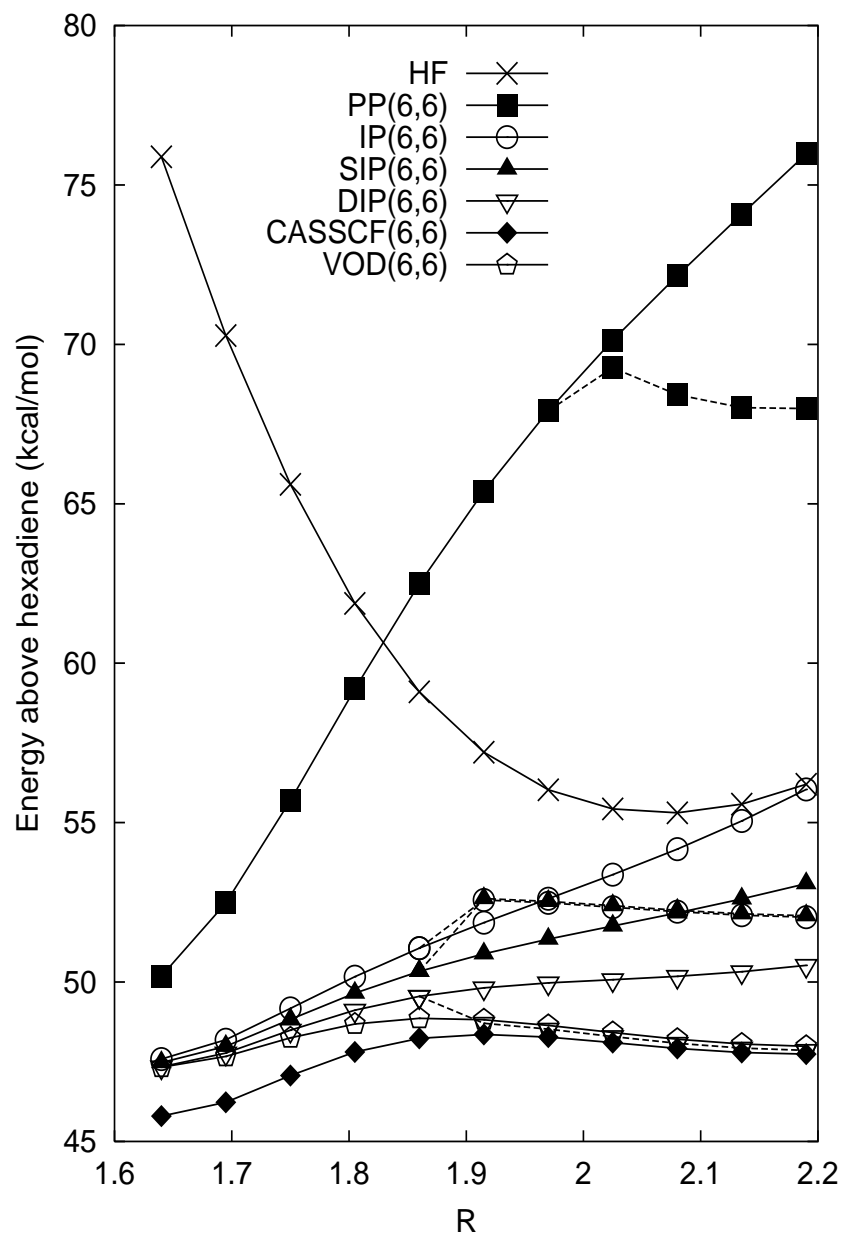
Figure 4.8: C_{2h} potential energy slice for the Cope Rearrangement in a (6,6) active space using the 6-31G* basis. IP, SIP, and DIP all use IP orbitals. PP and VOD use their own optimized orbitals. Dotted lines indicate a symmetry-broken solution. R refers to the interallyic distance.

be a couple of kcal/mol higher above 1,5-hexadiene than CASSCF, which reflects the difficulty in describing the diradical correctly with a restricted wave function. Nevertheless, VOD(6,6) is qualitatively correct and quantitatively in error by less than 2 kcal/mol. Using IP(6,6) orbitals instead of the true VOD orbitals introduces only a slight error (less than 1 kcal/mol) in the aromatic region. On the other hand, restricted HF predicts a very unphysical curve along this PES slice, failing in particular to describe the diradicaloid species at short bond-lengths.

It is reasonable to ask how the various local models do on this same slice of the PES. First, just like for benzene, in the aromatic region of the C_{2h} slice there exists a symmetry-breaking solution. First let us focus on the symmetry-preserving solution. Looking at Figure 4.8, we find that PP(6,6) is woefully inadequate. It has a non-parallelarity error of almost 25 kcal/mol and it recovers almost 96% of the VOD(6,6) correlation energy (with IP(6,6) orbitals) at 1.64 Å, but only half as much at 2.19 Å. Though PP is not totally unreasonable for the diradical, it cannot describe the aromatic region well.

Perhaps the inclusion of two-center correlations will improve the description. Though IP(6,6) performs far better, it still introduces an error of roughly 8 kcal/mol, recovering 99.4% (1.64 Å) and 86% (2.19 Å) of the VOD(6,6) correlation energies respectively. Using IP(6,6) orbitals, SIP(6,6) and DIP(6,6) improve further on IP. In fact, DIP(6,6) recovers 99.8% and 96% of the VOD(6,6) correlation energies and gives a relative error of only 2-3 kcal/mol in the aromatic region relative to VOD(6,6). As in benzene, the improved model describes the aromatic behavior better. If one allows the local wave functions to break symmetry in the aromatic region, the energy gaps are reduced, though at the cost of introducing a cusp into the potential energy surface around 2.0 Å and destroying the symmetry of the wave function. Clearly the additional two-center terms are very important in approximating the non-local VOD or CASSCF potential energy surfaces and demonstrate the improvement DIP provides over even IP or SIP. However, the symmetry-breaking issue is indicative of the challenge systems with multiple resonance structures provide for these simplified models.

4.4.4 Molecular structure prediction

Finally, another important use of static correlation methods is for the determination of molecular structures. GVB-PP does fairly well at predicting structures, but for some molecules (particularly those containing halogens and requiring the coupling of lone pair electrons with the bonding ones) it does not perform as well corresponding non-local models do. We investigate whether including the majority of the valence correlation energy provides a sufficiently accurate description of PES minima. For PP, IP, and VOD (using the truly optimized orbitals) analytical gradients are available and used. For SIP and DIP, IP orbitals are used and the geometries are optimized by finite difference. A test set of 22 small molecules with 34 unique bond lengths[121, 120] shows that although it is improved over HF in general, PP in particular overestimates bond lengths by roughly 0.03-0.10 Angstroms for molecules like F_2, ClF, Cl_2, HOF, and H_2O_2. Non-local VOD reduces this error to around 0.01 for these same cases.

As shown in Table 4.5, including more local correlations via IP, SIP, or DIP does statistically improve upon PP equilibrium geometries as compared to experiment or VOD. Whereas HF has a root-mean-square (RMS) error of 0.036 Å and PP has a root-mean-square (RMS) error of 0.033 Å versus experiment, IP/SIP/DIP (using IP orbitals) reduce this error to around 0.23 Å, and VOD has an RMS error of 0.016 Å. This error would be cut further in a larger basis set[120]. If instead we compare against VOD to avoid basis set effects, we see an even more noticeable improvement from HF (0.038 Å) to PP (0.026 Å), and another significant improvement upon going to the two-center models (0.015 Å). The advantage of the limited correlation over HF stands out more in the maximum errors, which are cut in half from HF (0.134 Å for Li_2) to the two-center models (0.064-0.069 Å for Cl_2). Comparing instead against VOD, the maximum errors do not change much, since none of the local models predict the Cl_2 bond distance very well, and HF still fails for Li_2, unlike the correlated models.

Comparing PP against the two-center models, particular improvement occurs in the cases mentioned previously where PP was inadequate. The largest exception is Cl_2, where despite the improvement, the error is still 0.069 Å for DIP. However, in

Table 4.5: Predicted equilibrium geometries in the 6-31G* basis.

	bond	HF	PP	IP	SIP[a]	DIP[a]	VOD	Expt.[b]
Li2	rLiLi	2.807	2.717	2.717	2.717	2.717	2.717	2.673
N2	rNN	1.078	1.103	1.110	1.111	1.112	1.112	1.098
F2	rFF	1.345	1.498	1.431	1.429	1.427	1.420	1.412
CO	rCO	1.114	1.127	1.138	1.139	1.139	1.139	1.128
CS	rCS	1.520	1.533	1.543	1.543	1.546	1.543	1.535
NP	rNP	1.455	1.503	1.502	1.504	1.505	1.503	1.491
P2	rPP	1.859	1.922	1.933	1.933	1.931	1.920	1.893
NaF	rNaF	1.885	1.889	1.899	1.899	1.899	1.899	1.926
ClF	rClF	1.612	1.693	1.661	1.656	1.658	1.653	1.628
Cl2	rClCl	1.989	2.070	2.054	2.052	2.057	1.992	1.988
HCN	rHC	1.059	1.073	1.075	1.077	1.076	1.079	1.065
	rCN	1.132	1.158	1.165	1.163	1.167	1.167	1.152
HNC	rHN	0.985	0.999	1.002	1.003	1.003	1.005	0.994
	rNC	1.154	1.170	1.179	1.177	1.181	1.183	1.169
HCP	rHC	1.063	1.086	1.083	1.083	1.083	1.085	1.069
	rCP	1.515	1.561	1.557	1.556	1.559	1.560	1.540
HNO	rHN	1.032	1.048	1.052	1.057	1.059	1.066	1.063
	rNO	1.175	1.219	1.221	1.222	1.220	1.217	1.212
HPO	rHP	1.431	1.450	1.449	1.457	1.460	1.470	-
	rPO	1.461	1.505	1.497	1.497	1.498	1.496	1.512
HOF	rHO	0.952	0.973	0.976	0.976	0.977	0.977	0.966
	rOF	1.376	1.488	1.444	1.441	1.443	1.438	1.442
HOCl	rHO	0.951	0.971	0.975	0.975	0.975	0.975	0.975
	rOCl	1.670	1.760	1.725	1.718	1.728	1.718	1.690
H2CO	rCO	1.184	1.212	1.217	1.215	1.216	1.214	1.208
	rCH	1.092	1.105	1.105	1.105	1.110	1.116	1.116
H2CS	rCS	1.597	1.631	1.642	1.635	1.635	1.629	1.611
	rCH	1.078	1.095	1.096	1.097	1.097	1.101	1.093
H2O2	rHO	0.949	0.969	0.973	0.973	0.974	0.974	0.965
	rOO	1.468	1.499	1.464	1.456	1.465	1.459	1.452
H2S2	rHS	1.327	1.350	1.355	1.355	1.357	1.352	1.345
	rSS	2.063	2.108	2.116	2.118	2.119	2.067	2.058
N2H2	rHN	1.014	1.033	1.036	1.038	1.039	1.042	1.028
	rNN	1.216	1.253	1.258	1.258	1.260	1.259	1.252
RMS error vs. Expt.		0.036	0.033	0.023	0.022	0.024	0.016	
Max. error vs. Expt.		0.134	0.086	0.066	0.064	0.069	0.044	
RMS error vs. VOD		0.038	0.026	0.015	0.014	0.015		
Max. error vs. VOD		0.090	0.078	0.062	0.060	0.065		

[a] used IP orbitals   [b] Ref. [121]

these examples there is little statistical improvement from IP to DIP. Nevertheless, these two-center methods do seem to provide an efficient approach for equilibrium structures with quite reasonable accuracy. Transition structures would likely prove more difficult and might separate the methods further, particularly in cases with multiple bonds being broken and formed simultaneously. Unfortunately, without an implementation of analytical gradients, transition state searching and characterization is computationally tedious and is beyond the scope of this chapter.

4.5 Conclusions

The contributions of different classes of excitations in a local electron-pair model were analyzed for some test systems, and the most promising truncations were identified as PP, IP, and a full two-center model (DIP). The three- and four-center terms suggest no obvious truncations, and their implementation would eliminate much of the efficiency of the two-center version relative to the non-local relatives; therefore, they were not investigated further.

These various models were applied to a variety of chemical reactions, potential energy surfaces, and geometry predictions. In general, all of these models improve upon HF. However, in a number of cases PP was found to be lacking, and IP provides the leading correction. Often, this leading correction is sufficient to improve upon the PP results. SIP/DIP usually provide the next largest correction, and DIP typically brings the recovery of the non-local correlation energy into the mid- or low-ninety percent range. However, in many cases, this additional recovery is insufficient to capture all of the effects missed by IP, as exemplified by benzene and the Cope rearrangement.

In this study, we used approximate orbitals for the ionic-pairing models instead of implementing the orbital gradients. Presumably the true orbitals would help these methods somewhat, particularly in cases where the approximate orbitals are heavily symmetry-broken. However, given the generally small improvements DIP provides over IP, it is unlikely that the correct orbitals could completely rectify any of the problems observed here.

These results suggest several directions for further research. First, there is the problem of symmetry-breaking in benzene, for example. This can be alleviated via the use of redundant excitations with non-orthogonal (NO) orbitals, as was demonstrated in a pilot implementation of NO-PP[105]. So in principle, this problem is not insurmountable.

Second, there was the slightly surprising result that doubly-ionic excitations played such a large role. This was attributable to atomic-like correlations and correlations between bonds in multiple-bonding systems. This character demonstrates a weakness of the orthogonal orbital electron-pair formulation: a non-orthogonal atomic-based model can include these same-atom atomic correlations with only a linear number of excitations.[6] In those models, only the singly-ionic excitations are important, whereas the doubly-ionic ones are negligible and more truly attributable to basis-set-superposition error. Moreover, the symmetry-breaking issues are eliminated, since for benzene for example, one would have six equivalent, atomic-like non-orthogonal orbitals. These atomic models also have the advantage that they can be defined for an arbitrary active space or even the full space. Furthermore, one might conceive of hybrid models where one correlates to some high excitation level on each atom and uses lower-level excitations (such as double) to account for inter-atomic correlations in a local fashion. Assuming that the orbital and amplitude equations can be converged efficiently, atomic orbital-based local schemes may prove overall more versatile than these pair models and offer superior wave function expansions for the correlation energy at low computational cost. However, no such efficient implementations have been made, so the extent of the trade-off in changing to redundant orbitals is not entirely clear.

Third, given the incompleteness of these models as stand-alone theories, one wonders how best to improve on them. For example, one might devise corrections to account for the remaining three- and four-center terms and the beyond valence dynamical correlations. Such corrections could come in the form of the (2) correction which has already been implemented for VOD and related methods, [5, 48, 49, 3] or one might explore other corrections using a different formulation of perturbation theory or even density functional theory. In the case of (2), it needs to be adapted to

take advantage of the local formulation. Any of these approaches raises the question of what is the best reference to correct from. Given the fact that valence correlation is typically only 30-40% of the total correlation energy, it may suffice to use a very simple model like PP which obtains more than half of the valence portion and captures the leading correlations, and correct from there. Or possibly the higher correlation energy recovery rate of DIP will prove important.

Finally, another interesting application of these approaches is for cases where the Hartree-Fock reference is poor due to orbital instabilities and symmetry-breaking effects, such as for radicals. In Chapter 2 we demonstrated that an improved reference in the context of a full-space OO-CCD helps to solve this problem in many cases. Open-shell versions of these models might offer a very inexpensive means of achieving a stable reference from which to introduce further correlations.

In conclusion, it is clear that the valence correlations are in general quite localizable, and very good approximations can be made that neglect all but a linear or quadratic number of cluster amplitudes in a VOD-like model. The models presented here have the most difficultly for systems involving multiple resonance structures or highly-delocalized electronic structure. However, for systems in the intermediate regime between those well-described by HF and those requiring an advanced method like VOD or CASSCF, the two-center local approximations provide a very nice balance between potential computational efficiency and accuracy.

Though there are clearly cases where the two-center or non-local models are preferable to PP, its utter simplicity makes it very useful as a reference wave function for a low-cost perturbative correction. Therefore, in the next chapter we generalize PP to an unrestricted formalism to enable bond dissociation to the proper atomic subunits. We will also present an extremely efficient algorithm that makes PP cost only a few times more than a HF calculation.

Chapter 5

Unrestricted perfect-pairing: The simplest wave-function-based model chemistry beyond mean field.

5.1   Introduction

Standard wave function-based electronic structure theory generally begins with a mean-field Hartree-Fock (HF) computation in order to obtain a qualitative description of the system of interest. Often, the HF description is sufficiently accurate that straightforward perturbative corrections such as second-order Møller-Plesset perturbation theory (MP2) enable the reliable prediction of chemical energetics and properties. The success of the MP2 approach, however, is predicated on obtaining a reasonable single-reference description of the system from HF. In systems with unusually strong static electron-electron correlations (radicals, diradicals, stretched bonds, and transition states, for example), the qualitatively correct wave function requires multiple determinants. In such cases, the HF wave function contains only the single most important electronic configuration and thereby biases subsequent electron-electron correlation treatments in favor of this determinant over other significant ones.

The usual approach in cases where HF behaves poorly is to construct a multi-configurational self-consistent-field (MCSCF) or complete-active-space self-consistent field (CASSCF) [41] wave function which captures the multiple significant configurations to provide a description of the static correlations in the system. For quantitative results, this wave function can be corrected with multi-reference perturbation theories. Given an active space that is sufficiently large, CASSCF, which incorporates all correlations in the active space, qualitatively describes even highly-correlated systems very well. Unfortunately, its computational cost grows factorially with the size of the active space, and feasible CASSCF calculations can only contain up to about fourteen electrons/orbitals at present. In MCSCF, this extreme computational cost is lowered by hand-selecting only those configurations that the practitioner deems important. However, its reliability and accuracy depends strongly on the skill of the user in identifying and including the important configurations. Additionally the result is not a well-defined theoretical model chemistry.[6]

The valence orbital-optimized coupled cluster doubles (VOO-CCD, or simply VOD) model was proposed several years ago to approximate CASSCF and make it applicable to a wider variety of systems.[46, 47] In VOD, a coupled cluster doubles (CCD)[17] calculation in the active space approximates the full configuration interaction (FCI) one in CASSCF, and the orbitals are optimized to minimize the energy, as in CASSCF. This approach breaks the link to the HF reference, and instead finds the reference determinant that minimizes the energy of the valence space CCD wave function. In many cases, VOD performs comparably to CASSCF at much lower cost (N^6), and it enables one to simply choose all valence orbitals as active for small to moderate-sized systems, thereby eliminating the problem of choosing the chemically relevant active space. Unfortunately, even this scaling is high, and it limits practical VOD computations to typically about fifteen non-hydrogen atoms (fewer in large basis sets).

There are some other approaches to replacing the HF reference function that should be mentioned. One promising approach that has recently been explored by Rassolov[135] uses an antisymmetrized product of strongly orthogonal geminals. Even simpler than this is to take just two functions to describe each geminal, which leads

to the so-called Generalized Valence Bond Perfect-Pairing (PP) approximation.[104, 54, 55, 56]

GVB-PP is an even simpler approximation to CASSCF which can be viewed as a strongly local restriction of VOD. Instead of including N^4 amplitudes in the wave function as in VOD, only a linear number of amplitudes are used to form a set of alpha-beta electron pairs, each of which resides in a spatial occupied orbital and which are then correlated with a single spatial virtual correlating orbital. Though somewhat crude, these alpha-beta pair correlations comprise the leading terms in the correlation energy expansion and are essential for describing the breaking of bonds correctly. It has also been applied to the study of certain classes of diradicals with much success.[136, 137, 137, 138] Importantly, it is potentially very inexpensive to calculate. Suggestions have also been made for generalizations of the PP approach that include interpair couplings,[76] or employ non-orthogonal orbitals.[105] Additionally there has been considerable effort devoted to the development of more sophisticated GVB wave functions that lift the restriction of orthogonality.[139, 140]

In this paper, we explore the feasibility of the simplest possible extension beyond HF theory, which is PP. We generalize the restricted PP approach to handle unrestricted open- and closed-shell molecules. Our point of view is that we want the result of bond-dissociation reactions to yield products that are exactly those which would be computed from separate calculations on the product fragments. For approximate treatments of electron correlation, such as PP, this can only be accomplished using unrestricted orbitals. Our implementation is based on an extremely efficient algorithm for PP utilizing the resolution of the identity approximation[141, 142, 143, 144] that introduces sub-millihartree errors in total energies. Turning from the methodology to chemistry, we will show that the simple PP model for correlation often provides a much improved starting point over HF for radicals and other difficult systems. This improved reference is a viable alternative to HF for many chemical problems and may be used subsequently for the treatment of additional correlations. Additionally, some of its limitations will be revealed in our series of test calculations.

5.2 Theory and implementation

5.2.1 Coupled-cluster perfect-pairing

In the perfect pairing model, the wave function for a closed-shell system is written as an antisymmetrized product of pair functions, g_i, and core orbitals ϕ_i,

$$|\Psi\rangle = |A[\phi_1 \phi_{\bar{1}} \phi_2 \phi_{\bar{2}} \cdots g_1 g_2 \cdots]\rangle \tag{5.1}$$

where the pair functions g_i are defined as,

$$g_i = A[\psi_i \psi_{\bar{i}} + t_i \psi_{i*} \psi_{\bar{i}*}]. \tag{5.2}$$

The core orbitals ϕ_i, the active orbitals ψ_i and amplitudes t_i are determined variationally.

The same wave function can be rewritten as a simplified coupled cluster wave function.[57, 58] In this case, it takes the form,

$$|\Psi_{PP}\rangle = e^{\hat{T}_{PP}}|\Phi_0\rangle \tag{5.3}$$

where $|\Phi_0\rangle$ is the reference determinant and $\hat{T}_{PP}$ is the cluster operator,

$$\hat{T}_{PP} = \sum_i^{pairs} t_{i\bar{i}}^{i*\bar{i}*} \hat{a}_{i*}^\dagger \hat{a}_i \hat{a}_{\bar{i}*}^\dagger \hat{a}_{\bar{i}} = \sum_i^{pairs} t_i \hat{a}_{i*}^\dagger \hat{a}_i \hat{a}_{\bar{i}*}^\dagger \hat{a}_{\bar{i}} \tag{5.4}$$

In this notation, i is an alpha active orbital, $\bar{i}$ is the beta active orbital paired with i, and $i*$ and $\bar{i}*$ are the virtual orbitals that correlate with i and $\bar{i}$, respectively. In the standard restricted version of PP, the alpha and beta spatial orbitals are identical. Here we explicitly allow them to differ.

In the unrestricted Ansatz, we separate the occupied space into three subspaces. The first two are the core space and the valence (or active) occupied pair space, containing equal numbers of alpha and beta orbitals. We typically choose all valence pairs to be active, but any chemically-reasonable number of active pairs can be used. In the active space, each alpha occupied orbital is paired with one beta occupied orbital, analogously to the restricted version. The third subspace contains unpaired, alpha-space singly-occupied molecular orbitals (SOMOs). Only the paired, active

occupied orbitals are correlated, meaning that the SOMO electrons are treated in a UHF-like fashion.

The virtual space contains only two subspaces: active and inactive. The singly-unoccupied molecular orbitals (SUMOs) are effectively part of the inactive virtual space. Of course, the orbitals in PP are completely optimized, meaning that the five alpha subspaces (and the four beta subspaces) mix freely to minimize the energy.

This uncorrelated treatment of the radical electrons is consistent with the PP treatment of electron pairs in a closed-shell species undergoing bond dissociation. Once the bond is stretched enough, the two electrons should localize to their respective atomic centers and have zero correlation amplitude between them. Any attempt to correlate these excess electrons would therefore be inconsistent with the closed-shell PP model.

The PP coupled-cluster equations are solved by constructing a coupled cluster Lagrangian,

$$L_{PP} = \langle \Phi_0 | (1 + \hat{\Lambda}_{PP}) e^{-\hat{T}_{PP}} \hat{H} e^{\hat{T}_{PP}} | \Phi_0 \rangle \tag{5.5}$$

where $\hat{\Lambda}_{PP} = \sum_i^{pairs} \lambda_i \hat{a}_i^\dagger \hat{a}_{i^*} \hat{a}_{\bar{i}}^\dagger \hat{a}_{\bar{i}^*}$ An efficient algorithm for solving this Lagrangian for the optimal t and λ amplitudes and variationally minimizing the energy with respect to the orbitals has been discussed previously,[132] so we will not repeat all the details here. Instead, we will focus only on the key steps and changes introduced by generalizing to unrestricted systems and the adaptation of the RI approximation. The interested reader is referred to Ref. [132] for further detail.

5.2.2   Initial guess

We start the UPP calculation from either a restricted (RHF or ROHF) or unrestricted (UHF) Hartree-Fock wave function. In the restricted cases, the occupied pairs are formed from electrons sharing a valence spatial orbital. In the unrestricted case, we identify the pair space using the corresponding orbital transformation to maximize the overlap of each alpha orbital with one beta orbital. The corresponding orbitals are formed by diagonalizing the alpha-beta spatial orbital overlap matrix,

$D_{ij}^{\alpha\beta} = \langle \phi_i^\alpha | \phi_j^\beta \rangle$, via the singular value decomposition,

$$D^{\alpha\beta} = (C^\alpha)^\dagger S C^\beta = U \Sigma V^\dagger, \tag{5.6}$$

where S is the atomic basis overlap matrix. Given the orthogonal transformations U and V, the corresponding orbitals $\tilde{C}$ are obtained as,

$$\tilde{C}^\alpha = C^\alpha U \qquad \tilde{C}^\beta = C^\beta V. \tag{5.7}$$

To ensure the selection of valence orbitals for the active space, the corresponding orbital transformation is performed separately in the core and active spaces. In all cases, the orbitals are then localized separately in each subspace. The localized alpha and beta orbitals are paired by maximum overlap. These alpha/beta occupied pair orbitals are then paired with virtual orbitals by maximizing the exchange overlap of the virtual with each occupied according to the procedure proposed by Sano.[145] The Sano algorithm is used separately for the alpha and beta components of each occupied pair and may draw in parts of the SUMO in forming the pairs. As described above, alpha SOMOs are not paired in any fashion and comprise a third alpha occupied subspace, leading to core, active, inactive SOMO subspaces.

5.2.3 Unrestricted energy evaluation and orbital optimization

For a given set of orbitals, the correlation energy is determined by a linear number of t amplitudes according to,

$$E_{PP} = E_{ref} + \sum_i^{pairs} t_i \langle i\bar{i} | i^* \bar{i}^* \rangle. \tag{5.8}$$

In effect, this means we treat unpaired, excess alpha electrons in a UHF fashion while correlating all the valence electron pairs. The advantage of the PP model not shared by any of the more complicated coupled cluster methods described above[46, 17, 127, 76, 106, 105] is that the amplitude equations completely decouple, allowing for the analytical solution of each amplitude via the solution of a quadratic equation

(choosing the root that gives the lowest energy):

$$\langle i\bar{\imath}|i^*\bar{\imath}^*\rangle + W_i t_i - \langle i\bar{\imath}|i^*\bar{\imath}^*\rangle t_i^2 = 0, \tag{5.9}$$

where,

$$W_i = f_{i^*i^*} + f_{\bar{\imath}^*\bar{\imath}^*} - f_{ii} - f_{\bar{\imath}\bar{\imath}} + \langle i\bar{\imath}|i\bar{\imath}\rangle + \langle i^*\bar{\imath}^*|i^*\bar{\imath}^*\rangle + \langle ii^*|i^*i\rangle - \langle ii^*|ii^*\rangle$$

$$+ \langle \bar{\imath}\bar{\imath}^*|\bar{\imath}^*\bar{\imath}\rangle - \langle \bar{\imath}\bar{\imath}^*|\bar{\imath}\bar{\imath}^*\rangle - \langle i\bar{\imath}^*|i\bar{\imath}^*\rangle - \langle \bar{\imath}i^*|\bar{\imath}i^*\rangle. \tag{5.10}$$

In this notation, the f_{pp} are diagonal elements of the fock matrix and $\langle pq|rs\rangle = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_p(\mathbf{r}_1)\phi_q(\mathbf{r}_2)\frac{1}{r_{12}}\phi_r(\mathbf{r}_1)\phi_s(\mathbf{r}_2)$. The λ amplitudes are also trivially obtainable as,

$$\lambda_i = -\frac{\langle i\bar{\imath}|i^*\bar{\imath}^*\rangle}{W_i - 2t_i\langle i\bar{\imath}|i^*\bar{\imath}^*\rangle}. \tag{5.11}$$

Thus, the energy and amplitudes can be obtained with only a linear number of molecular orbital (MO) basis two-electron integrals.

Having obtained the t and λ amplitudes, the coupled cluster Lagrangian can be written down according to Eq. 5.5, as presented in Table 5.1. Differentiating this Lagrangian with respect to a rotation Δ_q^p between orbitals p and q gives,

$$\frac{dL}{d\Delta_q^p} = \sum_\mu \left(\frac{\partial L}{\partial C_\mu^p} C_\mu^q - \frac{\partial L}{\partial C_\mu^q} C_\mu^p \right) \tag{5.12}$$

where C_μ^p is the molecular orbital coefficient for atomic orbital μ in molecular orbital p. In order to compute the gradient with respect to orbital rotations, we therefore need only the partial derivatives of the Lagrangian with respect to the molecular orbital coefficients, which requires a partial set of two-center integrals. For the unrestricted case, these derivatives must be evaluated separately for rotations in the alpha and beta subspaces. These derivatives are also listed in Table 5.1. Convergence of the orbital optimization procedure is enhanced by utilizing diagonal second derivatives in the geometric direct minimization (GDM) procedure.[146] In the interest of brevity, these second derivatives are not presented here. Computationally they require additional two-center integrals of the forms $\langle pq|pq\rangle$ and $\langle pq|qp\rangle$ where p is in the occupied or active virtual spaces, and q is any orbital. Because of the simple structure of the PP equations, computing the energy, amplitudes, and orbital derivatives is trivial

Table 5.1: The unrestricted perfect pairing Lagrangian and its orbital derivatives.

$$
\begin{aligned}
L_{PP} &= E_{ref} + \sum_i^{pairs} \lambda_i t_i \left(f_{i^*i^*} + f_{\bar{i}^*\bar{i}^*} - f_{ii} - f_{\bar{i}\bar{i}} \right) + \sum_i^{pairs} (t_i + \lambda_i - \lambda_i t_i^2)\langle i\bar{i}|i^*\bar{i}^*\rangle \\
&+ \sum_i^{pairs} \lambda_i t_i \Big(\langle i\bar{i}|i\bar{i}\rangle + \langle i^*\bar{i}^*|i^*\bar{i}^*\rangle + \langle ii^*|i^*i\rangle - \langle ii^*|ii^*\rangle + \langle \bar{i}\bar{i}^*|\bar{i}^*\bar{i}\rangle - \langle \bar{i}\bar{i}^*|\bar{i}\bar{i}^*\rangle \\
&\quad - \langle i\bar{i}^*|i\bar{i}^*\rangle - \langle \bar{i}i^*|\bar{i}i^*\rangle \Big)
\end{aligned}
$$

$$
\text{Orbital Gradient} \quad \frac{dL}{d\Delta_q^p} = \sum_\mu \left(\frac{\partial L}{\partial C_\mu^p} C_\mu^q - \frac{\partial L}{\partial C_\mu^q} C_\mu^p \right)
$$

$l \in$ inact. occ.
$$
\begin{aligned}
\frac{\partial L}{\partial C_\mu^l} =& \ 2f_{\mu l} \\
&- \sum_i^{pairs} 2\lambda_i t_i \Big(\langle i\mu|il\rangle - \langle i\mu|li\rangle + \langle \bar{i}\mu|\bar{i}l\rangle \\
&\quad - \langle i^*\mu|i^*l\rangle + \langle i^*\mu|li^*\rangle - \langle \bar{i}^*\mu|\bar{i}^*l\rangle \Big)
\end{aligned}
$$

$k \in$ act. occ.
$$
\begin{aligned}
\frac{\partial L}{\partial C_\mu^k} =& \ - \sum_i^{pairs} 2\lambda_i t_i \Big(\langle i\mu|ik\rangle - \langle i\mu|ki\rangle + \langle \bar{i}\mu|\bar{i}k\rangle \\
&\quad - \langle i^*\mu|i^*k\rangle + \langle i^*\mu|ki^*\rangle - \langle \bar{i}^*\mu|\bar{i}^*k\rangle \Big) \\
&+ (2 - 2\lambda_k t_k) f_{\mu k} + (t_k + \lambda_k - \lambda_k t_k^2)\langle \mu\bar{k}|k^*\bar{k}^*\rangle \\
&+ 2\lambda_k t_k \Big(\langle \mu\bar{k}|k\bar{k}\rangle + \langle \mu k^*|k^*k\rangle - \langle \mu k^*|kk^*\rangle - \langle \mu\bar{k}^*|k\bar{k}^*\rangle \Big)
\end{aligned}
$$

$k^* \in$ act. virt.
$$
\begin{aligned}
\frac{\partial L}{\partial C_\mu^{k^*}} =& \ 2\lambda_k t_k f_{\mu k^*} + (t_k + \lambda_k - \lambda_k t_k^2)\langle k\bar{k}|\mu\bar{k}^*\rangle \\
&+ 2\lambda_k t_k \Big(\langle \mu\bar{k}^*|k^*\bar{k}^*\rangle + \langle k\mu|k^*k\rangle - \langle k\mu|kk^*\rangle - \langle \bar{k}\mu|\bar{k}k^*\rangle \Big)
\end{aligned}
$$

$l^* \in$ inact. virt.
$$
\frac{\partial L}{\partial C_\mu^{l^*}} = \ 0
$$

compared to the time necessary to compute and transform the integrals. Thus, in the next section, we outline an efficient approach for computing the requisite molecular integrals utilizing the resolution of the identity approximation.

5.2.4 The resolution of the identity approximation

The primary algorithmic difference between the closed-shell PP code described previously[132] and the implementation here is the adaptation of the resolution of the identity (RI) or density-fitting approximation. In this approximation, a larger, higher-angular momentum auxiliary basis set $\{|K)\}$ is used to expand products of two Gaussian basis functions that typically occur in two-electron integrals,

$$|\rho) = |\mu\nu) \approx |\overline{\mu\nu}) = \sum_K^{aux} A_{\mu\nu}^K |K).$$ (5.13)

In this notation, $\mu, \nu, \ldots$ refer to atomic orbitals (AOs), $K, L, \ldots$ refer to auxiliary basis functions, $i, j, \ldots$ refer to occupied molecular orbitals (MOs), and $a, b, \ldots$ refer to virtual MOs. The fitting coefficients $A_{\mu\nu}^K$ are given by,

$$A_{\mu\nu}^K = \sum_L^{aux} (\mu\nu|L)(L|K)^{-1}.$$ (5.14)

A typical two-electron integral $(\mu\nu|\lambda\sigma) = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_\mu(\mathbf{r}_1)\phi_\lambda(\mathbf{r}_2)\frac{1}{r_{12}}\phi_\nu(\mathbf{r}_1)\phi_\sigma(\mathbf{r}_2)$ takes the form

$$\begin{aligned}
(\mu\nu|\lambda\sigma) \approx (\overline{\mu\nu}|\overline{\lambda\sigma}) &= \sum_{KLMN}^{aux} (\mu\nu|K)(K|L)^{-1}(L|M)(M|N)^{-1}(N|\lambda\sigma) & (5.15)\\
&= \sum_{KL}^{aux} (\mu\nu|K)(K|L)^{-1}(L|\lambda\sigma) & (5.16)\\
&= \sum_{KLM}^{aux} \left[(\mu\nu|K)(K|M)^{-\frac{1}{2}}\right]\left[(M|L)^{-\frac{1}{2}}(L|\lambda\sigma)\right] & (5.17)\\
&= \sum_M^{aux} B_{\mu\nu}^M B_{\lambda\sigma}^M, & (5.18)
\end{aligned}$$

where $B_{\mu\nu}^M = \sum_K^{aux}(\mu\nu|K)(K|M)^{-\frac{1}{2}}$.

Although there are typically several times as many auxiliary basis functions $\{|K)\}$ as there are primary AO basis functions, the approximation requires only three-center

integrals to be built via explicit integration, and the final four-center integrals are formed as matrix multiplies of the $B_{\mu\nu}^{K}$ matrices. It also facilitates the development of cubic disk storage algorithms, since all two-electron integrals can be formed as needed from the $B_{\mu\nu}^{K}$ matrices.

The details of our algorithm will be presented elsewhere, but the basic procedure is as follows. Following the approach outlined in Ref. [132], we wish to construct half-transformed coulomb- and exchange-like matrices,

$$J_{\mu\nu}^{[ii]} = \langle i\mu|i\nu\rangle = (ii|\mu\nu) \qquad K_{\mu\nu}^{[ii]} = \langle i\mu|\nu i\rangle = (i\mu|i\nu) \qquad (5.19)$$

$$J_{\mu\nu}^{[ii^*]} = \langle i\mu|i^*\nu\rangle = (ii^*|\mu\nu) \qquad K_{\mu\nu}^{[ii^*]} = \langle i\mu|\nu i^*\rangle = (i\mu|i^*\nu) \qquad (5.20)$$

$$J_{\mu\nu}^{[i^*i^*]} = \langle i^*\mu|i^*\nu\rangle = (i^*i^*|\mu\nu) \qquad K_{\mu\nu}^{[i^*i^*]} = \langle i^*\mu|\nu i^*\rangle = (i^*\mu|i^*\nu) \qquad (5.21)$$

and then transform the final two AO indices. In this case, the half-transformed integrals are constructed in an RI-fashion. That is, we:

1. Form $(M|L)^{-\frac{1}{2}}$ and $(\mu\nu|M)$. Contract to form $B_{\mu\nu}^{L} = \sum_M (\mu\nu|M)(M|L)^{-\frac{1}{2}}$. This needs to be done only once, and the resulting AO basis $B_{\mu\nu}^{L}$ can be stored on disk. All subsequent steps must be updated each iteration.

2. Transform first AO index μ to the MO basis for orbitals i and i^*, $e.g.$ $B_{i\nu}^{L} = \sum_{\mu} C_{\mu i} B_{\mu\nu}^{L}$.

3. Form $K_{\mu\nu}^{[ii]} = \sum_L B_{i\mu}^{L} B_{i\nu}^{L}$, along with $K_{\mu\nu}^{[ii^*]}$ and $K_{\mu\nu}^{[i^*i^*]}$.

4. Transform second AO index to the MO basis,$e.g.$ $B_{ii^*}^{L} = \sum_{\mu} C_{\nu i^*} B_{i\nu}^{L}$.

5. Form $J_{\mu\nu}^{[ii]} = \sum_L B_{ii}^{L} B_{\mu\nu}^{L}$, along with $J_{\mu\nu}^{[ii^*]}$ and $J_{\mu\nu}^{[i^*i^*]}$.

6. Transform the final to AO indices to form $J_{pp}^{[ii]}$, $K_{pp}^{[ii]}$, etc., where p is any MO. Alternatively, form the necessary three-quarter transformed integrals used in the orbital derivatives listed in Table 5.1.

All steps in this procedure, including the final formation of the half-transformed integrals, are performed in batches to maintain quadratic memory. This procedure

generates all the integrals necessary to compute the energy, the first orbital derivatives, and the diagonal second orbital derivatives. Finally, note that the Sano initial guess algorithm requires the formation of many exchange matrices $K_{\mu\nu}^{[ii]}$, which is also performed using the RI approximation as described above.

Overall, this algorithm is quartic, albeit with a small prefactor, requires cubic disk, and quadratic memory. Calculations on fifty carbon linear alkanes in the cc-pVDZ basis (more than 1200 basis functions in the primary basis)[64] are quite feasible on a modern personal computer. Further details on the algorithm, timings, and errors introduced by the RI approximation are beyond the scope of this paper and will be addressed elsewhere.[60]

5.2.5 Nuclear gradient

An efficient algorithm for nuclear gradients allows molecular geometries to be computed at the UPP level. Because the Lagrangian has been variationally minimized during the energy calculation, the Hellman-Feynman theorem permits an equation for its gradient to be obtained trivially. If $^{(x)}$ indicates partial differentiation with respect to nuclear displacement, then the Lagrangian becomes:

$$
\begin{aligned}
L_{PP}^{(x)} = E_{ref}^{(x)} \ &+ \ \sum_i^{pairs} \lambda_t t_i \left(f_{i^*i^*}^{(x)} + f_{\bar{\imath}^*\bar{\imath}^*}^{(x)} - f_{ii}^{(x)} - f_{\bar{\imath}\bar{\imath}}^{(x)} \right) \\
&+ \ \sum_i^{pairs} (t_i + \lambda_i - \lambda_i t_i^2) \langle i\bar{\imath}|i^*\bar{\imath}^* \rangle^{(x)} \\
&+ \ \sum_i^{pairs} \lambda_i t_i \Big(\langle i\bar{\imath}|i\bar{\imath} \rangle^{(x)} + \langle i^*\bar{\imath}^*|i^*\bar{\imath}^* \rangle^{(x)} + \langle ii^*|i^*i \rangle^{(x)} - \langle ii^*|ii^* \rangle^{(x)} \\
&\quad + \langle \bar{\imath}i^*|\bar{\imath}^*i \rangle^{(x)} - \langle \bar{\imath}i^*|\bar{\imath}i^* \rangle^{(x)} - \langle i\bar{\imath}^*|i i^* \rangle^{(x)} - \langle \bar{\imath}i^*|\bar{\imath}i^* \rangle^{(x)} \Big) \quad (5.22)
\end{aligned}
$$

Although the PP wave function is stable with respect to infinitesimal changes in the nuclear positions, the MO coefficients must vary to ensure that the orbitals remain orthogonal when the (atom-centered) basis functions are moved. Differentiating the MO coefficients implicit in Equation 5.22 yields a term in which an energy weighted density matrix, W, contracts with the overlap derivatives. [147, 148]

Our strategy for evaluating the two electron gradients efficiently is the same as before.[132] Half-transformed coulomb and exchange matrices are contracted with the set of effective density matrices listed in Table 5.2.5. Note that the effective density matrices used in Ref [132] have been redefined to accommodate the unrestriction. The resulting expression for the gradient is:

$$
\begin{aligned}
L_{PP}^{(x)} = \quad & E_{ref}^{(\tilde{x})} \\
+ \quad & \sum_i^{pairs} \lambda_i t_i \left(\quad f_{i^*i^*}^{(\tilde{x})} + f_{\bar{i}^*\bar{i}^*}^{(\tilde{x})} - f_{ii}^{(\tilde{x})} - f_{\bar{i}\bar{i}}^{(\tilde{x})} \right) \\
+ \quad & \sum_i^{pairs} \sum_{\mu\nu} \left(\quad P_{\mu\nu}^{[ii]} J_{\mu\nu}^{[ii](x)} + P_{\mu\nu}^{[\bar{i}\bar{i}]} J_{\mu\nu}^{[\bar{i}\bar{i}](x)} \right. \\
& \quad -Q_{\mu\nu}^{[ii]} K_{\mu\nu}^{[ii](x)} - Q_{\mu\nu}^{[\bar{i}\bar{i}]} K_{\mu\nu}^{[\bar{i}\bar{i}](x)} \\
& \quad +P_{\mu\nu}^{[ii^*]} J_{\mu\nu}^{[ii^*](x)} + P_{\mu\nu}^{[\bar{i}\bar{i}^*]} J_{\mu\nu}^{[\bar{i}\bar{i}^*](x)} \\
& \quad +P_{\mu\nu}^{[i^*i^*]} J_{\mu\nu}^{[i^*i^*](x)} + P_{\mu\nu}^{[\bar{i}^*\bar{i}^*]} J_{\mu\nu}^{[\bar{i}^*\bar{i}^*](x)} \\
& \quad \left. -Q_{\mu\nu}^{[i^*i^*]} K_{\mu\nu}^{[i^*i^*](x)} - Q_{\mu\nu}^{[\bar{i}^*\bar{i}^*]} K_{\mu\nu}^{[\bar{i}^*\bar{i}^*](x)} \right) \\
+ \quad & \sum_{\mu\nu} W_{\mu\nu} S_{\mu\nu}^{(x)},
\end{aligned}
\tag{5.23}
$$

where $^{(\tilde{x})}$ indicates partial differentiation of the one and two electron integrals, but not the MO coefficients.

The time necessary to compute these derivatives is comparable to that required for a single point energy evaluation. Batching the formation of the derivative matrices enables the gradient to be computed with only quadratic memory use. The speed of the gradient calculation is improved further by an RI implementation which will be described elsewhere.[60]

5.2.6   Orbital optimization convergence

In typical closed-shell organic molecules (i.e. with single bonds connecting the different atoms) near their equilibrium structures, the PP algorithm described here and in Ref. [132] converges within 10-30 iterations, which is comparable to the convergence rate of HF. However, in more complicated cases like radicals, the convergence is much slower and can take hundreds of iterations (though with GDM it always converges eventually). Moreover, just as different initial guesses often lead to different

Table 5.2: The derivative matrices and effective density matrices used for computing the nuclear gradient for the unrestricted perfect pairing Lagrangian. For the coulomb and exchange derivative matrices, differentiation has been intentionally limited to the external indices.

$$J_{\mu\nu}^{[ii](x)} = \langle i\mu^{(x)}|i\nu\rangle + \langle i\mu|i\nu^{(x)}\rangle$$

$$K_{\mu\nu}^{[ii](x)} = \langle i\mu^{(x)}|\nu i\rangle + \langle i\mu|\nu^{(x)}i\rangle$$

$$J_{\mu\nu}^{[ii^*](x)} = \langle i\mu^{(x)}|i^*\nu\rangle + \langle i\mu|i^*\nu^{(x)}\rangle$$

$$K_{\mu\nu}^{[ii^*](x)} = \langle i\mu^{(x)}|\nu i^*\rangle + \langle i\mu|\nu^{(x)}i^*\rangle$$

$$J_{\mu\nu}^{[i^*i^*](x)} = \langle i^*\mu^{(x)}|i^*\nu\rangle + \langle i^*\mu|i^*\nu^{(x)}\rangle$$

$$K_{\mu\nu}^{[i^*i^*](x)} = \langle i^*\mu^{(x)}|i^*\nu\rangle + \langle i^*\mu|i^*\nu^{(x)}\rangle$$

$$P_{\mu\nu}^{[ii]} = \lambda_i t_i(C_{\mu i}C_{\nu i} - C_{\mu i^*}C_{\nu i^*} - C_{\mu\bar{i}^*}C_{\nu\bar{i}^*})$$

$$Q_{\mu\nu}^{[ii]} = \lambda_i t_i C_{\mu i^*}C_{\nu i^*}$$

$$P_{\mu\nu}^{[ii^*]} = (\lambda_i + t_i - \lambda_i t_i^2)C_{\mu\bar{i}}C_{\nu\bar{i}^*}$$

$$P_{\mu\nu}^{[i^*i^*]} = \lambda_i t_i(C_{\mu\bar{i}^*}C_{\nu\bar{i}^*} - C_{\mu i}C_{\nu i} - C_{\mu\bar{i}}C_{\nu ib})$$

$$Q_{\mu\nu}^{[i^*i^*]} = \lambda_i t_i C_{\mu i}C_{\nu i}$$

$$W_{\mu\nu}^{\alpha} = -\frac{1}{2}\sum_{i,j}^{occ,act\alpha}\sum_{\sigma}C_{\mu i}\frac{\partial L}{\partial C_{\sigma i}}C_{\sigma j}C_{j\nu}^T$$

SCF solutions in more challenging species, we often find multiple PP solutions depending on the initial localization scheme used. In general, Pipek-Mezey orbitals[26] provide the fastest convergence, but sometimes the Boys orbitals[25] lead to lower energy solutions.

In order to use PP widely, it might be helpful to adapt Pulay's direct inversion of the iterative subspace (DIIS) to the PP problem[149, 150] to provide an alternative convergence scheme that could be faster when starting far away from the orbital minimum. Also, an implementation of complete orbital second derivatives to allow for stability analysis on converged stationary points in certain cases might also be helpful. Perhaps other initial guesses than the Sano guess from HF orbitals could be devised as well. Nevertheless, we were able to converge every molecule and made a serious effort (by varying the initial guess) to ensure that the variationally lowest energy solution was found for all of the results reported herein.

5.3   Results and discussion

The RI-PP algorithm for open- and closed-shell systems has been implemented in a developmental version of Q-Chem, and all calculations herein were performed using Q-Chem.[85] Unless otherwise specified, all PP calculations use the RI approximation with the auxiliary basis sets developed for RIMP2/cc-pVXZ calculations,[151, 64] and all valence orbitals were correlated. The active spaces in VOD always match those used in PP, with one correlating orbital for each pair of valence electrons. Furthermore, unless otherwise specified, all calculations were spin unrestricted and break symmetry whenever it lowers the energy.

5.3.1   Bond-breaking

To begin, we revisit the bond-breaking problem in the context of the N_2 molecule dissociating to two quartet state nitrogen atoms. Figure 5.1 compares HF, PP and VOD against FCI results (with the frozen core approximation)[152] in the cc-pVDZ basis. It comes as no surprise that RHF rises in energy much too quickly and dis-
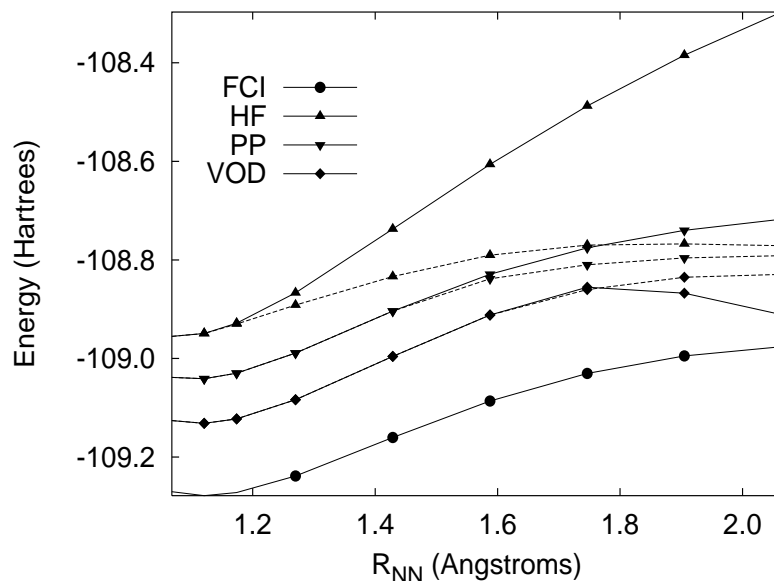
Figure 5.1: Restricted (solid lines) and unrestricted (dotted lines) N_2 bond breaking in the cc-pVDZ basis. The correlated methods employ a frozen core approximation, and the FCI results were obtained from Ref [3].

sociates to a high asymptote, leading to a PES that very non-parallel to the FCI one. RPP is significantly more parallel to FCI than RHF and does not exhibit the non-variational collapse associated with standard restricted MP2 or coupled-cluster doubles methods. For example, restricted VOD turns over around 1.75 Å. However, this parallelism of RPP does come at a cost: the dissociation limit is still too high as compared to a PP description of the atoms. Unrestricting enables the wave function to reach localized atomic limits, with the alpha electron in a bonding pair localizing to one atom and the beta electron to the other.

As we dissociate, the five valence pairs (1 σ bond, 2 π bonds and 2 lone pairs at equilibrium) should asymptote to two equivalent $2s$ pairs with non-zero correlation energy contribution and three amplitudes with magnitude zero corresponding to the non-existent interaction between the three alpha unpaired $2p$ electrons on one N atom and the three beta $2p$ unpaired electrons on the other atom at infinite separation. However, because the PP orbitals are variationally optimized and we are considering

five pairs, the energy can be lowered by correlating the $1s$ core electrons together or with the unpaired $2p$ electrons. Though the correlation energy contributions from these interactions are small, they are non-zero. This means that the wave function will change character discontinuously on the way to dissociation unless either all electrons are made active or the core electrons are not allowed to mix with the valence ones.

Note, however, that this issue of the variationally best active space changing qualitatively across the potential energy surface is in principle present in all variationally orbital-optimized active space methods, such as VOD or CASSCF. Of course, the more inter-pair correlations present in the model and/or the larger the active space, the less likely that "core" orbitals will be drawn into the active space. In N_2, for example, a more complete treatment of the correlations would correlate the unpaired electrons in quartet nitrogen atom, thereby obviating core electron correlations.

In this case, given the relative unimportance of the core $1s$ electrons on each nitrogen atom, we have frozen those orbitals (based on their HF definitions) and prevented them from mixing with the active orbitals in PP and VOD. The FCI data also employs the frozen core approximation. Freezing the core orbitals in PP instead of leaving them as inactive provides a minimal energetic penalty (less than 0.1 kcal/mol across the regime examined here) and ensures that our pairs remain valence in nature.

Looking at the UPP curve, we see both that the unrestriction occurs noticeably later than for UHF and that the difference between the restricted and unrestricted solutions is much less significant at the PP level. Not too surprisingly, VOD maintains spin symmetry even further away from equilibrium due to its more complete description of the correlations.[93] Notice too that UHF overshoots the atomic limits and has turned over slightly by 2.0 Å. In contrast, UPP and UVOD rise monotonically toward the dissociation limit for N_2. Because there is less correlation energy in the separate atoms than the molecule, the asymptotic limits of UHF, UPP, and UVOD are all much closer in energy than at the equilibrium geometry. This leads to an increased non-parallelarity error versus FCI, which in the case of UPP, is worse than RPP. Nevertheless, UPP is much more parallel than UHF, and its correct asymptotic behavior of UPP should be helpful in combination with the inclusion of dynamical correlation effects. Overall, UPP captures much of the energetic benefit of UVOD at

much lower cost.

The issue of unusual unrestricted pairing leads to especially odd behavior in F_2^+. Unlike UHF in the 6-31G* basis, which predicts that F_2^+ is unbound, UPP successfully predicts F_2^+ to be bound, though by only 17 kcal/mol (versus 48 kcal/mol at the VOD level). The PP PES minimum and curvature are also somewhat in error, leading to an erroneous bond length and harmonic frequency, as will be discussed below. More importantly in the current context, as the bond is stretched to about 1.8 Å, the standard ground state,

$$\Psi_{^2\Pi_g} = \cdots (\sigma_g)^2 (\pi_u)^4 (\pi_g^*)^3, \qquad (5.24)$$

crosses with an excited state with configuration,

$$\Psi = \cdots (\sigma_g)^2 (\pi_u)^4 (\pi_g^*)^2 (\sigma_u^*). \qquad (5.25)$$

This second state is bound by about 9 kcal/mol at its optimal bond length of 2.01 Å. Thus, the adiabatic ground state has a second, unphysical minimum. This second state crosses the ground state when it becomes energetically favorable to form a pair with $(\sigma_g \beta)(\sigma_u^* \alpha)$ instead of the conventional $(\sigma_g)^2$ pair. Fortunately, F_2^+ seems to be an unusually severe case, and problems of this extent have not been observed in other systems. Of course, in the broader context, where UHF does not even predict a bound state in this and many other basis sets, this failure on the part of PP 0.4 Å from equilibrium is more acceptable.

5.3.2 Geometries of open- and closed-shell species

Because molecular geometries generally do not depend too much on the description of the correlation energy, one of the best uses for HF is molecular structure prediction. However, it is known that the absence of correlation typically leads HF to underestimate bond lengths. To this end, we can assess to what extent including a limited description of pair correlations improves these structures. Therefore, we assess the reliability of PP in predicting open- and closed-shell geometries. We shall consider the two types of molecules separately in the 6-311G** basis. In this particular instance, we did not utilize the RI approximation for PP, though it introduces only

Table 5.3: Errors in predicted bond-lengths (in Å) versus experiment in the 6-311G**
basis for various small closed-shell molecules. All methods are unrestricted, and the
RI approximation was not used for PP.

		Expt[a]	HF	PP	VOD	B3LYP
C_2H_2	rCH	1.063	-0.008	0.008	0.013	0.000
	rCC	1.203	-0.020	-0.001	0.014	-0.005
C_2H_4	rCH	1.081	-0.004	0.013	0.019	0.004
	rCC	1.334	-0.018	0.011	0.015	-0.007
$^1CH_2$	rCH	1.107	-0.022	-0.002	0.023	-0.010
CH_4	rCH	1.086	-0.002	0.018	0.019	0.005
Cl_2	rClCl	1.988	0.015	0.095	0.016	0.069
ClF	rClF	1.628	-0.009	0.091	0.041	0.051
CO	rCO	1.128	-0.023	-0.010	0.002	-0.001
CO_2	rCO	1.160	-0.025	-0.014	0.001	0.000
CS	rCS	1.535	-0.019	-0.007	0.004	0.007
F_2	rFF	1.412	[b]	0.131	0.008	-0.005
H_2	rHH	0.741	-0.006	0.015	0.015	0.001
H_2CO	rCH	1.116	-0.021	-0.009	0.003	-0.006
	rCO	1.208	-0.029	0.002	0.000	-0.008
H_2CS	rCH	1.093	-0.015	0.003	0.009	-0.004
	rCS	1.611	0.017	0.021	0.013	0.005
H_2O	rOH	0.957	-0.016	0.002	0.004	0.005
H_2O_2	rOH	0.965	-0.023	-0.003	0.001	-0.003
	rOO	1.452	-0.059	0.043	0.004	0.009
H_2S_2	rSH	1.345	-0.016	0.007	0.000	0.002
	rSS	2.058	0.066	0.125	0.073	0.116
HCN	rCH	1.065	-0.007	0.007	0.013	0.002
	rCN	1.153	-0.026	-0.002	0.008	-0.005
HCP	rCH	1.069	-0.004	0.011	0.017	0.004
	rCP	1.540	0.019	0.009	0.016	0.000
HF	rHF	0.917	-0.021	-0.003	-0.002	0.005
HNC	rNH	0.994	-0.010	0.004	0.011	0.006
	rNC	1.169	-0.020	-0.005	0.007	-0.001
HNO	rHN	1.063	-0.037	-0.016	0.001	-0.002
	rNO	1.212	-0.021	0.001	-0.003	-0.009
HOCl	rOH	0.975	-0.031	-0.011	-0.008	-0.007
	rOCl	1.690	-0.017	0.085	0.036	0.047
HOF	rOH	0.966	-0.020	0.000	0.004	0.005
	rOF	1.442	-0.080	0.046	-0.014	-0.008
Li_2	rLiLi	2.673	0.257	0.016	0.016	0.032

Table 5.3: (continued)

		Expt[a]	HF	PP	VOD	B3LYP
N_2	rNN	1.098	-0.028	-0.002	0.006	-0.003
N_2H_2	rNH	1.028	-0.015	0.003	0.012	0.007
	rNN	1.252	-0.040	-0.003	0.002	-0.011
NaF	rNaF	1.926	-0.014	-0.008	0.003	-0.009
NH_3	rNH	1.012	-0.011	0.008	0.010	0.003
NNO	rNN	1.127	-0.042	-0.024	-0.003	-0.001
	rNO	1.185	-0.016	0.017	0.012	-0.001
NP	rNP	1.491	0.038	0.005	0.007	-0.002
NSF	rNS	1.448	-0.023	-0.011	-0.001	0.003
	rSF	1.643	-0.017	0.016	0.060	0.069
P_2	rPP	1.893	0.072	0.018	0.020	0.004
SO2	rSO	1.485	-0.078	-0.066	-0.041	-0.027
SO3	rSO	1.420	-0.024	-0.013	0.011	0.026
Median Error[c]			-0.017	0.004	0.009	0.000
RMS Error[c]			0.048	0.033	0.020	0.026

[a] Reference   [b] unbound   [c] excludes F_2

tiny errors into the predicted structures (typically on the order of $10^{-3} - 10^{-4}$Å). In all cases, unrestricted wave functions were used and symmetry was broken whenever possible. Unrestricted B3LYP results are presented as well, since it is probably the most-widely used method for molecular structure determination.

We consider first a group of closed-shell species containing first and second-row atoms. The test set consists of 34 molecules containing 49 unique bond lengths, and contains most of the species examined in references [121, 120] along with others found in references [153, 63, 87, 154]. The errors are presented in Table 5.3 and pictorially using box plots in Figure 5.2. These plots mark the median error with a white line inside the black box. The box then extends to include the central 50% of the data. The whiskers extend to include any data within a range extending up to 1.5 times the size of the box in each direction, and points lying beyond this range are denoted as outliers and are marked separately.

Looking at Figure 5.2, we see that, as expected, including some correlation tends to lengthen bonds slightly on average, shifting the median error (excluding F_2, for

which HF is unbound as discussed below) from -0.017 Å for HF to 0.004 Å for PP, 0.009 Å for VOD, and 0.000 Å for B3LYP. Indeed, the limited correlation in PP helps to correct for the underestimated bond lengths in HF. Also, the PP, VOD, and B3LYP root-mean-square (rms) errors of 0.033, 0.020, and 0.026 Å are noticeably improved over HF (0.048 Å). The largest improvements over HF with the correlated methods is observed for Li_2, which goes from 0.26 Å too long to 0.016 Å in error with PP or VOD, and F_2, which is not even bound at the HF level. Unfortunately, as observed previously [121, 120] and in Chapter 4, PP predicts much too long of a bond length for F_2 and other halogenated species, for which inter-pair correlations are important. Many of the species for which PP has the most difficulty, H_2S_2 ($\Delta r_{SS} = 0.125$ Å), Cl_2 (0.095 Å), ClF (0.091 Å), and HOCl($\Delta r_{OCl} = 0.085$ Å) are also problematic for VOD and B3LYP. In general, however, if one does not use PP on halogenated species, it predicts the geometries of these closed-shell species very reasonably, with only slightly larger errors than VOD. Although B3LYP is in general slightly more accurate than VOD for these structures, it exhibits some significant outliers that err more than their VOD counterparts.

One might expect that PP would have trouble in hyper-valent species, for which the electron pairing is perhaps more complicated than in standard octet-rule-obeying structures. The test set includes three hyper-valent species, NSF, SO_2, and SO_3. PP clearly makes improvements over HF on these species, and its geometry predictions for these species are not particularly worse than for the other species. On the other hand, we have not tested PP in systems such as transition-metal complexes which exhibit even more complicated bonding patterns.

Unrestricted PP also performs well for open-shell molecules. Table 5.4 and Figure 5.3 present UHF, UPP, and UVOD, and UB3LYP results for 29 diatomic and triatomic open-shell species containing 32 unique bonds.[155] Most of these species have doublet ground states, but BN, $^3CH_2$, NF, NH, O_2, and OH^+ exhibit triplet ground states. Overall, if we exclude F_2^+ (which UHF predicts to be unbound), PP reduces the rms error for the open-shelled species from 0.031 Å at the HF level to 0.015 Å. Furthermore, PP is statistically on par with VOD, which has an rms error of 0.014 Å, though for any given species the results differ moderately. The PP and
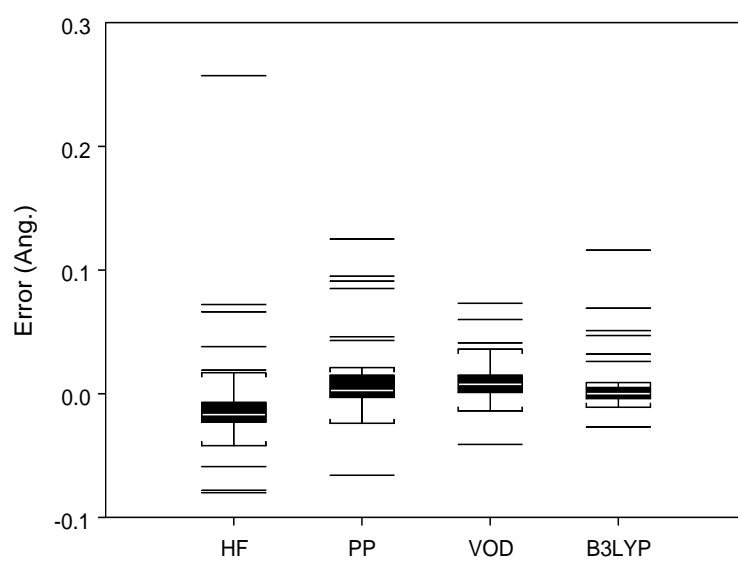
Figure 5.2: Errors in 6-311G** predicted bond lengths versus experiment for a set of small, closed-shell molecules.
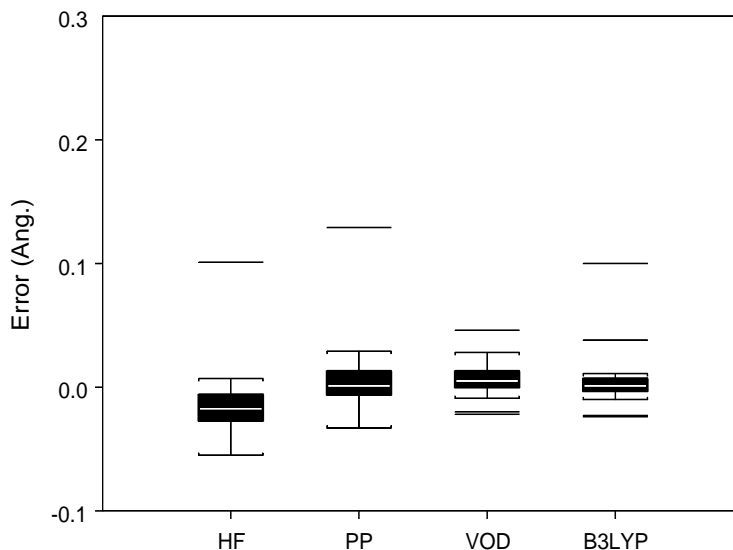
Figure 5.3: Errors in 6-311G** predicted bond lengths versus experiment for a set of small, doublet- and triplet-state open-shell molecules. Note: HF excludes F_2^+, since it is unbound.

VOD geometries are also slightly improved over B3LYP, which has an RMS error of 0.021 Å. As in the closed-shell species, the inclusion of static correlation increases the bond lengths versus HF and brings the median error closer to zero.

The largest improvements PP provides over HF are for C_2^+ and for F_2^+, for which UHF predicts no binding due to severe symmetry-breaking effects. On the other hand, F_2^+ is the only significant outlier in this data set for PP, with an error of 0.13 Å. This is not too surprising, however, since F_2^+, like F_2, requires a description of inter-pair correlations to be reliable, as discussed in Chapter 4.[120] The fact that VOD has no trouble with F_2^+ substantiates this reasoning. In any case, the fact that PP binds F_2^+ at all, unlike HF, is a notable success, even if the bond length is 10% too long.

Combining these two data sets, we see that overall PP predicts fairly reliable geometries for open- and closed-shelled systems (typically accurate to within a 0.01 Å) with some improvement over HF at much lower cost than VOD. However, the limited

Table 5.4: Errors in predicted bond-lengths (in Å) versus experiment in the 6-311G** basis for various small doublet- and triplet-state molecules. All methods are unrestricted, and the RI approximation was not used for PP.

		Expt[a]	HF	PP	VOD	B3LYP
BeF	rBeF	1.361	0.003	0.011	0.022	0.011
BeH	rBeH	1.343	0.000	0.025	0.025	0.000
BH^+	rBH	1.215	-0.027	-0.006	-0.009	-0.010
BH_2	rBH	1.181	0.004	0.026	0.028	0.007
BN	rBN	1.281	0.007	0.029	0.046	0.038
BO	rBO	1.205	-0.024	-0.011	0.002	-0.003
C_2^+	rCC	1.301	0.101	0.000	0.007	0.100
CF	rCF	1.272	-0.015	0.013	0.009	0.007
CH	rCH	1.120	-0.017	0.014	0.018	0.008
$^3CH_2$	rCH	1.078	-0.006	0.012	0.016	0.002
CH_2	rCH	1.079	-0.006	0.011	0.013	0.001
CN	rCN	1.172	-0.018	-0.006	0.005	-0.006
CO^+	rCO	1.115	-0.027	-0.013	0.000	-0.005
F_2^+	rFF	1.322	[b]	0.129	-0.022	-0.023
FH^+	rFH	1.001	-0.023	-0.006	-0.001	0.008
HCO	rHC	1.125	-0.016	-0.004	0.011	0.002
	rCO	1.175	-0.023	-0.002	0.005	-0.001
HNF	rHN	1.060	-0.047	-0.026	-0.020	-0.024
	rNF	1.370	-0.036	0.016	0.010	0.002
HO_2	rOH	0.977	-0.030	-0.013	-0.004	-0.002
	rOO	1.335	-0.036	0.023	0.002	-0.007
N_2^+	rNN	1.116	-0.008	0.006	0.006	-0.007
NF	rNF	1.317	-0.026	0.016	0.011	0.001
NH^+	rNH	1.070	-0.016	0.003	0.006	0.010
NH	rNH	1.036	-0.013	0.010	0.014	0.009
NH_2	rNH	1.024	-0.012	0.008	0.013	0.007
NO	rNO	1.151	-0.034	-0.008	0.002	-0.003
O_2^+	rOO	1.116	-0.055	-0.019	-0.006	-0.010
O_2	rOO	1.208	-0.051	-0.033	-0.003	-0.002
OH^+	rOH	1.028	-0.021	0.000	0.003	0.009
OH	rOH	0.970	-0.018	0.001	0.004	0.006
OH_2^+	rOH	0.999	-0.019	-0.001	0.005	0.007
Median Error[c]			-0.018	0.001	0.006	0.002
RMS Error[c]			0.031	0.015	0.014	0.021
RMS doublets[c]			0.032	0.013	0.012	0.021
RMS triplets[c]			0.026	0.020	0.021	0.016

[a] Reference   [b] unbound   [c] excludes F_2^+

description of correlation effects in PP is likely to cause trouble in systems with halogen atoms, multiple resonance structures, or other odd bonding patterns. In those cases, a more complete treatment of valence correlation that does not bias so strongly in favor or individual electron pairs, like the other models discussed in Chapter 4, is necessary. For typical systems, of course, B3LYP generally remains as accurate or better than VOD or PP. In complicated systems with very strong static correlation effects, however, density functional theory can fail miserably, and an inexpensive alternative like PP for structure prediction can be very useful.

5.3.3   Radicals and symmetry-breaking

Symmetry-breaking effects in radicals are known to produce spurious potential energy surfaces and properties. In the worst-case scenarios, such as F_2^+, UHF is unbound in certain basis sets (6-31G*, for example). Even if they are bound, symmetry-breaking leads to asymmetric electronic spin distributions, anomalous vibrational frequencies, etc. Byrd and co-workers demonstrated that standard methods like MP2 and CCSD(T) performed far below their standard, closed-shell system levels of accuracy on the geometries and vibrational frequencies of various small radicals.[66] It was later demonstrated in Chapter 2 that these problems are primarily linked to spin and spatial symmetry-breaking effects in the underlying HF reference, and that property predictions are substantially improved by improving upon the HF reference. In particular, this latter study demonstrated that methods like orbital-optimized coupled cluster doubles with perturbative tripled (OD(T)) or KS-CCSD(T), which uses Kohn-Sham orbitals as a more stable reference than HF, predicted vibrational frequencies of small diatomic radicals faithfully compared to experiment. In both cases, the new reference determinants are far more stable against symmetry-breaking.

Unfortunately, even KS-CCSD(T) is much too computationally expensive for applicability in systems beyond a few atoms, so it is desirable to explore to what extent the limited correlations included in PP can overcome these problems. With this in mind, we revisit the set of twelve diatomic radicals studied in Chapter 2 using unrestricted HF, PP, and VOD in the cc-pVTZ (and the auxiliary basis set for cc-pVTZ

Table 5.5: Harmonic vibrational frequencies (in cm^{-1}) for various diatomic radicals in the cc-pVTZ basis set as compared to experiment.

	Expt.	HF	PP	VOD	CCSD[a]
CH	2858.5^d	215	-59	-63	-6
OH	3737.8^c	301	-30	-11	-40
FH^+	3090^b	233	90	62	76
BO	1886^b	195	113	40	44
CN	2068.6^c	-54	19	31	89
CO^+	2169.8^c	239	187	68	124
N_2^+	2207^b	-326	-81	42	127
CF	1308^b	110	-5	-5	39
NO	1904.2^c	319	62	62	86
O_2^+	1904.7^c	591	153	123	128
OF	1053^b	157	-231	6	52
F_2^+	1104^d	479	-370	104	123
MAD (%)		14.7	7.7	2.8	4.4

[a] Chapter 2 [b] Reference [80]. [c] Reference [81].

[d] Reference [84].

in the PP case). These results are presented in Table 5.5. We see that for these species, HF is typically 50-600 cm^{-1} in error in this basis set, with the worst cases percentage-wise being F_2^+ (43%), O_2^+ (32%), NO (17%), and OF and N_2^+ (15%). The UPP model, on the other hand systematically improves virtually every one of these frequencies.

Some of the most difficult cases from previous studies are the isoelectronic pairs CO^+, CN, and N_2^+, which suffer from severe spin contamination, and O_2^+ and NO, and F_2^+ and OF, all of which suffer from spatial symmetry-breaking or rapid changes in the wave function for small displacements in the nuclei (for NO and OF, which technically do not have left-right symmetry). This explains the inclusion of most of these in the list above of the worst percentage errors at the HF level. In all of these cases except for F_2^+ and OF, PP predicts significantly more reasonable frequencies, dropping the percentage errors to below 10%. In most cases, PP even recovers the majority of the improvement offered by VOD (see particularly CH, OH, CN, CF, NO, and O_2^+), which makes no local approximation.

Table 5.6: Reference determinant $\langle S^2 \rangle$ for various diatomic radicals in the cc-pVTZ basis set.

	HF	PP
BO	0.794	0.776
CN	1.076	0.822
CO^+	0.908	0.807
N_2^+	1.163	0.794
OF	0.753	0.760
F_2^+	0.766	0.772

The worst frequencies (as compared against experiment percentage-wise) at the PP level are for BO (6%), CO^+ (9%), O_2^+ (8%), OF (22%), and F_2^+ (34%). The first three of these are all noticeably better than UHF and all correspond to systems with sizable spin-contamination at the UHF level. Table 5.6 lists all of the radicals from this set with UHF $\langle S^2 \rangle$ values greater than 0.77. For these doublet radicals the exact ground state should have $\langle S^2 \rangle = 0.75$. From this table, we see that UPP usually somewhat decreases the degree of spin contamination, as indicated by the lower $\langle S^2 \rangle$ values, particularly for the most spin-contaminated cases of CN, CO^+, and N_2^+. For all of the diatomics not listed in Table 5.6, UPP does reduce the spin-contamination, but it is already relatively minor at the UHF level and the improvement provided by UPP is not particularly significant. The sizable improvement in spin-contamination helps to explain the improvement in the computed frequencies for a number of species, particularly for BO, CN, CO^+, and N_2^+. These effects would become even more pronounced if dynamical correlation were included.

OF and F_2^+ are unique in that PP gives slightly larger reference $\langle S^2 \rangle$ values. Those two also stood out in their frequency predictions, with each being hundreds of wavenumbers in error. Furthermore, the OF frequency is actually worse at the UPP level than at the UHF level, though neither is particularly accurate. However, the difficulty for these isoelectronic species is a mixture of their being among the most difficult diatomic radicals in terms of symmetry-breaking with the general difficulty PP has treating halogens, for which inter-pair correlations are important (See Ref [120] or Chapter 4). On the other hand, even if UPP does not particularly improve upon HF

for these frequencies, it is worth noting once again that while UHF fails to bind F_2^+ in many basis sets (e.g. 6-31G*[78] or 6-311G**, as discussed above), UPP predicts a bound structure, albeit with only a qualitatively correct PES near equilibrium.

In summary, UPP substantially improves the predicted harmonic vibrational frequencies over UHF in these challenging radicals, cutting the mean-absolute percent deviation almost in half. Though UPP does not perform as well as VOD for these species, it is much less expensive to compute. The UPP wave function should then make for a significantly better reference for treating the dynamical correlation perturbatively and avoid the pathological frequency predictions characteristic of MP2 from the UHF reference.

Of course, the simplicity of the PP Ansatz dictates that it perform only moderately compared to more complete correlation treatments like VOD. Consider, for example, the allyl radical. The π system contains three electrons delocalized over the three carbons. This species exhibits the classic competition between the symmetric, delocalized three-electron π system with both carbon-carbon bonds equal in length and the symmetry-broken wave function that localizes the radical electron on one end of the molecule and the double bond on the other end. Restricted open-shell HF (ROHF) favors the symmetry-broken solution, but UHF actually prefers the symmetric one. The ROHF and UHF HOMO and SOMO orbitals for allyl at the symmetric, UB3LYP/cc-pVDZ optimized geometry are plotted in Figure 5.4. At this point on the PES, ROHF orbitals are symmetry-broken, with a slight shift in the HOMO towards the left carbon atom and the SOMO towards the right carbon atom. UHF, on the other hand, exhibits perfectly symmetrical orbitals. Therefore, correlation methods like MP2 that rely on the UHF reference will preserve the symmetry. In contrast, the open-shell PP model described here correlates the electron pair but treats the radical electron in a UHF-like fashion. This asymmetry acts as a driving force for the electron pair to localize and maximally separate itself from the radical electron, tipping the scales in favor of the symmetry-broken solution. The UPP orbitals are also plotted in Figure 5.4, and although the beta HOMO is delocalized reasonably over the molecule, the alpha HOMO and SOMO are extremely symmetry-broken. As expected, the electron pair localizes to a carbon-carbon double bond, and the radical

Table 5.7: Charge and spin symmetry Mulliken populations for the left and right atoms in selected diatomic radicals in the cc-pVTZ basis set.

	HF		PP	
	Charge	Spin	Charge	Spin
N_2^+	0.393/0.606	-0.817/1.817	0.434/0.566	0.130/0.870
O_2^+	0.500/0.500	0.500/0.500	0.498/0.502	0.494/0.506
F_2^+	0.500/0.500	0.500/0.500	0.702/0.298	-0.051/1.051

electron primarily occupies the other carbon atom.

A similar picture emerges energetically. Figure 5.5 plots the slice of the potential energy surface for deforming C-C bond from the UB3LYP/cc-pVDZ optimized symmetric structure (R(C-C) = 1.386 Å). All other degrees of freedom were held fixed. By slightly less than 2 kcal/mol, UPP favors the symmetry-broken structure with one C-C bond 0.1 Å longer than the other. In contrast, VOD, which correlates all three π electrons equivalently, maintains symmetry. Likewise, UHF favors the symmetric solution.

Revisiting the three homonuclear diatomics from our test set above, we do see that for F_2^+ and O_2^+, HF has not yet broken spatial symmetry at the PES minimum in this basis set, as shown by the Mulliken charge and spin population analysis results in Table 5.7. In contrast, PP is already heavily symmetry-broken for these species. This helps to explain the difficulty PP has with these species. On the other hand, for N_2^+, HF is far more symmetry-broken than PP, and PP behaves much better than HF. Overall, PP will sometimes help with symmetry-breaking phenomena, particularly spin contamination. However, it is clear that the asymmetry in the description of radical electrons and pairs will provide impetus for additional spatial symmetry-breaking, as in allyl radical.

5.4   Conclusions

Perfect-Pairing provides the leading correlation correction beyond Hartree-Fock at only slightly higher cost (some 3-5 times that of the HF calculation). We have extended the restricted coupled cluster Ansatz of PP to unrestricted wave functions

α SOMO

α HOMO

β HOMO

(a)   (b)   (c)
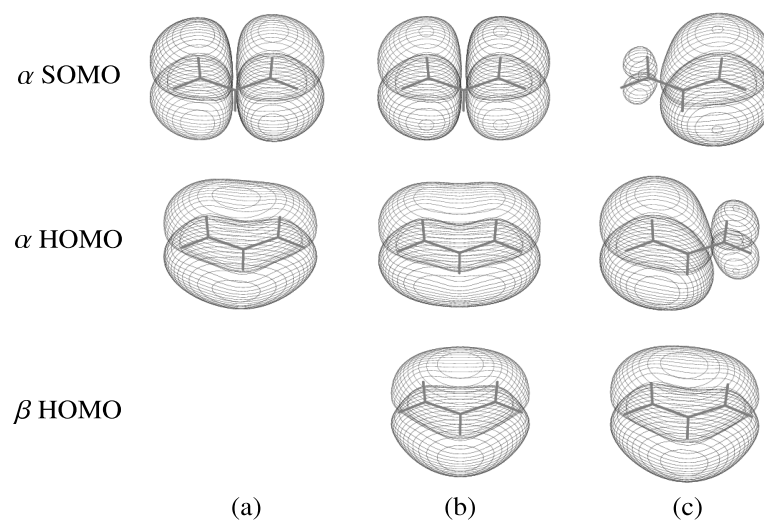
Figure 5.4: Occupied π molecular orbitals for allyl at the (a) ROHF, (b) UHF, and (c) UPP levels in the cc-pVDZ basis, using the symmetric UB3LYP/cc-pVDZ-optimized structure.
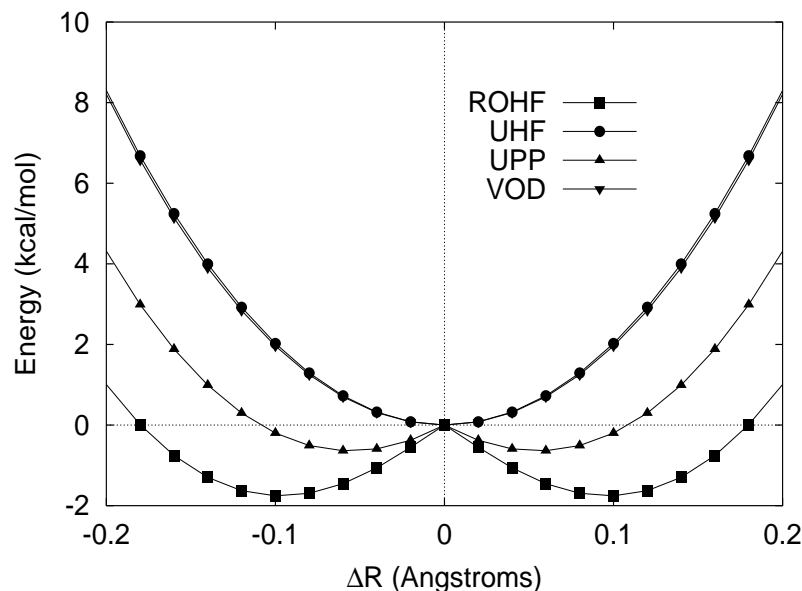
Figure 5.5: Symmetry-breaking in the allyl radical in the cc-pVDZ basis. The deformation ΔR is relative to the UB3LYP/cc-pVDZ optimized symmetric structure. The UHF and VOD curves are virtually coincident on the energy scale plotted here.

in a straightforward manner that treats unpaired electrons in a UHF-like fashion and correlates only the valence electron pairs. This formulation preserves the decoupling of the cluster amplitude equations with a linear number of correlation amplitudes that is characteristic of closed-shell PP and enables UPP to correctly describe the separation of electron pairs.

We have demonstrated that, in practice, this simple treatment of the electron correlations makes noticeable improvements over HF for a variety of systems and properties. Molecular structures are slightly improved, and symmetry-breaking effects in radicals are often reduced relative to HF. Furthermore, unlike HF or most other high-level correlation methods, PP can qualitatively correctly describe bond-breaking, diradicals, etc.

However, PP itself does suffer from various weaknesses. Though it substantially improves upon HF, it comes nowhere near eliminating the symmetry-broken solutions found in HF. For the challenging diatomic radicals studied here, $\langle S^2 \rangle$ values are

noticeably reduced, and F_2^+ is bound, unlike at the UHF level in many basis sets. A more complete correlation model like VOD, in contrast, more significantly reduces the symmetry-breaking in these molecules (though of course, it may not be completely impervious to symmetry breaking either).

On the other hand, the asymmetry in how UPP treats the unpaired electrons versus the paired ones leads to some strange chemical predictions in certain classes of systems. Singlet-triplet gaps, for example, are bound to be poor, as the singlet is preferentially stabilized relative to the triplet through the correlation of an additional electron pair. The use of orthogonal, localized orbitals causes PP to slightly favor D_{3h} symmetry over D_{6h} for benzene, or likewise to predict that the allyl radical has two unequal carbon-carbon bonds rather than equivalent ones. In cases like these, additional correlation effects must be included to remove the symmetry breaking, such as with a non-orthogonal formalism[105].

In the future, it will be very interesting to see what wave function method offers the best compromise between cost and accuracy in order to become established as the smallest useful step beyond the HF model as a general-purpose reference wave function. PP is one contender given its general improvement over HF combined with its affordability. Artifactual symmetry breaking of the type discussed above is clearly its main weakness, as may be valence-shell expansion in heavier elements. Although more elaborate methods (for instance CASSCF or VOD) are more reliable, they cost significantly more and often force the user to choose a limited number of active electrons. PP describes a narrower range of phenomena well, but can do so without requiring any such input from the user and at a very low cost. Its simplicity also recommends it as a tractable starting point for a perturbative description of the remaining correlations.

We will explore just such a combination, the PP(2) method, in Chapter 7. First, however, we digress to explore an analytic tool for understanding the nature of electron-electron correlations in many body wave functions. This technique was inspired by the simplicity of the PP wave function and the desire to rewrite more complicated wave functions in a similarly intuitive fashion.

Chapter 6

Extracting dominant pair correlations from many-body wave functions.

6.1 Introduction

Modern wave-function-based electronic structure calculations offer the straightforward and often accurate prediction of molecular properties for relatively well-behaved systems. Methods such as second-order Møller-Plesset perturbation theory (MP2) or coupled cluster singles and doubles (CCSD) are widely used for this purpose, and the vastness of the scientific literature involving such methods vouches for their success.

All of these approaches can be described as having a configuration interaction (CI)-like wave function,

$$|\Psi\rangle = (1 + \hat{C}_1 + \hat{C}_2 + \cdots)|\Phi_0\rangle, \tag{6.1}$$

where $\hat{C}_n$ refers to an operator that generates the set of n-tuply-excited configurations, though the specific terms used in the expansion and their method of calculation differ. For example, in the context of CCSD, $\hat{C}_2 = \hat{T}_2 + \frac{1}{2}\hat{T}_1^2$, and the individual t amplitudes are solved for projectively.[17]

However, analyzing the wave function obtained from a CI-like electronic structure

calculations can prove very difficult due to the long and complex configuration expansion. Looking at the expansion, it is not always obvious which physical electron correlations contribute dominantly to the total molecular correlation. It is indeed these dominant correlations that often correspond to interesting chemistry, such as bond-formation or diradicaloid behavior. One may analyze the largest double excitation amplitudes from a CCD calculation, for example, but these are usually expressed in terms of the one-particle molecular orbitals (MOs) obtained from a mean-field Hartree-Fock (HF) calculation. Such orbitals are by no means guaranteed to provide a good (compact) basis for the description of the pair (or higher order) correlations. In fact, the unphysically-large scaling of canonical MP2, CCSD, CCSD(T), etc., is due to the inability to write the corrections to the mean-field wave function efficiently in the basis of MOs.[24, 28]

Alternatively, based on the known local nature of electronic correlations, one might represent the wave function in terms of localized orbitals. This has the tendency to significantly shorten the expansion for a large molecule (*i.e.* localized orbitals form a better basis than canonical MOs). Unfortunately, it is not totally clear what form these localized orbitals should take. Some methods, such as Perfect-Pairing and related methods,[104, 54, 55, 56, 58] utilize bonding and anti-bonding pairs of orthogonal bond-centered, localized orbitals.[25, 26] In contrast, some non-orthogonal local correlation approaches utilize atomic centered orbitals.[156, 40, 31] Both approaches certainly show promise for the understanding of the intrinsic structure of the wave function beyond their commonly-cited computational efficiency. However, despite much progress in the field, local correlation approaches have not yet become widely accepted in practice and require further development.

Yet another choice are the so-called natural orbitals, which are those orbitals that diagonalize the one-particle reduced density matrix (1-PDM).[157] They are sometimes used as an alternative to the standard HF orbitals. For example, natural orbitals have been used to characterize unpaired electrons in the wave function.[158, 159, 160, 161] or to select configurations for CI expansions. However, this expansion is based entirely on one-particle information, whereas we are really interested in two-particle or higher correlations. In practice, despite conventional wisdom to

the contrary, natural orbitals do not form a very compact expansion for the correlation energy of the molecule, at least compared to localized orbitals (though they are still typically more efficient than canonical MOs).[162] Moreover, the eigenvalues corresponding to each natural orbital determines its occupancy in the wave function. In complicated wave functions, this occupancy may come from a variety of different excitations from the reference determinant, but the natural orbital analysis gives us no clear way of differentiating among these physically distinct correlations. One wants both to understand how important a correlating orbital is and to associate that orbital with certain physical correlations.

For the purpose of wave function analysis (as distinct from wave function computation), we desire an optimal basis in which to represent the wave function as computed with coupled cluster theory or some other CI-like approach. In particular, we believe it is interesting to ascribe the contributions to some particular n-body effect, such as pair correlations. The method for finding such a basis should have the following properties:

1. It should represent the wave function in terms of a few dominant terms and many smaller corrections. For example, in a local representation based on bond-centered orbitals, the perfect-pairing-like (e.g. $\sigma \to \sigma^*$) correlations overwhelmingly dominate any other individual term in the wave function expansion (though the combined contribution of those other terms is non-trivial). The dominant terms should describe the most significant correlations that define the interesting correlation behavior.

2. It should be well-defined for any chosen Ansatz and any level of excitation. For example, to a large extent, pair correlations in MP2, CCSD, and CCSDTQ are similar and ought to emerge as such. Similarly, local approximations or valence space methods[46, 76] should not deter the analysis. Furthermore, an approach that is readily generalizable from double to quadruple excitations enables the facile analysis of whatever correlations happen to be relevant for the system of interest.

3. Its results ought to be independent of the initial choice of wave function representation, whether it has been written in terms of localized orbitals, atomic orbitals or canonical HF orbitals.

4. Its results should avoid unphysical basis set dependencies. Although the convergence of the correlation energy with basis set size is somewhat slow, the qualitative physical nature of the dominant correlations are unlikely to change significantly as the basis grows (assuming the original basis provides suitable flexibility for a qualitative description of the system).

5. It should be asymptotically size intensive.[163, 164] The nature of correlations in butane and octane are similar, though the total correlation energy will approximately double in size.

6. It should preserve size consistency—the description of the correlations for two non-interacting, identical subunits ought to be the same as for each monomer (assuming that the initial wave function was also size consistent).

7. It should be computationally inexpensive relative to the method used for determining the wave function.

For practical purposes, we focus on pair correlations in the CCD Ansatz. The straightforward approach for pair correlations is to look at the two-particle density matrix (2-PDM) and perhaps to construct the "natural geminals" by diagonalizing it. For example, such analysis has been done for the beryllium atom using primitive wave functions[165, 166, 167, 168]. However, the construction of the full 2-PDM is often computationally expensive and is typically unnecessary for the computation of the energy and wave function. One would prefer to utilize information that is already available from the wave function calculation.

To this end, we wish to transform the operator describing pair correlations into the basis of two-particle functions, or geminals, (since we are talking about pairs of electrons, we must abandon one-particle functions) in a way that provides an optimal expansion. A standard approach in the fields of data compression, signal processing,

and numerical analysis to obtain such an optimal expansion is to use the singular value decomposition (SVD),

$$A = U\Sigma V^\dagger = \sum_p \mathbf{u}_p \sigma_p \mathbf{v}_p^\dagger. \tag{6.2}$$

Geometrically, any $m \times n$ matrix A maps the unit hypersphere in $\Re^n$ to a hyperellipse in $\Re^m$.[169] The right singular vectors v_i form the basis of vectors for the unit hypersphere, and the left singular vectors u_i form the basis of unit vectors describing the hyperellipse. The singular values express the length along each unit vector in the hyperellipse as induced by the matrix A. A theorem regarding the SVD states that the best approximation to A of rank r is the r-th partial sum from above (low-rank approximation). Geometrically, this amounts to saying that the best one-dimensional approximation to an ellipse is a line along the major axis. Analogously, the best two-dimensional approximation to a three-dimensional hyperellipse is the ellipsoid spanned by the two-largest axes, etc. Thus, each subsequent term in the above summation provides the leading correction to the approximation for A. It is in this sense that we consider the SVD expansion as providing the optimal basis for wave function representation.

For CCD in particular, we wish to decompose the T_2 vector, which is fundamental to the description of pair correlations, using the SVD and analyze the singular values and their corresponding vectors for insights into the nature of pair-correlations in the calculated wave function. The SVD has been recently applied to the $\hat{T}_2$ operator by Kinoshita and co-workers.[170] However, their primary goal was to reap computational savings by decomposing an approximate T_2 vector and reducing the dimensionality of the problem via the SVD's data compression abilities. In contrast, we have been evaluating the SVD for its utility in wave function analysis.

Clearly this analysis is by no means unique to the $\hat{T}_2$ operator. One could just as easily decompose the $\hat{C}_n$ operator for any excitation level n and any Ansatz. However, because of the typically dominant nature of the double excitations and their preponderance in standard electronic structure methods, we shall focus primarily on the double excitation operator in this chapter. Perhaps more importantly, the pair correlations are special in that they alone contribute to the correlation energy (when

Brillouin's theorem applies), and all other correlations couple indirectly to the correlation energy via the double excitations. This means that the principal-component analysis of the pair correlations is also roughly a principal-component analysis of the correlation energy (though subject to variations due to the values of the individual integrals). One can obtain these double excitation operators from perturbation theory, coupled cluster theory, local correlation models, etc. Because of its widespread use and general quality, we will primarily study the $\hat{T}_2$ operator from coupled cluster theory.

The generalization to CCSD is straightforward: the singles terms generally contribute to pair correlations (and the correlation energy) directly as the product of two single excitations,

$$\hat{\tau}_2 = \hat{T}_2 + \frac{1}{2}\hat{T}_1^2, \tag{6.3}$$

and $\hat{\tau}_2$ can be decomposed as above. Because the weight of $\frac{1}{2}\hat{T}_1^2$ is generally small, this correction will make little difference in general. However, it is the formally correct way to account for pair-correlation effects in the CCSD wave function and will be used whenever singles are included in the calculation.

The rest of the chapter is organized as follows: First we will explain some of the theoretical aspects of the SVD in application to the T_2 vector from doubles-level coupled cluster theories. Secondly, we will examine properties of the SVD-T_2 for some model systems in order to learn about its behavior. Finally, we will apply the analysis to some more complex systems to automatically extract the interesting correlations.

6.2 Theory

Since we will focus on the $\hat{T}_2$ operator in the coupled cluster Ansatz, we briefly review the standard coupled cluster doubles equations.[13] The wave function is assumed to take the form:

$$|\Psi_{CCD}\rangle = e^{\hat{T}_2}|\Phi_0\rangle, \tag{6.4}$$

where $|\Phi_0\rangle$ is some single-determinantal reference state that can be either the Hartree-Fock (HF) state or perhaps a reference obtained from orbital optimization at a higher

level of theory, and $\hat{T}_2$ is a double excitation operator,

$$\hat{T}_2 = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_i \hat{a}_b^\dagger \hat{a}_j. \tag{6.5}$$

Solving the CCD energy and amplitude equations,

$$E = \langle \Phi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle \tag{6.6}$$

$$0 = \langle \Phi_{ij}^{ab} | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi_0 \rangle \tag{6.7}$$

provides the converged t_{ij}^{ab} amplitudes.

From this point on, the converged $\hat{T}_2$ operator is written as an $O^2 \times V^2$ matrix and denoted simply as T_2 (no hat), where O is the number of occupied spin orbitals and V is the number of virtual spin orbitals. From Perfect-Pairing (PP) theory,[104, 54, 55, 56, 58] we know that the dominant correlation contributions typically come from a pair of electrons in the same region of space correlating only with one pair of virtual orbitals (holes) that are localized in the same region of space. In the somewhat arbitrary definition of T_2 as an $O^2 \times V^2$ matrix, we are implicitly viewing correlations as a pair of electrons exciting from an occupied geminal to a virtual geminal. We will return to this point momentarily.

We can perform the singular value decomposition and rewrite the operator as:

$$\hat{T}_2 = \sum_P \tilde{t}_P^{P*} \Big(\frac{1}{2} \sum_{ab} V_{(ab),P}^\dagger \hat{a}_a^\dagger \hat{a}_b^\dagger \Big) \Big(\frac{1}{2} \sum_{ij} U_{(ij),P} \hat{a}_j \hat{a}_i \Big) = \sum_P \tilde{t}_P^{P*} \hat{C}_P^\dagger \hat{A}_P, \tag{6.8}$$

where $C_P^\dagger$ and A_P are two-particle creation and annihilation operators, and the parentheses in $U_{(ij)}$ indicate a collective index (ij). This notation is the same as was used by Kinoshita *et al.*[170] These two-particle operators are in fact a superposition of products of their one-particle counterparts $\hat{a}_i$ and $\hat{a}_a^\dagger$. In other words, they transform products of two one-particle MO functions (what might be termed "simple geminals") into the new non-separable basis of geminals that provides the optimal expansion. The column vectors in U and the row vectors in $V^\dagger$ provide the transformations for the occupied and virtual spaces obtained from the SVD, respectively. This enables us to examine the occupied and virtual geminals involved in the pair correlation. The new amplitudes $\{\tilde{t}_i\}$ characterize the weight of each pair excitation in the wave function.

Notice too that this representation is somewhat analogous to PP theory (without of course the PP restriction to the valence space). The coupled cluster formulation of PP also uses the CCD wave function, except the $\hat{T}_2$ operator is redefined as:

$$\hat{T}_2 = \sum_i t_{i\bar{i}}^{i^*\bar{i}^*} \hat{a}_{i^*}^\dagger \hat{a}_i \hat{a}_{\bar{i}^*}^\dagger \hat{a}_{\bar{i}}. \tag{6.9}$$

However, unlike in PP theory, the SVD representation of CCD actually forms each possible pair of electrons (PP makes an assumption about the way electrons pair, matching each individual electron with exactly one other electron of opposite spin), regardless of spatial locality or other considerations, in the same way standard CCD does. It is an exact transformation to within machine precision, as we have not explicitly truncated any singular values. Importantly, the number of nonzero singular values can never be more than the smallest dimension of the original matrix, which in practice means that their are never more than O^2 singular values, regardless of how large the basis set becomes.

As hinted above, one might have written the four-dimensional $\hat{T}_2$ operator in several possible different ways. For example, it could also be written as an $OV \times OV$ matrix, thereby viewing pair correlations as two simultaneous single excitations. The transformation vectors obtained by the SVD of this alternative form for $\hat{T}_2$ would each describe an occupied to virtual transition. It was found empirically that this second approach led to much longer optimal expansions with many more significant terms than did the $O^2 \times V^2$ matrix, making the latter more desirable for analysis. The interpretation of the singular vectors is also less intuitive, as they would correspond to a superposition of products of occupied and virtual orbitals. One might also write $\hat{T}_2$ as an $O \times OV^2$ or $V \times O^2V$, as was used in the $D2$ diagnostic[164]. This would generate one-particle functions that might be easier to interpret but that seem less suited to describing the pair- or higher-order correlations. Also, these three-index by one-index matrices potentially couple non-interacting blocks of the $\hat{T}_2$ matrix (e.g. $\alpha\alpha\alpha\alpha$ and $\alpha\beta\alpha\beta$) unless one separates out the blocks *a priori*.

Because we wish to learn about the nature of pair correlations, we must work with geminal functions, which involve six spatial dimensions. Unfortunately, such functions are difficult to visualize. Therefore, we have considered two possible approaches for

analyzing these geminals. The information regarding the geminal functions is contained in the left- and right-hand singular vectors. First, one could take the column vectors of U or the row vectors of $V^\dagger$, which have composite indexing (ij,1) or (1,ab), and write them as a matrix (i,j) or (a,b). The SVD may be applied once again to these matrices to determine the one-particle molecular orbitals that contribute principally to each geminal. However, in test cases tried it seems that there are often several important singular values to these matrices, and that each one corresponds to a mixture of several molecular orbitals, making this analysis comparable in difficulty to the initial, untransformed wave function. In fact, this observation further supports the notion that the one-particle MOs are a poor basis for the representation of the intrinsically largest pair correlations. This approach can be used, however, to complement the next approach which will be used extensively: we construct and plot the 1-PDM (i.e. averaged over all positions of the second electron) associated with each occupied and virtual geminal. These 1-PDMs are not the true coupled cluster relaxed densities which would involve the the left-hand coupled-cluster wave function, $\langle\Phi_0|(1+\Lambda)$. They are, however, adequate and useful for the intended wave function analysis.

In order to generate the density, we operate the P-th pair-wise creation or annihilation operator from the $\hat{T}_2$ operator onto the reference determinant. For the P-th occupied geminal, the occupied-occupied block of the 1-PDM for the wave function

$$|\Phi_P\rangle = \hat{A}_P|\Phi_0\rangle = \frac{1}{2}\sum_{ij} U_{(ij),P}|\Phi_{ij}\rangle, \tag{6.10}$$

where the occupied geminal P has been annihilated from the N-electron HF reference is given by:

$$d_{rs}^{\Phi_P} = \langle\Phi_P|\hat{a}_r^\dagger\hat{a}_s|\Phi_P\rangle = \frac{1}{4}\sum_{ijkl} U_{(ij),P}^\dagger U_{(kl),P}\langle\Phi_0|\hat{a}_i^\dagger\hat{a}_j^\dagger\hat{a}_r^\dagger\hat{a}_s\hat{a}_l\hat{a}_k|\Phi_0\rangle. \tag{6.11}$$

Notice that $|\Phi_P\rangle$ results from annihilating a superposition of occupied orbitals from the HF reference to give an $N-2$ electron state. It is therefore a multi-determinantal expansion in the MO basis. All other blocks of the 1-PDM are zero in this case. The 1-PDM for the P-th geminal is then obtained by subtracting d^{Φ_P} from the HF 1-PDM

d^{HF},

$$d^P = d^{HF} - d^{\Phi_P} = 2U_P^\dagger U_P \tag{6.12}$$

where U_P is the matrix form $U_{i,j}^P$ of the P-th column vector in $U_{(ij),P}$. This density matrix satisfies

$$Tr(d^P) = -2, \tag{6.13}$$

corresponding to the two holes created by the pair-annihilation operator $\hat{A}_P$. Alternatively, by inverting the sign of d^P, we obtain a pair density for the occupied geminal that has been annihilated.

Analogously, the virtual geminal 1-PDM is formed from constructing the virtual-virtual block of the 1-PDM for the state vector

$$|\Phi^{P^*}\rangle = \hat{C}_{P^*}^\dagger|\Phi_0\rangle = \frac{1}{2}\sum_{ab} V_{(ab),P}^\dagger|\Phi^{ab}\rangle, \tag{6.14}$$

which is an $N + 2$ electron state. One obtains,

$$d_{rs}^{\Phi^{P^*}} = \langle\Phi^{P^*}|\hat{a}_r^\dagger\hat{a}_s|\Phi^{P^*}\rangle = \frac{1}{4}\sum_{abcd} V_{(ab),P}V_{(cd),P}^\dagger\langle\Phi_0|\hat{a}_b\hat{a}_a\hat{a}_r^\dagger\hat{a}_s\hat{a}_c^\dagger\hat{a}_d^\dagger|\Phi_0\rangle. \tag{6.15}$$

Since we care only about the two created electrons, we calculate only the virtual-virtual block of the 1-PDM. It is given by:

$$d^{P^*} = 2V_P V_P^\dagger, \tag{6.16}$$

where $V_{P^*}^\dagger$ is formed from the P-th row vector of $V_{(ab),P}^\dagger$. This density for the created virtual geminal satisfies

$$Tr(d^{P^*}) = 2, \tag{6.17}$$

These geminal functions provide the spatial and physical meaning behind the SVD amplitudes. Recall that the SVD writes the wave function such that each hole geminal correlates with only one particle geminal, and the SVD amplitude gives the strength of the correlation. More specifically, the SVD amplitude reflects the weight of that particular configuration in the wave function, analogously to the untransformed amplitudes. We can also separate the excitations by their spin class (e.g. $\alpha\alpha \rightarrow \alpha\alpha$, $\beta\beta \rightarrow \beta\beta$, or $\alpha\beta \rightarrow \alpha\beta$) Combined with visualization of the hole and particle densities

for a given excitation (by mapping the 1-PDM onto the real-space density), one may learn what physical and spatial correlations dominate the wave function. Conceptually, these geminals densities resemble the attachment/detachment densities used in the analysis of complex excited-state wave functions.[171] Like natural orbitals, these geminals transform according to the point-group symmetry of the molecule, so they are in general delocalized around the molecule. However, in cases with lower symmetry or where a certain type of correlation dominates the wave function, more localized correlations are interpretable from the geminals, as will be demonstrated in the results. Unlike natural orbitals from the 1-PDM, these two-particle quantities provide information on associations between hole and particle densities, which is fundamental to describing correlations.

The efficient calculation of the SVD of the $\hat{T}_2$ operator is not entirely trivial. The SVD of an $O^2 \times V^2$ matrix requires O^4V^2 work. Substantial further savings are achieved by utilizing the block structure of T_2 and treating the $T_2(^{\alpha\alpha}_{\alpha\alpha})$, $T_2(^{\alpha\beta}_{\alpha\beta})$, and $T_2(^{\beta\beta}_{\beta\beta})$ blocks separately. Finding the SVD block-wise reduces the cost by $2^6 = 64$ and eliminates the redundancy inherent in the full matrix due to permutational symmetry. Eliminating this redundancy scales each of the singular values by $\frac{1}{2}$ and assigns them the proper physical meaning (i.e. they typically range from 0 to 1 under intermediate normalization, as expected).

Though the SVD calculation scales less than CCSD, because of the sizeable prefactor it can cost a non-trivial fraction of the full CCSD calculation. However, due to the low-rank approximation for the SVD and our primary interest in the dominant correlations, we may calculate only the largest handful of singular values and analyze their geminals to learn much about molecular correlations. Therefore, we wish to use iterative techniques designed to calculate only the largest singular values of a matrix. Unfortunately, to our knowledge, no such standard algorithms for the SVD exist in major, public-domain linear algebra libraries. However, the SVD is readily related to the eigenvalue problem and we have adapted the Davidson diagonalization[172, 173] procedure for this purpose, as described in Appendix B. This approach is both computationally inexpensive (scaling only O^2V^2 per root and costing essentially nothing compared to the underlying CCSD calculation) and easy to implement, since the

Davidson diagonalization is well-known and widely-used in the electronic structure community.

Before discussing the results, we should emphasize once again that although we have presented the formalism and the examples below in terms of pair correlations from CCD or CCSD, it is readily applicable to *any* single reference CI-type wave function and *any* level of many-body correlation, which was the second criterion listed in the introduction (for multi-reference methods, one must decide how to treat excitations from both the primary and secondary references on equal footing). Actually, we have already demonstrated that the SVD satisfies several of the criteria described above. It is also independent of the representation of the wave function (criterion #3), since any unitary transformation of the wave function will not alter the singular value spectrum (analogously to diagonalization). The singular vectors will differ nominally, though only in their representation and not in their direction or magnitude. If both sets were transformed to a common basis (the atomic orbital basis for example), they would be the same. Finally, it meets the last criterion for being computationally inexpensive. In practice, the iterative SVD for the largest few singular values costs only a tiny fraction of the full CC calculation.

6.3   Results and discussion

To begin, the results from the SVD of the T_2 vector as obtained from CCD are presented for some simple examples in order to demonstrate the mechanics of the procedure. All calculations were performed using a developmental version of Q-Chem.[85] If a CCSD calculation is used instead, the single excitations are included by their contribution to the doubles as $\frac{1}{2}\hat{T}_1^2$. Unless otherwise indicated, all results use the 6-31G* basis.[1, 2] Some basic properties of the SVD-T_2 have already been discussed in Ref [170], so we will only briefly reiterate the relevant properties here. Also, it is instructive to look at the fraction of the correlation energy obtained by each amplitude, so the integrals have also been transformed into the geminal basis to enable the

energy calculation:

$$E_{corr} = \frac{1}{4} \sum_{ijab} \langle ij \| ab \rangle t_{ij}^{ab} = \frac{1}{4} \sum_P \sum_{ijab} U_{(ij),P} \langle ij \| ab \rangle V_{(ab),P}^{\dagger} \tilde{t}_P^{P^*} = \frac{1}{4} \sum_P \langle P \| P^* \rangle \tilde{t}_P^{P^*}. \quad (6.18)$$

First, the decomposition spectrum of N_2 at its equilibrium geometry, along with each amplitude's contribution to the total correlation energy, is presented in Figure 6.1. Immediately one notices that a small handful of amplitudes account for a disproportionate share of the correlation energy. The first five amplitudes account for more than 42% of the correlation energy, and the largest one accounts for 18%! A perfect-pairing calculation on N_2 containing all the valence orbitals would also utilize five amplitudes, but it recovers only 28% of the correlation energy. This suggests that the SVD amplitudes are recovering some of the coupling between localized pairs that would be found in Imperfect Pairing (or other generalizations of PP)[76, 46] as well as aspects of beyond valence (core and dynamical) correlation (as in Chapter 3). The full CCD wave function contains 5063 permutationally unique amplitudes (above a 10^{-9} threshold). Approximating the wave function using the rank 1 (i.e. approximating T_2 by only the first SVD amplitude) or rank 5 wave function generates an estimate for 1065 (21%) and 3041 (60%) of the amplitudes, respectively. Clearly the SVD transformation compresses the wave function dramatically and stores much information within each SVD amplitude.

Further singular value decomposition of the top five geminal pairs (all of which come from the $\alpha\beta \to \alpha\beta$ block) individually indicates that there are primarily valence contributions to each excitation, but that there are also some minor beyond-valence effects. The 1-PDMs for the first five geminal pairs are plotted in Figure 6.2. The dominant excitation, $\tilde{t} = 0.17$ (Figure 2(a)), corresponds to a mixture of all valence excitations: σ, π, and lone pair electrons going to their virtual counterparts. This characterization of the largest amplitude being a sort of "molecular valence correlation" is quite typical, particularly in systems with high symmetry. It means that all valence correlations in the molecule are roughly comparable in strength. Of course, equilibrium N_2 has relatively uninteresting electronic structure according conventional wisdom. These dominant singular values will play a more significant role in more interesting examples below.

Figure 6.1: SVD-T$_2$ spectrum for N$_2$ (r_{NN} = 1.1 Å) at the CCD/6-31G* level. Both the 91 amplitudes (solid line, left axis) and the corresponding integrated percent correlation energy (dashed line, right axis) are plotted.
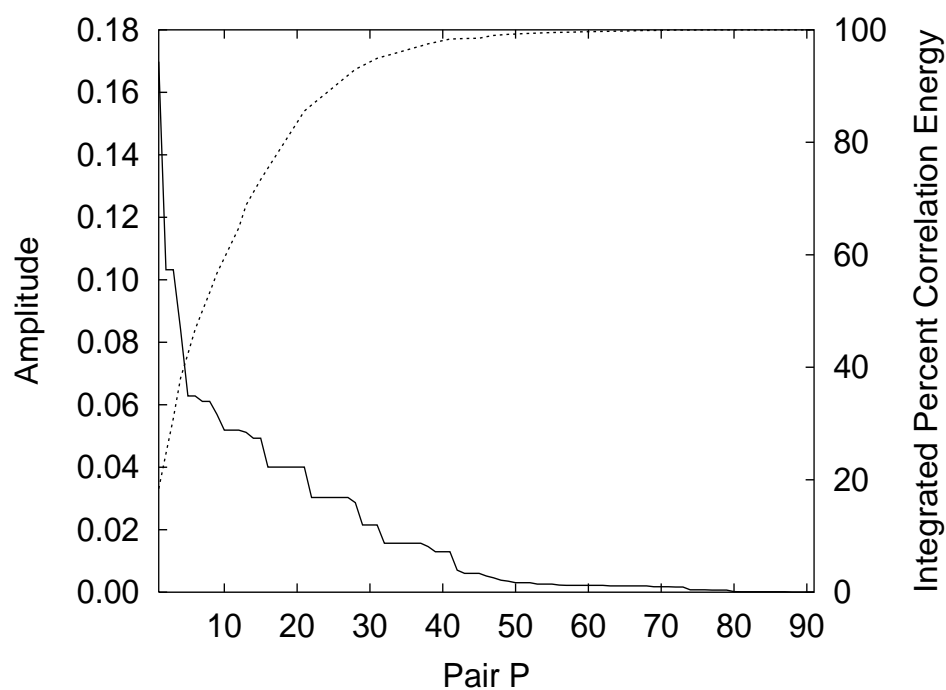
Figure 6.2: Top five geminal pairs for the CCD/6-31G* SVD of T_2 for N_2 at $r_{NN} = 1.1$ Å. The hole (occupied) geminal density is plotted on the left, and the particle (virtual) geminal density on the right. Note that excitation (b) is doubly degenerate.
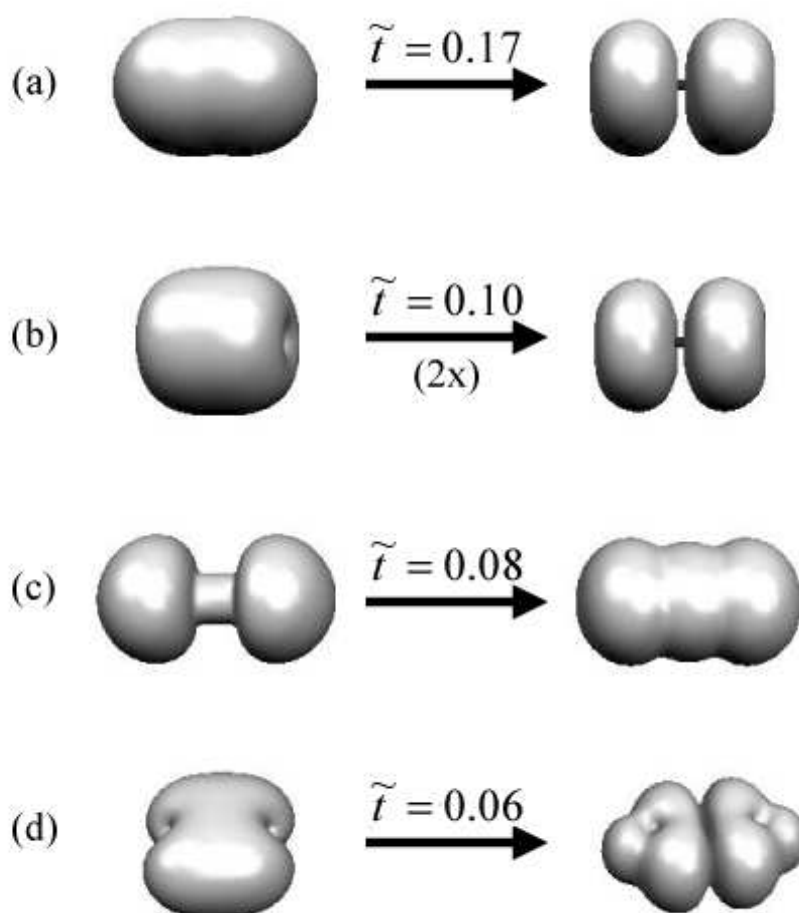
Secondary contributions come from two degenerate $\pi_x \to \pi_x^*$ and $\pi_y \to \pi_y^*$ excitations ($\tilde{t} = 0.10$, Figure 2(b)), and the next two excitations ($\tilde{t} = 0.08$ and $\tilde{t} = 0.06$ and Figures 2(c) and 2(d), respectively) involve lone pair, core, and further σ correlations. Because the $\pi \to \pi^*$ excitations are doubly degenerate, we have the freedom to mix them arbitrarily. The SVD mixes them to form the fully symmetric combination of π_x and π_y, though they could equally-well be represented as individual π_x and π_y excitations. At this point it may seem that the SVD has made the description of the correlations far more complicated than the initial wave function. But this is only the case because of the simple and well-understood nature of N_2's electronic structure in terms of the canonical MOs. We shall look at more interesting examples shortly that will demonstrate the interpretive simplicity of this approach. It should be noted, though, that this general pattern for the distribution of the singular values holds for all systems examined thus far.

Moreover, the dominant singular value amplitudes in N_2 were all of the type $\alpha\beta \to \alpha\beta$, as is typical for closed-shell molecules, where the largest same-spin correlations often occur well below the dominant paired-spin correlations. In the nitrogen molecule, the largest same-spin excitations have amplitude $\tilde{t}^{(\alpha\alpha)} = \tilde{t}^{(\beta\beta)} = 0.05$ and occur after nine larger $\alpha\beta \to \alpha\beta$ excitations. In open-shell molecules like triplet O_2, however, the same-spin correlations play a more important role. The largest two excitations at the UCCD/6-31G* level are a somewhat typical $\tilde{t}^{(\alpha\beta)} = 0.1016$ and the more surprising $\tilde{t}^{(\beta\beta)} = 0.1014$. This unusual and large second excitation looks primarily like a $\pi \to \pi^*$ excitation in the β spin space, exciting from the highest doubly-occupied orbitals to the singly-occupied orbitals. Like the original amplitudes, the SVD decomposition enables ready spin-specific analyses.

Clearly the SVD satisfies the first criterion of producing a compact expansion for the wave function. Secondly, as was previously demonstrated,[170] it is roughly independent of basis set size. The dominant correlations contained in the expansion do not change significantly as the basis set grows, but the smaller terms on the tail of the distribution each increase slightly to account for the additional correlation energy gained by the larger basis set. Significantly, enlarging the basis set leaves the fundamental length of the SVD expansion unchanged. Molecular nitrogen always

Table 6.1: Largest singular value for growing alkane chains C_nH_{2n} at the CCD/6-31G* level.

n	$\tilde{t}_{max}$	n	$\tilde{t}_{max}$
1	0.1248	5	0.1319
2	0.1293	6	0.1323
3	0.1307	7	0.1325
4	0.1314	8	0.1326

requires only 91 (the number of possible electron pairs) SVD configurations to describe the pair correlations, though not all of these configurations will necessary contribute significantly. This is another advantage of the SVD analysis over natural orbitals or natural geminals. The number of natural functions grows with the basis set size (though those extra functions may or may not have significant occupations). Thirdly, the SVD analysis preserves size consistency in size-consistent wave functions. The spectrum of singular values for two non-interacting subunits simply doubles in size, with each monomeric singular value occurring twice.

The final criterion from our original list of requirements is that the method should provide comparable descriptions for the correlations that are inherently similar, despite differences in system size. Table 6.1 lists the largest singular value for a growing linear alkane chain, from methane to octane, and we observe that this amplitude changes somewhat from methane to ethane (there is now a C-C bond), but once the chain gets longer and most of the C-H bonds and C-C bonds exists in roughly identical environments, the largest amplitude stabilizes around $\tilde{t} = 0.133$. The fact that the correlation energy grows is compensated for by filling in the spectrum more densely, so there are more terms of a similar size that contribute. In other words, the wave function has many more correlations of similar strength, rather than a constantly increasing correlation strength. Thus, the SVD is properly asymptotically size intensive and thereby satisfies all seven criteria initially laid out.

Having demonstrated that the procedure is sound, we now focus on more interesting chemical examples. To begin, we look at single, double and triple bond breaking in the context of ethane, ethylene, and acetylene. Figure 6.3 plots the largest singular

Figure 6.3: Largest singular value $\tilde{t}_{max}$ for bond-breaking in C_2H_6, C_2H_4, and C_2H_2 with restricted (solid) and unrestricted (dashed) CCD/6-31G*.
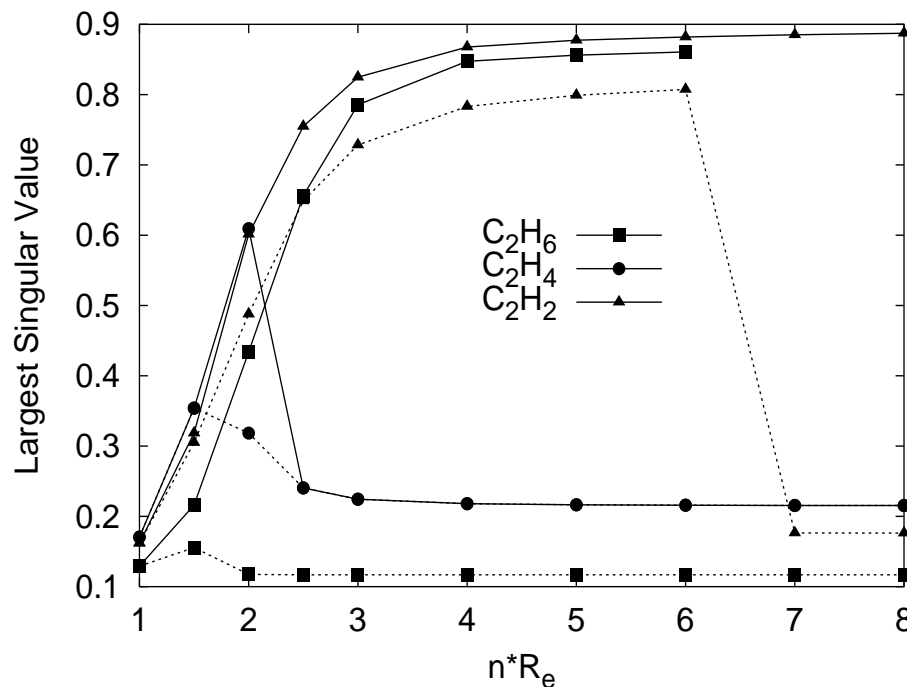
value for each of these species using both restricted and unrestricted HF references. For the double and triple bond-breaking, there are two or three dominant singular values, respectively—one for each highly-correlated electron pair. These multiple singular values are listed in Table 6.2.

For the restricted cases, the largest amplitude (spin type $\alpha\beta \rightarrow \alpha\beta$) for both ethane and acetylene rises quickly as the bond is stretched (unfortunately, the restricted CCD calculation for ethane would not converge after $7.0*R_e$, but this does not fundamentally alter the analysis). $\tilde{t}_{max}$ for ethylene also rises quickly, but it then turns over at $2.0*R_e$ and asymptotes to $\tilde{t}_{max} = 0.22$. The largest amplitude for the single and triple bond cases behaves roughly as expected—the electrons in the breaking bond become highly correlated and the corresponding amplitude grows substantially, heading towards 1.0 as the species becomes a pure diradical. For the unrestricted case, the ethane and acetylene curves turn over and asymptote to much lower values than the highest point on their corresponding restricted curves, and the

Table 6.2: Primary singular values associated with restricted bond breaking.

	C_2H_6	C_2H_4		C_2H_2		
$n * R_e$	$\tilde{t}_1$	$\tilde{t}_1$	$\tilde{t}_2$	$\tilde{t}_1$	$\tilde{t}_2$	$\tilde{t}_3$
1.0	0.129	0.170	0.104	0.162	0.110	0.110
1.5	0.216	0.354	0.167	0.319	0.244	0.244
2.0	0.434	0.609	0.370	0.601	0.212	0.198
2.5	0.656	0.240	0.236	0.755	0.185	0.184
3.0	0.786	0.224	0.224	0.825	0.180	0.180
4.0	0.847	0.218	0.218	0.868	0.179	0.177
5.0	0.856	0.216	0.216	0.877	0.178	0.177
6.0	0.861	0.216	0.216	0.882	0.178	0.177
7.0		0.216	0.216	0.885	0.178	0.177
8.0		0.215	0.215	0.887	0.178	0.177

unrestricted ethylene case turns over even sooner than the restricted case, though its asymptote is the same.

Examining the geminal pairs corresponding to the largest-amplitude excitation in ethane, which are plotted in Figures 6.4 and 6.5, will clarify the reason for these turnovers for the unrestricted case. At equilibrium, there is no unrestricted solution, and the dominant correlation behaves like a molecular valence bonding to antibonding excitation, analogous to what was observed in N_2. However, in the restricted case, stretching the bond shifts the correlation such that by $1.5*R_e$ the dominant correlation now appears to be purely along the C-C bond as a $\sigma \rightarrow \sigma^*$ excitation. The dominance of this correlation becomes even more pronounced by $2.5*R_e$ or $3.0*R_e$. Once the two fragments get far enough apart, very little density occurs in the middle of the bond, but the electrons remain highly correlated due to the restricted formalism.

The unrestricted case behaves similarly up to $1.5*R_e$, but beyond this point (when the dominant $\tilde{t}$ amplitude has completely turned over), the dominant correlation now occurs along the C-H bonds. The unrestriction allows the C-C bond electrons to localize onto their respective centers, and they are no longer intensely correlated. By $3.0*R_e$ the two fragments are far enough apart that they act mostly like two non-interacting methyl radicals, at least in terms of the most significant correlation effects.

Figure 6.4: Geminal pairs corresponding to the largest restricted SVD-T_2 amplitude $\tilde{t}_{max}$ for C_2H_6 with C-C bond lengths of (a) $1.0*R_e$, (b) $1.5*R_e$, (c) $2.0*R_e$, and (d) $3.0*R_e$ at the CCD/6-31G* level.
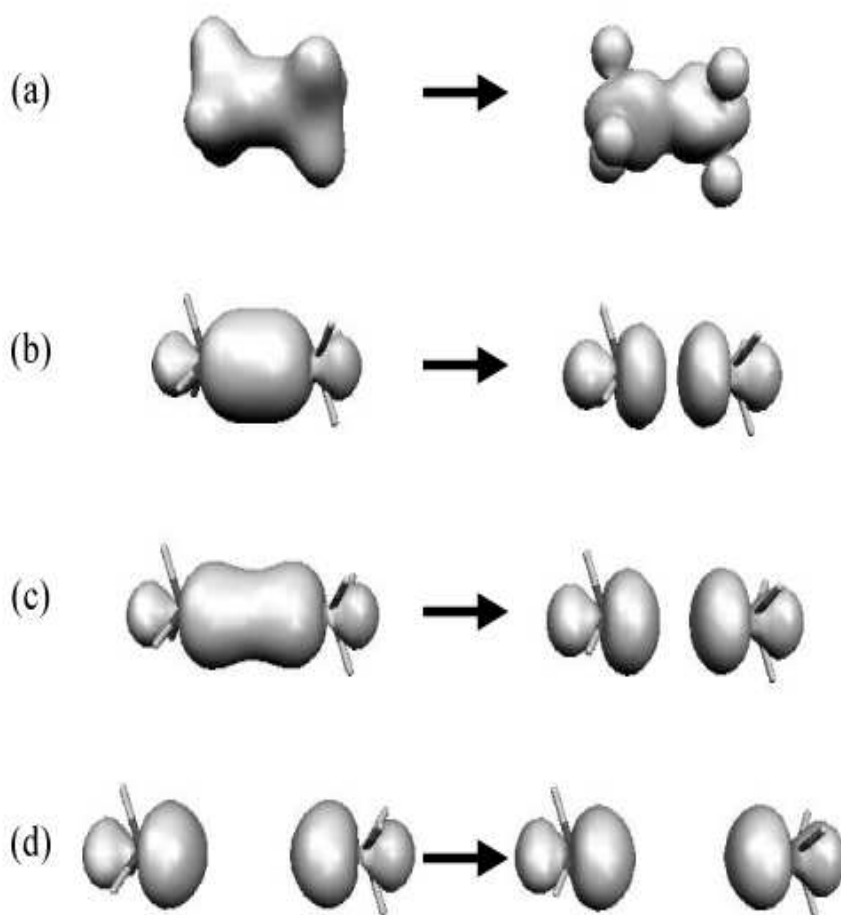
Figure 6.5: Geminal pairs corresponding to the largest unrestricted SVD-T_2 amplitude $\tilde{t}_{max}$ for C_2H_6 with C-C bond lengths of (a) $1.0*R_e$, (b) $1.5*R_e$, (c) $2.0*R_e$, and (d) $3.0*R_e$ at the CCD/6-31G* level.
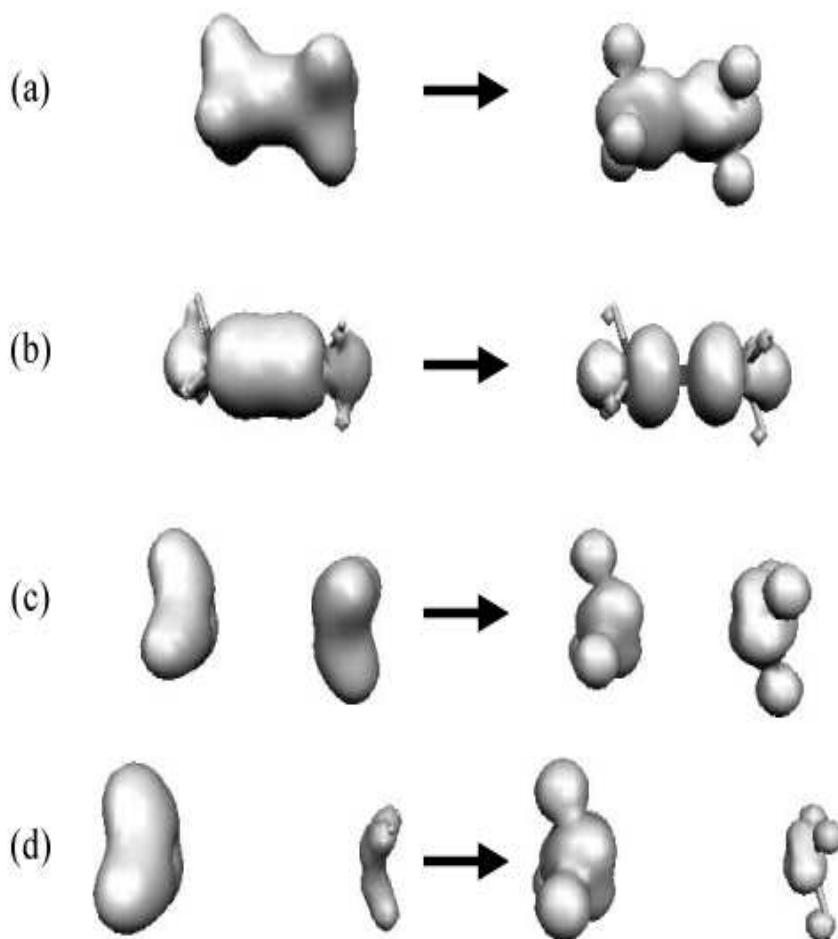
The two largest amplitudes are virtually degenerate, each localizing onto one of the two centers (at degeneracy, one is free to take any linear combination of the singular vectors to localize them completely).

Returning to the ethylene stretch and examining the two leading amplitudes, we see that the primary excitation $\tilde{t}_1$ corresponds to a $\pi \rightarrow \pi^*$ excitation, and the secondary one corresponds to the $\sigma \rightarrow \sigma^*$ along the C-C bond. Both the restricted and unrestricted curves for ethylene turned over because when stretching a double bond, the electrons recouple to the other free electron on the same site rather than with their now distant former partners. The system then behaves like two singlet methylene units rather than ethylene for a moderate stretch. This recoupling can happen in both the restricted and unrestricted cases. The two dominant correlations become $sp^2 \rightarrow p$ excitations on each fragment's carbon atom. In the triple bond case, a recoupling happens between one pair of π electrons and the σ electrons around $2.0*R_e$, analogously to ethylene. This recoupling leaves only a dominant $\pi \rightarrow \pi^*$ excitation, $\tilde{t}_1$. The power of this SVD approach has become clear—it has extracted the fundamental behavior of the electrons in these systems in an automatic fashion that enables the facile understanding of bond-breaking.

In molecules with lower symmetry, one can expect to obtain useful information from multiple singular values and vectors. Consider the molecule HOF. In this bent molecule at equilibrium, the largest singular value ($\tilde{t} = 0.17$) corresponds to correlation primarily along the O-F bond, while the second largest correlation ($\tilde{t} = 0.10$) corresponds to correlation mainly along the O-H bond. Looking at H_2O and F_2O we observe that the correlation energy per electron is 0.0198 E_h for H_2O and 0.0214 E_h for F_2O, suggesting that the O-F bond does indeed involve stronger correlations at equilibrium. Stretching the O-H bond to $1.5*R_e$ leaves the O-F correlation strength almost unchanged at $\tilde{t} = 0.16$, but it more than doubles the O-H correlation amplitude to $\tilde{t} = 0.21$, thereby reversing the order of the correlation strengths (though the respective geminal densities remain qualitatively the same). Water stretched to $1.5*R_e$ has a correlation energy of 0.0248 E_h per electron, again confirming the ordering of the excitations.

As a final example of diradicaloid behavior, we consider the planar ring BOBO

capped with four hydrogens. This model compound and others have been studied in the context of controlling diradical character based on neighboring group interactions using natural orbital analysis.[174] As with most diradicals, the diradicaloid correlation arises primarily from a HOMO-LUMO transition. Using the SVD approach on the VOD wave function (which should provide a better description than CCD),[46, 93] the largest $\alpha\beta \to \alpha\beta$ singular value of $\tilde{t} = 0.54$ dominates all other values (the second largest amplitude is $\tilde{t} = 0.12$). As one would expect, the geminal functions corresponding to the largest excitation are essentially the densities associated with the HOMO and LUMO orbitals described in Ref [174]. Once again, the SVD analysis has extracted the interesting correlation from the wave function automatically.

In all of the cases described thus far, the excitations were predominantly based on a single pair of electrons. In such situations, natural orbital analysis of the wave function is fairly straightforward, and the associations between natural orbitals can often be extracted from their occupation numbers, since there will usually be only one significantly populated virtual and a correspondingly-depleted occupied orbital. In cases with more complicated correlations, however, it is not always clear how certain electrons are correlated. With this problem in mind, we look at the case of benzene using OD, which is a variant of CCD in which the orbitals are variationally optimized during the CCD calculation.[47] In the natural orbital analysis, the top five most depleted occupied orbitals consist of various π and σ orbitals (see Table 6.3), with the two most depleted orbitals being one of each. The virtual orbitals provide similar difficulties. Although some of the most heavily populated virtual natural orbitals resemble anti-bonding forms of the most-depleted occupied natural orbitals, there is clearly no direct correspondence between the two sets. In particular, one cannot in general map the population depletion of a single occupied orbital with the population generation in a particular virtual orbital.

In contrast, if we apply the SVD to $\hat{T}_2$ for this system and examine the five largest excitations and their dominant character, which are listed in Table 6.4, we find that the largest corresponds to a fully symmetric $\pi \to \pi^*$ correlation with amplitude $\tilde{t} = 0.17$. The next four strongest correlations are somewhat smaller and form two doubly degenerate sets of $\sigma \to \sigma^*$ excitations with amplitudes $\tilde{t} = 0.12$ and $\tilde{t} = 0.10$,

Table 6.3: Natural spin orbital occupation numbers for benzene from the OD/6-31G* wave function.

	occupation	character	occupation	character	
HOMO	0.959	π	LUMO	π^*	0.035
HOMO-1	0.959	π	LUMO+1	π^*	0.035
HOMO-2	0.974	σ	LUMO+2	π^*	0.018
HOMO-3	0.979	σ	LUMO+3	π^*	0.013
HOMO-4	0.979	σ	LUMO+4	σ^*	0.013

Table 6.4: Largest five SVD-T_2 amplitudes and their dominant characters for benzene from the OD/6-31G* wave function. All excitations come from the $\alpha\beta \rightarrow \alpha\beta$ spin block.

	$\tilde{t}$	character
$\tilde{t}_1$	0.167	$\pi \rightarrow \pi^*$
$\tilde{t}_2$	0.119	$\sigma \rightarrow \sigma^*$
$\tilde{t}_3$	0.119	$\sigma \rightarrow \sigma^*$
$\tilde{t}_4$	0.102	$\sigma \rightarrow \sigma^*$
$\tilde{t}_5$	0.102	$\sigma \rightarrow \sigma^*$

respectively. Further down the list occur additional correlations of both types to account for all the possible electron pairs. Nevertheless, the analysis of benzene is substantially facilitated using the SVD analysis over natural orbitals.

The geminals formed from the SVD also verify the commonly-held understanding of "in-out" correlation. The Hartree-Fock wave function typically places the electrons too close to the nucleus on average, over-emphasizing the electron-electron repulsions. Introducing dynamical correlation into the wave function scatters the electrons into larger virtual orbitals which has the general effect of reducing electron-electron repulsion and thereby lowering the energy. To see this effect, we examine the geminals corresponding to the largest singular value ($\tilde{t} = 0.07$) for the neon atom as a function of distance from the nucleus (the densities are spherically symmetric). Both the geminal densities, $\rho_{hole}(r)$ and $\rho_{particle}(r)$, and their corresponding radial distribution functions, $4\pi\rho(r)$ are plotted in Figures 6.6 and 6.7. As seen in Figure 6.6 the leading correlation depletes the density at intermediate distances and enhances it at the nucleus and at longer distances. Weighting the density by the appropriate volume elements to construct the radial distribution function (Figure 6.7) provides a clear demonstration of "in-out" correlation. The excitation depletes the probability of finding an electron near 1.0 bohr and increases it substantially at longer distances.

Looking at a simple molecule like ethane, we see a very similar pattern for the leading correlation. Figures 6.8 and 6.9 plot density contours for the HCCH plane of the leading correlation in ethane at equilibrium. The correlation depletes density in the bond regions (along the C-C bonding axis in particular) and adds it near the hydrogen atoms and away from the other bonds near the carbon atoms, thereby extending the density distribution. This excitation exhibits the molecular version of "in-out" correlation.

6.4   Conclusions

The decomposition of pair correlations from CCD wave functions using the singular value decomposition has been presented as an efficient means of extracting the dominant pair correlations from many-body wave functions. This technique can also

Figure 6.6: Densities corresponding to the largest SVD excitation for the neon atom.
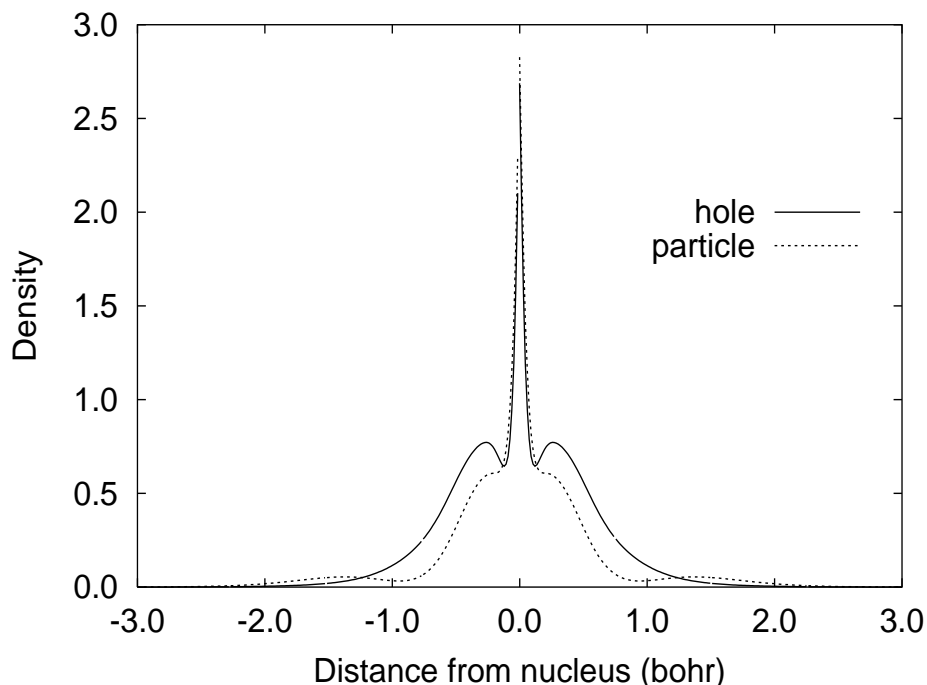
be generalized to higher-order correlations. For pair correlations, the SVD rewrites the wave function exactly in terms of one double excitation per pair of electrons and thereby associates an occupied and virtual geminal pair. Analysis of these geminals provides insight into the nature of the correlations, while the magnitude of the corresponding singular value determines the weight of this geminal in the wave function. That one obtains associations between pairs of occupied and virtual electrons is perhaps the principal advantage of this approach over natural orbital analysis, which, as a one-particle method, can only tell us the total occupation of an orbital, but not necessarily from which correlations that occupation comes from. Using the iterative Davidson SVD algorithm described here, the calculation of these dominant correlations requires almost trivial computational effort (as compared to the wave function determination), making it applicable to almost any molecular system accessible by standard correlation methods.

In general the dominant correlation resembles a "molecular valence" excitation that typically involves bonding to anti-bonding excitations among all the valence

Figure 6.7: Radial distribution function corresponding to the largest SVD excitation for the neon atom.
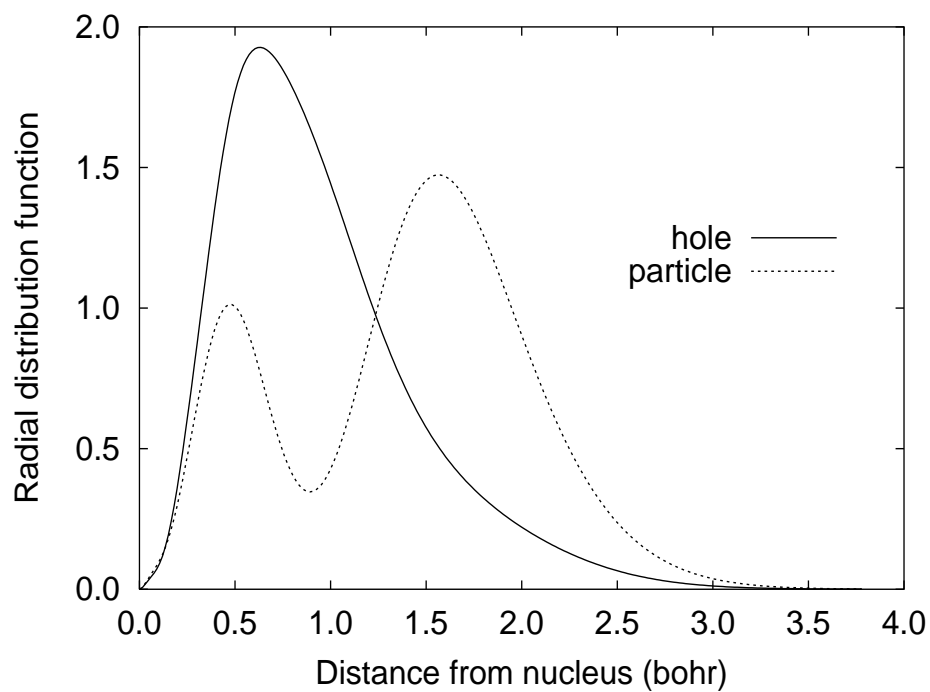
Figure 6.8: Contour plot in the HCCH plane of the occupied geminal corresponding to the leading SVD excitation in ethane. Contours are separated by 0.01 probability.
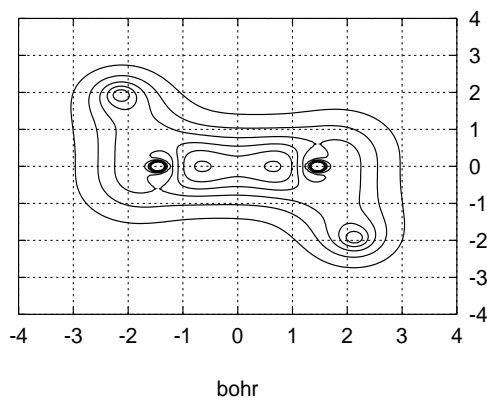
Figure 6.9: Contour plot in the HCCH plane of the virtual geminal corresponding to the leading SVD excitation in ethane. Contours are separated by 0.01 probability.
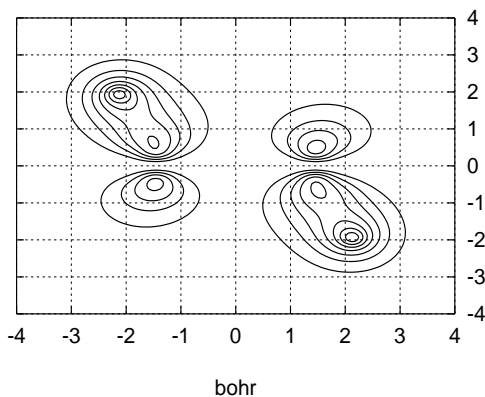
orbitals. In cases with more interesting correlation, however, this largest excitation can identify diradicaloid behavior or other unusual correlation effects. As an example, the SVD approach provided a clear analysis of the changes in the wave function during restricted and unrestricted bond-breaking with very little effort. We were also able to provide a simpler description of the correlations for benzene using the SVD analysis instead of natural orbital analysis. It is for these latter cases that the SVD analysis presented here seems most useful. Finally, the SVD geminals nicely demonstrated the notion of "in-out" correlation in a couple of simple examples.

The approach described herein is lacking in two regards. First, given the known local nature of electronic correlations, one would like a decomposition that describes the correlations locally. For example, the correlations in a particular functional group ought to remain similar in a variety of different molecules. Except in the case of a very strong correlation localized to one region of the molecule, the SVD will return delocalized descriptions of the correlations. Unfortunately, it does not seem likely to obtain a unique, well-defined decomposition that does not also adhere to molecular point-group symmetry. Any attempt to partition the excitations into localized blocks

for decomposition will depend dramatically on the partitioning scheme and could also miss non-local correlation effects.

Second, we have decomposed the wave function in terms of its dominant pair excitations. One might instead prefer to obtain the decomposition of the energy in terms of its dominant contributions. Though these two decompositions will in general be somewhat similar, the contributions of the SVD amplitudes to the energy are modulated by the integrals with which they contract. Empirically, it seems that the dominant energetic contributions typically arise from the dominant SVD amplitudes, but we are not guaranteed this result. A similar difficulty arose in the study of natural geminals.[165, 168] Of course, one can simply apply similar techniques to the energy itself rather than the wave function if one truly desires an energy decomposition.

Finally, in this chapter, we have not placed too much significance on the actual size of the largest correlation in the SVD wave function. We know, however, that stronger correlations correspond to a breakdown in the single reference assumption, suggesting that the calculated wave function requires higher-order excitations (either in the context of multi-reference approaches or longer CC expansions). The difficulty is knowing when such approaches are necessary or, conversely, how bad an approximation the single reference approach is. Perhaps this largest singular value can provide a key to measuring the reliability of a CCSD or related calculation with respect to the single reference character of the system. In particular, the SVD could extract this indicator from the wave function regardless whether or not the initial wave function representation made this multi-reference character obvious. Thus, the SVD technique presented here might make a useful diagnostic tool for the validity of an approximate wave function.

Finally, in the next chapter we return to our original stated goal of developing an efficient method for the treatment of strong correlations. We apply a second-order perturbative correction, (2), to the perfect pairing wave function. This very affordable method will be assessed in a variety of interesting chemical systems.

Chapter 7

An inexpensive electronic structure method for the treatment of highly-correlated systems.

7.1   Introduction

Reliably describing the electronic structure of species with significant diradicaloid character, such as bond-breaking, transition states, and novel organic materials, requires a careful treatment of the electron-electron correlations not present in standard density functional theory (DFT) or wave function methods (e.g. second-order Møller-Plesset perturbation theory, MP2). Instead, researchers must use computationally expensive methods like complete active space self-consistent field (CASSCF), to properly describe the static correlation effects,[41] followed by multi-reference perturbation theories such as CASPT2,[42] for these systems. Unfortunately, the factorial scaling of CASSCF limits its applicability to no more than fourteen active electrons. Clearly less expensive methods for studying large systems are desirable.

The simplest model chemistry wave function beyond the mean-field approximation is Generalized Valence Bond Perfect Pairing (PP), which in a coupled cluster

formalism writes the wave function as,[57, 58]

$$|\Psi_{PP}\rangle = e^{\hat{T}_{PP}}|\Phi_0\rangle, \tag{7.1}$$

where $|\Phi_0\rangle$ is a single-determinantal reference wave function and

$$\hat{T}_{PP} = \sum_i^p t_i \hat{a}_{i*}^\dagger \hat{a}_i \hat{a}_{\bar{i}*}^\dagger \hat{a}_{\bar{i}}, \tag{7.2}$$

where p is the number of electron pairs (usually all valence pairs). This model can be viewed as a strongly local approximation to a valence active space coupled cluster doubles calculation[46] that correlates the active electrons one pair at a time, with a single correlating amplitude per pair. This linear number of pair correlations typically comprises the leading terms in the correlation energy expansion, and PP generally recovers 50-80% of the full valence space doubles correlation energy (which requires a quartic number of amplitudes[6, 59]). The PP wave function serves as a significant improvement over Hartree-Fock (HF) wave functions, improving structural predictions[53] and even overcoming some symmetry-breaking effects and other pathologies found in HF.[59] We have recently developed an extremely efficient implementation of PP using auxiliary basis sets and the resolution of the identity (RI) approximation (also known as density fitting)[141, 142, 143, 144] that makes the cost of a PP calculation only a factor of several times that of a HF one with negligible additional errors.[60] With this implementation, single point PP energy calculations on systems with a thousand basis functions and tens of non-hydrogen atoms can be performed in a day or two on a modern personal computer.

Despite the advantages of the PP model, the limited description of the correlation energy with only a linear number of retained amplitudes is inadequate for accurate energetics and properties. We need to account for the multitude of individually-less-important but cumulatively-significant missing correlation terms. To this end, previous researchers have extended the treatment of correlations in various ways. One straightforward approach is to utilize configuration interaction (CI) expansions from the PP orbitals.[123, 124, 125, 126, 127, 128] As with all CI expansions, they generally must be truncated at some relatively low order to ensure that the number of configurations remains manageable. These expansions have been used to accurately

predict singlet-triplet gaps, for example,[175] but they are not size-consistent. Size consistent coupled cluster expansions similar to these restricted CI ones have also been developed,[76, 106] but these only improve the treatment of the valence correlations and do not account for the rest of the dynamical correlation.

Alternatively, based on the success of MP2 for treating dynamical correlation in systems for which HF provides a good reference wave function, one might use perturbation theory from the PP reference. Multi-reference Møller-Plesset perturbation theory has been applied to GVB wave functions with some success.[43, 44, 129, 130, 131] Similarly, perturbative corrections to the slightly more general antisymmetrized product of strongly orthogonal geminals (APSG)[135] have been developed.[176]

We want to apply the perturbative correction to our efficient coupled-cluster formulation of PP, so we have adapted the similarity-transformed perturbation theory of Gwaltney and Head-Gordon[48, 5, 49, 3] to PP wave functions. We have implemented the second-order correction, again using the RI approximation, which we term PP(2). As will be discussed below, this correction is MP2-like in computational cost and can be applied to systems with tens of heavy atoms and many hundreds of basis functions.

As mentioned above, PP is a strongly local approximation to valence orbital optimized coupled cluster doubles (VOD)[46] that includes only one correlation amplitude per valence pair.[53] For this reason, PP(2) can be viewed as a much simpler approximation to VOD(2).[5] We will place some emphasis on how PP(2) differs from VOD(2) in addition to how PP(2) compares to MP2 and other standard methods.

7.2 Theory

7.2.1 The formalism

Similarity-transformed perturbation theory has been presented elsewhere.[48, 5, 49] Here, we will review it only briefly and emphasize how the application differs for PP(2). In this Ansatz, the similarity-transformed Hamiltonian,

$$\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}, \tag{7.3}$$

which is represented in the basis of all many-body determinants, is partitioned into two spaces, $\mathbf{p}$, which is the space in which the coupled cluster equations have been solved (e.g. the space of determinants spanned by the PP wave function), and $\mathbf{q}$, which is the remainder of the space,

$$\bar{H} = |\mathbf{p}\rangle\langle\mathbf{p}|\bar{H}|\mathbf{p}\rangle\langle\mathbf{p}| + |\mathbf{p}\rangle\langle\mathbf{p}|\bar{H}|\mathbf{q}\rangle\langle\mathbf{q}| + |\mathbf{q}\rangle\langle\mathbf{q}|\bar{H}|\mathbf{p}\rangle\langle\mathbf{p}| + |\mathbf{q}\rangle\langle\mathbf{q}|\bar{H}|\mathbf{q}\rangle\langle\mathbf{q}|. \qquad (7.4)$$

A perturbative expansion for the exact ground state energy and wave function for $\bar{H}$ is then written, taking the reference coupled cluster wave function to be zeroth order. Because $\bar{H}$ is not Hermitian, it has different right- and left-hand eigenvectors which form a biorthogonal set, each of which we expand in orders of the perturbation:

$$\mathbf{R_p} = \mathbf{R_p}^{(0)} + \mathbf{R_p}^{(1)} + \mathbf{R_p}^{(2)} + \cdots \qquad (7.5)$$

and

$$\mathbf{L_p} = \mathbf{L_p}^{(0)} + \mathbf{L_p}^{(1)} + \mathbf{L_p}^{(2)} + \cdots . \qquad (7.6)$$

With this partitioning,

$$E^{(0)} = \langle 0|\mathbf{L_p}^{(0)}\bar{\mathbf{H}}_{\mathbf{pp}}^{(0)}\mathbf{R_p}^{(0)}|0\rangle \qquad (7.7)$$

and the first non-zero correction arises at second order,

$$E^{(2)} = \langle 0|\mathbf{L_p}^{(0)}\bar{\mathbf{H}}_{\mathbf{pq}}^{(1)}(E^{(0)}\mathbf{1} - \bar{\mathbf{H}}_{\mathbf{qq}}^{(0)})^{-1}\bar{\mathbf{H}}_{\mathbf{qp}}^{(1)}\mathbf{R_p}^{(0)}|0\rangle. \qquad (7.8)$$

To ensure that the zeroth order energy and wave function are those from PP, we choose $\mathbf{L_p}^{(0)} = (1+\Lambda)$ and $\mathbf{R_p}^{(0)} = 1$, in accordance with the standard coupled cluster Lagrangian. Substituting these in to Eqs 7.7 and 7.8, we find:

$$E^{(0)} = \langle 0|(1 + \Lambda)\bar{\mathbf{H}}_{\mathbf{pp}}^{(0)}|0\rangle = E_{PP} \qquad (7.9)$$

and

$$E^{(2)} = \langle 0|(1 + \Lambda)\bar{\mathbf{H}}_{\mathbf{pq}}^{(1)}(E^{(0)}\mathbf{1} - \bar{\mathbf{H}}_{\mathbf{qq}}^{(0)})^{-1}\bar{\mathbf{H}}_{\mathbf{qp}}^{(1)}|0\rangle. \qquad (7.10)$$

In the application of the (2) correction to PP, we made two significant changes from VOD(2). First, as will be discussed further below, we compute the quadruple excitation terms exactly, without the aid of the factorization approximation used

in other (2) methods. Because of the simplicity of the PP wave function, only a small subset of the full quadruples vector is included, and the quadruples are therefore almost trivial to compute in comparison to the doubles and triples, so no such approximation is necessary.

Second, the nature of the PP wave function makes it prudent to alter our choice of the zeroth-order Hamiltonian slightly. Unlike VOD, PP is not invariant to orbital rotations within the active occupied or active virtual subspaces. This means that changes of orbital representation in the active space will generally increase the number of amplitudes in the wave function from linear to quartic, thereby destroying the computational efficiency of PP(2).

The zeroth-order Hamiltonian used in VOD(2) is:

$$\bar{H}^{(0)} = |\mathbf{p}\rangle\langle\mathbf{p}|\bar{H}|\mathbf{p}\rangle\langle\mathbf{p}| + |\mathbf{q}\rangle\langle\mathbf{q}|(E^{(0)} + \bar{F}_{\bullet j'}^{i'} + \bar{F}_{\bullet J}^{I} + \bar{F}_{\bullet B}^{A} + \bar{F}_{\bullet b'}^{a'})|\mathbf{q}\rangle\langle\mathbf{q}|, \qquad (7.11)$$

where $\bar{F}_{\bullet j'}^{i'}$ and $\bar{F}_{\bullet J}^{I}$ represent the inactive and active occupied-occupied blocks, and $\bar{F}_{\bullet B}^{A}$ and $\bar{F}_{\bullet b'}^{a'}$ represent the active and inactive virtual-virtual blocks of the one-electron piece of $\bar{H}$, respectively.

Using this partitioning, it is necessary to semi-canonicalize $\bar{F}_{\bullet j}^{i}$ and $\bar{F}_{\bullet b}^{a}$ in order to facilitate evaluating the energy denominator $(E^{(0)}\mathbf{1} - \bar{\mathbf{H}}_{\mathbf{qq}}^{(0)})^{-1}$ in Eq 7.10. However, doing so will destroy the compactness of the PP wave function. Instead, we wish to utilize the PP orbitals (with the inactive blocks semi-canonicalized in terms of the bare Fock matrix), and therefore we choose only the diagonal of $\bar{F}$ for the second piece in the our zeroth-order Hamiltonian,

$$\bar{H}^{(0)} = |\mathbf{p}\rangle\langle\mathbf{p}|\bar{H}|\mathbf{p}\rangle\langle\mathbf{p}| + |\mathbf{q}\rangle\langle\mathbf{q}|(E^{(0)} + \bar{F}_{\bullet i'}^{i'} + \bar{F}_{\bullet I}^{I} + \bar{F}_{\bullet A}^{A} + \bar{F}_{\bullet a'}^{a'})|\mathbf{q}\rangle\langle\mathbf{q}|, \qquad (7.12)$$

The fact that we do not semi-canonicalize the orbitals in terms of $\bar{F}$ also means that it is not necessary to work in a biorthogonal basis, as was the case for VOD(2), which simplifies the implementation. In numerical tests with VOD(2) on small molecules, this alternate partitioning typically changed the total energy at the sub-millihartree level, which is small compared with the overall error.

When applied to a non-local coupled cluster doubles t vector, this second-order correction scales formally with the ninth power of system size (though in practice,

a factorization approximation in the quadruples reduces this scaling overall to the seventh power) . In contrast, PP(2) scales only with the fifth power with no factorization approximation, which is the same as canonical MP2. The computational savings arise from the PP-model-imposed sparsity of the doubles amplitude vectors t and λ, which each have only a linear number of non-zero elements. For example, whereas the formation of the two-body pieces of $\bar{H}$ for VOD(2) requires sixth-order work *after* the integrals and amplitudes have been obtained, it requires only linear work that depends on the number of PP pairs in PP(2). Similarly, the triples and quadruple excitation terms in the (2) correction simplify dramatically. The PP(2) equations are presented in spin-orbital form in Table 7.1. The relevant pieces of $\bar{H}$ which serve as intermediates are found in Appendix C.

7.2.2 Interpretation of the terms

The PP(2) correction includes six classes of terms. Representative examples of each class are listed in Table 7.2. The discussion of scaling contained herein assumes that we have more virtual orbitals than occupied ones, as is usually the case. The first terms are the MP2 singles and the Lambda singles. These two sets of terms account for single excitations, which are completely absent in PP. However, they tend to be small and do not play a major role in the overall correction. Evaluating these terms requires only OV work for the MP2 singles and pOV work for the Lambda singles terms, where O is the number of occupied orbitals, V is the number of virtual orbitals, and p is the number of PP pairs. They are therefore computationally trivial to compute once we have the requisite two-electron integrals. Though the details differ, the physics imparted by these terms is not significantly different from the singles in VOD(2).

Next come the MP2 doubles and Lambda Doubles terms. The MP2 doubles, which dominate the PP(2) correction both energetically and computationally, look exactly like the standard canonical MP2 expression, except that the integrals and energy denominators are replaced by the corresponding dressed elements of $\bar{H}$. These terms provide an MP2-like treatment of all pair correlations neglected by PP, and

Table 7.1: Spin-orbital equations for PP(2). These equations were obtained by applying the sparsity of the amplitude vectors to the VOD(2) equations[5].

$$\text{MP2 Singles} = \sum_{ia} \frac{\bar{F}_{\bullet a}^{i} \bar{F}_{\bullet i}^{a}}{\bar{\varepsilon}_i - \bar{\varepsilon}_a}$$

$$\text{MP2 Doubles} = \frac{1}{4} \sum_{ijab} \frac{\bar{W}_{\bullet\bullet ab}^{ij} \bar{W}_{\bullet\bullet ij}^{ab}}{\bar{\varepsilon}_i + \bar{\varepsilon}_j - \bar{\varepsilon}_a - \bar{\varepsilon}_b}$$

$$\text{Lam Singles} = \sum_{I} \lambda_I \left(\frac{\bar{F}_{\bullet\bar{I}}^{\bar{I}*} \bar{F}_{\bullet I}^{I*}}{\bar{\varepsilon}_I - \bar{\varepsilon}_{I*}} + \frac{\bar{F}_{\bullet I}^{I*} \bar{F}_{\bullet\bar{I}}^{\bar{I}*}}{\bar{\varepsilon}_{\bar{I}} - \bar{\varepsilon}_{\bar{I}*}} \right)$$

$$+ \sum_{Ia} \lambda_I \left(\frac{\bar{W}_{\bullet\bullet a\bar{I}}^{I*\bar{I}*} \bar{F}_{\bullet I}^{a}}{\bar{\varepsilon}_I - \bar{\varepsilon}_a} - \frac{\bar{W}_{\bullet\bullet aI}^{I*\bar{I}*} \bar{F}_{\bullet\bar{I}}^{a}}{\bar{\varepsilon}_{\bar{I}} - \bar{\varepsilon}_a} \right)$$

$$- \sum_{Ij} \lambda_I \left(\frac{\bar{W}_{\bullet\bullet I\bar{I}}^{j\bar{I}*} \bar{F}_{\bullet j}^{I*}}{\bar{\varepsilon}_j - \bar{\varepsilon}_{I*}} - \frac{\bar{W}_{\bullet\bullet I\bar{I}}^{jI*} \bar{F}_{\bullet j}^{\bar{I}*}}{\bar{\varepsilon}_j - \bar{\varepsilon}_{\bar{I}*}} \right)$$

$$+ \sum_{Ija} \frac{\lambda_I t_I \bar{F}_{\bullet j}^{a}}{\bar{\varepsilon}_j - \bar{\varepsilon}_a} \left(\bar{W}_{\bullet\bullet\bar{I}*a}^{\bar{I}*j} + \bar{W}_{\bullet\bullet I*a}^{I*j} - \bar{W}_{\bullet\bullet\bar{I}a}^{\bar{I}j} - \bar{W}_{\bullet\bullet Ia}^{Ij} \right) \qquad (7.13)$$

$$\text{Lam Doubles} = \sum_{Ib'} \lambda_I \left(\frac{\bar{F}_{\bullet b'}^{\bar{I}*} \bar{W}_{\bullet\bullet I\bar{I}}^{I*b'}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{\bar{I}} - \bar{\varepsilon}_{I*} - \bar{\varepsilon}_{b'}} - \frac{\bar{F}_{\bullet b'}^{I*} \bar{W}_{\bullet\bullet I\bar{I}}^{\bar{I}*b'}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{\bar{I}} - \bar{\varepsilon}_{\bar{I}*} - \bar{\varepsilon}_{b'}} \right)$$

$$- \sum_{Ij'} \lambda_I \left(\frac{\bar{F}_{\bullet\bar{I}}^{j'} \bar{W}_{\bullet\bullet Ij'}^{I*\bar{I}*}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{j'} - \bar{\varepsilon}_{I*} - \bar{\varepsilon}_{\bar{I}*}} - \frac{\bar{F}_{\bullet I}^{j'} \bar{W}_{\bullet\bullet\bar{I}j'}^{I*\bar{I}*}}{\bar{\varepsilon}_{\bar{I}} + \bar{\varepsilon}_{j'} - \bar{\varepsilon}_{I*} - \bar{\varepsilon}_{\bar{I}*}} \right)$$

$$+ \frac{1}{4} \sum_{Iab} \frac{\lambda_I \bar{W}_{\bullet\bullet ab}^{I*\bar{I}*} \bar{W}_{\bullet\bullet I\bar{I}}^{ab}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{\bar{I}} - \bar{\varepsilon}_a - \bar{\varepsilon}_b} + \frac{1}{4} \sum_{Ijk} \frac{\lambda_I \bar{W}_{\bullet\bullet I\bar{I}}^{jk} \bar{W}_{\bullet\bullet jk}^{I*\bar{I}*}}{\bar{\varepsilon}_j + \bar{\varepsilon}_k - \bar{\varepsilon}_{I*} - \bar{\varepsilon}_{\bar{I}*}}$$

$$- \sum_{Ija} \lambda_I \left(\frac{\bar{W}_{\bullet\bullet Ib}^{j\bar{I}*} \bar{W}_{\bullet\bullet j\bar{I}}^{I*b}}{\bar{\varepsilon}_j + \bar{\varepsilon}_{\bar{I}} - \bar{\varepsilon}_{I*} - \bar{\varepsilon}_b} - \frac{\bar{W}_{\bullet\bullet Ib}^{jI*} \bar{W}_{\bullet\bullet j\bar{I}}^{\bar{I}*b}}{\bar{\varepsilon}_j + \bar{\varepsilon}_{\bar{I}} - \bar{\varepsilon}_{\bar{I}*} - \bar{\varepsilon}_b} \right.$$

$$\left. - \frac{\bar{W}_{\bullet\bullet\bar{I}b}^{j\bar{I}*} \bar{W}_{\bullet\bullet jI}^{I*b}}{\bar{\varepsilon}_j + \bar{\varepsilon}_I - \bar{\varepsilon}_{I*} - \bar{\varepsilon}_b} + \frac{\bar{W}_{\bullet\bullet\bar{I}b}^{jI*} \bar{W}_{\bullet\bullet jI}^{\bar{I}*b}}{\bar{\varepsilon}_j + \bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}*} - \bar{\varepsilon}_b} \right)$$

$$- \frac{1}{2} \sum_{Ijab} \lambda_I t_I \left(\frac{\bar{W}_{\bullet\bullet ab}^{\bar{I}j} \bar{W}_{\bullet\bullet\bar{I}j}^{ab}}{\bar{\varepsilon}_{\bar{I}} + \bar{\varepsilon}_j - \bar{\varepsilon}_a - \bar{\varepsilon}_b} + \frac{\bar{W}_{\bullet\bullet ab}^{Ij} \bar{W}_{\bullet\bullet Ij}^{ab}}{\bar{\varepsilon}_I + \bar{\varepsilon}_j - \bar{\varepsilon}_a - \bar{\varepsilon}_b} \right)$$

$$- \frac{1}{2} \sum_{Ijka} \lambda_I t_I \left(\frac{\bar{W}_{\bullet\bullet\bar{I}*a}^{jk} \bar{W}_{\bullet\bullet jk}^{\bar{I}*a}}{\bar{\varepsilon}_j + \bar{\varepsilon}_k - \bar{\varepsilon}_{\bar{I}*} - \bar{\varepsilon}_a} - \frac{\bar{W}_{\bullet\bullet I*a}^{jk} \bar{W}_{\bullet\bullet jk}^{I*a}}{\bar{\varepsilon}_j + \bar{\varepsilon}_k - \bar{\varepsilon}_{I*} - \bar{\varepsilon}_a} \right) \qquad (7.14)$$

Table 7.1: (continued)

$$
\begin{aligned}
\text{Lam Triples} \;=\;& \sum_{Ikc} \frac{\lambda_I \bar{F}^k_{\bullet c} \bar{W}^{I^*\bar{I}^*c}_{\bullet\bullet\bullet I\bar{I}k}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{\bar{I}} + \bar{\varepsilon}_k - \bar{\varepsilon}_{I^*} - \bar{\varepsilon}_{\bar{I}^*} - \bar{\varepsilon}_c} \\
+\;& \frac{1}{2} \sum_{Ibck} \frac{\lambda_I \bar{W}^{\bar{I}^*k}_{\bullet\bullet bc} \bar{W}^{I^*bc}_{\bullet\bullet\bullet I\bar{I}k}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{\bar{I}} + \bar{\varepsilon}_k - \bar{\varepsilon}_{\bar{I}^*} - \bar{\varepsilon}_b - \bar{\varepsilon}_c} \\
-\;& \frac{1}{2} \sum_{Ibck} \frac{\lambda_I \bar{W}^{I^*k}_{\bullet\bullet bc} \bar{W}^{\bar{I}^*bc}_{\bullet\bullet\bullet I\bar{I}k}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{\bar{I}} + \bar{\varepsilon}_k - \bar{\varepsilon}_{\bar{I}^*} - \bar{\varepsilon}_b - \bar{\varepsilon}_c} \\
-\;& \frac{1}{2} \sum_{Ijkc} \frac{\lambda_I \bar{W}^{jk}_{\bullet\bullet \bar{I}c} \bar{W}^{I^*\bar{I}^*c}_{\bullet\bullet\bullet Ijk}}{\bar{\varepsilon}_I + \bar{\varepsilon}_j + \bar{\varepsilon}_k - \bar{\varepsilon}_{I^*} - \bar{\varepsilon}_{\bar{I}^*} - \bar{\varepsilon}_c} \\
-\;& \frac{1}{2} \sum_{Ijkc} \frac{\lambda_I \bar{W}^{jk}_{\bullet\bullet Ic} \bar{W}^{I^*\bar{I}^*c}_{\bullet\bullet\bullet \bar{I}jk}}{\bar{\varepsilon}_{\bar{I}} + \bar{\varepsilon}_j + \bar{\varepsilon}_k - \bar{\varepsilon}_{I^*} - \bar{\varepsilon}_{\bar{I}^*} - \bar{\varepsilon}_c} \tag{7.15}
\end{aligned}
$$

$$
\text{Lam Quads} \;=\; \frac{1}{4} \sum_{Iklcd} \frac{\lambda_I \bar{W}^{kl}_{\bullet\bullet cd} \bar{W}^{I^*\bar{I}^*cd}_{\bullet\bullet\bullet\bullet I\bar{I}kl}}{\bar{\varepsilon}_I + \bar{\varepsilon}_{\bar{I}} + \bar{\varepsilon}_k + \bar{\varepsilon}_l - \bar{\varepsilon}_{I^*} - \bar{\varepsilon}_{\bar{I}^*} - \bar{\varepsilon}_c - \bar{\varepsilon}_d} \tag{7.16}
$$

there are O^2V^2 such terms. The Lambda doubles occur for a smaller set of pair correlations, those involving at least one orbital that is a member of a PP correlation pair, $\{I\bar{I}\} \rightarrow \{I^*\bar{I}^*\}$. In this notation, capital letters are orbitals that are correlated by PP; i, j, k, ... (lower case) are any occupied orbitals in the reference determinant; and a, b, c,... (lower case) are any virtual orbitals in the reference determinant. In general, the Lambda doubles occur with the opposite sign of the MP2 doubles and are largest when the PP t and λ amplitudes are large. In other words, they help to compensate for the overestimation of the correlation by the MP2 doubles in cases with strong correlations. These terms are also similar to their VOD(2) counterparts, except that the non-local valence doubles absent in PP are treated by the (2) correction, instead of iteratively at the VOD level.

Finally come the Lambda triple and quadruple excitations. These terms provide dynamical correlation for the excited configurations generated in the PP wave function. The triple excitations, of which there are pOV^2, take the form $\{I\bar{I}k\} \rightarrow \{I^*ab\}$, which can also be viewed as a double excitation $\{\bar{I}^*k\} \rightarrow \{ab\}$ from the configuration in which $\{I\bar{I}\}$ has been replaced by $\{I^*\bar{I}^*\}$. Other triples like $\{Ijk\} \rightarrow \{I^*\bar{I}^*a\}$ take

the form of a double excitation from the $\{I^*\bar{I}^*\}$ configuration that reoccupies orbital $\bar{I}$ and one other virtual orbital a, $\{jk\} \rightarrow \{\bar{I}a\}$.

Only a small subset of the full quadruples arise in PP(2), and they primarily describe excitations from the reference such as $\{I\bar{I}J\bar{J}\} \rightarrow \{I^*\bar{I}^*J^*a\}$, which looks like the excitation of $\{J\bar{J}\} \rightarrow \{J^*a\}$ after PP has generated $\{I\bar{I}\} \rightarrow \{I^*\bar{I}^*\}$. Notice that only p^2V such terms exist, so the quadruples are very inexpensive to evaluate. The quadruples also estimate the connected quadruple contribution, e.g. $\{I\bar{I}J\bar{J}\} \rightarrow \{I^*\bar{I}^*J^*\bar{J}^*\}$, to balance the occurrence of disconnected quadruples arising as products of two pairs in the PP wave function. Notice that PP(2) includes no quadruples in the one pair case, since they require at least two active pairs for these terms. This is no different than VOD(2), however.

The triples and quadruples provide the most significant difference between PP(2) and VOD(2). Because the PP wave function produces many fewer doubly-excited determinants than the VOD one, far fewer semi-internal triples and quadruples are included by PP(2). For example, the quadruples in VOD(2) provide pair correlation for a doubly-excited configuration where two of the orbitals in the second correlating pair can lie outside the active space. In PP(2), only one of these orbitals can be outside the active space, and all the other three orbitals must correspond to the same pair.

Thus, at first glance, PP(2) behaves like a hybrid doubles level treatment of the electron correlation, where a linear number of the most important terms are treated at the coupled cluster level using PP, and the rest of the doubles are treated at the MP2 level. However, the MP2-like treatment couples to the correlation of the potentially large PP amplitudes through the similarity transformation and the counter-balancing Lambda singles and doubles terms. It is of course the coupling of amplitudes that differentiates CCSD from MP2 and makes the former much more reliable. These singles and doubles terms on top of PP could then be said to provide a simple and inexpensive approximation to CCSD.[52] PP(2) goes further, though. The triples and quadruples terms then provide dynamical correlation for those doubly excited configurations contained in the PP wave function. This makes it more suitable to the treatment of systems with strong electron-electron correlations than MP2.

Table 7.2: Representative terms for the correlations included in the PP(2) model. Capital letters refer to active pairs in the PP calculation, and lower case letters refer to any occupied (i, j, k) or virtual (a, b, c) orbitals.

Term	Correlation
MP2 singles	$i \rightarrow a$
MP2 doubles	$ij \rightarrow ab$
Lambda singles	$I \rightarrow a$
	$j \rightarrow I^*$
Lambda doubles	$Ij \rightarrow ab$
	$jk \rightarrow I^*b$
Lambda triples	$I\bar{I}j \rightarrow I^*ab$
	$Ijk \rightarrow I^*\bar{I}^*a$
Lambda quadruples	$I\bar{I}J\bar{J} \rightarrow I^*\bar{I}^*J^*\bar{a}$
	$I\bar{I}J\bar{k} \rightarrow I^*\bar{I}^*J^*\bar{J}^*$

We should be clear, however, that PP(2) does not treat all of the configurations equivalently and is therefore inherently a single-reference method. Whereas all double excitations from the reference configuration are included, only the limited set of correlations from the excited PP configurations are included via the triples and quadruples terms. For this reason, we have implemented PP(2) within an unrestricted formalism to ensure correct separation of species. On the other hand, the model is likely to capture the most important pair correlations from the excited configurations and provide much of their dynamical correlation.

7.2.3   Implementation

Because all of the steps after obtaining the integrals scale no worse O^2V^2 (the cost of evaluating the MP2 doubles), the computation of PP(2) is dominated by the fifth-order steps involved in the construction and transformation of the molecular orbital basis integrals, just like MP2. Unlike MP2, for which only $\langle ij\|ab \rangle$ integrals are needed, PP(2) requires most of the full set of two-electron integrals over all spin orbitals, $\langle pq\|rs \rangle$. This, combined with the fact that the integrals must be dressed with a linear or quadratic number of terms (to form $\bar{H}$) and that there are many more types of terms than in standard MP2 theory, PP(2) will scale the same as MP2 but

with a larger prefactor. On the other hand, PP(2) is dramatically less expensive than VOD(2), for which there are sixth and seventh-order steps involved in computing the (approximate) second-order energy.

The computation of an electronic structure method that is dominated by the construction of the two-electron integrals is the ideal situation for employing the RI approximation. In addition to the computational savings obtained with minimal loss of accuracy, the RI approximation facilitates the development of efficient algorithms. The standard RI-MP2 algorithm[141] and our own RI-PP(2) algorithm form the integral $\langle ij|ab \rangle$ as,

$$\langle ij|ab \rangle = (ia|jb) = \sum_K^{aux} B_{ia}^K B_{jb}^K. \tag{7.17}$$

Thus, we are able to form and store the B_{pq}^K matrices on disk using only a cubic amount of storage (without even exploiting sparsity), instead of the typical quartic storage required to store the full set of two-electron integrals. The integrals can be formed as needed by reading in blocks of B_{pq}^K in batches and performing matrix multiplies. The integrals are then used and discarded.

Consider the construction of the necessary integrals and the evaluation of the MP2 doubles terms, which dominates the computation. Our algorithm is nearly identical to an RI-MP2 one,[141] except for the integral dressing steps. It proceeds as written in Table 7.3. Our PP(2) algorithm is therefore fifth order, dominated by the matrix multiplication to form the integrals from the B_{pq}^K. It also requires quartic I/O to read the B_{pq}^K from disk. Finally, the loops are designed to utilize only quadratic memory, but they could be optionally batched to take advantage of all of the available memory.

In fact, the Lambda doubles terms require many of the same integrals, particularly the dressed $\bar{W}_{\bullet\bullet Ij}^{ab}$, where orbital I is active. Rather than rebuilding these integrals later or storing them on disk, we simply evaluate the Lambda doubles energy contribution simultaneously with the MP2 doubles. While this complicates the coding slightly, it does not alter the algorithm fundamentally.

The Lambda triples terms are also computationally intensive. Though there are only pOV^2 triples, instead of O^2V^2 terms for the MP2 doubles, their cost is comparable to that of the MP2 doubles since we usually take all valence orbitals to be active.

Table 7.3: MP2 doubles algorithm using the RI approximation.

```
Dress and store F̄ᵖ•q in memory
loop over i
   Read B(X,a,i) for all X and a from disk
   loop over j
      Read B(X,b,j) for all X and b from disk
      Form (ia|jb) = Σₓ^Aux B(X,a,i)B(X,b,j), antisymmetrize
      Dress (ia‖jb) to form W̄ab••ij, W̄ij••ab
      loop over a
         loop over b
         E⁽²⁾MP2 Doubles += ¼W̄ab••ij W̄ij••ab/(F̄ⁱ•i + F̄ʲ•j − F̄ᵃ•a − F̄ᵇ•b)
         end loop
      end loop
   end loop
end loop
```

In contrast, the singles and quadruples terms require minimal computational effort. For $C_{20}H_{42}$ in the cc-pVDZ basis and its corresponding auxiliary basis,[64, 151] the singles and quadruples consumed only 4% of the computational time, while the doubles and triples terms consumed 84% of the time, and the rest went to the formation of B_{pq}^K and other overhead.

The most difficult aspect in implementing PP(2) efficiently is the proper treatment of the dressing terms to minimize I/O bottlenecks. In a loop structure like the one given above, forming $\bar{W}_{\bullet\bullet ij}^{ab}$ for all a, b with a given i, j will include dressing terms for one particular a, b element where a and b correspond to some particular active orbitals from PP pairs. A naive implementation of these terms requires two disk seeks to load two vectors to form a single integral, which is extremely wasteful (since a seek takes roughly 10^4 times longer than reading a double precision number). However, there are usually no more than a quadratic number of such terms, so an efficient algorithm must build these in advance using fewer disk seeks and store them in memory or on disk until they are needed. Our implementation is partially optimized with regard to this issue, though further work could be done to bring down the computational time. Nevertheless, our current implementation demonstrates the affordability of PP(2) as compared to MP2 or other coupled cluster methods.

7.3   Results and discussion

PP(2) was implemented in a developmental version of Q-Chem[85], and all calculations reported here were performed using this version of Q-Chem. All PP and PP(2) calculations use the RI approximation. The correlation consistent basis sets cc-pVXZ[64] used in PP and PP(2) are combined with their corresponding auxiliary basis sets[151] for the RI approximation. Here we will examine a variety of cases with PP(2) and show that it significantly improves upon MP2 in many of them. First, however, we examine the efficiency of PP(2).
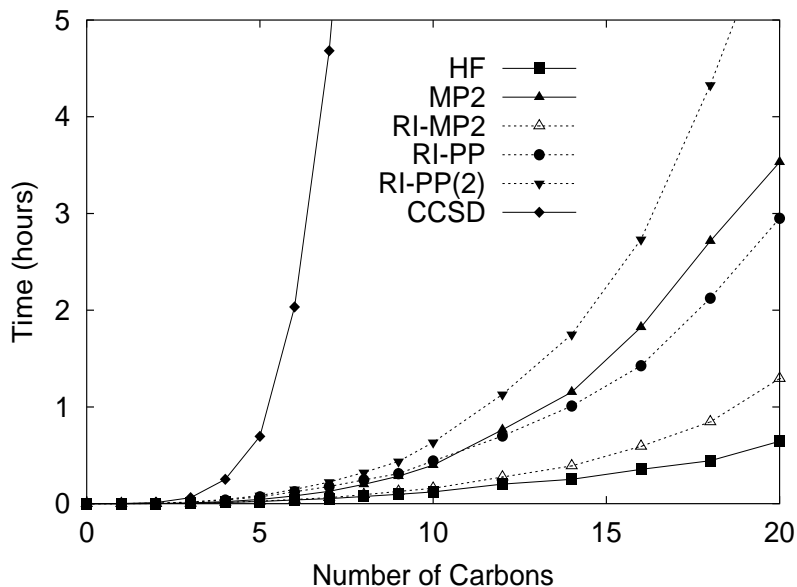
Figure 7.1: Total CPU timings on linear alkane chains comparing PP(2) against other standard methods. In this basis set, $C_{20}H_{42}$ corresponds to 490 basis functions.

7.3.1 Timings

CPU timings on linear alkanes on a 2.0 GHz Mac G5 XServe with 2 gigabytes of RAM are presented in Figure 7.1. All calculations are unrestricted and ignore point-group symmetry. Though not plotted explicitly here, an RI-(2) calculation (excluding HF and PP) is more expensive than a comparable RI-MP2 one (excluding HF) because of the additional terms in PP(2). On the other hand, for alkanes up to around $C_{18}H_{38}$ (roughly 450 basis functions), RI-(2) costs less than canonical MP2.

Comparing the time for the full PP(2) job versus a full MP2 one in Figure 7.1, we see that PP(2) costs a little more than canonical MP2. This is because the cost of the PP calculation is several times beyond that of the initial HF calculation. However, for $C_{20}H_{42}$, the total PP(2) time was only 76% more than that for MP2. PP(2) also costs far less than CCSD or other standard coupled cluster methods. Clearly, PP(2) is very affordable for systems with tens of heavy atoms and up to the high hundreds of basis functions.

7.3.2 Reaction energies

One of the most important uses of electronic structure theory is the prediction of reaction energetics. Table 7.4 lists the energetics for a set of thirteen isogyric reactions from Ref. [112] using geometries obtained from Ref. [63]. It compares these energies against CCSD(T) in the cc-pVTZ basis. For reaction energetics, the static correlation methods PP and VOD are no better than HF, since they lack the important dynamical correlation terms. All three of them have mean absolute deviation (MAD) versus CCSD(T) of at least 15 kcal/mol. Adding in the dynamical correlation through MP2 or (2) significantly improves the predicted energies. MP2 predictions have a MAD of only 4.6 kcal/mol, and PP(2) improves this to 3.8 kcal/mol, virtually the same as CCSD! Of course, VOD(2), with its more complete treatment of valence correlation and triples and quadruples (by having a more complete reference wave function), performs even better, with a MAD of 2.4 kcal/mol versus CCSD(T).

The worst case for PP(2) is the reaction:

$$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O. \tag{7.18}$$
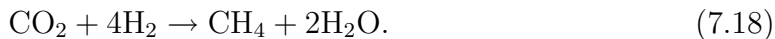
Its large error of 13 kcal/mol is primarily associated with CO_2, for which valence inter-pair correlations are important. PP recovers only 38% of the VOD correlation energy for CO_2 instead of the more typical range of 50-75%.[53] A simple self-consistent treatment of interpair correlations using the imperfect pairing (IP) model[76] raises this recovery to 84% of VOD, which is quite typical for IP. Given the poor PP reference in this molecule, it is not surprising that PP(2) misbehaves. In general, PP(2) behaves well as long as PP provides at least a reasonable description of the system, similarly to how the quality of MP2 depends strongly on the appropriateness of the HF description.

Table 7.4: Errors in reaction energies (in kcal/mol) as compared against CCSD(T) for a set of isogyric reactions in the cc-pVTZ basis. All calculations are unrestricted.

	HF	PP	VOD	MP2	PP(2)	CCSD	VOD(2)	CCSD(T)
$CO+H_2 \rightarrow CH_2O$	6.8	0.9	7.1	0.4	1.0	-0.3	-1.3	3.7
$N_2+3H_2 \rightarrow 2NH_3$	4.3	14.4	17.6	0.5	1.1	-2.0	2.0	-36.5
$C_2H_2+H_2 \rightarrow C_2H_4$	-0.9	0.9	12.6	2.1	1.4	-0.9	3.0	-50.2
$CO_2+4H_2 \rightarrow CH_4+2H_2O$	0.2	-4.1	16.8	2.4	-13.0	-4.3	2.5	-55.9
$CH_2O+2H_2 \rightarrow CH_4+H_2O$	0.7	9.8	9.0	-1.3	-0.2	-1.6	0.4	-61.1
$CO+3H_2 \rightarrow CH_4+H_2O$	7.5	10.7	16.1	-1.7	0.7	-1.9	-0.9	-64.9
$HCN+3H_2 \rightarrow CH_4+NH_3$	-0.8	10.6	20.2	1.2	2.5	-2.5	3.4	-76.6
$H_2O_2+H_2 \rightarrow 2H_2O$	-36.3	11.3	-11.0	-10.1	-4.6	-10.0	-2.4	-127.2
$HNO+2H_2 \rightarrow H_2O+NH_3$	-6.7	6.3	-0.1	-5.6	-4.0	-3.4	-1.1	-173.7
$C_2H_2+3H_2 \rightarrow 2CH_4$	-2.4	7.6	21.1	0.9	3.5	-2.5	4.6	-109.8
$CH_2(^1A_1)+H_2 \rightarrow CH_4$	24.2	22.8	18.6	-4.5	1.5	1.5	-0.2	-130.5
$F_2+H_2 \rightarrow 2HF$	-13.3	11.4	-10.0	-7.6	-4.0	-3.4	-0.8	-131.7
$2CH_2(^1A_1) \rightarrow C_2H_4$	49.8	38.8	28.8	-7.9	0.9	4.7	-1.3	-201.4
Mean Abs. Deviation	15.4	15.0	18.9	4.6	3.8	3.9	2.4	

7.3.3   Bond-breaking in N_2

One particular strength of the PP model is its behavior during bond-breaking. Unlike restricted MP2 (RMP2) or many other approximate restricted coupled cluster methods, restricted PP does not exhibit pathological non-variational behaviors as we stretch the N_2 bond, as shown in Figure 7.2. We also present total energies for selected methods in Table 7.5 to facilitate the reproduction of our work. Full Configuration Interaction (FCI) results were obtained from Ref [152]. RMP2 turns over only slightly beyond the equilibrium region. RCCSD(T) behaves well further out, but still turns over eventually as the underlying RHF reference cannot describe the bond-breaking correctly. RCCSD is the most robust of the standard single-reference methods, but it has levelled out and will turn over just beyond the range of the plot. In contrast to these methods, RPP (with all electrons correlated to ensure that the same PP pairs are correlated across the potential energy surface) is almost perfectly parallel to FCI. Clearly, however, the PP wave function is much too high in energy, and it does not dissociate to the proper atomic products. Unlike the other methods with dynamical correlation, restricted PP(2) also does not diverge at large bond lengths. However, with the restricted wave function PP(2) cannot dissociate to the proper atomic products, so it overestimates the binding energy.

The difficulty faced by restricted wave functions in breaking chemical bonds is well known, so perhaps a fairer test is to consider the unrestricted models which properly dissociate to the atomic species, as plotted in Figure 7.3 and listed in Table 7.5. Now all of the methods behave reasonably at long distances and near equilibrium, though UMP2 in particular predicts a very inaccurate potential energy surface in the intermediate regime. UPP now underbinds N_2, since it unrestricts to lower the energy at long range but not near equilibrium. On the other hand, UPP(2) is much more parallel to FCI than RPP(2) and much improved over UMP2. Though the absolute error versus FCI for UPP(2) is larger than for UCCSD or UCCSD(T), the relative errors are similar. Clearly PP(2) is vastly superior to MP2 for this challenging bond-breaking problem, and its low computational cost makes it potentially more useful than higher-level coupled cluster methods.

Table 7.5: Total energies for N_2 bond-breaking in the cc-pVDZ basis.

| R(N-N)	HF	PP [7 pairs]	PP(2)
Restricted			
1.0679	-108.955234	-109.038539	-109.232187
1.1208	-108.949377	-109.041257	-109.238721
1.1737	-108.928479	-109.029990	-109.228364
1.2700	-108.866830	-108.989045	-109.186919
1.4288	-108.737382	-108.903786	-109.094884
1.5875	-108.606251	-108.829215	-109.006341
1.7463	-108.487612	-108.775669	-108.936385
1.9050	-108.384780	-108.740261	-108.888575
2.0638	-108.297078	-108.717405	-108.858744
Unrestricted			
1.0679	-108.955234	-109.038577	-109.234279
1.1208	-108.949377	-109.041257	-109.238721
1.1737	-108.930114	-109.029990	-109.228364
1.2700	-108.891633	-108.989045	-109.186919
1.4288	-108.833680	-108.903786	-109.094884
1.5875	-108.790279	-108.837260	-109.009774
1.7463	-108.769959	-108.809995	-108.955919
1.9050	-108.767548	-108.796189	-108.915940
2.0638	-108.771051	-108.791646	-108.909681
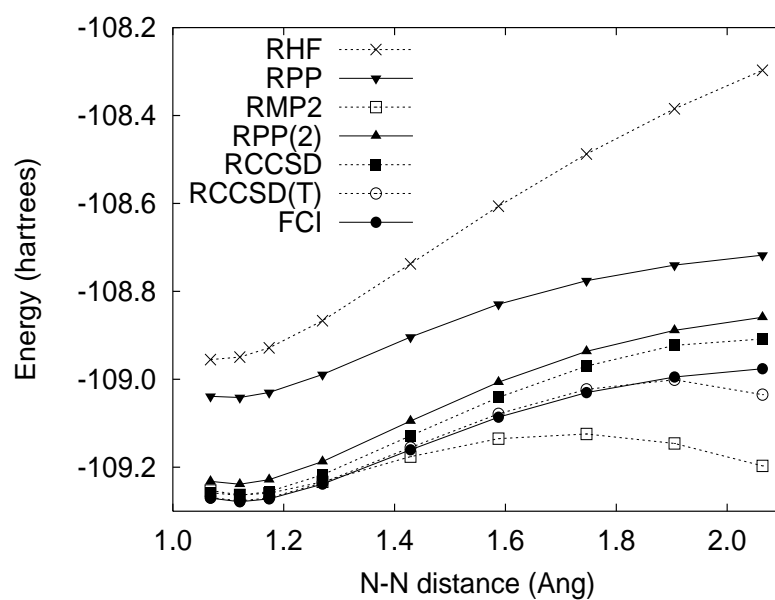
Figure 7.2: N_2 bond-breaking using restricted wave functions in the cc-pVDZ basis.
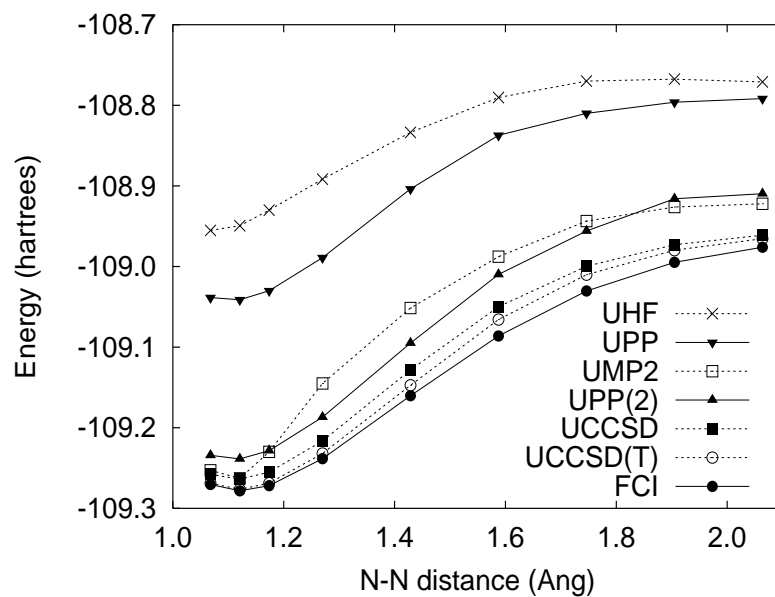
Figure 7.3: N_2 bond-breaking using unrestricted wave functions in the cc-pVDZ basis.

7.3.4   G2 set atomization energies

A more challenging test than reaction energies is posed by the prediction of atomization energies, which expose the large difference between atomic and molecular correlations. We have computed the atomization energies for the 54 small first- and second- row species in the G2 set[177] using the VDZ(d) basis set (which has existing auxiliary basis sets available for all atoms present in the G2 set)[178, 143] and compared against CCSD(T) results. All calculations were unrestricted and broke symmetry whenever it was energetically favorable.

As expected, HF substantially underbinds molecules, since the neglected correlation energy forms a larger fraction of the molecular total energy than the atomic ones. Root-mean-square (rms) atomization energy errors exceed 50% with HF. Not surprisingly, PP also underbinds, though by a fair amount less than HF, with an rms error of 37 kcal/mol, or 28%. The increased binding arises in large part because PP includes only opposite spin correlations. This preferentially stabilizes molecules over atoms, the latter of which tend to have higher spin states and fewer electron pairs. To get a balanced treatment of atomization energies, dynamical correlation is clearly necessary. MP2 significantly improves on HF and PP, with an rms error of 8.0 kcal/mol. PP(2) is actually slightly worse energy-wise, with an rms error of 8.3 kcal/mol. On the other hand, if we look at percentage rms errors, MP2 is almost twice as bad as PP(2) with an error of 15.1% instead of 8.2% because MP2 makes sizable errors on many of the more weakly-bound species, whereas PP(2) errors are much more even percentage-wise. CCSD remains somewhat better, with rms errors around 4.3 kcal/mol and 3.9%. These results suggest that PP(2), though it behaves well as a bond is stretched, is not always as reliable in the atomization energy as it was for N_2.

In many of these species the correlations are not particularly strong, meaning that MP2 should behave reasonably well. Furthermore, because of the unbalanced treatment of high- versus low-spin states in PP which PP(2) must overcome, this test is one of the most difficult for PP(2). However, the advantage of PP(2) over MP2 is clearest in a case like F_2, which has the highest diradical character in this set with a

Table 7.6: Root-mean-square errors versus CCSD(T) atomization energies for G2 set molecules in the VDZ(d) basis.

RMS Error	HF	PP	MP2	PP(2)	CCSD
(kcal/mol)	64.5	36.9	8.0	8.3	4.3
(%)	54.2	27.9	15.1	8.2	3.9

lowest-unoccupied MO (LUMO) occupation number of 0.14 (compared to ~ 0.04 for a typical equilibrium carbon-carbon single bond). CCSD(T) predicts a well depth of 28.6 kcal/mol. MP2 underbinds F_2 by almost 22 kcal/mol (i.e. it is barely bound!), whereas PP(2) predicts it only 1.5 kcal/mol too small. This example is also nice because PP has particular trouble with a species like F_2, underbinding it by almost 23 kcal/mol. F_2 is an example in which interpair correlations play and important role,[120, 53] and it seems that PP(2) is able to overcome this particular deficiency in its reference wave function in this instance.

7.3.5 Radicals

We have recently demonstrated the success of UPP when applied to the challenging prediction of harmonic vibrational frequencies for various diatomic doublet-state radicals which suffer from serious spin and spatial symmetry-breaking effects.[59] In a set of twelve diatomic radicals, UPP in the cc-pVTZ basis cut the errors in UHF predicted frequencies versus experiment by almost half, with a mean absolute deviation of 7.7% instead of 14.4%. UPP also significantly reduced the spin contamination, particularly for the most heavily spin-contaminated species like CN, CO^+, and N_2^+. It also reduced the extent of spatial symmetry breaking in N_2^+. OF and F_2^+, the most challenging species for UPP because of their notorious symmetry-breaking and the importance of interpair correlations, were the only species for which UPP did not improve the predicted frequencies.

The primary goal of improving the reference wave function for these challenging radicals is to enable a stable perturbative correction. As demonstrated in Refs [66, 50], MP2 and CCSD(T) make very inaccurate predictions because of the high degree of symmetry-breaking and spin contamination in these radicals. However, we demon-

strated that improved references significantly stabilized the corrections and made them much more reliable.[50] Here, we apply PP(2) to see if similar improvements can be obtained at much lower cost than at the VOD(2) or CCSD(2) levels. MP2, PP(2), and VOD(2) results, along with those from Ref. [59] are presented in Table 7.7.

Notice first how the symmetry-breaking effects in the UHF reference explode at the MP2 level.[95] MP2 errors range from roughly 40 cm^{-1} in the best two cases to over 1300 cm^{-1} for NO, with half of the frequencies off by more than 300 cm^{-1}! The mean absolute percent error is almost 23%, one and a half times larger than that for HF. UPP makes a significant improvement over UHF, cutting the mean percentage error almost in half. Furthermore, PP(2) uniformly outperforms MP2 and improves upon PP. Whereas PP has a mean absolute percent error of almost 8% (or less than 4% if we exclude the clear aberrations of F_2^+ and OF), PP(2) reduces it to around 5% (or less than 3% if we exclude F_2^+).

The most notable gains are made for the worst-case molecules. For example, the errors in the predicted frequencies for the three most heavily spin-contaminated species, CO$^+$, CN, and N$_2^+$, drop from more than 700 cm^{-1} (33, 40, and 45%, respectively) to 125-150 cm^{-1} (6, 7, and 6%, respectively). Even more impressively, PP(2) reduces the error in the NO frequency from 1324 cm^{-1} (70%) with MP2 to 146 cm^{-1} (8%). Interestingly, although PP itself was very close to VOD for NO, PP(2) is actually much worse than VOD(2) or OD(T) for this species. Even OF, for which PP was not very well-behaved, cleans up nicely at the PP(2) level, with an error of only 9 cm^{-1} (0.9%). Unfortunately, PP(2) does not significantly improve upon F_2^+, with an error of 245 cm^{-1} (22%). The difficulty of UPP with F_2^+ and OF due to the importance of interpair correlations in these species has been noted previously.[59] So it is not surprising that PP(2) does not fare as well in a case where PP is not clearly better than HF.

Comparing instead against VOD(2) or experiment, we see that it is for these poorly behaved molecules that PP(2) fares the worst. Despite the substantial improvements PP and PP(2) provide over HF and MP2 for species like CN, CO$^+$, N$_2^+$, and NO, PP omits significant correlations (primarily those involving the radical electron, and secondarily those between pairs) and cannot correct the wave function as successfully

Table 7.7: Harmonic vibrational frequencies (in cm^{-1}) for various diatomic radicals in the cc-pVTZ basis set as compared to experiment. All calculations are spin unrestricted.

	Expt.	HF	PP	VOD	MP2	PP(2)	VOD(2)	CCSD[a]
CH	2858.5[d]	215	-59	-63	109	17	-31	-6
OH	3737.8[c]	301	-30	-11	98	-1	-17	-40
FH$^+$	3090[b]	233	90	62	79	38	73	76
BO	1886[b]	195	113	40	43	6	19	44
CN	2068.6[c]	-54	19	31	839	128	-4	89
CO$^+$	2169.8[c]	239	187	68	707	151	52	124
N$_2^+$	2207[b]	-326	-81	42	981	137	-5	127
CF	1308[b]	110	-5	-5	46	31	19	39
NO	1904.2[c]	319	62	62	1324	146	8	86
O$_2^+$	1904.7[c]	591	153	123	-383	39	3	128
OF	1053[b]	157	-231	6	313	-9	15	52
F$_2^+$	1104[d]	479	-370	104	-218	245	27	123
MAD (%)		14.7	7.7	2.8	22.6	4.7	1.1	4.4

[a] Reference [50] [b] Reference [80]. [c] Reference [81].

[d] Reference [84].

as VOD. The better description of these types of correlation in VOD(2), along with a more complete description of triple and quadruple excitations, reduces their errors to about 1%. Of course, given the vast difference in computational costs between VOD(2) and PP(2), and the relatively modest difference between the costs of PP(2) and MP2, this approximation appears to provide a good balance between cost and reliability. Furthermore, when one considers the CCSD results, we see that CCSD also has some difficulty with these hardest species (CO$^+$, CN, N$_2^+$, O$_2^+$, and F$_2^+$), suggesting that a significant part of the discrepancy between PP(2) and VOD(2) is due to the additional triples and quadruples described by VOD(2). In fact, PP(2) is fairly comparable with CCSD overall here, but at a lower cost. The improvement of PP(2) over MP2 recommends it for systems for which these more reliable advanced coupled cluster methods are not affordable.

7.3.6 Benzene and symmetry-breaking

One notorious short-coming of the PP model is its failure to predict the D_{6h} symmetric structure of benzene. Instead, it favors the D_{3h} structure with alternating single and double bonds because it is impossible to treat all six bonds equivalently with only three orthogonal pairs describing the π space. There are two ways to eliminate this symmetry breaking. The first is to switch to non-orthogonal orbitals, though this makes the theories more complex.[105] Alternatively, one can include interpair correlations to restore the orbital invariance found in VOD or CCSD. Including even a limited set of interpair correlations substantially reduces the degree of symmetry-breaking, though only a complete set of doubles correlations for the six orbitals/electrons will completely restore symmetry.[76, 53]

Here, we use PP(2) to include those missing correlations and find that PP(2) correctly favors the D_{6h} symmetric structure, as plotted in Figure 7.4. In fact, the PP(2) potential energy slice is very similar to the MP2 one. There is, of course, a small artifactual cusp at the symmetric structure due to the symmetry-breaking in the PP wave function. Nevertheless, PP(2) clearly helps overcome this deficiency in PP.

PP also breaks symmetry for the allyl radical, C_3H_5, which has three equivalent electrons in the π space.[59] The fact that PP can only correlate two of the three electrons leads it to break symmetry by a little less than 1 kcal/mol, with one bond about 0.05Å longer than the other in the cc-pVDZ basis. It does break symmetry less severely than does restricted open shell HF, which prefers a structure with a 0.1Å difference in bond length by almost 2 kcal/mol. On the other hand, UHF and UMP2 (or UCCSD) have no trouble predicting a symmetric structure. When we apply dynamical correlation to PP for this species, it is not quite able to predict the symmetric structure, though the deviation from symmetric is only 0.02Å and less than 0.1 kcal/mol. Unsurprisingly, PP(2) also exhibits a cusp at the symmetric point where the two symmetry-broken solutions cross.
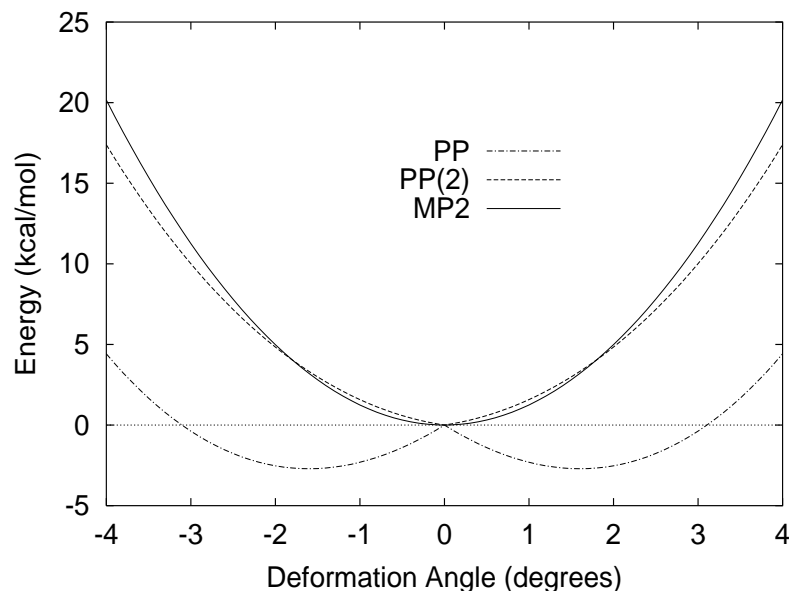
Figure 7.4: The deformation of benzene, in the cc-pVDZ basis. PP uses only three pairs corresponding to the π space.

7.3.7   H + H$_2$ → H$_2$ + H

Of particular interest among strongly-correlated systems are transition states, which can be quite difficult to predict reliably. We examine the simplest chemical reaction, H + H$_2$ → H$_2$ + H. All calculations are unrestricted, and the resulting potential energy surface along the reaction coordinate is plotted in Figure 7.5. Despite its apparent simplicity, this is a challenging case for PP. The limited correlations allowed in PP means that it cannot treat the three electrons equivalently at the transition state. Instead, proceeding along the reaction as written above, the wave function must change discontinuously from pairing the electrons on the right two hydrogen atoms to pairing those on the left two hydrogen atoms, leading to a cusp at the PP transition state. The barrier is also roughly a factor of two too-high. In contrast, HF and VOD predict a smooth barrier. In particular, VOD is able to correlate both possible pairs across the potential energy surface, so it can transition smoothly between the two species.

When we apply dynamical correlation via MP2, PP(2), or VOD(2), the barrier
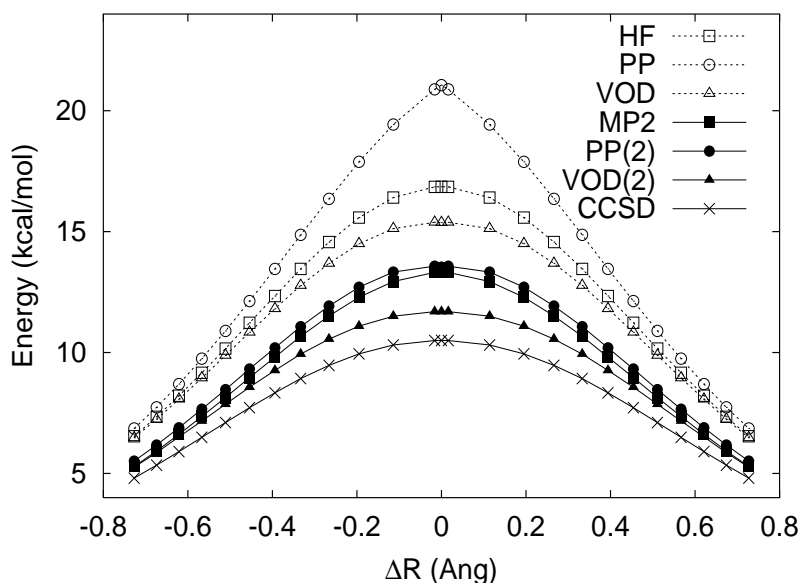
Figure 7.5: Potential energy surface for the $H + H_2 \rightarrow H_2 + H$ reaction in the cc-pVDZ basis along the CCSD reaction coordinate. The energy scale is relative to the reactants at infinite separation.

drops significantly. The MP2 barrier is still a few kcal/mol above the VOD(2) one, which itself is about 2 kcal/mol above the CCSD(T) one (not shown). PP(2) makes a valiant effort to obtain the correct PES despite the woeful PP reference, bringing the barrier height very close to that predicted by MP2. Unfortunately, PP(2) dips slightly and retains a small cusp at the transition state. In fact, this result is completely analogous to the allyl radical described above, which also suffers at the PP and PP(2) levels because of its three equivalent electrons.

7.3.8 Highly-correlated species: the $[\text{TCNE}]^-$ dimer.

Finally, we consider an interesting organic radical, the tetracyanoethylene [TCNE] radical anion. The species is ethylene with the hydrogens replaced by cyanide groups, and an excess electron in the π system. In the solid state, this species dimerizes in several ways, but the most interesting one is a π stacking-like interaction that forms an unusually long four-center, two-electron carbon-carbon bond.[179]

This dimer is particularly challenging to describe because this long carbon-carbon bond is very diradicaloid, with a PP occupation number in the lowest-unoccupied MO (LUMO) of just under 0.20 at 2.6Å and 0.36 at 2.9Å, the experimental dimer separation. In addition, dispersion effects between the two π systems are important in the binding. This means that standard wave function methods like MP2 cannot be trusted for this problem due to the very strong correlation, and density functional theory with existing functionals cannot be used because it fails to capture dispersion interactions. Finally, its large size, $C_{12}N_8$, makes more elaborate methods prohibitively expensive.

Jung and Head-Gordon recently showed that one-pair PP describing the dimer bond produces a fully repulsive binding curve, emphasizing the importance of dispersion effects.[138] They used a model system with a single dimer and capped it with two potassium counter ions to stabilize the structures. When they combined this PP wave function with multi-reference perturbation theory, they predicted a dimer bound by -17.3 kcal/mol in the 6-31G* basis at a distance of 2.6Å. Counter-poise-correcting for basis set superposition error (BSSE), these numbers were revised to -11.2 kcal/mol and 2.7Å, which is in reasonable agreement with the solid-state experimental data that suggests a binding enthalpy of -8.8 kcal/mol at 2.9Å.[180]

We repeated their calculations with PP(2) using one pair and restricting all orbitals except those in the pair, to enable them to dissociate to the separate atomic species to reach twice the restricted open-shell HF (ROHF) and RI-RMP2 monomer energies. We also used the VDZ(d) basis for which an auxiliary fitting set is available for the dimer and potassium counter ions. Doing so, we find that, as expected, RI-PP is completely repulsive. Likewise, PP(2) results predicted a binding energy of -15.3 kcal/mol at an optimal dimer separation of 2.6Å, in excellent agreement with the multi-reference perturbation results. Having demonstrated that PP(2) can reproduce the correct behavior of the TCNE dimer, in the future it would be interesting to examine how it behaves with all electron pairs correlated and using an unrestricted formalism. In this case, PP was used to generate a two-configuration wave function, and is therefore equivalent to CASSCF(2,2) (ignoring the RI approximation). However, we could correlate all 83 electron pairs using PP(2) and investigate what ef-

fect this has on the potential energy surface. Such calculations are utterly impossible with CASSCF and multi-reference perturbation theories or other common correlation methods, and they will be the subject of future studies.

7.4   Conclusions

We have presented a low-cost perturbative correction for perfect-pairing coupled cluster wave functions based on similarity-transformed perturbation theory. The computational requirements for PP(2) scale with the fifth power of system size, the same as canonical MP2, and overall it costs only up to a factor of several times more than MP2. This makes it applicable to systems with hundreds of active electrons and tens of non-hydrogen atoms.

We have demonstrated that PP(2) significantly improves upon MP2 in many difficult problems, including those of bond-breaking, the treatment of radicals suffering from symmetry-breaking effects, and reaction energetics. It is also able to treat highly-correlated species such as the TCNE dimer. This last example is particularly significant, because of the inability of DFT and MP2 to treat it reasonably. PP(2) is the simplest single reference method which can correctly predict the potential energy surface for the TCNE dimer.

Despite the large number of problems for which PP(2) is likely to be successful, it needs to be tested more widely. For example, it is likely to have trouble with many transition states. We saw the difficulty PP had with the reaction of H_2 + H, and we know that certain classes of highly-delocalized transition states involving multiple electron pairs, such as those found in organic electrocyclic reactions like the Cope rearrangement, are very difficult for pair-based local correlation models.[53]. It is not clear that PP(2) can adequately compensate for the deficiencies in PP in such systems. Further study in this regard is necessary

Nevertheless, there exist large classes of problems for which DFT with current functionals is inadequate due to either very strong correlations or dispersion effects. Similarly, MP2 fails outright for highly-correlated systems. More robust coupled cluster methods may perform better in such situations, but they are usually too

computationally expensive. It is exactly this niche which may be filled by PP(2).

Chapter 8

Conclusions

Though mean-field Hartree-Fock theory (HF) is extremely successful in many well-behaved molecules near their equilibrium structures, when the electron correlations in the molecule become more important, the performance of HF degrades. In our case study on the diatomic radicals (Chapter 2), we saw just such an example where the HF reference provided a poor starting point for further correlations. This degradation can lead to highly erroneous predictions from even nominally very accurate post-HF methods, since they are still based on HF orbitals.

Clearly, more elaborate multireference methods like CASSCF and CASPT2 can overcome the limitations of HF, but their extreme computational cost limits them to very small systems with no more than 14 active electrons. With the goal of developing a method that can describe strong correlations while remaining highly computationally affordable, we have developed a two-step approach PP(2).

PP(2) replaces the HF piece with an inexpensive perfect pairing (PP) calculation that incorporates a linear number of the leading pair correlations into the reference wave function to account for the static correlation. In many cases, these few terms are enough to overcome the deficiencies in HF and enable PP to provide a more stable reference wave function. Applying a second-order perturbative correction, (2), to account for dynamical correlation effects makes PP(2) quantitative. Our efficient implementation of PP(2) using the resolution of the identity approximation makes it cost only a few times more than canonical MP2, and PP(2) is applicable to systems

with tens of non-hydrogen atoms.

We have demonstrated that PP(2) often performs extremely well, particularly given its simplicity. For example, it reduces errors in harmonic vibrational frequencies in highly symmetry-broken species by up to an order of magnitude over MP2. Likewise, a small set of standard reaction energies was predicted very well by PP(2). It also provides a very good description of diradicaloid species and bond-breaking. The tetracyanoethylene dimer dianion provided the perfect case for the need for a method like PP(2). The dimer forms a very diradicaloid bond, which HF and MP2 cannot describe well. The bonding is also strongly influenced by π dispersion interactions between the two monomers, meaning that DFT (with current functionals) cannot describe it either.

On the other hand, we have also probed some of the limits of PP and PP(2). Due to its constrained description of correlation effects, PP exhibits a tendency to break symmetry in species with delocalized electronic structure. We have seen how poorly PP behaved for benzene and the Cope rearrangement. It also has trouble with three-electron systems like the transition state for the reaction $H + H_2 \rightarrow H_2 + H$ or the allyl radical, for example. In both cases, the inability of PP to correlate three electrons equivalently leads to spurious predictions that PP(2) can not quite overcome.

Clearly, further testing of PP and PP(2) is necessary to understand for which systems it behaves reasonably. This is the trade-off we face for the extreme computational efficiency obtained by neglecting so many correlation terms in PP and PP(2). As we saw in Chapter 4, however, much better recovery of the valence correlation energy is obtained by IP and DIP with only a quadratic number of correlating amplitudes. Unrestricted IP, for example, would likely eliminate many of the flaws in UPP for radicals, since it would correlate the unpaired electrons. Combined with dynamical correlation, IP(2) could likely provide a more robust treatment than PP(2). On the other hand, a full implementation of IP(2) should scale with the sixth power of system size, instead of the fifth power in PP(2), making it far less affordable for large systems. Of course, one could approximate the expensive triple and quadruple excitation terms in IP(2) in some fashion to reduce this scaling.

Alternatively, one might prefer to develop an efficient, non-orthogonal formulation of PP.[105] Using redundant orbitals would enable the uniform treatment of equivalent centers, as in allyl or benzene. This could overcome the troubling symmetry-breaking cases in PP and make it much more widely usable. On the other hand, developing an efficient non-orthogonal PP method with robust convergence presents its own set of challenges. Even if one of these proposals or some other affordable method eventually supersedes PP(2) for the problems where PP has some difficulty, it is likely that the low cost of PP(2) will make it a method of choice in the near future for treating the many strongly-correlated systems for which it can make useful predictions.

Appendix A

Restricted, closed-shell doubly-ionic pairing (DIP) amplitude equations

The coupled cluster energy and $\hat{T}$ amplitude equations for closed-shell DIP are presented below. They were derived by eliminating the singles and non-local doubles terms from the restricted CCSD equations presented in Ref [133]. The SIP equations can be obtained trivially by setting $t_{ii}^{j^*j^*} = 0$ for all i and j. The IP and PP equations have been presented previously.[58, 76]

- The energy equation:

$$
\begin{aligned}
E_{corr} &= \sum_{i}\langle ii|i^*i^*\rangle t_{ii}^{i^*i^*} + \sum_{ij}(1-\delta_{ij})\Big\{[2\langle ij|i^*j^*\rangle - \langle ij|j^*i^*\rangle]t_{ij}^{i^*j^*} \\
&\quad + [2\langle ij|j^*i^*\rangle - \langle ij|i^*j^*\rangle]t_{ij}^{j^*i^*} + 2\langle ii|i^*j^*\rangle t_{ii}^{i^*j^*} + 2\langle ij|i^*i^*\rangle t_{ij}^{i^*i^*} \\
&\quad + \langle ii|j^*j^*\rangle t_{ii}^{j^*j^*}\Big\}
\end{aligned}
\tag{A.1}
$$

- $\hat{T}$-amplitude equations. For the $t_{ij}^{i^*j^*}$, $t_{ij}^{j^*i^*}$, $t_{ij}^{i^*i^*}$, $t_{ii}^{i^*j^*}$, and $t_{ii}^{j^*j^*}$ equations, $i \neq j$, and P_{ij} is defined as $P_{ij}(x_{ij}) = x_{ij} + x_{ji}$. All sums run over active electron pairs only.

$$
\begin{aligned}
D_{ii}^{i^*i^*}t_{ii}^{i^*i^*} &= \langle ii|i^*i^*\rangle - (a_{ii}^{ii} + \langle i^*i^*|i^*i^*\rangle)t_{ii}^{i^*i^*} \\
&\quad + \sum_{k}\Big\{(1-\delta_{ki})[2a_{ii}^{ik}t_{ik}^{i^*i^*} + a_{ii}^{kk}t_{kk}^{i^*i^*} + 2\langle i^*i^*|i^*k^*\rangle t_{ii}^{i^*k^*} \\
&\quad + 2\langle k^*k^*|i^*i^*\rangle t_{ii}^{k^*k^*} + 2(f_{k^*}^{i^*}t_{ii}^{i^*k^*} - 2f_i^k t_{ik}^{i^*i^*})] - 2e_{k^*}^{i^*}t_{ii}^{i^*k^*}
\end{aligned}
$$

$$-2e_i^k t_{ik}^{i^*i^*} + (2j_{ik^*}^{i^*k} - k_{ik^*}^{ki^*})(2t_{ik}^{i^*k^*} - t_{ik}^{k^*i^*}) - 3k_{ik^*}^{ki^*} t_{ik}^{k^*i^*}$$
$$+ (1 - \delta_{ik})[(2j_{ik^*}^{i^*i} - 4k_{ik^*}^{ii^*})t_{ii}^{i^*k^*} + (2j_{ii^*}^{i^*k} - 4k_{ii^*}^{ki^*})t_{ik}^{i^*i^*}]\} \tag{A.2}$$

$$
\begin{aligned}
D_{ij}^{i^*j^*} t_{ij}^{i^*j^*} =\ & \langle ij|i^*j^*\rangle + (\langle i^*j^*|i^*j^*\rangle + a_{ij}^{ij} - e_{i^*}^{i^*} - e_{j^*}^{j^*} - e_i^i - e_j^j)t_{ij}^{i^*j^*} \\
& + (\langle i^*j^*|j^*i^*\rangle + a_{ij}^{ji})t_{ij}^{j^*i^*} + \langle i^*j^*|i^*i^*\rangle t_{ij}^{i^*i^*} \\
& + a_{ij}^{ii} t_{ii}^{i^*j^*} + \langle i^*j^*|j^*j^*\rangle t_{ji}^{j^*j^*} + a_{ij}^{jj} t_{jj}^{j^*i^*} \\
& + (f_{j^*}^{i^*} t_{ji}^{j^*j^*} + f_{i^*}^{j^*} t_{ij}^{i^*i^*} - f_i^j t_{jj}^{j^*i^*} - f_j^i t_{ii}^{i^*j^*}) \\
& - (e_{j^*}^{i^*} t_{ji}^{j^*j^*} + e_{i^*}^{j^*} t_{ij}^{i^*i^*} + e_i^j t_{jj}^{j^*i^*} + e_j^i t_{ii}^{i^*j^*}) \\
& + P_{ij}\Bigg\{ \sum_k \frac{1}{2}\Big[(2j_{ik^*}^{i^*k} - k_{ik^*}^{ki^*})(2t_{kj}^{k^*j^*} - t_{kj}^{j^*k^*}) - k_{ik^*}^{ki^*} t_{kj}^{j^*k^*} \\
& + (1 - \delta_{kj})\Big[2(j_{ij^*}^{i^*k} - k_{ij^*}^{ki^*})t_{jk}^{j^*j^*} + 2(j_{ik^*}^{i^*j} - k_{ik^*}^{ji^*})t_{jj}^{j^*k^*}\Big]\Big] \\
& - k_{ij^*}^{ij^*} t_{ij}^{i^*j^*} - k_{ii^*}^{ij^*} t_{ij}^{i^*i^*} - k_{ij^*}^{jj^*} t_{jj}^{j^*i^*} - k_{ii^*}^{jj^*} t_{jj}^{i^*i^*}\Bigg\} \tag{A.3}
\end{aligned}
$$

$$
\begin{aligned}
D_{ij}^{j^*i^*} t_{ij}^{j^*i^*} =\ & \langle ij|j^*i^*\rangle + (\langle i^*j^*|i^*j^*\rangle + a_{ij}^{ij} - e_{j^*}^{j^*} - e_{i^*}^{i^*} - e_j^j - e_i^i)t_{ij}^{j^*i^*} \\
& + (\langle i^*j^*|j^*i^*\rangle + a_{ij}^{ji})t_{ij}^{i^*j^*} + (a_{ij}^{ii} - f_j^i - e_j^i)t_{ii}^{j^*i^*} + (a_{ij}^{jj} - f_i^j - e_i^j)t_{jj}^{j^*i^*} \\
& + (\langle i^*i^*|j^*i^*\rangle + f_{i^*}^{j^*} - e_{i^*}^{j^*})t_{ij}^{i^*i^*} + (\langle j^*j^*|j^*i^*\rangle + f_{j^*}^{i^*} - e_{j^*}^{i^*})t_{ij}^{j^*j^*} \\
& + P_{ij}\Bigg\{ \frac{1}{2}(2j_{ij^*}^{j^*i} - k_{ij^*}^{ij^*})(2t_{ij}^{j^*i^*} - t_{ij}^{i^*j^*}) + (j_{ij^*}^{j^*j} - k_{ij^*}^{jj^*})t_{jj}^{j^*i^*} \\
& + (j_{ii^*}^{j^*i} - k_{ii^*}^{ij^*})t_{ij}^{i^*i^*} + (j_{ii^*}^{j^*j} - k_{ii^*}^{jj^*})t_{jj}^{i^*i^*} - \frac{1}{2}k_{ij^*}^{ij^*} t_{ij}^{i^*j^*} \\
& - \sum_k \Big[k_{ik^*}^{ki^*} t_{kj}^{j^*k^*} + (1 - \delta_{kj})\Big(k_{ij^*}^{ki^*} t_{kj}^{j^*j^*} + k_{ik^*}^{ji^*} t_{jj}^{j^*k^*}\Big)\Big]\Bigg\} \tag{A.4}
\end{aligned}
$$

$$
\begin{aligned}
D_{ij}^{i^*i^*} t_{ij}^{i^*i^*} =\ & \langle ij|i^*i^*\rangle + a_{ij}^{ii} t_{ii}^{i^*i^*} + \sum_k (1 - \delta_{ik})\Big[(a_{ij}^{ik} + a_{ij}^{ki})t_{ik}^{i^*i^*} + a_{ij}^{kk} t_{kk}^{i^*i^*}\Big] \\
& + (\langle i^*i^*|i^*j^*\rangle + f_{j^*}^{i^*} - e_{j^*}^{i^*})t_{ij}^{i^*j^*} - (f_i^j + e_i^j)t_{jj}^{i^*i^*} + (\langle i^*i^*|j^*i^*\rangle + f_{j^*}^{i^*} \\
& - e_{j^*}^{i^*})t_{ij}^{j^*i^*} + (\langle i^*i^*|i^*i^*\rangle - 2e_{i^*}^{i^*} - e_i^i - e_j^j)t_{ij}^{i^*i^*} + \langle i^*i^*|j^*j^*\rangle t_{ji}^{j^*j^*} \\
& - \sum_k (1 - \delta_{jk})(f_j^k + e_j^k)t_{ik}^{i^*i^*} + \frac{1}{2}(2j_{ij^*}^{i^*i} - k_{ij^*}^{ii^*})(2t_{ij}^{j^*i^*} - t_{ij}^{i^*j^*}) \\
& - \frac{3}{2}k_{ij^*}^{ii^*} t_{ij}^{i^*j^*} + (j_{ii^*}^{i^*i} - 2k_{ii^*}^{ii^*})t_{ij}^{i^*i^*} + (j_{ij^*}^{i^*j} - 2k_{ij^*}^{ji^*})t_{jj}^{i^*i^*} \\
& + (j_{ii^*}^{i^*j} - 2k_{ii^*}^{ji^*})t_{jj}^{i^*i^*} + \sum_k \Bigg\{ \frac{1}{2}(2j_{jk^*}^{i^*k} - k_{jk^*}^{ki^*})(2t_{ik}^{i^*k^*} - t_{ik}^{k^*i^*}) \\
& - \frac{3}{2}k_{jk^*}^{ki^*} t_{ik}^{k^*i^*} + (1 - \delta_{ik})[(j_{jk^*}^{i^*i} - 2k_{jk^*}^{ii^*})t_{ii}^{i^*k^*} + (j_{ji^*}^{i^*k} - 2k_{ji^*}^{ki^*})t_{ik}^{i^*i^*}]\Bigg\} \tag{A.5}
\end{aligned}
$$

$$
\begin{aligned}
D_{ii}^{i^*j^*} t_{ii}^{i^*j^*} =\ & \langle ii|i^*j^*\rangle + (f_{i^*}^{i^*} - e_{i^*}^{i^*} + \langle i^*i^*|j^*i^*\rangle)t_{ii}^{i^*i^*} + (a_{ii}^{ii} - e_{i^*}^{i^*} - e_{j^*}^{j^*} - 2e_i^i)t_{ii}^{i^*j^*} \\
& + a_{ii}^{jj} t_{jj}^{j^*i^*} + (a_{ii}^{ij} - f_i^j - e_i^j)t_{ij}^{i^*j^*} + (a_{ii}^{ji} + f_i^j - e_i^j)t_{ij}^{j^*i^*} + (f_{j^*}^{i^*} - e_{j^*}^{i^*})t_{ii}^{j^*j^*}
\end{aligned}
$$

$$+ \sum_k \Big\{ (1-\delta_{ik})(\langle i^*j^*|i^*k^*\rangle + \langle i^*j^*|k^*i^*\rangle + f_{i^*}^{j^*} - e_{i^*}^{j^*})t_{ii}^{i^*k^*}$$

$$+\langle i^*j^*|k^*k^*\rangle t_{ii}^{k^*k^*}\Big\} + \frac{1}{2}(2j_{ii^*}^{i^*j} - k_{ii^*}^{ji^*})(2t_{ij}^{j^*i^*} - t_{ij}^{i^*j^*}) - \frac{3}{2}k_{ii^*}^{ji^*}t_{ij}^{i^*j^*}$$

$$+(j_{ii^*}^{i^*i} - 2k_{ii^*}^{ii^*})t_{ii}^{i^*j^*} + (j_{ij^*}^{i^*j} - 2k_{ij^*}^{ji^*})t_{ji}^{j^*j^*} + (j_{ij^*}^{i^*j} - 2k_{ij^*}^{ii^*})t_{ii}^{j^*j^*}$$

$$+\sum_k \Big[\frac{1}{2}(2j_{ik^*}^{j^*k} - k_{ik^*}^{kj^*})(2t_{ik}^{i^*k^*} - t_{ik}^{k^*i^*}) - \frac{3}{2}k_{ik^*}^{kj^*}t_{ik}^{k^*i^*}$$

$$+(1-\delta_{ki})\Big[(j_{ik^*}^{j^*i} - 2k_{ik^*}^{ij^*})t_{ii}^{i^*k^*} + (j_{ii^*}^{j^*k} - 2k_{ii^*}^{kj^*})t_{ik}^{i^*i^*} \Big] \Big] \tag{A.6}$$

$$D_{ii}^{j^*j^*}t_{ii}^{j^*j^*} = \langle ii|j^*j^*\rangle + \sum_k [a_{ii}^{kk}t_{kk}^{j^*j^*} + (a_{ii}^{jk} + a_{ii}^{kj})(1-\delta_{jk})t_{jk}^{j^*j^*}$$

$$+\langle j^*j^*|k^*k^*\rangle t_{ii}^{k^*k^*} + 2\langle j^*j^*|i^*k^*\rangle(1-\delta_{ik})t_{ii}^{i^*k^*}] + 2(f_{i^*}^{j^*} - e_{i^*}^{j^*})t_{ii}^{i^*j^*}$$

$$-2(f_i^j + e_i^j)t_{ij}^{i^*i^*} - 2(e_{j^*}^{j^*} + e_i^i)t_{ii}^{j^*j^*} + (2j_{ii^*}^{j^*j} - k_{ii^*}^{jj^*})(2t_{ij}^{j^*i^*} - t_{ij}^{i^*j^*})$$

$$-3k_{ii^*}^{jj^*}t_{ij}^{i^*j^*} + (2j_{ii^*}^{j^*i} - 4k_{ii^*}^{ij^*})t_{ii}^{i^*j^*} + (2j_{ij^*}^{j^*j} - 4k_{ij^*}^{jj^*})t_{ji}^{j^*j^*}$$

$$+(2j_{ij^*}^{j^*i} - 4k_{ij^*}^{ij^*})t_{ii}^{j^*j^*} \tag{A.7}$$

where the following intermediates are defined:

$$a_{ii}^{ij} = a_{ii}^{ji} = \langle ij|ii\rangle + \sum_k [\langle ij|k^*i^*\rangle + \langle ij|i^*k^*\rangle(1-\delta_{ik})]t_{ii}^{i^*k^*} \tag{A.8}$$

$$(\text{for } i \neq j) \quad a_{ij}^{ik} = \langle ik|ij\rangle + \langle ik|i^*i^*\rangle t_{ij}^{i^*i^*} + \langle ik|i^*j^*\rangle t_{ij}^{i^*j^*} + \langle ik|j^*i^*\rangle t_{ij}^{j^*i^*}$$

$$+\langle ik|j^*j^*\rangle t_{ij}^{j^*j^*} \tag{A.9}$$

$$(\text{for } i \neq j) \quad a_{ij}^{ki} = \langle ki|ij\rangle + \langle ki|i^*i^*\rangle t_{ij}^{i^*i^*} + \langle ki|i^*j^*\rangle t_{ij}^{i^*j^*} + \langle ki|j^*i^*\rangle t_{ij}^{j^*i^*}$$

$$+\langle ki|j^*j^*\rangle t_{ij}^{j^*j^*} \tag{A.10}$$

$$(\text{for } i \neq j) \quad a_{ij}^{kk} = \langle ki|jk\rangle + \langle ki^*|j^*k\rangle(t_{ij}^{i^*j^*} + t_{ij}^{j^*i^*}) + \langle ki^*|i^*k\rangle t_{ij}^{i^*i^*}$$

$$+\langle kj^*|j^*k\rangle t_{ji}^{j^*j^*} \tag{A.11}$$

$$(\text{for } i \neq j) \quad a_{ii}^{jk} = \langle ij|ki\rangle + \sum_m \Big[\langle m^*k|jm^*\rangle t_{ii}^{m^*m^*} + (1-\delta_{im})(\langle jk|i^*m^*\rangle$$

$$+\langle jk|m^*i^*\rangle)t_{ii}^{i^*m^*} \Big] \tag{A.12}$$

$$e_i^j = \langle ji|i^*i^*\rangle t_{ii}^{i^*i^*} + \sum_k (1-\delta_{ik})\Big\{ [2\langle jk|i^*k^*\rangle - \langle jk|k^*i^*\rangle]t_{ik}^{i^*k^*} +$$

$$[2\langle jk|k^*i^*\rangle - \langle jk|i^*k^*\rangle]t_{ik}^{k^*i^*} + [\langle ji|i^*k^*\rangle + \langle ji|k^*i^*\rangle]t_{ii}^{i^*k^*}$$

$$+\langle jk|k^*k^*\rangle t_{ki}^{k^*k^*} + \langle jk|i^*i^*\rangle t_{ik}^{i^*i^*} + \langle ji|k^*k^*\rangle t_{ii}^{k^*k^*} \Big\} \tag{A.13}$$

$$e_{j^*}^{i^*} = \langle ii|j^*i^*\rangle t_{ii}^{i^*i^*} + \sum_k (1-\delta_{ik})\Big\{ [2\langle ik|j^*k^*\rangle - \langle ik|k^*j^*\rangle]t_{ik}^{i^*k^*}$$

$$+[2\langle ik|k^*j^*\rangle - \langle ik|j^*k^*\rangle]t_{ik}^{k^*i^*} + [\langle ik|j^*i^*\rangle + \langle ki|j^*i^*\rangle]t_{ik}^{i^*i^*}$$
$$+\langle kk|j^*k^*\rangle t_{kk}^{k^*i^*} + \langle ii|j^*k^*\rangle t_{ii}^{i^*k^*} + \langle kk|j^*i^*\rangle t_{kk}^{i^*i^*}\Big\} \tag{A.14}$$

$$
\begin{aligned}
j_{ij^*}^{k^*i} &= \langle k^*i|ij^*\rangle + \frac{1}{2}\delta_{ki}\sum_m\Big\{ -\langle im|j^*m^*\rangle t_{im}^{m^*i^*} \\
&+[2\langle im|j^*m^*\rangle - \langle im|m^*j^*\rangle]t_{im}^{i^*m^*} \\
&+(1-\delta_{mi})[\langle im|j^*i^*\rangle - \langle im|i^*j^*\rangle]t_{im}^{i^*i^*}\Big\} \\
&+\frac{1}{2}(1-\delta_{ki})\Big\{ -\langle ik|j^*i^*\rangle t_{ik}^{i^*k^*} + [2\langle ik|j^*i^*\rangle - \langle ik|i^*j^*\rangle]t_{ik}^{k^*i^*} \\
&+[\langle ik|j^*k^*\rangle - \langle ik|k^*j^*\rangle]t_{ik}^{k^*k^*}\Big\} \tag{A.15}
\end{aligned}
$$

$$
\begin{aligned}
j_{ki^*}^{i^*j} &= \langle i^*j|ki^*\rangle + \frac{1}{2}\delta_{ki}\sum_m\Big\{ -\langle jm|i^*m^*\rangle t_{km}^{m^*i^*} \\
&+[2\langle jm|i^*m^*\rangle - \langle jm|m^*i^*\rangle]t_{km}^{i^*m^*} \\
&+(1-\delta_{mi})[\langle ji|i^*m^*\rangle - \langle ji|m^*i^*\rangle]t_{ki}^{i^*m^*}\Big\} \\
&+\frac{1}{2}(1-\delta_{ki})\Big\{ -\langle ji|i^*k^*\rangle t_{ki}^{k^*i^*} + [2\langle ij|k^*i^*\rangle - \langle ij|i^*k^*\rangle]t_{ik}^{k^*i^*} \\
&+ + [\langle jk|i^*k^*\rangle - \langle jk|k^*i^*\rangle]t_{kk}^{k^*i^*}\Big\} \tag{A.16}
\end{aligned}
$$

$$
\begin{aligned}
j_{ij^*}^{i^*k} &= \langle i^*k|ij^*\rangle + \frac{1}{2}\sum_m\Big\{ -\langle km|j^*m^*\rangle t_{im}^{m^*i^*} \\
&+[2\langle km|j^*m^*\rangle - \langle km|m^*j^*\rangle]t_{im}^{i^*m^*} \\
&+(1-\delta_{mi})\Big([\langle ki|j^*m^*\rangle - \langle ki|m^*j^*\rangle]t_{ii}^{i^*m^*} \\
&+[\langle km|j^*i^*\rangle - \langle km|i^*j^*\rangle]t_{im}^{i^*i^*}\Big)\Big\} \tag{A.17}
\end{aligned}
$$

$$
\begin{aligned}
(\text{for } i \neq jk)\quad j_{jk^*}^{i^*i} &= \langle i^*i|jk^*\rangle + \frac{1}{2}\langle ii|k^*j^*\rangle(t_{ij}^{j^*i^*} - t_{ij}^{i^*j^*}) \\
&+\frac{1}{2}(\langle ij|k^*j^*\rangle - \langle ij|j^*k^*\rangle)t_{jj}^{j^*i^*} \\
&+\frac{1}{2}(\langle ij|k^*i^*\rangle - \langle ij|i^*k^*\rangle)t_{jj}^{i^*i^*} \tag{A.18}
\end{aligned}
$$

$$
\begin{aligned}
j_{ji^*}^{k^*i} &= \langle k^*i|ji^*\rangle + \frac{1}{2}\delta_{jk}\sum_m\Big[-\langle im|i^*m^*\rangle t_{jm}^{m^*j^*} \\
&+(2\langle im|i^*m^*\rangle - \langle im|m^*i^*\rangle)t_{jm}^{j^*m^*} \\
&+(1-\delta_{jm})(\langle im|i^*j^*\rangle - \langle im|j^*i^*\rangle)t_{jm}^{j^*j^*} \\
&+(1-\delta_{jm})(\langle ij|i^*m^*\rangle - \langle ij|m^*i^*\rangle)t_{jj}^{j^*m^*})\Big] \\
&+\frac{1}{2}(1-\delta_{jk})\Big[-\langle ik|i^*j^*\rangle t_{jk}^{j^*k^*} + (2\langle ik|i^*j^*\rangle - \langle ik|j^*i^*\rangle)t_{jk}^{k^*j^*} \\
&+(\langle ij|i^*j^*\rangle - \langle ij|j^*i^*\rangle)t_{jj}^{j^*k^*} + (\langle ik|i^*k^*\rangle - \langle ik|k^*i^*\rangle)t_{kj}^{k^*k^*}
\end{aligned}
$$

$$+(\langle ij|i^*k^*\rangle - \langle ij|k^*i^*\rangle)t_{jj}^{k^*k^*}] \tag{A.19}$$

$$
\begin{aligned}
j_{ii^*}^{j^*k} &= \langle j^*k|ii^*\rangle + \frac{1}{2}\delta_{ij}\sum_m\Big[-\langle km|i^*m^*\rangle t_{im}^{m^*i^*}\\
&\quad +(2\langle km|i^*m^*\rangle - \langle km|m^*i^*\rangle)t_{im}^{i^*m^*}\\
&\quad +(1-\delta_{mi})(\langle ki|i^*m^*\rangle - \langle ki|m^*i^*\rangle)t_{ii}^{i^*m^*}\Big]\\
&\quad +\frac{1}{2}(1-\delta_{ij})\Big[(\langle kj|i^*j^*\rangle - \langle kj|j^*i^*\rangle)t_{ji}^{j^*j^*}\\
&\quad +\frac{1}{2}\langle kj|i^*i^*\rangle(t_{ij}^{j^*i^*} - t_{ij}^{i^*j^*})\\
&\quad +\frac{1}{2}(\langle ki|i^*j^*\rangle - \langle ki|j^*i^*\rangle)t_{ii}^{j^*j^*}\Big] \tag{A.20}
\end{aligned}
$$

$$
\begin{aligned}
k_{ij^*}^{ki^*} &= \langle ki^*|ij^*\rangle - \frac{1}{2}\sum_m\Big\{\langle km|m^*j^*\rangle t_{im}^{m^*i^*} + (1-\delta_{mi})[\langle ki|m^*j^*\rangle t_{ii}^{i^*m^*}\\
&\quad +\langle km|i^*j^*\rangle t_{im}^{i^*i^*}]\Big\} \tag{A.21}
\end{aligned}
$$

$$
\begin{aligned}
k_{ik^*}^{ij^*} &= \langle ij^*|ik^*\rangle - \frac{1}{2}\delta_{ij}\sum_m\Big\{\langle im|m^*k^*\rangle t_{im}^{m^*i^*} + (1-\delta_{mi})[\langle ii|m^*k^*\rangle t_{ii}^{i^*m^*}\\
&\quad +\langle im|i^*k^*\rangle t_{im}^{i^*i^*}]\Big\} - \frac{1}{2}(1-\delta_{ij})\Big[\langle ij|i^*k^*\rangle t_{ij}^{i^*j^*} + \langle ij|j^*k^*\rangle t_{ji}^{j^*j^*}\\
&\quad +\langle ii|i^*k^*\rangle t_{ii}^{i^*j^*} + \langle ii|j^*k^*\rangle t_{ii}^{j^*j^*}\Big] \tag{A.22}
\end{aligned}
$$

$$
\begin{aligned}
k_{ik^*}^{kj^*} &= \langle ik^*|kj^*\rangle - \frac{1}{2}\delta_{ij}\sum_m\Big\{\langle km|m^*k^*\rangle t_{im}^{m^*i^*}\\
&\quad +(1-\delta_{mi})[\langle ki|m^*k^*\rangle t_{ii}^{i^*m^*} + \langle km|i^*k^*\rangle t_{im}^{i^*i^*}]\Big\}\\
&\quad -\frac{1}{2}(1-\delta_{ij})\Big[\langle kj|i^*k^*\rangle t_{ij}^{i^*j^*} + \langle ki|i^*k^*\rangle t_{ii}^{i^*j^*} + \langle kj|j^*k^*\rangle t_{ji}^{j^*j^*}\\
&\quad +\langle ki|j^*k^*\rangle t_{ii}^{j^*j^*}\Big] \tag{A.23}
\end{aligned}
$$

$$
\begin{aligned}
k_{ii^*}^{kj^*} &= \langle ii^*|kj^*\rangle - \frac{1}{2}\delta_{ij}\sum_m\Big\{\langle km|m^*i^*\rangle t_{im}^{m^*i^*} + (1-\delta_{mi})[\langle ki|m^*i^*\rangle t_{ii}^{i^*m^*}\\
&\quad +\langle km|i^*i^*\rangle t_{im}^{i^*i^*}]\Big\} - \frac{1}{2}(1-\delta_{ij})[\langle kj|i^*i^*\rangle t_{ij}^{i^*j^*} + \langle ki|i^*i^*\rangle t_{ii}^{i^*j^*}\\
&\quad +\langle kj|j^*i^*\rangle t_{ji}^{j^*j^*} + \langle ki|j^*i^*\rangle t_{ii}^{j^*j^*}] \tag{A.24}
\end{aligned}
$$

$$
\begin{aligned}
(\text{for } i\neq jk)\quad k_{jk^*}^{ii^*} &= \langle ii^*|jk^*\rangle - \frac{1}{2}(1-\delta_{ij})[\langle ii|j^*k^*\rangle t_{ij}^{i^*j^*} + \langle ii|i^*k^*\rangle t_{ij}^{i^*i^*}\\
&\quad +\langle ij|j^*k^*\rangle t_{jj}^{j^*i^*} + \langle ij|i^*k^*\rangle t_{jj}^{i^*i^*}] \tag{A.25}
\end{aligned}
$$

$$
\begin{aligned}
k_{ji^*}^{ki^*} &= \langle ki^*|ji^*\rangle - \frac{1}{2}\delta_{ij}\sum_m\Big[\langle km|m^*i^*\rangle t_{im}^{m^*i^*}\\
&\quad +(1-\delta_{im})(\langle km|i^*i^*\rangle t_{im}^{i^*i^*} + \langle ki|m^*i^*\rangle t_{ii}^{i^*m^*})\Big]
\end{aligned}
$$

$$-\frac{1}{2}(1-\delta_{ij})[\langle ki|j^*i^*\rangle t_{ji}^{j^*i^*} + \langle kj|j^*i^*\rangle t_{jj}^{j^*i^*} + \langle ki|i^*i^*\rangle t_{ij}^{i^*i^*}$$
$$+\langle kj|i^*i^*\rangle t_{jj}^{i^*i^*}] \tag{A.26}$$

Appendix B

Iterative singular value decomposition via Davidson diagonalization

As discussed above, it turns out that one typically only cares about the largest or largest few singular values for the T_n matrix. Moreover, the number of elements in the T_n matrix scales as $O^n V^n$ and therefore a full SVD, scaling as $O^{2n} V^n$, quickly becomes expensive. An ideal algorithm should be computationally inexpensive compared to the underlying coupled cluster calculation, making such analysis feasible for all systems that can be studied with CC theory. A stable, inexpensive, iterative approach for calculating a small number of singular values along the lines of the Davidson diagonalization meets these criteria and is presented here. The computational effort associated with this calculation is almost trivial compared to the underlying CCSD calculation.

As explained in most graduate-level books on numerical matrix computations,[169] the stable way to write the SVD of $T_n = U \Sigma V^\dagger$ as an eigenvalue problem is as the Hermitian matrix,

$$H = \begin{pmatrix} 0 & T_n^\dagger \\ T_n & 0 \end{pmatrix} \tag{B.1}$$

which has the following eigenvalue equation:

$$\begin{pmatrix} 0 & T_n^\dagger \\ T_n & 0 \end{pmatrix} \begin{pmatrix} V & V \\ U & -U \end{pmatrix} = \begin{pmatrix} V & V \\ U & -U \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix}. \tag{B.2}$$

The matrix H can be diagonalized using the Davidson diagonalization (or using any other iterative diagonalization procedure.). Obviously, one does not want to form the full H matrix, as it is extremely large and would quickly become impossible to store in memory. Fortunately, the Davidson procedure requires only the formation of matrix-vector products $\delta_i = H\mathbf{b}_i$ and inner products between vectors in the sets $\{b_i\}$ and $\{\delta_j\}$. These matrix-vector products can be formed by multiplying only the proper rows and columns of the full T_2 matrix without ever constructing the full H matrix. This most expensive step scales as O^2V^2, which is substantially less than the O^2V^4 scaling of CCSD.

Secondly, standard Davidson diagonalization as applied to Configuration Interaction utilizes a preconditioner based on the diagonal elements of the Hamiltonian, which is typically diagonal-dominant. In this case, the H matrix is definitely not diagonal dominant (the diagonal elements are all zero), but nevertheless, we empirically find convergence within roughly 20 iterations using the standard Davidson or Davidson-Liu (to solve for multiple roots simultaneously)[181] procedures. Therefore, the diagonal non-dominance of the H matrix here does not seem to be a major obstacle in the application of this procedure.

Appendix C

Blocks of the similarity-transformed Hamiltonian $\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$ required for PP(2).

In the notation used here, $i, j, k, \ldots$ refer to any occupied orbital, $a, b, c, \ldots$ to any virtual orbital, and capital letters to the orbitals in the pair: I and $\bar{I}$, the alpha and beta occupied orbitals in pair I, and I^* and $\bar{I}^*$, the corresponding alpha and beta virtual orbitals in the pair. Summations run over all occupied, virtuals, or pairs, respectively.

One-body operator $\bar{F}^p_{\bullet q}\{p^\dagger q\}$:

$$\bar{F}^a_{\bullet i} = f^a_{\bullet i} + \delta_{I^*A}f^{\bar{I}}_{\bullet\bar{I}^*}t_I + W^{aI}_{\bullet\bullet I^*\bar{I}^*}t_I - \delta_{J^*A}W^{J\bar{J}}_{\bullet\bullet i\bar{J}^*}t_J \tag{C.1}$$

$$\bar{F}^a_{\bullet b} = f^a_{\bullet b} - \delta_{J^*A}W^{J\bar{J}}_{\bullet\bullet b\bar{J}^*}t_J \tag{C.2}$$

$$\bar{F}^i_{\bullet j} = f^i_{\bullet j} + W^{i\bar{J}}_{\bullet\bullet J^*\bar{J}^*}t_J \tag{C.3}$$

$$\bar{F}^i_{\bullet a} = f^i_{\bullet a} \tag{C.4}$$

- Two-body operator $\bar{W}^{pq}_{\bullet\bullet rs}\{p^\dagger q^\dagger sr\}$:

$$\begin{aligned}
\bar{W}^{ab}_{\bullet\bullet ij} = {} & W^{ab}_{\bullet\bullet ij} + \delta_{I\bar{J}}P(ab)\left(\delta_{\bar{I}^*B}f^a_{\bullet I^*} - \delta_{I^*B}f^a_{\bullet\bar{I}^*}\right)t_I \\
& -\delta_{AK^*}\delta_{A\bar{B}}P(ij)\left(\delta_{J\bar{K}}f^K_{\bullet i} - \delta_{JK}f^{\bar{K}}_{\bullet i}\right)t_K \\
& +P(ab)P(ij)\left(\delta_{AI^*}W^{\bar{I}b}_{\bullet\bullet\bar{I}^*j} - \delta_{A\bar{I}^*}W^{\bar{I}b}_{\bullet\bullet I^*j}\right)t_I
\end{aligned}$$

$$+\delta_{I\bar{J}}W^{ab}_{\bullet\bullet I^*\bar{I}^*}t_I + \delta_{A\bar{B}}\delta_{AK^*}W^{K\bar{K}}_{\bullet\bullet ij}t_K$$
$$+P(ij)P(ab)\delta_{AI^*}\delta_{\bar{B}\bar{J}^*}W^{\bar{I}J}_{\bullet\bullet\bar{I}^*J^*}t_It_J + \delta_{I\bar{J}}\delta_{A\bar{B}}\delta_{AK^*}W^{K\bar{K}}_{\bullet\bullet I^*\bar{I}^*}t_It_K$$
$$-P(ij)\delta_{A\bar{B}}\delta_{BJ^*}W^{\bar{I}\bar{J}}_{\bullet\bullet I^*\bar{I}^*}t_It_J - P(kj)\delta_{I\bar{J}}\delta_{BJ^*}\delta_{AK^*}W^{K\bar{K}}_{\bullet\bullet\bar{K}^*\bar{J}}t_It_K \quad \text{(C.5)}$$

$$\bar{W}^{ab}_{\bullet\bullet ci} = W^{ab}_{\bullet\bullet ci} + \left(\delta_{AI^*}\delta_{A\bar{B}} - \delta_{I^*\bar{A}}\delta_{\bar{A}B}\right)f^{\bar{I}}_{\bullet c}t_I - \delta_{A\bar{B}}\delta_{AK^*}W^{K\bar{K}}_{\bullet\bullet ic}t_K$$
$$+P(ab)\left(\delta_{BI^*}W^{a\bar{I}}_{\bullet\bullet c\bar{I}^*} - \delta_{B\bar{I}^*}W^{a\bar{I}}_{\bullet\bullet c I^*}\right)t_I \quad \text{(C.6)}$$

$$\bar{W}^{ia}_{\bullet\bullet jk} = W^{ia}_{\bullet\bullet jk} - \delta_{J\bar{K}}\left(\delta_{a\bar{J}^*}f^i_{\bullet J^*} - \delta_{aJ^*}f^i_{\bullet\bar{J}^*}\right)t_J - \delta_{J\bar{K}}W^{ai}_{\bullet\bullet J^*\bar{J}^*}t_J$$
$$+P(jk)\left(\delta_{AK^*}W^{i\bar{K}}_{\bullet\bullet j\bar{K}^*} - \delta_{A\bar{K}^*}W^{i\bar{K}}_{\bullet\bullet jK^*}\right)t_K \quad \text{(C.7)}$$

$$\bar{W}^{ab}_{\bullet\bullet cd} = W^{ab}_{\bullet\bullet cd} + \delta_{I^*A}\delta_{A\bar{B}}W^{I\bar{I}}_{\bullet\bullet cd}t_I \quad \text{(C.8)}$$

$$\bar{W}^{kl}_{\bullet\bullet ij} = W^{kl}_{\bullet\bullet ij} + \delta_{I\bar{J}}W^{kl}_{\bullet\bullet I^*\bar{I}^*}t_I \quad \text{(C.9)}$$

$$\bar{W}^{ia}_{\bullet\bullet jb} = W^{ia}_{\bullet\bullet jb} + \left(\delta_{AJ^*}W^{i\bar{J}}_{\bullet\bullet\bar{J}^*b} - \delta_{A\bar{J}^*}W^{i\bar{J}}_{\bullet\bullet J^*b}\right)t_J \quad \text{(C.10)}$$

$$\bar{W}^{ai}_{\bullet\bullet bc} = W^{ai}_{\bullet\bullet bc} \quad \text{(C.11)}$$

$$\bar{W}^{ij}_{\bullet\bullet ka} = W^{ij}_{\bullet\bullet ka} \quad \text{(C.12)}$$

$$\bar{W}^{ij}_{\bullet\bullet ab} = W^{ij}_{\bullet\bullet ab} \quad \text{(C.13)}$$

$$\text{(C.14)}$$

- Required blocks of the three-body operator $\bar{W}^{abc}_{\bullet\bullet\bullet ijk}\{a^\dagger b^\dagger c^\dagger ijk\}$:

$$\bar{W}^{I^*\bar{I}^*c}_{\bullet\bullet\bullet Ijk} = (1-\delta_{I^*c})(1-\delta_{\bar{I}^*c})(1-\delta_{Ij})(1-\delta_{Ik})t_I$$
$$\times\left[-\bar{W}^{\bar{I}c}_{\bullet\bullet jk} + P(jk)\delta_{j\bar{I}}(\bar{W}^{\bar{I}^*c}_{\bullet\bullet\bar{I}^*k} + \bar{W}^{I^*c}_{\bullet\bullet I^*k} - \bar{W}^{Ic}_{\bullet\bullet Ik})\right]$$
$$+\sum_{M\neq I}^{pairs}t_M(\bar{W}^{\bar{I}^*I^*}_{\bullet\bullet\bar{M}^*I} + t_i\bar{F}^{\bar{I}}_{\bullet\bar{M}^*})\delta_{cM^*}(\delta_{j\bar{M}}\delta_{kM} - \delta_{jM}\delta_{k\bar{M}}) \quad \text{(C.15)}$$

$$\bar{W}^{I^*\bar{I}^*c}_{\bullet\bullet\bullet\bar{I}jk} = (1-\delta_{I^*c})(1-\delta_{\bar{I}^*c})(1-\delta_{\bar{I}j})(1-\delta_{\bar{I}k})t_I$$
$$\times\left[\bar{W}^{Ic}_{\bullet\bullet jk} - P(jk)\delta_{jI}(\bar{W}^{I^*c}_{\bullet\bullet\bar{I}^*k} + \bar{W}^{I^*c}_{\bullet\bullet I^*k} - \bar{W}^{\bar{I}c}_{\bullet\bullet\bar{I}k})\right]$$
$$+\sum_{M\neq I}^{pairs}t_M(\bar{W}^{\bar{I}^*I^*}_{\bullet\bullet M^*\bar{I}} - t_i\bar{F}^I_{\bullet M^*})\delta_{c\bar{M}^*}(\delta_{jM}\delta_{k\bar{M}} - \delta_{j\bar{M}}\delta_{kM}) \quad \text{(C.16)}$$

$$\bar{W}^{I^*bc}_{\bullet\bullet\bullet I\bar{I}k} = (1-\delta_{Ik})(1-\delta_{\bar{I}k})(1-\delta_{I^*b})(1-\delta_{I^*c})t_I$$
$$\times\left[\bar{W}^{bc}_{\bullet\bullet\bar{I}^*k} + P(bc)\delta_{\bar{I}^*b}(\bar{W}^{I^*c}_{\bullet\bullet I^*k} - \bar{W}^{\bar{I}c}_{\bullet\bullet\bar{I}k} - \bar{W}^{Ic}_{\bullet\bullet Ik})\right]$$
$$-\sum_{M\neq I}^{pairs}t_M(\bar{W}^{\bar{M}I^*}_{\bullet\bullet\bar{I}I} - t_i\bar{F}^{\bar{M}}_{\bullet\bar{I}^*})\delta_{kM}(\delta_{cM^*}\delta_{b\bar{M}^*} - \delta_{c\bar{M}^*}\delta_{bM^*}) \quad \text{(C.17)}$$

$$\bar{W}^{\bar{I}^*bc}_{\bullet\bullet\bullet I\bar{I}k} = (1-\delta_{Ik})(1-\delta_{\bar{I}k})(1-\delta_{\bar{I}^*b})(1-\delta_{\bar{I}^*c})t_I$$
$$\times\left[-\bar{W}^{bc}_{\bullet\bullet I^*k} - P(bc)\delta_{I^*b}(\bar{W}^{\bar{I}^*c}_{\bullet\bullet\bar{I}^*k} - \bar{W}^{\bar{I}c}_{\bullet\bullet\bar{I}k} - \bar{W}^{Ic}_{\bullet\bullet Ik})\right]$$

$$-\sum_{M\neq I}^{pairs} t_M(\bar{W}^{M\bar{I}^*}_{\bullet\bullet\bar{I}I} + t_i\bar{F}^M_{\bullet I^*})\delta_{kM}(\delta_{cM^*}\delta_{b\bar{M}^*} - \delta_{c\bar{M}^*}\delta_{bM^*}) \quad \text{(C.18)}$$

- Required blocks of the four-body operator $\bar{W}^{abcd}_{\bullet\bullet\bullet\bullet ijkl}\{a^\dagger b^\dagger c^\dagger d^\dagger ijkl\}$:

$$
\begin{aligned}
\bar{W}^{I^*\bar{I}^*cd}_{\bullet\bullet\bullet\bullet I\bar{I}kl} &= P(cd)P(kl)\sum_{M\neq I}^{pairs} t_I t_M \\
&\times \Big\{(1-\delta_{I^*c})(1-\delta_{\bar{I}^*c}) \\
&\times[\delta_{d\bar{M}^*}\delta_{kM}\delta_{l\bar{M}}\delta_{c\in\alpha}(\bar{W}^{\bar{I}^*c}_{\bullet\bullet\bar{I}^*M^*} + \bar{W}^{I^*c}_{\bullet\bullet I^*M^*} - \bar{W}^{\bar{I}c}_{\bullet\bullet\bar{I}M^*} - \bar{W}^{Ic}_{\bullet\bullet IM^*})] \\
&+ (1-\delta_{I^*c})(1-\delta_{\bar{I}^*c}) \\
&\times[\delta_{dM^*}\delta_{k\bar{M}}\delta_{lM}\delta_{c\in\beta}(\bar{W}^{\bar{I}^*c}_{\bullet\bullet\bar{I}^*\bar{M}^*} + \bar{W}^{I^*c}_{\bullet\bullet I^*\bar{M}^*} - \bar{W}^{\bar{I}c}_{\bullet\bullet\bar{I}\bar{M}^*} - \bar{W}^{Ic}_{\bullet\bullet I\bar{M}^*})] \\
&+ (1-\delta_{Ik})(1-\delta_{\bar{I}k}) \\
&\times[\delta_{l\bar{M}}\delta_{cM^*}\delta_{d\bar{M}^*}\delta_{k\in\alpha}(\bar{W}^{\bar{I}M}_{\bullet\bullet\bar{I}k} + \bar{W}^{IM}_{\bullet\bullet Ik} - \bar{W}^{M\bar{I}^*}_{\bullet\bullet k\bar{I}^*} - \bar{W}^{MI^*}_{\bullet\bullet kI^*})] \\
&+ (1-\delta_{Ik})(1-\delta_{\bar{I}k}) \\
&\times[\delta_{lM}\delta_{c\bar{M}^*}\delta_{dM^*}\delta_{k\in\beta}(\bar{W}^{\bar{I}\bar{M}}_{\bullet\bullet\bar{I}k} + \bar{W}^{I\bar{M}}_{\bullet\bullet Ik} - \bar{W}^{\bar{M}\bar{I}^*}_{\bullet\bullet k\bar{I}^*} - \bar{W}^{\bar{M}I^*}_{\bullet\bullet kI^*})]\Big\} \quad \text{(C.19)}
\end{aligned}
$$

Bibliography

[1] W. J. Hehre, R. Ditchfield, and J. A. Pople. Self-consistent molecular orbital methods. xii. further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.*, 56:2257–2261, 1972.

[2] P. C. Hariharan and J. A. Pople. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta*, 28:213–222, 1973.

[3] Steven R. Gwaltney, Edward F. C. Byrd, Troy Van Voorhis, and Martin Head-Gordon. A perturbative correction to the quadratic coupled-cluster doubles method for higher excitations. *Chem. Phys. Lett.*, 353:359–367, 2002.

[4] Srinivasan Parthiban, Glênisson de Oliveira, and Jan. M.L. Martin. Benchmark *ab initio* energy profiles for the gas-phase $s_N 2$ reactions $y^- + ch_3 \rightarrow ch_3 y + x^-$ (x,y=f,cl,br). validation of hybrid dft methods. *J. Phys. Chem. A*, 105:895–904, 2001.

[5] Steven R. Gwaltney, C. David Sherrill, Martin Head-Gordon, and Anna I. Krylov. Second order perturbation corrections to singles and doubles coupled-cluster methods: General theory and application to the valence optimized doubles model. *J. Chem. Phys.*, 113:3548–3560, 2000.

[6] Martin Head-Gordon, Troy Van Voorhis, Gregory J.O. Beran, and Barry D. Dunietz. Local correlation methods. *Computational Science–ICCS 2003, Pt IV, Proceedings Lecture Notes in Computer Science*, 2660:96–102, 2003.

[7] Robert G. Parr and Weitao Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1990.

[8] John P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33:8822–8824, 1986.

[9] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, 1988.

[10] C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, 1988.

[11] A. D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.*, 98:5648–5652, 1993.

[12] Chr. Møller and M. S. Plesset. Note on an approximate treatment for many-electron systems. *Phys. Rev.*, 46:618–622, 1934.

[13] T. Daniel Crawford and Henry F. Schaefer III. An introduction to coupled cluster theory for computational chemists. *Rev. Comp. Chem.*, 14:33–136, 2000.

[14] Jesper Wisborg Krogh and Jeppe Olsen. A general coupled cluster study of the n_2 molecule. *Chem. Phys. Lett.*, 393:578–596, 2001.

[15] Lea Thøgersen and Jeppe Olsen. A coupled cluster and full configuration interaction study of cn and cn$^-$. *Chem. Phys. Lett.*, 393:36–43, 2004.

[16] Troy Van Voorhis and Martin Head-Gordon. Benchmakr variational coupled cluster doubles results. *J. Chem. Phys.*, 113:8873–8879, 2000.

[17] G. D. Purvis, III and Rodney J. Bartlett. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.*, 76:1910–1918, 1982.

[18] K. Raghavachari, G. Trucks, J. A. Pople, and Martin Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.*, 157:479–483, 1989.

[19] John D. Watts, Jürgen Gauss, and Rodney J. Bartlett. Coupled-cluster methods with non-iterative triple excitations for restricted open-shell Hartree-Fock and other general single determinant reference functions. energies and analytical gradients. *J. Chem. Phys.*, 98(11):8718–8733, 1993.

[20] Chris A. White, B.G. Johnson, Peter M. W. Gill, and Martin Head-Gordon. Linear scaling density functional calculations via the continuous fast multipole method. *Chem. Phys. Lett.*, 253:268–278, 1996.

[21] Eric R. Schwegler, Matt Challacombe, and Martin Head-Gordon. Linear scaling computation of the fock matrix. ii. rigorous bounds on exchange integrals and incremental fock build. *J. Chem. Phys.*, 106:9708–9717, 1997.

[22] Christian Ochsenfeld, Chris A. White, and M. Head-Gordon. Linear and sub-linear scaling formation of hartree-fock exchange-type matrices. *J. Chem. Phys.*, 109:1663–1669, 1998.

[23] Yihan Shao, Chandra Saravanan, Martin Head-Gordon, and Chris A. White. Curvy steps for density matrix-based energy minimization: application to large-scale self-consistent-field calculations. *J. Chem. Phys.*, 118:6144–6151, 2003.

[24] Peter Pulay. Localizability of dynamic electron correlation. *Chem. Phys. Lett.*, 100:151–154, 1983.

[25] S. F. Boys. *Quantum Theory of Atoms, Molecules, and the Solid State*, pages 253–280. Academic, New York, 1968.

[26] J. Pipek and P. G. Mezey. A fast intrinsic localization procedure applicable for ab initio and semi-empirical linear combination of atomic orbital wave functions. *J. Phys. Chem.*, 90:4916–4926, 1989.

[27] Svein Saebø and Peter Pulay. Orbital-invariant formulation and second-order gradient evaluation in møller-plesset perturbation theory. *Theor. Chim. Acta*, 69:357–368, 1986.

[28] Svein Saebø and Peter Pulay. Local treatment of electron correlation. *Ann. Rev. Phys. Chem.*, 44:213–236, 1993. and references cited therein.

[29] Robert B. Murphy, Michael D. Beachy, Richard A. Friesner, and Murco N. Ringnalda. Pseudospectral localized møller-plesset methods: Theory and calculation of conformational energies. *J. Chem. Phys.*, 103:1481–1490, 1995.

[30] Richard A. Friesner, Robert B. Murphy, Michael D. Beachy, Murco N. Ringnalda, W. Thomas Pollard, Barry D. Dunietz, and Yixiang Cao. Correlated ab initio electronic structure calculations for large molecules. *J. Phys. Chem. A.*, 103:1913–1928, 1999.

[31] Philippe Y. Ayala and Gustavo E. Scuseria. Linear scaling second-order møller plesset theory in the atomic orbital basis for large molecules. *J. Chem. Phys.*, 110:3660–3671, 1999.

[32] Martin Schütz, Georg Hetzer, and Hans-Joachim Werner. Low-order scaling local electron correlation methods. i. linear scaling local mp2. *J. Chem. Phys.*, 111:5691–5705, 1999.

[33] Svein Saebø and Peter Pulay. A low-scaling method for second order møller-plesset calculations. *J. Chem. Phys.*, 115:3975–3983, 2001.

[34] Claudia Hampel and Hans-Joachim Werner. Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.*, 104:6286–6297, 1996.

[35] Martin Schütz and Hans-Joachim Werner. Low-order scaling local electron correlation methods. iv. linear scaling local coupled-cluster (lccsd). *J. Chem. Phys.*, 114:661–681, 2001.

[36] Martin Schütz and Hans-Joachim Werner. Local perturbative triples correction (t) with linear cost scaling. *Chem. Phys. Lett.*, 318:370–378, 2000.

[37] Martin Schütz. Low-order scaling local electron correlation methods. iii. linear scaling local perturbative triples correction (t). *J. Chem. Phys.*, 113:9986–10001, 2000.

[38] Nicholas J. Russ and T .Daniel Crawford. Potential energy surface discontinuities in local correlation methods. *J. Chem. Phys.*, 121:691–696, 2004.

[39] Paul E. Maslen and Martin Head-Gordon. Noniterative local second order møller-plesset theory: Convergence with local correlation space. *J. Chem. Phys.*, 109:7093–7099, 1998.

[40] Michael S. Lee, Paul E. Maslen, and Martin Head-Gordon. Closely approximating second-order møller-plesset perturbation theory with a local triatomics in molecules model. *J. Chem. Phys.*, 112:3592–3601, 2000.

[41] Björn O. Roos. The complete-active-space self-consistent-field method and its applications in electronic structure calculations. *Adv. Chem. Phys.*, 69:399–345, 1987.

[42] Kerstin Andersson, Per-Åke Malmqvist, Björn O. Roos, Andrzej J. Sadlej, and Krzysztof Wolinksi. Second-order perturbation theory with a casscf reference function. *J. Phys. Chem.*, 94:5483–5488, 1990.

[43] Robert B. Murphy and Richard P. Messmer. Generalized møller-plesset perturbation theory applied to general mcscf reference wave functions. *Chem. Phys. Lett.*, 183:443–448, 1991.

[44] Robert B. Murphy and Richard P. Messmer. Generalized møller-plesset and epstein-nesbet perturbation theory applied to multiply-bonded molecules. *J. Chem. Phys.*, 97:4170–4184, 1992.

[45] Robert B. Murphy and Richard P. Messmer. Correlation in first-row transition metal atoms using the generalized møller-plesset perturbation theory. *J. Chem. Phys.*, 97:4974–4985, 1992.

[46] Anna I. Krylov, C. David Sherrill, Edward F. C. Byrd, and Martin Head-Gordon. Size-consistent wave functions for non-dynamical correlation energy: The valence active space optimized orbital coupled-cluster doubles model. *J. Chem. Phys.*, 109(24):10669–10678, 1998.

[47] C. David Sherrill, Anna I. Krylov, Edward F. C. Byrd, and Martin Head-Gordon. Energies and analytic gradients for a coupled-cluster doubles model using variational brueckner orbitals: application to symmetry breaking in o_4^+. *J. Chem. Phys.*, 109(11):4171–4181, 1998.

[48] Steven R. Gwaltney and Martin Head-Gordon. A second-order correction to singles and doubles coupled-cluster methods based on a perturbative expansion of a similarity-transformed hamiltonian. *Chem. Phys. Lett.*, 323:21–28, 2000.

[49] Steven R. Gwaltney and Martin Head-Gordon. A second order perturbative corrections to the coupled-cluster singles and doubles method:ccsd(2). *J. Chem. Phys.*, 115:2014–2021, 2001.

[50] Gregory J. O. Beran, Steven R. Gwaltney, and Martin Head-Gordon. Approaching closed-shell accuracy for radicals using coupled cluster theory with perturbative triple substitutions. *Phys. Chem. Chem. Phys.*, 5:2488–2493, 2003.

[51] David C. Graham, Gregory J. O. Beran, Martin Head-Gordon, Gemma Christian, Robert Stranger, and Brian F. Yates. Nitrogen activation via three-coordinate molybdenum complexes: comparison of density functional theory performance with wave function based methods. *J. Phys. Chem. A*, page in press, 2005.

[52] Gregory J. O. Beran, Steven R. Gwaltney, and Martin Head-Gordon. Can coupled cluster singles and doubles be approximated by a valence active space model? *J. Chem. Phys.*, 117:3040–3048, 2002.

[53] Gregory J. O. Beran and Martin Head-Gordon. The localizability of valence space electron-electron correlations in pair-based coupled cluster models. *Mol. Phys.*, page submitted, 2005.

[54] W. J. Hunt, P. J. Hay, and W. A. Goddard III. Self-consistent procedures for generalized valence bond wave functions. applications h_3, bh, h_2o, c_2h_6, and o_2. *J. Chem. Phys.*, 57:738–748, 1972.

[55] W. A. Goddard III and L. B. Harding. The description of chemical bonding from *ab initio* calculations. *Ann. Rev. Phys. Chem.*, 29:363–396, 1978.

[56] F. B. Bobrowicz and W. A. Goddard, *Methods of Electronic Structure Theory 3*, ed. H. F. Schaefer III, Plenum Press (1977).

[57] I. I. Ukrainskii. New variational function in theory of quasi-one-dimensional metals. *Theoret. Math. Phys.*, 32:816–822, 1977.

[58] John Cullen. Generalized valence bond solutions from a constrained coupled cluster method. *Chem. Phys.*, 202:217–229, 1996.

[59] Gregory J. O. Beran, Brian Austin, Alex Sodt, and Martin Head-Gordon. *J. Phys. Chem. A*, page submitted, 2005.

[60] Alex Sodt, Gregory J. O. Beran, and Martin Head-Gordon. *in preparation*, 2005.

[61] Gregory J. O. Beran and Martin Head-Gordon. Extracting dominant pair correlations from many-body wave functions. *J. Chem. Phys.*, 121:78–88, 2004.

[62] Gregory J. O. Beran, Alex Sodt, and Martin Head-Gordon. *in preparation*, 2005.

[63] Keld L. Bak, Jürgen Gauss, Poul Jørgensen, Jeppe Olsen, Trygve Helgaker, and John F. Stanton. The accurate determination of molecular equilibrium structures. *J. Chem. Phys.*, 114:6548–6556, 2001.

[64] Thom H. Dunning, Jr. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90:1007–1023, 1989.

[65] J. Russell Thomas, Bradley J. DeLeeuw, George Vacek, T. Daniel Crawford, Yukio Tamaguchi, and Henry F. Schaefer III. The balance between theoretical method and basis set quality: A systematic study of equilibrium geometries, dipole moments, harmonic vibrational frequencies, and infrared intensities. *J. Chem. Phys.*, 99:403–416, 1993.

[66] Edward F. C. Byrd, C. David Sherrill, and Martin Head-Gordon. The theoretical prediction of molecular radical species: a systematic study of equilibrium geometries and harmonic vibrational frequencies. *J. Phys. Chem. A*, 105:9736–9747, 2001.

[67] Gustavo E. Scuseria and Henry F. Schaefer III. The optimization of molecular orbitals for coupled cluster wave functions. *Chem. Phys. Lett.*, 142(5):354–358, 1987.

[68] R. A. Chiles and C. E. Dykstra. An electron pair operator approach to coupled cluster wave functions. application to he_2, be_2, and mg_2 and comparison with cepa methods. *J. Chem. Phys.*, 74:4544–4556, 1981.

[69] Nicholas C. Handy, John A. Pople, Martin Head-Gordon, Krishnan Raghavachari, and Gary W. Trucks. Size-consistent brueckner theory limited to double substitutions. *Chem. Phys. Lett.*, 164:185–198, 1989.

[70] Petr Bouř. Configuration interaction with kohn-sham orbitals and their relation to excited electronic states. *Chem. Phys. Lett.*, 345:331–337, 2001.

[71] R. J. Needs, P. R. C. Kent, A. R. Porter, M .D. Towler, and G. Rajagopal. Properties, dynamics, and electronic structure of condensed systems and clusters quantum monte carlo calculations for ground and excited states. *Int. J. Quant. Chem.*, 86:218–225, 2002.

[72] Gustavo E. Scuseria. On the connections between brueckner-coupled-cluster, density-dependent hartree-fock, and density functional theory. *Int. J. Quant. Chem.*, 55:165–171, 1995.

[73] Andreas Heßelmann and Georg Jansen. First-order intermolecular interaction energies from coupled-cluster brueckner orbitals. *J. Chem. Phys.*, 112:6949–6952, 2000.

[74] Ingvar Lindgren and Sten Salomonson. Brueckner orbitals and density-functional theory. *Int. J. Quant. Chem.*, 90:294–308, 2002.

[75] Troy Van Voorhis and Martin Head-Gordon. The quadratic coupled cluster doubles model. *Chem. Phys. Lett.*, 330:585–594, 2000.

[76] Troy Van Voorhis and Martin Head-Gordon. The imperfect pairing approximation. *Chem. Phys. Lett.*, 317:575–580, 2000.

[77] C. David Sherrill, Michael S. Lee, and Martin Head-Gordon. On the performance of density functional theory for symmetry-breaking problems. *Chem. Phys. Lett.*, 302:425–430, 1999.

[78] Barry D. Dunietz and Martin Head-Gordon. Manifestations of symmetry breaking in self-consistent field electronic structure calculations. *J. Phys. Chem. A*, 107:9160–9167, 2003.

[79] John F. Stanton, Jürgen Gauss, John D. Watts, and Rodney J. Bartlett. A direct product decomposition approach for symmetry exploitation in many-body methods. I. energy calculations. *J. Chem. Phys.*, 94(6):4334–4345, 1991.

[80] G. Herzberg. *Molecular Spectra and Molecular Structure I. Spectra of Diatomic Molecules.* Van Nostrand, New York, 1950.

[81] K. P. Huber and G. Herzberg. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules.* Van Nostrand Reinhold, New York, 1979.

[82] R. P. Tuckett, A. R. Dale, D. M. Jaffey, P. S. Jarrett, and T. Kelly. The $a^2\pi_u/x^2\pi_g$ electronic emission-spectrum of the fluorine molecular ion f_2^+ studied in a supersonic beam. *Mol. Phys.*, 49:475–486, 1983.

[83] J. B. Burkholder, P. D. Hammer, C. J. Howard, and A. R. W. McKellar. High-resolution fourier transform spectroscopy of the 1-0 and 2-0 infrared bands of the fo radical ($^2\pi_{3/2}$, $^2\pi_{1/2}$). *J. Mol. Spectrosc.*, 118:471–480, 1986.

[84] A. J. Cormack, A. J. Yencha, R. J. Donovan, K. P. Lawley, A. Hopkirk, and G. C. King. High-resolution threshold photo-electron spectroscopy of molecular fluorine. *Chem. Phys.*, 213:439–448, 1996.

[85] Jing Kong, Christopher A. White, Anna I. Krylov, David Sherrill, Ross D. Adamson, Thomas R. Furlani, Michael S. Lee, Aaron M. Lee, Steven R. Gwaltney, Terry R. Adams, Christian Ochsenfeld, Andrew T. B. Gilbert, Gary S. Kedziora, Vitaly A. Rassolov, David R. Maurice, Nikhil Nair, Yihan Shao, Nicholas A. Besley, Paul E. Maslen, Jeremy P. Dombroski, Holger Daschel, Weiman Zhang, Prakashan P. Korambath, Jon Baker, Edward F. C. Byrd, Troy Van Voorhis, Manabu Oumi, So Hirata, Chao-Ping Hsu, Naoto Ishikawa, Jan Florian, Areih Warshel, Benny G. Johnson, Peter M. W. Gill, Martin Head-Gordon, and John A. Pople. Q-Chem 2.0: a high-performance *ab initio* electronic structure program package. *J. Comp. Chem.*, 21(16):1532–1548, 2000.

[86] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, M. L. Leininger, R. A. King, S. T. Brown, C. L. Janssen, E. T. Seidl, Y. Yamaguchi, W. D. Allen, Y. Xie, G. Vacek, T. P. Hamilton, C. B. Kellogg, R. B. Remington, and H. F. Schaefer III. PSI 3.0, development version, PSITECH, Inc., Watkinsville, GA 30677, U.S.A., 2000.

[87] Jan M. L. Martin. On the performance of correlation consistent basis sets for the calculation of total atomization energies, geometries, and harmonic frequencies. *J. Chem. Phys.*, 100:8186–8193, 1994.

[88] Filip Pawłowski, Asger Halkier, Poul Jørgensen, Keld L. Bak, Trygve Helgaker, and Wim Klopper. Accuracy of spectroscopic constants of diatomic molecules. *J. Chem. Phys.*, 118:2539–2549, 2003.

[89] Frank Jensen. A remarkable large effect of spin contamination on calculated vibrational frequencies. *Chem. Phys. Lett.*, 169:519–528, 1990.

[90] Peter M. W. Gill and Leo Radom. Deceptive convergence in møller-plesset perturbation energies. *Chem. Phys. Lett.*, 132:16–22, 1986.

[91] Ross H. Nobes, John A. Pople, Leo Radom, Nicholas C. Handy, and Peter J. Knowles. Slow convergence of the møller-plesset perturbation series: the dissociation energy of hydrogen cyanide and the electron affinity of the cyano radical. *Chem. Phys. Lett.*, 138:481–485, 1987.

[92] Peter M. W. Gill, John A. Pople, Leo Radom, and Ross H. Nobes. Why does unrestricted møller-plesset perturbation theory converge so slowly for spin-contaminated wave functions? *J. Chem. Phys.*, 89:7307–7314, 1988.

[93] Anna I. Krylov. Spin contamination of coupled cluster wave functions. *J. Chem. Phys.*, 113:6052–6062, 2000.

[94] Jon Baker, Andrew Scheiner, and Jan Andzelm. Spin contamination in density functional theory. *Chem. Phys. Lett.*, 216:380–388, 1993.

[95] T. Daniel Crawford, John F. Stanton, Wesley D. Allen, and Henry F. Schaefer III. Hartree-fock orbital instability envelopes in highly correlated single-reference wave functions. *J. Chem. Phys.*, 107:10626–10632, 1997.

[96] T. Daniel Crawford and John F. Stanton. Some surprising failures of brueckner coupled cluster theory. *J. Chem. Phys.*, 112:7873–7879, 2000.

[97] Ryan D. Cohen and C. David Sherrill. The performance of density functional theory for equilibrium properties of symmetry-breaking molecules. *J. Chem. Phys.*, 114:8257–8269, 2001.

[98] Galina Orlova and John D. Goddard. Is density functional theory free of spatial symmetry breaking? the case of the linear carbon radical cations: C_3^+, c_5^+, c_7^+, and c_9^+. *Chem. Phys. Lett.*, 363:486–491, 2002.

[99] John F. Stanton, Jürgen Gauss, and Rodney J. Bartlett. On the choice of orbitals for symmetry breaking problems with application to no$_3$. *J. Chem. Phys.*, 97:5554–5559, 1992.

[100] Leslie A. Barnes and Roland Lindh. Symmetry breaking in o$_4^+$: An application of the brueckner coupled-cluster method. *Chem. Phys. Lett.*, 223:207–214, 1994.

[101] Martin Head-Gordon. Quantum chemistry and molecular processes. *J. Chem. Phys.*, 100:13213–13225, 1996.

[102] Peter Pulay, Svein Saebø, and K. Wolinski. Efficient calculation of canonical mp2 energies. *Chem. Phys. Lett.*, 344:543–552, 2001.

[103] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen. *Molecular Electronic-Structure Theory*. John Wiley and Sons, Chichester, England, 2000.

[104] A. C. Hurley, J. Lennard-Jones, and J. A. Pople. The molecular orbital theory of chemical valency. xvi. a theory of paired-electrons in polyatomic molecules. *Proc. Roy. Soc. A*, 220:446–455, 1953.

[105] Troy Van Voorhis and Martin Head-Gordon. A nonorthogonal approach to perfect pairing. *J. Chem. Phys.*, 112:5633–5638, 2000.

[106] Troy Van Voorhis and Martin Head-Gordon. Connections between coupled cluster and generalized valence bond theories. *J. Chem. Phys.*, 115:7814–7821, 2001.

[107] Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 10/12/01, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institue for the U.S. Department of Energy under

contract DE-AC06-76RLO 1830. Contact David Feller or Karen Schuchardt for further information.

[108] David. E. Woon and Thom H. Dunning, Jr. Gaussian basis sets for use in correlated molecular calculations. v. core-valence basis sets for boron through neon. *J. Chem. Phys.*, 103:4572–4585, 1995.

[109] Thom H. Dunning, Jr., K. A. Peterson, and A. K. Wilson. Gaussian basis sets for use in correlated molecular calculations. x. the atoms aluminum through argon revisited. *J. Chem. Phys.*, 114:9244–9253, 2001.

[110] R. Krishnan, J.S. Binkley, R. Seeger, and J.A. Pople. Self-consistent molecular orbital methods. xx. a basis set for correlated wave functions. *J. Chem. Phys.*, 72:650–654, 1980.

[111] A.D. McLean and G.S. Chandler. Contracted gaussian basis sets for molecular calculations. i. second row atoms, z=11-18. *J. Chem. Phys.*, 72:5639–5648, 1980.

[112] Keld L. Bak, Poul Jørgensen, Jeppe Olsen, Trygve Helgaker, and Wim Klopper. Accuracy of atomization energies and reaction enthalpies in standard and extrapolated electronic wave function/basis set calculations. *J. Chem. Phys.*, 112:9229–9242, 2000.

[113] M. J. S. Dewar and C. Jie. Mechanism of the cope rearrangement. *J. Am. Chem. Soc.*, 109:5893–5900, 1987.

[114] Michel Dupuis, Christopher Murray, and Ernest R. Davidson. The cope rearrangement revisited. *J. Am. Chem. Soc.*, 113:9756–9759, 1991.

[115] David. A. Hrovat, Keiji Morokuma, and Weston Thatcher Borden. The cope rearrangement revisited again. results of *ab initio* calculations beyond the casscf level. *J. Am. Chem. Soc.*, 116:1072–1076, 1994.

[116] P. M. Kozlowski, Michel Dupuis, and Ernest R. Davidson. The cope rearrangement revisited with multi-reference perturbation theory. *J. Am. Chem. Soc.*, 117:774–778, 1995.

[117] Weston Thatcher Borden and Ernest R. Davidson. The importance of including dynamic electron correlation in *ab initio* calculations. *Acc. Chem. Res.*, 29:67–75, 1996.

[118] Viktor N. Staroverov and Ernest R. Davidson. The cope rearrangement in theoretical retrospect. *J. Mol. Struct. (Theochem)*, 573:81–89, 2001.

[119] Ernest R. Davidson. The spatial extent of the v state of ethylene and its relation to dynamic correlation in the cope rearrangement. *J. Phys. Chem.*, 100:6161–6166, 1996.

[120] Edward F.C. Byrd, Troy Van Voorhis, and Martin Head-Gordon. Quadratic coupled cluster doubles: Implementation and assessment of perfect-pairing optimized geometries. *J. Phys. Chem. B*, 106:8070–8077, 2002.

[121] Youliang Wang and Raymond A. Poirier. Generalized valence bond molecular equilibrium geometries. *J. Mol. Struct.*, 340:1–14, 1995.

[122] Robert C. Ladner and William A. Goddard III. Improved quantum theory of many-electron systems. v. the spin-coupling optimized gi method. *J. Chem. Phys.*, 51:1073–1087, 1969.

[123] Lawrence B. Harding and William A. Goddard III. The generalized valence bond description of the low-lying states of formaldehyde. *J. Am. Chem. Soc.*, 97:6293–6299, 1975.

[124] Emily A. Carter and William A. Goddard III. Electron correlation, basis sets, and the methylene singlet-triplet gap. *J. Chem. Phys.*, 86:862–865, 1987.

[125] Emily A. Carter and William A. Goddard III. Correlation-consistent configuration interaction: Accurate bond dissociation energies from simple wave functions. *J. Chem. Phys.*, 88:3132–3140, 1988.

[126] Robert B. Murphy and Richard P. Messmer. A generalized valence bond representation of complete-active-space self-consistent-field (casscf) wave functions. *J. Chem. Phys.*, 98:7957–7968, 1993.

[127] Francesco Faglioni and William A. Goddard III. Gvb-rp: A reliable mcscf wave function for large systems. *Int. J. Quant. Chem.*, 73:1–22, 1999.

[128] John Cullen. Is gvb-ci superior to casscf? *J. Comp. Chem.*, 20:999–1008, 1999.

[129] Robert B. Murphy, W. Thomas Pollard, and Richard A. Friesner. Psuedospectral localized generalized møller-plesset methods with a generalized valence bond reference wave function: Theory and calculation of conformational energies. *J. Chem. Phys.*, 106:5073–5084, 1997.

[130] Mihir Sejpal and Richard P. Messmer. Calculations using generalized valence bond based møller-plesset perturbation theory. *J. Chem. Phys.*, 114:4796–4804, 2001.

[131] Mihir Sejpal and Richard P. Messmer. On the size extensivity of second-order møller-plesset perturbation theory based on a perfectly-paired generalized valence bond reference wave function. *Chem. Phys.*, 270:237–243, 2001.

[132] Troy Van Voorhis and Martin Head-Gordon. Implementation of generalized valence bond-inspired coupled cluster theories. *J. Chem. Phys.*, 117:9190–9201, 2002.

[133] Gustavo E. Scuseria, Curtis L. Janssen, and Henry F. Schaefer III. An efficient reformulation of the closed-shell coupled cluster single and double excitation (ccsd) equations. *J. Chem. Phys.*, 89(12):7382–7387, 1988.

[134] Andreas F. Voegele, Christofer S. Tautermann, Thomas Loerting, and Klaus R. Liedl. Reactions of hocl + hcl + nh$_2$o and hocl + hbr and nh$_2$o. *J. Phys. Chem. A*, 106:7850–7857, 2002.

[135] Vitaly A. Rassolov. A geminal model chemistry. *J. Chem. Phys.*, 117:5978–5987, 2002.

[136] James Shoemaker, Larry W. Burggraf, and Mark S. Gordon. An *ab initio* cluster study of the structure of the si(001) surface. *J. Chem. Phys.*, 112:2994–3005, 2003.

[137] Yousung Jung and Martin Head-Gordon. How diradicaloid is a stable diradical? *Chem. Phys. Chem.*, 4:522–525, 2003.

[138] Yousung Jung and Martin Head-Gordon. What is the nature of the long bond in the [tcne]$_2^{2-}$ π dimer? *Chem. Phys. Phys. Chem.*, 6:2008–2010, 2004.

[139] David L. Cooper, Joseph Gerratt, and Mario Raimondi. Applications of spin-coupled valence bond theory. *Chem. Rev.*, 91:929–964, 1991.

[140] Mario Raimondi and David L. Cooper. Ab initio moden valence bond theory. *Topics in Current Physics*, 203:105–120, 1999.

[141] Martin W. Feyereisen, George Fitzgerald, and Andrew Komornicki. Use of approximate integrals in ab initio theory. an application in mp2 energy calculations. *Chem. Phys. Lett.*, 208:359–363, 1993.

[142] O. Vahtras, J. Almlöf, and M. W. Feyereisen. Integral approximations for lcao-scf calculations. *Chem. Phys. Lett.*, 213:514–518, 1993.

[143] Karin Eichkorn, Oliver Treutler, Holger Öhm, Marco Häser, and Reinhart Ahlrichs. Auxiliary basis sets to approximate coulomb potentials. *Chem. Phys. Lett.*, 240:283–289, 1995.

[144] Karin Eichkorn, Florian Weigend, Oliver Treutler, and Reinhart Ahlrichs. Auxiliary basis sets for main row atoms and transition metals and their use to approximate coulomb potentials. *Theor. Chem. Acc*, 97:119–124, 1997.

[145] T. Sano. Elementary jacobi rotation method for generalize valence bond perfect-pairing calculations combined with a simple procedure for generating reliable initial orbitals. *J. Mol. Struct. (Theochem)*, 528:177–191, 2000.

[146] Troy Van Voorhis and Martin Head-Gordon. A geometric approach to direct minimization. *Molecular Physics*, 100:1713–1721, 2002.

[147] J. A. Pople, K. Raghavachari, H. B. Schlegel, and J. S. Binkley. Derivative studies in hartree-fock and møller-plesset theories. *Int. J. Quant. Chem. Symp.*, 13:225–241, 1979.

[148] J. Gauss. *Modern methods and algorithms of quantum chemistry, NIC series*, volume 3, pages 541–592. John von Neumann Institute for Computing, Jülich, 2000.

[149] R. P. Muller, J. M. Langlois, M. N. Ringnalda, R.A. Friesner, and W.A. Goddard III. A generalized direct inversion in the iterative subspace approach for generalized valence bond wave functions. *J. Chem. Phys.*, 100:1226–1235, 1994.

[150] Irina V. Ionova and Emily A. Carter. Orbital-based direct inversion in the iterative subspace for the generalized valence bond method. *J. Chem. Phys.*, 102:1251–1256, 1995.

[151] Florian Weigend, Andreas Köhn, and Christof Hättig. Efficient use of correlation consistent basis sets in the resolution of the identity mp2 calculations. *J. Chem. Phys.*, 116:3175–3183, 2002.

[152] Helena Larsen, Jeppe Olsen, Poul Jørgensen, and Ove Christiansen. Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of n_2. *J. Chem. Phys.*, 113:6677–6686, 2000. cc-pVDZ basis.

[153] Trygve Helgaker, Jürgen Gauss, Poul Jørgensen, and Jeppe Olsen. The prediction of molecular equilibrium structures by the standard electronic wave functions. *J. Chem. Phys.*, 106:6430–6440, 1997.

[154] Warren J. Hehre, Leo Radom, Paul v. R. Schleyer, and John A. Pople. *Ab initio Molecular Orbital Theory*. Wiley, New York, 1986.

[155] Leslie Farnell, John A. Pople, and Leo Radom. Structural predictions for open-shell systems: a comparative assessment of ab initio procedures. *J. Phys. Chem.*, 87:79–82, 1983.

[156] Michael S. Lee and Martin Head-Gordon. Extracting polarized atomic orbitals from molecular orbital calculations. *Int. J. Quant. Chem.*, 76:169–184, 2000.

[157] Per-Olov Löwdin. Quantum theory of many-particle systems. i. physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configuration interaction. *Phys. Rev.*, 97:1474–1489, 1955.

[158] K. Takatsuka, T. Fueno, and K. Yamaguchi. Distribution of odd electrons in ground-state molecules. *Theor. Chim. Acta*, 48:175–183, 1978.

[159] Viktor N. Staroverov and Ernest R. Davidson. Diradical character of the cope rearrangement transition state. *J. Am. Chem. Soc.*, 122:186–187, 2000.

[160] Viktor N. Staroverov and Ernest R. Davidson. Distribution of effectively unpaired electrons. *Chem. Phys. Lett.*, 330:161–168, 2000.

[161] Martin Head-Gordon. Characterizing unpaired electrons from the one-particle density matrix. *Chem. Phys. Lett.*, 372:508–511, 2003.

[162] Laimutis Bytautis, Joseph Ivanic, and Klaus Ruedenberg. Split-localized orbitals can yield stronger configuration interaction convergence than natural orbitals. *J. Chem. Phys.*, 119:8217–8224, 2003.

[163] Ida M. B. Nielsen and Curtis L. Janssen. New diagnostics for coupled-cluster and møller-plesset perturbation theory. *Chem. Phys. Lett.*, 290:423–430, 1998.

[164] Ida M. B. Nielsen and Curtis L. Janssen. Double-substitution-based diagnostics for coupled-cluster and møller-plesset perturbation theory. *Chem. Phys. Lett.*, 310:568–576, 1999.

[165] Gene P. Barnett and Oscar R. Platas. Reduced-density-matrix theory: The energy properties of the be 2-matrix. *J. Chem. Phys.*, 48:4265–4271, 1968.

[166] Pedro L. Olympia, Jr. and Darwin W. Smith. Density matrix study of atomic ground and excited states. i. beryllium ground state. *J. Chem. Phys.*, 52:67–87, 1970.

[167] Pedro L. Olympia, Jr. and Darwin W. Smith. Density matrix study of atomic ground and excited states. ii. beryllium ^{1}s excited state. *J. Chem. Phys.*, 57:4018–4028, 1972.

[168] Pedro L. Olympia, Jr., Robert C. Morrison, and Darwin W. Smith. Density matrix study of atomic ground and excited states. iii. geminal energy analysis of an accurate beryllium ground-state wave function. *Phys. Rev. A*, 6:1767–1771, 1972.

[169] Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*. SIAM, Philadelphia, 1997.

[170] Tomoko Kinoshita, Osamu Hino, and Rodney J. Bartlett. Singular value decomposition approach for the approximate coupled cluster method. *J. Chem. Phys.*, 119:7756–7762, 2003.

[171] Martin Head-Gordon, A. M. Graña, David Maurice, and Christopher A. White. Analysis of electronic transitions as the difference of electron attachment and detachment densities. *J. Phys. Chem.*, 99:14261–14270, 1995.

[172] Ernest R. Davidson. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comp. Phys.*, 17:87–94, 1975.

[173] Matthew L. Leininger, C. David Sherrill, Wesley D. Allen, and Henry F. Schaefer III. Systematic study of selected diagonalization methods for configuration interaction matrices. *J. Comp. Chem.*, 22:1574–1589, 2001.

[174] Yousung Jung and Martin Head-Gordon. Controlling the extent of diradical character by utilizing neighboring group interactions. *J. Phys. Chem. A*, 107:7475–7481, 2003.

[175] Emily A. Carter and William A. Goddard III. Correlation-consistent singlet-triplet gaps in substituted carbenes. *J. Chem. Phys.*, 88:1752–1763, 1988.

[176] Vitaly A. Rassolov, Feng Xu, and Sophya Garashchuk. Geminal model chemistry ii. perturbative corrections. *J. Chem. Phys.*, 120:10385–10394, 2004.

[177] Larry A Curtiss, Krishnan Raghavachari, Gary W. Trucks, and John A. Pople. Gaussian-2 theory for molecular energies of first- and second-row compounds. *J. Chem. Phys.*, 94:7221–7230, 1991.

[178] Ansgar Schäfer, Hans Horn, and Reinhart Ahlrichs. Fully optimized contracted gaussian basis sets for atoms li to kr. *J. Chem. Phys.*, 97:2571–2577, 1992.

[179] Juan J. Novoa, Pilar Lafuente, Rico E. Del Sesto, and Joel S. Miller. Exceptionally long (≥ 2.9å) c-c bonds between [tcne]$^-$ ions: Two-electron, four-center π^*-π^* c-c bonding in π-[tcne]$_2^{2-}$. *Angew. Chem., Int. Ed. Engl.*, 40:2540–2545, 2001.

[180] Jian-Ming Lu, Sergiy V. Rosokha, and Jay K. Kochi. Stable (long-bonded) dimers via the quantitative self-association of different cationic, anionic, and uncharged π-radicals: Structures, energetics, and optical transitions. *J. Am. Chem. Soc.*, 125:12161–12171, 2003.

[181] Liu, B. The simultaneous expansion method for the iterative solution of several of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices, Technical Report LBL-8158, Lawrence Berkeley Laboratory, University of California, Berkeley, 1978.