Unrestricted Perfect Pairing: The Simplest Wave-Function-Based Model Chemistry beyond Mean Field

Gregory J. O. Beran, Brian Austin, Alex Sodt, and Martin Head-Gordon*

Department of Chemistry, University of California, and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720-1460

Received: July 9, 2005; In Final Form: August 4, 2005

The perfect pairing (PP) approximation from generalized valence bond theory is formulated in an unrestricted fashion for both closed- and open-shell systems using a coupled cluster ansatz. In the model chemistry proposed here, active electron pairs are correlated, but the unpaired or radical electrons remain uncorrelated, leading to a linear number of decoupled cluster amplitudes which can be solved for analytically. The alpha and beta spatial orbitals are variationally optimized independently. This minimal treatment of electron−electron correlation noticeably improves upon symmetry-breaking problems and other pathologies in Hartree−Fock (HF) theory and may be computed using the resolution of the identity approximation at only a factor of several times more effort than HF itself. PP also generally predicts improved molecular structures over HF. This compact, correlated wave function potentially provides a useful starting point for dynamical correlation corrections.

I. 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 multiconfigurational self-consistent field (MCSCF) or complete active space self-consistent field (CASSCF)[1] 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 multireference 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 14 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 depend 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.[2]

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.[3,4] In VOD, a coupled cluster doubles (CCD)[5] 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 a 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 computational scaling is high, and it limits practical VOD computations to about 15 non-hydrogen atoms (fewer in large basis sets).

There are some other approaches for replacing the HF reference function that should be mentioned. One promising approach that has been recently explored by Rassolov[6] 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.[7−10]

GVB-PP is a simple 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 correlated with a single spatial virtual correlating orbital. Though the model is 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. PP has also been applied to the study of certain classes of diradicals with much success.[11−13] Importantly, it is potentially very inexpensive to calculate. For example, the pseudospectral approximation has been applied

10.1021/jp053780c CCC: $30.25
Published on Web 09/17/2005

to speed the integral formation.[14] Suggestions have also been made for generalizations of the PP approach that include interpair couplings[15] or employ nonorthogonal orbitals.[16] Additionally, there has been considerable effort devoted to the development of more sophisticated GVB wave functions that lift the restriction of orthogonality.[17,18]

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[19−22] that introduces errors on the order of tens of microhartrees per atom 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.

II. Theory and Implementation

A. 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,

$$|\Psi\rangle = |A[\phi_1\phi_{\bar{1}}\phi_2\phi_{\bar{2}}...g_1g_2...]\rangle \tag{1}$$

where the pair functions, g_i, are defined as

$$g_i = A[\psi_i\psi_{\bar{i}} + t_i\psi_{i^*}\psi_{\bar{i}^*}] \tag{2}$$

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

The same wave function can be rewritten as a simplified coupled cluster (CC) wave function.[23,24] In this case, it takes the form

$$|\Psi_{\mathrm{PP}}\rangle = \mathrm{e}^{\hat{T}_{\mathrm{PP}}}|\Phi_0\rangle \tag{3}$$

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

$$\hat{T}_{\mathrm{PP}} = \sum_i^{\mathrm{pairs}} t_{i\bar{i}}^{i^*\bar{i}^*}\hat{a}_{i^*}^\dagger\hat{a}_i\hat{a}_{\bar{i}^*}^\dagger\hat{a}_{\bar{i}} = \sum_i^{\mathrm{pairs}} t_i\hat{a}_{i^*}^\dagger\hat{a}_i\hat{a}_{\bar{i}^*}^\dagger\hat{a}_{\bar{i}} \tag{4}$$

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 this formalism, the t amplitudes are solved using projective techniques as in most CC theories, rather than variationally. 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, analogous 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 an unrestricted HF (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 separately 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 will localize to their respective atomic centers and have zero correlation amplitude between them. Any attempt to correlate these unpaired electrons would therefore be inconsistent with the closed-shell PP model.

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

$$L_{\mathrm{PP}} = \langle\Phi_0|(1 + \hat{\Lambda}_{\mathrm{PP}})\mathrm{e}^{-\hat{T}_{\mathrm{PP}}}\hat{H}\mathrm{e}^{\hat{T}_{\mathrm{PP}}}|\Phi_0\rangle \tag{5}$$

where $\hat{\Lambda}_{\mathrm{PP}} = \sum_i^{\mathrm{pairs}} \lambda_i\hat{a}_i^\dagger\hat{a}_{i^*}\hat{a}_{\bar{i}}^\dagger\hat{a}_{\bar{i}^*}$. 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,[25] so we will not repeat all of 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 25 for further detail.

B. Initial Guess. We start the unrestricted PP (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

$$D^{\alpha\beta} = (C^\alpha)^\dagger S C^\beta = U\Sigma V^\dagger \tag{6}$$

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

$$\tilde{C}^\alpha = C^\alpha U \qquad \tilde{C}^\beta = C^\beta V \tag{7}$$

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.[26] The Sano algorithm is used separately for the alpha and beta components of each occupied pair and may draw in parts of the HF SUMO in forming the pairs. As described above, alpha

TABLE 1: Unrestricted Perfect Pairing Lagrangian and Its Orbital Derivatives

$L_{\mathrm{PP}} = E_{\mathrm{ref}} + \sum_i^{\mathrm{pairs}} \lambda_i t_i (f_{i^*i^*} + f_{\bar{i}^*\bar{i}^*} - f_{ii} - f_{\bar{i}\bar{i}}) + \sum_i^{\mathrm{pairs}} (t_i + \lambda_i - \lambda_i t_i^2)\langle i\bar{i}|i^*\bar{i}^*\rangle + \sum_i^{\mathrm{pairs}} \lambda_i t_i (\langle i\bar{i}|i\bar{i}\rangle + \langle i^*\bar{i}^*|i^*\bar{i}^*\rangle + \langle ii^*|i^*i\rangle - \langle ii^*|ii^*\rangle + \langle \bar{i}\bar{i}^*|\bar{i}^*\bar{i}\rangle - \langle \bar{i}\bar{i}^*|\bar{i}\bar{i}^*\rangle - \langle i\bar{i}^*|\bar{i}i^*\rangle - \langle \bar{i}i^*|i\bar{i}^*\rangle)$

orbital gradient $(\mathrm{d}L/\mathrm{d}\Delta_q^p) = \Sigma_\mu [(\partial L/\partial C_\mu^p)C_\mu^q - (\partial L/\partial C_\mu^q)C_\mu^p]$

$l \in$ inact. occ. $(\partial L/\partial C_\mu^l) = 2f_{\mu l} - \sum_i^{\mathrm{pairs}} 2\lambda_i t_i (\langle i\mu|il\rangle - \langle i\mu|li\rangle + \langle \bar{i}\mu|\bar{i}l\rangle - \langle i^*\mu|i^*l\rangle + \langle i^*\mu|li^*\rangle - \langle \bar{i}^*\mu|\bar{i}^*l\rangle)$

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

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

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

SOMOs are not paired in any fashion and comprise a third alpha occupied subspace, leading to core, active, and inactive SOMO subspaces.

C. 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

$$E_{\mathrm{pp}} = E_{\mathrm{ref}} + \sum_i^{\mathrm{pairs}} t_i \langle i\bar{i}|i^*\bar{i}^*\rangle \tag{8}$$

In effect, this means we treat unpaired, excess alpha electrons in a UHF fashion while correlating all of the valence electron pairs. The advantage of the PP model not shared by any of the more complicated coupled cluster methods described above[3,5,15,16,27,28] 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):

$$\langle i\bar{i}|i^*\bar{i}^*\rangle + W_i t_i - \langle i\bar{i}|i^*\bar{i}^*\rangle t_i^2 = 0 \tag{9}$$

where

$$W_i = f_{i^*i^*} + f_{\bar{i}^*\bar{i}^*} - f_{ii} - f_{\bar{i}\bar{i}} + \langle i\bar{i}|i\bar{i}\rangle + \langle i^*\bar{i}^*|i^*\bar{i}^*\rangle + \langle ii^*|i^*i\rangle - \langle ii^*|ii^*\rangle + \langle \bar{i}\bar{i}^*|\imath^*\bar{\imath}\ - \langle \bar{i}\bar{i}^*|\bar{i}\bar{i}^*\rangle - \langle i\bar{i}^*|i\bar{i}^*\rangle - \langle \bar{i}i^*|\bar{i}i^*\rangle \tag{10}$$

In this notation, the f_{pp} are diagonal elements of the fock matrix $f_{pq} = h_{pq} + \sum_j^{\mathrm{occ}} \langle pj|qj\rangle - \langle pj|jq\rangle$ and $\langle pq|rs\rangle = \int \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2\, \phi_p(\mathbf{r}_1)\, \phi_q(\mathbf{r}_2)(1/r_{12})\, \phi_r(\mathbf{r}_1)\, \phi_s(\mathbf{r}_2)$. The λ amplitudes are also trivially obtainable as

$$\lambda_i = -\frac{\langle i\bar{i}|i^*\bar{i}^*\rangle}{W_i - 2t_i\langle i\bar{i}|i^*\bar{i}^*\rangle} \tag{11}$$

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, as presented in Table 1. Differentiating this Lagrangian with respect to a rotation, Δ_q^p, between orbitals p and q gives

$$\frac{\mathrm{d}L}{\mathrm{d}\Delta_q^p} = \sum_\mu \left(\frac{\partial L}{\partial C_\mu^p} C_\mu^q - \frac{\partial L}{\partial C_\mu^q} C_\mu^p \right) \tag{12}$$

where C_μ^p is the molecular orbital coefficient for atomic orbital μ in molecular orbital p. 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 1. Convergence of the orbital optimization procedure is enhanced by utilizing diagonal second derivatives in the geometric direct minimization (GDM) procedure.[29] 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 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.

D. Resolution of the Identity Approximation. The primary algorithmic difference between the closed-shell PP code described previously[25] 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

$$|\rho) = |\mu\nu) \approx |\overline{\mu\nu}) = \sum_K^{\mathrm{aux}} A_{\mu\nu}^K |K) \tag{13}$$

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

$$A_{\mu\nu}^K = \sum_L^{\mathrm{aux}} (\mu\nu|L)(L|K)^{-1} \tag{14}$$

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

$$(\mu\nu|\lambda\sigma) \approx$$

$$(\overline{\mu\nu}|\overline{\lambda\sigma}) = \sum_{KLMN}^{\mathrm{aux}} (\mu\nu|K)(K|L)^{-1}(L|M)(M|N)^{-1}(N|\lambda\sigma) \tag{15}$$

$$= \sum_{KL}^{\mathrm{aux}} (\mu\nu|K)(K|L)^{-1}(L|\lambda\sigma) \tag{16}$$

$$= \sum_{KLM}^{\mathrm{aux}} [(\mu\nu|K)(K|M)^{-1/2}][(M|L)^{-1/2}(L|\lambda\sigma)] \tag{17}$$

$$= \sum_M^{\mathrm{aux}} B_{\mu\nu}^M B_{\lambda\sigma}^M \tag{18}$$

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

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 multiplications of the $B^K_{\mu\nu}$ matrices. It also facilitates the development of cubic disk storage algorithms, since all two-electron integrals can be formed as needed from the $B^K_{\mu\nu}$ matrices.

The details of our algorithm will be presented elsewhere, but the basic procedure is as follows. Following the approach outlined in ref 25, we wish to construct half-transformed Coulomb- and exchangelike matrices

$$J^{[ii]}_{\mu\nu} = \langle i\mu|i\nu\rangle = (ii|\mu\nu) \qquad K^{[ii]}_{\mu\nu} = \langle i\mu|\nu i\rangle = (i\mu|i\nu) \tag{19}$$

$$J^{[ii^*]}_{\mu\nu} = \langle i\mu|i^*\nu\rangle = (ii^*|\mu\nu) \quad K^{[ii^*]}_{\mu\nu} = \langle i\mu|\nu i^*\rangle = (i\mu|i^*\nu) \tag{20}$$

$$J^{[i^*i^*]}_{\mu\nu} = \langle i^*\mu|i^*\nu\rangle = (i^*i^*|\mu\nu) \qquad K^{[i^*i^*]}_{\mu\nu} = \langle i^*\mu|\nu i^*\rangle = (i^*\mu|i^*\nu) \tag{21}$$

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

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

(2) We transform the first AO index μ to the MO basis for orbitals i and i^*, for example, $B^L_{i\nu} = \sum_\mu C_{\mu i} B^L_{\mu\nu}$.

(3) We form $K^{[ii]}_{\mu\nu} = \sum_L B^L_{i\mu} B^L_{i\nu}$, along with $K^{[ii^*]}_{\mu\nu}$ and $K^{[i^*i^*]}_{\mu\nu}$.

(4) We transform the second AO index to the MO basis, for example, $B^L_{ii^*} = \sum_\mu C_{\nu i^*} B^L_{i\nu}$.

(5) We form $J^{[ii]}_{\mu\nu} = \sum_L B^L_{ii} B^L_{\mu\nu}$, along with $J^{[ii^*]}_{\mu\nu}$ and $J^{[i^*i^*]}_{\mu\nu}$.

(6) We transform the final to AO indices to form $J^{[ii]}_{pp}$, $K^{[ii]}_{pp}$, and so forth, where p is any MO. Alternatively, we form the necessary three-quarter-transformed integrals used in the orbital derivatives listed in Table 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 of 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, $K^{[ii]}_{\mu\nu}$, which is also performed using the RI approximation as described above.

Overall, this algorithm is quartic, albeit with a small prefactor, requires a cubic amount of disk storage, and has quadratic memory requirements. Calculations on 50 carbon linear alkanes in the cc-pVDZ basis (more than 1200 basis functions in the primary basis)[30] 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.[31]

E. 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 Hellmann−Feynman theorem permits an equation for its gradient to be obtained trivially. If the superscript "(x)" indicates partial differentiation with respect to nuclear displacement, then the Lagrangian becomes

$$L^{(x)}_{\mathrm{PP}} = E^{(x)}_{\mathrm{ref}} + \sum_i^{\mathrm{pairs}} \lambda_i t_i (f^{(x)}_{i^*i^*} + f^{(x)}_{\bar{\imath}^*\bar{\imath}^*} - f^{(x)}_{ii} - f^{(x)}_{\bar{\imath}\bar{\imath}}) + \sum_i^{\mathrm{pairs}} (t_i + \lambda_i - \lambda_i t_i^2)\langle i\bar{\imath}|i^*\bar{\imath}^*\rangle^{(x)} + \sum_i^{\mathrm{pairs}} \lambda_i t_i (\langle i\bar{\imath}|i\bar{\imath}\rangle^{(x)} + \langle i^*\bar{\imath}^*|i^*\bar{\imath}^*\rangle^{(x)} + \langle ii^*|i^*i\rangle^{(x)} - \langle ii^*|ii^*\rangle^{(x)} + \langle \bar{\imath}\bar{\imath}^*|\bar{\imath}^*\bar{\imath}\rangle^{(x)} - \langle \bar{\imath}\bar{\imath}^*|\bar{\imath}\bar{\imath}^*\rangle^{(x)} - \langle i\bar{\imath}^*|i\bar{\imath}^*\rangle^{(x)} - \langle \bar{\imath}i^*|\bar{\imath}i^*\rangle^{(x)}) \tag{22}$$

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 eq 22 yields a term in which an energy weighted density matrix, W, contracts with the overlap derivatives.[32,33]

Our strategy for evaluating the two-electron gradients efficiently is the same as before.[25] Half-transformed coulomb and exchange matrices are contracted with the set of effective density matrices listed in Table 4. Note that the effective density matrices used in ref 25 have been redefined to accommodate the unrestriction. The resulting expression for the gradient is

$$L^{(x)}_{\mathrm{PP}} = E^{(\tilde{x})}_{\mathrm{ref}} + \sum_i^{\mathrm{pairs}} \lambda_i t_i (f^{(\tilde{x})}_{i^*i^*} + f^{(\tilde{x})}_{\bar{\imath}^*\bar{\imath}^*} - f^{(\tilde{x})}_{ii} - f^{(\tilde{x})}_{\bar{\imath}\bar{\imath}}) + \sum_i^{\mathrm{pairs}} \sum_{\mu\nu} (P^{[ii]}_{\mu\nu} J^{[ii](x)}_{\mu\nu} + P^{[\bar{\imath}\bar{\imath}]}_{\mu\nu} J^{[\bar{\imath}\bar{\imath}](x)}_{\mu\nu} - Q^{[ii]}_{\mu\nu} K^{[ii](x)}_{\mu\nu} - Q^{[\bar{\imath}\bar{\imath}]}_{\mu\nu} K^{[\bar{\imath}\bar{\imath}](x)}_{\mu\nu} + P^{[ii^*]}_{\mu\nu} J^{[ii^*](x)}_{\mu\nu} + P^{[\bar{\imath}\bar{\imath}^*]}_{\mu\nu} J^{[\bar{\imath}\bar{\imath}^*](x)}_{\mu\nu} + P^{[i^*i^*]}_{\mu\nu} J^{[i^*i^*](x)}_{\mu\nu} + P^{[\bar{\imath}^*\bar{\imath}^*]}_{\mu\nu} J^{[\bar{\imath}^*\bar{\imath}^*](x)}_{\mu\nu} - Q^{[i^*i^*]}_{\mu\nu} K^{[i^*i^*](x)}_{\mu\nu} - Q^{[\bar{\imath}^*\bar{\imath}^*]}_{\mu\nu} K^{[\bar{\imath}^*\bar{\imath}^*](x)}_{\mu\nu}) + \sum_{\mu\nu} W_{\mu\nu} S^{(x)}_{\mu\nu} \tag{23}$$

where the superscript "$(\tilde{x})$" 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.[31]

F. Orbital Optimization Convergence. In typical closed-shell organic molecules (i.e., those with single bonds connecting the atoms) near their equilibrium structures, the PP algorithm described here and in ref 25 converges within 10−30 iterations, which is comparable to the convergence rate of HF. However, in more complicated cases such as 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 SCF solutions in more challenging species, we often find multiple PP solutions depending on the initial localization scheme used. In general, Pipek−Mezey orbitals[34] provide the fastest convergence, but sometimes the Boys orbitals[35] lead to lower energy solutions.

To use PP widely, it might be helpful to adapt Pulay's direct inversion of the iterative subspace (DIIS) to the PP problem[36,37] to provide an alternative convergence scheme that could be faster when starting far away from the orbital minimum. 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

TABLE 2: 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)

$$J_{\mu\nu}^{[ii](x)} = \langle i\mu^{(x)}|i\nu\rangle + \langle i\mu|i\nu^{(x)}\rangle$$
$$K_{\mu\nu}^{[ii](x)} = \langle i\mu^{(x)}|\nu i\rangle + \langle i\mu|\nu^{(x)}i\rangle$$
$$J_{\mu\nu}^{[ii^*](x)} = \langle i\mu^{(x)}|i^*\nu\rangle + \langle i\mu|i^*\nu^{(x)}\rangle$$
$$K_{\mu\nu}^{[ii^*](x)} = \langle i\mu^{(x)}|\nu i^*\rangle + \langle i\mu|\nu^{(x)}i^*\rangle$$
$$J_{\mu\nu}^{[i^*i^*](x)} = \langle i^*\mu^{(x)}|i^*\nu\rangle + \langle i^*\mu|i^*\nu^{(x)}\rangle$$
$$K_{\mu\nu}^{[i^*i^*](x)} = \langle i^*\mu^{(x)}|i^*\nu\rangle + \langle i^*\mu|i^*\nu^{(x)}\rangle$$
$$P_{\mu\nu}^{[ii]} = \lambda_i t_i (C_{\mu i}C_{\nu i} - C_{\mu i^*}C_{\nu i^*} - C_{\mu\bar{i}^*}C_{\nu\bar{i}^*})$$
$$Q_{\mu\nu}^{[ii]} = \lambda_i t_i C_{\mu i^*}C_{\nu i^*}$$
$$P_{\mu\nu}^{[ii^*]} = (\lambda_i + t_i - \lambda_i t_i^2)C_{\mu\bar{i}}C_{\nu\bar{i}^*}$$
$$P_{\mu\nu}^{[i^*i^*]} = \lambda_i t_i (C_{\mu\bar{i}^*}C_{\nu i^*} - C_{\mu i}C_{\nu i} - C_{\mu\bar{i}}C_{\nu\mathrm{ib}})$$
$$Q_{\mu\nu}^{[i^*i^*]} = \lambda_i t_i (C_{\mu i}C_{\nu i}$$
$$W_{\mu\nu}^{\alpha} = -{}^1/_2 \sum_{i,j}^{\mathrm{occ,act}\alpha} \sum_{\sigma} C_{\mu i}(\partial L/\partial C_{\sigma i})C_{\sigma j}C_{j\nu}^T$$

that the variationally lowest energy solution was found for all of the results reported herein.

III. 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.[38] Unless otherwise specified, all PP calculations use the RI approximation with the auxiliary basis sets developed for RIMP2/cc-pVXZ calculations,[30,39] 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.

A. Bond Breaking. To begin, we revisit the bond-breaking problem in the context of the N_2 molecule dissociating into two quartet-state nitrogen atoms. Figure 1 compares the HF, PP, and VOD results against FCI results (with the frozen core approximation)[40] in the cc-pVDZ basis. It comes as no surprise that RHF rises in energy much too quickly and dissociates to a high asymptote, leading to a potential energy surface (PES) that is very nonparallel to the FCI one. RPP is significantly more parallel to FCI than RHF and does not exhibit the nonvariational 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 (one σ bond, two π bonds, and two lone pairs at equilibrium) should asymptote to two equivalent 2s pairs with nonzero correlation energy contribution and three amplitudes with magnitude zero corresponding to the nonexistent interaction between the three alpha unpaired 2p electrons on one N atom and the three beta unpaired 2p 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 nonzero. 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 interpair 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 a 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

TABLE 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	*r*CH	1.063	−0.008	0.008	0.013	0.000
	*r*CC	1.203	−0.020	−0.001	0.014	−0.005
C_2H_4	*r*CH	1.081	−0.004	0.013	0.019	0.004
	*r*CC	1.334	−0.018	0.011	0.015	−0.007
$^1CH_2$	*r*CH	1.107	−0.022	−0.002	0.023	−0.010
CH_4	*r*CH	1.086	−0.002	0.018	0.019	0.005
Cl_2	*r*ClCl	1.988	0.015	0.095	0.016	0.069
ClF	*r*ClF	1.628	−0.009	0.091	0.041	0.051
CO	*r*CO	1.128	−0.023	−0.010	0.002	−0.001
CO_2	*r*CO	1.160	−0.025	−0.014	0.001	0.000
CS	*r*CS	1.535	−0.019	−0.007	0.004	0.007
F_2	*r*FF	1.412	*b*	0.131	0.008	−0.005
H_2	*r*HH	0.741	−0.006	0.015	0.015	0.001
H_2CO	*r*CH	1.116	−0.021	−0.009	0.003	−0.006
	*r*CO	1.208	−0.029	0.002	0.000	−0.008
H_2CS	*r*CH	1.093	−0.015	0.003	0.009	−0.004
	*r*CS	1.611	0.017	0.021	0.013	0.005
H_2O	*r*OH	0.957	−0.016	0.002	0.004	0.005
H_2O_2	*r*OH	0.965	−0.023	−0.003	0.001	−0.003
	*r*OO	1.452	−0.059	0.043	0.004	0.009
H_2S_2	*r*SH	1.345	−0.016	0.007	0.000	0.002
	*r*SS	2.058	0.066	0.125	0.073	0.116
HCN	*r*CH	1.065	−0.007	0.007	0.013	0.002
	*r*CN	1.153	−0.026	−0.002	0.008	−0.005
HCP	*r*CH	1.069	−0.004	0.011	0.017	0.004
	*r*CP	1.540	0.019	0.009	0.016	0.000
HF	*r*HF	0.917	−0.021	−0.003	−0.002	0.005
HNC	*r*NH	0.994	−0.010	0.004	0.011	0.006
	*r*NC	1.169	−0.020	−0.005	0.007	−0.001
HNO	*r*HN	1.063	−0.037	−0.016	0.001	−0.002
	*r*NO	1.212	−0.021	0.001	−0.003	−0.009
HOCl	*r*OH	0.975	−0.031	−0.011	−0.008	−0.007
	*r*OCl	1.690	−0.017	0.085	0.036	0.047
HOF	*r*OH	0.966	−0.020	0.000	0.004	0.005
	*r*OF	1.442	−0.080	0.046	−0.014	−0.008
Li_2	*r*LiLi	2.673	0.257	0.016	0.016	0.032
N_2	*r*NN	1.098	−0.028	−0.002	0.006	−0.003
N_2H_2	*r*NH	1.028	−0.015	0.003	0.012	0.007
	*r*NN	1.252	−0.040	−0.003	0.002	−0.011
NaF	*r*NaF	1.926	−0.014	−0.008	0.003	−0.009
NH_3	*r*NH	1.012	−0.011	0.008	0.010	0.003
NNO	*r*NN	1.127	−0.042	−0.024	−0.003	−0.001
	*r*NO	1.185	−0.016	0.017	0.012	−0.001
NP	*r*NP	1.491	0.038	0.005	0.007	−0.002
NSF	*r*NS	1.448	−0.023	−0.011	−0.001	0.003
	*r*SF	1.643	−0.017	0.016	0.060	0.069
P_2	*r*PP	1.893	0.072	0.018	0.020	0.004
SO2	*r*SO	1.485	−0.078	−0.066	−0.041	−0.027
SO3	*r*SO	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.

TABLE 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	*r*BeF	1.361	0.003	0.011	0.022	0.011
BeH	*r*BeH	1.343	0.000	0.025	0.025	0.000
BH^+	*r*BH	1.215	−0.027	−0.006	−0.009	−0.010
BH_2	*r*BH	1.181	0.004	0.026	0.028	0.007
BN	*r*BN	1.281	0.007	0.029	0.046	0.038
BO	*r*BO	1.205	−0.024	−0.011	0.002	−0.003
C_2^+	*r*CC	1.301	0.101	0.000	0.007	0.100
CF	*r*CF	1.272	−0.015	0.013	0.009	0.007
CH	*r*CH	1.120	−0.017	0.014	0.018	0.008
$^3CH_2$	*r*CH	1.078	−0.006	0.012	0.016	0.002
CH_2	*r*CH	1.079	−0.006	0.011	0.013	0.001
CN	*r*CN	1.172	−0.018	−0.006	0.005	−0.006
CO^+	*r*CO	1.115	−0.027	−0.013	0.000	−0.005
F_2^+	*r*FF	1.322	*b*	0.129	−0.022	−0.023
FH^+	*r*FH	1.001	−0.023	−0.006	−0.001	0.008
HCO	*r*HC	1.125	−0.016	−0.004	0.011	0.002
	*r*CO	1.175	−0.023	−0.002	0.005	−0.001
HNF	*r*HN	1.060	−0.047	−0.026	−0.020	−0.024
	*r*NF	1.370	−0.036	0.016	0.010	0.002
HO_2	*r*OH	0.977	−0.030	−0.013	−0.004	−0.002
	*r*OO	1.335	−0.036	0.023	0.002	−0.007
N_2^+	*r*NN	1.116	−0.008	0.006	0.006	−0.007
NF	*r*NF	1.317	−0.026	0.016	0.011	0.001
NH^+	*r*NH	1.070	−0.016	0.003	0.006	0.010
NH	*r*NH	1.036	−0.013	0.010	0.014	0.009
NH_2	*r*NH	1.024	−0.012	0.008	0.013	0.007
NO	*r*NO	1.151	−0.034	−0.008	0.002	−0.003
O_2^+	*r*OO	1.116	−0.055	−0.019	−0.006	−0.010
O_2	*r*OO	1.208	−0.051	−0.033	−0.003	−0.002
OH^+	*r*OH	1.028	−0.021	0.000	0.003	0.009
OH	*r*OH	0.970	−0.018	0.001	0.004	0.006
OH_2^+	*r*OH	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^+.

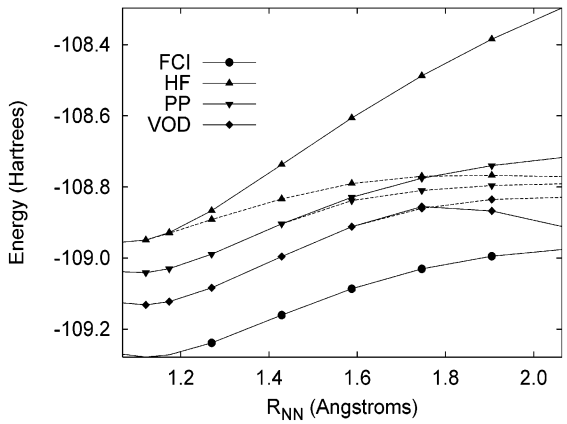

Figure 1. Restricted (solid lines) and unrestricted (dashed lines) N_2 bond breaking in the cc-pVDZ basis. The correlated methods employ a frozen core approximation, and the FCI results were obtained from ref 57.

with the active orbitals in PP and VOD. The FCI data also employ 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 that for the UHF curve 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.[41] 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 nonparallelity 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 a 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

$$\Psi_{^2\Pi_g} = ...(\sigma_g)^2(\pi_u)^4(\pi_g^*)^3 \tag{24}$$

crosses with an excited state with configuration

$$\Psi_{^2\Pi_u} = ...(\sigma_g)^2(\pi_u)^4(\pi_g^*)^2(\pi_u^*) \tag{25}$$

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 palatable.

B. 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 wish to know 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 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 refs 42 and 43 along with others

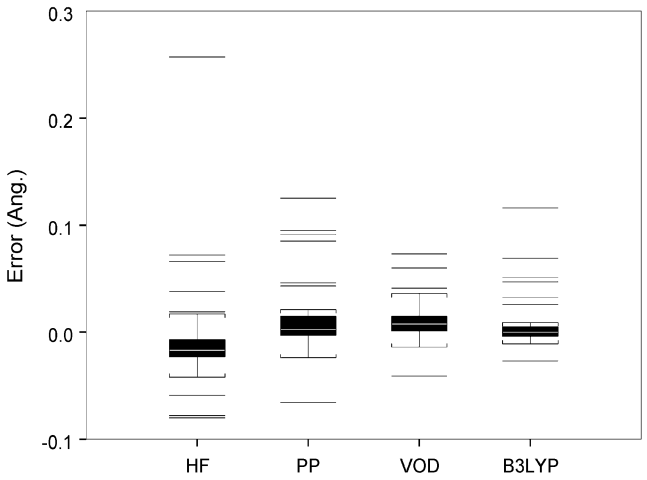

Figure 2. Errors in 6-311G** predicted bond lengths versus experiment for a set of small, closed-shell molecules.

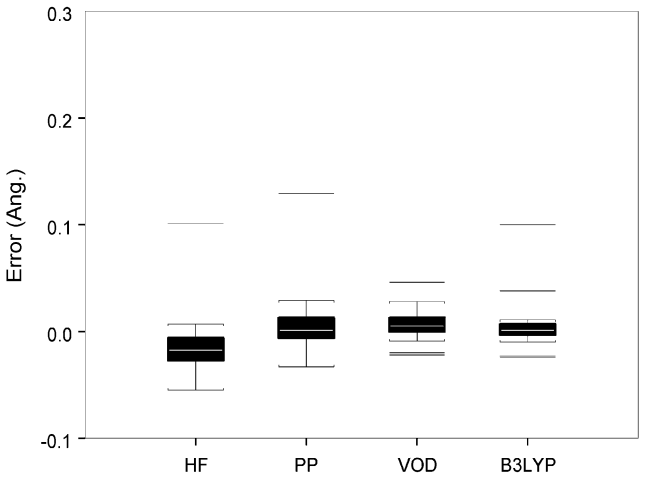

Figure 3. Errors in 6-311G** predicted bond lengths versus experiment for a set of small, doublet- and triplet-state open-shell molecules. Note that HF excludes F_2^+, since it is unbound.

found in refs 44−47. The errors are presented in Table 3 and pictorially using box plots in Figure 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 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,[42,43,48] PP predicts much too long of a bond length for F_2 and other halogenated species, for which interpair correlations are important. Many of the species for which PP has the most difficulty—H_2S_2 (Δr_{SS} = 0.125 Å), Cl_2 (0.095 Å), ClF (0.091 Å), and HOCl (Δr_{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 hypervalent species, for which the electron pairing is perhaps more complicated than in standard octet-rule-obeying structures. The test set includes three hypervalent 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 those 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 4 and Figure 3 present UHF, UPP, UVOD, and UB3LYP results for 29 diatomic and triatomic open-shell species containing 32 unique bonds.[49] 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 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 interpair correlations to be reliable.[43,48] 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-shell systems (typically accurate to within 0.01 Å) with some improvement over HF at a much lower cost than VOD. However, the limited 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 of individual electron pairs is necessary.[48] 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.

C. 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, and so forth. Byrd and co-workers demonstrated that standard methods such as MP2 and CCSD(T) performed far below their standard, closed-shell system levels of accuracy on the geometries and vibrational frequencies of various small

TABLE 5: Errors in 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] Reference 51. [b] Reference 54. [c] Reference 55. [d] Reference 56.

TABLE 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

TABLE 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

radicals.[50] It was later demonstrated 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.[51] In particular, this latter study demonstrated that methods such as orbital-optimized coupled cluster doubles with perturbative triples (OD(T)) or KS-CCSD(T), which uses Kohn−Sham orbitals as a more stable reference than the HF ones, 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 12 diatomic radicals studied in ref 51 using unrestricted HF, PP, and VOD in the cc-pVTZ basis (and the auxiliary basis set for cc-pVTZ in the PP case). These results are presented in Table 5. We see that, for these species, HF is typically 50−600 cm^{-1} in error in this basis set, with the worst cases percentagewise 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, 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 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.

The worst frequencies (as compared against experiment percentagewise) 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 4 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 4, 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 challenging diatomic radicals in terms of symmetry breaking with the general difficulty PP has treating halogens, for which interpair correlations are important.[43,48] 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*[52] 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 than the UHF one 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 such as 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 highest occupied molecular orbital (HOMO) and SOMO orbitals for allyl at the symmetric, UB3LYP/cc-pVDZ-optimized geometry are plotted in Figure 5. At this point on the PES,

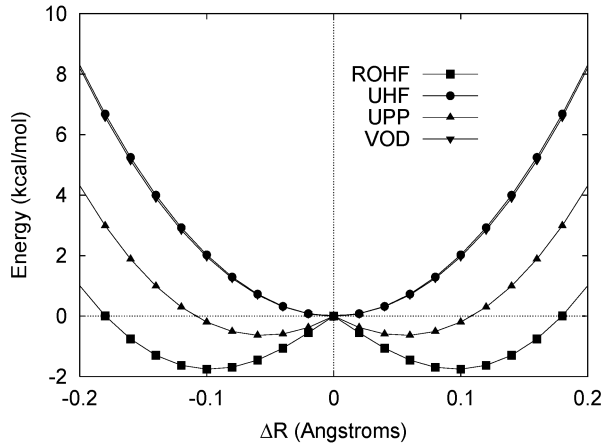

Figure 4. Symmetry breaking in the allyl radical in the cc-pVDZ basis. The deformation, ΔR, is relative to the UB3LYP/cc-pVDZ-optimized symmetric structure. The UHF and VOD curves are virtually coincident on the energy scale plotted here.

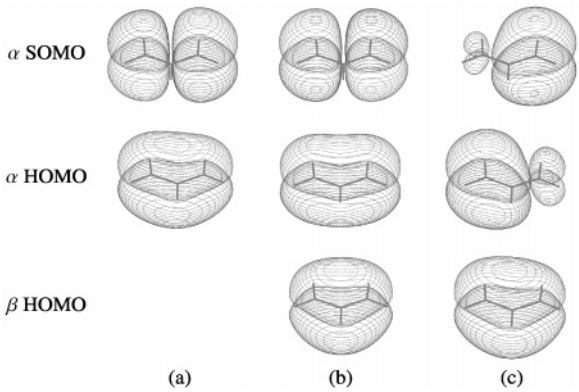

Figure 5. Occupied π molecular orbitals for allyl at the (a) ROHF, (b) UHF, and (c) UPP levels in the cc-pVDZ basis, using the symmetric UB3LYP/cc-pVDZ-optimized structure.

ROHF orbitals are symmetry-broken, with a slight shift in the HOMO toward the left carbon atom and the SOMO toward the right carbon atom. UHF, on the other hand, exhibits perfectly symmetrical orbitals. Therefore, correlation methods such as 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, 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 electron primarily occupies the other carbon atom.

A similar picture emerges energetically. Figure 4 plots the slice of the potential energy surface for deforming the 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 about 0.65 kcal/mol, UPP favors the symmetry-broken structure with one C−C bond about 0.06 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 4. 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 can provide the impetus for additional spatial symmetry breaking, as in the allyl radical.

IV. Conclusions

Perfect pairing provides the leading correlation correction beyond Hartree−Fock at only a 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 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.

In this article, we have not addressed how the coupled cluster formulation of UPP used here differs from the alternative unrestricted GVB one. The primary difference between the two formulations is that the amplitudes are solved for projectively in the CC version and variationally in the GVB version.[24] The projective approach is known to lead to spurious nonvariational behavior in multiple-bond breaking using restricted CCD that disappears if the CCD equations are solved variationally.[53] The fact that the CC version of PP treats each pair separately and solves for the pair amplitude in a CI-like fashion suggests that it should always be well-behaved. In no case, restricted or unrestricted, have we ever observed such nonvariational behaviors with PP. Therefore, we do not anticipate significant differences between the two formalisms. Nevertheless, investigating the unrestricted GVB formalism would make for an interesting future study.

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, and so forth.

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).

The use of orthogonal, localized orbitals causes PP to slightly favor D_{3h} symmetry over D_{6h} symmetry for benzene or likewise to predict that the allyl radical has two unequal carbon−carbon bonds rather than equivalent ones. In cases such as these, additional correlation effects must be included to remove the symmetry breaking, such as with a nonorthogonal formalism.[16]

In addition, the asymmetry in how UPP treats the unpaired electrons versus the paired ones can lead to some strange chemical predictions in certain classes of systems. Singlet−triplet gaps, for example, may be poor, as the singlet is

preferentially stabilized relative to the triplet through the correlation of an additional electron pair. This does counteract the tendency of HF to favor the triplet due to the nature of the exchange interaction, but it may overcompensate. For example, in the 6-31G* basis, $^3F^+$ is only 3 kcal/mol more stable than the singlet with UPP. In contrast, VOD prefers the triplet by 66 kcal/mol.

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 report on just such a combination, the PP(2) method, soon.

Acknowledgment. This work was supported by the Department of Energy, Office of Basic Energy Sciences, SciDAC Computational Chemistry Program (Grant No. DE-FG0201-ER403301).

References and Notes

(1) Roos, B. O. *Adv. Chem. Phys.* **1987**, *69*, 399.
(2) Head-Gordon, M.; Voorhis, T. V.; Beran, G. J.; Dunietz, B. D. Computational Science-ICCS 2003, Pt IV, *Proceedings Lecture Notes in Computer Science* **2003**, *2660*, 96.
(3) Krylov, A. I.; Sherrill, C. D.; Byrd, E. F. C.; Head-Gordon, M. *J. Chem. Phys.* **1998**, *109*, 10669.
(4) Sherrill, C. D.; Krylov, A. I.; Byrd, E. F. C.; Head-Gordon, M. *J. Chem. Phys.* **1998**, *109*, 4171.
(5) Purvis, G. D., III; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910.
(6) Rassolov, V. A. *J. Chem. Phys.* **2002**, *117*, 5978.
(7) Hurley, A. C.; Lennard-Jones, J.; Pople, J. A. *Proc. R. Soc. London, Ser. A* **1953**, *220*, 446.
(8) Hunt, W. J.; Hay, P. J.; Goddard, W. A., III. *J. Chem. Phys.* **1972**, *57*, 738.
(9) Goddard, W. A., III; Harding, L. B. *Annu. Rev. Phys. Chem.* **1978**, *29*, 363.
(10) Bobrowicz, F. B.; Goddard, W. A. In *Methods of Electronic Structure Theory 3*; Schaefer, H. F., III, Ed.; Plenum Press: 1977.
(11) Shoemaker, J.; Burggraf, L. W.; Gordon, M. S. *J. Chem. Phys.* **2003**, *112*, 2994.
(12) Jung, Y.; Head-Gordon, M. *ChemPhysChem* **2003**, *4*, 522.
(13) Jung, Y.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2008.
(14) Langlois, J.-M.; Muller, R. P.; Coley, T. R.; Goddard, W. A., III; Ringnalda, M. N.; Won, Y.; Friesner, R. A. *J. Chem. Phys.* **1990**, *92*, 7488.
(15) Van Voorhis, T.; Head-Gordon, M. *Chem. Phys. Lett.* **2000**, *317*, 575.
(16) Van Voorhis, T.; Head-Gordon, M. *J. Chem. Phys.* **2000**, *112*, 5633.
(17) Cooper, D. L.; Gerratt, J.; Raimondi, M. *Chem. Rev.* **1991**, *91*, 929.
(18) Raimondi, M.; Cooper, D. L. *Top. Curr. Phys.* **1999**, *203*, 105.
(19) Feyereisen, M. W.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359.
(20) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514.
(21) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283.
(22) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119.
(23) Ukrainskii, I. I. *Theor. Math. Phys.* **1977**, *32*, 816.
(24) Cullen, J. *Chem. Phys.* **1996**, *202*, 217.
(25) Van Voorhis, T.; Head-Gordon, M. *J. Chem. Phys.* **2002**, *117*, 9190.
(26) Sano, T. *THEOCHEM* **2000**, *528*, 177.
(27) Faglioni, F.; Goddard, W. A., III. *Int. J. Quantum Chem.* **1999**, *73*, 1.
(28) Van Voorhis, T.; Head-Gordon, M. *J. Chem. Phys.* **2001**, *115*, 7814.
(29) Van Voorhis, T.; Head-Gordon, M. *Mol. Phys.* **2002**, *100*, 1713.
(30) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(31) Sodt, A.; Beran, G. J. O.; Head-Gordon, M. Manuscript in preparation, 2005.
(32) Pople, J. A.; Raghavachari, K.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem. Symp.* **1979**, *13*, 225.
(33) Gauss, J. In *Modern methods and algorithms of quantum chemistry*; Grotendorst, J., Ed.; NIC Series; John von Neumann Institute for Computing: Jülich, Germany, 2000; Vol. 3, pp 541−592.
(34) Pipek, J.; Mezey, P. G. *J. Phys. Chem.* **1989**, *90*, 4916.
(35) Boys, S. F. In *Quantum Theory of Atoms, Molecules, the Solid State*; Löwdin, P. O.; Academic: New York, 1968; pp 253−280.
(36) Muller, R. P.; Langlois, J. M.; Ringnalda, M. N.; Friesner, R.; Goddard, W., III. *J. Chem. Phys.* **1994**, *100*, 1226.
(37) Ionova, I. V.; Carter, E. A. *J. Chem. Phys.* **1995**, *102*, 1251.
(38) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Van Voorhis, T.; Oumi, M.; Hirata, S.; Hsu, C.-P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **2000**, *21*, 1532.
(39) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175.
(40) Larsen, H.; Olsen, J.; Jørgensen, P.; Christiansen, O. *J. Chem. Phys.* **2000**, *113*, 6677 (cc-pVDZ basis).
(41) Krylov, A. I. *J. Chem. Phys.* **2000**, *113*, 6052.
(42) Wang, Y.; Poirier, R. A. *J. Mol. Struct.* **1995**, *340*, 1.
(43) Byrd, E. F.; Van Voorhis, T.; Head-Gordon, M. *J. Phys. Chem. B* **2002**, *106*, 8070.
(44) Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *J. Chem. Phys.* **1997**, *106*, 6430.
(45) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. *J. Chem. Phys.* **2001**, *114*, 6548.
(46) Martin, J. M. L. *J. Chem. Phys.* **1994**, *100*, 8186.
(47) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*; Wiley: New York, 1986.
(48) Beran, G. J. O.; Head-Gordon, M. *Mol. Phys.*, submitted for publication, 2005.
(49) Farnell, L.; Pople, J. A.; Radom, L. *J. Phys. Chem.* **1983**, *87*, 79.
(50) Byrd, E. F. C.; Sherrill, C. D.; Head-Gordon, M. *J. Phys. Chem. A* **2001**, *105*, 9736.
(51) Beran, G. J. O.; Gwaltney, S. R.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2488.
(52) Dunietz, B. D.; Head-Gordon, M. *J. Phys. Chem. A* **2003**, *107*, 9160.
(53) Van Voorhis, T.; Head-Gordon, M. *J. Chem. Phys.* **2000**, *113*, 8873.
(54) Herzberg, G. *Molecular Spectra and Molecular Structure I. Spectra of Diatomic Molecules*; Van Nostrand Reinhold: New York, 1950.
(55) Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules*; Van Nostrand Reinhold: New York, 1979.
(56) Cormack, A. J.; Yencha, A. J.; Donovan, R. J.; Lawley, K. P.; Hopkirk, A.; King, G. C. *Chem. Phys.* **1996**, *213*, 439.
(57) Gwaltney, S. R.; Byrd, E. F. C.; Van Voorhis, T.; Head-Gordon, M. *Chem. Phys. Lett.* **2002**, *353*, 359.