Considerations in constructing a multireference second-order perturbation theory

Pawel M. Kozlowski and Ernest R. Davidson

Department of Chemistry, Indiana University, Bloomington, Indiana 47405

(Received 13 September 1993; accepted 16 November 1993)

Several possible definitions for a multireference second-order perturbation theory are suggested. These are tested against some standard test problems from the literature.

I. INTRODUCTION

For a closed shell ground state with no near-degeneracy effects, a Hartree–Fock calculation followed by second-order many-body perturbation theory, MBPT2, provides reasonably accurate results at low cost in computer time. Hence, for a multireference second-order perturbation theory, we would like a method that reduces to the MBPT2 method for a single closed shell reference state and which retains the important features of MBPT2 for multireference cases. At the same time, we seek a method that overcomes the major defect of MBPT2—it does not apply to excited states or nearly degenerate ground states. Consequently, it does not give uniform accuracy over much of the ground state potential energy surface.

MBPT2 has many special properties, some of which must be lost in any generalization, so it is important to realize which features are essential for accuracy and low cost. Among the attributes of closed-shell single reference MBPT2, that are sometimes claimed to lead to its success, are the following.

1. There exists a “one-body” operator, $H_0$, given as the sum of the Fock operators for each electron, $F(i)$. The zeroth-order reference state is an eigenfunction of this one-body operator

$$H_0\Psi_0 = E_0 \Psi_0,$$

(1)

2. When formulated as a variational-perturbation theory, the MBPT2 energy is given by

$$E_1 = \langle \Psi_0 | H_1 | \Psi_0 \rangle,$$

(2)

$$E_2 = \langle \Psi_0 | H_2 | \Psi_0 \rangle.$$  

3. The energy is size consistent. The energy of a dilute gas is the sum of the energies of the monomers. The energy of a nearly infinite polymer, liquid, or periodic solid is linear in the number of monomers.

4. The canonical orbital energies for occupied orbitals correspond to negatives of ionization energies. They are defined using the repulsion of $N-1$ electrons. The orbital energies of virtual orbitals correspond to negatives of electron affinities. They are defined using the repulsion from $N$ electrons. While this would actually seem to be a defect of MBPT2 theory since differences of these orbital energies are poor approximations to excitation energies, empirical tests have shown that the energy denominators in MBPT2 theory give a better second-order energy in most cases than any other method that has been suggested.

5. The canonical orbital energies for occupied orbitals correspond to negatives of ionization energies. They are defined using the repulsion of $N-1$ electrons. The orbital energies of virtual orbitals correspond to negatives of electron affinities. They are defined using the repulsion from $N$ electrons. While this would actually seem to be a defect of MBPT2 theory since differences of these orbital energies are poor approximations to excitation energies, empirical tests have shown that the energy denominators in MBPT2 theory give a better second-order energy in most cases than any other method that has been suggested.

6. The wave function in every order of the expansion is a spin eigenfunction. This is because both $H$ and $H_0$ commute with $S^2$.

It is not immediately clear which of these features of MBPT2 is responsible for its accuracy and low cost, and which are irrelevant. Extension to a spin-unrestricted Hartree–Fock (UHF) reference wave function and different Fock operators for different spin loses only feature (6), yet the convergence is somewhat slower. It is not clear, however, whether this is due to an inherent problem with broken-spin symmetry, or is just due to the higher density of states in open shell systems. As in closed shell theory, the occupied orbitals feel a $V_{N-1}$ potential and the virtual orbitals a $V_N$ potential. Certainly use of MBPT2 for closed shell systems where configurations other than the Hartree–Fock have appreciable coefficients also leads to slow convergence.

Several recent extensions of MBPT2 to the situation of a high-spin spin-restricted open shell Hartree–Fock (ROHF) reference wave function have been proposed (i.e., a wave function in which the space part of each occupied spin-down orbital is also used to form an occupied spin-up orbital). None of these retain all of the features of MBPT2 mentioned above, but each proposal has selected certain features to be retained. The methods differ mainly in the definition of “canonical” orbitals and in the definition of $H_0$. Knowles et al. and Lauderdale et al. proposed that spin-up and spin-down Fock operators be evaluated with the spin-up and spin-down ROHF densities and then the occupied and virtual blocks of the operators be separately diagonalized to define a set of orbitals and orbital energies.
This transformation leaves the ROHF wave function unchanged, but makes the orbitals as much like UHF as possible with that constraint. The unperturbed Hamiltonian is chosen to be

$$H_0 = |\alpha\rangle F_\alpha \langle \alpha | + |\beta\rangle F_\beta \langle \beta |.$$

(4)

This $H_0$ is not diagonal in the basis of Slater determinants formed from these orbitals, so a second perturbation expansion is then made to express $(H_0 - E_0)^{-1}$ as a series involving only its diagonal elements in the denominator. The energy in second order is then given by the same equation as UHF but off-diagonal elements of $(H_0 - E_0)^{-1}$ appear in higher order. While $\Psi_0$ is a spin eigenfunction in this approximation, $H_0$ and $H_1$ are spin dependent, as is $\Psi_1$. Also, $\Psi_0$ is not an eigenfunction of this $H_0$. In the limit that the ROHF wave function is identical to the UHF wave function, this method reduces to UMP2. In a slight variation on this, Amos et al. earlier proposed diagonalizing the alpha and beta Fock matrices in the three subspaces corresponding to doubly occupied, singly occupied, and virtual orbitals. These methods seem to work mostly because they preserve feature (5) listed above (occupied orbital energies see a $V_{N-1}$ potential and virtual orbitals see a $V_N$ potential).

Others have retained the spin restricted features in the choice of canonical orbitals. If the spin-averaged Fock operator is defined as

$$F = F_s + \frac{1}{2} (F_a + F_b),$$

(5)

then Hubač and Čarsky defined canonical orbitals and orbital energies by diagonalizing $2F_s - F_a$ in the doubly occupied space, $F_s$ in the singly occupied space, and $2F_s - F_b$ in the virtual space. $H_0$ was then defined as in Eq. (3) with no spin dependence. The doubly occupied orbitals "see" a $V_{N-1}$ potential, the singly occupied see a $V_{N-1/2}$ potential, and the virtual orbitals see a $V_N$ potential. Because of this, excitations involving the singly occupied subspace tend to have smaller denominators than in the ROHF/UMP2 approach. This definition did not lead to good convergence.

Murray and Davidson defined two methods. In OPT1, the canonical orbitals are defined by diagonalizing $F$ separately in the doubly occupied, singly occupied, and virtual subspaces. $H_0$ was then defined as in Eq. (3) using the resulting eigenvalues as orbital energies. Alternatively, this $H_0$ can be expressed using projection operators $P_x$ ($x = d, s, u$ for doubly occupied, single occupied, or virtual subspaces) as

$$H_0 = F - \frac{1}{2} P_d (F_a - F_b) P_d - \frac{1}{2} P_s (F_a - F_b) P_s + \frac{1}{2} P_s (F_a - F_b) P_d + \frac{1}{2} P_d (F_a - F_b) P_s.$$

(6)

This gives a $V_{N-1}$, $V_{N-1/2}$ and $V_N$ potential in the doubly occupied, singly occupied, and virtual subspaces, respectively. Like the Hubač and Čarsky definition, this gives some small denominators and poor convergence for some cases. Murray and Davidson defined a second method, called OPT2, that modified the orbital energies for the canonical orbitals by adding a population dependent correction so that (effectively)

$$H_0 = \sum n_k \epsilon_k - \sum m_{m=1}^{n_m-1} (m | \frac{1}{2} (F_a - F_b) | m).$$

(7)

where the second sum is only over the singly occupied orbitals. When the population of orbital $m$ changes from 1 in the reference configuration to 0 in configuration $K$, this contributes

$$- (m | \frac{1}{2} (F_a - F_b) | m) - \epsilon_m = - (m | F_\alpha | m)$$

(8)

to the denominator $E_K - E_0$. Conversely, if the population changes to 2 in configuration $K$, this orbital contributes

$$2\epsilon_m - (m | \frac{1}{2} (F_a - F_b) | m) - \epsilon_m = + (m | F_\beta | m)$$

(9)

to the denominator. This scheme gave convergence similar to ROHF/UMP2.

These ROHF/MBPT2 schemes provide a further guide to the form desired for a multireference (MR) MBPT2. In fact, these are special cases of complete active space self-consistent-field (CASSCF) wave functions for which many methods have been recently suggested. A CAS wave function includes all possible arrangements of the "active" electrons among the active orbitals. While the choice of active space is arbitrary, it includes as a minimum all orbitals whose occupations change during the process under consideration—be it excitation or chemical reaction. This space of Slater determinants may also be called a "complete model space" even if the orbitals are not chosen by a CASSCF optimization. There are several of these (MR)MBPT2 methods in the literature, with new ones appearing frequently. Here we mention only a few to illustrate some types of approach to the problem. We have chosen to organize the methods into two main families depending on whether the perturbing effect of configurations outside the CAS is considered before or after diagonalization of the Hamiltonian over the CAS configurations. Many of these methods work, also, for an importance selected multifreference space that is an "incomplete" model space.

The first family of methods chooses to treat the CAS space by diagonalization to give a $\psi_0$ (where $\psi_0$ may refer to an excited state). The space that interacts with $\psi_0$ may be generated by applying elementary single and double excitation operators to $\psi_0$ to generate linear combinations of Slater determinants. Wolinski et al. chose $H_0$ to have the form

$$H_0 = P_0 F P_0 + P_3 F P_3 + P_6 F P_6 + \cdots,$$

(10)

where the $P_k$ are projectors onto $\psi_0$, the singly excited interacting space, the doubly excited interacting space, etc. As in most of the MR methods, $\psi_0$ is not an eigenfunction of $F$ and $E_0$ is defined as $\langle \psi_0 | H_0 | \psi_0 \rangle$. The Fock operator $F$ is defined as

$$F = |\alpha\rangle F_\alpha \langle \alpha | + |\beta\rangle F_\beta \langle \beta |,$$

(11)
where \( F_\alpha \) and \( F_\beta \) are evaluated using \( \rho_\alpha \) and \( \rho_\beta \) formed from \( \psi_0 \). The subspaces are generated by acting on \( \psi_0 \) with excitation operators so that they contain only eigenfunctions of \( S^2 \) and \( S_z \) with the desired eigenvalues. Hence, \( H_0 \) commutes with \( S^2 \) even though \( F \) does not. The first-order wave function and second-order energy are found from equations like Eq. (2) so these results are independent of the choice of orbitals within the active and virtual subspaces.

Roos and co-workers\textsuperscript{19-21} use a very similar approach except that \( F \) is \( \frac{1}{2}(F_\alpha + F_\beta) \) for both the \( |\alpha\rangle \) and \( |\beta\rangle \) spin orbitals. For the case of ROHF, the OPT1 method of Murray and Davidson\textsuperscript{14} is a special case of this method introduced earlier by Roos. A disadvantage to both the Roos and Pulay methods is that a very large set of linear equations must be solved.

Hirao\textsuperscript{25-28} has suggested a conceptually simpler method. He chooses the same spin-averaged \( F \) as Roos, but then defines \( H_0 \) to be

\[
H_0 = \sum n_i \epsilon_i \tag{12}
\]

with \( \epsilon_i \) the diagonal elements of \( F \) in a particular basis. His basis consists of the natural orbitals of \( \psi_0 \) within the active space. Hirao expands \( \psi_i \) in the full set of Slater determinants formed as single and double excitations from the Slater determinants in the reference set. For a large CAS, this space is much larger than the interacting space used by Pulay and Roos, but has the advantage that \( H_0 \) is diagonal in this space so \((H_0 - E_0)^{-1}\) is easily formed in the SD subspace. Hirao recognized that this approach did not work well for the high spin ROHF case, so he also suggested modifying \( H_0 \) by defining a high spin \( F_{\text{hS}} \) as

\[
F_{\text{hS}} = \frac{1}{2} (F_\alpha + F_\beta) - \frac{1}{2} \sum n_i^0 K_{ii} \tag{13}
\]

\[
- \frac{1}{2} \sum_{n_i > 1} (2-n_i^0) K_{ii},
\]

where \( n_i^0 \) is the average occupation of orbital \( i \) in the reference function \( \psi_0 \) and \( K_{ii} \) is the self-exchange integral. For a high spin ROHF \( \psi_0 \), this gives \( F = F_\alpha \). In general, \( F \) corresponds to an electrostatic potential with \( N - n_i^0 \) electron repulsions for an active orbital, while in \( F_{\text{hS}} \) it would be \( N - n_i^0 \) for \( n_i^0 < 1 \) and \( N - 1 \) for \( n_i^0 > 1 \).

Versions of (MR)MBPT2 with an arbitrary reference space\textsuperscript{29-32} have been a part of MELD\textsuperscript{37} and other perturbation-selected CI programs for many years. MELD expands \( \psi_i \) in spin adapted linear combinations of Slater determinants generated, like Hirao, as single and double excitations of the model space. Unlike the methods mentioned above, MELD uses a modified Epstein–Nesbet \( H_0 \) formed from diagonal elements of \( H \) averaged over families of Slater determinants related by spin flips.\textsuperscript{38,39} MELD also uses \( \langle \psi_0 | H | \psi_0 \rangle \) for \( E_0 \). In MELD, this method is used as a basis for selecting important configurations outside of the model space to be treated by configuration interaction, CI, and for extrapolation of the energy to estimate the effect of configurations not included in the CI. The three subspace philosophy of this approach is very similar to the CIPSI method recently advocated by Malrieu and co-workers.\textsuperscript{34,35} Malrieu has noted, however, that use of \( \langle \psi_0 | H | \psi_0 \rangle \) instead of \( \langle \psi_0 | H | \psi_0 \rangle \) for \( E_0 \) is more nearly size consistent. He also prefers a Møller–Plesset \( H_0 \). The experience with MELD has been that the Epstein–Nesbet \( H_0 \) gives about a 15% overestimate of the magnitude of the correlation energy. The choice of orbitals to use with an Epstein–Nesbet method is also crucial and various choices have been tried.\textsuperscript{40,41} Generally, canonical Hartree–Fock orbitals or improved virtual orbitals\textsuperscript{40} make \( H \) more nearly diagonal and make MBPT2 more accurate. At the same time, \( K \) orbitals\textsuperscript{41} or other approximate natural orbitals make the CI more compact, and hence more effective, but diminish the utility of the perturbation estimates.

The other family of methods treats the multireference space as quasidegenerate with diagonalization after perturbation correction to give an effective Hamiltonian \( H_{\text{eff}} \). Here again, there are two approaches. The one we will discuss follows a Löwdin\textsuperscript{42-44} partitioning to construct a \( H_{\text{eff}} \) with only one eigenvalue approximating an eigenvalue of \( H \). The other possible approach follows Bloch\textsuperscript{45-55} and constructs a \( H_{\text{eff}} \) with all eigenvalues approximating eigenvalues of \( H \). The amount of effort required in second order is about the same in each approach but the Löwdin partitioning should allow second order to be a better approximation in the common case where only one eigenvalue is of interest. Again, MELD\textsuperscript{56-59} has incorporated this approach for many years using an Epstein–Nesbet \( H_0 \). In its Brillouin–Wigner form, this method is identical to the Shavitt\textsuperscript{60,61} \( B_K \) method, but the Rayleigh–Schrödinger form is more nearly size consistent.

II. METHOD

In this paper, we adopt a general strategy for a multireference second-order perturbation theory. Within this strategy, many methods are possible that differ in small details. We apply a few of the possible methods to some example systems in an effort to locate the optimal scheme. The steps in the general strategy are as follows.

1. A canonical spin-restricted molecular orbital basis is defined based on some preliminary calculations. This basis is supposed to be partitioned into core, inactive, active, and virtual orbitals. Only the active orbitals were allowed to have occupation numbers other than 0 or 2 in the preliminary calculations. In all of the calculations considered here, the charge-density matrix from the preliminary calculation was used to define

\[
F = h + J(\rho) - \frac{1}{2} K(\rho) \tag{14}
\]

and the canonical orbitals separately block diagonalized this in the core plus inactive, active, and virtual subspaces. Preliminary calculations tested in the examples include RHF, ROHF, and CASSCF wave functions. The active spaces in the following perturbation calculations were often larger than the active space of fractionally occupied orbitals in the preliminary calculations.
(2) Based on some partitioning of the orbital space into core, inactive, and active, and virtual orbitals a CAS space of Slater determinants was formed from all possible arrangements of the active electrons among the active orbitals. Frozen core and inactive orbitals are doubly occupied in these determinants and virtual orbitals are empty. An initial zeroth-order wave function \( \psi_0 \) was formed by diagonalization of the CI matrix in this CAS space. From this wave function, a new spin-free charge density was formed and used to define a new \( F \) as in Eq. (14). Diagonal elements of this \( F \) in the canonical MO basis were defined to be the orbital energies

\[
e_k = \langle \phi_k | F | \phi_k \rangle. \tag{15}\]

In some calculations, the zeroth-order energy was defined as

\[
E_0 = \langle \psi_0^0 | \sum F(i) | \psi_0^0 \rangle. \tag{16}\]

The zeroth-order Hamiltonian was then defined as

\[
H_0 = \sum n_k e_k + \sum f(n_k - n_k^0), \tag{17}\]

\[
n_k = a^{1*}_{k\alpha} a_{k\alpha} + a^{1\alpha}_{k\delta} a_{k\delta}, \quad n^0_k = \langle \psi_0^0 | n_k | \psi_0^0 \rangle.
\]

Various correction functions \( f(n_k - n_k^0) \) were used to make this more like the ROHF OPT2 method of Murray and Davidson. Without this added term, \( H_0 \) would be the same as used by Hirao or Roos. Hirao also noted the need for a modified definition for high spin cases, but his modification is different from ours.

In addition to setting \( f \) to zero (the MROPT1 choice in the following tables), we have tried the following choices for \( f \) in the calculations in this paper (labeled MROPT2, MROPT3, MROPT4 in the tables):

\[
\langle \Phi_L | f_2(n_k) | \Phi_L \rangle = \begin{cases} \frac{1}{2} (n_k^0 - n_k) n_k^0 K_{kk}, & n^0_k > n_k^0 \\ \frac{1}{2} (n_k^0 - n_k) (2 - n_k) K_{kk}, & n_k^0 > n_k^0 \end{cases} \tag{18}\]

\[
\langle \Phi_L | f_3(n_k) | \Phi_L \rangle = \frac{1}{2} [n_k^0 - n_k^0] n_k^0 (2 - n_k) K_{kk} \tag{19}\]

\[
\langle \Phi_L | f_4(n_k) | \Phi_L \rangle = \frac{1}{2} [n_k^0 - n_k^0]^2 n_k^0 (2 - n_k) K_{kk}. \tag{20}\]

Here, \( K_{kk} \) is the self-exchange integral and \( \Phi_L \) is a Slater determinant.

Each of these choices for \( f \) gives corrections only for active orbitals (for which \( n_k^0 \) is not 0 or 2). Each choice gives the same result as the Murray and Davidson OPT2 for ROHF with one singly occupied orbital. For more open shells, this choice omits exchange between open shell orbitals, so it differs a little from OPT2. The \( f_2 \) choice gives an effective orbital energy equivalent to seeing \( N-1 \) electrons for decreasing the population in orbital \( k \) and \( N \) electrons for increasing the population. The \( f_3 \) choice has the same effect for \( n_k^0 = 0.1 \), but differs when \( n_k^0 \) is nonintegral. Similarly, \( f_4 \) agrees with \( f_2 \) for \( n_k^0 = 0.1 \) (because \( n_k^0 \) must be 0, 1, or 2) but differs for fractional values of \( n_k^0 \). Choices \( f_3 \) and \( f_4 \) are attempts to provide a smoother interpolation of \( f_2 \). The results from these three choices are quite similar for the tests in this paper.

Because our strategy will involve construction of an effective Hamiltonian in this CAS space and diagonalization to produce a new function \( \psi_0 \) expanded in this space, there is the possibility of iterating this procedure with \( \psi_0 \) replaced by \( \psi_0 \). In the few cases where \( \psi_0 \) and \( \psi_0 \) differ greatly, this would seem to be a logical procedure. One could carry this iteration all the way back to step 1 and use \( \psi_0 \) to define the preliminary wave function, but we have not tested this possibility, as it would involve repeating the transformation of the integrals from atomic orbitals to molecular orbitals.

It will be noted that neither \( \psi_0^0 \) nor \( \psi_0 \) are eigenfunctions of \( H_0 \) (unless all the active orbital energies are equal). Consequently, we will usually define

\[
E_0 = \langle \psi_0^0 | H_0 | \psi_0^0 \rangle, \tag{19}\]

\[
E_0 + E_1 = \langle \psi_0^0 | H | \psi_0^0 \rangle.
\]

As Huron et al.\textsuperscript{34} have noted, this choice of \( E_0 \) reduces size-inconsistency errors.

(3) The space of functions that interact with \( \psi_0 \) is spanned by the set of all single and double excitations from the CAS space. With our choice of \( H_0 \), the matrix of \( H_0 \) in this SD(CAS) space is diagonal. Hence, we follow Hira\textsuperscript{27} and expand \( \psi_j \) in this SD(CAS) space. This space is larger than the space of single and double excitation operators acting on \( \psi_0 \) used by Andersson et al.\textsuperscript{20} We prefer this larger space because (a) with our definition of \( H_0 \), \( (H_0 - E_0)^{-1} \) is easily formed in the space and (b) we assume \( \psi_0 \) is unknown and will only be defined after perturbation corrections.

We follow the philosophy of diagonalization after perturbation. This requires construction of a \( H_{\text{eff}} \) over the CAS space. Following Löwdin,\textsuperscript{42} the second-order estimate of \( H_{\text{eff}} \) is then

\[
(H_{\text{eff}})_{PQ} = \frac{\sum_K H_{KQ} E_{K} - E_0}{E_K - E_0} P, Q \in \text{CAS}. \tag{20}\]

Here \( K \) is a Slater determinant in the space SD(CAS) so it is an eigenfunction of \( H_0 \),

\[
H_0 \Phi_K = E_K \Phi_K, \quad E_K = \sum n_i^K e_i + \sum f(n_i^K - n_i^0). \tag{21}\]

Also, note that

\[
\langle \Phi_P | H_1 - E_1 | \Phi_K \rangle = 0 \tag{22}\]

because

\[
\langle \Phi_P | H_0 | \Phi_K \rangle = 0. \tag{23}\]

This \( H_{\text{eff}} \) can be diagonalized to define \( \psi_0 \) and \( E_{\text{eff}} \) correct through second order

\[
H_{\text{eff}} \psi_0 = E_{\text{eff}} \psi_0. \tag{24}\]

Alternatively, the expectation value

\[
\bar{E} = \langle \psi_0 | H_{\text{eff}} | \psi_0 \rangle \tag{25}\]
should correspond to the usual energy obtained by diagonalization before perturbation since this gives
\[ E_2 = \sum_{K} \frac{\langle \Phi_K | H | \Phi_K \rangle^2}{E_K^2 - E_0^2}, \quad E = E_0 + E_1 + E_2. \] (26)

In implementing this scheme, we have assumed that the SD(CAS) space will normally be very large for applications to real problems in chemistry (i.e., > 10^6 functions) and the \( H_{PK} \) matrix connecting the CAS and SD(CAS) spaces will be sparse. Hence, we have organized the program so that the outermost loop runs sequentially through the following possible categories of \( K \).

1. inactive \( \rightarrow \) active
   A. \( a \rightarrow a \)
   B. \( b \rightarrow b \)
2. inactive \( \rightarrow \) virtual
   A. \( a \rightarrow a \)
   B. \( b \rightarrow b \)
3. active \( \rightarrow \) virtual
   A. \( a \rightarrow a \)
   B. \( b \rightarrow b \)
4. inactive \( \rightarrow \) inactive', active,
   A. \( a \rightarrow a \)
   B. \( b \rightarrow b \)
   C. \( \alpha \rightarrow \alpha \)
   D. \( \beta \rightarrow \beta \)
   E. \( \alpha \rightarrow \beta \)
   F. \( \beta \rightarrow \alpha \)
5. inactive \( \rightarrow \) inactive', active,
   A. \( a \rightarrow a \)
   B. \( b \rightarrow b \)
   C. \( a \rightarrow b \)
   D. \( b \rightarrow a \)
   E. \( \alpha \rightarrow a \)
   F. \( \beta \rightarrow b \)
6. inactive \( \rightarrow \) inactive', virtual, virtual',
   A. \( a \rightarrow a \)
   B. \( b \rightarrow b \)
   C. \( a \rightarrow b \)
   D. \( b \rightarrow a \)
   E. \( \alpha \rightarrow a \)
   F. \( \beta \rightarrow b \)
7. active, active' \( \rightarrow \) virtual, virtual',
   A. \( a \rightarrow a \)
   B. \( b \rightarrow b \)
   C. \( a \rightarrow b \)
   D. \( b \rightarrow a \)
   E. \( \alpha \rightarrow a \)
   F. \( \beta \rightarrow b \)

The next loops run over virtual and inactive orbitals. Inside of that, loops run over pretabulated distributions of electrons among active orbitals. If there are \( N_a \) spin-up electrons and \( N_b \) spin-down electrons in the active space in the CAS, then these distributions are needed for \( N_a, N_a \pm 1, N_b, N_b \pm 1 \), and \( N_b \pm 2 \). This loop structure generates the SD(CAS) configurations in a convenient order. Inside of this there are loops that apply appropriate single and double (de)excitation operators to a given \( \Phi_K \) to generate connected configurations in the CAS space.

The program can doubtless be improved. At present, it makes use of three additional features to simplify the calculation.

1. The integrals are presorted so that when a given \( K \) is reached, all of the integrals required for \( H_{PK} \) for all \( P \) in the CAS are contained in the next contiguous set of integrals encountered on the integral file.
2. A CAS is used with the configurations in dictionary order. Hence when a member, \( \Phi_P \), of the CAS is generated by acting on \( \Phi_K \) with (de)excitation operators, the index \( P \) can be computed. The calculation of \( P \) involves only addition of pretabulated binomial coefficients. If an importance selected incomplete reference space were used, then it would be necessary to search the configuration list to determine \( P \).
3. Most of the SD(CAS) configurations are in category 6 above (inactive \( \rightarrow \) virtual). For these, only one \( \Phi_P \) connects to each \( \Phi_K \) so there are only contributions to the diagonal of \( H_{PP} \). The contribution from all spin combinations for a given set of space orbitals is easily summed so that the contribution to \( H_{PP} \) from \( cc' \rightarrow vv' \) is just
\[ \sum \left( \frac{\langle cc' | vv' \rangle^2 - \langle cc' | vv' \rangle \langle cc' | vv' \rangle}{\epsilon_v + \epsilon_v - \epsilon_c - \epsilon_c + E_P^2 - E_0^2} \right) \] (27)

Thus, the predominant category requires little effort.

It will be noted that this method loses many of the attributes of closed shell MP2 theory. Neither \( \psi_0 \) nor \( \psi_0 \) are eigenfunctions of \( H_0 \). The energy is not invariant to orbital rotation if a new \( H_0 \) is adopted that is again diagonal after rotation. The energy denominators can be adjusted to be like MP2 in the use of electron affinities when electrons are added and ionization energies when electrons are removed. Thus the feasibility and accuracy are expected to be comparable to MP2. The wave functions will be eigenfunctions of \( S^2 \) and \( S_z \) and will belong to irreducible representations of the space group as long as each subspace (active, virtual, etc.) of molecular orbitals is closed under the operators in the space group.

The most important difficulty is the lack of size consistency. This shows itself in several ways. As Huron et al.34 noted, the excitation spectra of one molecule in a noninteracting diffuse gas of other molecules should not be affected by their presence. If the active space were just the orbitals involved in the excitation then the effect of corelation in the other molecules would be given by Eq. (27). The numerator in this is independent of \( P \). In the calculation of \( E_2, \) a \( cc' \rightarrow vv' \) would involve the sum over the coefficients \( C_P^0 \) in \( \psi_0 \)
\[ \sum \left( C_P^0 \right)^2 \frac{1}{\Delta \epsilon + E_P^2 - E_0} = \frac{1}{\Delta \epsilon} \frac{1}{\Delta \epsilon} \sum \left( C_P^0 \right)^2 (E_P^2 - E_0) + \cdots \] (28)

If we choose \( E_0 = \sum (C_P^0)^2 E_P^2 \), the second term vanishes but there are still contributions from higher terms.

For the quasidegenerate case, it is conceivable to replace all orbital energies in the active space by their weighted average, as was done by Brandon50 in his quasidegenerate perturbation theory. Then \( E_P^2 = E_0 \) and \( \psi_0 \) is an eigenfunction of \( H_0 \). This makes the computed exci-
tation spectrum exactly independent of spectator molecules. It has the disadvantage, however, that the average orbital energy could change drastically if the active space is changed to include more orbitals. The recent perturbation theory for excited states introduced by Foresman et al. uses \( E_0^2 + \Delta \) for \( E_0 \), where \( E_0^2 \) is the ground state sum of orbital energies, but \( \Delta \) is the Cl excitation energy. For the reasons noted here, this method is not size consistent.

Another way that size consistency is violated is that the total energy of well-separated monomers is not the sum of their energies calculated separately. For identical monomers, this could be fixed by the same trick of replacing the orbital energies of all active orbitals by their average. This assumes that equivalent active orbital sets are used for each monomer and the active set for the diffuse gas is the union of these monomer sets. For nonidentical monomers, this will not work since the average orbital energy of the active orbitals of one monomer will differ from the average over all monomers.

These difficulties with size consistency can also be exactly avoided by adopting a modified \( H_{\text{eff}} \) intermediate between the Löwdin and Bloch approaches. This is suggested by Eq. (28) which shows that Eq. (27) becomes exactly size consistent if the division by the denominator in Eq. (28) is simply replaced by multiplication by

\[ \frac{1}{\Delta \epsilon} = \frac{\Delta \epsilon - E_0^2 + E_0}{\Delta \epsilon^2}. \]  

This expression can be derived by considering “dressed” reference states

\[ \Phi_{\text{sd}} = \Phi_F - \sum K \langle \Phi_K \rangle^{-1} H_{PK} \Phi_K. \]  

Notice that the energy denominator in the dressed state is appropriate for each \( \Phi_F \) being \( \Phi_0^2 \) for a different state of interest. Expansion of \( \Phi_0 \) in this set of dressed states leads to a nonorthogonal eigenvalue problem. Truncating \( \langle \Phi_F \mid [H_0^2] \rangle \) and \( \langle \Phi_F \mid \Phi_0 \rangle \) to second order and collecting all second order pieces of the secular equation into \( H_{\text{eff}} \) gives

\[ (H_{\text{eff}})_{PO} = HPO - \sum K \langle \Phi_K \rangle^{-1} H_{PK} \langle \Phi_0 \rangle, \]  

\[ P, Q \in \text{CAS}. \]  

For \( E_0 \) defined by Eq. (19), the energy, \( E_i \), given by the expectation value, Eq. (25), of this \( H_{\text{eff}} \) is easily shown to be size consistent. This formula is close to the second-order Van Vleck equation

\[ (H_{\text{eff}})_{PO} = HPO - \sum \langle \Phi_K \rangle^{-1} H_{PK} \langle \Phi_0 \rangle. \]  

The Van Vleck energy eigenvalue, as well as energy expectation value, is size consistent because the individual matrix elements are size-consistent.

### III. NUMERICAL ANALYSIS

In this section, we present a series of preliminary calculations performed using different forms of multireference second-order perturbation theory. First, we show a study of the size-consistency problem for two model systems. We then apply perturbation theory demonstrating applications to different problems, including the potential energy surface for the nitrogen molecule in its ground state, electron affinity of cyano radical, vertical excitations in formaldehyde, and singlet-triplet separation in \( \text{CH}_2 \) radical. Most of the results presented here must be characterized as numerical analysis, since a limited basis set was used to make comparison with full CI results possible. To obtain the above results, the current version of the program was interfaced with the MELD suite of programs to obtain a RHF or ROHF wave function as the starting point for perturbation theory. The CASCCF wave function was obtained using HONDO 8.4 and the vectors were transformed to a form suitable for MELD. For convenience of presentation, we introduce the following notations for different forms of perturbation theory: \( [\text{Ne}^-/\text{Morb}] \) describes a CAS model space with \( N \) electrons distributed between \( M \) active orbitals. To distinguish between different choices of \( \sum \), we refer to the generalization of OPT1 as MROPT1 while the generalizations of OPT2 to multireference as given by Eq. (18) are denoted by MROPT2, MROPT3, and MROPT4. Barycentric \( E_0 \) is defined according to Eq. (19).

#### A. Size-consistency test

In the first set of test calculations, we analyzed the size-consistency problem of different forms of second-order multireference perturbation theory. We studied two model systems: the first model contains two beryllium atoms separated by 400 a.u.; in the second model, one beryllium atom was replaced by boron positive ion. In both cases we assumed that the interaction is small and can be neglected. A split valence double-zeta basis set was used to make comparison with full CI results. Results for each monomer are summarized in Table I. For both monomers, the active space was chosen as a \([2s, 2p_\sigma, 2p_\pi, 2p_\rho]\) set of canonical (occupied and virtual) orbitals from a single configuration SCF wave function, while for dimers the active space was the superposition of the active spaces of the two beryllium atoms or the beryllium and boron ion. In Table II, we present the size-consistency error, i.e., \( \langle AB \rangle - \langle A + B \rangle \) for each case considered.

According to Table I, the scheme with average orbital energies, as well as the Van Vleck form are in better agreement with full CI than other forms of perturbation theories. MROPT1 and MROPT2 also seem to be better choices of \( \sum \) than MROPT3 or MROPT4. Table II shows that not all forms of multireference perturbation theory presented here are size consistent. The option with equal orbital energies in the active space is only size consistent for the homogenous system. The expectation value of the effective Hamiltonian is size consistent for the scheme given by Eq. (31) intermediate between Löwdin and Bloch. After iteration, the energy \( E_{\text{iter}} \) is the same whether
TABLE I. Energy of Be and B+ atoms. SCF energy for Be = -14.570 906, SCF energy for B+ = -24.233 384.

<table>
<thead>
<tr>
<th>Method</th>
<th>Be</th>
<th>B+</th>
</tr>
</thead>
<tbody>
<tr>
<td>MROPT1a</td>
<td>-14.605 855</td>
<td>-24.286 284</td>
</tr>
<tr>
<td>MROPT2a</td>
<td>-14.605 065</td>
<td>-24.285 435</td>
</tr>
<tr>
<td>MROPT3a</td>
<td>-14.604 841</td>
<td>-24.285 250</td>
</tr>
<tr>
<td>MROPT4a</td>
<td>-14.604 108</td>
<td>-24.284 584</td>
</tr>
<tr>
<td>Averageb</td>
<td>-14.613 721</td>
<td>-24.293 680</td>
</tr>
<tr>
<td>MROPTc</td>
<td>-14.605 855</td>
<td>-24.286 284</td>
</tr>
<tr>
<td>MROPT2c</td>
<td>-14.605 909</td>
<td>-24.286 360</td>
</tr>
<tr>
<td>MROPT4c</td>
<td>-14.604 194</td>
<td>-24.284 703</td>
</tr>
<tr>
<td>MROPTd</td>
<td>-14.598 713</td>
<td>-24.278 550</td>
</tr>
<tr>
<td>MROPT2d</td>
<td>-14.599 038</td>
<td>-24.278 475</td>
</tr>
<tr>
<td>MROPT3d</td>
<td>-14.597 928</td>
<td>-24.276 334</td>
</tr>
<tr>
<td>MROPT4d</td>
<td>-14.596 614</td>
<td>-24.273 053</td>
</tr>
<tr>
<td>MROPTe</td>
<td>-14.614 827</td>
<td>-24.295 844</td>
</tr>
<tr>
<td>MROPT2e</td>
<td>-14.615 224</td>
<td>-24.296 325</td>
</tr>
<tr>
<td>MROPT3e</td>
<td>-14.614 747</td>
<td>-24.295 901</td>
</tr>
<tr>
<td>MROPT4e</td>
<td>-14.615 382</td>
<td>-24.297 293</td>
</tr>
<tr>
<td>Full Cl</td>
<td>-14.615 569</td>
<td>-24.291 596</td>
</tr>
</tbody>
</table>

*E0 defined by Eq. (16), H$_{ef}$ defined by Eq. (20).
*Active orbital energies are equal, replaced by their $\bar{\psi}$ population weighted average, H$_{ef}$ defined by Eq. (20), E$_0$ defined by Eq. (16).
*Barycentric definition of E$_0$, Eq. (19), H$_{ef}$ defined by Eq. (20).
*Intermediate scheme between Löwdin and Bloch. E$_0$ defined by Eq. (31), barycentric E$_0$.
*Van Vleck form of perturbation theory, Eq. (32).

For perturbation theory, with H$_0$ defined according to MROPT1-4, E$_0$ defined by Eq. (16) and E defined by Eq. (24). Six electrons were distributed among the six 2p orbitals (3 occupied and 3 virtual from the SCF) to form the active space. Because the nitrogen potential surface represents a multireference problem, the CASSCF wave function would be a more logically correct starting point for perturbation theory. Results of calculations are displayed in Fig. 1. Upon examining these energy curves, one can conclude that, regardless of which option was used to define H$_0$, the overall shape remains in very good agreement with the correct potential energy surface.

In spite of starting from RHF orbitals, the absolute energies are also fairly good. At 2.068 Å, MROPT1 gives -109.289, compared with the best estimate of -109.3 1. At 6.0 Å, MROPT1 gives -108.975, compared with the best estimate of -108.99, in spite of the very poor SCF energy.

B. Nitrogen molecule potential energy curve

The second set of test calculations were performed for the ground state of the nitrogen molecule as a function of internuclear separation. Determination of the correct potential energy surface for multiple bonded molecules represents a very difficult problem and most standard ab initio techniques such as configuration interaction methods, many-body perturbation theory, and coupled cluster methods are not adequate for correct description, as was demonstrated by Laidig et al. The potential energy curve for the nitrogen molecule was also analyzed by Andersson et al. and by Hirao using multireference Møller-Plesset perturbation theory as well as by Murphy and Messmer. In our calculations we used the same basis set as Laidig et al., a Huzinaga-Dunning double-zeta (9s5p/4s2p) set with the normal 4,1p contraction replaced by a 3,1,1 contraction and augmented with six Cartesian d polarization functions of exponent 0.900. We used the RHF canonical occupied and virtual orbitals as the starting point for perturbation theory, with H$_0$ defined according to MROPT1-4, E$_0$ defined by Eq. (16) and E defined by Eq. (24). Six electrons were distributed among the six 2p orbitals (3 occupied and 3 virtual from the SCF) to form the active space.
TABLE II. Size-consistency error between dimers and sum of monomers in $mE_a$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$E_{sa}$</th>
<th>$C_0^2 H_{sa} C_0$</th>
<th>$E_{ext}$</th>
<th>$E_{sa}$</th>
<th>$C_0^2 H_{sa} C_0$</th>
<th>$E_{ext}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MROPT1**</td>
<td>0.364</td>
<td>0.457</td>
<td>0.461</td>
<td>0.725</td>
<td>0.781</td>
<td>0.187</td>
</tr>
<tr>
<td>MROPT2</td>
<td>0.494</td>
<td>0.576</td>
<td>0.640</td>
<td>1.462</td>
<td>1.387</td>
<td>1.517</td>
</tr>
<tr>
<td>MROPT3</td>
<td>1.270</td>
<td>1.340</td>
<td>1.533</td>
<td>0.379</td>
<td>0.448</td>
<td>0.457</td>
</tr>
<tr>
<td>MROPT4</td>
<td>2.373</td>
<td>2.423</td>
<td>2.765</td>
<td>0.451</td>
<td>0.803</td>
<td>0.528</td>
</tr>
<tr>
<td>Average**</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.129</td>
<td>1.147</td>
<td>1.152</td>
</tr>
<tr>
<td>MROPT1b*</td>
<td>0.364</td>
<td>0.457</td>
<td>0.461</td>
<td>0.725</td>
<td>0.781</td>
<td>0.187</td>
</tr>
<tr>
<td>MROPT2</td>
<td>0.562</td>
<td>0.655</td>
<td>0.687</td>
<td>0.731</td>
<td>0.789</td>
<td>0.180</td>
</tr>
<tr>
<td>MROPT3</td>
<td>1.287</td>
<td>1.356</td>
<td>1.546</td>
<td>0.747</td>
<td>0.795</td>
<td>0.847</td>
</tr>
<tr>
<td>MROPT4</td>
<td>2.355</td>
<td>2.322</td>
<td>2.741</td>
<td>0.772</td>
<td>0.816</td>
<td>0.878</td>
</tr>
<tr>
<td>MROPT1b*</td>
<td>0.180</td>
<td>0.000</td>
<td>0.116</td>
<td>1.383</td>
<td>0.000</td>
<td>0.213</td>
</tr>
<tr>
<td>MROPT2</td>
<td>0.731</td>
<td>0.992</td>
<td>0.055</td>
<td>1.408</td>
<td>0.000</td>
<td>0.229</td>
</tr>
<tr>
<td>MROPT3</td>
<td>2.163</td>
<td>1.184</td>
<td>0.616</td>
<td>1.563</td>
<td>0.000</td>
<td>0.326</td>
</tr>
<tr>
<td>MROPT4</td>
<td>4.027</td>
<td>4.028</td>
<td>1.446</td>
<td>2.380</td>
<td>0.000</td>
<td>0.464</td>
</tr>
<tr>
<td>MROPT1b*</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MROPT2</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MROPT3</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MROPT4</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MROPT1b*</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>MROPT2</td>
<td>0.366</td>
<td>0.213</td>
<td>0.730</td>
<td>0.093</td>
<td>0.337</td>
<td>0.004</td>
</tr>
<tr>
<td>MROPT3</td>
<td>0.093</td>
<td>0.337</td>
<td>0.004</td>
<td>0.336</td>
<td>0.407</td>
<td>0.805</td>
</tr>
<tr>
<td>MROPT4</td>
<td>0.336</td>
<td>0.407</td>
<td>0.805</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

*a* $E_0$ defined by Eq. (16).
*b* Delocalized canonical orbitals for Be*.
*c* Löwdin perturbation theory, Eq. (20).
*d* Active orbital energies are replaced by population weighted average.

The size-consistency error between dimers and sum of monomers in $mE_a$ converges faster to the exact energy than OPT1, although OPT1 gives a better estimation of energy in second order. Results of our calculations are summarized in Table III. In these calculations we tried different model spaces as well as different choices of $H^0$. Upon comparing the numerical values with the exact ones, we conclude that a model space where five electrons are distributed between six canonical ROHF orbitals is an adequate size for all geometries considered. Among all results presented, MROPT1 and MROPT2 remain in better agreement with the exact values of the correlation energy, than the MROPT3 and MROPT4 options. On the other hand, the barycentric choice of $E_0$ seems to be better, particularly for the stretched geometries.

D. Electron affinity of cyano radical

The electron affinity (EA) of the cyano radical and convergence of the Möller–Plesset perturbation series (spin unrestricted and spin restricted) has been extensively studied theoretically by Nobes *et al.* The estimated value of the EA is basically in poor agreement with experiment, even when large basis sets are used. This is due to the rather slow UMP convergence for the radical. In this section we will discuss this problem with use of multireference perturbation theory. We use the STO-3G basis set to make comparison with full CI result. We use the same geometry as in Ref. 64, i.e., $r_{CN}$ equals 1.1607 Å for CN⁻ and $r_{CN}$

![Image of a graph showing the potential energy in Hartrees vs bond length in bohrs.](J. Chem. Phys., Vol. 100, No. 5, 1 March 1994)
TABLE III. Correlation energy of the NH2 radical with a 6-31G basis set.

<table>
<thead>
<tr>
<th>Reference</th>
<th>MROPT1</th>
<th>MROPT2</th>
<th>MROPT3</th>
<th>MROPT4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1e^-/1orb$</td>
<td>-0.089 631</td>
<td>-0.087 331</td>
<td>-0.087 331</td>
<td>-0.087 331</td>
</tr>
<tr>
<td>$5e^-/5orb$</td>
<td>-0.091 405</td>
<td>-0.088 908</td>
<td>-0.088 908</td>
<td>-0.088 908</td>
</tr>
<tr>
<td>$5e^-/6orb$</td>
<td>-0.091 838</td>
<td>-0.088 286</td>
<td>-0.088 412</td>
<td>-0.089 363</td>
</tr>
<tr>
<td>Exact</td>
<td>-0.103 087</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$R = 1.0130 \, \text{Å}, \, \alpha = 103.2^\circ$ (ROHF = -55.530 177, full Cl$^b$ = -55.633 264)

$R = 1.5195 \, \text{Å}, \, \alpha = 103.2^\circ$ (ROHF = -55.367 730, full Cl$^b$ = -55.526 661)

$R = 2.0260 \, \text{Å}, \, \alpha = 103.2^\circ$ (ROHF = -55.181 593, full Cl$^b$ = -55.441 568)

$1e^-/1orb$ | -0.125 615 | -0.122 514 | -0.122 514 | -0.122 514 |
| $5e^-/5orb$ | -0.148 066 | -0.140 977 | -0.142 279 | -0.141 975 |
| $5e^-/6orb$ | -0.150 064 | -0.143 134 | -0.144 611 | -0.144 513 |
| Exact | -0.158 931 | | | |

Starting point for perturbation theory was the ROHF wave function and canonical ROHF orbitals, $H_\text{er}$ defined by Eq. (20) and $E$ chosen as $E_\text{er}$, Eq. (24) not iterated. The six orbital reference space contained two $a_1$, three $b_2$, and one $b_1$ orbitals.

FULL Cl with frozen core.

$H_0$ defined by Eqs. (17) and (18).

$E_\text{er}$ defined by Eq. (16).

Barycentric definition of $E_\text{er}$, Eq. (19).

TABLE IV. Total energies$^a$ (in $E_b$) and electron affinity of CN radical calculated with the STO-3G basis set.

<table>
<thead>
<tr>
<th>Method</th>
<th>CN$^-$</th>
<th>CN$^-$</th>
<th>EA (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCF</td>
<td>-90.937 66</td>
<td>-90.997 52</td>
<td>-157</td>
</tr>
<tr>
<td>MP2</td>
<td>-91.071 43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPT1</td>
<td>-91.153 46</td>
<td>[1e$^-$/1orb]</td>
<td>-234</td>
</tr>
<tr>
<td>OPT2</td>
<td>-91.153 46</td>
<td>[1e$^-$/1orb]</td>
<td>-215</td>
</tr>
<tr>
<td>MROPT1</td>
<td>-91.153 46</td>
<td>[2e$^-$/2orb]</td>
<td>-245</td>
</tr>
<tr>
<td>MROPT4</td>
<td>-91.153 46</td>
<td>[2e$^-$/2orb]</td>
<td>-240</td>
</tr>
<tr>
<td>MROPT4</td>
<td>-91.153 46</td>
<td>[2e$^-$/2orb]</td>
<td>-240</td>
</tr>
<tr>
<td>MROPT4</td>
<td>-91.153 46</td>
<td>[2e$^-$/2orb]</td>
<td>-240</td>
</tr>
</tbody>
</table>

$^a$Starting point for perturbation theory was the ROHF wave function and canonical ROHF orbitals, $H_\text{er}$ defined by Eq. (20) and $E$ chosen as $E_\text{er}$, Eq. (24) not iterated.

$^b$Full Cl with frozen core.

E. Vertical excitation energies for formaldehyde

We have calculated the vertical excitation energies for formaldehyde using the same geometry as Foresman et al. (i.e., CO 1.2122 Å, CH 1.044 Å, HCO 121.94°). We used also the same basis set, i.e., augment the 6-31+G* basis set with an additional diffuse sp shell with the exponent 0.02, called 6-31+G*R. Numerical values of calculations for the first few transition energies are summarized in Table V. Two different approaches were used. In the first approach, we performed ROHF calculations for each excited state separately, and then applied perturbation...
TABLE V. Vertical excitation energies for formaldehyde calculated using different form of perturbation theory.  

<table>
<thead>
<tr>
<th>Transition</th>
<th>$E_{\text{ref}}$</th>
<th>MROPT1</th>
<th>MROPT2</th>
<th>MROPT3</th>
<th>MROPT4</th>
<th>Expt.</th>
<th>Ref. 33</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^3A_2$ $n-\pi^*$</td>
<td>3.15</td>
<td>3.56</td>
<td>3.92</td>
<td>3.48</td>
<td>3.85</td>
<td>3.50</td>
<td>4.163</td>
</tr>
<tr>
<td>$^1A_2$ $n-\pi^*$</td>
<td>3.48</td>
<td>4.04</td>
<td>4.37</td>
<td>3.92</td>
<td>4.27</td>
<td>4.10</td>
<td>4.580</td>
</tr>
<tr>
<td>$^3A_1$ $\pi-\pi^*$</td>
<td>5.38</td>
<td>5.97</td>
<td>6.70</td>
<td>6.09</td>
<td>6.74</td>
<td>6.00</td>
<td>6.812</td>
</tr>
<tr>
<td>$^3B_2$ $n-3s$</td>
<td>6.34</td>
<td>7.09</td>
<td>7.32</td>
<td>7.15</td>
<td>7.41</td>
<td>7.09</td>
<td>6.792</td>
</tr>
<tr>
<td>$^1B_1$ $n-3s$</td>
<td>6.38</td>
<td>7.19</td>
<td>7.45</td>
<td>7.26</td>
<td>7.51</td>
<td>7.13</td>
<td>6.695</td>
</tr>
</tbody>
</table>

Using SCI' orbitals for each state

<table>
<thead>
<tr>
<th>Transition</th>
<th>$E_{\text{ref}}$</th>
<th>MROPT1</th>
<th>MROPT2</th>
<th>MROPT3</th>
<th>MROPT4</th>
<th>Expt.</th>
<th>Ref. 33</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^3A_2$ $n-\pi^*$</td>
<td>5.77</td>
<td>2.52</td>
<td>3.34</td>
<td>2.39</td>
<td>3.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^1A_2$ $n-\pi^*$</td>
<td>6.11</td>
<td>3.14</td>
<td>3.89</td>
<td>4.69</td>
<td>3.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^3A_1$ $\pi-\pi^*$</td>
<td>7.65</td>
<td>4.93</td>
<td>5.90</td>
<td>4.79</td>
<td>6.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^3B_2$ $n-3s$</td>
<td>8.87</td>
<td>6.82</td>
<td>7.34</td>
<td>6.92</td>
<td>7.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^1B_1$ $n-3s$</td>
<td>8.89</td>
<td>6.90</td>
<td>7.40</td>
<td>7.00</td>
<td>7.42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using IVO orbitals from ground state

<table>
<thead>
<tr>
<th>Transition</th>
<th>$E_{\text{ref}}$</th>
<th>MROPT1</th>
<th>MROPT2</th>
<th>MROPT3</th>
<th>MROPT4</th>
<th>Expt.</th>
<th>Ref. 33</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^3A_2$ $n-\pi^*$</td>
<td>3.92</td>
<td>3.48</td>
<td>4.37</td>
<td>3.92</td>
<td>4.27</td>
<td>4.10</td>
<td>4.580</td>
</tr>
<tr>
<td>$^1A_2$ $n-\pi^*$</td>
<td>4.37</td>
<td>3.92</td>
<td>6.70</td>
<td>6.09</td>
<td>6.74</td>
<td>6.00</td>
<td>6.812</td>
</tr>
<tr>
<td>$^3A_1$ $\pi-\pi^*$</td>
<td>6.34</td>
<td>7.09</td>
<td>7.32</td>
<td>7.15</td>
<td>7.41</td>
<td>7.09</td>
<td>6.792</td>
</tr>
<tr>
<td>$^3B_2$ $n-3s$</td>
<td>6.38</td>
<td>7.19</td>
<td>7.45</td>
<td>7.26</td>
<td>7.51</td>
<td>7.13</td>
<td>6.695</td>
</tr>
</tbody>
</table>

Reference space (n,\pi,\pi^*,3s,3p_1,3p_2,3p_3) with four electrons. Energy given by $E_{\text{ref}}$, Eq. (24), barycentric definition of $E_0$, Eq. (19). In the first half of the table, canonical orbitals were separately defined for each state. In the second half, a common set of IVOs were used.

bEnergy obtained after diagonalization of model space.

cLöwdin $H_{\text{ef}}$, Eq. (20).

dIntermediate scheme between Löwdin and Bloch, Eq. (31).

eReference space was CI with all single excitations from ground state. Second order perturbation theory as defined in Ref. 33.

F. Singlet–triplet splitting in CH₂ radical

As a final test, we have studied the singlet-triplet splitting in the methylene radical. The basis set is the DZP basis defined by Bauschlicher and Taylor. Results of calculations can be compared to the full CI results of Bauschlicher and Taylor and CAS second-order perturbation calculations of Anderson, Malmqvist and Roos as

TABLE VI. The singlet–triplet separation in CH₂ radical.  

<table>
<thead>
<tr>
<th>Methodb,c</th>
<th>$E(1 \alpha_1)$</th>
<th>$E(1 \beta_1)$</th>
<th>$\Delta$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCF</td>
<td>-38.868 297</td>
<td>38.927 947</td>
<td>26.14</td>
</tr>
<tr>
<td>(1010) 2×2 CASSCF</td>
<td>-38.907 660</td>
<td>-38.927 947</td>
<td>12.73</td>
</tr>
<tr>
<td>(3210) 6×6 CASSCF</td>
<td>-38.954 529</td>
<td>-38.965 954</td>
<td>12.82</td>
</tr>
<tr>
<td>(4220) 6×8 CASSCF</td>
<td>-39.968 726</td>
<td>-39.982 741</td>
<td>8.79</td>
</tr>
</tbody>
</table>

aThe geometry and basis set were taken from Ref. 66.

bBarycentric definition of $E_0$, Eq. (19), $H_{\text{ef}}$ from Eq. (20), $E$ from $E_{\text{ef}}$, Eq. (24).

bThe numbers within parentheses are the number of active orbitals used in the CASSCF and multireference perturbation calculations given in symmetry order a, b, c, and a₂.

dFull CI with frozen core.
as those of Hirao. Three sets of calculations were performed with different CAS(CI) levels of approximation and results are summarized in Table VI. The SCF separation in is error by 14.17 kcal/mol, relative to the full CI, while the two-configurational treatment of the singlet state reduces the error to 0.76 kcal/mol. One would expect that the result should only change slightly after perturbation, however, results of calculations indicate different trends. This is due to the fact that the two states \(1_A1\) and \(2_B1\) are treated in an unbalanced way at the correlated level of calculation. Extension of the size of the active space does not reduce the error. After application of perturbation theory, results strongly depend on the choice of \(H_0\). The scheme based on the MROPT1 definition of \(H_0\) seems to be the worst one, while MROPT2 is the best one. Although the extension of the size of the model space does not reduce the error in the singlet-triplet gap, perturbation corrections systematically reduce the error. Calculations with the third active space, particularly with MROPT2, are in very good agreement with full CI, being different by 0.21 kcal/mol for the singlet-triplet gap, even though the singlet and triplet energies are separately still in error by over 6 kcal/mol.

IV. CONCLUSIONS

The Van Vleck formula for \(H_{\text{eff}}\) is not suitable because it often has instabilities caused by intruder states. Both the Löwdin form, Eq. (20), and the modified form, Eq. (31), work fairly well. The barycentric definition of \(E_0\), Eq. (19), is usually better than Eq. (16). The MROPT1 and MROPT2 definitions of \(H_0\) work about equally well. The best choice is not always the same one in this set of calculations. The energies given by \(E_{\text{eff}}\), Eq. (24), \(\tilde{E}\), Eq. (25) are of comparable accuracy in most cases. If \(\psi_0^0\) and \(\psi_0\) differ greatly, \(E_{\text{eff}}\) is more accurate than \(\tilde{E}\).

ACKNOWLEDGMENT

This work was supported by Grant No. CHE 90 07393 from the National Science Foundation.