Applications of multiconfigurational coupled-cluster theory

Ajit Banerjee and Jack Simons[a]

Chemistry Department, University of Utah, Salt Lake City, Utah 84112

A coupled-cluster method which permits the use of multiconfiguration reference states has recently been developed in this laboratory. In the present work, it is applied to several states of H_2 ($^1\Sigma_g^+$), Li(^2S), HeH_2($^1A_1$), and CH_2($^3B_1$,$^1A_1$), which include both open and closed shells. These applications are made within an approximation in which the cluster operator (T) is truncated at T_2, $T \simeq T_1 + T_2$ and the expansion of $e^{-T}He^T$ is truncated at the double-commutator level. For cases where a single configuration function ceases to be a good starting point, it is found that a single configuration based truncated coupled-cluster procedure may exhibit serious difficulties. In such cases we find it possible to choose a multiconfigurational reference state for which our coupled-cluster procedure converges reasonably rapidly. This paper contains several illustrations of such convergence characteristics.

I. INTRODUCTION

In recent years, it has become ever more common to apply the methods of molecular quantum mechanics to numerous chemical phenomena in which knowledge of the potential energy surface at the equilibrium and distorted geometries of the molecule is essential. So-called size-consistent (SC)[1,2] methods are especially desirable in such applications. Rayleigh–Schrödinger or many-body perturbation theories and certain coupled-cluster (CC) methods are examples of SC theories. CC theory based on the exponential operator [exp(T)] ansatz has been developed using a Hartree–Fock (HF) single determinant reference function for closed-shell systems.[3-5]

Recently, we developed a CC method[6] which admits a multiconfiguration (MC) reference wave function $\Phi = \sum_K C_K \chi_K$, where the $\{\chi_K\}$ label the configurations. In this method, closed- and open-shell problems can be solved using the same formalism. The theory is cast in terms of the generators of the unitary group, which leads to fast computation of requisite matrix elements. Although CC methods for open-shell systems have previously been discussed[7,8] and extensively used in a single determinant framework[12] by others, the results presented here represent the first *applications* of an open- and closed-shell theory which employ a multiconfigurational self-consistent field (MCSCF) reference function.

To understand why we feel that it is desirable to have available a MCSCF-based CC method, consider the simple reaction Be(^1S) + H_2($^1\Sigma_g$) → Be H_2 ($^1A_1$), which we assume to occur via a concerted bond insertion step which preserves C_{2v} symmetry. The dominant configuration of the reactants is $1a_1^2 2a_1^2 3a_1^2$, where $1a_1 = 1s_{Be}$, $2a_1 = \sigma_{g,H_2}$, and $3a_1 = 2s_{Be}$. For the products BeH_2, the dominant configuration is $1a_1^2 2a_1^2 1b_2^2$, where $1a_1 \cong 1s_{Be}$, $2a_1 \cong \sigma_{BeH}$, and $1b_2 \cong \sigma_{BeH}$. Clearly, there will be an avoided configuration crossing along the reaction path of the ground state $^1A_1$ surface being considered here.

Now consider how a single configuration based CC

wave function attempts to describe the $^1A_1$ wave function along the reaction path (RP). If the $1a_1^2 2a_1^2 3a_1^2$ configuration is chosen to be the reference function $|\Phi\rangle$ on which exp(T) operates to yield the correct wave function $|^1A_1\rangle = \exp(T) |\Phi\rangle$, we expect the amplitudes occuring in the T operator (see later) to be small when we are on the reactant side of the RP. As we approach the region of the avoided configuration crossing, we expect amplitudes in T, which describe double excitations (T_2) out of $|\Phi\rangle$, to approach unity in magnitude [so that $\exp(T) |\Phi\rangle \cong |\Phi\rangle + T |\Phi\rangle + \cdots$ will contain both of the above configurations with large amplitude]. Also, if the orthonormal basis orbitals are obtained from a single configuration SCF calculation on the $1a_1^2 2a_1^2 3a_1^2$ reference determinant, we expect amplitudes in T, which describe single excitations (T_1) out of $|\Phi\rangle$, to be relatively large (~ 0.1) (because they describe orbital relaxation effects which should be significant when one attempts to use SCF orbitals of $|\Phi\rangle$ in situations where $|^1A_1\rangle$ contains two dominant configurations). To implement a single configuration CC theory, we therefore need to be able to deal with these large T amplitudes.

Actually, the situation becomes even more difficult as we move along the RP toward the product (BeH_2) region. Here, $|^1A_1\rangle$ is *dominated* by the $1a_1^2 2a_1^2 1b_2^2$ configuration. $|\Phi\rangle = |1a_1^2 2a_1^2 3a_1^2\rangle$ is only a minor component of $|^1A_1\rangle$. Hence, the T_2 amplitudes are very large (≫ 1.0) because $|^1A_1\rangle = |\Phi\rangle + T |\Phi\rangle + \cdots$ must contain relatively little $|\Phi\rangle$ and a large amplitude of $T_2 |\Phi\rangle$. Moreover, in the product region, the SCF orbitals of $|\Phi\rangle$ are very bad for use in describing BeH_2; the $1b_2$ orbital of $|\Phi\rangle$ is an SCF virtual orbital of the $1a_1^2 2a_1^2 3a_1^2$ charge density which (for products) in no way corresponds to the $1s_{Be}^2 \sigma_{BeH}^2 \sigma_{BeH}^2$ charge density. Therefore, we also expect the T_1 amplitudes to be vary large in the product region of the RP.

From the above analysis of a not atypical chemical reaction, we see that single configuration based CC methods can work, but only if they are capable of dealing with T amplitudes which range from small (~ 0.01) to large (≫ 1.0) values. Of course, some of the difficulties pointed out above could be removed by using the $1a_1^2 2a_1^2 3a_1^2$ configuration as a reference function on the reactant side of the RP and the $1a_1^2 2a_1^2 1b_2^2$ configuration

on the product side. Such "solutions" are, however, not satisfactory. What do you do near the avoided crossing and where, along the RP, do you switch from one reference configuration to the other?

We feel that the problems outlined above are better dealt with by using a multiconfiguration reference function ($1a_1^2 2a_1^2 3a_1^2$ plus $1a_1^2 2a_1^2 1b_2^2$ for this example) so that $\exp(T)$ will always contain small ($\lesssim 0.1$) amplitudes. We admit that it is possible to use a single configuration reference but only if the following two conditions are met:

(i) One must not, in deriving a set of CC equations which govern the T amplitudes, discard higher powers of T in comparison to lower powers of T. If T is large this would be entirely incorrect. As we will see later, this means that one can not throw away the cubic and quartic commutators [see Eq. (6)] which arise in $\exp(-T) H \exp(T)$; they may be as large or larger than the lower commutators if the T amplitudes are large. Only in special cases ($T \approx T_2$) will the contributions made by the cubic and quartic terms to the energy and to the equations which govern the T amplitudes vanish identically.

(ii) One must be able to numerically solve the equations for T using an algorithm which is capable of finding large T amplitudes. Procedures which are based upon iterative solutions starting with a perturbation theory estimate of the T amplitudes may be dangerous to use. Rather than face the formidable difficulties which arise in using the single configuration reference function, we have chosen to explore (as an *alternative*) the use of a MCSCF reference function within a CC framework. The remainder of this paper deals with our findings about how this MCSCF-based approach performs.

We first apply the CCMC method to several model systems [for which full configuration interaction (CI) calculations can be done] to explore the effect of the choice of reference function on the computational accuracy and efficiency. Based upon the experience gained from these model calculations, we have also applied the method to a chemically more important system (CH_2, $^3B_1$, and $^1A_1$). Section II gives a brief sketch of the formal development carried out in Ref. (6). Section III describes aspects of the computational implementation of the method, and Sec. IV contains our results.

II. DESCRIPTION OF OUR CCMC METHOD

We begin by letting $|0\rangle$ define a "core function" in which all orbitals are doubly occupied. These orbitals remain passive in the sense that they remain doubly occupied in *all* configurations $\{\chi_k\}$ of the multiconfigurational (MC) reference and in all determinants arising in $\exp(T)\Phi$. That is, the T operator does not excite electrons from these core orbitals. The choice of the core is dictated by the chemical problems at hand. In the reaction $Be + H_2 \rightarrow BeH_2 \rightarrow BeH + H$, it may be acceptable to choose the $1s$ orbitals of Be to be the core. In studying $Ni + CO \rightarrow NiCO$, it could be appropriate to include the CO σ_{CO} and n_O (lone pair on oxygen) orbitals in the

core (along with the obvious $1s_O$, $1s_C$, $1s_{Ni}$, $2s_{Ni}$, $2p_{Ni}$), even though these orbitals are energetically not widely separated from the "active" (e.g., n_c, π_{CO}, π_{CO}^*) orbitals. That is, the core orbitals *need not* include only the low-energy inner-shell orbitals. The chemical process under study *defines* the active orbitals; the remaining orbitals can then be treated as core orbitals as far as that particular process is involved. We emphasize that the reduction in the size of the problem effected by a particular choice of the core (passive) orbitals is merely a matter of convenience. One can, of course, always choose the core $|0\rangle$ to involve no electrons, thereby treating all orbitals as active. Those orbitals which occur in the various configurations of the MC reference function but which are *not* always doubly occupied we refer to as *valence* orbitals. Our CC method, therefore, deals only with how to treat electron correlation among the valence electrons.

The multiconfiguration (MC) reference wave function used in Ref. (6) is expressed in the form

$$\Phi(C, \chi) = \sum_K \chi_K C_K, \quad \chi_K = \Omega_K^+ |0\rangle, \tag{1}$$

where the configurations χ_K are written in terms of operators which add *valence* electrons to $|0\rangle$. That is, Ω_K^+ contains only valence creation operators (in normal order). These valence orbitals are designated by Greek subscripts $\alpha, \beta, \ldots$. As was done in Ref. (6), we assume that the MC function Φ is expressed as a linear combination of *all* configurations $\{\chi_k\}$ which can be formed by distributing the N valence electrons among the M valence orbitals. For this reason, we say that Φ consists of a full-valence CI wave function. For the $Be + H_2$ example used in Sec. I, the valence orbitals could be the $2a_1$, $3a_1$, and $1b_2$ orbitals (with the $1a_1$ orbital being a core orbital). Φ would then contain all valence configurations ($2a_1^2 3a_1^2$, $2a_1^2 1b_2^2$, $3a_1^2 1b_2^2$, $2a_1 3a_1 1b_2^2$) consistent with $^1A_1$ symmetry (the $1a_1^2$ component of each configuration was suppressed). The precise form of the core and valence ($1a_1, 2a_1, 3a_1, 1b_2$) orbitals is determined via the MCSCF process. In this MCSCF step, the amplitudes $\{C_K\}$ of the above four configurations are determined as are the full set of core, valence, and excited MCSCF orbitals.

We designate the (*excited*) orbitals not occupied anywhere in Φ by Roman subscripts $p, q, r, s, \ldots$. The subscripts i, j, k, and l are used to denote arbitrary orbitals. Within this notation, the cluster operators, which will eventually be allowed to act on the MC reference function Φ, are expressed as follows:

$$T = \sum_{s=1}^{N} T_s,$$

with

$$T_1 = \sum_{r\alpha} t_{r\alpha} e_{r\alpha}, \quad T_2 = \frac{1}{2} \sum_{r\alpha s\beta} t_{r\alpha s\beta} e_{r\alpha s\beta}, \ldots. \tag{2}$$

The amplitudes $t_{r\alpha}$, $t_{r\alpha s\beta}$, $\ldots$ are the fundamental quantities to be evaluated in the CC calculation. Here, $e_{ijkl} = e_{ij} e_{kl} - \delta_{jk} e_{il}$, and $e_{ij} = \sum_\sigma a_{i\sigma}^+ a_{j\sigma}$ are the generators of the unitary group, with $a_{i\sigma}^+$ being a Fermion creation operator for an electron in spatial orbital ϕ_i

and spin state $\sigma(\alpha$ or $\beta)$.

In writing the T operator as in Eq. (2), we have introduced one of the fundamental approximations of our method. We restrict T to excite electrons from the set of valence orbitals to the excited orbitals. T contains no terms of the form $t_{\alpha\beta}e_{\alpha\beta}$ which would cause excitations among the valence orbitals, or $t_{\alpha\beta s\gamma}e_{\alpha\beta s\gamma}$ which would cause "semi-internal" excitations from valence orbital β to valence orbital α, and from γ to excited orbital s. Because our reference function Φ is of the full-valence CI form, excitations among the valence orbitals above $(t_{\alpha\beta}e_{\alpha\beta})$ would generate no new configurations when acting in Φ (and hence would only allow for adjustments in the C_K mixing coefficients) which is expected to be of little importance because these valence orbitals have been MCSCF optimized. Likewise, by neglecting the semi-internal excitation terms $(t_{\alpha\beta s\gamma}e_{\alpha\beta s\gamma})$ we are making an approximation, but one which we feel the FVCI reference function makes physically reasonable. For example, in the BeH$_2$ case, the effect of a $t_{1b_2 3a_1 2b_2 3a_1}$ term would be to generate from the $2a_1^2 3a_1^2$ configuration a $2a_1^2 1b_2 2b_2$ configuration (properly spin coupled). A term $t_{2a_1 3a_1 4a_1 3a_1}$ would give a configuration $2a_1 4a_1 1b_2^2$ when operating on $3a_1^2 1b_2^2$. Both of these two resultant configurations could instead be generated by valence-to-excited T operators which *are* included in our calculations: $2a_1^2 1b_2 2b_2$ arises from $e_{2b_1 1b_2}$ operating on $2a_1^2 1b_2^2$. Moreover, the importance of these configurations is expected to be small because they correspond to valence orbital relaxation configurations. If the MCSCF process has already optimized the valence orbitals to adequate precision, further relaxation should be minimal. In summary, we feel that the use of the FVCI reference function whose orbitals have been MCSCF optimized makes our *approximation* (restricting T to contain only valence-to-excited excitations) quite reasonable.

The Hamiltonian can also be written in terms of the above unitary generators as

$$H = \langle 0 | H | 0 \rangle + \sum_{ij} f_{ij} N[e_{ij}] + \frac{1}{2} \sum_{ijkl} V_{ijkl} N[e_{ik} e_{jl}] , \qquad (3)$$

where

$$f_{ij} = h_{ij} + \sum_{c \in |0\rangle} (2 V_{icjc} - V_{iccj}) , \qquad (4)$$

and $c \in |0\rangle$ denotes that the orbitals ϕ_c run over the core orbitals. Here, h_{ij} and V_{ijkl} are the usual one- and two-electron integrals in the Hamiltonian. Because of the normal ordering of the operators in H, and because T contains no core-orbital operators, any time H operates on a function of the form $T^n | \Phi \rangle$, the terms involving core-orbital integrals vanish. That is, the indices i, j, k, l in the last two terms of Eq. (3) need not be summed over the core when H operates on any function containing a frozen core.

Because our T operator only contains operators of the form $e_{r\alpha}$, $e_{r\alpha s\beta}$, etc. (which involve only valence-to-excited orbital promotions), the usual intermediate normalization property of $\exp(T)\Phi$

$$\langle \Phi | e^T \Phi \rangle = \langle \Phi | \Phi \rangle + \sum_{r\alpha} t_{r\alpha} \langle \Phi | e_{r\alpha} | \Phi \rangle + \cdots = 1 , \qquad (5)$$

and the (exact) truncation of the commutator expansion

$$\tilde{H} \equiv e^{-T} H e^T = H + [H, T] + \frac{1}{2!} [[H, T], T]$$

$$+ \frac{1}{3!} [[[H, T], T], T] + \frac{1}{4!} [[[[H, T], T], T], T] \qquad (6)$$

are straightforward to verify.

In a single-configuration based CC theory, the above [Eq. (6)] quartic truncation of the commutator expansion also occurs. Moreover, if one approximates T by T_2 (i.e., if one neglects T_1, T_3, etc.), Then the cubic and quartic terms in Eq. (6) contribute nothing (in the single configuration case) to the equations [Eq. (9) and (10)] governing the t amplitudes and the energy. Such a $T = T_2$ approximation is made in the so-called coupled pair many-electron theory[4] (CPMET). It is our opinion that if one insists on using a single configuration reference function then, as we illustrated in Sec. I for the Be + H$_2$ example, one can not neglect the orbital relaxation effects described by T_1, in which case Eq. (6) contains all four commutator terms. Certainly if one has good reason to believe that the single configuration reference function dominates the true wave function (i.e., that the t amplitudes are small), it makes good sense to then (as an approximation) neglect the cubic and quartic commutators in Eq. (6). However, to do so when the single configuration reference function represents a nondominant or even minor contribution to the true wave function is not proper because in such cases the t amplitudes (including T_1) will be large ($\gtrsim 1.0$).

In the work presented here, we have ignored the cubic and quartic commutators in Eq. (6) because we have reason to believe that our MCSCF reference function already contains the dominant contributions to the true wave function, and hence that the t amplitudes are small.

Assuming that $\exp(T)$ is the operator which, when acting on Φ, gives an eigenfunction of H

$$\Psi = e^T \Phi , \qquad (7)$$

with

$$H\Psi = E\Psi , \qquad (8)$$

then the unknown t amplitudes $(t_{m\alpha}, t_{m\alpha n\beta}, \ldots)$ can be calculated by projecting Eq. (8) against a sufficient set of "excited" functions $\Phi_{r\alpha s\beta\cdots} = e_{r\alpha s\beta\cdots} | \Phi \rangle$, and using Eqs. (5) and (6) to give

$$\langle \Phi_{r\alpha s\beta\cdots} | \tilde{H} | \Phi \rangle = 0 . \qquad (9)$$

Once the t amplitudes are known, the total energy can be calculated from

$$\langle \Phi | \tilde{H} | \Phi \rangle = E , \qquad (10)$$

Although the configuration mixing coefficients C are presumed (in Ref. 6) to have already been determined, and hence held fixed, it is possible to further extend the range of the theory by calculating the C coefficients by projecting Eq. (8) against the configurations of the reference space $\{\chi_L\}$

$$\langle \chi_L | \tilde{H} | \Phi \rangle = E C_L . \qquad (11)$$

Because Eqs. (8) and (11) both contain reference to the C and t coefficients, the calculation of t and C would then require an iterative scheme. In the present applications of the theory developed in Ref. 6, we choose to compute C via a small multiconfigurational self-consistent field calculation and to then calculate t via Eq. (9). This approximation procedure is in line with that put forth earlier,[6] and it avoids the need for iterative determination of t and C.

In Ref. 6, we give detailed expressions for the matrix elements arising in Eq. (9), subject to the following approximations: (i) the T operator manifold is truncated to include one- and two-body operators (T_1 and T_2), and (ii) the expansion for the effective Hamiltonian $\bar{H}$ [Eq. (6)] is truncated at the double commutator level. Approximation (i) is generally accepted to account for most of the electron correlation effects in atoms and molecules. Approximation (ii), whose justification was discussed briefly above, is common to almost all CC calculations which have been performed to date[13(b)] (all of which have involved only single-determinant reference states). Adoption of this second approximation assumes that the t amplitudes are small ($\ll 1$) for the chosen reference function Φ. Otherwise, terms arising from the third and fourth commutators in the expansion for $\bar{H}$ [Eq. (6)] would be important and the truncation of Eq. (6) at the quadratic terms would lead to meaningless results.

In most of the applications presented here, we have also explored the use of one further approximation which is dictated by computational considerations. Even the linear and quadratic commutators occurring in $\bar{H}$ [Eq. (6)] lead to three-body density matrix elements (see Ref. 6) in the equations [Eq. (9)] which determine the t amplitudes (but *not* in the equation giving the energy). Only density matrices whose indices simultaneously involve both valence and excited orbitals occur in these equations. However, the lists of these three-body density matrix elements is too large to be efficiently incorporated into our CCMC implementation. In CC methods which are based upon a *single-determinant* reference state, the evaluation of these three-body density matrices is substantially easier than in our (MC) case because, as is well known, all such density matrix elements are either zero or unity. In the present work, we chose to carry out two sets of calculations; one in which these higher density matrices are ignored and another in which they are properly treated. In this way, we attempt to test the hypothesis that they can often be safely ignored. In some of the results presented below (those involving only two *active* electrons—H_2 and certain CH_2 cases), the three-body density matrices vanish identically; hence, they present no problem. In other cases, we have evaluated the *full* effect of the higher density matrices as well as the results obtained in their absence. As is shown by the results given below, it seems that the neglect of these specific higher-body density matrices may often be justified. It is our feeling that, although these initial results are rather encouraging, further systematic exploration of this question is called for.

Before moving on to investigate, via numerical application, how well the proposed CCMC method works, we first turn our attention to the computational implementation of the technique.

III. COMPUTATIONAL CONSIDERATIONS

We have implemented, on the Utah DEC 2060 computer, the CCMC method developed by us in Ref. 6. The computer implementation of the CCMC method involves both calculation of the MC reference wave function and subsequent evaluation of the cluster amplitudes, using the CCMC equations given in Ref. 6. The calculational details of the procedure undertaken are dictated by our decision to use the unitary group approach.[9–11] As seen in Ref. 6, the matrix elements of H and $\bar{H}$ can be written as sums of products of integrals and density matrix elements $f_{ij}\langle\Phi|e_{\alpha\beta}|\Phi\rangle$, $V_{ijkl}\langle\Phi|e_{\alpha\beta\mu\nu}|\Phi\rangle$. For any nontrivial problem, neither of these two lists can be kept in the computer's fast memory. Because we have chosen to write all matrix elements in terms of density matrix elements of the reference function Φ, all $\langle e_{\alpha\beta}\rangle$ and $\langle e_{\alpha\beta\mu\nu}\rangle$ involve only the valence orbitals of Φ. As a result, it is most efficient to permit the flow of the program to be driven by the (longer) list of integrals. That is, for each block of integrals, we read in and sort the entire $\langle e_{\alpha\beta}\rangle$ and $\langle e_{\alpha\beta\mu\nu}\rangle$ lists and we update the appropriate matrix elements of H or $\bar{H}$. The machinery of the unitary group approach provides very efficient algorithms[11] for the calculation of density matrix elements through a nonredundant and "global" organization of the configurations $\{\chi_K\}$. In addition, since the H and T operators are explicitly constructed to be totally (spin and space) symmetric, the symmetry of Φ is preserved throughout the calculations. Two steps involved in the calculation of the cluster amplitudes require further elaboration.

A. Solution of the nonlinear equations

As can be seen in Ref. 6, retaining up through the double commutator in the expansion of $\bar{H}$ yields nonlinear (quadratic) equations for the t amplitudes

$$F(t) \cong A + Bt + Ctt = 0 . \tag{12}$$

Here, the matrices A, B and C arise, respectively, from the zero-, one-, and two-commutator terms of Eq. (6). The dimension (m) of this problem, which is the number of t amplitudes, can become so large that the list of A, B, and C matrix elements cannot be kept in the computer's high speed memory. In the literature,[13] one can find efficient solution procedures for handling large dimensional *linear* equations $A + Bt = 0$, as well as discussion[13] of how one might treat the nonlinear CC equations via linearization procedures. In what follows, we present our own approach to this nonlinear equation problem. Our approach begins with the above [Eq. (12)] set of quadratic equations, and results in a small dimensional set of quadratic equations which can be solved in the computer's high speed memory.

In most such procedures, the matrices are projected onto a (smaller dimensional) space spanned by a new set of variables $t_i = \sum_j^m \alpha_{ji} t_j$, $i = 1, \ldots, p \ll m$. To solve the above equations [Eq. (12)], we need to use a pro-

cedure appropriate to the nonlinear case. Let us represent the desired solution vector t as

$$t = \alpha_0 t_0 + \alpha_1 (t_1 - t_0) + \alpha_2 (t_2 - t_1) + \cdots \equiv \sum_{j=0} \alpha_j \Delta t_j , \qquad (13)$$

where Δt_j is the difference between the jth and $(j-1)$th approximations to t, and $\Delta t_0 = t_0$. As a criterion for convergence, we require $\alpha_p \to 0$ (hopefully for $p \ll m$). Substituting Eq. (13) into Eq. (12) and projecting against $\langle \Delta t_i |$, we obtain

$$\langle \Delta t_i | \mathbf{A} \rangle + \sum_j \alpha_j \langle \Delta t_i | \mathbf{B} | \Delta t_j \rangle + \sum_{jk} \alpha_j \alpha_k \langle \Delta t_i | \mathbf{C} | \Delta t_j \Delta t_k \rangle = 0 ,$$

or, in matrix notation

$$\tilde{\mathbf{A}} + \alpha \tilde{\mathbf{B}} + \tilde{\mathbf{C}} \alpha \alpha = 0 , \qquad (14)$$

where the $\{\alpha_j\}$ are now viewed as the unknown coefficients. These are a set of nonlinear equations in the (smaller dimension) projected space which can be solved in the computer's high speed core memory by (as in our calculations) the Newton–Raphson (NR) procedure or its variant $\tilde{\mathbf{B}} \alpha^{(i+1)} = -\tilde{\mathbf{A}} - \tilde{\mathbf{C}} \alpha^{(i)} \alpha^{(i)}$, which yields an updated $\alpha^{(i+1)}$ from a previous $\alpha^{(i)}$.

The convergence of this procedure obviously depends upon the choice of the initial vector[13] t_0, as well as on the method used to generate the successive t_j's (or Δt_j's) which are needed to form the $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, and $\tilde{\mathbf{C}}$ arrays in Eq. (14).

Given some choice of t_0, further Δt_j's can be obtained from the following algorithm based on iterating Eq. (12)

$$[\mathbf{B}_d + (\mathbf{C}t_j)_d] t_{j+1} = -\mathbf{A} - \mathbf{B}_n t_j - (\mathbf{C}t_j)_n t_j . \qquad (15)$$

Here, the subscripts d and n refer to the diagonal and off-diagonal parts of the respective matrices. By subtracting $[\mathbf{B}_d + (\mathbf{C}t_i)_d] t_j$ from this equation, we obtain an expression which can be used to iteratively obtain the $\{\Delta t_j\}$

$$[\mathbf{B}_d + (\mathbf{C}t_j)_d] \Delta t_{j+1} = -\mathbf{F}(t_j) . \qquad (16)$$

By starting with a t_0, Eq. (16) then generates successive Δt_j's, which can be orthogonalized, and used in Eq. (14) to construct the (smaller dimensional) projected-space problem. The solution of Eq. (14) then gives values of $\{\alpha_j\}$, $j = 1, \ldots, p$. If for some p, $|\alpha_p| \to 0$ within a tolerance, we say that the process has reached convergence. The *final* t amplitude array is then given in terms of the α by Eq. (13). If the dimension (p) of these reduced equations becomes larger than can be handled, one can simply take a new t_0 to be given by the "current" t [Eq. (13)] and start the procedure over again with this t_0.

Of course, it is important to address how one obtains t_0 at the start of the entire calculation. An obvious choice of t_0 is the one satisfying the linear CCMC equation $\mathbf{A} + \mathbf{B} t_0 = 0$. This starting value for t_0 has been successfully used[12] in cases where the single-determinant reference function Φ dominates the exact function $\exp(T)\Phi$. In such cases, one expects the magnitudes of the t amplitudes to be small ($\ll 1.0$), and hence the $\mathbf{C} tt$ term should be smaller than the $\mathbf{B}t$ and $\mathbf{A}$ terms. However, in such single-determinant-based CC theo-

ries, it is not at all clear that this starting point for t_0 will be successful[13(b)] when Φ does not dominate the exact $\exp(T)\Phi$. In some cases, Φ might even be a *negligible* component to $\exp(T)\Phi$. For such problem cases, one could achieve an initial estimate for the t amplitudes by carrying out a CI calculation (the CI expansion coefficients then being used to approximate t). By then developing iterative solutions to Eq. (12) in the neighborhood of *this* t_0, one might be able to efficiently converge to a solution. Alternatively, one can (as in our CCMC method), by employing the MC reference function, maximize the chances that all t amplitudes remain small (since the large contributions to the exact state are already in Φ). In this case, the linearized solution $\mathbf{A} + \mathbf{B} t_0 = 0$ should represent a useful starting point.

B. Calculation of core-orbital contributions

The core orbitals are those which, by definition, are not correlated by either the reference wave function or the CC method. They are assumed to be doubly occupied in *all* configurations of Φ and of $T|\Phi\rangle$. Moreover, as pointed out after Eq. (3), the Hamiltonian need not contain reference to integrals involving these core orbitals whenever H operates on a function of the form $T^n |\Phi\rangle$. Based on these facts, it can be shown[23] that the core contributions (i.e., the core–core, core-valence, and core-excited orbital interactions) can be grouped into a simple effective one-particle potential (f) of the form given in Eq. (4). Moreover, f need be calculated only once at the beginning of the CCMC calculation. The use of this potential then obviates any further reference to core orbitals in the construction of the matrix elements of H and $\bar{H}$. This fact is especially helpful since it reduces the size of the lists of two-electron integrals and density matrix elements which need to be randomly accessed from a peripheral device in the construction of the $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ matrix elements.

IV. RESULTS AND DISCUSSION

Applications of the above outlined CCMC theory have been carried out on four systems, $H_2(^1\Sigma_g^+)$, $Li(^2S)$, $HeH_2(^1A_1)$, and $CH_2(^3B_1)$ which we view as *model problems* because the basis sets used are sufficiently small to permit a full CI calculation to be done so that the "exact" answers can be known. This fact allows us to gauge the accuracy of CCMC results obtained by starting from any of three different MC reference functions which we have explored: the single-determinant restricted Hartree–Fock (RHF) wave function, a CI-type wave function expressed in terms of RHF orbitals, and a MCSCF-type wave function. For the closed-shell systems, we have also compared our results with the CPMET ($T \cong T_2$, Φ equal to a closed-shell single-determinant Hartree–Fock function) results. For such cases, our CCMC theory exactly reduces to the CPMET.

When considering various reference functions for use in a CCMC calculation, it should be kept in mind that a good reference function not only should have the potential of giving good total energies; it should also produce

small t amplitudes so that our CCMC iteration process will converge rapidly and so that our neglect of the cubic and quartic commutators of Eq. (6) is justified. An MCSCF reference function which includes the qualitatively essential configurations for the description of the exact state[14] is shown below to have these attractive attributes. For such an MCSCF reference function, the T_1 contributions are expected to be negligible whenever the nonlinear contributions from T_1 and T_2 are small. This follows from the generalized Brillouin theorem (GBT)[15] which, for our definitions of Φ and the T operators, takes the form

$$\langle \Phi_{r\alpha} - \Phi_{\alpha r} | H | \Phi \rangle = \langle \Phi_{r\alpha} | H | \Phi \rangle = 0 . \tag{17}$$

Thus, if $T = T_1$, the GBT requires $A = 0$, from which the *linear* approximation to the CC equations yields $A + Bt = 0$ or $t = 0$, so that we expect the T_1 amplitudes to be small.

As mentioned earlier, we also examined whether one could, as has been proposed,[7,8] obtain reliable results by neglecting the small subset of higher body density matrices which arise in the equation governing the t amplitudes. This was done, as described below, by carrying out calculations with and without these higher density matrices.

Based on experience gained from the model calculations whose results are presented below, we also applied the CCMC method to a more realistic calculation of the $^3B_1$ ground and $^1A_1$ excited states of CH_2 using a larger basis set. These calculations were performed to permit us to evaluate the accuracy of the CCMC method on a problem for which numerous other state-of-the-art theoretical calculations have been done.

The results of all of our CCMC calculations are given in Tables I–VIII. In each table, the first column describes the reference function (HF, CI, or MCSCF). The next column describes the order of cluster excita-

tions (T_1, T_2) included in the calculation. Column three gives the total energy as a sum of various contributions $E_{tot} = E_0 + E_{T_1} + E_{T_1 - T_1} + E_{T_2}$, according to Eq. (15a) of Ref. (6). The fourth column gives the number of iterations needed to obtain convergence to a tolerance of 10^{-10} in the t amplitudes and the fifth column gives the largest t amplitude obtained from the converged CCMC calculation. This last information is a measure of how much correlation lies outside the MC reference space. In *all* of the applications presented here, we have assumed a truncation of Eq. (6) at the quadratic term. The cubic and quartic commutators are not evaluated. As a result, we may encounter difficulties in solving the CC equations in the single configuration reference function case which might disappear if one were to include all terms[13] in Eq. (6). We *choose* not to calculate the cubic and quartic commutators but to instead use a MCSCF reference function. Let us now analyze in some detail the convergence characteristics demonstrated by these CCMC results.

A. $H_2(^1\Sigma_g^+)$

Tables I to III describe the results of our CCMC calculations on the lowest $^1\Sigma_g^+$ state of H_2 for various reference functions at internuclear distances of $R = 1.4$, 4.0, and 6.0 bohr. These bond lengths cover values of R near equilibrium, when the bond is essentially broken, and when the bond is nearly half-formed. Our results were obtained by using a double-zeta contracted Gaussian basis set.[17] For this closed-shell two-electron system, our intention is to compare the performance of our method with the CPMET to which our method reduces if we take $T = T_2$ and $\Phi = \Phi_{HF}$. In this case, the CCMC method also reproduces exactly the doubles-CI (DCI) result, for which our three-particle density matrix approximation clearly has no effect (i.e., the result of a CI calculation comprised only of doubly excited

TABLE I. $H_2(^1\Sigma_g^+)$, $R = 1.4$ bohr, CCMC energies and t amplitudes for RHF, CI, and MCSCF reference functions.

Reference wave function	Calculation type	Energy contributions (hartree)					Number of iterations	Largest t amplitudes
		E_0	E_{T_1}	$E_{T_1 - T_1}$	E_{T_2}	E_{tot}		
1. Φ_{HF}	T_1	-1.126815	10^{-11}	10^{-13}		-1.126815	1	10^{-4}
2. Φ_{HF}	T_2(CPMET)	-1.126815			-0.0258	-1.152669	3	-0.06
3. Φ_{HF}	$T_1 + T_2$	-1.126815	10^{-8}	10^{-5}	-0.0	-1.152760	3	-0.06
4. $\Phi = C_1 \sigma_g^2 + C_2 \sigma_u^2$ $C_1 = 0.9987,\ C_2 = -0.05088$	T_1	-1.129904	-0.0042	-0.0015		-1.135709	5	-0.468
5. $\Phi = C_1 \sigma_g^2 + C_2 \sigma_u^2$ $C_1 = 0.9987,\ C_1 = -0.05088$	T_2	-1.129904			-0.0137	-1.143593	4	0.268
6. $\Phi = C_1 \sigma_g^2 + C_2 \sigma_u^2$ $C_1 = 0.9987,\ C_1 = -0.05088$	$T_1 + T_2$	-1.129904	-0.0101	-0.0003	-0.01227	-1.152699	4	0.269
7. $\Phi = C_1 \sigma_g^2 + C_2 \sigma_u^2$ MCSCF $C_1 = 0.9934,\ C_2 = -0.1145$	T_1	-1.146921	10^{-7}	10^{-8}		-1.146921	2	10^{-3}
8. $\Phi = C_1 \sigma_g^2 + C_2 \sigma_u^2$ MCSCF $C_1 = 0.9934,\ C_2 = -0.1145$	T_2	-1.146921			-0.0057	-1.152672	4	0.1
9. $\Phi = C_1 \sigma_g^2 + C_2 \sigma_u^2$ MCSCF $C_1 = 0.9934,\ C_2 = -0.1145$	$T_1 + T_2$	-1.146921	10^{-6}	-10^{-5}	-0.0058	-1.152749	4	0.1
10. DCI						-1.152669		
11. FCI						-1.152764		

TABLE II. $H_2(^1\Sigma_g^+)$, $R = 4.0$ bohr, CCMC energies and t amplitudes for RHF, CI, and MCSCF reference functions.

Reference wave function	Calculation type	Energy contributions (hartree)					Number of iterations	Largest t amplitudes
		E_0	E_{T_1}	$E_{T_1-T_1}$	E_{T_2}	E_{tot}		
1. Φ_{HF}	T_1	-0.901978	10^{-13}	10^{-14}		-0.901978	1	10^{-6}
2. Φ_{HF}	T_2(CPMET)	-0.901978			-0.0998	-1.001851	6	$0.54(t_{\sigma_u^2})$
3. Φ_{HF}	T_1+T_2	-0.901978	10^{-7}	10^{-3}	-0.1081	-1.009691	6	0.54
4. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ $C_1=0.8768,\ C_2=-0.4807$	T_1	-0.989253	-0.019	10^{-3}		-1.009035	3	$0.20\,\sigma_u^2\to\sigma_u\sigma_u'$ amplitude
5. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ $C_1=0.8768,\ C_2=-0.4807$	T_2	-0.989253			-0.0005	-0.989774	2	0.01
6. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ $C_1=0.8768,\ C_2=0.4807$	T_1+T_2	-0.989253	-0.0196	-0.0003	-0.0003	-1.009699	3	0.19
7. $\Phi=C_1\sigma^2+C_2\sigma_u^2$ MCSCF $C_1=0.08596,\ C_2=-0.5109$	T_1	-1.00956	10^{-6}	10^{-10}		-1.009569	2	10^{-3}
8. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ MCSCF $C_1=0.08596,\ C_2=-0.5109$	T_2	-1.00956			-0.0001	-1.009704	2	0.03
9. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ MCSCF $C_1=0.08596,\ C_2=-0.5109$	T_1+T_2	-1.00956	10^{-6}	10^{-8}	-0.0001	-1.00708	3	0.03
10. DCI						-1.001851		
11. FCI						-1.009709		

configurations). Table I displays the CCMC results for $R=1.4$ bohr (which is near the equilibrium bond length). With a RHF reference function ($\Phi=\sigma_g^2$), the results (rows numbered 1–3) show that energy contributions from T_1 alone are negligibly small ($\sim 10^{-10}$ a.u.) due to the Brillouin theorem (BT), and that they increase slightly ($\sim 10^{-5}$ a.u.) when T_2 is added, due to the T_1-T_2 coupling. The resulting CC energy is within 4×10^{-5} hartree of the full-CI (FCI) result (which is, of course, the exact result in any theory CC, CI, etc.), the discrepancy being due to the neglected commutators involving T_1^3, $T_1^2T_2$, etc., which are small because t_1 is small. The CPMET (i.e., $T=T_2$) result clearly reproduces the DCI value, as it must.

At this geometry, σ_g^2 is the only dominant configuration. The σ_u^2 configuration, which is needed for the cor-

rect dissociation to two $H(^2S)$ fragments, has an amplitude of -0.11 in the MCSCF reference function at $R=1.4$ bohr. Using a two configuration $(C_1\sigma_g^2+C_2\sigma_u^2)$ MCSCF reference function, the CCMC energy (rows numbered 6–8) for $T=T_1+T_2$ comes within 1×10^{-5} a.u. of the FCI value, and the T_1 contributions are found to be negligible (due to the GBT). For a two-configuration CI reference function (using RHF orbitals), the T_1 contribution is not negligible since the σ_u orbital is not optimized (as it was in the MCSCF case). Hence, the $\sigma_u\sigma_u'$ configurations resulting from single excitations contribute significantly via T_1.

Difficulties in the CC calculations based upon the RHF reference function [with a quadratic truncation of Eq. (6)] begin to emerge as we move toward the dissociation limit. As shown in Tables II and III for $R=4.0$, 6.0

TABLE III. $H_2(^1\Sigma_g^+)$, $R=6.0$ bohr, CCMC energies and t amplitudes for RHF, CI, and MCSCF reference functions.

Reference wave function	Calculation type	Energy contributions (hartree)					Number of iteration	Largest t amplitudes
		E_0	E_{T_1}	$E_{T_1-T_1}$	E_{T_2}	E_{tot}		
1. Φ_{HF}	T_1	-0.808908	10^{-13}	10^{-15}		-0.808908	1	10^{-6}
2. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ $C_1=0.7532,\ C_2=-0.6577$	T_1	-0.976505	-0.0198	10^{-5}		-0.996396	3	0.1
3. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ $C_1=0.7532,\ C_2=0.6577$	T_2	-0.976505			-10^{-5}	-0.976508	2	10^{-2}
4. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ $C_1=0.7532,\ C_2=-0.6577$	T_1+T_2	-0.976505	-0.0199	10^{-5}	-10^{-4}	-0.996926	2	0.12
5. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ MCSCF $C_1=0.7433,\ C_2=-0.6689$	T_1	-0.99701	10^{-6}	10^{-10}		-0.997016	2	10^{-3}
6. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ MCSCF $C_1=0.7433,\ C_2=-0.6689$	T_2	-0.99701			-10^{-5}	-0.997018	3	0.02
7. $\Phi=C_1\sigma_g^2+C_2\sigma_u^2$ MCSCF $C_1=0.7433,\ C_2=-0.6689$	T_1+T_2	-0.99701	-10^{-6}	10^{-10}	-10^{-5}	-0.997018	3	0.02
8. DCI						-0.983244		
9. FCI						-0.997018		

bohr, respectively, the RHF wave function σ_g^2 becomes a poorer representation of the ground state. The proper wave function in this region is a combination of σ_g^2 and σ_u^2. For $R = 4.0$ bohr, the solution of our CC equations via the Newton–Raphson process showed convergence difficulty. The final solution to which the NR process converged depended strongly upon the initial choice of t and the damping factors used in the NR iterative process. The final solutions[18] shown in the 2nd and 3rd rows of Table II were obtained by monitoring the magnitude of the $t_{\sigma_u^2}$ amplitude and by comparing the resultant CCMC energy with the FCI energy (which is, of course, available only for such "model" calculations). Certainly, it might be possible[13] to devise iterative schemes based upon solving the quartic CC equations which can converge in such large-t cases. However, within our approach to the CC problem the appearance of large t amplitudes makes the NR solution of the CC equations very difficult. At $R = 6.0$ bohr, we were unable to converge the quadratic (not quartic) CCMC equations except by taking the T_2 amplitudes from a doubles CI wave function. Upon doing so, we were able to converge to the DCI (which equals CPMET) energy. For starting t amplitudes which were far from the converged values, our truncated CC equations displayed significant convergent trouble.

Even though it is *possible* to devise iterative solutions[13] of the CC equations which do converge and hence allow this CPMET calculation to reach the "correct" DCI solution, we find that the DCI result itself yields a progressively poorer result with increasing R value, the energy difference ($E_{FCI} - E_{DCI}$) being -0.00009, -0.008, -0.013, and -0.015 hartree, respectively, for R values of 1.4, 4.0, 6.0, and 10.0 bohr. The reason, of course, is that the dissociating fragments [$H(^2S)$ in this case] are themselves *open shell* systems which require both double excitations and single excitations (to relax the orbitals).

In contrast to the above single-configuration case, CCMC calculations using CI or MCSCF reference functions which include the essential configurations (σ_g^2 and σ_u^2) showed *no* convergence difficulties. The CCMC results with the MCSCF reference function (which implicitly includes the orbital relaxation effects to first order) with $T = T_2$ gave essentially the FCI energy values at $R = 4.0$, 6.0, and 10.0 bohr. On the other hand, use of the CI reference function, which also yields a CCMC energy close to the FCI value, gives rise (at $R = 6$ bohr) to large T_1 amplitudes, which attempt to account for the orbital relaxation which is absent in the CI function, but which was already present in the MCSCF reference function.

For either an RHF or MCSCF reference function, it is expected that T_1 contributions will be small whenever the $T_1 - T_2$ coupling is small. This result is a consequence of the BT or GBT, and does not necessarily hold for the CI-type reference functions which use RHF or natural orbitals. Therefore, by using an MCSCF reference space containing the dominant (essential) configurations, it can be virtually assured that T_1 contributions can be made small, which then allows one to

eliminate further consideration of T_1. CC solutions which are related to the reference solution via large (of the order of 1) t amplitudes are very difficult to locate and are poorly represented in the "truncated" commutator expansion which we choose to use here. The CI- or MCSCF-type reference function, which include dominant configurations, give CCMC energies comparable to the corresponding FCI energies. It is found that the CCMC energies obtained from linear CCMC equations ($A + Bt = 0$) generally lie *below* the FCI energies. This has also been observed for the case of a closed-shell single-determinant reference function in CPMET by Cizek[4] and Pople *et al.*,[12] and Mukherjee[16] has shown that the linear CC energy for such cases with $T = T_2$, *always* lies below the corresponding DCI energy.

B. Li(^2S)

Table IV describes the results of our CCMC calculations on the ^2S ground state of Li using five s-type contracted Gaussian functions. For Li(^2S), the RHF configuration[24] $\Phi_{RHF} = 1s^2 2s$ is an open-shell configuration. Some workers have extended the scope of their closed-shell RHF-based CC formalism to include such open shell cases by using spin-unrestricted Hartree–Fock (UHF) single-determinant reference functions[12,22] in which, of course, the wave functions are not exact spin eigenfunctions. In our formalism, such calculations can be performed *without* violating spin symmetry (and hence the CC function corresponds exactly to a spin eigenvalue). All of the results, except those given in rows 4 and 11, involve neglect of the three-body density matrices. Rows 4 and 11 give results for the full treatment of all density matrices. The CC results for a single configuration reference state (rows numbered 1–4) are found to be within 0.5×10^{-3} hartree of the FCI value since, in *this* 5s basis, the RHF configuration is the dominant configuration. Inclusion in the reference function itself of all of the configurations constructed within the orbital space $\{1s, 2s, 3s\}$ allows the CCMC calculation to capture the residual higher order correlation effects and reproduce the FCI value to within 0.7×10^{-4} hartree with T_2 alone. In this case, the MCSCF reference wave function itself captured almost all (99.9%) of the available correlation energy. The subsequent CCMC calculations, not surprisingly, give very small cluster contributions with $t_1 \cong 0$, $E_{T_1} \cong E_{T_1 - T_1} = 0$, and $E_{T_2} \cong -0.00003$ hartree. Since Li is a three-electron system, the three-body density matrices have a chance to affect our results. The results given in rows 5–10 were obtained by neglecting the set of three-body density matrices which arise (the one- and two-body matrices are, of course, treated at the MC level). In row 4, we give results in which the three-body density matrices are properly treated when Φ is taken to be the $1s^2 2s$ reference. Again, comparison of rows 4 and 3 provides some support for neglecting these higher density matrix contributions. In row 11, we give the result of a CCMC calculation in which Φ is the eight-configuration MCSCF function $T \cong T_1 + T_2$, and *all* of the three-body density matrix elements are calculated properly. Comparison with the results of row

TABLE IV. Li(^2S), CCMC energies and t amplitudes for RHF, CI, and MCSCF reference functions.

| Reference wave function | Calculation type | Energy contributions (hartrees) | | | | | Number of iterations | Largest t amplitudes |
		E_0	E_{T_1}	$E_{T_1-T_1}$	E_{T_2}	E_{tot}		
1. $\Phi_{HF} = 1s^2 2s$	T_1	-7.36869	-10^{-6}	10^{-8}		-7.368690	2	10^{-5}
2. $\Phi_{HF} = 1s^2 2s$	T_2	-7.36869			-0.01565	-7.384347	3	-0.02
3. $\Phi_{HF} = 1s^2 2s$	$T_1 + T_2$	-7.36869	-10^{-5}	10^{-6}	-0.01568	-7.384371	3	-0.02
4. $\Phi_{HF}: 1s^2 2s$	$T_1 + T_2$ (full)	-7.36869	-10^{-5}	10^{-6}	-0.01593	-7.384621	3	-0.02
5. $\Phi_{CI}:$ eight configurations within the orbital space of $\{1s, 2s, 3s\}$	T_1	-7.36872	-10^{-6}	10^{-8}		-7.368723	2	10^{-3}
6. $\Phi_{CI}:$ eight configurations within the orbital space of $\{1s, 2s, 3s\}$	T_1	-7.36872			-0.014844	-7.383568	3	-0.06
7. $\Phi_{CI}:$ eight configurations within the orbital space of $\{1s, 2s, 3s\}$	$T_1 + T_2$	-7.36872	-10^{-6}	10^{-6}	-0.014859	-7.383583	3	-0.06
8. $\Phi_{MCSCF}:$ eight configurations within the orbital space of $\{1s, 2s, 3s\}$	T_1	-7.38503	-10^{-8}	-10^{-10}		-7.385035	2	10^{-3}
9. $\Phi_{MCSCF}:$ eight configurations within the orbital space of $\{1s, 2s, 3s\}$	T_2	-7.38503			-0.00001	-7.385045	2	0.06
10. $\Phi_{MCSCF}:$ eight configurations within the orbital space of $\{1s, 2s, 3s\}$	$T_1 + T_2$	-7.38503	-10^{-7}	-10^{-6}	-0.00003	-7.385070	2	0.06
11. $\Phi_{MCSCF}:$ Eight configurations within the orbital space of $\{1s, 2s, 3s\}$	$T_1 + T_2$ (full)	-7.38503	-10^{-7}	10^{-6}	-0.00003	-7.385078	2	-0.06
12. FCI						-7.385117		

10 show that neglect of the higher density matrix was, in this case, accurate.

C. $CH_2(^3B_1)$ model calculation

CCMC calculations on the $^3B_1$ ground state of methylene were performed at its equilibrium geometry[20] ($\theta_{HCH} = 132.4$, $R_{CH} = 1.082$ Å) using a double-zeta contracted Gaussian[17] (C:$4s$, $2p$/H:$2s$) basis set. The results of these CCMC calculations are displayed in Table V. In Sec. IVE, we present the results of CCMC calculations on the $^3B_1$ and $^1A_1$ states using a larger basis set; here, we only describe the pertinent points of this small-basis *model* study for which full-CI results are available. The triplet RHF wave function

can be described in terms of the C_{2v} symmetry of the molecule as $\Phi_{RHF}(^3B_1) = |0\rangle 3a_1 1b_1$, with the (frozen) core being given by $|0\rangle = |1a_1^2 2a_1^2 1b_2^2\rangle$. For this particular choice of the core, which gives the most simple approximation to the active orbital space, the calculation reduces to an effective two-valence-electron problem. Thus, our CCMC calculation, using the RHF reference function and $T = T_2$, should exactly reproduce the corresponding DCI (within the active orbital space) result. The fact that the largest t amplitude arising in the RHF-based CC calculation is 0.04 indicates that the RHF configuration is the dominant configuration for the description of this triplet state. Thus, it is not surprising that the two-configuration CCMC calculation produces a very small change is the total energy. The

TABLE V. $CH_2(^3B_1)$, CCMC energies and t amplitudes for RHF and MCSCF reference functions.

| $CH_2(^3B_1)$ Reference wave function | Calculation type | Energy contributions (hartree) | | | | | Number of iterations | Largest t amplitudes |
		E_0	E_{T_1}	$E_{T_1-T_1}$	E_{T_2}	E_{tot}		
1. $\Phi_{HF} = 3a_1 1b_1$	T_1	-38.90851	-10^{-6}	10^{-8}		-38.908514	2	10^{-3}
2. $\Phi_{HF} = 3a_1 1b_1$	T_2	-38.90851			-0.0032	-38.911763	3	0.04
3. $\Phi_{HF} = 3a_1 1b_1$	$T_1 + T_2$	-38.90851	10^{-5}	10^{-6}	-0.0032	-38.911781	3	0.04
4. $\Phi_{MCSCF} = C_1(3a_1 1b_1) + C_2(4a_1 2b_1)$ $C_1 = 0.9986$, $C_2 = 0.0052$	T_2	-38.91177			10^{-5}	-38.911781	2	0.006
5. DCI						-38.911763		
6. FCI						-38.911780		

TABLE VI. $HeH_2(^1A_1)$, CCMC energies and t amplitudes for RHF and MCSCF reference functions.

Reference function	Calculation type	Energy contributions (hartree)			Number of iterations	Largest t amplitudes
		E_0	E_{T_2}	E_{tot}		
1. $\Phi_{HF} = 1a_1^2 1b_2^2$	$T_2(CPMET)$[a]	-2.52846	-0.144986	-2.673449	22	
2. $\Phi_{HF} = 1a_1^2 1b_2^2$	T_2 (full)	-2.52846	-0.144986	-2.673449	4	-0.35
3. Φ_{MCSCF}: four configurations within the model orbital space of $\{1a_1 2a_1, 1b_2\}$	T_2	-2.67374	-0.00523	-2.678975	3	-0.17
4. Φ_{MCSCF}: eight configurations within the model orbital space of $\{1a_1, 2a_1, 3a_1, 1b_2\}$	T_2	-2.67962	-0.00262	-2.682252	2	-0.009
5. Φ_{MCSCF}: eight configurations within the model orbital space $\{1a_1, 2a_1, 3a_1, 1b_2\}$	T_2 (full)	-2.67962	-0.00269	-2.682322	2	-0.009
6. FCI				-2.683104		

[a]Reference 19.

final energy is found to be within 10^{-6} hartree of the corresponding FCI (with the *same* frozen core) value.

D. $HeH_2(^1A_1)$

The CCMC calculations on $HeH_2(^1A_1)$ were performed using an approximately double-zeta quality basis set (H: $2s$/He: $3s$) of Gaussian functions[25] at the C_{2v} geometry $\theta_{H-He-H} \approx 170°$, $R_{He-H} = 1.0$ Å. Certainly, the lowest energy configuration of HeH_2 is of $1a_1^2 2a_1^2 (^1A_1)$ symmetry at large He-to-H_2 distances. However, at near-linear geometries with He inserted between the two H atoms, the lowest energy configuration is of $1a_1^2 1b_2^2(^1A_1)$ symmetry. This configuration involves the bonding $1s_H + 1s_{He} + 1s_{H'}$ interaction which gives the $1a_1$ orbital and the nonbonding $1s_H - 1s_{H'}$ interaction. We chose to carry out test CCMC calculations near the linear geometry where the $1a_1^2 1b_2^2$ configuration dominates the wave function. For this four-electron closed-shell system, we can compare the performance of our CCMC theory with the results of the CPMET calculations obtained independently for us by Krishnan.[19] Our CCMC theory for a closed-shell *single*-determinant RHF reference function reduces to the CPMET model when the three-body density matrix elements are included, and when we take $T \cong T_2$. As the results of Table VI (rows 1 and 2) indicate, the CPMET energies obtained in our work and those of Krishnan are identical. However, when the CCMC reference function is improved (but the three-body density matrices are neglected) to include the $(1b_2^2 - 1a_1'^2)$ configuration (which had the largest t amplitude in the single-configuration CC study), the CCMC energy (row 3) is found to lie 0.006 a.u. below the CPMET value. The reason for this improvement is that T operators when acting on a larger reference function give higher-order correlation effects in the CCMC equations. To underscore this point and the inherent flexibility of the CCMC method, we also included in the reference space configurations of the form $(1b_2^2 - a_1'a_1'')$ corresponding to the largest t amplitudes (0.17) occurring in the above CCMC calculation. The resulting CCMC calculation (row 4) reproduced the FCI energy to within 0.0008 a.u. Using this prescription

for successively expanding the reference function's configuration list, one can progressively improve the quality of the CCMC calculation. In row 5 of Table VI, we give the CCMC energy, which is obtained using the same MC reference function as used for row 4 *but* with a full evaluation of the three-body density matrices, which, in row 4, were neglected. The fact that the energies of rows 4 and 5 are identical to within 7.0×10^{-5} a.u. further supports the neglect of the higher-body density matrices.

The main conclusions to be drawn from these *model* CC calculations can be summarized as follows. The RHF configuration reference state gives good convergence of the (NR) iterative solution to the (quadratic truncated) CC equations only when the RHF function is the only dominant configuration in the correlated state. When there exist configurations outside the reference space that give rise to large (~ 0.2) t amplitudes, the equations determining the CC amplitudes are poorly represented within the truncated commutator [see Eq. (6)] expansion. As a result, the truncated quadratic equations may not yield meaningful results (no matter how one iteratively solves them). An MCSCF reference function which includes *all* dominant configurations gives CC energies that are better than the RHF-based (quadratic truncated) CC energies, produces fewer convergence difficulties, and yields T_1 contributions which are often negligible (because the GBT is nearly obeyed). A CI-type reference function (which also includes all the dominant configurations) is also a potentially good starting point. However, the T_1 contribution may not be small in this case.

E. Lowest singlet and triplet states of methylene: Moderate basis

Based on the experience gained from the above model calculations, we applied the CCMC method to calculate the adiabatic splitting between the lowest $^3B_1$ and $^1A_1$ states of CH_2, using a larger basis set. This choice of system is motivated by the large number of theoretical calculations already existing in the literature,[20-22]

TABLE VII. $CH_2(^1A_1, ^3B_1)$, CCMC energies and t amplitudes for RHF and MCSCF reference functions.

$CH_2(^1A_1)$: Reference wave function	Calculation type	Energy contributions (hartrees)					Number of iterations	Largest t amplitudes
		E_0	E_{T_1}	$E_{T_1-T_1}$	E_{T_2}	E_{tot}		
1. Φ_{HF}	T_1	-38.876642	-10^{-7}	$+10^{-9}$		-38.876642	2	10^{-3}
2. Φ_{HF}	T_2	-38.876642			-0.02721	-38.903853	4	0.05
3. $\Phi = C_1 3a_1^2 + C_2 1b_1^2$, MCSCF $C_1 = 0.980278,\ C_2 = -0.197622$	T_1	-38.897532	-10^{-7}	-10^{-8}		-38.897532	2	10^{-3}
4. $\Phi = C_1 3a_1^2 + C_2 1b_1^2$, MCSCF $C_1 = 0.980278,\ C_2 = -0.197622$	T_2	-38897532			-0.01153	-38.909061	3	0.06
$CH_2(^3B_1)$:								
1. $\Phi_{HF} = 3a_1 1b_1$	T_1	-38.916984	-10^{-7}	$+10^{-8}$		-38.916984	2	10^{-3}
2. $\Phi_{HF} = 3a_1 1b_1$	T_2	-38.916984			0.00618	-38.923174	3	0.04
3. $\Phi = C_1 3a_1 1b_1 + C_2 4a_1 2b_1$, MCSCF $C_1 = 0.99866,\ C_2 = -0.051702$	T_2	-38.902570			0.00031	-38.923722	3	0.08

which allows us to gauge the performance of our CCMC method. In addition, since the $^1A_1$ state is closed shell and the $^3B_1$ state is open shell, it gives us a chance to compare our method with others which use RHF and UHF descriptions for these two states. We employed the $(4s, 2p, 1d/2s)$ contracted Gaussian basis set used by Kenney et al.[22] for their MBPT calculation on CH_2. The d-polarization functions have been shown to be important for the description of many such carbenes.[20-22] The equilibrium geometries used for the $^1A_1$ and $^3B_1$ states were those determined by Bauschlicher and Shavitt[20] ($^1A_1: \theta_{HCH} = 102.4°$, $R_{CH} = 1.116$ Å, $^3B_1: \theta_{HCH} = 132.4°, R_{CH} = 1.082$ Å). The triplet RHF wave function can be described in terms of the C_{2v} symmetry of the molecule as $\Phi_{RHF}(^3B_1) = |0\rangle\, 3a_1 1b_1$ and $\Phi_{RHF}(^1A_1) = |0\rangle 3a_1^2$, with the core chosen to be $|0\rangle = |1a_1^2 2a_1^2 1b_2^2\rangle$. It has been shown in the calculations of Meadows and Schaefer[21] and Bauschlicher and Shavitt[20] that the correlation is almost identical for both of these states and for the above choice of core. Hence, the decision to treat the $1a_1$, $2a_1$, and $1b_2$ orbitals as frozen and doubly occupied is probably quite valid. We could, of course, simply decrease the definition of "core" to include only the $1a_1(1s_c)$ to test this hypothesis; however, Refs. 20 and 21 already did this for us. Finally, because there are only two active electrons in both the $^1A_1$ and $^3B_1$ states, we have no three-body density matrices to worry about.

With the choice of core and valence orbitals made, CCMC calculations were performed using both RHF and MCSCF reference states. In performing the calculations whose results are displayed in Table VII, we included six a_1, three b_1, three b_2, and one a_2 orbitals in the active space (i.e., the space in which T operates).

For the $^1A_1$ state, use of the $3a_1^2$ RHF reference configuration in the CC theory yields a large t_2 amplitude, corresponding to the double excitations $3a_1^2 \rightarrow b_1'^2$.[2] This $(b_1')^2$ configuration was subsequently appended to the MCSCF reference wave function, thereby leading to a two-configuration MCSCF reference function $C_1 3a_1^2 + C_2 1b_1^2$, with $C_1 = 0.98$, $C_2 = -0.20$, whose CCMC energy lies 0.006 a.u. below that obtained when the RHF reference function was used.

Analogous RHF- and MCSCF-based CC calculations

were performed on the $^3B_1$ state. To assure that all of the important configurations of the $^3B_1$ state are included in the reference space, we carried out an RHF-based CC calculation and found (see Table VII) the largest t amplitude (0.04) to correspond to the $3a_1 1b_1 \rightarrow a_1' b_1'$ excitation. The resultant two-configuration MCSCF calculation $(C_1 3a_1 1b_1 + C_2 4a_1 2b_1)$ gave a relatively weakly occupied second configuration $(C_2 = -0.05)$ and led to a negligible decrease (0.3 kcal/mol) in the CCMC energy.

Table VIII shows the singlet–triplet energy splittings obtained in our calculations as well as those obtained by other authors. Our converged CCMC value for the $^1A_1$–$^3B_1$ splitting (9.2 kcal/mol) can be compared to the value of 13.1 kcal/mol obtained by Kenney et al.[22] using many-body perturbation theory (MBPT) with the same basis set. Part of this discrepancy of 3.9 kcal/mol is due to the spin symmetry breaking in Kenney's spin-unrestricted MBPT calculation for the triplet state. Some of the difference is also due to the fact that we are using a MCSCF reference function, whereas Ref. 22 used a single configuration based MBPT. Because our single (RHF) configuration based CC energy splitting is 12.1 kcal/mol, it may be more appropriate to compare this result to the 13.1 kcal/mol obtained in Ref. 22. We would then ascribe, at most, 1.0 kcal/mol to the spin contamination (UHF) problems. Of course, a more detailed comparison is impossible because the MBPT calculations of Ref. 12 did not include the nonlinear T_2^2

TABLE VIII. $CH_2(^3B_1 - {}^1A_1)$ difference energies.

	ΔE (kcal/mol)
ΔRHF	25.31
$\Delta[^3B_1(RHF) - {}^1A_1(2MC)]$	12.20
$\Delta[^3B_1(2MC) - {}^1A_1(2MC)]$	14.45
$\Delta[^3B_1(T_2:\ RHF) - {}^1A_1(T_2:\ RHF)]$	12.12
$\Delta[^3B_1(T_2:2MC) - {}^1A_1(T_2:2MC)]$	9.20
Reference 20	10.6
Reference 21	10.9
Reference 22	13.1

terms which are present in our RHF (and MC) based CC results.

CONCLUSION

We have demonstrated that our CCMC procedure renders itself a practical tool for the study of electronic energy levels of molecular systems. By admitting an MCSCF reference function, the method allows one to treat on equal footing configurations which are essential at any geometry along the potential energy surface or reaction path. As we stressed in Sec. I, the ability to do so is especially important when one wishes to treat concerted reactions in which reaction barriers arise from avoided configuration crossings.

ACKNOWLEDGMENTS

[1]J. A. Pople, J. S. Binkley, and R. Seeger, Int. J. Quantum Chem. Symp. 10, 1 (1976).

[2]R. J. Bartlett, I. Shavitt, and G. D. Purvis, J. Chem. Phys. 71, 281 (1979).

[3]F. Coester, Nucl. Phys. 7, 421 (1958).

[4]J. Cizek, J. Chem. Phys. 45, 4256 (1966); Adv. Chem. Phys. 24, 35 (1969).

[5]O. Sinanoglu, J. Chem. Phys. 36, 706, 3198 (1962).

[6]A. Banerjee and J. Simons, Int. J. Quantum Chem. 14, 207 (1981).

[7]D. Mukherjee, R. J. Moitra, and A. Mukhopadhyay, Mol. Phys. 6, 1861 (1975); 33, 955 (1977).

[8]I. Lindgren, Int. J. Quantum Chem. Symp. 11, 403 (1977); S. Salomonson, I. Lindgren, and A. Martensson, Phys. Scr. 21, 351 (1980).

[9]J. Moshinsky, Group Theory and the Many-Body Problem (Gordon and Beach, New York, 1968).

[10]J. Paldus, in Group Theoretical Methods in Physics, edited by W. Beiglboeck, A. Bohm, and E. Takasugi (Springer, Heidelberg, 1979).

[11]I. Shavitt, Int. J. Quantum Chem. Symp. 12, 5 (1978); Annual Report on New Methods in Computational Quantum Chemistry and Their Application on Modern Super Computers, Battelle Columbus Laboratories (1979).

[12]J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. ibid. Quantum Chem. 14, 545 (1978); R. J. Bartlett and C. D. Purvis, ibid. 14, 561 (1978).

[13](a) J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. Symp. 13, 225 (1979); G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. 75, 1284 (1981); (b) At the NRCC sponsored workshop on perturbation theory held in July of 1981, Dr. G. D. Purvis reported how one can use the so-called reduced linear equations (RLE) method to solve the full quartic Eq. (6) for the t amplitudes. As mentioned earlier in the present paper, the full quartic equation must be solved whenever one suspects that the t amplitudes are large ($\gtrsim 1.0$). Such will, of course, often be the case in single configuration based CC theories (e.g., as in our $Be + H_2$ example). Given that one can, as Purvis has shown, write down the quartic Eq. (6) to be solved, it is still essential that one employ a numerical algorithm which will converge to the desired solution. The RLE method advocated by Purvis is indeed attractive. However, it requires, as does any iterative method, that one make a reasonable "guess" for the t amplitudes to begin the iteration procedure. We admit that one can perhaps guess the t amplitudes to use in a single configuration CC theory by looking at the results (expansion coefficients) of a small CI or MCSCF calculation. The RLE method might then converge to the proper result. However, we simply prefer to begin with the large contributions to the true wave function in our reference function so that our t amplitudes will be small as a result of which we can truncate Eq. (6) at its quadratic term.

[14]A. Banerjee and F. Grein, J. Chem. Phys. 66, 1054 (1977); R. Shepard, A. Banerjee, and J. Simons, J. Am. Chem. Soc. 101, 6174 (1979).

[15]B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968).

[16]D. Mukherjee, Chem. Phys. Lett. 79, 559 (1981).

[17]T. H. Dunning, J. Chem. Phys. 53, 2823 (1970).

[18]These solutions could not be reached by starting from the solution of the linear equation $\mathbf{A} + \mathbf{Bt} = 0$ as the initial guess.

[19]R. Krishnan (private communication).

[20]C. W. Bauschlicher and I. Shavitt, J. Am. Chem. Soc. 100, 739 (1978).

[21]J. H. Meadows and H. F. Schaefer, J. Am. Chem. Soc. 98, 4383 (1976).

[22]J. W. Kennery, J. Simons, G. D. Purvis, and R. J. Bartlett, J. Am. Chem. Soc. 100, 6930 (1978), and references therein.

[23]R. P. Hosteny, T. H. Dunning, Jr., R. R. Gilman, A. Pipano, and I. Shavitt, J. Chem. Phys. 62, 4764 (1975).

[24]By RHF wave function for open-shell systems, we mean a symmetry adapted open-shell configuration Φ which is constructed from orbitals that give minimum energy $E = \langle \Phi | H | \Phi \rangle / \langle \Phi | \Phi \rangle$.

[25]The basis set consisted of uncontracted s-type gaussian functions with exponents 10.0, 0.3, and 0.07 for the He atom and 5.0 and 0.8 for the H atom.