A potentially size-consistent multiconfiguration based coupled electron pair approximation

Mark R. Hoffmann[a] and Jack Simons
Department of Chemistry, University of Utah, Salt Lake City, Utah 84112

(Received 24 June 1988; accepted 30 November 1988)

A coupled electron pair approximation is derived and illustrative calculations are presented. The present approximation, which we refer to as the unitary coupled electron pair approximation (UCEPA), provides multiconfigurational (MC) reference capability and is as computationally tractable as multireference configuration interaction (MRCI). The method is capable of yielding size-consistent energies if the MC reference function is of the complete active space (CAS) variety. The coefficient matrix of the resultant set of simultaneous linear equations is evaluated using internal/external orbital space partitioning within a unitary group approach (UGA) treatment of the state space. We demonstrate the accuracy of the method on several small benchmark molecules for which full CI results are known, and on nontrivial studies on the singlet–triplet splitting in methylene and the electron affinity of the oxygen atom.

I. INTRODUCTION

In the present paper, we develop and implement a method which is similar in structure to the highly successful single-configuration based coupled electron pair approximation[1,2] (CEPA) but which admits multiconfigurational (MC) reference functions and which is based on a unitary ansatz. We call this our unitary CEPA (UCEPA) method. Although this development is by no means the first attempt to include MC reference functions within CEPA-like methods,[3] we believe that the computational efficiency of our approach make it especially attractive. Our present development expresses all computationally necessary quantities as matrix elements of the Hamiltonian in the usual Gelfand–Tsetlin basis utilized in unitary group approach (UGA) configuration interaction (CI) calculations.[4–6]

Section II details our theoretical developments, with the Appendix providing the proof that our final energy expression is size consistent if the so-called MC reference function is of the complete active space (CAS) type. In Sec. III, we report the results of our calculations on several small molecules which have been examined at the full-configuration interaction (CI) level by Handy, Schaefer, and co-workers[7]; in addition, we compare our results on the singlet–triplet splitting in CH_2 with the benchmark results of Bauschlicher and Taylor.[8] The UCEPA method is shown to yield quite accurate energies even in cases in which strong configuration mixing occurs. We also report the results of calculations with the UCEPA method on the singlet–triplet splitting of methylene using basis sets containing 48 and 71 contracted Gaussian-type basis functions, and we examine the electron affinity of the oxygen atom using a moderately large basis set optimized for correlated studies. Section IV summarizes our results.

II. THEORY

A. The UCEPA ansatz

In our ansatz, corrections to the multiconfiguration reference function $|\Phi\rangle$ are expressed in terms of the action of a unitary operator $\exp(G)$ acting on $|\Phi\rangle$

$$|\psi\rangle = \exp G\,|\Phi\rangle\ . \tag{2.1}$$

The anti-Hermitian operator G is, in turn, expressed in terms of a set of state transfer operators[9] $\{|n\rangle\langle\Phi| - |\Phi\rangle\langle n|\}$ and corresponding amplitudes $\{\tau_n\}$:

$$G = \sum_n \tau_n (|n\rangle\langle\Phi| - |\Phi\rangle\langle n|)\ . \tag{2.2}$$

The summation over n in Eq. (2.2) runs, *in principle*, over all CSFs which are *external* to the MCSCF space (i.e., lying outside the MCSCF space) and over the *orthogonal complement* states within the MCSCF space (i.e., *combinations* of the CSFs from which $|\Phi\rangle$ is formed which are orthogonal to $|\Phi\rangle$ and to one another). Restriction of the operator manifold in Eq. (2.2) to the so-called first-order interacting space can clearly be effected if one desires to do so. Alternatively, one may wish to allow a larger space (e.g., the so-called second-order space: those CSFs which do not have more than two electrons in the MCSCF virtual orbital space and more than two holes in the MCSCF core orbital space) to be included in the aforementioned summation. No particular choice of configuration state function space is dictated by the functional form of our theory. As within conventional multireference CI methods, the choice of configurations is left up to the individual researcher. The wisdom of any particular choice must be verified through computational means. Later, in Sec. III D, we illustrate this point by comparing the effects of using first- and second-order spaces in an investigation of the singlet–triplet splitting in methylene.

Given a particular choice of MCSCF reference function $|\Phi\rangle$ and the space of CSFs lying external and internal to the MC reference space, the parameterization of $|\psi\rangle$ provided in Eq. (2.1) is no more flexible than that contained in the con-

[a] Present address: Department of Chemistry, University of North Dakota, Box 7185 University Station, Grand Forks, ND 58202.

 0021-9606/89/073671-09$02.10

ventional configuration interaction (CI) wave function expression $|\psi\rangle = C_0|\Phi\rangle + \Sigma_n C_n|n\rangle$. However, the particular parameterization introduced here allows us to:

(1) Expand the C_0 and C_n amplitudes in terms of their "starting" MCSCF-level values ($C_0 = 1, C_n = 0$) plus "improvements" ($\delta C_0, \delta C_n$);

(2) Through the unitary nature of $\exp(G)$, maintain normalization of $|\psi\rangle$ without additional constraint equations being imposed;

(3) Achieve a final energy expression which is size consistent *if* the MC reference function $|\Phi\rangle$ is of the complete-active-space (CAS) variety (i.e., if the CSFs used to form $|\Phi\rangle$ are constructed by distributing M electrons among V valence orbitals in *all* possible ways consistent with the desired spin and space symmetry).

Our strategy is to use a variational energy expression $\langle\psi|H|\psi\rangle$ in which the variational parameters $\{\tau_n\}$ are contained in G, to expand $\langle\psi|H|\psi\rangle$ through second order in the $\{\tau_n\}$ to focus on a solution close to $|\Phi\rangle$ and its associated energy E_0, to then minimize the resultant second-order functional with respect to the $\{\tau_n\}$, and finally to use those "optimal" $\{\tau_n\}$ values to evaluate second-order energy corrections to E_0.

Substitution of Eq. (2.2) into the expression for the energy

$$E = \langle\Phi|\exp(-G)H\exp(G)|\Phi\rangle ,$$

followed by expansion of the exponential operators, gives[9]

$$E = [\cos^2 t\,]E_0 + 2\cos t\,\frac{\sin t}{t}\sum_n \tau_n\langle\Phi|H|n\rangle + \frac{\sin^2 t}{t^2}\sum_{n,m}\tau_n\langle n|H|m\rangle\tau_m , \tag{2.3}$$

where $t^2 = \Sigma_n \tau_n^2$. Thus far, our Eq. (2.3) is *entirely* equivalent to the variational energy expression for a configuration interaction wave function in a configuration space which spans the $\{|\Phi\rangle, |n\rangle\}$ functions.

We now introduce approximations through the manner by which we solve for the $\{\tau_n\}$ amplitudes whose values characterize our wave function and in the manner in which we estimate the energy of Eq. (2.3). Because the MCSCF wave function, by assumption, contains all of the essential configuration state functions (CSFs) and has energy-optimized MCSCF orbitals and CSF amplitudes, *all* of the τ_n amplitudes should be small. We are therefore motivated to seek solutions which lie close to the origin in τ space (i.e., values of τ_n near zero). To do so, we carry out a Taylor-series expansion of E about $\tau = 0$ and keep terms through second order in the $\{\tau_n\}$ parameter. Such an expansion of Eq. (2.3) gives

$$E = E_0\left(1 - \sum_n \tau_n^2\right) + 2\sum_n \tau_n\langle\Phi|H|n\rangle + \sum_{nn'}\tau_n\tau_{n'}\langle n|H|n'\rangle . \tag{2.4}$$

Eq. (2.4) is easily recognized to be of the functional form which appears in CEPA theory[1,2]; it is our MC generalization of the CEPA-like energy functional, and it forms the basis of our further developments. We think of this expression as providing a local (to $\tau = 0$) quadratic energy functional [Eq. (2.4)], (i) which approximates the correct CI energy of Eq. (2.3) through second order in the $\{\tau_n\}$, (ii) whose minimization leads to a set of linear equations (see below) for the $\{\tau_n\}$, (iii) whose evaluation at the resultant optimal $\{\tau_n\}$ values yields (see below) an expression for E which contains only linear terms in the $\{\tau_n\}$ amplitudes and which is size consistent if $|\Phi\rangle$ is a CAS-type MCSCF wave function (as shown in the Appendix).

B. UCEPA linear equations

Differentiating Eq. (2.4) with respect to the τ_n amplitudes and requiring this local quadratic energy to be stationary[10] gives a set of simultaneous linear equations of the unknown τ_n amplitudes,

$$-\langle\Phi|H|n\rangle = \sum_m \tau_m(\langle n|H|m\rangle - \delta_{mn}E_0) \tag{2.5a}$$

or symbolically,

$$-B_n = \sum_m A_{nm}\tau_m . \tag{2.5b}$$

The energy functional of Eq. (2.4), when evaluated using the values of $\{\tau_n\}$ which result from Eq. (2.5b), reduces to

$$E = E_0 + \sum_n \tau_n\langle\Phi|H|n\rangle . \tag{2.5c}$$

Because $|\Phi\rangle$ has itself been variationally optimized in the space spanned by $|\Phi\rangle$ and those $|\tilde{n}\rangle$ in the orthogonal complement space, off diagonal elements $\langle\tilde{n}|H|\Phi\rangle$ in the OCS space vanish. Thus, only those τ_n and $\langle\Phi|H|n\rangle$ which involve states $\{|n\rangle\}$ in the *external* space contribute to E.

The energy expression of Eq. (2.5c) and the form of the equations which determine the τ_n amplitudes deserve further elaboration. First, we note that the constituents of the matrices in Eq. (2.5) are identical to those required in a CI calculation performed within the same configuration state space, with the exception that the eigenvector and energy eigenvalue of the MCSCF *reference* function $|\Phi\rangle$ must be known beforehand in the case of the UCEPA calculation. We also note that, in the single-reference-configuration case, Eq. (2.5a) is identical to the cluster amplitude equations obtained in linearized coupled-cluster theory[1-3] (i.e., identical to the CEPA-0 equations). This further establishes a tie to CEPA theory. Finally, we note that the total energy E does not appear in Eq. (2.5a) which, because the CAS-MCSCF references energy E_0 is size extensive, allows the proof given in the Appendix to show that the equations for the τ_n are decoupled when applied to noninteracting subsystems. This, in turn, implies that the total energy expression of Eq. (2.5c) is additively separable and hence size consistent.

A complication seems to arise in the implementation of Eq. (2.5) due to the apparent need for the orthogonal complement MCSCF states. However, this can be circumvented using the projection technique introduced by Lengsfield[11] within the content of MCSCF wave function optimization. As a result, in the configuration state function (CSF) basis, denoted by capital roman letters, the matrices A and B enter-

ing into the simultaneous equations for the cluster amplitudes become

$$A_{I,J} = \langle I|H|J\rangle - E_0\delta_{IJ} + C_I C_J\,,$$
for I,J within the internal CSF space, (2.6a)

$$A_{I,J} = \langle I|H|J\rangle - C_I\langle J|H|\Phi\rangle,$$
for I internal and J external, (2.6b)

$$A_{I,J} = \langle I|H|J\rangle - E_0\delta_{IJ}, \quad \text{for } I \text{ and } J \text{ external}\,, \tag{2.6c}$$

$$B_I = 0, \quad \text{for } I \text{ internal}\,, \tag{2.7a}$$

and

$$B_I = \langle\Phi|H|I\rangle, \quad \text{for } I \text{ external}\,. \tag{2.7b}$$

The notation "internal" denotes those CSFs used to form $|\Phi\rangle$, while "external" labels those CSFs which do not appear in $|\Phi\rangle$.

Solution of the set of simultaneous equations (i.e., $-B_I = \Sigma_J A_{IJ}\tau_J$) allows the local quadratic energy [cf. Eq. (2.4)] to be written in the linear form analogous to Eq. (2.5c),

$$E = E_0 + \sum_I \tau_I\langle\Phi|H|I\rangle\,, \tag{2.8}$$

where the index I runs only over *external* CSFs because contributions from the MCSCF orthogonal complement space vanish. As shown in the Appendix, this energy expression is size extensive if $|\Phi\rangle$ is a CAS-type MCSCF function.

C. Computational aspects

In the preceding section, the coefficients of the set of simultaneous linear equations that define the amplitudes τ were expressed in terms of Hamiltonian matrix elements *in the CSF basis* modified by, at most, a simple product of (CSF basis) vector components. It is perhaps obvious that large-matrix techniques used efficaciously for the CI calculations can be straightforwardly adapted to the present case. In particular, the coefficient matrix A is sufficiently diagonally dominant that an iterative solution converges rapidly; moreover, the A_{IJ} coefficients are computationally simple enough to allow for regeneration as needed.

We solve the set of simultaneous linear equations for the amplitudes using a Galerkin, or reduced linear equation, framework.[12–14] Because this method has received considerable attention in the chemistry literature, our discussion will be brief. The k th approximation to the solution vector, $\tau^{(k)}$, is expanded in an *orthonormal* basis $\phi^{(i)}$

$$\tau^{(k)} = \sum_{i=1}^{k} a_i^{(k)}\,\phi^{(i)}\,, \tag{2.9}$$

where the basis vectors ($\phi^{(i)}, i = 1,k$) are known. The coefficients a_i are determined through solution of the simultaneous linear equations induced by projection on the subspace,

$$-(\mathbf{B},\phi^{(j)}) = \sum_i^k a_i^{(k)}(\mathsf{A}\phi^{(i)},\phi^{(j)})\,. \tag{2.10}$$

At this point, either $\tau^{(k)}$ is accepted as the converged solution τ or a new basis vector $\phi^{(k+1)}$ is generated via Jacobi prediction

$$\phi^{(k+1)} = \mathsf{D}^{-1}\tau^{(k)} \tag{2.11}$$

followed by Gram–Schmidt orthonormalization of $\phi^{(k+1)}$ to the extant $\phi^{(k)}$ *and* to the MC solution vector **C**. In Eq. (2.11), the matrix A is decomposed into diagonal D and off-diagonal Δ parts, i.e., $\mathsf{A} = \mathsf{D} + \Delta$. The next row and column (i.e., $k+1$) of the reduced matrix [cf. Eq. (2.10)] is then generated by matrix multiplication of the full coefficient matrix and the new basis vector (i.e., $\mathsf{A}\phi$, followed by appropriate inner products). In practice, the diagonal and off-diagonal matrix products are computed separately.

The use of orthonormal basis vectors $\phi^{(i)}$ is advantageous for two reasons: (1) the reduced coefficient matrix [i.e., $(\mathsf{A}\phi^{(i)}, \phi^{(j)})$] remains nonsingular as convergence is approached; and (2) the modifications to $\Delta\phi^{(i)}$ due to non-Hamiltonian terms is minimal. In particular, no modification is necessary for vector components in the external CSF space,

$$(\Delta\phi^{(i)})_E = \Sigma'_J\langle E|H|J\rangle\phi_J^{(i)}\,; \tag{2.12a}$$

while for the internal CSF space,

$$(\Delta\phi^{(i)})_I = \Sigma'_J\langle I|H|J\rangle\phi_J^{(i)} - C_I\left(C_I\phi_I^{(i)} + \sum_J^{\text{ext}}\langle\Phi|H|J\rangle\phi_J^{(i)}\right). \tag{2.12b}$$

In Eq. (2.12), the prime on the summation indicates that $J = I$ (or E) is to be excluded. Hence, the matrix-vector products in UCEPA are identical to the matrix-vector products in CI, with the additional requirement that a small number of components need minor postmultiplication processing.

The factorization of Hamiltonian matrix elements between CSFs with no more than two electrons in the virtual orbital space into the product of an internal and external contribution was first realized by Roos in 1972.[15] Henceforth, many CI codes calculate and store only the internal contributions to such Hamiltonian matrix elements, completing the external portions as needed on each construction of a new correction vector σ. Within the unitary group approach (UGA), such factorization was made possible with Shavitt's so-called segment values.[16] As noted by Siegbahn in the first direct CI (i.e., using internal/external partitioning) from a general MC reference,[17] very little evidence of the particular method used to evaluate the internal Hamiltonian matrix elements is present in the iterative solution cycle (i.e., the time-consuming step of the direct CI). Hence, we choose to vary from the usual graphical implementation of the unitary group approach (GUGA), and instead retain the so-called step vector representation over internal portions of the CSFs (see, e.g., the review by Shavitt, Ref. 5). The use of step vectors facilitates the description of MC reference spaces of the generalized CAS (also called restricted CI) variety (i.e., those CSF spaces in which specified numbers of electrons are distributed among specified sets of orbitals in *all* possible ways).

Evaluation of internal contributions to the off-diagonal generator and generator product matrix elements occurs before the start of the iterative solution. External storage (e.g.,

TABLE I. Results for water at various bond lengths.

	Energy (a.u.)		
	$1.0R_e$	$1.5\,R_e$	$2.0\,R_e$
MC[a]	− 76.062 878	− 75.924 342	− 75.827 219
CISD	− 76.155 840	− 76.012 305	− 75.903 264
UCEPA	− 76.158 648	− 76.015 355	− 75.905 610
FCI[b]	− 76.157 866	− 76.014 521	− 75.905 247
MRLCCM[c]	− 76.157 90	− 76.013 73	− 75.904 15
UCC[d]	− 76.156 234	− 76.013 827	− 75.905 141

[a] $1a_1^2 2a_1^2 1b_1^2 (3a_1 4a_1 1b_2 2b_2)^4$.
[b] Reference 7.
[c] The multireference linearized coupled-cluster results by Laidig and Bartlett (see Ref. 21).
[d] The so-called unitary coupled-cluster theory result of Ref. 32, based on the same MC reference function and orbital basis.

disk) of the so-called *internal* formula tape is not a limiting factor. We utilized Payne's 1982 segment values[6] for the internal contribution to matrix elements. We derived and implemented external contributions which are consistent with Payne; readers interested in greater detail are welcome to contact the authors directly.

Construction of the requisite matrix elements for our UCEPA theory and the solution of the (large) matrix problem have been shown to proceed analogously to the corresponding CI problem. In particular, internal/external partitioning in a UGA framework and a computational Galerkin method are realized. Thus, in principle, any problem that is computationally accessible to a modern CI program is also computationally tractable by our UCEPA method. Of course, the relative accuracy of UCEPA must be judged from chemical studies as the following section attempts to address.

III. RESULTS

We performed illustrative calculations using our UCEPA method on H_2O, HF, NH_3, and BH, using the same basis sets and geometries as used in the benchmark full-CI calculations of Handy, Schaefer, and co-workers.[7] We also studied the singlet–triplet separation in CH_2 using the basis sets and geometries of Bauschlicher and Taylor.[8] Our results for H_2O, BH, and CH_2 are summarized in Tables I–III, respectively; our HF and NH_3 results are not reported here because they were essentially identical, in their comparison with the full-CI data, to those reported for H_2O and BH.

To demonstrate the applicability of our UCEPA method to more realistic chemical studies, we also examined the singlet–triplet separation in CH_2 using two larger atomic basis sets, and we examined the electron affinity of the oxygen atom, a well-known difficult system,[18–20] using a moderately large unpublished basis of Taylor.[18] The results of our studies of methylene using the larger basis sets are presented in Table IV and V. Table VI summarizes the results for the oxygen atom.

A. Water

A double-zeta basis was used for calculations on the lowest $^1A_1$ state of H_2O at three nuclear geometries. The MCSCF reference function is identical to that used by Laidig and Bartlett[21] in their multireference linearized coupled cluster method (MR-LCCM) studies, and corresponds to distributing four electrons among two a_1 orbitals and two b_2 orbitals while keeping two a_1 orbitals and one b_1 orbital doubly occupied in all configurations. This procedure yields 12 CSFs of $^1A_1$ symmetry in the internal space. A full spin configuration singles and doubles (i.e., second-order) space was chosen for this example; this yields a total of 4141 CSFs. The second-order space is somewhat larger than the approximate singles and doubles interacting (or first-order) space used

TABLE II. Results for boron hydride.

	(Valence correlation)[a] Energy (a.u.)	(Core + valence correlation) Energy (a.u.)
SCF		− 25.125 260
MC[b]		− 25.177 887
CISD	− 25.212 910	− 25.227 092
UCEPA	− 25.213 632	− 25.228 119
FCI	− 25.213 185	− 25.227 627[c]

[a] The canonical orbitals with the lowest and highest eigenvalues of the core Fock operator were kept doubly occupied and unoccupied, respectively, in these calculations.
[b] $1a_1^2 (2a_1 3a_1 4a_1 1b_1 1b_2)^4$.
[c] Reference 7.

TABLE III. Results for methylene in the (421/21) basis.

	$E(^3B_1)$ (a.u.)	$E(^1A_1)$ (a.u.)	Δ (kcal/mol)	Error Δ (kcal/mol)
MC[a]	− 38.927 947	− 38.907 660	12.73	0.76
CISD[b]	− 39.041 602	− 39.022 156	12.20	0.23
UCEPA	− 39.047 715	− 39.028 718	11.92	− 0.05
FCI[b]	− 39.046 260	− 39.027 183	11.97	...

[a] SCF reference for $^3B_1$: $1a_1^2 2a_1^2 1b_2^2 3a_1 1b_1$; TCSCF reference for $^1A_1$: $1a_1^2 2a_1^2 1b_2^2 (3a_1 1b_1)^2$.
[b] Reference 8.

TABLE IV. Results for methylene in the (642/32) basis.

	$E(^3B_1)$ (a.u.)	$E(^1A_1)$ (a.u.)	Δ(kcal/mol)
MC[a]	− 38.932 792	− 38.914 999	11.16
CISD[b]	− 39.062 945	− 39.046 041	10.61
UCEPA[b]	− 39.070 540	− 39.053 845	10.48
CISD[c]	− 39.063 638	− 39.046 326	10.86
UCEPA[c]	− 39.071 505	− 39.054 226	10.84

[a] See footnote a of Table III.
[b] Calculation using first-order (i.e., interacting) space.
[c] Calculation using second-order space.

TABLE V. Results for methylene in the (9621/52) basis.

	$E(^3B_1)$ (a.u.)	$E(^1A_1)$ (a.u)	Δ(kcal/mol)
MC[a]	− 39.933 428	− 38.916 581	10.57
CISD[b]	− 39.071 183	− 39.056 013	9.52
UCEPA[b]	− 39.079 372	− 39.064 528	9.32
CISD[c]	− 39.072 036	− 39.056 337	9.85
UCEPA[c]	− 39.080 559	− 39.064 964	9.79

[a] See footnote for Table IV.
[b] Same as footnote a.
[c] Same as footnote a.

by Laidig and Bartlett, which gives 3277 CSFs. However, the energy difference due to the use of these two different spaces in a CI calculation is less than 0.03 mhartree.

Examination of Table I reveals that our UCEPA method reproduces the full CI energy to within 0.5 kcal/mol for all three geometries studied, although the UCEPA energy is not truly variational (see Table I) once the exactly variational energy function of Eq. (2.3) is expanded to second order to obtain Eq. (2.4). The corresponding (i.e., singles and doubles) CI calculations are only accurate to within 1.4 kcal/mol. The UCEPA energies also lie within 1 kcal/mol of both the MR-LCCM energies of Ref. 21 and our UCC energies[32] at all three geometries.

Gdanitz and Ahlrichs[3] studied H_2O at these same geometries using their multiconfiguration-based averaged coupled-pair functional (ACPF) method. They employed a double-zeta plus polarization basis, so a direct comparison with their results is impossible. Within their basis, the FCI energies are − 76.256 62, − 76.071 41, and − 75.952 27 hartrees; their ACPF energies lie below the FCI energies by 0.12 and 0.22 mhartree at R_e and 1.5 R_e and above the FCI value by 0.18 mhartree at 2.0 R_e when they use a 12-configuration reference function. When they employ a 55-configuration reference, their respective FCI deviations are − 0.40, − 0.79, and − 0.78 mhartree. Our respective FCI deviations are − 0.78, − 0.83, and − 0.36 mhartree.

Further assessment of the UCEPA method may be made at the equilibrium bond length of water, where it is appropriate to use a single configuration reference (i.e., the dominant CSF in the 12 MC has an amplitude of 0.986). Using a single configuration reference (i.e., SCF), there are 361 CSFs in both the first- and second-order spaces. We obtain a UCEPA correlation energy (i.e., $E_{\mathrm{UCEPA}} - E_{\mathrm{SCF}}$) of 0.146 719 hartree, which may be compared to the full CI result[7] of 0.148 028 hartree. The correlation energy results of other approximation treatments include CISD: 0.140 177 hartree, SDTQ MBPT(4)[22]: 0.147 040 hartree, CCSD[22]: 0.146 240 hartree, CPF[23]: 0.145 02 hartree, UCC[32]: 0.146 396 hartree and CISDTQ[7]: 0.147 765 hartree. The UCEPA method appears to give substantial improvements over the CISD results without increase in computational effort also in cases where a single reference configuration is appropriate.

TABLE VI. Results for oxygen.

	E(O) (a.u.)	$E(O^-)$ (a.u.)	EA (a.u.)
SCF[a]	− 74.809 108	− 74.789 307	− 0.019 802
CISD	− 74.979 535	− 75.011 601	+ 0.032 066
UCEPA	− 74.987 371	− 75.029 218	+ 0.041 847
FCI[b]	− 74.985 694	− 75.028 584	+ 0.042 890

[a] Degenerate SCF result (see the text for details).
[b] Reference 18.

B. Boron hydride

For our study of BH, we used a double-zeta plus polarization basis. An MCSCF calculation involving the valence space (i.e., 2*s*, 2*p* on boron, and 1*s* on hydrogen) results in 19 internal CSFs of $^1A_1$ symmetry (i.e., using the C_{2v} point group). Examination of the CSF amplitudes for the converged MCSCF wave function shows that there are five significant configurations: $2a_1^2 3a_1^2$ (0.959), $2a_1^2 1b_{1,2}^2$ (− 0.159), and $2a_1^1 3a_1^1 1b_{1,2}^2$ (0.092). We employed the full 19 CSF MCSCF wave functions in our studies of boron hydride.

Harrison and Handy[7] report two sets of calculations for BH, one with all orbitals active in the expansion space and one with the energetically lowest and highest lying orbitals doubly occupied and unoccupied, respectively, for all expansion CSFs; we likewise examine BH in both approximations.

The second-order space corresponding to including excitations from the core and to the highest virtual orbital was spanned by 7303 CSFs. A MRCI calculation in this space gave a correlation energy ($E - E_{\mathrm{SCF}}$) of 0.101 831 hartree, compared to the full CI result[7] of 0.102 367 hartree. Our UCEPA method gives a correlation energy of 0.102 859 hartree. For this molecule then, the MRCI and UCEPA results bracket the full CI answer, each being in error by about 0.3 kcal/mol.

Restricting the $1a_1$ orbital to be doubly occupied and the $12a_1$ orbital to be unoccupied, as Ref. 7 also did, results in 912 CSFs in the second-order space. A MRCI calculation in this space gives a correlation energy of 0.087 649 hartree, and our UCEPA method gives a result of 0.088 371 hartree. We had need to perform a new full CI calculation in this case, since the calculation of Harrison and Handy used $1a_1$ and $12a_1$ orbitals from an SCF calculation. We obtained a full CI (3036 CSFs) correlation energy of 0.087 925 hartree, a value 0.161 mhartree less than the single-configuration orbitals result. Again the MRCI and UCEPA results are within 0.3 kcal/mol of the FCI result; however, in this case, the MRCI result is about 40% closer to the FCI than the UCEPA result.

C. Methylene

The $^1A_1$–$^3B_1$ splitting in methylene has been the subject of considerable theoretical and experimental work for the past two decades. Several excellent reviews[24–26] exist in the literature. The question of the appropriateness of our UCEPA method in obtaining chemically meaningful results for this system is addressed in Sec. III D. In the present section we compare the results of UCEPA calculations using a modest basis set (DZP) with exact (full CI) results.

Recently, Bauschlicher and Taylor[8] reported full CI results for the singlet–triplet splitting using a double-zeta polarization basis set, thereby providing a small, yet difficult, model problem requiring *uniform* treatment of two qualitatively different states. Heuristically, the UCEPA method is expected to work well when the *qualitative* chemistry is embodied in the MC reference wave function. The preponderance of theoretical work indicates that a single CSF suffices in the zero-order description of the $^3B_1$ state but that a two configuration wave function, $1a_1^2 2a_1^2 1b_2^2 (3a_1 1b_1)^2$ must be used to describe the $^1A_1$ state. These choices of reference wave functions lead to a second-order expansion space of 2232 CSFs for the $^3B_1$ and 1590 CSFs for the $^1A_1$ state, where the $1a_1$ orbital has been kept doubly occupied in all CSFs as in Bauschlicher and Taylor's FCI work.

The singlet–triplet energy gap using the reference wave functions (i.e., without additional correlation) is 0.020 287 hartree, a result 0.001 210 hartree, or 0.76 kcal/mol, larger than the FCI result[8] of 0.019 077 hartree. Parenthetically we note that the FCI result corresponds to $1a_1$ orbitals obtained from SCF calculations (and not from a $^3B_1$ SCF and $^1A_1$ TCSCF calculation); Bauschlicher and Taylor have also performed FCI calculations using CASSCF orbitals and report[8] a change of less than 0.01 kcal/mol in the splitting; thus the particular form of the $1a_1$ orbital appears to be unimportant at the $\gtrsim$ 0.01 kcal/mol level of accuracy. Our UCEPA prediction of the energy splitting is 0.018 998 hartree, which is in error by less than 0.05 kcal/mol.

In comparison, the error in the gap, using a first-order multireference CISD calculation[8] for the $^1A_1$ state is 0.000 369 hartree, or 0.23 kcal/mol. Consideration of the calculated separation using Davidson's correction[8] (i.e., 0.06 kcal/mol), to approximately account for disconnected clusters, supports the reliability of the UCEPA splitting and argues against a fortuitous result.

D. Methylene using larger basis sets

Accurate calculation of the $^1A_1$–$^3B_1$ splitting in CH_2 is challenging because of the differential correlation and basis set effects for the two states in question. Final (i.e., extrapolated to basis set and correlation limits) theoretical predictions for the singlet–triplet splitting in methylene were obtainable because of smooth trends in both basis set size and size of configuration state function expansion space in the CI framework. A very valid question then for a new method, such as our UCEPA, is whether systematic improvements on a prediction can be made; this is an especially relevant question for a theory such as UCEPA, which is not strictly variational. In this section, we investigate the predicted singlet–triplet energy gap using two basis sets larger than the DZP of the preceding section. The smaller of the two may be designated (642/32) and would be in the "extended basis plus polarization" (EBP) category of Ref. 24. The larger of the two, (9621/52), falls in the "very large" (VL) category of Ref. 24, and so is comparable to the very largest basis sets used to study methylene.

We pattern our investigation on the earlier work of Werner and Reinsch.[27] For the EBP study, the 11*s* primitive *C* basis of Salez and Veillard[28] was contracted as (521 111), while the 6*p* primitive basis of Huzinaga[29] was contracted as (3111). Cartesian double polarization functions (*d*-type) with exponents 1.3 and 0.4 were added to the carbon basis.[27] Huzinaga's 6*s* primitive hydrogen basis[29] was contracted according to (321) and double *p*-type polarization functions with exponents 1.0 and 0.25 were added.[27] In total, 48 CGTO's were used. The geometry used by Taylor[30] was adopted (for the $^3B_1$ state: $H=0$, $\pm 1.865\,798\,61$, 0.824 860 94; for the $^1A_1$ state: $H=0$, $\pm 1.625\,838\,28$ 1.330 733 21).

As in the benchmark CH_2 study with a DZP basis, a single reference configuration was used to describe the $^3B_1$ state and a two-configuration reference was used for the $^1A_1$ state. The calculated splitting using these (reference) wave functions was 11.16 kcal/mol. Use of a second-order CSF space in which to describe the dynamic correlation via our UCEPA formalism reduces this value to 10.84 kcal/mol. A CI calculation spanning the identical second-order space gives a splitting of 10.86 kcal/mol. The second-order space for the $^3B_1$ state has 9121 space- and spin-adapted CSFs, while the corresponding space for the $^1A_1$ state is spanned by 6327 CSFs.

An approximation to the full second-order space, (i.e., the state space spanned by all configurations with not more than two holes in the core and two electrons in the virtual orbital space) consists of those configurations in the second-order space which interact in first order with the reference wave function; i.e., those configuration state functions, $|I\rangle$, for which $\langle\Phi|H|I\rangle \neq 0$. The use of such spaces has been a most useful approximation in CI calculations. We are therefore interested in the ability of these so called first-order interacting spaces to reproduce second-order results in the context of our UCEPA method. A first-order interacting-space[31] UCEPA calculation on the singlet–triplet splitting in methylene yields 10.48 kcal/mol, which is in disagreement with the second-order calculation by 0.36 kcal/mol. 3458 CSFs were included in the approximate first-order space for the $^3B_1$ state while 4755 CSFs were used to describe the $^1A_1$ state. The equivalent first-order interacting-space CI calculation gives 10.61 kcal/mol, which is in disagreement with the second-order result by 0.25 kcal/mol.

A yet larger study of the splitting in methylene was undertaken to investigate whether the best theoretical predictions of the splitting could be reproduced by our new UCEPA method. For this, our largest study, we employed the same 11*s* *C* primitive basis as before, but we only contracted the three Gaussians with the largest exponents, giving nine *s*-type functions. The earlier 6*p* primitives of Huzinaga were replaced by the 7*p* primitives of Salez and Veillard,[28] of which only the tightest two were contracted. The same first polarization functions were added, but now a Cartesian second (*f*-type) polarization function, with exponent 0.65, was included.[27] For hydrogen, the 6*s* primitive basis of Huzinaga[29] was replaced by his 7*s*, and only the first three were contracted. *p* functions identical with the smaller basis study were used. In total, 71 CGTOs were used to describe CH_2. The geometries used are identical to those of Werner and Reinsch[27]

(for $^3B_1$: $H = 0, \pm 1.867\,341\,84, 0.813\,881\,73$;

for $^1A_1$: $H = 0, \pm 1.628\,120\,79, 1.318\,426\,22$).

As before, the reference description of the $^3B_1$ state is a single configuration, and the $^1A_1$ state is described by a TCSCF reference wave function. The reference splitting with the larger basis is reduced to 10.57 kcal/mol, a value 0.59 kcal/mol less than the result with the smaller basis set. In this large basis, the second-order spaces grew to dimensions of 21 319 and 14 637 for the $^3B_1$ and $^1A_1$ states, respectively. The singlet–triplet splitting predicted by our UCEPA method is 9.79 kcal/mol, in comparison with 9.85 kcal/mol obtained from a CI calculation in the same CSF space. Calculations in first-order interacting spaces[31] [7924 CSFs ($^3B_1$) and 10 927 CSFs ($^1A_1$)] yield results of 9.32 kcal/mol for UCEPA and 9.52 kcal/mol for CI. Again the discrepancies between first- and second-order calculations were similar for UCEPA and CI, 0.47 and 0.33 kcal/mol, respectively.

E. Electron affinity of the oxygen atom

Theoretical calculations of electron affinities (EAs) within 0.1 eV of experimental results, especially for molecules containing C, O, or F, require large basis sets (especially multiple higher polarization functions) and a description of dynamic correlation beyond the singles and doubles level in a balanced treatment of the neutral and the anion (see, e.g., the review article by Simons and Jordan[20]). The goal of our study here of the EA of the ^3P state of the oxygen atom is not to obtain the definitive value, but rather to investigate the ability of our new UCEPA method to reproduce the exact (i.e., full CI) result using nontrivial basis sets. The recently developed natural orbital based contraction procedure of Almlöf and Taylor[33] can, in general, reproduce correlated results which are obtained using much larger numbers of standard CGTOs. We use in our study Taylor's (unpublished) oxygen atom and oxygen anion contractions[18] from $14s9p6d\,5f$ to $4s3p2d\,1f$. [The (unpublished) full CI results within these basis sets were also made available to us by Taylor.[18]] Symmetry combinations of d- and f-type orbitals which produced s- and p-type orbitals, respectively, were eliminated.

A single reference (i.e., SCF) wave function was used to describe the ^3P state of the neutral as well as the ^2P state of the anion. Modifications of the molecular MCSCF programs available in this laboratory were made to ensure the degeneracy of the orbitals, yielding a correct SCF wave function. Second-order spaces of dimensionality 2052 and 1780 for the neutral and anion, respectively, were used for the correlated studies. As our UCEPA program can only make use of the symmetry relations of nondegenerate point groups, we restricted the correlated wave function to transform as an irreducible representation of D_{2h}. The oxygen $1s$ orbital was kept doubly occupied in all CSFs as it was in the FCI study.

An EA of -0.54 eV was computed using the reference wave functions (i.e., the electron is not bound at the SCF level). CI calculations in the state spaces given above yield an EA of $+0.87$ eV. The UCEPA result for the EA is $+1.14$ eV; and the FCI result[18] is $+1.17$ eV. The UCEPA method substantially improves the accuracy of the predicted EA of the oxygen atom, a system in which size extensity is a major concern.

IV. SUMMARY

A potentially size-consistent multiconfiguration-based unitary coupled electron pair approximation (UCEPA) was introduced in this paper, and its usefulness was illustrated by examining benchmark systems for which full CI calculations are available and by realistic studies of two well known species (e.g., CH_2 and O^-). Formally, the present development expresses the wave function in terms of a unitary exponential operator $\exp(G)$ acting on the MCSCF reference function. The state transfer operators which appear in G connect the (MC) reference and an expanded state space $\{|n\rangle\}$. Expansion of the resulting variational energy expectation value through second order in powers of the amplitudes $\{\tau_n\}$ which appear in G results in a set of simultaneous linear equations for the optimal $\{\tau_n\}$ with all requisite matrix elements given in the Gelfand–Tsetlin basis. This second-order expansion also destroys the rigorous variational character of the resulting energy expression. On the other hand, it allows the final energy expression to be size consistent if the MC reference wave function is of the CAS-type. We employ modifications of computational techniques originally developed for large-scale CI, making our newly developed UCEPA method as "open ended" (i.e., only limited by available CPU time) as large-scale CI techniques. Our UCEPA method may be viewed as an MC extension of CEPA(0)[2] compatible with the unitary group approach.

Application of UCEPA to the benchmark systems, H_2O, BH, and CH_2 and to larger basis set studies of CH_2 and O^-, indicates that UCEPA calculations are at least as accurate as CI calculations in the same space, and may be considerably more accurate. We wish to emphasize that a UCEPA calculation and a CI calculation in the same state space have nearly identical computational requirements, although the UCEPA energy can be size consistent if the MC reference function is chosen to be of the CAS-MCSCF variety.

ACKNOWLEDGMENTS

We acknowledge the financial support of the National Science Foundation (CHE-8511307 and CHE-8704779) and the donors of the Petroleum Research Office Grant (PRF 14446AC6) administered by the American Chemical Society. Some calculations associated with this work were done at the NSF supported San Diego Supercomputer Center. We also acknowledge several extremely valuable correspondences with Professor I. Shavitt which caused us, in particular, to look carefully into the size consistency of this method.

APPENDIX: SIZE EXTENSIVITY

The size extensivity of the energy expression given in Eqs. (2.5) or (2.8) can be demonstrated by considering an example consisting of two isolated (noninteracting) molecules which need not be chemically identical or in identical electronic states. The CAS-MCSCF reference wave function

for such a system can be decomposed into a product of the CAS-MCSCF wave functions for the individual molecules

$$|\Phi\rangle = \prod_{i=1}^{2} |\Phi_i\rangle \,. \tag{A1}$$

The orthogonal-complement-space functions $\{|\tilde{n}\rangle\}$ taken together with $|\Phi\rangle$ contain *all* CSFs which can be formed by distributing the $N_1 + N_2 = N$ electrons among all of the active orbitals of the two-molecule system. The "charge transfer" CSFs, which contain more than and less than N_i electrons on the ith molecule do not contribute to the wave function $|\Phi\rangle$ because Hamiltonian matrix elements which couple them with the (dominant) "nonionic" CSFs, vanish for the separated molecules.

As a result of being able to neglect such charge transfer CSFs, we can think of the two-molecule CAS space as consisting of direct products of CAS spaces, one for each molecule. This implies that the orthogonal complement space for the N-electron system $\{|\tilde{n}\rangle\}$ as well as the space of CSFs *external* to the CAS space $\{|n\rangle\}$ (for which we can also ignore charge transfer terms), can be labeled in terms of the orthogonal complement functions $\{|\tilde{n}_j\rangle; j = 1,2\}$ and external CSFs $\{|n_j\rangle; j = 1,2\}$ for each of the molecules. For example, $|\tilde{n}_1,\tilde{n}_2'\rangle$ is a function within the orthogonal complement space (OCS) having molecules 1 and 2 "excited" out of their Φ's into $\tilde{n}_1$ and $\tilde{n}_2'$, respectively; $|n_1,\tilde{n}_2\rangle$ is a member of the *external* CSF space (because it does not contain only Φ's and $\tilde{n}$'s) with molecule 2 excited to OCS function $\tilde{n}_2$ and molecule 1 excited to external function n_1.

Using this molecule-by-molecule labeling of the OCS and external spaces, we can decompose the matrices $B_n \equiv \langle\Phi|H|n\rangle = H_{0n}, H_{0\tilde{n}}, \langle n|H|m\rangle - E_0\delta_{mn} \equiv A_{nm}, A_{\tilde{n}\tilde{m}}$, and $A_{\tilde{n}m}$ and determine which elements are nonvanishing. This information can then be used to analyze the size consistency of the energy expansion of Eq. (2.5c) as well as Eqs. (2.5) which determine the τ_n values entering into Eq. (2.5c).

All of the CAS-space $B_{\tilde{n}}$ values vanish because $|\Phi\rangle$ (and each of the $|\Phi_i\rangle$) has been determined via the MCSCF variational method; the matrix elements $B_{\tilde{n}} = \langle\Phi|H|\tilde{n}\rangle$ are the CSF-space generalized Brillouin theorem elements, which vanish. The B_n values for $|n\rangle$ in the external CSF space do *not* vanish. The linear equations [Eq. (2.5)] which determine the $\tau_{\tilde{n}}$ and τ_n values can be partitioned and solved for the $\tau_{\tilde{n}}$ and τ_n to give

$$\sum_n \Big[(H_{mn} - E_0\delta_{mn}) - \sum_{\tilde{n}\tilde{m}} H_{m\tilde{m}} (H - E_0)^{-1}_{\tilde{m}\tilde{n}} H_{\tilde{n}n} \Big] \tau_n = -B_m \tag{A2a}$$

and

$$\tau_{\tilde{n}} = -\sum_{\tilde{m}n} (H - E_0 1)^{-1}_{\tilde{n}\tilde{m}} H_{\tilde{m}n} \tau_n \,. \tag{A2b}$$

Equation (A2a) determines the τ_n amplitudes for the *external* space, which are the only ones that contribute to the energy of Eq. (2.5c). Let us now focus on the matrices B_m and the partitioned matrix of Eq. (A2a) labeled by indices m and n.

The conventional Slater–Condon rules, combined with the "direct product" nature of $|\Phi\rangle$, $\{|\tilde{n}\rangle\}$, and $\{|n\rangle\}$, the result $\langle\Phi|H|\tilde{n}\rangle = 0$ and the fact that integrals over functions on different molecules vanish, allow us to conclude that the energy expression involves a sum over only *external*-space terms $\Sigma_n \langle\Phi|H|n\rangle \tau_n$. Moreover, only those external-space functions $|n\rangle$ in which a *single* molecule is excited (from its $|\Phi_j\rangle$ to its $|n_j\rangle$) contribute directly. For these functions the matrix elements reduce to $\langle\Phi|H|n\rangle = \langle\Phi_j|H|n_j\rangle$ and involve one-molecule matrix elements[34] which can be expressed in terms of one-molecule integrals.

Given that only the τ_n for $|n\rangle = |\Phi_1 n_2\rangle$ and $|n_1\Phi_2\rangle$ are needed to evaluate the energy, we now examine the equations which determine *these particular* τ_n's. If we can show that these equations yield values for these τ_n's which are independent of the presence of the other molecule (and thus are equal to the τ_{n_j}'s which one obtains in a calculation on a single molecule), we can show that E is size consistent. First, we observe that for $|n\rangle = |\Phi_1 n_2\rangle$, $H_{mn} - E_0\delta_{mn}$ vanishes unless $|m\rangle$ is also of the form $|\Phi_1 m_2\rangle$. Likewise if $|n\rangle = |n_1\Phi_2\rangle$, $|m\rangle$ must be $|m_1\Phi_2\rangle$. These particular elements reduce to $\langle m_j|H|n_j\rangle - \delta_{m_j n_j}E_j$, where $E_j = \langle\Phi_j|H|\Phi_j\rangle$ is the CAS-MCSCF energy of a single molecule. External space functions which contain one molecule (e.g., $j = 2$) excited externally *and* the other molecule excited to an OCS function (e.g. $|m\rangle = |\tilde{n}_1 m_2\rangle$) do not couple with the $|n\rangle = |\Phi_1 n_2\rangle$ even if $m_2 = n_2$ because their coupling matrix element $H_{mn} = \delta_{m_2 n_2}\langle\tilde{n}_1|H|\Phi_1\rangle$ vanishes because of the generalized Brillouin theorem for molecule 1.

Next, we observe that with $|m\rangle = |\Phi_1 m_2\rangle$ or $|m_1\Phi_2\rangle$, the only $H_{\tilde{m}m}$ elements which are nonzero are those in which $|\tilde{m}\rangle = |\Phi_1\tilde{n}_2\rangle$ or $|\tilde{n}_1\Phi_2\rangle$, respectively, and for these cases $H_{m\tilde{m}} = \langle m_j|H|\tilde{m}_j\rangle$. For $|\tilde{m}\rangle$ of this form, the matrix elements $H_{\tilde{m}\tilde{n}} - E_0\delta_{\tilde{m}\tilde{n}}$ vanish unless $|\tilde{n}\rangle$ is of the same form, in which case, $H_{\tilde{m}\tilde{n}} - E_0\delta_{\tilde{m}\tilde{n}} = \langle\tilde{m}_j|H|\tilde{n}_j\rangle = E_j\delta_{\tilde{m}_j\tilde{n}_j}$. The fact that the $H_{\tilde{m}\tilde{n}} - E_0\delta_{\tilde{m}\tilde{n}}$ matrix involves a "block" with $\tilde{m}$ and $\tilde{n}$ of the form described above, and that this block does not couple with other $|\tilde{m}\rangle$'s means that the inverse matrix $[H - E_0 1]^{-1}_{\tilde{m}\tilde{n}}$ is the inverse of this particular block of $(H - E_0 1)_{\tilde{m}\tilde{n}}$ each element of which $(H_{\tilde{m}_j\tilde{n}_j} - E_j\delta_{\tilde{m}_j\tilde{n}_j})$ is a single-molecule matrix element.

Thus far, we have concluded that if $|m\rangle$ is of the form $|\Phi_1 m_2\rangle$ or $|m_1\Phi_2\rangle$ then $\tilde{m}$ and $\tilde{n}$ must also be of this form. A similar analysis of $H_{\tilde{n}n}$ requires $|n\rangle$ to also be of the form $|\Phi_1 n_2\rangle$ or $|n_1\Phi_2\rangle$, in which case $H_{\tilde{n}n} = \langle\tilde{n}_j|H|n_j\rangle$.

Assembling all of these observations into Eq. (A2a) gives

$$\sum_n \Big[\langle m_j|H|n_j\rangle - E_j\delta_{m_j n_j} - \sum_{\tilde{m}_j\tilde{n}_j} \langle m_j|H|\tilde{m}_j\rangle [H - E_j]^{-1}_{\tilde{m}_j\tilde{n}_j} \langle\tilde{n}_j|H|n_j\rangle \Big] \tau_{n_j} = -\langle\Phi_j|H|m_j\rangle \,, \tag{A3}$$

for $j = 1$ and for $j = 2$, *separately*. This equation gives the amplitudes τ_{n_j} for the *two-molecule* external states of the form $|\Phi_1 n_2\rangle$ and $|n_1\Phi_2\rangle$. The matrix multiplying the $\{\tau_{n_j}\}$ vector and the vector $\{-\langle\Phi_j|H|m_j\rangle\}$ on the right-hand side of Eq. (A3) involve only *single-molecule* integrals. They

are exactly the same matrices which would arise if one were studying the single (jth) molecule.

We therefore conclude that for two noninteracting not necessarily identical molecules:

(i) The τ_{n_j}'s which correspond to *external* excitations of a single molecule [the only amplitudes needed to evaluate E via Eq. (2.5c)] are *identical* to those obtained in the absence of the other molecule.

(ii) Since the CAS-MCSCF energy E_0 itself is size extensive, the energy expression of Eq. (2.5c) is also size extensive becaue the sum $\Sigma_n \tau_n \langle \Phi | H | n \rangle$ reduces to $\Sigma_{j=1}^{2} \Sigma_{n_j} \tau_{n_j} \langle \Phi_j | H | n_j \rangle$. Note that although the proof of size consistency has been carried out using the $\{|\Phi\rangle, |\tilde{n}\rangle, |n\rangle\}$ function space, it remains equally valid in the CSF space because these two function classes span exactly the same spaces.

[1]J. Paldus, P. E. S. Wormer, F. Visser, and A. van der Avoid, J. Chem. Phys. **76**, 2458 (1982).

[2]W. Kutzelnigg, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer (Plenum, New York, 1977), pp. 129–188; W. Meyer, Int. J. Quantum Chem. **5**, 341 (1971); J. Chem. Phys. **58**, 1017 (1973).

[3]In the sense in which the linearized coupled cluster method is equivalent to what is termed CEPA-0, the linearized MC coupled cluster developments of the Bartlett group (MR-LCCM) represent earlier treatments of MC-based CEPA. See, for example, W. D. Laidig, P. Saxe, and R. J. Bartlett, J. Chem. Phys. **86**, 887 (1987); W. D. Laidig and R. J. Bartlett, Chem. Phys. Lett. **104**, 424 (1984). Moreover, R. J. Gdanitz and R. Ahlrichs [in Chem. Phys. Lett. **143**, 413 (1988)] recently reported the development and implementation of what they call the averaged coupled-pair functional (ACPF); this is also a MR-based extension of CEPA theory.

[4]J. Paldus, J. Chem. Phys. **61**, 5321 (1974).

[5]J. Paldus, in *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements*, edited by J. Hinze, Lecture Notes in Chemistry (Springer, Berlin, 1981), Vol. 22, pp. 1–50; I. Shavitt, *ibid.* pp. 51–99.

[6]P. W. Payne, Int. J. Quantum Chem. **22**, 1085 (1982).

[7]P. Saxe, H. F. Schaefer, and N. C. Handy, Chem. Phys. Lett. **79**, 202 (1981); R. J. Harrison and N. C. Handy, *ibid.* **95**, 386 (1983).

[8]C. W. Bauschlicher and P. R. Taylor, J. Chem. Phys. **85**, 6510 (1986).

[9]See Chap. 2, Sec. 2.B.1 of P. Jørgensen and J. Simons, *Second Quantization-Based Methods in Quantum Chemistry* (Academic, New York, 1981).

[10]The fact that the energy functional of Eq. (2.4) has been minimized can be seen by examining the A matrix which corresponds to the second derivative of E with respect to the $\{\tau_n\}$. The A matrix is positive definite as long as $\langle n | H | m \rangle$ has eigenvalues which all lie above E_0.

[11]B. H. Lengsfield and B. Liu, J. Chem. Phys. **75**, 478 (1981).

[12]C. A. J. Fletcher, *Computational Galerkin Methods* (Springer, Berlin, 1984).

[13]G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **75**, 1284 (1981).

[14]J. A. Pople, R. Krishnan, H. B. Schlegal, and J. S. Binkely, Int. J. Quantum Chem. Symp. **13**, 225 (1979).

[15]B. Roos, Chem. Phys. Lett. **15**, 153 (1972).

[16]I. Shavitt, Int. J. Quantum Chem. Symp. **12**, 5 (1978).

[17]P. E. M. Siegbahn, J. Chem. Phys. **72**, 1647 (1980).

[18]P. R. Taylor (private communication).

[19]D. Feller, L. E. McMurchie, W. T. Borden, and E.R. Davidson, J. Chem. Phys. **77**, 6134 (1982); D. Feller and E. R. Davidson, *ibid.* **82**, 4135 (1985); R. Krishnan, *ibid.* **82**, 4142 (1985).

[20]J. Simons and K. D. Jordan, Chem. Rev. **87**, 535 (1987).

[21]W. D. Laidig and R. J. Bartlett, Chem. Phys. Lett. **104**, 424 (1984).

[22]R. J. Bartlett, H. Sekino, and G. D. Purvis, Chem. Phys. Lett. **98**, 66 (1983).

[23]R. Ahlrichs, P. Scharf, and C. Ehrhardt, J. Chem. Phys. **82**, 890 (1985).

[24]I. Shavitt, Tetrahedron **41**, 1531 (1985).

[25]H. F. Schaefer, Science **231**, 1100 (1986).

[26]W. A. Goddard, Science **227**, 917 (1985).

[27]H.-J. Werner and E.-A. Reinsch, J. Chem. Phys. **76**, 3144 (1982).

[28]C. Salez and A. Veillard, Theor. Chim. Acta **11**, 441 (1968).

[29]S. Huzinaga, J. Chem. Phys. **42**, 1293 (1965).

[30]P. R. Taylor, J. Chem. Phys. **74**, 1256 (1981).

[31]Truncation of the CSF list to the *exact* first-order space is difficult with the UGA programs available in this laboratory. However, truncation to a space *slightly* larger (and including) the first-order space is straightforward and economical. The so-called first-order calculations in this paper are actually performed in this slightly larger space.

[32]M. R. Hoffmann and J. Simons, J. Chem. Phys. **88**, 993 (1988).

[33]J. Almlöf and P. R. Taylor, J. Chem. Phys. **86**, 4070 (1987).

[34]We use the symbol H to denote both the two-molecule Hamiltonian and the one-molecule Hamiltonian. This notation should not cause confusion because which Hamiltonian is involved will always be clear when one notes whether two-molecule $\{|\Phi\rangle, |n\rangle, |\tilde{n}\rangle\}$ or one-molecule $\{|\Phi_j\rangle, |n_j\rangle, |\tilde{n}_j\rangle\}$ functions are involved in the matrix elements.