5-1-1988

Hylleraas Variational Perturbation Theory: Application to Correlation Problems in Molecular Systems

Robert J. Cave
Harvey Mudd College

Ernest R. Davidson
Indiana University - Bloomington

Hylleraas variational perturbation theory: Application to correlation problems in molecular systems

Robert J. Cave and Ernest R. Davidson[a]

Department of Chemistry, Indiana University, Bloomington, Indiana 47405

Hylleraas variational perturbation theory is applied through second order in energy to estimate the correlation energy in several molecular systems. The specific choices for H_0 and V which are made lead to equations nearly identical to the multireference linearized coupled-cluster method of Laidig and Bartlett. The results obtained are in virtually exact agreement where comparisons have been made. Results from test calculations are presented for BeH_2, CH_2, and C_2H_4. In addition, the utility of perturbation theory for selecting correlating configurations is examined. This procedure is found to be quite accurate while significantly reducing the size of the system of linear equations to be solved.

I. INTRODUCTION

As the scope of quantum chemical applications increases, the problem of producing "size-consistent" results grows in importance. A size-consistent[1] method is one in which the calculated energy scales linearly with the number of particles. For example, a size-consistent method when applied to a system of n well-separated H_2 molecules obtains an energy equal to n times the result of a calculation on a single H_2. Restricted and unrestricted Hartree–Fock, some forms of complete active space SCF, and full CI are size-consistent methods. However, singles and doubles configuration interaction (SDCI) results are not. This is not merely a problem concerned with examining separated systems, but also arises for the estimation of correlation for individual many-electron systems. That is, the fraction of the correlation energy obtained in any CI treatment using a fixed level of excitation from a set of reference configurations decreases as the number of electrons increases.

Some perturbation theory approaches have the advantage that they can be size consistent.[2] Another method which can be size consistent is the coupled-cluster approach of Coester and Kummell,[3] first applied to chemical problems by Cizek[4] and Paldus.[4(b)] In the coupled-cluster method, the effects of higher excitations are approximated as products of lower excitations (rather than completely neglected as in truncated CIs). The simplest choice of reference space is the single SCF configuration. Pople *et al.*[5] and Bartlett and co-workers[6–8] have applied the method to systems with large basis sets and have produced results within the coupled-cluster doubles[5,6] and coupled-cluster singles and doubles approximations,[7] and even several variants of the inclusion of connected triples (i.e., triple excitations which cannot be written as products of lower excitations).[8] Related methods are the symmetry-adapted cluster and symmetry-adapted cluster CI approaches of Nakatsuji and co-workers.[9] The work of Bartlett and co-workers[8,10] has shown that the coupled-cluster method is able to compensate for relatively poor reference functions via the approximate inclusion of higher

excitations. For chemical accuracy, however, triple excitations are reasonably important when a poor reference function is being used.[8,10] Since the inclusion of triple excitations can be computationally intensive, multireference coupled-cluster methods have also been explored.

Jeziorski and Monkhorst,[11] Lindgren,[12] and Muhkerjee and co-workers[13] have explored the requirements for a multireference coupled-cluster approach. Applications of their results have been limited due to the computational complexity or to the presence of intruder states when a complete reference space is treated as a quasidegenerate zeroth-order space. However, a number of groups have examined simplified variants of these approaches.[10,14–16] In each case the procedure was implemented based on a complete active-space SCF zeroth-order wave function, and all single and double excitations relative to all configurations in the zeroth-order space were considered. In addition, the equations to generate the cluster amplitudes were linearized (except in Ref. 14, where the coupled-cluster equations were truncated at second-order commutators) to simplify the calculations. In the one-reference case it is known that the linearized coupled-cluster approach yields reasonable energies when no near degeneracies arise. In the multireference approach it is believed that the zeroth-order space will accurately treat interactions of near-degenerate low-lying states, thus linearization should not lead to serious errors.[10]

The method presented here is an application of Hylleraas variational perturbation theory to the electron correlation problem. It will be shown to be closely related to the multireference linearized coupled-cluster approaches mentioned above. The choice made for the partitioning of H into an H_0 and perturbation V yields equations which are essentially identical to those of Laidig and Bartlett's multireference linearized coupled-cluster approach.[10] The coefficients obtained are those which extremize the Hylleraas second-order energy expression. This expression can, in turn, be written in a form quite similar to the expression used to develop a perturbative correction to the SDCI energy expression. Thus, this expression manifests the similarity between the various approaches to the size-consistency problem.

In addition, the present method has been implemented

---

[a] Author to whom correspondence should be addressed.

with the option of using second-order Rayleigh–Schrö-
dinger perturbation theory to select correlating configura-
tions. It is clear that for size-consistent methods to be useful
they must be applied to problems where SDCI is expected to
give quantitatively wrong results, i.e., large systems. How-
ever, the number of double excitations grows so rapidly that
large systems will be untreatable if all double excitations
must be considered. A number of groups have used perturba-
tion-theory selection criteria in CI approaches with suc-
cess.[17] Nakatsuji et al.[9] have also used perturbation-theory
selection in the implementation of the symmetry-adapted
cluster and symmetry-adapted cluster CI methods. Below
we demonstrate that perturbation-theory selection also
yields reasonable accuracy within the present framework.

The systems examined below are: (1) the lowest elec-
tronic surface of BeH_2 as the two hydrogens are removed
symmetrically to form $Be + H_2$; (2) the lowest singlet and
triplet states of methylene; and (3) a number of low-lying
excited states of ethylene. The aim in the first is to compare
the present method with previous results of Laidig and Bart-
lett.[10(a)] In the second case we compare with full CI results of
Bauschlicher and Taylor.[18] In the third case we examine the
performance of the method for a well-known difficult exam-
ple of valence-Rydberg mixing.

II. VARIATIONAL PERTURBATION THEORY

The method is based[19,20] on the Rayleigh–Schrödinger
perturbation theory equations[21] for the correlation energy
and coefficients of the wave function. That is, the wave func-
tion, energy, and Hamiltonian are all considered functions of
the strength of the perturbation, which itself is parametrized
using the quantity λ as

$$H = H_0 + \lambda V,$$
$$\Psi = \Psi_0 + \lambda \Psi_1 + \lambda^2 \Psi_2 + \cdots,$$
$$E = E_0 + \lambda E_1 + \lambda^2 E_2 + \cdots. \tag{1}$$

Intermediate normalization is assumed, thus

$$\langle \Psi_0 | \Psi_0 \rangle = 1,$$
$$\langle \Psi | \Psi_0 \rangle = 1, \tag{2}$$
$$\langle \Psi_i | \Psi_0 \rangle = 0, \quad i \neq 0.$$

Inserting Eqs. (1) into the Schrödinger equation and col-
lecting all terms with the same λ dependence leads to the
familiar results of Eqs. (3):

$$(H_0 - E_0)\Psi_0 = 0, \tag{3a}$$
$$(H_0 - E_0)\Psi_1 = (E_1 - V)\Psi_0, \tag{3b}$$
$$(H_0 - E_0)\Psi_2 = (E_1 - V)\Psi_1 + E_2\Psi_0. \tag{3c}$$

The scalar product of Eqs. (3) with Ψ_0 yields

$$\langle \Psi_0 | V | \Psi_i \rangle = E_{i+1}. \tag{4}$$

The scalar product of Eq. (3b) with Ψ_1 yields

$$\langle \Psi_1 | H_0 - E_0 | \Psi_1 \rangle = - \langle \Psi_1 | V | \Psi_0 \rangle. \tag{5}$$

Adding Eq. (4) (with $i = 1$) and Eq. (5) one obtains

$$E_2 = 2\langle \Psi_0 | V | \Psi_1 \rangle + \langle \Psi_1 | H_0 - E_0 | \Psi_1 \rangle. \tag{6}$$

Equation (6) is identical to the result obtained for the coeffi-
cient of λ^2 in an expansion of $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$.

Equation (6) is the energy expression (to second order
in energy) for Hylleraas variational perturbation theory
(VPT).[19,20] Variation of E_2 with respect to changes in Ψ_1
can be shown to yield Eq. (3b); i.e., the solution of the Ray-
leigh–Schrödinger perturbation-theory equations for Ψ_1
also yields a stationary energy for the VPT second-order
energy expression. In addition, when Ψ_1 is the function
which yields a stationary value for E_2, Eq. (5) is true and Eq.
(6) reduces to Eq. (4). It can also be shown that any ap-
proximate Ψ_1 yields an upper bound for E_2 of Eq. (6) when-
ever E_0 is the lowest eigenvalue of H_0.[20] That is, whenever E_0
is the lowest eigenvalue of H_0, the exact Ψ_1 yields a minimum
for E_2 evaluated using Eq. (6). We now particularize the
choices of Ψ_0, H_0, and V to the case at hand.

We assume that Ψ_0 is some suitable n-configurational
zeroth-order guess, obtained either from a multiconfigura-
tion SCF calculation, an iterative natural orbital calculation,
or the dominant configurations obtained in a preliminary CI
or previous variational perturbation theory calculation on
the state of interest. We write

$$\Psi_0 = \sum_{i=1}^{n} d_0^i \phi_i. \tag{7}$$

d_0^i may be obtained from diagonalizing H over the zeroth-
order space, or chosen on the basis of a prior calculation over
the entire configuration space to obtain a specific zeroth-
order state. In either case we choose $E_0 = \langle \Psi_0 | H_0 | \Psi_0 \rangle$. We
define projection operators $\mathbf{P} = |\Psi_0\rangle\langle\Psi_0|$ and $\mathbf{Q}$, with $\mathbf{Q}$
given by

$$\mathbf{Q} = \mathbf{Q}_0 + \mathbf{Q}_1 = \sum_{k=1}^{n-1} |\Phi_k\rangle\langle\Phi_k| + \sum_{j=n+1}^{N} |\phi_j\rangle\langle\phi_j| \tag{8}$$

and

$$\Phi_k = \sum_{i=1}^{n} d_k^i \phi_i, \quad k = 1,...,n-1. \tag{9}$$

That is, the Φ_k are the orthogonal complement of Ψ_0 in the
space of the n reference configurations. The individual ϕ_i are
taken as orthonormal, as are Ψ_0 and $n-1$ Φ_k, and we treat
d_k^i and d_0^i as fixed throughout the calculation. The remain-
ing $N-n$ configurations (which comprise $\mathbf{Q}_1$) are selected
single and double excitations relative to the reference space.
H_0 and V are then defined as

$$H_0 = \mathbf{PHP} + \mathbf{QHQ}, \tag{10a}$$
$$V = \mathbf{PHQ} + \mathbf{QHP}. \tag{10b}$$

Note that H_0 contains the full Hamiltonian matrix coupling
all configurations other than Ψ_0. With these definitions we
write Ψ_1 as

$$\Psi_1 = \sum_{m=1}^{n-1} C_m \Phi_m + \sum_{m=n+1}^{N} C_m \Phi_m. \tag{11}$$

If Ψ_1 is substituted into Eq. (6) and the variation of E_2 with
respect to C_m is set equal to zero, one obtains

$$\mathbf{Q}(H - E_0)\mathbf{QC} + \mathbf{QHP} = 0, \tag{12}$$

where $\mathbf{C}$ is the column vector of coefficients defining Ψ_1.
Once Ψ_1 is obtained one has E_2 via Eq. (4). Below we desig-

nate $E_0 + E_2$ as E_{VPT}. The numerical method for solving Eq. (12) is a simple modification for the inhomogeneous problem of a standard procedure for the solution of sparse-matrix homogeneous linear equations.[22] It is also quite similar to the method developed by Purvis and Bartlett.[23] The procedure is briefly summarized in the Appendix.

Equation (12) is equivalent to the final equation obtained by Laidig and Bartlett[10] in their multireference linearized coupled-cluster (MRLCC) approach when their definition of $\mathbf{Q}$ contains the orthogonal complement functions within the reference space Φ_k. Alternatively, defining our $\mathbf{Q} = \mathbf{Q}_1$ (with Ψ_0 an eigenfunction of H within the reference space) causes Eq. (12) to be exactly that solved in Ref. 10 for the MRLCC method. (Where total energies from MRLCC calculations are presented below we denote them as E_{MRLCC}.) Kutzelnigg[24] has obtained entirely equivalent formal results using a more general partitioning of H than that used above. He also obtained expressions for higher order corrections to the energy and Ψ.

We noted above that the expression for E_2 can be rewritten in a more suggestive form. Using the fact that E_0 is chosen to be $\langle \Psi_0 | H | \Psi_0 \rangle$, one can rearrange the VPT energy expression to yield

$$E_2 = E_{TOT} - E_0 = \langle \Psi_0 + \Psi_1 | H - E_0 | \Psi_0 + \Psi_1 \rangle. \quad (13)$$

The SDCI expression for the correlation energy (E_{SD}) can similarly be written as

$$E_{SD} = \langle \Psi_0 + \Psi_{SD} | H - E_0 | \Psi_0 + \Psi_{SD} \rangle / (1 + |\Psi_{SD}|^2). \quad (14)$$

Expanding the denominator in Eq. (14) one obtains

$$E_{SD} = (1 - |\Psi_{SD}|^2 + \cdots)$$

$$\times \langle \Psi_0 + \Psi_{SD} | H - E_0 | \Psi_0 + \Psi_{SD} \rangle. \quad (15)$$

For small systems an approximate size-consistency correction to E_{SD} can be made by discarding all but the leading term in Eq. (15), under the assumption that the main contribution of higher excitations would be to cancel the terms which follow. Formally Eq. (13) is identical to the leading term of Eq. (15). For large systems the post-CI size-consistency correction (i.e., adding $|\Psi_{SD}|^2$ $\langle \Psi_0 + \Psi_{SD} | H - E_0 | \Psi_0 + \Psi_{SD} \rangle$ to E_{SD}) deteriorates because $|\Psi_{SD}|^2$ and $\langle \Psi_0 + \Psi_{SD} | H - E_0 | \Psi_0 + \Psi_{SD} \rangle$ scale improperly with the number of electrons. However, in VPT one assumes that Eq. (13) is the correct expression for the energy from the outset and extremizes E_2. Thus, another way to view the present approximation is that one is optimizing the coefficients of a SDCI under the constraints of a simple size-consistency correction.[25,26] Ahlrichs[27] and Paldus et al.[28] have shown that when Ψ_0 is composed of a single function, the equations defining the extremum of E_2 in Eq. (12) reduce to the linearized coupled-cluster equations. The latter authors[28] also related the linearized coupled-cluster equations to VPT in the one-reference case.

As mentioned above, one of the aims in the development of a size-consistent method is to examine the effects of correlation in larger systems. However, as the system size increases one quickly reaches the point where treatment of all single and double excitations within a good quality basis set

becomes intractable. Thus, even though a method may be size consistent, large systems can remain out of reach. One could attempt to improve and then truncate the virtual space, performing the final calculation in a smaller MO basis. A second option, one which we have incorporated into our implementation of VPT, is to use second-order Rayleigh–Schrödinger perturbation theory to select a given subset of configurations which yield the largest energy lowering at second order. That is, for each configuration outside the reference space we calculate

$$\epsilon_2^i = |\langle \Psi_0 | H | \phi_i \rangle|^2 / (E_0 - E_i), \quad (16)$$

where E_i is the average expectation value of the various spin-adapted configurations corresponding to the particular orbital occupation in ϕ_i. Selecting the subset of configurations with the largest energy lowerings, the VPT equations are then solved treating only this set of configurations. The contribution of the remaining configurations is then estimated via a simple extrapolation formula:

$$(E_2)_{EXTRAP} = E_0 + (E_{VPT} - E_0)(1 + E_{PTD}/E_{PTK}), \quad (17)$$

where E_{VPT} is the VPT energy corresponding to the subset explicitly treated and E_{PTD} and E_{PTK} are the second-order Rayleigh–Schrödinger perturbation theory energy contributions obtained via Eq. (16) for those configurations discarded and kept in the VPT calculation, respectively. A similar procedure has been used by Nakatsuji and co-workers in their implementation of the symmetry-adapted cluster and symmetry-adapted cluster CI theories,[9] as well as by several groups in CI approaches.[17,29] It will be shown below to be a useful approximation for VPT and MRLCC calculations as well.

Finally, we comment on the conditions under which one can expect E_2 obtained via the VPT equations to be a minimum with respect to arbitrary variations in the C_i. It was noted above that E_2 [evaluated using Eq. (6)] will be an upper bound to the true E_2 when E_0 is the lowest eigenvalue of H_0. For applications to excited states within a given symmetry, where E_0 is clearly not the lowest eigenvalue of H_0, the VPT E_2 will not be a minimum. However, due to the choice of H_0, the VPT E_2 will most often not be a minimum even for the lowest state of a symmetry. This occurs because H_0 allows interactions between the excited configurations, and thus the lowest eigenvalue of $\mathbf{QHQ}$ can be below E_0, even though no diagonal elements of H_0 are lower than E_0. This will be true, in particular, for large systems since the correlation energy will become much larger than the separation between zeroth-order states. The fact that the VPT E_2 is not, in general, a minimum is not a drawback to the method, it is only noted here for completeness. However, were the present method applied with the usual choice for H_0 in Rayleigh–Schrödinger perturbation theory (i.e., H_0 diagonal), E_2 would be a minimum for the lowest state in each symmetry independent of the size of the total correlation energy.

III. RESULTS

All calculations presented below were performed using the MELD suite of electronic structure codes from this labo-

ratory.[29] The SCF calculations were from restricted Hartree–Fock, open-shell restricted Hartree–Fock, or two-configuration SCF (= TCSCF) calculations. In all cases except that of the $^1B_{1u}$ calculations on ethylene the virtual orbitals used in the correlation treatments were the canonical Hartree–Fock virtual orbitals. In the case of the $^1B_{1u}$ states of ethylene the MOs used were the average natural orbitals[30,31] of the two lowest states obtained in a "CI1" calculation.[32] Briefly, in a CI1 calculation all single excitations and all double excitations involving the π^* electron are included in the CI and the two lowest roots sought. The average natural orbitals so obtained yield a compact MO description of the *two* lowest states of $^1B_{1u}$ symmetry. We are in no way forced to utilize this procedure by the use of VPT. Rather, past experience[31] has shown that this method results in a more balanced description of the two lowest states in either a CI or perturbation-theory based approach.

The basis sets used are as follows. For the calculations on BeH$_2$ the basis set of Purvis and Bartlett[7] was used. It consists of a $(10s,3p)$ set on Be contracted to $(3s,1p)$ and a $(4s)$ set on H contracted to $(2s)$. This allows for simple comparison with the MRLCC results of Laidig and Bartlett[10] and the full CI results of Purvis *et al.*[33] For CH$_2$ and for ethylene the Dunning/Huzinaga[34] DZ basis sets for C and H were used. For CH$_2$ the polarization functions used were those used in the Bauschlicher and Taylor[18] full CI studies (H: $2p = 1.0$ for both states; C: $1\,^1A_1$, $3d = 0.51$; $1\,^3B_1$, $3d = 0.74$). For ethylene the polarization and Rydberg functions used were those of Brooks and Schaefer's study[35] (polarization: C $3d = 0.75$, H $2p = 1.0$; Rydberg: two $p\pi$ Rydberg functions with exponents 0.034 and 0.012 on each C).

In the calculations on the C_{2V} insertion pathway of Be into H$_2$ to yield BeH$_2$, the geometries were taken from the study of Purvis *et al.*[7] The geometries used for the two states of CH$_2$ were those of Ref. 18 and the geometry of ethylene[36] was taken from Ref. 35.

In the correlation treatments below, all single and double excitations were allowed from all reference functions, with the exceptions that no excitations were allowed from the C $1s$ orbitals of CH$_2$ and C$_2$H$_4$. Where a size-consistency

correction is applied to SDCI results below it is based on the expression of Refs. 25 and 26 and is of the form $(1 - c_0^2)\Delta E_{SD}$. Here c_0^2 is the sum of the squares of the coefficients of the reference space in the final SDCI wave function, and ΔE_{SD} is the energy lowering in the final SDCI relative to the reference wave function. We designate these corrected results as QCSDCI (quadruples corrected singles and doubles CI). When perturbation-theory selection is used to truncate the CI, ΔE_{SD} is based on the extrapolated SDCI energy, using an expression similar to Eq. (17) above.

A. BeH$_2$

In Table I results are shown from a variety of single-reference calculations along the path chosen by Purvis *et al.*[7] The results designated a_1^2 are based on the SCF performed having the orbital occupation $1a_1^2 2a_1^2 3a_1^2$, and those labeled b_2^2 arise from the SCF having the orbital occupation $1a_1^2 2a_1^2 1b_2^2$. Those results denoted LCC are from single-reference linearized coupled-cluster calculations (i.e., Ψ_0 is composed of a single configuration). It is seen from the SCF results that the nature of the lowest $^1A_1$ state changes along the reaction path, as would be expected from simple MO arguments. Thus, in comparison with the full CI results of Purvis *et al.*,[33] the b_2^2 results tend to be better at short $r(\text{BeH}_2)$ and the a_1^2 results tend to be more accurate at long $r(\text{BeH}_2)$. However, neither could be considered even a semi-quantitative representation of the true surface over the whole path.

A two-configurational description yields a better description of the insertion reaction along the whole path. Using this two-configuration SCF as the reference space we performed multireference singles and doubles CIs (MRSDCI), MRLCC, and VPT calculations; the results are presented in Table II. The SDCI results are presented with (QCSDCI) and without the quadruples correction. In general, all methods are reasonably close to the full CI results. The QCSDCI, MRLCC, and VPT results are quite close to, but generally overshoot the full CI values. Interestingly, the smallest variation in error relative to the full CI

TABLE I. BeH$_2$ single reference results,[a] C_{2v} geometries.

$r(\text{BeH}_2)$[b]	$r(\text{H}_2)$[b]	SCF(a_1^2)	ΔSDCI(a_1^2)[c]	ΔLCC(a_1^2)[c]	SCF(b_2^2)	ΔSDCI(b_2^2)[c]	ΔLCC(b_2^2)[c]	Full[d]
1.00	4.16	−15.0903	22.0	440.2	−15.6996	0.7	−0.3	−15.7372
2.00	3.24	−15.3190	20.9	240.5	−15.6284	1.3	−1.1	−15.6748
2.50	2.78	−15.4361	20.6	85.7	−15.5627	3.2	−5.2	−15.6229
2.75	2.55	−15.4885	14.1	28.4	−15.5212	11.4	52.1	−15.6029
3.00	2.32	−15.5365	6.5	−44.9	−15.4773	30.1	84.1	−15.6250
3.50	1.86	−15.6187	3.7	−5.8	−15.4020	32.2	220.0	−15.6932
4.0	1.4	−15.6699	3.5	−3.6	−15.3606	26.7	329.8	−15.7376
6.0	1.4	−15.6954	2.4	−4.2	−15.3254	24.7	425.5	−15.7609

[a] SCF and full CI energies in hartrees. The reference configuration is shown in parentheses.
[b] Distances in bohr. $r(\text{BeH}_2)$ is the distance from Be to the H$_2$ midpoint. $r(\text{H}_2)$ is the H$_2$ bond length.
[c] Energies relative to the full CI result, in millihartrees. LCC denotes a single-reference linearized coupled-cluster calculation, SDCI denotes a single-reference singles and doubles CI.
[d] Results taken from Ref. 33.

TABLE II. BeH$_2$ two-reference results, C_{2v} geometries.

$r(BeH_2)$[a]	$r(H_2)$[a]	TCSCF[b]	Corr. E[c]	ΔSDCI[d]	ΔQCSDCI[d]	ΔMRLCC[d]	ΔVPT[d]
1.00	4.16	− 15.7054	31.8	0.3	− 0.4	− 0.4	− 0.5
2.00	3.24	− 15.6330	41.8	0.4	− 1.4	− 1.2	− 1.4
2.50	2.78	− 15.5696	53.3	0.9	− 2.8	− 2.6	− 3.2
2.75	2.55	− 15.5386	64.3	2.0	− 4.3	− 2.4	4.2
3.00	2.32	− 15.5583	66.7	3.1	− 4.8	− 5.5	− 5.5
3.50	1.86	− 15.6372	56.0	2.1	− 2.2	− 1.8	− 2.4
4.0	1.4	− 15.6872	50.4	2.5	− 0.7	− 0.3	− 0.9
6.0	1.4	− 15.7107	50.2	1.6	− 1.7	− 1.4	− 1.9

[a] Distances in bohr. The distances are defined as in Table I.
[b] Energy of the two-configuration SCF wave function, in hartrees.
[c] Correlation energy, in millihartrees, relative to the two-configuration SCF energy, using the full CI energies of Ref. 33.
[d] Error relative to the full CI energies of Ref. 33, in millihartrees. SDCI designates a singles and doubles CI (in this case a two-reference SDCI), QCSDCI indicates the size-consistency corrected SDCI result, MRLCC denotes multireference linearized coupled-cluster results, and VPT denotes variational perturbation theory results.

occurs for the uncorrected MRSDCI values (minimum error = 0.0003 hartree, maximum error = 0.0032 hartree). However, the VPT, MRLCC, and QCSDCI results give error variations of similar size.

One surprising result is the change in sign of the VPT error at $r(BeH_2) = 2.75b$, $r(H_2) = 2.55b$. In this case, the coefficient of the orthogonal complement of Ψ_0 in the final VPT wave function is quite large ($c^2 = 0.39$), indicating that the zeroth-order guess used is a poor one, even though the two configuration SCF wave function was used for Ψ_0. (Since the MRLCC result excludes the orthogonal complement "configuration" this strong mixing does not occur there.) Indeed, the ratios of the two configurations which make up Ψ_0 are quite different between the final MRSDCI, MRLCC, and VPT results. In addition, the VPT energy is above the MRLCC energy. This occurs due to the near degeneracy of Ψ_0 and its orthogonal complement in the reference space, so that E_0 is no longer the lowest eigenvalue of H_0. Thus, the VPT energy need no longer be monotonically decreasing as the expansion set is augmented.

A series of larger calculations were performed on BeH$_2$ at $r(BeH_2) = 2.75b$, $r(H_2) = 2.55b$ to examine the sensitivity of the results to increases in the reference space, the results of which are given in Table III. In the 15-reference function calculation of Table III, the reference space was composed of the dominant spin-adapted configurations[37] in the two-reference VPT result of Table II, the coefficients were obtained by diagonalizing H over these 15 functions. We also include QCSDCI results for these calculations. It is seen that the absolute errors in the VPT and the MRLCC results are similar to the two-reference function case. In the 12-reference case the reference space was composed of the 12 dominant spin-adapted configurations from the SDCI result[37] of Table II, the coefficients defining Ψ_0 were taken from the results of the two-reference CI calculation. The VPT, MRLCC, and QCSDCI results are quite close to the full CI result. Finally, the dominant 17 spin-adapted configurations[37] from the preceding SDCI were used as a reference space; the coefficients defining Ψ_0 were taken from the results of the preceding large SDCI. It is seen that the VPT and MRLCC results are again quite close to the full CI energy. In parallel, one sees that the ratio of the coefficients of the $1a_1^2 2a_1^2 3a_1^2$ and the $1a_1^2 2a_1^2 1b_2^2$ configurations (= $\times$) changes along with the form of Ψ_0. This ratio is − 0.849 in a

TABLE III. BeH$_2$ expanded reference space calculations.[a]

Reference[b]	Conf.[c]	ΔSDCI	$\times$[d]	ΔQCSDCI	ΔMRLCC	$\times$[d]	ΔVPT	$\times$[d]
2[e]	182	2.0	− 0.82	− 4.3	− 2.5	− 0.57	4.2	0.04
15[f]	741	1.0	− 0.84	− 0.7	1.9	− 0.54	− 3.4	− 0.75
12[g]	698	0.0	− 0.85	− 0.4	− 0.3	− 0.82	− 0.5	− 0.89
17[h]	821	0.0	− 0.85	− 0.2	− 0.3	− 0.85	− 0.3	− 0.85

[a] The results are for the point $r(BeH_2) = 2.75b$, $r(H_2) = 2.55b$. Energy differences (in millihartrees) relative to the full CI values are reported. The labels are defined in Table II.
[b] Number of spin-adapted configurations in the reference space.
[c] Number of spin-adapted configurations in the final wave function.
[d] Ratio of the coefficient of the $1a_1^2 2a_1^2 3a_1^2$ configuration to that of the $1a_1^2 2a_1^2 1b_2^2$ configuration in the final wave function.
[e] Results for the present geometry from Table II. The two-reference configurations are from the two-configuration SCF.
[f] The 15 spin-adapted reference configurations are taken from the dominant configurations (Ref. 37) in the VPT two-reference result.
[g] The 12 spin-adapted reference configurations are taken from the dominant configurations (Ref. 37) in the two-reference SDCI result. The zeroth-order wave function was defined by the coefficients of the two-reference SDCI.
[h] The 17 spin-adapted reference configurations are taken from the dominant configurations (Ref. 37) in the 12-reference SDCI results. The zeroth-order wave function was defined by the coefficients of the 12-reference SDCI.

TABLE IV. BeH$_2$ PT selected results,[a] C_{2V} geometries.

$r(BeH_2)$[b]	$r(H_2)$[b]	Conf.[c]	PTK[d]	δQCSDCI	δMRLCC	δVPT
1.00	4.16	113	>99.9	0.3	0.3	0.3
2.75	2.55	99	99.8	0.1	0.0	0.0
4.0	1.4	101	>99.9	0.0	0.0	0.0

[a] The entries marked δQCSDCI, δMRLCC, and δVPT are the differences (in millihartrees) between the perturbation-theory selected calculation and the corresponding unselected result of Table II.
[b] Distances in bohr. The quantities are as defined in Table II.
[c] Number of spin-adapted configurations treated explicitly.
[d] Percentage of the second-order Rayleigh–Schrödinger perturbation-theory energy accounted for by the configurations retained.

calculation allowing up to quadruple excitations from both reference functions from the two-configuration SCF, the energy of which agrees with the full CI energy to five decimal places.

The calculations listed in Table II are quite small, the two-reference calculations consider only 182 spin-adapted configurations. In this case perturbation-theory selection is clearly not needed, but we present results from perturbation-theory selected calculations in Table IV for comparison at several points. The scheme described in Sec. II was used, and all single excitations were included. The number of configurations, the percentage of the perturbation-theory energy kept, and the final energy relative to the full CI are included in Table IV. It is seen that use of perturbation-theory selection yields quite good agreement with the unselected results.

B. CH$_2$

Results are presented in Table V from SDCI, QCSDCI, MRLCC, and VPT calculations along with the full CI results of Ref. 18 for the $^3B_1$ and $^1A_1$ states of methylene. The $^3B_1$ state is well described in zeroth order by the HF configuration, thus all calculations concerning it are based on this single-reference function. Results from both one- and two-reference calculations are presented for the $^1A_1$ state, since the zeroth-order description is known to be markedly improved by including correlation in the C lone pair. It is seen that the QCSDCI, MRLCC, and VPT results all overshoot the full CI energies. The variations in the MRLCC and VPT result in proceeding from the single-reference to the two-

reference calculations on the $^1A_1$ state are somewhat less than the uncorrected SDCI values, but the QCSDCI variation is smaller still. The excitation energies obtained are in quite good agreement with the full CI values, but the variation among the four techniques employed here is admittedly rather small.

The sensitivity of the $^1A_1$ energies to the size of the reference space was also examined. In Table V results are shown from a calculation where the reference space was expanded to include 11 spin-adapted configurations, chosen as those having the largest coefficients in the two-reference VPT calculation.[37] The coefficients defining Ψ_0 were obtained from diagonalizing H over these 11 configurations. It is seen that the change in energy is largest for the SDCI results, the change in the QCSDCI, MRLCC, and VPT energies being less than 1 mhartree, even though E_0 changed by over 18 mhartree.

C. C$_2$H$_4$

Results from second-order Rayleigh–Schrödinger perturbation theory, SDCI, QCSDCI, MRLCC, and VPT calculations on ethylene are shown in Table VI. The excitation energies based on these calculations are given in Table VII. Several calculations are presented for the 1 ^1A_g state for the purposes of examining the sensitivity of the final energy to the size of the reference space and the use of perturbation-theory selection. In the four-reference calculations the four configurations having the largest coefficients in the single-reference VPT calculation[37] were used as the reference space, the coefficients of the reference space being those obtained from diagonalizing H over the reference space. In the two-configuration calculations the MOs were obtained from a two-configuration SCF calculation on the π-electron pair.

In the results for the 1 ^1A_g state the VPT and MRLCC methods are least sensitive to variations in the reference space size, or to the use of perturbation-theory selection. Other than the second-order Rayleigh–Schrödinger perturbation-theory results, the SDCI results are most sensitive to variations in the choice of reference space.

For the $^1B_{1u}$ states, a two-reference configuration Ψ_0 was used in each case, the two configurations being those that principally describe the two lowest $^1B_{1u}$ states in zeroth order in the average natural orbital basis. The coefficients for Ψ_0 in the MRLCC and VPT calculations were obtained via

TABLE V. Results for CH$_2$, C_{2V} symmetry.

State	Reference[a]	ΔSDCI[b]	ΔQCSDCI[b]	Δ(MR)LCC[b]	ΔVPT[b]	Full CI[c]
1 $^3B_1$	1	4.7	− 0.7[d]	− 1.3	− 1.3[e]	− 39.0463
1 $^1A_1$	1	8.9	0.0[d]	− 2.9	− 2.9[e]	− 39.0272
1 $^1A_1$	2	5.0	− 0.8[d]	− 1.3	− 1.3	$\cdots$
1 $^1A_1$	11	2.6	− 0.8[d]	− 1.9	− 2.1	

[a] The number of spin-adapted configurations in the reference space.
[b] Errors relative to the full CI, in millihartrees. The labels are defined in Table II.
[c] Results from Ref. 17, in hartrees.
[d] The QCSDCI results were calculated to three decimal places.
[e] In the one-reference case, the LCC and VPT methods are equivalent.

TABLE VI. Results for C_2H_4,[a] D_{2h} symmetry.

State	Reference[b]	PTK[c]	Conf.[d]	RSPT	SDCI	QCSDCI	(MR)LCC	VPT
$1\,^1A_g$	1	100	5 252	-78.3863	-78.3272	-78.353	-78.3615	-78.3615
$1\,^1A_g$	1	98.8	3 243	-78.3863	-78.3270	-78.353	-78.3610	-78.3610
$1\,^1A_g$	4	100	23 762	-78.3784	-78.3330	-78.356	-78.3616	-78.3618
$1\,^1A_g$	4	96.5	3 242	-78.3784	-78.3319	-78.354	-78.3563	-78.3610
$1\,^1A_g$	2	100	10 226	-78.3778	-78.3356	-78.357	-78.3610	-78.3622
$1\,^1B_{1u}$	2	100	21 062	-78.1398	-78.0287	-78.053	-78.0603	-78.0603
$1\,^1B_{1u}$	2	91.8	5 814	-78.1398	-78.0283	-78.052	-78.0582	-78.0582
$2\,^1B_{1u}$	2	100	21 062	-78.0994	-77.9911	-78.015	-78.0217	-78.0216
$2\,^1B_{1u}$	2	91.7	5 814	-78.0994	-77.9907	-78.014	-78.0195	-78.0195
$1\,^3B_{1u}$	1	100	18 904	-78.2045	-78.1689	-78.190	-78.1942	-78.1942
$1\,^3B_{1u}$	1	86.2	5 287	-78.2045	-78.1695	-78.188	-78.1922	-78.1922
$1\,^3B_{1u}$	28	99.6	20 010	-78.2016	-78.1703	-78.189	-78.1925	-78.1929

[a] All energies in hartrees.

[b] Number of spin-adapted configurations in the reference space.

[c] Percentage of the second-order Rayleigh–Schrödinger perturbation theory energy accounted for by the configurations retained. All single excitations were kept.

[d] Number of spin-adapted configurations treated beyond second-order Rayleigh–Schrödinger perturbation theory.

an iterative procedure, whereby the initial guess at Ψ_0 in either case was obtained from the output of the two-reference SDCI result. In the next iteration the final coefficients from the preceding VPT calculation were used to define Ψ_0. In general, we iterate this procedure until convergence, in this case one iteration, was sufficient. The results for the excitation energies of the QCSDCI, MRLCC, and VPT calculations are in excellent agreement, and are essentially those of Ref. 35. The SDCI excitation energies are somewhat higher, and the Rayleigh–Schrödinger perturbation-theory results are clearly in error. Use of perturbation-theory selection yields quite good agreement with the complete calculations, with changes in total energies of either $^1B_{1u}$ state of less than 2 mhartree. The number of spin-adapted configurations explicitly considered has decreased from 21 062 in the complete calculation to 5814 in the perturbation-theory selected case.

Calculations are also presented for the $^3B_{1u}$ state of ethylene. Comparison is made in Table VI of results from a single-reference calculation (complete and perturbation-theory selected) and an eight-reference perturbation-theory selected calculation. In all cases the results are similar for all the calculational methods employed, in this case the VPT

and MRLCC results are somewhat more sensitive than the other methods, but the variation is at most 2 mhartree.

IV. DISCUSSION

From the results presented above a number of general statements can be made concerning the utility of VPT and MRLCC. First, as expected from the analysis given in Sec. II, the energies obtained are generally quite close to the QCSDCI results. Similar to the QCSDCI results, the total energies were somewhat lower than the full CI results for CH_2 and for all but a single point for BeH_2. Nevertheless, the variation with geometry or the calculated excitation energies were in good agreement with full CI results where available.

Second, the energies were found to be nearly insensitive to the size of the zeroth-order space. For the $1\,^1A_g$ of ethylene and the $^1A_1$ state of methylene the MRLCC and VPT results were less sensitive than the SDCI results, but for the $^3B_{1u}$ state of ethylene the reverse was true. However, some of the variation in the $^3B_{1u}$ state's energy in the MRLCC and VPT cases appears to be due to the use of perturbation-theory selection, as can be seen by comparing the one-reference results with and without perturbation-theory selection. For

TABLE VII. Excitation energies for B_{1u} states of C_2H_4.[a]

State	Reference[b]	RSPT	SDCI	QCSDCI	(MR)LCC	VPT
unselected						
$1\,^1B_{1u}$	2	6.5	8.4	8.3	8.2	8.2
$2\,^1B_{1u}$	2	7.6	9.4	9.3	9.2	9.3
$1\,^3B_{1u}$	1	4.7	4.5	4.5	4.5	4.6
PT selected						
$1\,^1B_{1u}$	2	6.5	8.4	8.3	8.2	8.3
$2\,^1B_{1u}$	2	7.6	9.4	9.3	9.3	9.3
$1\,^3B_{1u}$	1	4.7	4.5	4.6	4.6	4.6

[a] All energies are in eV and are relative to the two-reference calculation for the $1\,^1A_g$ state of Table VI.

small molecules one expects the differences between SDCI and the potentially size-consistent methods to be less in evidence, due to the small number of electrons and thus smaller correlation energy.

Third, the use of the perturbation-theory selection criterion yields results in good agreement with those obtained from the complete calculations. Certainly this approach is not needed for the systems described here, but in examinations of extended systems, where such size-consistent methods will be of greatest use, it may mean the difference between performing and not performing the calculation.

The accuracy obtained for excitation energies was quite good. The results on the $^1B_{1u}$ states of ethylene are particularly intriguing. There it is shown that both VPT and MRLCC are capable of describing two states of the same symmetry that are quite close in energy (within 1.2 eV).

Concerning comparisons between VPT and MRLCC, it is clear that in most instances presented above the distinction is minute. The one case where a true difference arises concerns the anomalous point on the BeH_2 potential [$r(BeH_2) = 2.75b$, $r(H_2) = 2.55b$]. In this case the MRLCC results yield a smoother potential, one for which the variation relative to the full CI result is smaller. On the other hand, this result may be somewhat artifactual. In Table III it was seen that even though the VPT and MRLCC energies were reasonably close to the full CI value, the ratios of the coefficients ($\times$) of the two dominant configurations were quite far from the SDTQ value until the reference space was expanded significantly. This implies that in the two-reference case the MRLCC wave function was still in error, even though no severe discontinuities arose in the surface at this point. In the VPT case large values were obtained for the norm of $Q_0\Psi_1$. In fact, the norm of $Q_0\Psi_1$ can be used as a signal for the possible inaccuracy of Ψ_0, due to large changes in the relative weights of the reference configurations brought on by correlation.

V. CONCLUSIONS

An application of Hylleraas variational perturbation theory is presented for the treatment of the many-electron correlation problem in molecular systems. The choice made for the partitioned Hamiltonian leads to equations similar to the MRLCC approach of Laidig and Bartlett.[10] The method has been implemented with the use of a perturbation-theory selection criterion for the configurations to be explicitly treated.

Results are presented for three systems, BeH_2, CH_2, and C_2H_4. In general, the results are in good agreement with either full CI or good truncated CI results for these systems. The method is shown to be relatively insensitive to expansion of the zeroth-order space and the perturbation-theory selected results are in close agreement with the unselected results, suggesting that the method will be useful for large systems where the unselected calculation is intractable. The excitation energies obtained are in good agreement with past results and it is shown that excited states within a given symmetry can be obtained with no greater effort than the lowest state of a symmetry.

Note added in proof: In further applications of variational perturbation theory to the calculation of one-electron properties we have shown that inclusion of the orthogonal complement functions of the reference space causes the method not to be strictly size consistent. In test calculations this size inconsistency manifests itself most strongly in changes in the one-electron properties, while the total energy remains very nearly size consistent. In a forthcoming article on one-electron properties we will examine the origins of this size inconsistency for variational perturbation theory.

APPENDIX

The method used to solve the inhomogeneous linear equations of Eq. (12) is briefly summarized here. It is entirely analogous to the method used in this group to obtain the lowest eigenvalues of large real-symmetric matrices.[22] We rewrite Eq. (12) as

$$MC = -B \tag{A1}$$

with $B = QHP$ $(N - 1 \times 1)$ and $M = Q(H - E_0)Q$ $(N - 1 \times N - 1)$. At present it is assumed that Q is in terms of the Φ_k $(k = 1, n - 1)$ and the $\phi_i (i = n + 1, N)$. We write

$$C = \sum_{i=1}^{r} a_i X_i, \tag{A2}$$

where X_i are normalized vectors of length $N - 1$, where $N - 1$ is large. For $r \ll N - 1$ one forms the scalar product of Eq. (A1) with each X_i and obtains a set of r linear inhomogeneous equations that can then be solved using some standard in-core algorithm for the solution of simultaneous equations. To obtain the correction vector X_{r+1} we proceed as follows. Assume C_r is the present approximant to C_{exact} and δC the difference between C_r and C_{exact}. Then Eq. (A1) can be rearranged to give

$$M\delta C = -(B + MC_r). \tag{A3}$$

Of course, M is too large to invert, otherwise one could immediately obtain δC. Instead we approximate M on the left-hand side by a matrix (M_D) containing only the diagonal elements of M, and obtain

$$\delta C = -M_D^{-1}(B + MC_r). \tag{A4}$$

Orthogonalization of δC to Ψ_0 and the preceding X_i and normalization of the result yields X_{r+1}. One then projects with X_{r+1} on HX_i $(i = 1, r+1)$ and solves the small set of linear equations to obtain a new set of a_i. This process is repeated until convergence is reached.

In practice, we have implemented the above procedure to be used with the output of our CI program, which produces H $(N \times N)$ in untransformed form over ϕ_i $(i = 1, N)$. To avoid transforming H into the form corresponding to the

configurations Ψ_0, Φ_k ($k = 1, n - 1$), and ϕ_i ($i = n + 1, N$) we have made the following modifications. First, the $n - 1$ Φ_k are treated as the first $n - 1$ X_i, M, and M_D are taken to be in their untransformed form, and B is taken in the form $H\Psi_0$, i.e., it is half-transformed. Since C_r is written in terms of the $n - 1$ Φ_k and any additional expansion vectors, Eq. (A4) may be viewed as a half-transformed version of the system of equations from Eq. (13). One can show that the only components of δC in Eq. (A4) that would differ from those obtained using a fully transformed B, M, and M_D are those of the Φ_k and possible contributions from Ψ_0. However, each δC is orthogonalized to all preceding X_i and to Ψ_0. Thus, all components of Φ_k and Ψ_0 are removed from δC, and no errors arise through the use of the untransformed M.

[1] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).

[2] B. H. Brandow, Rev. Mod. Phys. **39**, 771 (1967).

[3] F. Coester and H. Kummel, Nucl. Phys. **17**, 477 (1960).

[4] (a) J. Cizek, J. Chem. Phys. **45**, 4256 (1966); (b) J. Cizek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).

[5] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. **14**, 545 (1978).

[6] R. J. Bartlett and G. D. Purvis, Int. J. Quantum Chem. **14**, 561 (1978).

[7] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[8] J. Noga, R. J. Bartlett, and M. Urban, Chem. Phys. Lett. **134**, 126 (1987).

[9] (a) H. Nakatsuji and K. Hirao, J. Chem. Phys. **68**, 2053 (1978); (b) H. Nakatsuji, Chem. Phys. Lett. **67**, 334 (1979).

[10] (a) W. D. Laidig and R. J. Bartlett, Chem. Phys. Lett. **104**, 424 (1984); (b) W. D. Laidig, P. Saxe, and R. J. Bartlett, J. Chem. Phys. **86**, 887 (1987).

[11] B. Jeziorski and H. J. Monkhorst, Phys. Rev. A **24**, 1668 (1981).

[12] I. Lindgren, Int. J. Quantum Chem. Quantum Chem. Symp. **12**, 33 (1978).

[13] (a) D. Mukherjee, Chem. Phys. Lett. **125**, 207 (1986); (b) D. Sinha, S. K. Mukhopadhyay, M. D. Prasad, and D. Mukherjee, *ibid.* **125**, 213 (1986).

[14] (a) A. Banerjee and J. Simons, Int. J. Quantum Chem. **19**, 207 (1981); (b) J. Chem. Phys. **76**, 4548 (1982).

[15] H. Baker and M. A. Robb, Mol. Phys. **50**, 1077 (1983).

[16] K. Tanaka and H. Terashima, Chem. Phys. Lett. **106**, 558 (1984).

[17] R. J. Buenker and S. D. Peyerimhoff, Theor. Chim. Acta **35**, 33 (1974).

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

[19] E. A. Hylleraas, Z. Phys. **65**, 209 (1930).

[20] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems* (Springer, Berlin, 1957), p. 122.

[21] L. I. Schiff, *Quantum Mechanics*, 3rd ed. (McGraw–Hill, New York, 1968), Chap. 8.

[22] E. R. Davidson, J. Comp. Phys. **17**, 87 (1975).

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

[24] W. Kutzelnigg, Chem. Phys. Lett. **35**, 283 (1975).

[25] (a) S. R. Langhoff and E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974); (b) E. R. Davidson and D. W. Silver, Chem. Phys. Lett. **53**, 403 (1977).

[26] R. J. Buenker, S.-K. Shih, and S. D. Peyerimhoff, Chem. Phys. **36**, 97 (1979).

[27] R. Ahlrichs, Comput. Phys. Commun. **17**, 31 (1979).

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

[29] The MELD series of electronic structure codes was developed by L. E. McMurchie, S. T. Elbert, S. R. Langhoff, and E. R. Davidson, and was extensively modified by D. Feller and D. C. Rawlings.

[30] C. F. Bender and H. F. Schaefer III, J. Chem. Phys. **55**, 4798 (1971).

[31] R. J. Cave and E. R. Davidson, J. Phys. Chem. **91**, 4481 (1987).

[32] L. E. McMurchie and E. R. Davidson, J. Chem. Phys. **66**, 2959 (1977).

[33] G. D. Purvis III, R. Shepard, F. B. Brown, and R. J. Bartlett, Int. J. Quantum Chem. **23**, 835 (1983).

[34] T. H. Dunning, Jr., J. Chem. Phys. **53**, 2823 (1970).

[35] B. R. Brooks and H. F. Schaefer III, J. Chem. Phys. **68**, 4839 (1978).

[36] K. Kuchitsu, J. Chem. Phys. **44**, 906 (1966).

[37] The dominant configurations are defined as follows. The configurations were ordered on the basis of the size of the square of the coefficients in the final wave function. Where two or more spin couplings exist for a given orbital occupation the sum of the squares of all spin couplings was used to order the configuration. Thus, it is more accurate to say that the dominant n-electron orbital products were selected, and all spin-adapted configuration state functions associated with them were included in the reference space.