

AD-A130 040

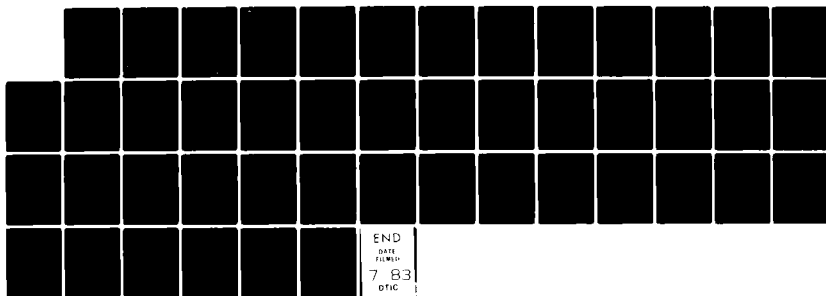
MOLECULAR INTERACTIONS WITH MANY-BODY METHODS(U)
FLORIDA UNIV GAINESVILLE QUANTUM THEORY PROJECT
R J BARTLETT 30 DEC 82 AFOSR-TR-83-0551 AFOSR-82-0026

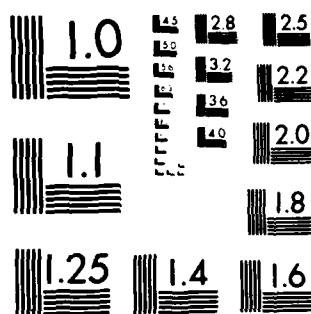
1/1

UNCLASSIFIED

F/G 20/5

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 83-0551

3

ADA130040

ANNUAL TECHNICAL REPORT

AFOSR-
Grant Number 82-0026

MOLECULAR INTERACTIONS WITH
MANY-BODY METHODS

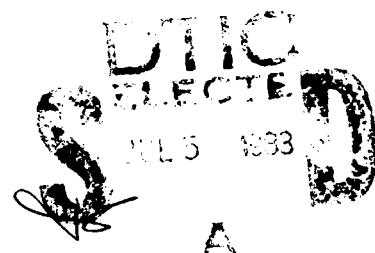
TO

U. S. AIR FORCE
OFFICE OF SCIENTIFIC RESEARCH

December 30, 1982

Principal Investigator:
Rodney J. Bartlett

Quantum Theory Project
362 Williamson Hall
University of Florida
Gainesville, FL 32611



Approved for public release;
distribution unlimited.

DTIC FILE COPY

83 07 01 063

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0551	2. GOVT ACCESSION NO. AD-A130040	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MOLECULAR INTERACTIONS WITH MANY-BODY METHODS		5. TYPE OF REPORT & PERIOD COVERED Annual Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Rodney J. Bartlett		8. CONTRACT OR GRANT NUMBER(s) Grant # 82-0026 AFOSR-82-0026
9. PERFORMING ORGANIZATION NAME AND ADDRESS Quantum Theory Project University of Florida Gainesville, FL 32611		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2301/A4
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE December 30, 1982
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Dr. Ralph Kelley Air Force Office of Scientific Research Building 410 Bolling AFB, D. C. 20332		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Many-body perturbation theory, <u>ab initio</u> quantum mechanical methods, chemical lasers, plume detection, couple-cluster approaches		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Modern military technology has become highly dependent on a detailed knowledge of atom-molecule and molecule-molecule interactions. This type of information is required in diverse defense applications including chemical laser development, in the detection and modeling of plumes, and in the decomposition of energetic materials. The description of forces governing molecular reactions is provided by potential energy surfaces. These surfaces are the crucial first step in dynamics calculations that provide required information about state-to-state cross-sections and rate constants. Since potential energy surfaces are not generally available		

DD FORM 1 JAN 73 1473

UNCLASSIFIED

~~UNCLASSIFIED~~

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. ABSTRACT

from experiment, the most reliable approach to their determination lies in the development and application of predictive ab initio quantum mechanical methods. The following annual report describes our research on the development of many-body perturbation theory (MBPT) and related infinite-order coupled-cluster (CC) methods for potential energy surfaces.

In the past year, several things have been accomplished. The coupled-cluster model has been generalized to include all contributions from single and double excitation operators, (CCSD), which means the model accurately includes quadruple excitations and some triple excitation effects. The other fourth-order triple excitation terms have also been added [CCSD+T(4)]. This very highly correlated wavefunction shows excellent promise for describing energy surfaces accurately. Comparisons made with full CI, which includes all possible excitations, demonstrate that CCSD+T(4) obtains all of the available correlation energy surface for BeH₂ even shows that CCSD can correctly describe highly degenerate bond breaking processes.

In other work, we have (1) investigated localized orbital descriptions of molecules in the CCSD model to begin to develop approaches that are more suited to large molecules; (2) began the development of multi-reference MBPT/CC method to better describe energy surfaces; and (3) initiated a study of infra-red intensities.

The importance of the infra-red signature to plume identification and detection makes it important to develop techniques for accurately computing the infra-red intensities of molecules, particularly those of the transient free-radicals that occur in plumes. This year we have developed and applied our MBPT/CC techniques to intensities for a series of molecules including diatomic inter-halogens and halogen hydrides that arise in plumes as components of spent fuel.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

TABLE OF CONTENTS

	<u>Page</u>
I. Description of Problem	1
II. Review of Research Accomplishments	4
III. Summary of Research Problems	8
IV. Presentations and Publications	13
References	15
Appendix A: C_{2v} Insertion Pathway for BeH_2 : A Test Problem for the Coupled-Cluster Single and Double Excitation Model	18
Appendix B: SCF and Localized Orbitals in Ethylene: MBPT/CC Results and Comparisons With One-Million Configuration CI.	19



Distribution/	
Availability Code	
Dist	Special
A	

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
NOTICE ON REPRODUCTION OF DTIC
This technical report is approved for public release and distribution.
Distribution Statement: Approved for public release; distribution is unlimited.
MATTHEW J. KEMPER
Chief, Technical Information Division

-- 1 --

ANNUAL REPORT

on

MOLECULAR INTERACTIONS WITH MANY-BODY METHODS

to

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

from

Quantum Theory Project
University of Florida
Gainesville, Florida 32611

I. DESCRIPTION OF PROBLEM

In a wide variety of Air Force applications, highly detailed information about atoms, molecules, and their interactions is required[1-3]. This information is necessary in problems ranging from chemical laser development, to the detection and identification of rocket plumes, to metal clustering and aerosol formations, and even to nuclear weapons effects[1-3].

Probably the most crucial component needed to understand molecular reactions is the potential energy surfaces that serve to describe the attractions among the atoms and molecules[1]. However, such information is not easy to obtain. A certain amount of information about the molecular forces near equilibrium in a bound molecule is available from spectroscopy. Some information about the potential energy surface even in the absence of binding can be provided from crossed molecular-beam experiments. But, in general, potential energy surfaces are not amenable to experimental determination. Instead, other types of experimental observations such as kinetics experiments, coupled with very simple theoretical

models for a surface, are used to infer pieces of information about the parameters of the model such as what the activation barrier might be.

The most direct approach to obtaining detailed information about a potential energy surface is offered by predictive, ab initio quantum mechanical calculations. However, to make it feasible to calculate accurate energy surfaces for molecules, much better and more computationally efficient methods must still be developed.

One such approach, namely many-body perturbation theory (MBPT) [4-15] and its infinite-order extensions termed coupled-cluster (CC) methods [11,16-22] offer a number of attractive features that the more traditional configuration interaction approaches lack[21]. Under AFOSR support, we have generalized the theory and developed very efficient computer codes to perform MBPT/CCM calculations and employed these for the first time in large-scale ab initio calculations of potential energy surfaces[11,21]. The successes of this effort have been substantial. These include the determination of a complete force-field for the H_2O molecule, including all force-constants through fourth-order, that is sufficiently accurate that once improved experiments were carried out after our calculations, many of the previously accepted values for the force constants were revised to be more consistent with our predictions[22]. Also, a study of the binding energies of the molecules $B_2H_6 \rightarrow 2BH_3$, $H_3BNH_3 \rightarrow BH_3 + NH_3$, and $H_3BCO \rightarrow BH_3 + CO$ was made that predict these binding energies to within 1 kcal/mole of the accepted experimental values for diborane and borane carbonyl, and make a prediction in the case of borazane in the absence of an experiment[14]. Earlier experiments which gave much higher values for the binding energies of diborane and borane carbonyl than we computed are now completely discounted. Similar successes with studies of the isomerization energy and activation barrier of $HNC \rightarrow HCN$ [23], and $CH_3NC \rightarrow CH_3CN$ [24], the

photodissociation of formaldehyde[25], and various studies of methanol, methoxy, and the formyl radical[26] attest to the reliability of our MBPT/CC methods.

Building upon this work supported by the AFOSR, we carried out extensive studies of the potential energy surface for the two inelastic collisions, $O(^3P) + H_2O$ and $O(^3P) + CO_2$, under contract to the Air Force Rocket Propulsion Laboratory, for the purpose of obtaining vibrational excitation cross-sections that are needed in actual detection devices[27].

Under our new AFOSR grant starting November 1, 1981, we have initiated three exciting new research categories that represent major additions to our earlier pioneering work with many-body methods for energy surfaces. These are:

- Localized orbital descriptions of molecules and their energy surfaces. This work is directed toward developing the theory required for accurate predictions about large molecules.
- The development of the theory and application of multi-reference MBPT/CC techniques. This effort is geared to better describing energy surfaces when several configurations are quite important in an accurate wavefunction and for accurate predictions of excitation energies.
- Predictions of properties like dipole moments and polarizabilities and their derivatives with nuclear displacements. Such derivative quantities are responsible for infra-red and Raman intensities and thereby, offer crucial information for plume signature detection. These properties are also very difficult to predict accurately, placing extreme demands on the accuracy of any quantum mechanical method.

This past year we have made much progress in each of these areas. A survey of some specific accomplishments is presented in Section II, while Section III elaborates on these topics as well as describing a few other new results. Section IV lists the presentations and publications supported by AFOSR from 1981 and 1982. Two of the papers we have published this year are included as Appendices to describe in more technical terms some of our research results.

II. Review of Research Accomplishments

The overall objectives of this research program include the following:

- 1) Develop new, more accurate and more efficient ab initio quantum mechanical methods based upon MBPT and CCM for determining molecular properties, and particularly, potential energy surfaces for molecular interactions.
- 2) Implement these methods in highly efficient, transportable computer codes, to enable computations on potential energy surfaces to be made on an almost routine basis.
- 3) Apply these techniques to a variety of problems that are of interest to AFOSR, and that serve to establish the range of accuracy for MBPT and CCM methods.

In line with these overall objectives, a number of accomplishments have been made under our new AFOSR grant. The main focus for our effort this last year has been the development of formal methods and new computer programs designed to generalize MBPT and CCM theory. In particular, the inclusion of monoexcited clusters in a full CCSD model have been developed and efficiently implemented within the CCM mode. Also fourth-order triple excitation diagram terms including the diagrams required for non-SCF orbitals have been implemented. Using this highly accurate, correlated model, comparisons have been made with full CI (i.e. the best possible solution in a basis set) and other very accurate CI calculations that include higher

categories of excitations. Since CCSD is invariant to wide varieties of transformations, localized orbital techniques for molecules are being investigated to develop methods for large molecules. A summary of achievements follows.

- A. For the first time, the equations for the full CCSD model, built upon the wavefunction, $\psi_{\text{CCSD}} = \exp(T_1 + T_2) | \phi_0 \rangle$, have been obtained and implemented on the computer. Initial applications were to H_2O and BeH_2 where full CI solutions were available. In both cases, CCSD recovers better than 98% of the correlation energy in the basis set[28], while the addition of triple excitations makes the agreement almost exact.
- B. An approximate reaction path for the insertion of Be into H_2 was also determined at the CCSD level and compared with full CI. This problem is very difficult to describe with a single reference function since both Be and H_2 have at least two important configurations along this path. In spite of this difficulty, CCSD agrees with the full CI result to within 1 kcal/mole (i.e. .04 e.v.) for the full reaction path. This result attests to the stability of the CCSD model even in a difficult multi-reference environment[29]. (See Appendix A for the paper.)
- C. In the treatment of large molecules, localized orbitals would appear to be necessary to eliminate the highly delocalized nature of the usual SCF orbitals. This reduces the number of integrals in a calculation down to a manageable size. Furthermore, such orbitals are more conceptual and offer better prospects for transferability to other molecules. This year, we have made the first

many-body calculations with localized orbitals, using the CCSD model and easily generated hybrid bond orbitals. Comparisons with SCF orbital calculations on CH_4 and H_2O are excellent, with the localized orbital CCSD calculation differing from the SCF orbital CCSD result by less than 1 kcal/mole[30].

- D. In a substantial achievement, Schaefer, et. al., have carried out a CI calculation involving more than one million single, double, triple and quadruple excitations on C_2H_4 [31]. We have recently completed a CCSD calculation on the same system which is an improvement over the CI result by 3 kcal/mole. Of this, triple excitations account for about 5 kcal/mole of correlation energy. In addition, CCSD requires 1-2 orders of magnitude less computing time than the CI[32]. (See Appendix B for the manuscript.)
- E. Two full CI calculations have recently been accomplished by Harrison and Handy [33], one for H_2O at geometries displaced from equilibrium and one for the BH molecule. For these benchmark results we have performed full fourth-order perturbation theory SDTQ-MBPT(4) and CCSD+T(4) where the fourth-order triple excitations are added. Our CCSD+T(4) results agree to within 0.27 kcal/mole for H_2O at equilibrium, and 1.0 kcal/mole for a displacement of one-half the OH bond length. For BH in a larger basis set, the difference is only 0.73 kcal/mole.
- F. Preliminary correlated calculations on the derivatives of the dipole moment with respect to displacements of the nuclei in H_2O , demonstrates that SDQ-MBPT(4) (the fourth-order approximation to CCSD) can describe these properties to within a mean error of 7%,

while all previous calculations (none included correlation) have an error ranging from above 30%[34].

A number of other accomplishments involving on-going projects are discussed in the next section.

III. Summary of Research Problems

In the past year we have initiated a localized-orbital many-body method directed toward large molecules and we hope, eventually, to address the transferability of parts of molecules into other bonding environments. Our initial efforts have used very simple bonding orbitals composed of hybrid orbitals on the various atomic centers. In this way, we have studied CH_4 , C_2H_4 , C_2H_2 , C_2H_6 , and H_2O , of which two papers are published [30,32] and a third one under preparation. Each of these molecules have a well-defined hybridization scheme, which is why we chose these systems for an initial study. However, we have implemented the Del'Re procedure [35] to define hybrid orbitals in an unbiased way regardless of bonding environment for more complicated cases.

Our procedure for making simple localized bond orbitals, which is similar to that used in the semi-empirical PCILO scheme [36], is to define an atomic orbital basis set, hybridize the atomic orbitals to establish directionality, and then combine the hybrids together with other hybrid or atomic orbitals to give normal chemical bonds. For CH_4 , these bond orbitals have the general form

$$b = C(sp^3) + \lambda H(1s)$$

where the weight factor, λ , reflects the relative charge distribution. By doubly occupying the $C(1s)$ core orbitals, and the four C-H bond orbitals, a crude zeroth-order localized reference function is constructed. Similarly an antibonding orbital is formed from the negative combination

$$b = C(sp^3) - \lambda H(1s)$$

The remaining polarization functions in the basis set are incorporated by suitable orthogonalization techniques discussed elsewhere [30].

In the CCSD model, the wavefunction is

$$\psi_{\text{CCSD}} = \exp(T_1 + T_2) | \phi_o \rangle$$

where

$$T_1 = \sum_i t_1^a a_i^\dagger$$

$$T_2 = \sum_{\substack{i < j \\ a < b}} t_{ij}^{ab} a_i^\dagger a_j^\dagger b a_b$$

Since T_1 and T_2 commute,

$$\exp(T_1 + T_2) | \phi_o \rangle = \exp(T_2) \exp(T_1) | \phi_o \rangle$$

and any single determinant function ϕ_o can be inter-related to any other Φ_o by $\exp(T_1)$, via Thouless' relation [37]

$$| \Phi_o \rangle = \exp(T_1) | \phi_o \rangle$$

Consequently, even with a poor choice of ϕ_o as given (e.g.) by relatively crude localized bond-orbitals, once the T_1 amplitudes (t_1^a) are converged, one recovers a good single reference function, Φ_o .

However, if it is desirable to use low-order perturbation methods instead of converging to infinite order as in the CCSD model, better perturbation approximations are possible when a better reference function, ϕ_o , is chosen. One way to do this while maintaining the other advantages of localized orbitals is to first do an SCF calculation and then localize the orbitals using a technique like that of Foster and Boys [38] or Edmiston

and Ruedenberg [39]. This has the advantage that the computational reference energy is optimum since the localization procedure just amounts to a unitary transformation of the canonical SCF result. Another advantage is that this type of procedure is completely independent of geometrical considerations which make it easy to follow a molecule along a reaction path while redefining the localized orbitals on the path. We have not yet pursued this localized orbital variation but we plan to this coming year.

Another research area which we are actively pursuing is the development of the multi-reference MBPT/CC model. As discussed in Section II and Appendix A, we have studied the C_{2v} insertion of Be into H_2 because this problem has an extremely difficult multi-reference character. The Be atom has a near degeneracy between its 2s and 2p orbitals. Similarly, when H_2 is separated to permit the Be to insert, the two H_2 configurations, $1\sigma_g^2$ and $1\sigma_u^2$ become nearly degenerate. Consequently, it is extremely difficult to see how one can describe this insertion without including all of the important configurations from the separated products into the reference space for the problem. For the molecule, the multi-reference character is manifested by the fact that two BeH_2 configurations, $1a_1^2 2a_1^2 3a_1^2$ and $1a_1^2 2a_1^2 3b_2^2$ are dominant at different points along the reaction path for the insertion, causing an ambiguity in which to choose as a reference for any single reference approach.

By changing the reference function along the path, CCSD was shown to accurately reproduce the full CI result for this problem[29]. However, having to change reference functions is undesirable, so for many problems of this type, a multi-reference technique is requisite.

We have addressed the multi-reference problem this year in two ongoing efforts: (1) a finite-order small reference MBPT approach; and, (2) an

infinite-order large reference linearized multi-configuration coupled-cluster method (L-MCCCM). Each approach has its advantages and limitations. Since almost nothing is known about the numerical characteristics of such methods, it is first necessary to explore alternatives. Besides the BeH_2 application, we are also studying the potential energy curve for the N_2 molecule, which is an extremely difficult case. Each of these multi-reference approaches also should enable the calculation of excitation energies, as well, which is a goal of this research program.

In our work on properties other than the energy, we have focused on infra-red intensities due to their pertinence in plume detection and identification. Infra-red intensities are obtained from the derivative of a molecular dipole moment with respect to bond and angle displacements and offer an exceedingly difficult property to compute accurately with even the best quantum mechanical methods. Our initial applications have been to the prototype molecule H_2O , where, even for such a simple system, the discrepancies between experiment and theory have been severe. In our calculations, SDQ-MBPT(4) has been able to predict these dipole derivatives to within a mean error of 7% while other calculations have been in error by over 30% [34]. We are currently investigating the effect of infinite-order single excitations via the CCSD model on those predictions and the effect of triple excitations. Other applications have been made to inter-halogen molecules and the hydrogen halides because of their relevance to plume signatures. This work includes a prediction of the experimentally unknown dipole moment of the IF molecule.

In related applications, the highly unusual HIF inorganic free-radical molecule [40] which is stable to decomposition to $\text{HI}+\text{F}$ or $\text{IF}+\text{H}$ has been studied to locate its transition state for the exoergic decomposition to $\text{HF}+\text{I}$.

This very unusual system has two nearly equal minima at an HIF angle of 127° and 80° suggesting the possibility of the H atom tunnelling through the double-well potential. We expect to publish this work soon.

Also, we have studied several electronic states of the HCF molecule, which is a common chemiluminescent constituent of plumes. A number of interesting questions pertain to the relative stability of the singlet and triplet form of this free-radical and the location of other low-lying excited states[41]. Our calculations are the most accurate available for this unusual species.

IV. PRESENTATIONS AND PUBLICATIONS

Presentations (1982)

- R.J. Bartlett, "Progress in Many-Body Methods for Electron Correlation in Molecules," Invited Speaker, Fourth International Congress of Quantum Chemistry, Uppsala, Sweden, June 14-19, 1982.
- R.J. Bartlett, "The Coupled-Cluster Single and Double Excitation Model with Localized Orbitals," Invited Speaker, Meeting on Electron Correlation in Molecular Systems: Methods and Applications, Orenäs, Sweden, June 21-23, 1982.
- R.J. Bartlett, "Many-Body Methods for Electron Correlation in Molecules," Invited Speaker, American Chemical Society Symposium on "Electron Correlation in Molecules," Las Vegas, Nevada, March 29-April 2, 1982.
- R.J. Bartlett, "A Perspective on Many-Body Methods in Quantum Chemistry," Invited Speaker, Southeastern Theoretical Chemistry Conference, Birmingham, Alabama, May 21-22, 1982.
- R.J. Bartlett, "Many-Body Theory and Its Applications to Molecules", Ohio State University, February 11, 1982.
- G.D. Purvis, "Non-Hartree-Fock Terms in Coupled Cluster Theory," Invited Speaker, Meeting on Electron Correlation in Molecular Systems: Methods and Applications, Orenäs, Sweden, June 21-23, 1982.
- W.D. Laidig, "Localized Orbitals in Coupled-Cluster Singles and Doubles Model," Sanibel Meeting on the Quantum Theory of Matter, Palm Coast, Florida, March 1982.
- G.D. Purvis, "Coupled-Cluster Singles and Doubles Model," Sanibel Meeting on the Quantum Theory of Matter, Palm Coast, Florida, March 1982.
- G.D. Purvis, "Orbital Optimization in Coupled-Cluster Methods," Southeastern Theoretical Chemistry Conference, Birmingham, Alabama, May 21-22, 1982.
- W.D. Laidig, "Localized Orbitals in Coupled-Cluster Singles and Doubles Model," Southeastern Theoretical Chemistry Conference, Birmingham, Alabama, May 21-22, 1982.

Presentations (1981)

- R.J. Bartlett, "Vibrational Excitation Cross Sections for $O(^3P) + H_2O$ and $O(^3P) + CO_2$ and Plume Detection," Invited Speaker, Sanibel Symposium on the Quantum Theory of Matter, March 9, 1981.
- R.J. Bartlett, "Contribution of Many-Body Methods to the Study of Energetic Materials," Invited Speaker, Workshop on Fundamental Research Direction for the Decomposition of Energetic Materials, Berkeley, California, January 20, 1981.

- R.J. Bartlett, "Molecular Applications of Many-Body Methods," Department of Chemistry, University of Florida, Gainesville, Florida, March 16, 1981.
- R.J. Bartlett, "Multireference Many-Body Methods for Potential Energy Surfaces," Invited Speaker, NRCC-Sponsored Meeting on Perturbation Theory, Seattle, Washington, July 7, 1981.
- G.D. Purvis, "Computational Implementation of the Coupled-Cluster Model," Invited Speaker, NRCC-Sponsored Meeting on Perturbation Theory, Seattle, Washington, July 7, 1981.
- R.J. Bartlett, "Molecular Applications of Many-Body Methods," Department of Chemistry, Michigan Technological University, Houghton, Michigan, April 6, 1981.

Publications (1982)

- G.D. Purvis and R.J. Bartlett, "Coupled-Cluster Single and Double Excitation Model: The Inclusion of Disconnected Triples," J. Chem. Phys. 76, 1910 (1982).
- W.D. Laidig, G.D. Purvis, and R.J. Bartlett, "Localized Orbitals in the Coupled-Cluster Singles and Doubles Model," Int. J. Quantum Chem., Symp. 16, 561 (1982).
- G.D. Bent, G.F. Adams, R.H. Bartram, G.D. Purvis, and R.J. Bartlett, "Many-Body Perturbation Theory Electronic Structure Calculations for the Methoxy Radical. I. Determination of Jahn-Teller Energy Surfaces, Spin-Orbit Splitting, and Zeeman Effect," J. Chem. Phys. 76, 4144 (1982).
- R.J. Bartlett and G.D. Purvis, "Many-Body Methods for Molecular Properties: Dipole Derivatives for H₂O" submitted, J. Chem. Phys.
- G.D. Purvis, R. Shepherd, F. Brown, and R.J. Bartlett, "The Insertion of Be into H₂: A Comparative Study of Many-Body Methods for Energy Surfaces", Proceedings of Fourth International Congress of Quantum Chemistry, Int. J. Quantum Chem., in press.

Publications (1981)

- W.D. Laidig, G.D. Purvis and R.J. Bartlett, "SCF and Localized Orbitals in Ethylene: MBPT and Coupled-Cluster Results and Comparison with One-Million Configuration CI, submitted, Chem. Phys. Letters.
- R.J. Bartlett, "Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules," Annual Reviews of Physical Chemistry 32, 359 (1981).
- R.J. Bartlett and G.D. Purvis, "Electron Correlation in Large Molecules with Many-Body Methods," Annals N.Y. Academy of Sciences 367, 62 (1981).

REFERENCES

1. L. Wilson, "Report of Workshop on State-to-State Molecular Dynamics," AFOSR Physics 77-111, June, 1976.
2. T.C. Collins and D. Stewart, "Report of Conference on Chemistry and Physics of Plumes," AFOSR Physics 77-062, July, 1976.
3. "Rocket-Plume High Altitude Intra-Red Workshop," August 3-4, 1978, Air Force Geophysics Laboratory, sponsored by AFRPL.
4. K.A. Brueckner, Phys. Rev. 97, 1353 (1955); 100, 36 (1955).
5. J. Goldstone, Proc. Royl Soc. A239, 267 (1957).
6. H.P. Kelly, Advan. Chem. Phys. 14, 129 (1969); Phys. Rev. 131, 684 (1963); 136B, 896 (1964); 144, 39 (1966).
7. E.S. Chang, R.T. Pu, and T.P. Das, Phys. Rev. 174, 16 (1968); N.C. Dutta, C. Matsubara, R.T. Pu, and T.P. Das, Phys. Rev. 177, 33 (1969).
8. J. Cizek and J. Paldus, Advan. Quantum Chem. 9, 105 (1975).
9. R.J. Bartlett and D.M. Silver, Intern. J. Quantum Chem. 95, 183 (1975).
10. R.J. Bartlett and I. Shavitt, Chem. Phys. Letters 50, 190 (1977); Int. J. Quantum Chem. 11S, 165 (1977).
11. R.J. Bartlett and G.D. Purvis, Proceedings of American Conference on Theoretical Chemistry, Int. J. Quantum Chem. 14, 561 (1978).
12. M.A. Robb, Chem. Phys. Letters 20, 274 (1973); S. Prime and M.A. Robb, Chem. Phys. Letters 35, 86 (1975).
13. D.L. Freeman and M. Karplus, J. Chem. Phys. 64, 2641 (1976).
14. L.T. Redmon, G.D. Purvis, and R.J. Bartlett, J. Am. Chem. Soc. 101, 2856 (1976).
15. J.A. Pople, J.S. Binkley, and R. Seeger, Int. J. Quantum Chem. 10S, 1 (1976).
16. J.A. Pople, J.S. Binkley, and R. Krishnan, Int. J. Quantum Chem. 11S, 149 (1977).
17. H. Kummel, Lectures on the Many-Body Problems, ed. E.R. Caianiello, Academic Press, New York, 1962.
18. J. Cizek, J. Chem. Phys. 45, 4256 (1966); Adv. Chem. Phys. 14, 35 (1969).
19. J. Cizek and J. Paldus, Int. J. Quantum Chem. 5, 359 (1971).

20. J. Paldus, J. Cizek, and I. Shavitt, Phys. Rev. A5, 50 (1972).
21. R.J. Bartlett and G.D. Purvis, Physica Scripta 21, 255 (1980).
22. R.J. Bartlett, I. Shavitt, and G.D. Purvis, J. Chem. Phys. 71, 281 (1979).
23. L.T. Redmon, G.D. Purvis, and R.J. Bartlett, J. Chem. Phys. 72, 986 (1980).
24. L.T. Redmon, G.D. Purvis, and R.J. Bartlett, J. Chem. Phys. 69, 5386 (1978).
25. G. Adams, G. Bent, R.J. Bartlett, and G.D. Purvis, J. Chem. Phys. 75, 834 (1981).
26. G. Adams, G. Bent, G.D. Purvis and R. J. Bartlett, J. Chem. Phys. 71, 3697 (1979); Chem. Phys. Letters 87, 311 (1982).
27. R.J. Bartlett and M.J. Redmon, "Vibrational Excitation Cross Section for the O + H₂O and O + CO₂ Collisions," Final Report, AFRPL, Contract #F04611-79-C-0024.
28. G.D. Purvis and R.J. Bartlett, J. Chem. Phys. 76, 1910 (1982).
29. G.D. Purvis, R. Shepherd, F. Brown, and R.J. Bartlett, Proceedings of Fourth International Congress of Quantum Chemistry, Int. J. Quantum Chem., in press. (This paper is included as Appendix A to this report.)
30. W.D. Laidig, G.D. Purvis, and R.J. Bartlett, Int. J. Quantum Chem. Symp. 16, 561 (1982).
31. P. Saxe, D.F. Fox, H.F. Schaefer and N.C. Handy, J. Chem. Phys., in press.
32. W.D. Laidig, G.D. Purvis, and R.J. Bartlett, Chem. Phys. Letters, in press.
33. R.J. Harrison and N.C. Handy, Chem. Phys. Letters, in press.
34. R.J. Bartlett and G.D. Purvis, to be published J. Chem. Phys.
35. G. Del Re, Theor. Chim. Acta 1, 188 (1963).
36. S. Diner, J.P. Malrieu, and P. Claverie, Theor. Chim. Acta 13, 1 (1969).
37. D.J. Thouless, Nucl. Phys. 21, 225 (1960); 22, 78 (1961).
38. J.M. Foster and S.F. Boys, Rev. Mod. Phys. 32, 300 (1960).
39. C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. 35, 457 (1963); J. Chem. Phys. S43, 97 (1965); in Quantum Theory of Atoms, Molecules, and the Solid State, ed. P.O. Löwdin. Academic Press, N.Y. p. 263 (1966).

40. James J. Valentini, M.J. Coggiola, and Y.T. Lee, J. Am. Chem. Soc. 98, 853 (1976); Discussions Faraday Soc. 62, 232 (1977).
41. C.W. Matthews, J. Chem. Phys. 45, 1068 (1966); V. Staemmler, Theor. Chim. Acta 35, 309 (1974).

APPENDIX A

C_{2v} Insertion Pathway for BeH_2 : A Test
Problem for the Coupled-Cluster Single
and Double Excitation Model.

C_{2v} Insertion Pathway for BeH₂: A Test Problem for the Coupled-Cluster Single and Double Excitation Model*

GEORGE D. PURVIS III

Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32611, U.S.A.

RON SHEPARD

Chemistry Division, Theoretical Chemistry Group, Argonne National Laboratory, Argonne, Illinois 60439, U.S.A.

FRANKLIN B. BROWN

Department of Chemistry, Ohio State University, Columbus, Ohio 43210, U.S.A.

RODNEY J. BARTLETT

Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32611, U.S.A.

Abstract

The coupled-cluster single and double excitation model (CCSD) is applied to an energy path for the insertion of Be into H₂ and compared to the full configuration interaction (FCI) and full valence-multiconfiguration self-consistent field (FV-MCSCF) results. This model problem is a severe test of a single-reference-function correlated method since two configurations are heavily weighted in the FCI description. CCSD is demonstrated to describe the FCI results using a single reference function which, however, changes orbital characteristics along the sampling path. In this case CCSD gives excellent agreement with the FCI results.

1. Introduction

A number of applications of coupled-cluster methods (CCM) [1-4] and the closely related many-body perturbation theory (MBPT) [5-8] for molecules and their energy surfaces have been made in the last few years.[†] In essentially every case [10], the calculation used a single-reference configuration, usually of SCF form as a restricted (RHF) or unrestricted Hartree-Fock (UHF) function. With a single reference function, MBPT/CCM techniques offer uniform, size-extensive, and highly efficient procedures for the treatment of electron correlation.

Although a number of approaches to the multireference MBPT/CCM problem exist [11-16], the applications of these techniques have been severely limited [17-19], due to the comparative complexity of the multidimensional approach. As a consequence,

* Supported by the U.S. AFOSR under Grant No. 82-0026.

† For a recent review of CCM/MBPT and their applications, see Ref. 9.

the predominant criticism of MBPT/CCM has been the practical limitation to a single reference function, particularly in a region of bond breaking when more than one configuration is highly weighted in a configuration-interaction (CI) wave function. No doubt such a criticism is justified in many cases, however, the problem deserves closer investigation to accurately assess the limitations of the single-reference approach. For example, the coupled-cluster doubles (CCD) model has been shown to correctly model some severely degenerate problems [20] as has the coupled-cluster single and double excitation model (CCSD) [21].

Furthermore, the multireference procedure is not a universal solution. Besides the comparative complexity, the principal objection to multireference techniques is the ambiguity in the choice of reference configuration. Typically, single and double excitations from the full reference space will be chosen to form the CI wave function. However, to separate a polyatomic molecule like

C

$D-A-B$

into its appropriate fragments, requires that the multireference space contain all configurations that are needed to describe the component subspecies. Consequently, if we are breaking the $A-B$ bond, we typically need to include the configurations that describe those fragments, often even including the united atom configuration(s) [22]. However, when we choose to break another bond, like $A-C$ or $A-D$, we need a new set of reference configurations which yield a new set of single and double excitations to be included in the CI. Furthermore, since a typical objective is to compare reaction paths for the different channels, the different levels of approximation partially invalidate such numerical comparisons. The alternative of including all possible reference configurations for the plausible reaction paths, which would eliminate the ambiguity, is usually not feasible for polyatomic systems since the size of the reference space becomes prohibitive.

Faced with this dilemma, single-reference approaches which allow different reaction pathways to be treated within the same well-defined model should be examined. After all, regardless of the reference function chosen, the corresponding full CI calculation will give the ultimate answer within the basis set. If one can introduce the important correlation effects by a suitable treatment of higher excitations, one should be able to treat even some very difficult problems with single-reference-function coupled-cluster techniques. (When degeneracies are important, the infinite-order CCM can be expected to be more stable than finite-order MBPT [3,8], so we will mainly focus on the use of CCSD in this paper.)

We have previously considered the BeH_2 system with CCSD at a fixed geometry [21]. In this paper, we have expanded this study to include a quasireaction path for the perpendicular C_{2v} insertion of Be into H_2 to form BeH_2 , where several configurations should be important along the energy surface. This system is small enough that the full CI calculation [23,24] can be performed to provide rigorous comparisons. In addition, we report full valence-multiconfiguration self-consistent field (FV-

MCSCF) calculations, which use all of the expected important configurations while simultaneously optimizing the molecular orbitals to obtain the best possible variational fit for that set of configurations on the energy surface [25,26].

2. Synopsis of CCSD

The CCSD theory is presented in detail elsewhere [21]; here we present only a brief summary. The CCSD wave function is

$$\psi_{\text{CCSD}} = \exp(T_1 + T_2)|\phi_0\rangle,$$

where

$$T_1 = \sum_{i,a} t_i^a a^\dagger i,$$

$$T_2 = \sum_{\substack{i,j \\ a>b}} t_{ij}^{ab} a^\dagger i b^\dagger j.$$

We make the convention that the indices i, j, k, l represent orbitals occupied in the reference configuration ϕ_0 , while the indices a, b, c, d are for excited orbitals, with the corresponding creation $a^\dagger$ and annihilation i operators. The two-particle amplitudes are antisymmetrized

$$t_{ij}^{ab} = -t_{ji}^{ab} = -t_{ij}^{ba} = t_{ji}^{ba},$$

where t_i^a are the one-particle amplitudes.

By considering the projections of the Schrödinger equation for the trial wave function ψ_{CCSD} onto the space of single and double excitations, a set of nonlinear equations sufficient to determine amplitudes t_i^a and t_{ij}^{ab} are obtained. Once the equations are solved for the amplitudes, the energy is given by

$$E = \langle \phi_0 | H \exp(T_1 + T_2) | \phi_0 \rangle,$$

$$E = \sum_{\substack{i,j \\ a>b}} t_{ij}^{ab} \langle ab || ij \rangle + \sum_{i,a} f_i^a t_i^a + \sum_{\substack{i,j \\ a>b}} \langle ab || ij \rangle (t_i^a t_j^b - t_j^a t_i^b).$$

The quantity $\langle ab || ij \rangle = \langle ai | bj \rangle - \langle aj | bi \rangle$ is the antisymmetrized two-electron integral, and f_i^a is the general Fock matrix element. Our formulation is completely general [21], admitting non-SCF orbitals where f_i^a , f_j^b , or f_a^b may be nonzero. This generality has found a useful implementation in coupled-cluster calculations using localized orbitals [10].

A consideration of the expansion of the operator $\exp(T_1 + T_2)$ demonstrates the inclusion of all single and double excitations and the disconnected contributions of triple, quadruple, and higher excitations. In the quadruple excitation case, the disconnected contribution $\frac{1}{2} T_2^2 |\phi_0\rangle$ is clearly the dominant part, making the neglect of T_4 a good approximation [3,27]. Based upon perturbation theory arguments, with SCF orbitals the connected triple excitation parts that would arise from T_3 are expected to be larger than those from the disconnected $T_1 T_2$ and T_1^3 contributions [28].

Projection of ψ_{CCSD} onto just single and double excitations decouples the CCSD equations to terms quartic in T_1 , quadratic in T_2 , and the cubic term $T_1^3 T_2$. However, the wave function is *not* truncated, retaining its exponential form. These nonlinear equations are solved iteratively using Padé approximants [29] and a recently proposed reduced linear equations method [30] to assist the convergence. The lowest order iterations correspond to second- and third-order perturbation theory and the fourth-order model limited to single, double, and quadruple excitations SDQ-MBPT(4) [3].

3. Results and Discussion

The insertion of Be into H_2 can be a complicated process. The Be atom is well known for the quasidegeneracy of the $2s$ and $2p$ orbitals, so it is usually thought to have two important configurations: the $1s^2 2s^2$ and the $1s^2 2p^2$ set. The comparative weights of these two configurations are very much a consequence of the orbital set used, with natural orbitals [31], e.g., weighting the second configuration higher than is the case with SCF orbitals. However, only the p^2 components allow the BeH_2 bond to be formed, according to orbital symmetry rules.

The insertion of Be into H_2 also requires that the H_2 bond be effectively broken. Yet, at large internuclear separation, the $1\sigma_u^2$ configuration for H_2 becomes nearly degenerate with the $1\sigma_g^2$ configuration, causing both configurations to be almost equally

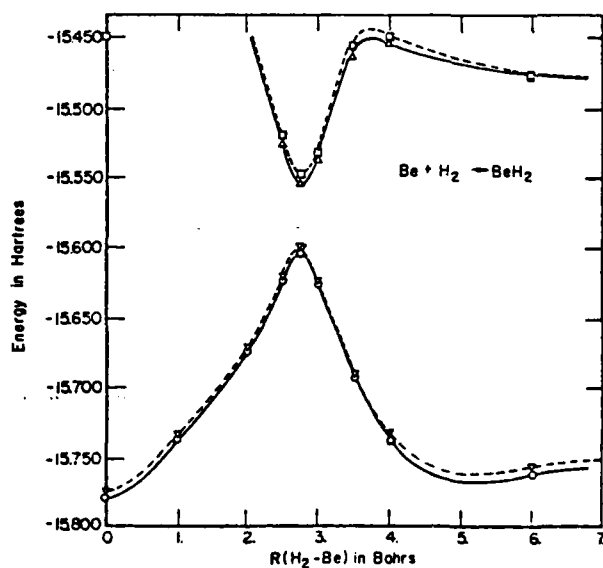


Figure 1. Comparison of FV-MCSCF and FCI for the two lowest 1A_1 states of BeH_2 in C_{2v} symmetry. (∇) Full valence MCSCF1 1A_1 ; ($\square$) full valence MCSCF2 1A_1 ; ($\circ$, Δ) FCI.

TABLE I. Contracted Gaussian basis used for the ten orbital BeH₂ model problem.

		Exponent	Contraction Coefficient
Be	1s	1267.07	.001940
		190.356	.014786
		43.2959	.071795
		12.1442	.236348
		3.80923	.471763
		1.26847	.355183
	1s'	5.693380	-0.028876
		1.555630	-0.177565
		0.171855	1.071630
	1s''	0.057181	1.000
	2p	5.693330	.004836
		1.555630	.144045
		0.171855	.949692
H	1s	19.2406	.032828
		2.8992	.231208
		0.6534	.817238
	1s'	0.17760	1.000

weighted in the CI description of the molecule. Consequently, a rational first approximation to the computation of the interaction energy along a reaction path for BeH₂ would seem to require at least the configurations obtained by coupling together these four configurations of the two components. In BeH₂, though, there are even more complications.

Since the perpendicular insertion of Be into H₂ requires *p*-orbital participation on Be, there is a promotion from Be(2s²) to Be(2p²) near the critical geometry. This causes the principal configuration in BeH₂ to change from 1a²2a²3a² to 1a²2a²1b². In a single reference model, one of these two very important configurations must be treated in the complementary space, placing a severe burden on any single reference model to describe the insertion reaction. To emphasize the multiconfigurational aspect of this problem, at the peak on Figure 1, which corresponds to a Be to H₂ distance of 2.75 bohr, configuration *b* (with SCF orbitals from *b*) has a coefficient of 0.724, while the *a* coefficient is 0.560. In addition, the configuration 1a, ²2a, ²3a, 5a also has a large coefficient of 0.239. When Be is moved to 3.0 bohr, *a* has the larger coefficient of 0.823 compared to 0.294 for *b*. A final complication is that both of these RHF reference functions are unstable since there are lower-lying UHF solutions.

The contracted Gaussian basis set is defined in Table I. The sampling path for the

TABLE II. Coordinates of points along the sampling path^a for the perpendicular insertion of Be into H₂.

Point	Coordinates for H ₂ (X, Y, Z) ^b
A	(0, ± 2.54, 0.0)
B	(0, ± 2.08, 1.0)
C	(0, ± 1.62, 2.0)
D	(0, ± 1.39, 2.5)
E	(0, ± 1.275, 2.75)
F	(0, ± 1.16, 3.0)
G	(0, ± 0.93, 3.5)
H	(0, ± 0.70, 4.0)
I	(0, ± 0.70, 6.0)
J	(0, ± 0.70, 20.0)

^a This sampling path is selected to allow simple comparisons of calculated energies along a path which samples points near the reaction path.

^b Be is located at (0.0,0.0,0.0).

C_{2v} insertion energy surface is defined in Table II. The z coordinate is the distance of Be from the midpoint of the H₂ bond. The y coordinate defines the location of the two H atoms. This sampling path has been determined as follows. MCSCF calculations were performed to locate the vicinity of the transition state for the perpendicular insertion of Be into H₂. On a graph where the Be—H₂ distance is the z axis and the H—H distance is the y axis, a straight line was drawn between the BeH₂ equilibrium point and the transition state geometry. A second straight line was drawn parallel to the z axis at the H₂ equilibrium distance and the intersection with the other line defines the corner of the sampling path. Clearly, this is *not* the reaction path, but has the advantage of permitting a comparison of energies determined with different models at the same point for all methods while passing through regions of the potential energy surface near the reaction path. The true reaction path for the unconstrained insertion of Be into H₂ need not have C_{2v} symmetry.

Before discussing the CCSD results, we show, in Figure 1, a comparison of the full CI with a full valence (FV-MCSCF) (also called a complete active space [32] or full reaction space [33]) calculation for the lowest two 1A_1 states of BeH₂ [25,26]. The active space for the FV-MCSCF includes the $2a_1$, $3a_1$, $4a_1$, $5a_1$, $1b_2$, $2b_2$, and $1b_1$ orbitals, leading to 70 configuration state functions. The wave function is then optimized in terms of the orbital coefficients and configuration state function mixing coefficients. This would appear to be a systematic well-defined approximation that exploits the multireference approach to the fullest, short of adding the remaining CI functions.

The agreement between the FV-MCSCF and full CI (FCI) is quite good, although there is an energy difference varying from 2.4 to 5.4 mhartree that is noticeable even on the 50 mhartree scale of Figure 1. There are also different errors for different parts

of the sampling path, with the greatest errors, surprisingly, being near the end points for the ground state.

Alternatively, we can take a single, albeit poor, reference and see whether CCSD can introduce the requisite corrections. To do so, we first have to define the reference function, and this will itself introduce ambiguities.

Since this is a singlet molecule, we will choose to use a spin-restricted Hartree-Fock determinant as the reference determinant. We thus avoid some of the well-known problems associated with the use of unrestricted Hartree-Fock determinants to describe bond formation and breaking. However, the use of space and spin restricted Hartree-Fock reference determinants here introduces its own symmetry problems. While the symmetry problems associated with UHF functions are well known, the use of restricted Hartree-Fock in the present C_{2v} insertion reaction forces too much symmetry. In particular, the RHF method allows the crossing of 1A_1 determinants $1a_1^2 2a_1^2 3a_1^2$ (hereafter called *a*) and $1a_1^2 2a_1^2 1b_1^2$ (called *b*) configurations, resulting in two ways to obtain an energy for the 1A_1 state. This crossing is allowed at the RHF level because of the extra symmetry element imposed by separating electrons into orbitals. Since the full CI is invariant to the choice of molecular orbitals and allows both configurations to determine their weight in the CI solution, the full CI naturally

TABLE III. Comparison of CCSD, MBPT, and FV-MCSCF with full CI results for the perpendicular insertion of Be into H_2 (C_{2v} symmetry).^a

Point	SCF	SDQ-MBPT (4)	ΔE^b	FV-MCSCF	ΔE^b	CCSD	ΔE^b	FCI
A ^c	-15.7417	-15.7776	1.0	-15.7766	1.0	-15.7788	0.26	-15.7792
B	-15.6996	-15.7357	1.0	-15.7327	2.8	-15.7368	0.26	-15.7372
C	-15.6284	-15.6716	1.0	-15.6709	2.5	-15.6763	0.32	-15.6768
D	-15.5627	-15.6155	4.7	-15.6195	2.2	-15.6217	0.73	-15.6229
E ^d	-15.5212	-15.5806	16.1	-15.5993	2.2	-15.5973	3.6	-15.6029
E ^e	-15.5215	-15.5861	11.8	-15.5993	2.2	-15.6033	-0.25	-15.6076
F ^f	-15.4885	-15.5719	19.5	-15.5993	2.2	-15.5992	2.3	-15.6029
F ^g	-15.5365	-15.6138	7.0	-15.6229	2.3	-15.6272	0.48	-15.6250
G	-15.6187	-15.6673	3.7	-15.6699	2.1	-15.6929	0.21	-15.6932
H	-15.6699	-15.7323	2.8	-15.7313	3.6	-15.7365	0.12	-15.7367
I	-15.6956	-15.7586	2.6	-15.7556	3.3	-15.7608	0.33	-15.7609
J	-15.6979	-15.7580	2.5	-15.7576	3.3	-15.7629	0.41	-15.7629

^a There were 1574 configuration functions used in the FCI expansion, and 141 amplitudes in the $1a_1^2 2a_1^2 3a_1^2$ CCSD expansion. All numbers are in a.u.

^b Energy difference from the FCI is in kcal/mol.

^c MBPT/CCSD energies for points A $\rightarrow$ D are determined from the SCF reference configuration $1a_1^2 2a_1^2 1b_1^2$.

^d Energy obtained from pure symmetry SCF reference $1a_1^2 2a_1^2 1b_1^2$.

^e Energy obtained from a broken symmetry SCF reference configuration. In this case there are 381 amplitudes.

^f Energy obtained from pure symmetry SCF reference $1a_1^2 2a_1^2 3a_1^2$.

^g MBPT/CCSD energies for points F $\rightarrow$ J are determined from the SCF reference configuration $1a_1^2 2a_1^2 3a_1^2$.

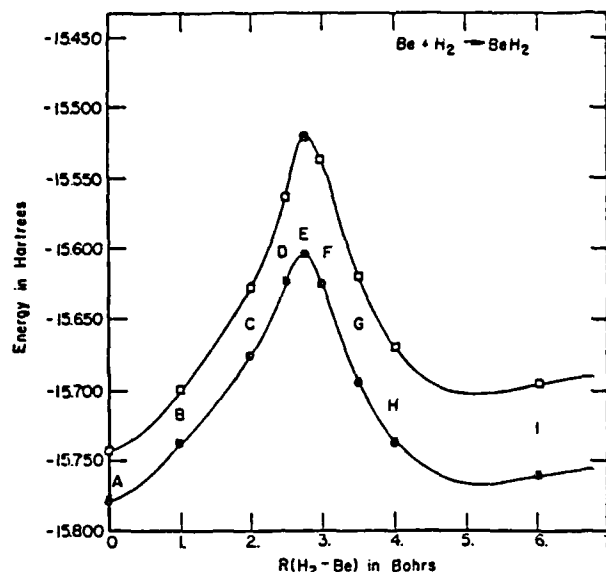


Figure 2. Comparison of (□) SCF and (X) CCSD results with (O) FCI for the ground 1A_1 state of BeH₂ in C_{2v} symmetry. Different RHF reference functions are used at different points on the sampling path. See Table III for detailed specifications. (□) SCF symmetry broken; (X) CCSD symmetry broken.

resolves this problem. MCSCF does too, since one set of optimum orbitals is obtained with participation from several important configurations.

In any nonoptimized orbital limited CI, though, this problem is not resolved. Unlike full CI, a truncated CI is not invariant to this kind of orbital modification, leading to different solutions depending upon which reference function is chosen to define the orbitals. Since CCSD is also a function of its reference configuration, in spite of many other attractive invariance properties [21], it is also possible to obtain separate CCSD results for each choice of reference configuration.

In Table III and Figure 2, we use the reference function *b* for points A–D and *a* for points E–J, whose geometries are defined in Table I. Besides point E, which will be discussed separately, this always corresponds to the lowest CCSD energy. This also reflects the lowest RHF energy except at point D, where the *a* configuration SCF energy is lower by 0.126 hartree while the CCSD result with the *b* configuration is slightly lower by 0.2 mhartree.

As illustrated in Figure 2 and Table III, the CCSD results are very close to the full CI, differing by less than 1 kcal/mol except at point E. This is a surprisingly good result for a single reference approach in a problem with such obvious multiconfigurational character. CCSD has 141 coefficients compared to 1574 configurations in the full CI. Furthermore, the CCSD energies are significantly better than the FV-MCSCF energies, the latter of which results from taking full advantage of the multiconfigurational character within its 70 configuration reference space. Previous calculations [25] in-

APPENDIX B

SCF and Localized Orbitals in Ethylene:
MBPT/CC Results and Comparisons With
One-Million Configuration CI.

SCF AND LOCALIZED ORBITALS IN ETHYLENE: MBPT/CC RESULTS AND COMPARISON
WITH ONE MILLION CONFIGURATION CI*

William D. Laidig, George D. Purvis, and Rodney J. Bartlett

Quantum Theory Project
University of Florida
Gainesville, FL 32611

Abstract

A series of Many-Body Perturbation Theory (MBPT) and Coupled-Cluster (CC) calculations were performed on the ethylene molecule employing both the standard canonical SCF and localized orbitals (LO). The geometry and (4s2p1d/2slp) basis set are identical to that used in the over one million configuration "bench mark" CI calculation of Saxe et. al. Using SCF orbitals we computed full 4th order MBPT, (i.e. SDTQ-MBPT (4)), CC doubles (CCD) and CC singles and doubles (CCSD) energies and compare with their CI results. In addition, LOs were generated in a manner similar to the semi-empirical PCILO procedure and SDQ-MBPT(4) and CCSD calculations were carried out within this basis. Though the SCF and LO reference determinant energies differ by 0.29706 hartrees, the CCSD energy difference is just 0.00170 hartrees. Comparing with Saxe et. al. our most extensive SCF orbital calculation, CCSD plus 4th order triples, was found to be lower in energy than the CI result by -0.00533 hartrees.

*This research is supported by the AFOSR under Grant No. 82-0026 to the University of Florida.

Introduction

Orbitals localized on one or two atomic centers possess a number of advantages over the typical delocalized canonical Self-Consistent-Field (SCF) orbitals for the study of molecular potential energy hypersurfaces[1]. They are conceptual; they are concentrated largely on a single atom or two nearest neighbor atoms; and, they approximate traditional chemical bonding concepts. For example, within an LO basis we might hope to develop bond additivity schemes that would apply for molecular properties such as the total energy and the dipole moment. Transferability of calculated properties of molecular fragments between different molecules might also be expected to be more accurate in a localized basis. Lastly, but most important from a computational point-of-view, is the reduction in the number of molecular orbital (MO) integrals required in a correlated energy calculation. In a typical delocalized SCF basis the number of non-zero integrals grows as the fifth power of n where n is the number of MOs included. However, in a localized basis the number of MO integrals increases only linearly with the number of bonds for extended systems[2]. This reduces the required time for both the integral transformation and the MBPT/CC calculation substantially.

LOs can be generated in a variety of ways, such as the energy invariant transformations among SCF orbitals as in the procedures of Foster and Boys [3] and of Edmiston and Ruedenberg [4]. From the viewpoint of perturbation theory, these procedures have the advantage that the single reference energy is optimum, suggesting better approximations in low-order perturbation theory. Also, the prescription for forming the localized-orbitals is independent of geometrical considerations, providing a well-defined procedure when following a bond formation, for example. However,

the SCF procedure is more complex and requires the generation of additional molecular integrals. From an infinite order viewpoint, models like configuration interaction singles and doubles (CISD) and coupled cluster singles and doubles (CCSD) are invariant to these localized orbital transformations, since only occupied or excited orbitals are mixed among themselves, so the final energies will be the same as when using SCF orbitals provided the MO basis is not truncated.

A more primitive category of localized orbitals are LOs constructed to directly represent the chemical bonds in a molecule as in the PCILO procedure[5]. These methods, built on hybrid and bond orbitals, are usually computationally quite simple. No SCF calculation is required, however, unlike localized SCF orbitals, the CCSD or CISD energy will not be the same as in the SCF orbital case. In this approach a minimum basis atomic orbital set is used to form the LOs, hence, even if a large basis (DZP or better) is used to define the AOs, the resultant single determinant energy does not benefit at all from the polarization functions in the basis. It has been shown that this energy difference is of the same order of magnitude as the correlation energy itself[2], thereby placing a greater burden on the procedure for incorporating corrections to the single-determinant reference function. However, we found in a preliminary study using these simple LO's for CH_4 and H_2O that if extensive correlated calculations are performed, such as CCSD, very good agreement with analogous calculations employing SCF orbitals is obtained even including an accurate description of bond stretching[2]. In the current study we extend this treatment to C_2H_4 .

Correlation effects in ethylene beyond the CISD level have recently been examined by Saxe, Fox, Schaefer and Handy (SF5H)[6], using a (4s2p1d/2slp) contracted Gaussian basis set. In addition to all single and

double excitations from the 1A_g reference, a substantial subset of all triple and quadruple excitations were included in a massive 1,046,758 configuration CI calculation. This set should include the most important triple and quadruple excitations and should give a very good estimate to both the CI containing 1 through 4 fold excitations (CISDTQ) and the full CI results. Since this CI provides some values for higher categories of excitation, some interesting comparisons with MBPT/CC results are possible, which we report.

Comparison of MBPT/CC Results With CI

In our study of ethylene we employed the same basis set and geometry as SFSH[6]. The contracted Gaussian set used is the standard Huzinaga-Dunning[7] double zeta basis set augmented with a set of d functions on each carbon and a set of p functions on each hydrogen. Precisely, the DZP basis sets for carbon and hydrogen are designated (9s5p1d/4s2p1d) and (4s1p/2s1p) respectively. A factor of 1.2 was used to scale the DZ hydrogen basis functions. Six cartesian d functions were placed at each carbon center and each has an exponent of 0.75. The hydrogen p exponent is 1.0. The geometry of ethylene was chosen to be $r(C-C) = 1.330A$, $r(C-H) = 1.076A$, $\angle HCH = 116.6$ which reproduces SFSH's nuclear repulsion energy of 33.51358956 to all reported digits[6].

Using this basis set and geometry a number of energy calculations using LO's and SCF orbitals were performed on the 1A_g ground state of C_2H_4 . All calculations involving the SCF orbitals were done in full D_{2h} symmetry and correspond to the configuration,

$$\dots\dots\dots 3a_g^2 2b_{1u}^2 1b_{2u}^2 1b_{3g}^2 1b_{3u}^2 \quad .$$

The SCF energy is computed to be -78.050530 hartrees which agrees with the reported value of SFSH. Beyond the SCF level we computed the full 4th order MBPT (designated SDTQ-MBPT(4) where SDTQ represents all single, double, triple and quadruple excitation diagram contributions)[8,9], coupled-cluster doubles (CCD)[8] and CCSD[10] energies using the SCF orbitals. (See Reference [11] for the procedure for evaluating the triple excitation diagrams). Consistent with SFSH, all of our correlated calculations kept the two carbon 1s core orbitals "frozen". These results are presented in Table 1, along with selected CI results of SFSH.

A short summary about the SFSH calculation is in order. Initially they carried out a standard one reference CI including all single and double excitations (CISD) in the SCF orbital basis. One natural orbital (NO) iteration was next performed in this 5057 configuration space to obtain a "better" set of orbitals for the remaining CI calculations. To go beyond CISD the orbital space was partitioned into two groups: the valence space and the remainder. Their valence space is composed of the occupied orbitals plus the $1b_{2g}$, $3b_{1u}$, $2b_{2u}$, 4_{ag} , $2b_{3g}$ and $4b_{1u}$ virtual orbitals and corresponds to the minimum basis subspace with the core frozen. This partitioning is unsatisfactory for SCF orbitals since the gap between the orbital energies of the highest valence and lowest non-valence orbital is quite small. This implies that certain energetically close orbitals will be treated in an unequal manner. Using the natural orbitals instead, yields a larger separation and a more satisfactory partitioning.

Two very large CI calculations were carried out using this partitioning. In the smaller calculation the primary natural configuration and all valence space single excitations were used as reference functions (this reference calculation is termed CIS) and then all singles and doubles

from the reference functions were used in a CI (CIS+SD). This procedure yields 109,473 configurations including all single, double and a subset of triple excitations with respect to the NO reference. In their largest CI all single and double excitations in their valence space plus the primary configuration were selected as reference functions and all single and double excitations were generated out of this subspace (CISD+SD). This 1,046,768 configuration basis is composed of all single and double excitations and a subset of all triple and quadruple excitations from the NO reference.

In their preliminary CISD calculations a correlation energy of -0.27758 was found using SCF orbitals, and -0.27748 with NO orbitals. Comparing our MBPT/CC results, these values are close to our second order MBPT energy, E_2 , of -0.27222 hartrees. Adding in E_3 , the third order contribution, a total correlation energy of -0.29655 hartrees is obtained. In their most complete calculation, CISD+SD, their total energy is -78.35451 hartrees yielding a correlation energy of 0.30398 hartrees. From our fourth-order calculation, SDQ-MBPT(4) and infinite order CCSD result, we compute correlation energies of 0.30004 and 0.30116 hartrees, respectively, or about 98.7% and 99.1% of the CISD+SD correlation energy.

In CCSD and its fourth-order approximation, SDQ-MBPT(4) [8], the contribution from the triple excitation diagrams is neglected. With SCF orbitals these diagrams first appear in MBPT in the 4th order (see Ref. [8,9]) and we designate this energy contribution, E_{4T} . Computing E_{4T} we obtain an energy of -0.00815 hartrees. This is effectively identical with the E_{4T} contribution of Frisch et. al. of -0.00815 hartrees in the slightly smaller 6-31G* basis set [12]. Subtracting the CIS+SD and CISD energies yields an estimate of the importance of triple excitations in the

CI framework of -0.00701 hartrees, which is similar to our MBPT result.

Adding E_{4T} to the SDQ-MBPT(4) and CCSD energies our two best estimates of the correlation energy are -0.30840 and -0.30931 hartrees respectively. The corresponding total energies for the SDTQ-MBPT(4) and CCSD+T(4) models are -78.35873 and -78.35984 hartrees. These energies are below the CISD+SD value by 4.4 and 5.3 mh, respectively. Since the triple excitation selection can account for about 1 mh of this difference, it would appear that as much as 3-4 mh arises from the subset of quadruple excitation configurations chosen.

To analyze this in a bit more detail, we can make some approximate comparisons for the contribution of quadruple excitations to the CI. As we have discussed in detail elsewhere[13], the contribution of quadruple excitations to CI is approximately given by the difference of the fourth-order linked and unlinked quadruple excitation diagrams of

$$E_{CI}(Q) \approx E_{4Q} + E_2\Delta \quad (1)$$

The value of the unlinked diagrams depends upon $\Delta = \langle \phi_1 | \phi_1 \rangle$, which is the overlap of the first-order perturbed wavefunction. This formula forms the basis for the widely used Davidson's approximation [14] for $E_{CI}(Q)$ (where $E_{CI}(Q) \approx E_2\Delta$) as we proved elsewhere [15,9]. In our calculation, $E_2\Delta = -0.02647$ hartrees. Using E_{4Q} from Table 1, we obtain $E_{CI}(Q) \approx -20.8$ mh. An appropriate estimate for $E_{CI}(Q)$ from the CI calculation is given by the difference between CISD+SD(NO) and CIS+SD(NO) which is -19.5 mh, which is in good agreement. Although this number arises largely due to the addition of quadruple excitations into the CI, some additional triple excitation terms are also included. Again using an estimate ~1 mh for the additional triple excitations in CISD+SD gives ~2-3mh as the energy lost in the quadruple selection, which is consistent with our above observation. Clearly, any such estimates mixing infinite-order and finite-order calculations as well

as different (but similar SCF and natural) orbitals are not rigorous, but appear to be suggestive of the size of the selection error.

The over one million configuration CISD+SD calculation of SFSH was performed primarily as a demonstration of the feasibility of such large computations and used code largely designed for just such a purpose, hence timing comparisons are not entirely appropriate. Recognizing this, we can make a few very approximate comparisons. The CISD+SD calculation was carried out on a Harris 800 minicomputer and required just over 100 hours of CPU time (13 hours per iteration). The SDQ-MBPT(4) and CCSD timings by comparison took 0.2 and 2.2 hours, respectively, on an IBM 3033N computer. The IBM runs electronic structure calculations at between 2 to 5 times the speed of the Harris model thus suggesting about 20 hours for the CI on the 3033N, or roughly 100 times as long as the SDQ-MBPT(4) and about 10 times as long as the CCSD calculations. Computing the E_{4T} energy is inherently harder than in the SDQ-MBPT(4) or CCSD cases since finding E_{4T} is proportional to n^7 while the latter are proportional to n^6 operations where n is the number of orbitals. Even so, E_{4T} required almost 1.0 hour to compute. Our best calculations, SDTQ-MBPT(4) and CCSD+T(4) therefore, require 1.2 and 3.2 hours, respectively, or 6.0% and 16.0% of the CISD+SD time. The time savings occurs essentially due to the fact that only a number of coefficients equal to all single and double excitations are required in CCSD[10], although non-linear terms are considered, compared to the over 1 million coefficients in the CI. CISD has a similarly small number of coefficients and requires less computer time than the MBPT/CC calculations.

Comparison of Localized Orbital and SCF Orbital Results

Our second series of calculations on ethylene used simple bond-antibond localized orbitals. These LOs were constructed as follows. (A more detailed discussion is in ref. [2].) First, the 2s, 2px and 2pz atomic functions on each carbon atom were hybridized to form a set of sp^2 functions directed at the opposing carbon and the two nearest hydrogen atoms. Next, linear combinations of hybrid orbitals were formed to yield bonding and antibonding orbitals. Precisely, the non-orthogonal C-H, sigma (C-C) and pi(C-C) bonds and antibonds are $[C(sp^2) \pm H(1s)]$,

$[C_1(sp^2) \pm C_2(sp^2)]$ and $[C_1(p_y) \pm C_2(p_y)]$, respectively, which forces equal weighing even for the CH bond. Non-polar bonds are not necessary but chosen purely for convenience in the present calculation. (See Ref. [16] for the effect of non-equal weights in the PCILO scheme and Ref. [17] for an ab-initio study.) All polarization functions in the DZP basis are introduced into the calculation as excited orbitals in addition to the anti-bonding orbitals.

In the PCILO method [5], the zero-differential overlap approximation is normally used, however an ab-initio, non-minimum basis calculation requires an orthogonalization procedure. Instead of carrying out a Lowdin symmetric orthonormalization [18] on the complete localized set, the basis is first divided into three subsets. These are the core plus bonding orbitals, the antibonding orbitals, and the remaining excited orbitals. The third category is necessary to accommodate larger than minimum AO basis sets. (The term "excited" orbitals refers to antibonding orbitals plus virtual orbitals.) Next each subset is individually symmetrically orthonormalized and the antibonding block is Schmidt orthonormalized to the bonding orbitals. Lastly, the virtual orbitals are Schmidt orthonormalized to both the bonding and anti-

bonding subsets. There are several advantages for employing this orthonormalizing procedure. Perhaps the most important is that the bonding orbitals contain no contamination from the excited orbitals. This feature should facilitate potential transferability among different calculations. In addition, since the CCSD energy is invariant to occupied-occupied and excited-excited orbital rotations, use of either localized or delocalized virtual orbitals would result in energetically equivalent results.

The reference energy for the single determinant formed by doubly occupying the two 1s core and bonding orbitals is -77.75347 hartrees. Within this L0 basis a number of correlated calculations are performed, including SDQ-MBPT(4) and CCSD. As in the calculations employing SCF orbitals, the two core orbitals are kept frozen.

Comparing first the SCF and L0 reference determinant energies a difference of -0.29706 hartrees is found. At first this difference seems large since the L0s should lead to a reasonable guess at the reference determinant. However, since our L0s are formed from linear combinations of the valence atomic functions, the polarization functions in the DZP set contribute nothing to the reference determinant energy, nor is there any opportunity for the comparative weights in the bond orbitals to change to more accurately reflect the correct charge distribution. Both of these features are introduced into the calculation via the single excitation term in the CCSD wavefunction[10] (i.e.

$T_1 = \sum_i t_i^a X_a^\dagger X_i$). The CCSD wavefunction,

$$\Psi_{\text{CCSD}} = \exp(T_1 + T_2) | \phi_0 \rangle = \exp(T_2) \exp(T_1) | \phi_0 \rangle \quad (2)$$

benefits from the fact that any single determinant wavefunction, ϕ_0 can be interchanged with any other ϕ_0 via Thouless' relation [19],

$$| \phi_0 \rangle = \exp(T_1) | \phi_0 \rangle \quad (3)$$

Hence the difference in energy (ΔE_{pol} between $\phi_{Local} = \phi_0$ and a localized orbital representation of the SCF solution, ϕ_{SCF} , is largely taken into account by

$$\Delta E_{pol} = \langle \phi_0 | H \exp(T_1) | \phi_0 \rangle = \sum_{i,a} f_{ia} t_i^a + \sum_{\substack{i < j \\ a < b}} \langle ij || ab \rangle (t_i^a t_j^b - t_i^b t_j^a) \quad (4)$$

for the converged amplitudes, $\{t_i^a\}$, where f_{ia} is the off-diagonal Fock operator[10]. In the lowest-order iteration, $t_i^a = f_{ia}$ hence we can observe the low-order contributions to ΔE_{pol} by conventional perturbation theory arguments. The second-order contributions to ΔE_{pol} is -0.23008 hartrees giving a comparison value of -77.98355 hartrees, which is 0.06698 hartrees above E_{SCF} . At convergence, this is improved to -78.04475, or 0.00578 hartrees about E_{SCF} .

Our best estimate of the correlation energy, the SCF orbital CCSD+ E_{4T} value of -0.30931 hartrees, is only marginally larger than the SCF-LO reference determinant energy separation. Therefore, using this particular LO set, the CCSD procedure recovers twice as much energy as in the analogous SCF case. However, the rates of convergence of the LO and SCF orbital MBPT calculations should be different. Comparing E_2 first, we find that the LO value of -0.48326 hartrees is substantially larger than the SCF orbital value of -0.27222 hartrees, but represents 81.0% of the CCSD energy correction compared to 90.4% in the SCF orbital basis. Similarly, third-order LO improves this to 93.6% compared to 98.5% for the SCF case. Adding in the fourth-order terms due to single, double, and quadruple excitations obtains 97.7% in the LO case and 99.6% for SCF orbitals, still causing a discrepancy of 14 mh to 1 mh, respectively. However, as seen in Table 2, at convergence the CCSD energies differ by only 1.7 mh regardless of orbital basis set.

In C_2H_4 all bond orbitals are chosen to be equally weighted despite the electronegativity differences between C and H. We have made a study of how the CCSD results with LO's change as a function of the bond orbital weighting factors, which will be presented elsewhere [17]. The effect of triple excitations with localized orbitals requires the computation of non-Hartree-Fock triple excitation diagrams. We have presented the theory and procedure for this calculation[11] and are currently developing the computer code.

Summary

In conclusion, even with very crude localized orbital descriptions of C_2H_4 , CCSD is able to recover about the same energy correction as when using SCF orbitals. This is primarily due to the effect of T_1 in the coupled cluster ansatz, which is able to introduce the requisite polarization effects into the calculation. Clearly, with Foster-Boys[3] or Edmiston-Ruedenberg[4] LO's, much better low order MBPT approximations would be possible.

From comparisons between MBPT/CC and the $>10^6$ configuration CI study of C_2H_4 by SFHS, we show that CCSD and full fourth-order perturbation theory, SDTQ-MBPT(4), give quite good results. From the SDTQ-MBPT(4) and CCSD value, it is possible to suggest that configuration selection among the triple and quadruple excitations as accomplished by SFHS, results in only a small 2-5 mH error for this molecule.

References

1. See H. Weinstein, R. Pauncz, and M. Cohen, *Advances At. and Mol. Phys.* 7, 97 (1971), for a review.
2. W.D. Laidig, G.D. Purvis III, and R.J. Bartlett, *Int. J. Quantum Chem. Symp.* 16, 561 (1982).
3. J.M. Foster and S.F. Boys, *Rev. Mod. Phys.* 32, 300 (1960).
4. C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* 35, 457 (1963); *J. Chem. Phys.* 54, 97 (1965).
5. S. Diner, J.P. Malrieu, P. Claverie and F. Jordan, *Chem. Phys. Letters* 2, 319 (1968); S. Diner, J.P. Malrieu and P. Claverie, *Theor. Chim. Acta* 13, 1 (1969).
6. P. Saxe, D.J. Fox, H.F. Schaefer III and N.C. Handy, *J. Chem. Phys.*, to be published.
7. S. Huzinaga, *J. Chem. Phys.* 42, 1293 (1965); T.H. Dunning, *J. Chem. Phys.* 53, 2823 (1970).
8. R.J. Bartlett and G.D. Purvis III, *Int. J. Quantum Chem.* 14, 561 (1978).
9. R.J. Bartlett, *Annual Reviews Phys. Chem.* 32, 359 (1981).
10. G.D. Purvis III, and R.J. Bartlett, *J. Chem. Phys.* 76, 1910 (1982).
11. R.J. Bartlett and G.D. Purvis III, to be published.
12. M.J. Frisch, R. Krishnan and J.A. Pople, *Chem. Phys. Lett.* 75, 66 (1980).
13. R.J. Bartlett, I. Shavitt, and G.D. Purvis III, *J. Chem. Phys.* 71, 281 (1979).
14. E.R. Davidson, in *The World of Quantum Chemistry*, ed. R. Daudel and B. Pullman, Dordrecht, Holland (1974).
15. R.J. Bartlett and I. Shavitt, *Int. J. Quantum Chem. Symp.* 11, 165 (1978); 12, 543 (1978).
16. J.M. Cullen and M.C. Zerner, *Int. J. Quantum Chem.* 22, 497 (1982).
17. W.D. Laidig, G.D. Purvis III and R.J. Bartlett, to be published.
18. P.O. Löwdin, *Adv. Quantum Chem.* 5, 185 (1970); *J. Chem. Phys.* 18, 365 (1950).
19. D.J. Thouless, *Nucl. Phys.* 21, 225 (1960); 22, 78 (1961).

Table 1

C_2H_4 : Summary of SCF Orbital MBPT/CC Energy Calculations Including the NO
CI Calculations of Saxe, Fox, Schaefer and Handy [6]. (DZP Basis Set)

<u>Type of Calculation</u>	<u>Correlation Energy</u>	<u>Cumulative Total Energy</u>
E	-0.0	-78.05053
SCF		
E (NO)	-----	-78.04942
Ref		
CISD(SCF)	-0.27758	-78.32810
CIS+SD(NO)	-0.28448	-78.33501
CISD+SD(NO)	-0.30398	-78.35451
E_2	-0.27222	-78.32275
E_3	-0.02433	-78.34708
E_{4S}	-0.00163	-----
E_{4D}	-0.00751	-----
E_{4T}	-0.00815	-----
E_{4Q}	0.00563	-----
SDQ-MBPT(4)	-0.30004	-78.35057
SDTQ-MBPT(4)	-0.30840	-78.35873
CCD	-0.29919	-78.34972
CCSD	-0.30116	-78.35169
CCSD+ E_{4T}	-0.30931	-78.35984

Table 2

C_2H_4 : Localized Orbital MBPT/CCM Results (DZP Basis Set)

<u>Type of Calculation</u>	<u>Energy Correction</u>	<u>Cumulative Total Energy</u>
E SCF	-0.0	-78.05053
E Ref	-----	-77.75347
E_2	-0.48326	-78.23673
E_3	-0.07531	-78.31204
E_4 (SDQ)	-0.02409	-78.33612
SDQ-MBPT(4)	-0.58266	-78.33612
CCSD	-0.59652	-78.34999