ELSEVIER

11 February 1994

Chemical Physics Letters 218 (1994) 276-282

CHEMICAL
PHYSICS
LETTERS

Size-consistent self-consistent combination of selected CI and perturbation theory

J. Meller [1], J.L. Heully, J.P. Malrieu

Laboratoire de Physique Quantique, IRSAMC, Université Paul Sabatier, 118, route de Narbonne, 31062 Toulouse Cedex, France

Received 18 October 1993; in final form 13 December 1993

Abstract

The present work proposes a consistent combination of two recently reported computational schemes, namely a size-consistent selected CI and a non-divergent second-order perturbation evaluation. The method treats large doubles (and any set of higher excited determinants) in a variational manner and the small doubles in a perturbative mode, without any unlinked contributions and exactly treating the EPV terms. Strict separability into closed-shell subsystems is ensured if the localized MOs are used. The efficiency of our proposal is illustrated by a few calculations (NH_3 in the DZP basis set, H_2O in the DZ basis set and the water dimer $2(H_2O)$).

1. Introduction

The size-consistency problem is a major difficulty in the approximate treatments of electronic correlation. Coupled-cluster [1] algorithms solve this problem, at least for the closed-shell, single-reference case, but their computational cost is high, they face convergence difficulties when bonds are broken, and they are not flexible. All the double excitations are considered, which may become impossible for large problems, and then all triples (in approximations like CCSD(T)) [2] are also taken into account at the next stage, whereas only a few of them play a significant role. In principle, as shown by the linked-cluster theorem, the Møller–Plesset Rayleigh–Schrödinger perturbation expansion is size-consistent when starting from a single reference. But it diverges when chemical bonds are broken if one uses a closed-shell reference. Going to a spin unrestricted reference one faces in turn both spin contamination and poor convergence difficulties [3].

The quasi-degenerate perturbation theory [4] may in principle solve the divergence problem and ensure size extensivity when the model space is a complete active space (CAS) [5], but in practice these perturbative expansions face an intruder state problem and are almost useless in the context of quantum chemistry. Variational CI approaches do not diverge even for excited states and there is no problem with a multireference zeroth-order description. But full-CI is possible only in special cases and one must take care in selecting or truncating the CI expansion. Despite the progress in the dimension of CI expansions treated by direct selected CI algorithms, these approaches are never free from the size-consistency problem.

Recent efforts have brought simple solutions to two of the above difficulties. Both are based on the development of the intermediate Hamiltonian theory proposed by the quantum physics group of Toulouse [6]. For the time being they are restricted to the closed-shell ground state problem. Initial progress [7] led to

---

[1] On leave from Department of Computer Methods, Nicholas Copernicus University, ul. Grudziądzka 5, 87-100 Toruń, Poland.

SSDI 0009-2614(93)E1475-V

self-consistent, non-divergent evaluation of the second-order type correlation energies. The cost of this self-consistent second-order (SCPT2) procedure is a few times that of a classical MP2 calculation and the behaviour reported so far of the potential energy for single bond breaking is encouraging. However this simple method could not be applied to the breaking of multiple bonds.

Further progress [8] achieved the size extensivity of any selected CI by a proper shift of the diagonal energies of the CI matrix. This method has been called self-consistent size-consistent CI ($(SC)^2CI$). For SDCI this may be seen as an improved CEPA [8] version with an accurate treatment of all EPV (exclusion principle violating) contributions, which ensures the strict separability when localized MOs are used. However the method is completely general and may be applied to any selected CI including arbitrary fractions of the various classes of excitations.

In the present Letter, we would like to join these two self-consistent size-consistent processes. The proposed algorithm is only relevant when all doubles are not included in the variational CI step, otherwise it reduces to the $(SC)^2CI$ algorithm. At the other border it reduces to self-consistent second order (SCPT2) when the number of selected doubles falls to 1, so the method proceeds continuously from SCPT2 to $(SC)^2CI$ by increasing the number of the doubles included in the variational CI. But it is always possible to add the important triples, quadrupoles to this selected space. We propose to label the method $(SC)^2(CI+PT2)$. The general features of the method as well as comments about the practical implementation will be given in section 2. The formal properties and especially the strict separability will also be discussed. Section 3 will be devoted to a comparison with other popular combinations of variation and perturbation, and it will be shown that we may keep the full flexibility of this combination, whilst keeping also the size extensivity. The physical content and the relative cost of the present method and other MRPT2 methods will be discussed. Section 4 gives a few illustrative examples of the efficiency.

2. The method

The correlation energy may be written as

$$E=\sum_i c_i\langle\phi_0|H|\phi_i\rangle\ ,$$

where c_i is the coefficient of the determinant ϕ_i in the expansion of the eigenvector in the intermediate normalization

$$\Psi=\phi_0+\sum_i c_i\phi_i\ ,$$

ϕ_0 being the reference (for instance HF) determinant. Due to the bielectronic nature of H, it is sufficient to know the amplitudes of the singles and doubles. Then one may think of obtaining c_i by a perturbative technique if $|c_i|$ is small, and from a (large) variational CI if $|c_i|$ is larger than a certain threshold. The variational CI may include more than the so selected large doubles.

Of course we must avoid two drawbacks; namely the crudeness of the first-order evaluation of c_i for the small doubles due to its perturbative character and the impact of size inconsistency of the variational CI on the large c_i. To avoid the first difficulty one may think of determining c_i through a 2×2 matrix diagonalization, i.e. of the matrix spanned by ϕ_0 and ϕ_i,

$$\begin{array}{ccc} & \phi_0 & \phi_i \\ \phi_0 & \multicolumn{2}{c}{\multirow{2}{*}{$\begin{pmatrix} H_{00} & H_{0i} \\ H_{i0} & H_{ii} \end{pmatrix}$}} \\ \phi_i & & \end{array}\ ,$$

as done in the independent excitation pair approximation [9]. As shown elsewhere [7], a better evaluation is obtained by considering a slightly modified matrix

$$\begin{array}{ccc} & \phi_0 & \phi_i \\ \phi_0 & H_{00}+\delta_{00} & H_{0i} \\ \phi_i & H_{i0} & H_{ii}+\delta_{ii} \end{array}\ ,$$

where the diagonal energy shifts represent the effect of all the double excitations D_j^+ that are possible when acting on either ϕ_0, except D_i^+ if $\phi_i=D_i^+\phi_0$, since it is already considered,

$$\delta_{00}=\sum_{j\neq i} c_j\langle\phi_0|H|\phi_j\rangle\ ,$$

or on ϕ_i. Many double excitations D_j^+ become impossible on ϕ_i since they employ at least one hole or one particle also involved in ϕ_i, so that

$$\delta_{ii}=\sum_{\substack{j \\ D_j^+\phi_i\neq 0}} c_j\langle\phi_0|H|\phi_j\rangle\ .$$

These dressings (the shifts on the diagonal of the matrix) incorporate the effects of the doubles $\phi_j = D_j^+ \phi_0$ on ϕ_0 and of quadruples $D_j^+ \phi_i$ on ϕ_i which must cancel when both exist. From a practical point of view for the determination of c_i, it is sufficient to consider the shifted matrix

$$\begin{array}{cc} & \begin{array}{cc} \phi_0 & \phi_i \end{array} \\ \begin{array}{c} \phi_0 \\ \phi_i \end{array} & \begin{pmatrix} 0 & H_{0i} \\ H_{i0} & H_{ii} - H_{00} + \delta_{ii} - \delta_{00} \end{pmatrix} \end{array} .$$

Now

$$\delta_{ii} - \delta_{00} = - \sum_{\substack{j \\ D_j^+ \phi_i = 0}} c_j \langle \phi_0 | H | \phi_j \rangle + c_i \langle \phi_0 | H | \phi_i \rangle .$$

The dressing eliminates all unlinked terms and incorporates an infinite summation of EPV terms that actually go through the $D_j^+ \phi_i = 0$ processes.

Regarding the CI matrix we do the same, with the only difference that we should restrict our summation to avoid a double counting of the effect of the double excitations leading to configurations inside the variational space S. For instance we should not consider the contribution $c_j \langle \phi_0 | H | \phi_j \rangle$ on δ_{ii} if $D_j^+ \phi_i$ is a quadruple appearing among the selected determinants. Hence, for all determinants the dressing should be written

$$\delta_{ii} = \sum_{\substack{j \\ D_j^+ \phi_i \neq 0 \\ D_j^+ \phi_i \notin S}} c_j \langle \phi_0 | H | \phi_j \rangle .$$

Notice that it differs from the expression given in ref. [8] due to fact that at that time the summation was only extended over the doubles belonging to the S space. Now ϕ_0 is shifted by the effect of all small doubles, while this was not the case in ref. [8].

Here we have a unique and complete set of coefficients of all doubles, coming from either small 2×2 or large $n \times n$ diagonalizations of matrices which are themself shifted, in a self-consistent way, using the complete set of the c_is.

It is possible to demonstrate that the proposed algorithm provides *strict separability* of the energy of a supersystem AB composed of two non-interacting subsystems A and B

$$E_{AB} = E_A + E_B ,$$

provided that the MOs are localized on either A or B. The demonstration is quite easy and follows the same logic as that given in ref. [8] except for the fact that the small coefficients result from 2×2 diagonalizations. The generalisation of the proof of separability for an arbitrary selected space S also follows that of ref. [8].

Regarding the practical implementation, one does not perform explicit summations over all EPV contributions (the number of which is approximately proportional to N^3, N being the size of the basis set). One benefits from a trick proposed in a previous study on infinite summation of EPV diagrams [10]. One-, two- and three-index arrays store the contributions to the correlation energy of each orbital

$$e(r) = \sum_{\substack{i \\ D_i^+ \text{ involving the MO } r}} c_i \langle \phi_0 | H | \phi_i \rangle ,$$

of each pair of orbitals

$$e(r, s) = \sum_{\substack{i \\ D_i^+ \text{ involving the MOs } r \text{ and } s}} c_i \langle \phi_0 | H | \phi_i \rangle ,$$

and of each triplet of orbitals

$$e(r, s, p) = \sum_{\substack{i \\ D_i^+ \text{ involving the MOs } r, s \text{ and } p}} c_i \langle \phi_0 | H | \phi_i \rangle .$$

Using these quantities, the calculation of the diagonal dressing δ_{ii} becomes straightforward since it only requires summations over the holes and particles of ϕ_i. In practice all determinants are dressed by all the possible excitations.

3. Comparison with other methods

The method may first be compared with the usual single reference Møller–Plesset perturbative expansion. On the perturbative part (the small doubles), it includes the full-MP2, plus an infinite order summation over diagrams which normally leads to the definition of the Epstein–Nesbet zeroth-order Hamiltonian [11], plus another infinite order summation over EPVs for the small doubles. From the variational part (the large doubles plus some higher excited determinants if needed) it includes also the full

MP2 but also the MP3, ...MPn contributions which come from the interaction between large doubles, and if triples and quadruples are present the corresponding MP4 and higher-order corrections are included, without any undesirable unlinked terms.

More interesting is the comparison with some popular multireference MP2 methods. It should be pointed out that we are not comparing this method with multireference methods which use effective Hamiltonians [12] but rather with state selective multireference methods [13]. Standard MRMP2 methods proceed first by a diagonalization of the $P_S H P_S$ matrix and later perturb the multiconfigurational vector as a whole. Our method differs in two important aspects: the treatment of size-consistency which is of crucial importance in treating dissociation processes, and the contributions from the interactions of the model space excited determinants with the outer space. Size-consistency is obtained because before the diagonalization of $P_S H P_S$ each diagonal term of this matrix is shifted, removing any unlinked terms and including important EPV terms. Such a correction is not necessary when the model space is complete (CAS) as in the CASPT2 method of Andersson [13]. The second difference comes from the fact that obviously some effects treated in MRMP2 methods are neglected in our scheme, mainly the MP3 contribution incorporating the interactions between small and large doubles. The MRMP2 methods calculate all matrix elements between the model space determinants ϕ_i and the outer space determinants ϕ_α: $\langle\phi_i|H|\phi_\alpha\rangle$ whereas we only use explicitly the matrix elements of the first line $\langle\phi_0|H|\phi_\alpha\rangle$ because we only perturb the reference determinant and not the others. This is the reason why the computational cost of our procedure is simply that of the diagonalization plus that of an MP2, while in MRMP2 techniques the bottleneck becomes the perturbative step for which the cost increases rapidly with the number of reference determinants.

4. Test calculations

The efficiency of the algorithm proposed here has been tested through three sets of computations.

The first one concerns the NH_3 molecule in the DZP basis set for which an estimate of the full-CI energy was proposed a few years ago [14]. The results appear in Table 1 and Fig. 1 for different variational spaces selected according to the CIPSI scheme, and which contain from 5% to 85% of the doubles, plus the most important triples and quadruples. The error never exceeds 5 mE_h. The present procedure gives slightly better values than a previous approximation that consisted in dressing the variational CI by the large doubles only and adding the MP2 effect of the small doubles. The relative success of the previous calculation comes from the cancellation of two errors, namely terms between large and small double excitations. Comparing the results from ref. [8] and Table 1, one can notice that including 2633 doubles and 174 of the most important triples and quadruples in the S space, and treating the other 3182 doubles in a perturbative mode, one gets a better value for the energy (−0.205 au) than treating all doubles

Table 1

Results for ammonia in the DZP basis set. Geometry and basis set from Knowles and Handy [14], 1s orbital frozen. Energy differences in mE_h, with respect to the SCF energy. The full-CI energy is estimated as −209.9 mE_h [14]. n^{LD} means the number of large doubles, whereas n^{SD} is the number of small ones. E^{var} refers to undressed CI, $\tilde{E}$ to (SC)^{2}CI energy (dressed only by large doubles) and $\hat{E}$ refers to (SC)^2(CI+PT) energy. ϵ^{MP2} and ϵ^{SCPT} denote the usual MP2 and SCPT contributions of small doubles, respectively

Var. space		Var. energies		Small doubles			Energies	
dim	n^{LD}	E^{var}	$\tilde{E}$	n^{SD}	ϵ^{MP2}	ϵ^{SCPT}	$\tilde{E}+\epsilon^{MP2}$	$\hat{E}$
397	392	−135.6	−138.8	5426	−62.0	−73.5	−200.8	−210.9
1440	1414	−180.0	−185.0	4401	−17.3	−20.3	−202.3	−204.8
2807	2633	−194.4	−199.7	3182	−4.4	−5.1	−204.1	−204.7
3381	3061	−196.9	−202.2	2765	−2.3	−2.7	−204.5	−204.8
4957	3947	−200.1	−205.1	1861	−0.4	−0.5	−205.5	−205.6
13588	4892	−203.2	−207.3	918	0.0	0.0	−207.3	−207.3

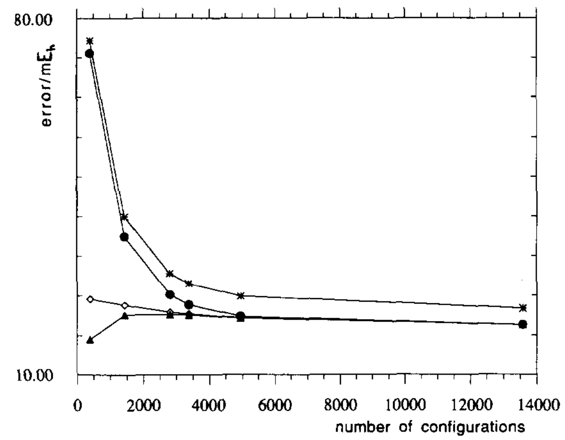

Fig. 1. NH_3 correlation energies for different dimensions of the variational space. (△) $(SC)^2(CI+PT)$; (◇) $(SC)^2CI+\epsilon^{MP2}$; (○) $(SC)^2CI$; (✱) undressed CI.

in the SDCI (−0.197 au) or in the $(SC)^2SDCI$ (−0.203 au). This example shows the value of including the most important higher excitations in the variational process while treating small doubles at a low level. As can be seen from Table 1 small doubles still give 5.1 mE_h. In cases with a much larger number of double excitations, where only a fraction of them could be treated variationally and where any MRMP2 would be difficult (if possible), the advantage of our combination of $(SC)^2CI$ and SCPT2 may be even more encouraging.

In order to test the behaviour of our proposal when bonds are broken, the H_2O molecule in the DZ basis set has been examined for three interatomic distances (r_e, $1.5r_e$, $2r_e$). Comparison is possible not only with the full-CI [15], but with other approximate algorithms combining variation and perturbation [16]. The results appear in Table 2. Regarding the compar-

Table 2
Results for water in the DZ basis set. Geometry and basis set from Saxe et al. [15]. Energy differences in mE_h, with respect to the SCF energy. The full-CI energies are −148.0, −211.0, −310.1 mE_h for r_e, $1.5r_e$ and $2r_e$, respectively [8]. n^{LD} means the number of large doubles, whereas n^{SD} is the number of small ones. E^{var} refers to undressed CI, $\hat{E}$ to $(SC)^2(CI+PT)$ energy. ϵ^{SCPT} denotes SCPT contributions of small doubles, whereas C^{MP} and C^{EN} denote Møller–Plesset and Epstein–Nesbet CIPSI energies. Single reference CISD and CISDTQ energies are also given for comparison

	Var. space			Small doubles		Energies		
	dim	n^{LD}	E^{var}	n^{SD}	ϵ^{SCPT}	C^{MP}	C^{EN}	$\hat{E}$
r_e	1	0	0	841	−177.0	−130.5	−183.3	−177.0
	8	7	−32.9	834	−139.7	−134.5	−161.5	−172.0
	32	31	−69.0	810	−90.9	−136.1	−150.9	−159.9
	82	81	−97.1	760	−50.5	−140.2	−147.3	−148.6
	767	364	−142.5	477	−1.6	−147.3	−148.0	−146.2
	6478	709	−147.4	132	−0.01	–	–	−147.9
	CISD = −140.2					CISDTQ = −147.8		
$1.5r_e$	1	0	0	841	−256.2	−191.1	−288.8	−256.2
	33	30	−112.5	811	−115.9	−199.5	−218.0	−226.5
	54	50	−136.6	791	−81.0	−201.5	−213.3	−218.0
	117	107	−160.8	734	−37.4	−204.1	−211.2	−202.6
	1329	366	−189.1	613	−6.6	−209.9	−210.6	−209.0
	6965	683	−209.2	158	−0.02	–	–	−210.1
	CISD = −188.6					CISDTQ = −209.9		
$2r_e$	1	0	0	841	−348.1	−257.3	−519.8	−348.1
	38	32	−212.1	809	−102.9	−303.3	−314.4	−307.8
	73	61	−234.0	780	−52.9	−302.3	−312.1	−294.4
	123	89	−253.0	752	−34.8	−303.3	−309.4	−297.5
	1483	360	−298.7	481	−1.2	−308.4	−309.4	−305.3
	6551	649	−304.6	192	−0.02	–	–	−307.2
	CISD = −249.6					CISDTQ = −305.7		

ison with the CIPSI method, one should notice that our proposal is free from the uncertainty that comes from the choice of H_0 (which has a large impact on the final energy) and it is much cheaper. The CIPSI MP2 values ensure an approximate size extensivity and give similar errors (in the absolute numbers) to the present results. However the perturbative CIPSI step is much more expensive. The method behaves satisfactorily when bonds are broken, giving with the lowest threshold of selection of the variational space S, errors of 0.1, 0.9 and 2.9 mE_h at r_e, $1.5r_e$ and $2r_e$ geometries, respectively (for comparison CISDTQ gives errors of 0.2, 1.1 and 4.4 mE_h, respectively, see ref. [14]). For stretched geometries some doubles and even quadruples have large coefficients and the dressing becomes crucial as may be seen from Table 2 – it brings in some cases more than 10 mE_h (for instance for $1.5r_e$, dim = 1329, the difference between $\hat{E}$ and $E^{\mathrm{var}}+\epsilon^{\mathrm{SCPT}}$ is about 13 mE_h, the effect of dressing is comparable to that difference).

We have verified the separability property by calculating two water molecules at very large distances, using localized MOs. When the selection only introduces doubles in the variational space, for the supersystem AB as well as for A and B, the variational space is the simple reunion of those coresponding to the subsystems and the additivity follows from the proof given in refs. [7,8]. A stronger and more surprising property concerns the separability in the case where, using the same thresholds for the selections, the variational space contains some intermolecular quadruples product of intramolecular doubles that are not contained in the subsystem selected spaces.

5. Conclusion

Although the test calculations have not been performed on real-scale problems since they concern medium basis sets and small numbers of electrons, they illustrate the efficiency and possibilities of the method. The reciprocal dressing of the variational and perturbative type treatments significantly improves the results when the contribution of the small doubles remains important and this should be the case for the relevant domains of application of the method.

The procedure is flexible and powerful first because it is size-consistent, but essentially because one may incorporate higher-order effects including the most important triples and quadruples. This should be important for many problems in quantum chemistry where a subset of highly correlated electrons appears, while the electronic correlation of the remainder may be treated at a lower level of accuracy. One may think of chemical reactions where only a few bonds are strongly affected, or of the π electrons above the σ electrons in conjugated molecules.

Another field of application may be the interaction between molecules (and any problem where size extensivity is required).

In comparison with MRMP2 methods, the present scheme may be seen as more approximate since it neglects the interaction between small and large doubles, but it is formally rigorous and much cheaper for a given size of the variational space since we do not generate all the determinants interacting with the selected configurations. The bottleneck is no longer the perturbative step and one may go to a large dimension of the CI space (especially if one uses direct selected algorithms).

Work to generalize the method to excited states is in progress.

6. Acknowledgement

The authors are indebted to J.P. Daudey and M.B. Lepetit for helpful advice. JM gratefully acknowledges the financial support of the Tempus Project.

7. Appendix

The scheme of the program for obtaining the $(\mathrm{SC})^2(\mathrm{CI}+\mathrm{PT})$ energy is given in Fig. A.1, where

(1) in practice two iterations of CIPSI were done, so that some doubles and most important triples and quadruples were included in S;

(6) for a given double $\phi_i = D_i^+\phi_0$ instead of explicit summation over all doubles impossible on ϕ_i in the expression for δ_{ii} one can sum over particles and holes in ϕ_i (using e-arrays) and then the corresponding diagonal element in 2×2 or CI matrix is modified;

(7) the undressing is done analogously to ref. [8];

(8) the contribution of small doubles is added to

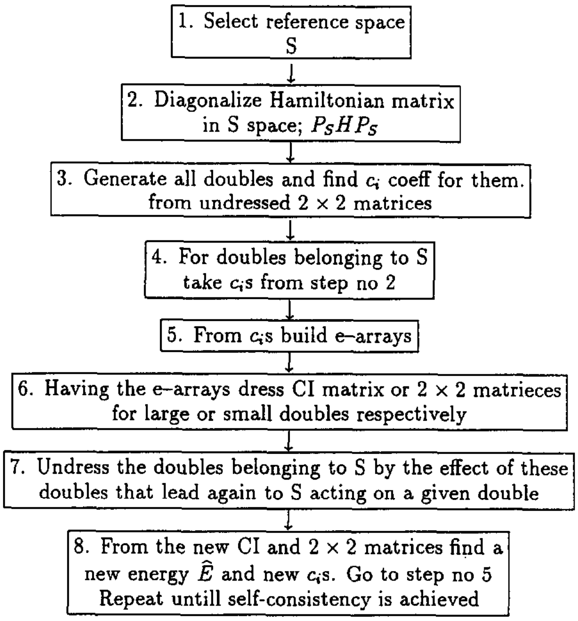

Fig. A.1.

the $\langle\phi_0|H|\phi_0\rangle$ element of the dressed-CI matrix (for every double $\phi_i \in \mathrm{S}$ the corresponding diagonal element $\langle\phi_i|H|\phi_i\rangle$ is modified by the effect of *all* other possible doubles), which is diagonalized to obtain the next value of $\hat{E}$ and the c_i coefficients of large doubles.

8. References

[1] J. Cizek and J. Paldus, Intern. J. Quantum Chem. Symp. 5 (1971) 359.

[2] K. Raghavachari, G.W. Trucks, J.A. Pople and M. Head-Gordon, Chem. Phys. Letters 157 (1989) 479;
T.J. Lee, A.P. Rendell and P.R. Taylor, J. Chem. Phys. 92 (1990) 5463.

[3] R.H. Nobes, J.A. Pople, L. Radom, N.C. Handy and P.J. Knowles, J. Chem. Phys. 85 (1986) 1469.

[4] Ph. Durand and J.P. Malrieu, in: Advances in Chemical Physics, Vol. 67, Ab initio methods in quantum chemistry, ed. K.P. Lawley (Wiley, New York, 1987) p. 321.

[5] I. Lindgren, J. Phys. A 7 (1974) 2441.

[6] J.P. Malrieu, Ph. Durand and J.P. Daudey, J. Phys. A 18 (1985) 809.

[7] M.B. Lepetit and J.P. Malrieu, Chem. Phys. Letters 208 (1993) 503.

[8] J.P. Daudey, J.L. Heully and J.P. Malrieu, J. Chem. Phys. 99 (1993) 1240;
J.L. Heully and J.P. Malrieu, Chem. Phys. Letters 199 (1992) 545.

[9] A.C. Hurley, Electron correlation in small molecules (Academic Press, New York, 1976).

[10] M.B. Lepetit and J.P. Malrieu, J. Chem. Phys. 87 (1987) 5937.

[11] P.S. Epstein, Phys. Rev. 28 (1926) 690.

[12] B.H. Brandow, Rev. Mod. Phys. 39 (1967) 771;
U. Kaldor, J. Chem. Phys. 62 (1975) 4634; 63 (1975) 2199.

[13] K. Andersson, P.-A. Malmquist and B.O. Roos, J. Chem. Phys. 96 (1992) 1218, and reference therein;
K. Hirao, Chem. Phys. Letters 190 (1991) 374;
R.B. Murphy and R.P. Messmer, Chem. Phys. Letters 183 (1991) 443;
K. Wolinski and P. Pulay, J. Chem. Phys. 90 (1989) 3647.

[14] P.J. Knowles and N.C. Handy, J. Chem. Phys. 91 (1989) 2396.

[15] P. Saxe, H.F. Schaefer III and N.C. Handy, Chem. Phys. Letters 79 (1981) 202;
R.J. Harrison and N.C. Handy, Chem. Phys. Letters 95 (1983) 386.

[16] S. Evangelisti, J.P. Daudey and J.P. Malrieu, Chem. Phys. 75 (1983) 91;
R.J. Harrison, J. Chem. Phys. 94 (1991) 5021.