

Multireference perturbation CI I. Extrapolation procedures with CAS or selected zero-order spaces

Celestino Angeli¹, Renzo Cimiraglia², Maurizio Persico³, Alessandro Toniolo³

¹ Scuola Normale Superiore di Pisa, p.za Cavalieri 7, I-56100 Pisa, Italy

² Dipartimento di Chimica, Università di Ferrara, Via Borsari 46, I-44100 Ferrara, Italy

³ Dipartimento di Chimica e Chim. Ind., Università di Pisa, v. Risorgimento 35, I-56126 Pisa, Italy
e-mail: mau@hermes.dcci.unipi.it

Received: 12 June 1997 / Accepted: 31 July 1997

Abstract. We discuss the “three class” approximation to full multireference perturbation CI, which greatly reduces the computational effort by restricting the summation of diagrams to determinants belonging to a subspace of the zero-order space. In the framework of the CIPSI algorithm, we propose a new extrapolation procedure allowing recovery of the full “two class” results. The new procedure is applied to complete active spaces (CAS) and to individually selected zero-order spaces. Comparison with a full two class calculation on a CAS shows a reduction of computer time of one or two orders of magnitude in the tests presented here, with an accuracy in the order of 0.1 kcal/mol. Our procedure can thus compete with the CASPT2 algorithm, specifically conceived to deal with CAS. In the case of selected zero-order spaces, the speed-up is less dramatic but the method still retains its advantages.

Key words: Multireference CI – Perturbation CI – CIPSI

1 Introduction

The evaluation of correlation energy still represents a challenge for computational quantum chemistry. Different approaches (both variational and perturbative) are applied routinely when a single determinant (or configuration) is a reasonable description for the state of the system. The ground state of several molecules satisfies this condition: in this case methods such as MP2, MP4 or SDCI have proved to be good approximations. These methods fail when applied to the study of states where several determinants are necessary for a first approximate description (static correlation) as in the case of bond breaking, of transition metal compounds and of excited states. Multireference perturbation CI methods are useful tools for the study of these systems.

One widely used multireference perturbation method is the CIPSI algorithm (configuration interaction by perturbation with multiconfigurational zeroth-order wave functions selected by iterative process) first proposed by Huron et al. [1] and further extended by the Toulouse group [2, 3] (see also [4] for a critical overview) and by ourselves [5, 6]. Second-order perturbation theory is applied to the eigenvalues of the CI Hamiltonian projected in a determinantal subspace \mathcal{S} . Two different partitions of the Hamiltonian are considered, Epstein–Nesbet (EN) and Møller–Plesset baricentric (MPB). Until recently, only the latter was implemented in the most efficient version of the program, based on diagrammatic theory [5]. The implementation of the EN partition in the diagrammatic code has been completed recently by one of us [7]. Similar strategies have been proposed by other authors, even in recent years (see [8–15] and references therein).

In the original CIPSI algorithm, the zero-order \mathcal{S} subspace is iteratively selected by inclusion of all the determinants whose contribution to the first-order wave function correction in the previous step is found to be larger than a certain threshold $\eta_{\mathcal{S}}$. Recently we have established a new way of selecting the \mathcal{S} space, based on the requirement that the first-order corrections to the wave functions should have the same norms for all states with all molecular geometries (see the second paper in this series [16]). The new selection produces a more balanced treatment of different states with different geometries. A popular alternative to individual selection of determinants in \mathcal{S} is to define the multireference space as a complete active space, CAS ($\mathcal{S} \equiv \text{CAS}$). Exploiting the CASSCF wave function properties leads to the very efficient CASPT2 algorithm [17].

The CIPSI method, especially in the diagrammatic formulation, can be made computationally very efficient by resorting to an approximate procedure, hereafter called a “three class” calculation. The basic idea was put forward by Evangelisti et al. [2] and later reformulated by us [6]. In this paper we present a study of this approximation. In the perturbation step of a three class calculation one takes into account only the single and double excitations which are generated from a subspace

\mathcal{G} of the variational space \mathcal{S} . Having to deal with \mathcal{S} spaces of increasing size, thanks to the improvement of the computational facilities and aiming at better accuracy, we have devised an extrapolation procedure which allows the results of a full perturbation treatment (“two class calculation”, $\mathcal{G} \equiv \mathcal{S}$) to be recovered.

The next section presents the method and the following ones a few test calculations using the N_2 and 2N_2 systems (for N_2 an FCI benchmark calculation is available [18]) and the CH_3NO molecule (at present under study in our laboratory). The test calculations illustrate the application of our extrapolation procedure in the cases of individually selected spaces and of CAS.

In the case of a CAS zero-order space the speed-up obtained with the extrapolation procedure is so large that our approach can compete with the CASPT2 one as has been implemented in the MOLCAS-3 package [19]. It is important to note that the latter algorithm was specifically conceived to deal with a CASSCF zero-order wave function, while the CIPSI method can treat one-electron functions and zero-order spaces however they are defined.

2 Method

In this section we recall first the CIPSI algorithm and the formulation of the three class calculation as it has been implemented by our group [6]. Briefly the original CIPSI method is based on the partitioning of the FCI space into two subspaces called \mathcal{S} and \mathcal{Q} (two class calculation). The \mathcal{S} space contains the most important determinants describing the electronic states under investigation. The Hamiltonian projected onto the \mathcal{S} space is diagonalized and the eigenstates are taken as the zero-order description of the true wave functions:

$$\hat{P} \hat{\mathcal{H}} \hat{P} |\Psi_m^{(0)}\rangle = E_m^{(MR)} |\Psi_m^{(0)}\rangle \quad (1)$$

$$\hat{P} = \sum_{N \in \mathcal{S}} |N\rangle \langle N| \quad |\Psi_m^{(0)}\rangle = \sum_{N \in \mathcal{S}} C_N^m |N\rangle \quad (2)$$

The contributions of the determinants $|I\rangle \in \mathcal{Q}$ are evaluated by a second-order Rayleigh-Schrödinger treatment: thus, only those $|I\rangle$ that are single or double excitations from some $|N\rangle \in \mathcal{S}$ contribute. Given a partition of the CI Hamiltonian:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{V}} \quad (3)$$

the second-order energy correction is:

$$E_m^{(2)} = - \sum_{I \in \mathcal{P}} \frac{\langle I | \hat{\mathcal{V}} | \Psi_m^{(0)} \rangle^2}{E_I - E_m^{(0)}} \quad (4)$$

or

$$E_m^{(2)} = - \sum_{M, N \in \mathcal{S}} C_M^m C_N^m \sum_{I \in \mathcal{P}} \frac{\langle M | \hat{\mathcal{V}} | I \rangle \langle I | \hat{\mathcal{V}} | N \rangle}{E_I - E_m^{(0)}} \quad (5)$$

Here \mathcal{P} is the subspace of \mathcal{Q} collecting all single and double excitations from \mathcal{S} . Equations (4) and (5) are totally equivalent at this level of the treatment; they

correspond to the traditional and to the diagrammatic implementation of the algorithm, respectively.

As pointed out before, two different partitions of the Hamiltonian (3) are implemented in our programs. Only the $\hat{\mathcal{H}}_0$ operator needs to be explicitly defined, while $\hat{\mathcal{V}}$ is simply $\hat{\mathcal{H}} - \hat{\mathcal{H}}_0$. We want the basis $\{|\Psi_m^{(0)}\rangle \in \mathcal{S}, |I\rangle \in \mathcal{P}\}$ to form a set of eigenstates of $\hat{\mathcal{H}}_0$:

$$\hat{\mathcal{H}}_0 |\Psi_m^{(0)}\rangle = E_m^{(0)} |\Psi_m^{(0)}\rangle \quad (6)$$

$$\hat{\mathcal{H}}_0 |I\rangle = E_I |I\rangle \quad (7)$$

In the MPB partition the eigenvalues of $\hat{\mathcal{H}}_0$ are the expectation values of the diagonal part of the Fock operator:

$$E_m^{(0)} = \langle \Psi_m^{(0)} | \hat{F}_{\text{diag}} | \Psi_m^{(0)} \rangle = \sum_{N \in \mathcal{S}} |C_N^m|^2 E_N \quad (8)$$

$$E_I = \langle I | \hat{F}_{\text{diag}} | I \rangle \quad (9)$$

E_I and E_N are sums of one-electron matrix elements F_{ii} over the occupied spin orbitals of the determinants $|I\rangle$ and $|N\rangle$. The F_{ii} coincide with the “orbital energies” when the orbitals are the canonical SCF ones.

In the EN partition, $\hat{\mathcal{H}}_0$ is the diagonal part of the matrix representation of $\hat{\mathcal{H}}$ in the basis $\{|\Psi_m^{(0)}\rangle \in \mathcal{S}, |I\rangle \in \mathcal{P}\}$

$$E_m^{(0)} = \langle \Psi_m^{(0)} | \hat{\mathcal{H}} | \Psi_m^{(0)} \rangle = E_m^{(MR)} \quad (10)$$

$$E_I = \langle I | \hat{\mathcal{H}} | I \rangle \quad (11)$$

In this case $E_m^{(0)}$ coincides with the multireference energy $E_m^{(MR)}$.

Each of the two partitions offers some advantages and but also involves problems (for discussions see refs. [4, 7, 20]). Briefly, with respect to EN, MPB has better formal properties of separability [4] and of independence on the one-electron basis (localization versus delocalization of the molecular orbitals [20]). The MPB partition may yield small denominators, which cause the summations (4–5) to blow up (“intruder states”), especially with large, non-selected, spaces; the same may occur with CASPT2 [21]. The EN partition is not subject to the intruder state problem; accumulated experience and theoretical analysis [22] indicate that the EN results converge more rapidly than the MPB ones to the full CI limit, when expanding the \mathcal{S} space. Generally it is very useful to compare the results of both partitions.

In this paper we will not describe the selection procedure for the \mathcal{S} space, as it is not really important in this context and will be the subject of a separate paper [16]. Suffice to say that a few CIPSI steps are usually performed and the \mathcal{S} space is iteratively enlarged including the determinants that have produced a large contribution in the perturbation step of the previous calculation.

In a two class CIPSI calculation there are two different approximations that introduce an error with respect to the FCI limit: the truncation of the CI space to $\mathcal{S} + \mathcal{P}$; the approximate evaluation of the contribution of the \mathcal{P} space, by second-order perturbation theory.

It is obvious that enlarging the \mathcal{S} space reduces the magnitude of both errors. Indeed, if more determinants are introduced in the \mathcal{S} space, their contributions are taken into account variationally, i.e. at the infinite order of perturbation theory. For the remaining determinants of \mathcal{P} , an improved zero-order wave function usually guarantees a more accurate second-order treatment. At the same time, the \mathcal{P} space is enlarged and $\mathcal{S} + \mathcal{P}$ approaches the FCI space. On the other hand, the computer time needed for the perturbation step increases approximately as $N_{\mathcal{S}}^2$, where $N_{\mathcal{S}}$ is the dimension of \mathcal{S} .

The three class calculation has been devised in order to achieve a compromise between a large \mathcal{S} space and a reasonably time-consuming calculation. The dimension of the \mathcal{P} space is reduced by choosing only the determinants generated as single and double excitations from a subspace $\mathcal{G} \subset \mathcal{S}$. \mathcal{G} contains the ‘‘most important’’ determinants of \mathcal{S} . We shall see later how to define the \mathcal{G} subspace properly. The three class approximation is not quite the same for the original algorithm and for the diagrammatic one. For the former, no further change is made, beyond truncating the summation in Eq. (4) because of the reduction of the \mathcal{P} space. In the diagrammatic implementation [5, 7], the summation on M and N is limited to the determinants in \mathcal{G} . Practically, this means neglecting all the interactions of the determinants $|I\rangle \in \mathcal{P}$ with the component of $|\Psi_m^{(0)}\rangle$ outside \mathcal{G} . As a consequence, the numerical results of the two versions of a three class calculation with the same \mathcal{G} space are slightly different.

In this paper we shall concentrate on the diagrammatic three class algorithm, because it is particularly efficient, the computing time being dependent only on the dimension of \mathcal{G} , $N_{\mathcal{G}}$, and almost independent of $N_{\mathcal{S}}$. Since Eq. (5) is a quadratic form in the C_N^m coefficients, it seems quite natural to suppose that the second-order correction to the energy is approximately linear with respect to the square norm of the projection of the zero-order wave function $|\Psi_m^{(0)}\rangle$ over the \mathcal{G} space. The ‘‘ \mathcal{G} weight’’ $W_{\mathcal{G}}^{(m)}$ of $|\Psi_m^{(0)}\rangle$ is

$$W_{\mathcal{G}}^{(m)} = \sum_{N \in \mathcal{G}} |C_N^m|^2 \quad (12)$$

Given a set of thresholds $\{\tau_1 < \tau_2 < \dots < \tau_n\}$ it is possible to define a set of \mathcal{G} spaces $\{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n\}$ as the minimal sets for which the relations

$$W_{\mathcal{G}_j}^{(m)} = \sum_{N \in \mathcal{G}_j} |C_N^m|^2 \geq \tau_j \quad \forall m \quad (13)$$

are satisfied. Note that $\dim(\mathcal{G}_{j-1}) \leq \dim(\mathcal{G}_j) \leq \dim(\mathcal{G}_{j+1})$. This definition of the \mathcal{G} space allows the magnitude of the approximation introduced with the three class scheme to be controlled, in a balanced way for several electronic states. To test our hypothesis we have modified the diagrammatic code in such a way that the results of the calculations with the n \mathcal{G}_j spaces of increasing size can be obtained with a single run. The extra cost with respect to a single calculation with the largest \mathcal{G}_j is negligible.

The following sections present test calculations using different molecules and with CAS or with individually

selected zero-order spaces. The results show that $E_m^{(2)}$ is in a good measure linear in $W_{\mathcal{G}}$ if this parameter is sufficiently close to 1. A weighted least-square linear fit of $E_m^{(2)}$ versus $W_{\mathcal{G}}$, followed by an extrapolation to $W_{\mathcal{G}} = 1$, allows recovery of the results of a two class calculation with excellent accuracy. The weights used in the fitting procedure have been chosen as $1/(1 - W_{\mathcal{G}_j})$. The efficiency ratio of this procedure with respect to a complete two class calculation strongly depends on the required accuracy and on the nature of the \mathcal{S} space (selected or CAS). In order to make meaningful comparisons of the computing times, all the calculations have been performed on an IBM RISC/6000 590 workstation.

3 Test calculations: N_2 and $2N_2$

The N_2 molecule has been used in a first series of test calculations because for this system an FCI study for a subset of the valence electrons has been published [18]. In this study the one-electron basis is provided by a CASSCF calculation involving six orbitals and six electrons. In the CI the $1s$ and $2s$ orbitals are frozen and only the six $2p$ electrons are fully correlated. For larger systems (four heavy atoms or more) no FCI calculations have been made, given the prohibitive growth of the FCI expansion with the dimension of the basis. Our interest is to test the extrapolation procedure here proposed on systems of this size, so we have considered two non-interacting N_2 molecules for which the FCI energy is obviously twice the FCI energy of one molecule.

The test calculations on N_2 and $2N_2$ have been performed using Bauschlicher’s DZP ($9s5p1d$)/[$4s2p1d$] modification [18] of Huzinaga’s basis set [23]. In order to allow a direct comparison of our results with the FCI ones, CASSCF orbitals have been obtained and the $1s$ and $2s$ orbitals are kept frozen in the CI procedure. Two different choices of the \mathcal{S} space are discussed: the CAS used in the CASSCF procedure and a space selected according to the CIPSI scheme [16].

We discuss first the case of a CAS zero-order space. The N_2 molecule is a rather small system for testing our extrapolation procedure. The CAS space considered by Bauschlicher and Langhoff [18] and in the CASPT2 [17] calculations consists of 56 determinants in the D_{2h} symmetry and either the CASPT2 or our two class calculation use only a few seconds of CPU time. Our MPB results and the CASPT2 ones are in good agreement. The errors with respect to FCI are $\Delta E_{MPB} = 0.004878$ and $\Delta E_{PT2} = 0.004914$ a.u. at $R_{NN} = 2.10$ bohr; $\Delta E_{MPB} = 0.000801$ and $\Delta E_{PT2} = 0.000264$ a.u. at $R_{NN} = 50$ bohr.

For the $2N_2$ system, two different geometries are considered: in both the distance between the two N_2 molecules is very large. In the first (called in the following ‘‘eq-eq’’) $R_{NN} = 2.10$ bohr for each of the two molecules, while in the second (‘‘diss-diss’’) $R_{NN} = 50$ bohr. The same distances were employed in the FCI [18] and in the CASPT2 calculations [17]. A 12-electron/12-orbital CAS space for the $2N_2$ system, consistent with the 6/6 CAS of a single N_2 molecule, contains 214 032 determinants: we ran our calculations

in the C_{2v} symmetry, with two co-linear N_2 molecules, in order to localize the orbitals on each N_2 subsystem. It is clear that this space contains a lot of meaningless charge transfer excitations: this may be considered an extreme example of the inconvenience which can be met when an a priori zero-order space is defined, namely the inclusion of a large number of useless determinants. In fact, when excluding the charge transfer determinants, the dimension of the \mathcal{S} space reduces to $N_{\mathcal{S}} = 4360$. We have employed this reduced space, which is in all respects equivalent to the CAS space, in the CIPSI calculations. It is therefore unfair to compare the CIPSI and CASPT2 computing times, but this shows the possible advantage of a procedure applicable to a completely general zero-order space.

For a discussion of the performance of three class calculations with extrapolation we only consider the “eq-eq” geometry, because in the CAS calculation the “diss-diss” wave function is described by a single configuration with 12 unpaired electrons. We show in Table 1 and in Fig. 1 a comparison between the full two class calculation and the extrapolated one. The size of the \mathcal{G} space, $N_{\mathcal{G}}$, even with $W_{\mathcal{G}}$ very close to 1, is drastically reduced with respect to $N_{\mathcal{S}}$: therefore we obtain a large speed-up (a factor of 12) with errors in the energies smaller than 10^{-4} a.u. Notice that the energy is not merely linear but practically constant in the range of $W_{\mathcal{G}}$ considered here. In fact, the simple three class procedure without extrapolation may yield very accurate results for CAS zero-order spaces. A three class calculation with $W_{\mathcal{G}} = 0.99999$ (1486 determinants) is still 4 times faster than the full two class treatment, with negligible errors: $-5 \cdot 10^{-6}$ a.u. (MPB) and $-6 \cdot 10^{-6}$ a.u. (EN).

The extrapolation procedure has also been tested in the case of individually selected \mathcal{S} spaces. The \mathcal{S} space for both eq-eq and diss-diss $2N_2$ was iteratively enlarged until the square norm of the first-order EN correction to the wave function was reduced to 0.0030. In spite of the smaller size $N_{\mathcal{S}}$, this choice of the \mathcal{S} space is inherently better than the previous one ($\mathcal{S} \equiv \text{CAS}$): it yields smaller perturbation corrections to the wave functions (as explicitly requested by the selection procedure [16]) and a closer agreement between MPB and EN results (see Table 1). Also the dissociation energies (computed as half the difference between diss-diss and eq-eq) are closer to the FCI limit [18]: the errors are -0.001748 and 0.001318 a.u. for MPB and EN respectively, while with $\mathcal{S} \equiv \text{CAS}$ we get -0.003312 and 0.003243 a.u. A six-point extrapolation procedure ($W_{\mathcal{G}} = 0.992, 0.993, 0.994,$

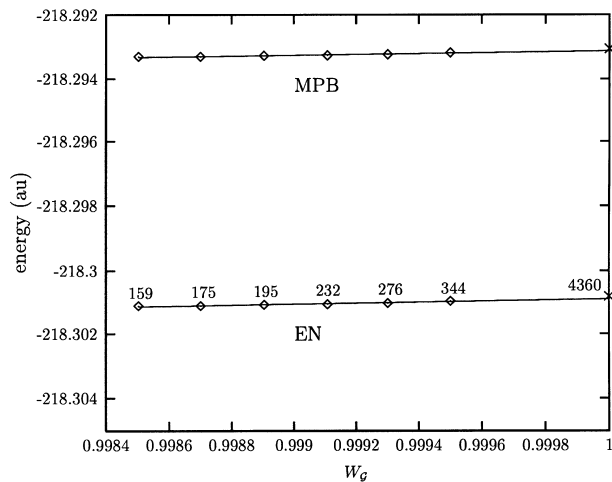


Fig. 1. $2N_2$ EN and MPB ground state energies (CIPSI with $\mathcal{S} \equiv \text{CAS}$, $N_{\mathcal{S}} = 4360$) with the eq-eq geometry (see text). The numbers near the points are the dimensions of the \mathcal{G}_j spaces. The points with $W_{\mathcal{G}} = 1$ correspond to a full two class calculation. \diamond = points used in the extrapolation; \times = points obtained with full calculation, shown for comparison

0.995, 0.996, 0.997) has been applied and the results compared with the full two class ones (see Table 1 and Fig. 2).

With a selected \mathcal{S} space the extrapolation procedure yields a lesser speed-up (about 3 times, with the parameters we have chosen) and a larger loss in accuracy, in comparison with the case $\mathcal{S} \equiv \text{CAS}$. Moreover, the slopes of the fitting straight lines are not negligible and not the same for the two geometries. This behaviour is due to the individual selection, which ensures that all the determinants in \mathcal{S} have non-negligible coefficients in the zero-order wave functions. Even so, the extrapolated results for the selected \mathcal{S} space differ from the full two class ones by only about 0.0005 a.u.

We note that the extrapolation improves the evaluation of dissociation energies, with respect to a three class calculation with a single \mathcal{G} space, because the slopes of the energy versus $W_{\mathcal{G}}$ plots are different for the eq-eq and the diss-diss geometries. The errors in dissociation energy, introduced by the three class + extrapolation procedure, with respect to a full two class calculation, are -0.000480 and -0.000466 a.u. for MPB and EN respectively: they are substantially smaller than the errors of two class perturbation versus FCI (see above).

Table 1. Comparison of full two class and extrapolated three class results for the $2N_2$ system for two different geometries (“eq-eq” = two N_2 molecules with $R_{\text{NN}} = 2.10$ bohr, “diss-diss” = four N atoms). Selected and CAS zero-order spaces are used. Energies in a.u., CPU times in seconds

	Full two class calculation			Three class + extrapolation ^a		
	MP	EN	Time	MP	EN	Time
$\mathcal{S} \equiv \text{CAS}^b$ (eq-eq)	-218.293058	-218.300812	1412	-0.000064	-0.000087	113
Sel - \mathcal{S}^c (eq-eq)	-218.303712	-218.302268	707	-0.000528	-0.000495	227
Sel - \mathcal{S}^c (diss-diss)	-217.657386	-217.657386	45	0.000432	0.000417	31

^a Energy difference versus full two class results

^b The \mathcal{S} space is the 12-electron/12-orbital CAS. The extrapolation is based on six \mathcal{G}_j subspaces defined by the thresholds $\tau = 0.9985, 0.9987, 0.9989, 0.9991, 0.9993, 0.9995$ (see Fig. 1)

^c Individually selected \mathcal{S} space (see text). The extrapolation is based on six \mathcal{G}_j subspaces defined by the thresholds $\tau = 0.992, 0.993, 0.994, 0.995, 0.996, 0.997$ (see Fig. 2)

4 Test calculations: CH₃NO

A second series of calculations has been performed using the CH₃NO molecule, which is at present under investigation in our laboratory in the framework of a study of the photochemistry of nitrosocompounds [24]. The 6-31G* Gaussian basis set [25] has been used. As in the previous section, we present the results of our

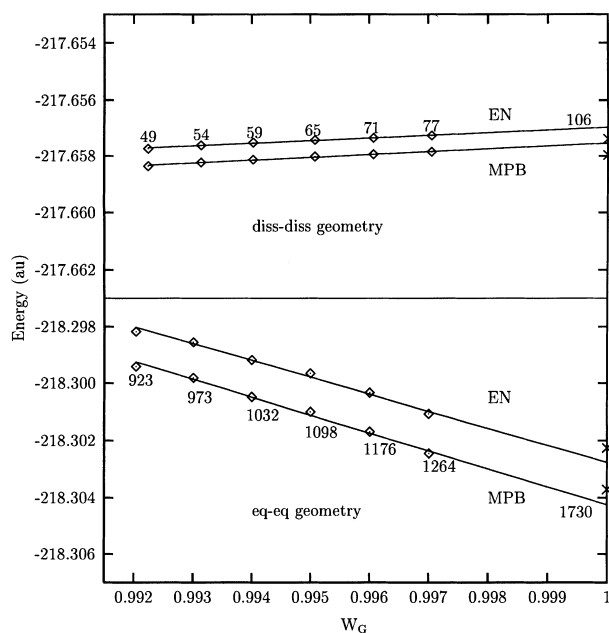


Fig. 2. $2N_2$ EN and MPB ground state energies (CIPSI with selected \mathcal{S} space). The \mathcal{S} space is selected in order to obtain a square norm of the first-order EN correction to the wave function of 0.003. The numbers near the points are the dimensions of the \mathcal{G}_j spaces. The points with $W_G = 1$ correspond to a full two class calculation. \diamond = points used in the extrapolation; \times = points obtained with full calculation, shown for comparison

extrapolation procedure compared with the corresponding two class calculation, both in the case of CAS and of selected \mathcal{S} spaces. Two different geometries are considered: the first is the fully optimized geometry for the ground state (at the CASSCF 10 electrons in 10 orbitals level), while in the second the C—N bond is broken and the geometries of the two non-interacting radicals are separately optimized at a consistent CASSCF level [26]. Besides the total energy of the states (ground state S_0 , first excited singlet S_1 and first excited triplet T_1), we consider in this section the C—N bond dissociation energy and the vertical excitation energies of the S_1 and T_1 states, as an indication of the accuracy of our extrapolation procedure with energy differences.

We report in Table 2 a summary of the results obtained with the absolute energies. With 10 electrons and 10 active orbitals (C_s symmetry), the CAS space consists of 32144 determinants. A full two class perturbation calculation using this space is rather expensive: the CPU time needed can be estimated in about 200 000 s. Given the results of the previous section in the case of a CAS zero-order space, we can assume that a three class calculation with $W_G = 0.99999$ is a very good approximation to a full two class one. As in the case of $2N_2$, even a W_G so close to 1 results in a \mathcal{G} space substantially smaller than \mathcal{S} ($N_G = 10893$ for equilibrium geometry and 2724 for the dissociation). The ground state energies and the CPU times are reported in the first two rows of Table 2. These energies are taken as a reference to test the extrapolation procedure (based on the calculations with $W_G = 0.997, 0.998, \text{ and } 0.999$). In Fig. 3 we show the total energies as functions of W_G . Again the straight lines are very flat: this means that even a simple three class calculation with a very small \mathcal{G} space (some hundreds of determinants in \mathcal{G} with an \mathcal{S} space of 32144 determinants) gives reasonable results as shown in the previous section. The accuracy of the extrapolation, versus the two class results, is better than 0.0003 a.u., with a dramatic reduction in computer time. MOLCAS-CASPT2

Table 2. Comparison of full two class and extrapolated three class results for three electronic states of the CH₃NO molecule (“eq” = equilibrium geometry, S_0 and S_1 are the first two singlets and T_1 the first triplet) and for the dissociated fragments (“diss” = dissociation of the C—N bond). Selected and CAS zero-order spaces are used. Energies in a.u., CPU times in seconds

	Full two class calculation			Three class + extrapolation ^a		
	MP	EN	Time	MP	EN	Time
$\mathcal{S} \equiv \text{CAS}^b$ (S_0 eq)	-169.295855	-169.325794	26136	-0.000041	-0.000063	991
$\mathcal{S} \equiv \text{CAS}^b$ (S_0 diss)	-169.235875	-169.271133	2247	-0.000239	-0.000272	493
Small selection ^c (S_0 eq)	-169.300493	-169.328003	170	0.000290	0.000398	106
Small selection ^c (T_1 eq)	-169.264949	-169.287373	458	-0.000380	-0.000277	310
Small selection ^c (S_1 eq)	-169.232109	-169.256648	458	-0.000160	-0.000070	310
Small selection ^c (S_0 diss)	-169.237071	-169.270880	78	0.000938	0.001126	60
Large selection ^d (S_0 eq)	-169.309155	-169.320327	1488	-0.000014	0.000000	893
Large selection ^d (T_1 eq)	-169.272806	-169.280492	3910	0.000385	0.000437	2484
Large selection ^d (S_1 eq)	-169.240314	-169.249789	3910	0.000506	0.000545	2484
Large selection ^d (S_0 diss)	-169.245501	-169.262388	327	0.000886	0.001045	237

^a Energy differences versus full two class results

^b The \mathcal{S} space is the 10-electron/10-orbital CAS. The full two class result is approximated as a three class one with $\tau = 0.99999$. The extrapolation is based on three \mathcal{G}_j subspaces defined by the thresholds $\tau = 0.997, 0.998, 0.999$ (see Fig. 3)

^c Individually selected \mathcal{S} space, such that the square norm of the EN correction to the wave function is about 0.100 (see text). The extrapolation is based on six \mathcal{G}_j subspaces defined by the thresholds $\tau = 0.970, 0.975, 0.980, 0.985, 0.990, 0.995$ (see Fig. 4)

^d Individually selected \mathcal{S} space, such that the square norm of the EN correction to the wave function is about 0.060 (see text). Extrapolation as before (see Fig. 5)

calculations on the same system give energies between the MPB and EN ones (-169.309354 at equilibrium and -169.247334 at dissociation), with a CPU time of about 300 s. When comparing computing times, one should remember that two diagrammatic codes exist, the one we are making use of [7], and one limited to MPB calculations [5], which is almost twice as fast. We can conclude that, when applied to a CAS zero-order space, our procedure is only slightly more expensive than CASPT2 [19], although the latter is specialized to treat the $\mathcal{S} \equiv \text{CAS}$ case.

We have also run tests with selected \mathcal{S} spaces. Two series of calculations have been performed. The first one, with smaller \mathcal{S} spaces, is of approximately the same quality as the calculations with $\mathcal{S} \equiv \text{CAS}$: we have re-

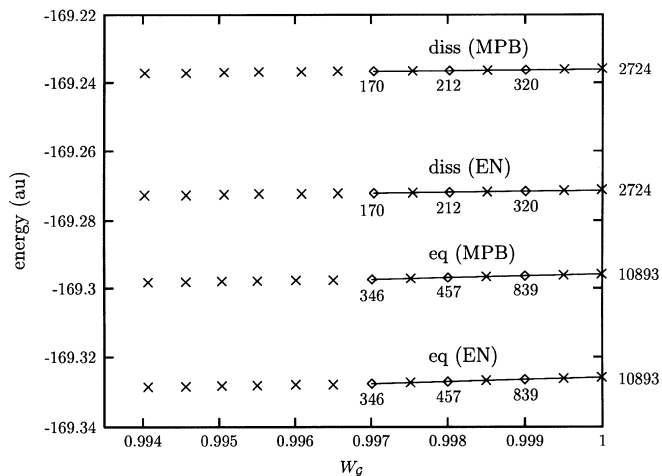


Fig. 3. CH_3NO EN and MPB ground state energies (CIPSI with $\mathcal{S} \equiv \text{CAS}$, $N_{\mathcal{G}} = 32144$). The numbers near the points are the dimensions of the \mathcal{G}_j spaces. The points for which $W_{\mathcal{G}}$ is closest to 1 correspond to $W_{\mathcal{G}} = 0.99999$. In this point $N_{\mathcal{G}} = 10893$ for the equilibrium geometry and $N_{\mathcal{G}} = 2724$ for dissociation. eq = optimized equilibrium geometry; diss = optimized geometry at the C–N dissociation. \diamond = points used in the extrapolation; \times = points shown for comparison

Table 3. Comparison of full two class and extrapolated three class results. Dissociation and vertical excitation energies (in kcal/mol) for the CH_3NO molecule. Same zero-order spaces and extrapolation procedure as in Table 2

	Full two class calc.		Three class + extrapol ^a	
	MP	EN	MP	EN
$\mathcal{S} \equiv \text{CAS}$ ($E_{\text{diss}}^{\text{b}}$)	37.64	34.30	-.13	-.13
Sel- \mathcal{S} ($E_{\text{diss}}^{\text{b}}$)	39.80	35.85	.41	.46
Sel- \mathcal{S} ($E_{S_0} \rightarrow T_1$) ^c	22.31	25.50	-.42	-.42
Sel- \mathcal{S} ($E_{S_0} \rightarrow S_1$) ^d	42.91	44.78	-.28	-.29
Sel- \mathcal{S} ($E_{\text{diss}}^{\text{b}}$)	39.94	36.36	.56	.66
Sel- \mathcal{S} ($E_{S_0} \rightarrow T_1$) ^c	22.81	25.00	.25	.27
Sel- \mathcal{S} ($E_{S_0} \rightarrow S_1$) ^d	43.20	44.26	.33	.34

^a Energy differences versus full two class results.

^b Ground state dissociation energy. Experimental [27, 28]: 39.0 ± 40.3 kcal/mol.

^c Vertical excitation energy to T_1

^d Vertical excitation energy to S_1 . Experimental [29, 30]: 42.4 kcal/mol.

quired that the square norm of the EN first-order correction to the wave function is about 0.10, the value obtained with $\mathcal{S} \equiv \text{CAS}$ using the equilibrium geometry. In the second series of calculations we have reduced the square norm of the first-order correction to the value of 0.060 (our standard in the study of the energy surfaces of CH_3NO).

The results are shown in Tables 2 and 3 and in Figs. 4 and 5. They confirm the conclusions of the previous

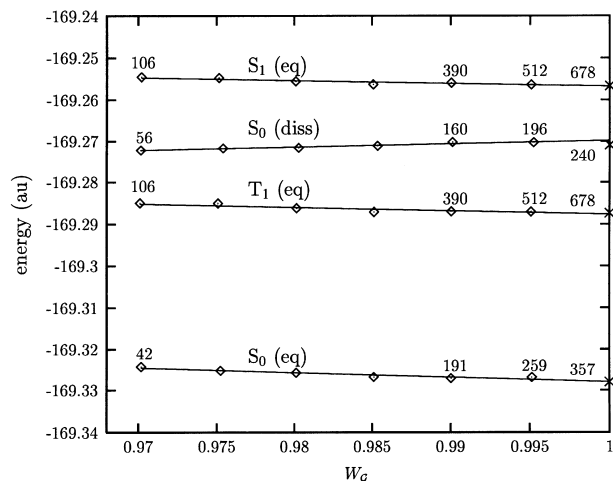


Fig. 4. CH_3NO ground state (S_0) and excited state (T_1 and S_1) EN energies (CIPSI with selected \mathcal{S} space). The \mathcal{S} space is selected, aiming to obtain a square norm of the first-order EN correction to the wave function of the same order as in the calculation using the CAS (0.100). The numbers near the points are the dimensions of the \mathcal{G}_j spaces. The points with $W_{\mathcal{G}} = 1$ correspond to a full two class calculation. \diamond = points used in the extrapolation; \times = points obtained with full calculation, shown for comparison

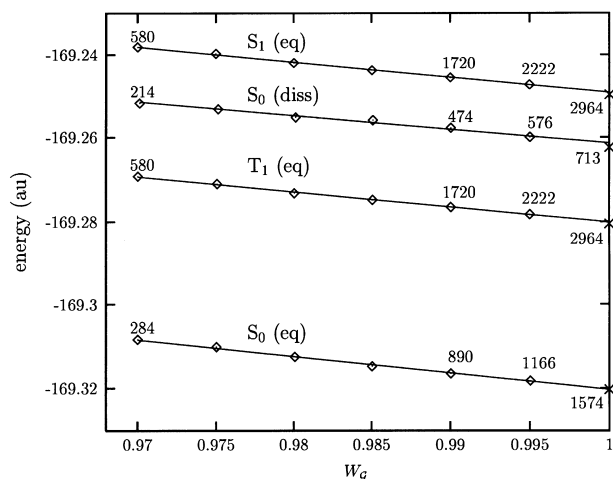


Fig. 5. CH_3NO ground state (S_0) and excited state (T_1 and S_1) EN energies (CIPSI with selected \mathcal{S} space). The \mathcal{S} space is selected in order to obtain a square norm of the first-order EN correction to the wave function equal to 0.060. The numbers near the points are the dimensions of the \mathcal{G}_j spaces. The points for with $W_{\mathcal{G}} = 1$ correspond to a full two class calculation. \diamond = points used in the extrapolation; \times = points obtained with full calculation, shown for comparison

section: the extrapolation procedure for selected spaces is less effective than in the case $\mathcal{S} \equiv \text{CAS}$, but is still convenient and reliable. In Fig. 4 we again find opposite slopes for the equilibrium geometry and the dissociated fragments, showing that the extrapolation may be important to determine energy differences. Excitation and dissociation energies are given in Table 3. The errors in the energy differences obtained with our extrapolation procedure with respect to the full two class results are practically negligible when $\mathcal{S} \equiv \text{CAS}$; they are of the order of 0.4 kcal/mol for selected \mathcal{S} spaces.

5 Conclusions

In this paper we have considered the approximation introduced in a three class CIPSI calculation. A new extrapolation procedure has been proposed to eliminate this error and therefore to recover the results of a full two class calculation. The extrapolation is effective and reliable when employed in conjunction with the usual CIPSI selection scheme, because it allows a considerable reduction in computational effort with errors in the energies of the order of few tenths of kcal/mol.

In the case of a CAS zero-order space the speed-up obtained with the three class + extrapolation procedure is more dramatic, because a very large number of highly excited determinants in the zero-order wave function can be safely neglected in the perturbation step. For the CH_3NO molecule the computer time drops from an estimated 200 000 s of a full two class calculation to 991 (equilibrium geometry) or 493 (dissociated fragments), without any practical loss in accuracy. On the other hand, the completeness of a CAS is exploited in the very efficient CASPT2 algorithm, which yields approximately the same results with comparable, although less, computational effort.

In conclusion, the progress presented here in the general purpose multireference perturbation CIPSI algorithm makes it competitive with CASPT2: more flexible as to the choice of the zero-order space and to the partition of the CI Hamiltonian, and comparable in performance for CAS wave functions.

Acknowledgements. The authors gratefully thank the Istituto di Chimica Quantistica ed Energetica Molecolare of Pisa for the computer facilities. This work has been supported by the Italian MURST within its 40% and 60% research funds.

References

- Huron B, Malrieu J-P, Rancurel P (1973) *J Chem Phys* 58: 5745
- Evangelisti S, Daudey J-P, Malrieu J-P (1983) *Chem Phys* 75: 91
- Spiegelmann F, Malrieu J-P (1984) *J Phys B* 17: 1235
- Malrieu J-P, Heully J-L, Zaitsevskii A (1995) *Theor Chim Acta* 90: 167
- Cimiraglia R (1985) *J Chem Phys* 83: 1746
- Cimiraglia R, Persico M (1987) *J Comp Chem* 8: 39
- Cimiraglia R (1996) *Int J Quantum Chem* 60: 167
- Wolinski K, Pulay P (1989) *J Chem Phys* 90: 3647
- Harrison RJ (1991) *J Chem Phys* 94: 5021
- Shavitt I (1992) *Chem Phys Lett* 192: 135
- Lisini A, Decleva P (1992) *Chem Phys* 168: 1
- Hoffmann MR (1992) *Chem Phys Lett* 195: 127
- Hirao K (1993) *Chem Phys Lett* 201: 59
- Kozłowski PM, Davidson ER (1994) *J Chem Phys* 100: 3672
- Davidson ER (1995) *Chem Phys Lett* 241: 432
- Angeli C, Persico M, *Theor Chem Acc* (in press)
- Andersson K, Malmqvist P-Å, Roos BO (1992) *J Chem Phys* 96: 1218
- Bauschlicher CW, Langhoff SR (1987) *J Chem Phys* 86: 5595
- Andersson K, Blomberg MRA, Fülscher MP, Kellö V, Lindh R, Malmqvist P-Å, Noga J, Olsen J, Roos BO, Sadlej AJ, Siegbahn PEM, Urban M, Widmark P-O (1994) *MOLCAS version 3*. University of Lund, Sweden
- Malrieu J-P, Spiegelmann F (1979) *Theor Chim Acta* 52: 55
- Roos BO, Andersson K (1995) *Chem Phys Lett* 245: 215
- Claverie P, Diner S, Malrieu J-P (1967) *Int J Quantum Chem* 1: 751
- Huzinaga S (1965) *J Chem Phys* 42: 1293
- Cimiraglia R, Persico M, Tomasi J (1985) *J Am Chem Soc* 107: 1617
- Hariharan PC, Pople JA (1973) *Theor Chim Acta* 28: 213
- Toniolo A (1997) Thesis, University of Pisa
- Jodkowski JT, Ratajczak E, Sillesen A, Pagsberg P (1993) *Chem Phys Lett* 203: 490
- Wolff T, Wagner HG (1988) *Ber Bunsen-Ges Phys Chem* 92: 678
- Dixon RN, Kroto HW (1965) *Proc R Soc Lond Ser A* 283: 423
- Ernsting NP, Pfab J, Römel J (1978) *J Chem Soc Faraday Trans II* 74: 2286