Theor Chem Acc (1997) 98:117–128

Theoretical
Chemistry Accounts

Multireference perturbation CI II. Selection of the zero-order space

Celestino Angeli[1], **Maurizio Persico**[2]

[1] Scuola Normale Superiore di Pisa, p.za Cavalieri 7, I-56100 Pisa, Italy
[2] Dipartimento di Chimica e Chim. Ind., Università di Pisa, v. Risorgimento 35, I-56126 Pisa, Italy

Received: 12 June 1997 / Accepted: 31 July 1997

Abstract. In this paper, the second of a series devoted to multi-reference perturbation CI, we tackle the problem of the appropriate selection of the zero-order space in CIPSI calculations. We propose a new selection procedure, explicitly devised in order to obtain a balanced description for different electronic states and nuclear geometries. To this aim, we define numerically the quality of the zero-order space by means of a suitable parameter σ, which is the square norm of the perturbative correction of the wavefunction. The zero-order space is expanded stepwise so as to obtain the same target σ for all states and geometries. This strategy is applied to the calculation of dissociation, activation and transition energies. It yields a much better convergence of perturbative and zero-order results, when compared with the selection procedure previously used.

Key words: Multi-reference CI – Perturbation CI – CIPSI

1 Introduction

Multi-reference perturbation CI methods are among the most powerful, general purpose tools for the investigation of energies and properties of electronic states (see Malrieu et al. for a recent review [1]). All of them are based on the definition of an appropriate zero-order CI subspace $\mathscr{S}$, which contains the most important Slater determinants or configurations for the description of the electronic states of interest. The methods proposed in the past and those currently employed differ as to important choices, such as the order of application of variational and perturbation theory. The option "perturbation first" implies the use of quasi degenerate perturbation theory (QDPT) in the basis of the determinants belonging to $\mathscr{S}$ ("uncontracted" procedure). It leads to the definition of intermediate Hamiltonians [2], a generalization of the concept of effective Hamiltonians (see also Malrieu et al. [3] for further developments).

The option "diagonalization first" involves computation of the eigenvectors of the CI Hamiltonian projected into the $\mathscr{S}$ subspace and perturbation of only a few of the zero-order wave functions thus obtained, corresponding to the electronic states of interest ("contracted" procedure). Examples of the latter strategy are the CIPSI [4] and the CASPT2 [5] algorithms, both based on second-order perturbation theory. The contraction of the zero-order wave functions has an advantage, in that a substantial energy gap is introduced between the target states and the perturbers. However, the simplest, state-by-state, applications of this strategy do not allow for the mixing of the states belonging to $\mathscr{S}$, under the influence of the interaction with outer configurations. This may be important in cases of quasi-degeneracy, such as avoided crossings and conical intersections. Such a problem can be solved to a large extent by a QDPT treatment in the basis of the zero-order target states [6], which is an option in our implementation of the method [7].

In CIPSI the determinants belonging to $\mathscr{S}$ are individually selected by means of an iterative procedure, while in CASPT2, $\mathscr{S}$ is a complete active space, i.e. a full CI space in a restricted set of molecular orbitals (MOs). Other strategies may be applied, for instance selecting configurations according to the class of excitation with respect to a reference wave function [8].

In all cases, the appropriate choice of the $\mathscr{S}$ subspace is crucial, because it determines the accuracy of the perturbation corrections. The basic requirement is that the quality of $\mathscr{S}$ should be uniform throughout the set of electronic states and the range of molecular geometries under consideration. None of the procedures mentioned above are explicitly devised to meet such a requirement, even those endowed with the important formal property of size consistency [1, 9, 10], for instance, CASPT2 [5] or effective Hamiltonian-based methods [3].

In this paper we propose a simple and automatic procedure to select $\mathscr{S}$ with a uniform quality, suitably defined as a numerical parameter, within the CIPSI

e-mail: mau@hermes.dcci.unipi.it
Correspondence to: M. Persico

method. Test calculations show the effectiveness of the new procedure in calculating dissociation, activation and excitation energies.

2 Method

The CIPSI method was reviewed in the first paper of this series [11] (see also Cimiraglia and Persico [7]). It rests on a partition of the CI Hamiltonian, usually according to the Epstein-Nesbet (EN) or Møller-Plesset baricentric (MPB) schemes:

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_0 + \hat{\mathscr{V}} \quad . \tag{1}$$

The second-order Rayleigh-Schrödinger perturbation theory is applied to the eigenstates of $\hat{\mathscr{H}}_0$ in the $\mathscr{S}$ subspace; in the EN partition:

$$\hat{\mathscr{H}}_0 \left|\Psi_m^{(0)}\right\rangle = E_m^{(MR)} \left|\Psi_m^{(0)}\right\rangle \quad . \tag{2}$$

Here $E_m^{(MR)}$ is the multi-reference zero-order energy of the m-th electronic state. The second-order correction to the energy is

$$E_m^{(2)} = -\sum_{I\in\mathscr{P}} \frac{\left\langle I\left|\hat{\mathscr{V}}\right|\Psi_m^{(0)}\right\rangle^2}{E_I - E_m^{(MR)}} \tag{3}$$

and the first-order correction to the wave function is

$$\left|\Psi_m^{(1)}\right\rangle = \sum_{I\in\mathscr{P}} C_{I,m}|I\rangle \quad , \tag{4}$$

with coefficients

$$C_{I,m} = -\frac{\left\langle I\left|\hat{\mathscr{V}}\right|\Psi_m^{(0)}\right\rangle}{E_I - E_m^{(MR)}} \quad . \tag{5}$$

Here $|I\rangle$ is a determinant belonging to the space $\mathscr{P}$ of all single and double excitations from determinants of $\mathscr{S}$ and E_I, its energy: higher classes of excitation do not contribute until the fourth-order. The total energies are the sum of zero- and second-order terms, $E_m^{(MR)} + E_m^{(2)}$. The Møller-Plesset formulae only differ in the denominators [4, 7, 11].

2.1 Selection based on wave function or energy thresholds

Two sources of error affect the results: one is the truncation of the total CI space considered, $\mathscr{S} \cup \mathscr{P}$; the other one is the use of second-order perturbation formulae instead of an exact diagonalization. A selective enlargement of the $\mathscr{S}$ subspace may reduce both sources of error. In the CIPSI algorithm, this is done iteratively, usually starting with an $\mathscr{S}$ subspace of a few determinants. At each step, the largest coefficients $C_{I,m}$ in the first-order correction of the wave functions are identified, and the corresponding determinants $|I\rangle$ are included in the enlarged $\mathscr{S}$ space for the next calculation. We shall call $\mathscr{M}$ the subset of newly selected determinants, and $\mathscr{S}' = \mathscr{S} \cup \mathscr{M}$ the new zero-order space. In this way, the most important contributions to the correlation energy are taken into account by diagonalization in the next step. Moreover, the total CI space is expanded by addition of the subspace $\mathscr{P}'''$, made of all single and double excitations with respect to $\mathscr{M}$, which could not be generated from $\mathscr{S}$ in the previous step. The new perturbation space is thus $\mathscr{P}' = \mathscr{P}'' \cup \mathscr{P}'''$, with $\mathscr{P}'' = \mathscr{P} - \mathscr{M}$ (see Fig. 1).

A strategy which has been applied over the past several years consists of setting a threshold $\eta^{(n)}$ for the selection of determinants at step n: then, all $|I\rangle \in \mathscr{P}$ with $\left|C_{I,m}\right| \geq \eta^{(n)}\left|C_m^{(\max)}\right|$ for at least one state m are included in the zero-order space $\mathscr{S}'$ in the next step, $n+1$. $C_m^{(\max)}$ is the largest coefficient in the zero-order eigenfunction of step n, $\Psi_m^{(0,n)}$. The modification of the threshold by the factor $\left|C_m^{(\max)}\right|$ introduces a state-by-state adaptation to the type of wave function expansion in the determinantal basis. This is done in order to select equivalent subspaces $\mathscr{S}$ for states whose dominant configurations differ in the number of open shells and associated determinants. The $\eta^{(n)}$ threshold is lowered step by step, and the $\mathscr{S}$ space is correspondingly expanded, in principle, until the desired accuracy of the computed properties is obtained. In the following, we shall call this procedure, based on the contributions to the first-order wave function, Ψ-selection.

In practice, the computer facilities available may limit the dimension of the $\mathscr{S}$ space, $N_{\mathscr{S}}$, and the number of correlated electrons, N_e. However, the standard of computation has increased over the years, as a result of methodological advances [11–13] and better hardware performance. Typical examples may be found in applicative works [14–17]. As more experience about the convergence properties of the CIPSI algorithm with larger $N_{\mathscr{S}}$ and N_e has accumulated, we have found that Ψ-selection may yield markedly different slopes of energy versus η plots, for various electronic states or nuclear geometries. The unbalanced selection leads to errors in dissociation and activation or transition energies, unless the full CI limit is approached. Examples will be given in Sect. 3. Notice that similar problems may be met with CASPT2, as in the study [5] of the dissociation energy of N_2.

An alternative to Ψ-selection is the analogous procedure based on the contributions to the energies $E_m^{(2)}$ (E-selection). In this case, the appropriate factor which multiplies the energy threshold $\theta^{(n)}$ is $(C_m^{(\max)})^2$: all determinants $|I\rangle$ with $\left|\left\langle I\left|\hat{\mathscr{V}}\right|\Psi_m^{(0)}\right\rangle^2/(E_I - E_m^{(MR)})\right| > \theta^{(n)}$ $(C_m^{(\max)})^2$ are selected and included in $\mathscr{S}$ for the next

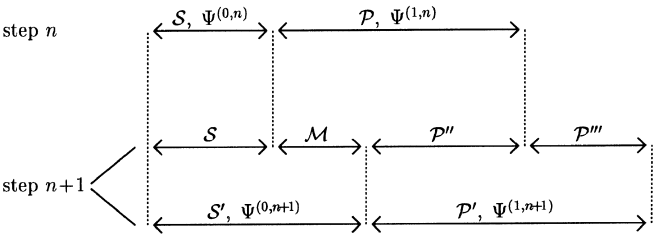


Fig. 1. Partition of the CI space in two successive CIPSI steps

calculation. A configuration selection based on energetic contributions is the standard in multireference single and double excitation CI (MRD-CI) [18, 19]. As we shall see in Sect. 3, E-selection suffers from the same problems as Ψ-selection.

2.2 Aimed selection

A good measure of the quality of perturbation CI calculations is the square norm of the first-order correction to the wave function:

$$\sigma_m = \sum_{I\in\mathscr{P}} C_{I,m}^2 = \sum_{I\in\mathscr{P}} \frac{\left\langle I\middle|\hat{\mathscr{V}}\middle|\Psi_m^{(0)}\right\rangle^2}{\left(E_I - E_m^{(MR)}\right)^2} \quad . \tag{6}$$

If large interactions $\left\langle I\middle|\hat{\mathscr{V}}\middle|\Psi_m^{(0)}\right\rangle$ and/or small denominators $E_I - E_m^{(MR)}$ occur in Eqs. (3) and (6), the results are not accurate, and σ_m is correspondingly large. When the $\mathscr{S}$ space is extended, the most important contributions in Eq. (6) are eliminated; therefore σ_m usually decreases, although new terms involving determinants $|I\rangle \in \mathscr{P}'''$ are added. We also recall that σ_m, or related parameters, play a central role in all attempts to define size-extensivity corrections to MRD-CI methods [10]. In the examples shown in Sect. 3, we shall see that the Ψ-selection procedure may lead to quite different σ values, e.g. for a molecule in its equilibrium geometry and for two dissociated fragments. The same happens with E-selection.

These observations have prompted us to devise an algorithm to select the $\mathscr{S}$ space so as to obtain a preset value of σ for all states and nuclear geometries. A slightly weaker requirement is that all σ_m values should be the same, although not equal to a target value. We shall call such a procedure "*aimed* selection". In order to set up an *aimed* selection, we must be able to predict the σ value we would obtain at the step $n+1$ of a CIPSI calculation, from the results of the step n, at least with some approximation.

In the following, for simplicity, we shall drop the state subscript m, if not strictly needed, and we shall add instead an index n or $n+1$ to distinguish the quantities computed at two subsequent steps. The theory applies to the EN perturbation results, because of the relationship we want to establish between diagonal energies and denominators in the perturbation formulae, Eqs. (10–13). Let us define the normalized first-order wave function $(\sigma^{(n)})^{-1/2}\Psi^{(1,n)}$. We shall call V its interaction with $\Psi^{(0)}$:

$$V = (\sigma^{(n)})^{-1/2}\left\langle \Psi^{(0,n)}\middle|\hat{\mathscr{V}}\middle|\Psi^{(1,n)}\right\rangle \quad . \tag{7}$$

From the relationship between $\Psi^{(1)}$ and $E^{(2)}$,

$$E^{(2,n)} = \left\langle \Psi^{(0,n)}\middle|\hat{\mathscr{V}}\middle|\Psi^{(1,n)}\right\rangle \tag{8}$$

we get:

$$V^2 = \frac{\left(E^{(2,n)}\right)^2}{\sigma^{(n)}} \quad . \tag{9}$$

The average energy associated with $\Psi^{(1,n)}$, taking $E^{(MR,n)}$ as a reference, is Δ:

$$\Delta = (\sigma^{(n)})^{-1}\left\langle \Psi^{(1,n)}\middle|\hat{\mathscr{H}}\middle|\Psi^{(1,n)}\right\rangle - E^{(MR,n)} \quad . \tag{10}$$

If we consider $(\sigma^{(n)})^{-1/2}\Psi^{(1)}$ as a single basis function, we can approximate the second-order energy correction as:

$$E^{(2,n)} = -\frac{V^2}{\Delta} \quad . \tag{11}$$

Then:

$$\Delta = -\frac{E^{(2,n)}}{\sigma^{(n)}} \quad . \tag{12}$$

Equations (9) and (12) allow us to compute the two parameters V^2 and Δ from the perturbation results $E^{(2,n)}$ and $\sigma^{(n)}$. Notice that one obtains formally the same formulae with the approximation that all denominators are equal in Eqs. (3) and (6). We can similarly treat the partial results $E_{\mathscr{M}}^{(2,n)}$ and $\sigma_{\mathscr{M}}^{(n)}$, which are obtained by taking into account only determinants $|I\rangle \in \mathscr{M}$ in Eqs. (3) and (6). We shall call $V_{\mathscr{M}}^2 = (E_{\mathscr{M}}^{(2,n)})^2/\sigma_{\mathscr{M}}^{(n)}$ and $\Delta_{\mathscr{M}} = -E_{\mathscr{M}}^{(2,n)}/\sigma_{\mathscr{M}}^{(n)}$ the parameters referring to the $\mathscr{M}$ subspace, while V^2 and Δ belong to the full $\mathscr{P}$ space.

In the next step, $n+1$, the CI matrix in the space $\mathscr{S}' = \mathscr{S} \cup \mathscr{M}$ will be diagonalized. We shall consider as basis functions $\Psi^{(0,n)}$ and $\Psi_{\mathscr{M}}$ (the normalized projection of $\Psi^{(1,n)}$ into $\mathscr{M}$), i.e. we shall neglect the inner flexibility of the $\mathscr{S}$ and $\mathscr{M}$ determinantal basis sets. It is clear that, in this attempt to predict the results of the step $n+1$, we cannot take into account the interactions between determinants of $\mathscr{P}$, which are never computed in step n. With this approximation, the CI matrix reduces to:

$$\begin{pmatrix} E^{(MR,n)} & V_{\mathscr{M}} \\ V_{\mathscr{M}} & E^{(MR,n)} + \Delta_{\mathscr{M}} \end{pmatrix} \quad . \tag{13}$$

The lowest eigenvalue of this matrix approximates the zero-order energy at step $n+1$:

$$\begin{aligned} E^{(MR,n+1)} &= E^{(MR,n)} - \frac{\sqrt{\Delta_{\mathscr{M}}^2 + 4V_{\mathscr{M}}^2} - \Delta_{\mathscr{M}}}{2} \\ &= E^{(MR,n)} + E_{\mathscr{M}}^{(2,n)}\, f\left(\sigma_{\mathscr{M}}^{(n)}\right) \quad , \end{aligned} \tag{14}$$

where $f(x)$ is the function

$$f(x) = \frac{\sqrt{1+4x}-1}{2x} \simeq 1 - x \quad . \tag{15}$$

The corresponding eigenstate is:

$$\left|\Psi^{(0,n+1)}\right\rangle = (1 + C_{\mathscr{M}}^2)^{-1/2}\left(\left|\Psi^{(0,n)}\right\rangle + C_{\mathscr{M}}\left|\Psi_{\mathscr{M}}\right\rangle\right) , \tag{16}$$

where

$$C_{\mathscr{M}}^2 = \sigma_{\mathscr{M}}^{(n)}\, f^2\left(\sigma_{\mathscr{M}}^{(n)}\right) \quad . \tag{17}$$

At step $n+1$, the perturbation involves two subspaces, $\mathscr{P}''$ and $\mathscr{P}'''$ (see Fig. 1). We consider first the normalized projection of $\Psi^{(1,n)}$ into $\mathscr{P}''$. Its interaction with $\Psi^{(0,n)}$, $V_{\mathscr{P}''}^2$, and the effective energy, $\Delta_{\mathscr{P}''}$, can be computed again from the results at step n, taking into account only

the terms due to determinants $|I\rangle \notin \mathscr{M}$ in Eqs. (3) and (6). We then get:

$$V^2_{\mathscr{P}''} = \frac{(E^{(2,n)} - E^{(2,n)}_{\mathscr{M}})^2}{\sigma^{(n)} - \sigma^{(n)}_{\mathscr{M}}} \quad , \tag{18}$$

$$\Delta_{\mathscr{P}''} = -\frac{E^{(2,n)} - E^{(2,n)}_{\mathscr{M}}}{\sigma^{(n)} - \sigma^{(n)}_{\mathscr{M}}} \quad . \tag{19}$$

The $\mathscr{P}''$ contribution to the square norm of the wave function correction $\sigma^{(n+1)}$ in the CIPSI step $n+1$ is:

$$\sigma^{(n+1)}_{\mathscr{P}''} = \frac{V^2_{\mathscr{P}''}}{[1 + C^2_{\mathscr{M}}]\left[\Delta_{\mathscr{P}''} - E^{(2,n)}_{\mathscr{M}} f(\sigma^{(n)}_{\mathscr{M}})\right]^2} \quad . \tag{20}$$

Here we have taken into account the normalization of wave function (16), which reduces the effective interaction by a factor $1/(1 + C^2_{\mathscr{M}})$, and the lowering of the reference energy, Eq. (14).

For the $\mathscr{P}'''$ subspace we have no direct information from the n-th calculation; therefore, we must resort to an "educated guess" as to its contribution in step $n+1$. We notice that the determinants in $\mathscr{P}'''$ only interact with those of $\mathscr{M}$, so we assume that the interaction between $\mathscr{P}'''$ and $\mathscr{S}'$ is the same as between $\mathscr{P}$ and $\mathscr{S}$, but for a factor $C^2_{\mathscr{M}}/(1 + C^2_{\mathscr{M}})$. The effective energy parameter $\Delta_{\mathscr{P}'''}$ is computed as $\Delta_{\mathscr{P}''}$, considering the hypothesis that $\sigma^{(n)}_{\mathscr{M}} = \sigma^{(n)}/2$. That is, we assume that the average energy of determinants of $\mathscr{P}'''$ is the same as in the upper half of $\mathscr{P}''$. In this way, the $\mathscr{P}'''$ contribution to $\sigma^{(n+1)}$ is estimated as:

$$\sigma^{(n+1)}_{\mathscr{P}'''} = \frac{V^2\, C^2_{\mathscr{M}}}{[1 + C^2_{\mathscr{M}}]\left[\Delta_{\mathscr{P}''} - E^{(2,n)}_{\mathscr{M}} f(\sigma^{(n)}_{\mathscr{M}})\right]^2} \quad . \tag{21}$$

The total square norm of the first-order correction is predicted, at step $n+1$, to be:

$$\sigma^{(n+1)} = \sigma^{(n+1)}_{\mathscr{P}''} + \sigma^{(n+1)}_{\mathscr{P}'''} \quad . \tag{22}$$

Normally $\sigma^{(n)}_{\mathscr{M}}$ is a rather small fraction of σ because we select a moderate number of determinants from the $\mathscr{P}$ space at each step: the $\sigma^{(n+1)}_{\mathscr{P}''}$ contribution is then more important than $\sigma^{(n+1)}_{\mathscr{P}'''}$ and the approximations made in evaluating the latter should be quite acceptable. In the limiting case where $\mathscr{M}$ is empty, $\sigma^{(n+1)}_{\mathscr{P}'''}$ vanishes and $\sigma^{(n+1)}_{\mathscr{P}''} = \sigma^{(n)}$. In the opposite case, that is including the whole $\mathscr{P}$ space in $\mathscr{S}'$, we have $\sigma^{(n+1)}_{\mathscr{P}''} = 0$. Therefore, the quantity $\sigma^{(n+1)}_{\mathscr{P}'''}$ computed with $\sigma^{(n)}_{\mathscr{M}} = \sigma^{(n)}$ is an estimate of the smallest $\sigma^{(n+1)}$ value we can obtain in step $n+1$, with the maximum extension of the $\mathscr{S}'$ space:

$$\sigma^{(n+1)}_{\min} = \frac{V^2\, \sigma^{(n)}\, f^2(\sigma^{(n)})}{[1 + \sigma^{(n)} f^2(\sigma^{(n)})][\Delta_{\mathscr{P}'''} - E^{(2,n)} f(\sigma^{(n)})]^2} \quad . \tag{23}$$

The *aimed* selection scheme is then implemented as follows:

1. Perform the CIPSI step n; write a file containing the most important determinants and their contributions to $E^{(2,n)}_m$ and $\sigma^{(n)}_m$.

2. Set a target value $\sigma^{(t)}$ for $\sigma^{(n+1)}$.

3. At the beginning the $\mathscr{M}$ space is empty but it will be incremented in steps 4 and 5.

4. Compute $\sigma^{(n+1)}_m$ for each state m, according to the updated definition of the $\mathscr{M}$ space, and compare it with the target. If $\sigma^{(n+1)}_m < \sigma^{(t)}$ for all states, the selection procedure ends; otherwise, the state m' with the largest $\sigma^{(n+1)}_m$ is singled out.

5. Increase the $\mathscr{M}$ space with the determinant $|I\rangle$ yielding the largest contribution to $\sigma^{(n)}_{m'}$.

6. Go back to point 4.

A variant of point 5 is optionally active, if one desires to include in $\mathscr{S}'$ the determinants with a given MO occupation and all possible spin functions, in order to produce eigenstates of $\mathscr{S}^2$ in the next step. In this case, all the contributions to $E^{(2,n)}$ and $\sigma^{(n)}$ of the required determinants are taken into account in the selection algorithm ("pure-spin" option). The *aimed* selection procedure is much faster than the corresponding perturbation CI calculation. The most demanding step is the evaluation of $\Delta_{\mathscr{P}''}$, which can be avoided if one neglects the $\sigma^{(n+1)}_{\mathscr{P}'''}$ contribution to $\sigma^{(n+1)}$. We shall see examples of both options in Sect. 3.

A further simplification of Eqs. (17–20) is obtained by neglecting terms of second or higher order in $\sigma^{(n)}$, and the energy lowering due to the enlargement of the $\mathscr{S}$ space. The consequent loss in accuracy has been found negligible in several examples, probably because of a cancellation of errors. The resulting formula is:

$$\sigma^{(n+1)}_{\mathscr{P}''} = \left(\sigma^{(n)} - \sigma^{(n)}_{\mathscr{M}}\right)\left(1 - \sigma^{(n)}_{\mathscr{M}}\right) \quad . \tag{24}$$

In summary, we have illustrated three options, from the simplest to the most elaborate: option 1, $\sigma^{(n+1)} = \sigma^{(n+1)}_{\mathscr{P}''}$ is computed by Eq. (24); option 2, $\sigma^{(n+1)} = \sigma^{(n+1)}_{\mathscr{P}''}$, Eqs. (17–20); option 3, also adds the $\sigma^{(n+1)}_{\mathscr{P}'''}$ contribution in Eqs. (17–22).

Other options are available, based on an improved evaluation of $\sigma^{(n+1)}_{\mathscr{P}''}$, with a little extra cost in the CIPSI calculation. We notice that Eq. (20) contains the unnecessary approximation of neglecting the inner flexibility of the projection of $\Psi^{(1,n)}$ into $\mathscr{P}''$. In fact, it is easy to compute the effect of any variation of the reference energy $E^{(MR)}$ on the value of σ and on the most important single determinant contributions to it, without approximations. We only need to apply a suitable set of shifts $\Delta E_1 \ldots \Delta E_k \ldots$ to the reference energy at the CIPSI step n, and store the results obtained with the altered denominators $E_I - E^{(MR,n)} + \Delta E_k$. By interpolation, the results are available as functions of the energy shift. In the selection procedure, for a given composition of the subspace $\mathscr{M}$, we evaluate the lowering of $E^{(MR,n+1)}$ with respect to $E^{(MR,n)}$ through Eq. (14); then we obtain the corresponding energy shifted $\sigma^{(n)}$ and $\sigma^{(n)}_{\mathscr{M}}$ and, by difference, $\sigma^{(n)}_{\mathscr{P}''}$. The latter, divided by the renormalization factor $(1 + C^2_{\mathscr{M}})$, yields $\sigma^{(n+1)}_{\mathscr{P}''}$.

The selection procedure based on Eqs. (17–20) and the more refined one, outlined above, have been compared in several test cases. The results are very similar, showing that the "rigid $\Psi^{(1,n)}$" approximation

is sufficiently accurate. In the next section, we shall present results based on options 1–3 only.

3 Test calculations

Test calculations on a few molecular systems of realistic size have been performed in order to show the advantages of the *aimed* selection, namely a faster and more stable convergence of the computed energy differences, and a better agreement between variational (zero-order) and perturbative (second-order) results. The processes we have considered involve single-, double- or triple-bond dissociations, double-bond twisting and electronic excitations.

$E^{(2)}$ and σ can be computed by means of the very fast diagrammatic algorithm (CIPDIAGR) [12, 13], made even more efficient by the extrapolation procedure described in the first paper of this series [11]. However, for the selection of the $\mathscr{S}$ space we still resort to the older program (CIPPI, CIPSI-Pisa), which calculates the single-determinant contributions. Therefore, we normally run a few CIPPI steps with $\mathscr{S}$ spaces of moderate size, and we perform a CIPDIAGR calculation only for the last and more demanding step. For the purpose of testing, for example, in some of the studies presented in Sect. 3, after the last CIPPI step we have performed several selections of increasingly large $\mathscr{S}$ spaces and the corresponding CIPDIAGR calculations.

3.1 Single-bond dissociation and $n \rightarrow \pi^*$ excitation: CH_3NO and $CH_3N{=}NCH_3$

We have studied the C-N bond dissociation energy of two small organic molecules, CH_3NO (nitrosomethane) and $CH_3N{=}NCH_3$ (trans-azomethane), and the $S_0 - S_1$ vertical transition energy in CH_3NO. In both cases the basis set was 6-31G*, i.e. split-valence with polarization functions [20]. The MOs were obtained by diagonalization of an average density matrix for the first two singlet states, after a CASSCF calculation (approximate natural orbitals). For CH_3NO, the CAS space involved 12 electrons and 8 MOs and for $CH_3N{=}NCH_3$, 6 electrons and 6 MOs. The ground-state structures for the two molecules and for the dissociated fragments were obtained by state-specific CASSCF geometry optimizations [11, 21].

Table 1 summarizes the results of CH_3NO. Let us first consider the square norms of the first-order wave function corrections, σ_m, which are obtained with the different selection procedures. We find that Ψ-selection and E-selection yield larger σ values at the equilibrium geometry (σ_{eq}) than at dissociation (σ_{diss}). In Fig. 2 we show a graphical comparison of the two values, σ_{eq} and σ_{diss}, for a sequence of several selection steps. For the $n \rightarrow \pi^*$ excited state S_1 we obtain σ_{exc} values not far from those of the ground state at the same geometry. The *aimed* selection brings σ_{eq}, σ_{diss} and σ_{exc} in good agreement already in the first steps, i.e. with small $\mathscr{S}$ spaces (see Table 1).

As shown in Fig. 3, the complete Eq. (22) predicts rather well the σ values which will be obtained in the next perturbation step. Neglecting the contribution of

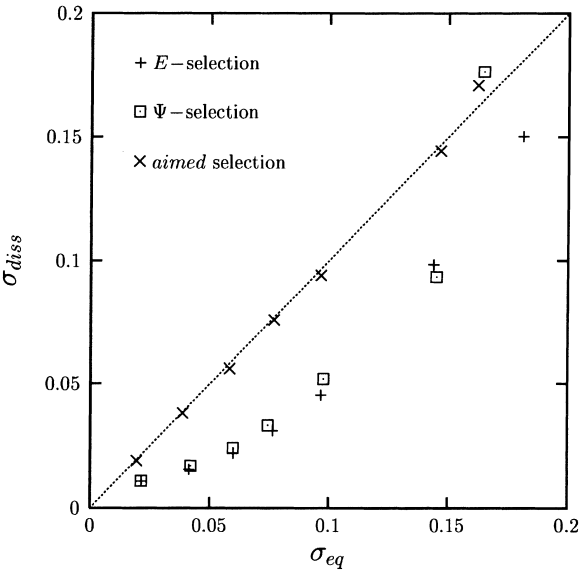


Fig. 2. Comparison of the Epstein-Nesbet (EN) square norms of the first-order correction to the wave function in CIPSI calculations for CH_3NO (σ_{eq}) and $CH_3 + NO$ (σ_{diss}), obtained at different steps of three selection procedures

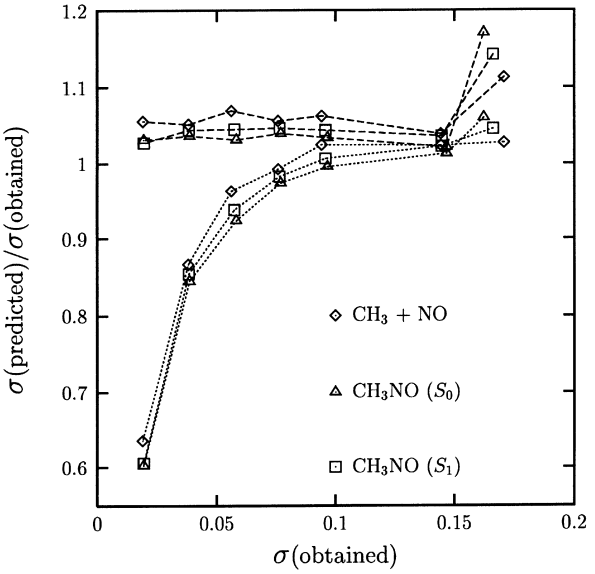


Fig. 3. Accuracy in the prediction of the EN square norms of the first-order correction to the wave function in CIPSI calculations for CH_3NO (S_0 and S_1) states and $CH_3 + NO$. *Dashed lines* contributions of both $\mathscr{P}''$ and $\mathscr{P}'''$ subspaces (option 3, see text); *dotted lines* $\mathscr{P}''$ only (option 2)

the $\mathscr{P}'''$ space leads to sizeable errors when one drastically increases the dimension of the $\mathscr{S}$ space, $N_{\mathscr{S}}$, starting with a calculation of moderate quality ("long steps" in the iterative selection). However, using this approximation we still obtain a close agreement of the three values σ_{eq}, σ_{diss} and σ_{exc}, even if they cannot be accurately programmed in advance.

These differences in the selection of the $\mathscr{S}$ spaces have important effects on the computed energies. Figures 4 and 5 show the variation of the computed dissociation

Table 1. CIPSI results for CH_3NO, C-N bond dissociation and $S_0 - S_1$ vertical excitation. The headings *eq*, *diss* and *exc* refer to the ground state equilibrium geometry, the dissociated CH_3 and NO fragments, and the $n \rightarrow \pi^*$ excited singlet, respectively. The heading *var* refers to variational, i.e. zero-order results. Dissociation and excitation energies are in kcal/mol

Ψ-selection												
η	$N_{\mathscr{S}}$			σ, EN			Dissociation energy			Excitation energy		
	eq	diss	exc	eq	diss	exc	var	EN	MPB	var	EN	MPB
–	3	3	2	0.3636	0.3204	0.3627	18.9	49.2	37.7	39.9	46.1	36.8
0.0450	15	13	24	0.1648	0.1763	0.1723	33.1	29.2	35.9	47.8	40.6	40.3
0.0180	38	251	126	0.1452	0.0935	0.1366	−22.1	49.3	36.6	36.3	46.8	41.7
0.0090	324	841	688	0.0977	0.0519	0.0949	−35.6	42.5	33.7	40.8	45.5	43.0
0.0060	807	1727	1672	0.0743	0.0331	0.0725	−40.9	36.9	31.4	41.2	44.5	43.8
0.0043	1533	2607	3270	0.0598	0.0241	0.0568	−37.7	35.5	32.5	39.4	44.7	43.3
0.0028	3423	4263	7078	0.0420	0.0169	0.0401	−22.2	33.7	33.1	40.4	44.3	43.3
0.0012	12364	9005	24654	0.0213	0.0109	0.0206	9.1	35.4	36.1	42.4	44.2	43.7
E-selection												
θ,	$N_{\mathscr{S}}$			σ, EN			Dissociation energy			Excitation energy		
millihartree	eq	diss	exc	eq	diss	exc	var	EN	MPB	var	EN	MPB
–	3	3	2	0.3636	0.3204	0.3627	18.9	49.2	37.7	39.9	46.1	36.8
2.0000	19	55	40	0.1812	0.1502	0.1614	13.6	43.1	39.7	42.8	47.7	43.9
0.7500	46	211	106	0.1440	0.0985	0.1399	−17.1	48.4	36.9	39.9	44.8	40.7
0.1800	344	1119	790	0.0967	0.0453	0.0919	−49.1	41.6	31.7	38.4	45.6	42.7
0.1000	769	1961	1640	0.0764	0.0309	0.0737	−48.9	37.4	31.2	40.2	44.8	42.7
0.0530	1558	3025	3346	0.0599	0.0220	0.0565	−40.6	34.8	31.9	39.2	44.8	43.2
0.0230	3655	4961	7604	0.0414	0.0154	0.0391	−21.9	33.6	33.4	40.4	44.4	43.4
0.0050	12450	9561	24776	0.0214	0.0108	0.0207	9.9	35.2	36.0	42.7	44.3	43.8
Aimed selection												
$\sigma^{(t)}$	$N_{\mathscr{S}}$			σ, EN			Dissociation energy			Excitation energy		
	eq	diss	exc	eq	diss	exc	var	EN	MPB	var	EN	MPB
–	3	3	2	0.3636	0.3204	0.3627	18.9	49.2	37.7	39.9	46.1	36.8
0.1900	18	21	32	0.1623	0.1707	0.1663	31.9	30.5	36.6	46.0	41.8	40.4
0.1500	36	53	84	0.1468	0.1443	0.1447	26.1	34.2	37.1	42.2	44.4	41.2
0.1000	336	235	650	0.0967	0.0941	0.0958	28.8	36.7	40.4	43.0	44.7	42.9
0.0800	726	407	1396	0.0769	0.0757	0.0764	31.1	36.4	40.4	43.0	44.5	42.9
0.0600	1646	743	3120	0.0582	0.0561	0.0574	31.4	37.2	40.7	43.3	44.7	43.7
0.0400	4102	1435	7640	0.0386	0.0380	0.0383	33.7	36.0	40.0	43.7	44.4	43.7
0.0200	14623	3651	27046	0.0194	0.0189	0.0195	36.5	37.3	40.0	44.5	44.4	44.0

energy of CH_3NO with increasing $N_{\mathscr{S}}$ (see also Table 1). The contrast between the Ψ-selection and *aimed* selection zero-order results is striking. The Ψ-selection procedure favours the dissociated fragments unduly, selecting too many determinants (hence we get $\sigma_{diss} < \sigma_{eq}$). As a result, the variational or zero-order energy of the fragments decreases faster than that of the undissociated molecule, when expanding the $\mathscr{S}$ space. For selected spaces of about $40 \div 4000$ determinants at the equilibrium geometry, the dissociation energy computed as a difference of zero-order energies is very low, or even negative. Only for the largest $\mathscr{S}$ spaces ($N_{\mathscr{S}} \simeq 12\,000$) do we again get a positive dissociation energy. The same problem is met with *E*-selection, whereas this disappears almost completely when using *aimed* selection. It is then clear that the zero-order wave functions and energies obtained with *aimed* selection are a better starting point for the perturbation calculation. Actually, the dissociation energies obtained after perturbation with all three selection procedures are acceptable.

However, the *aimed* selection results are much more stable with respect to successive expansions of $\mathscr{S}$ and show a better agreement between MPB and EN energies for smaller $\mathscr{S}$. They also get closer to the experimental [22, 23] dissociation energy, $39 \div 40.3$ kcal/mol. As may be expected from the σ values, the calculation of vertical transition energies is less problematic. Also in this case, however, the convergence of variational, MPB and EN results is faster and more stable with *aimed* selection than with Ψ-selection or *E*-selection (see Table 1).

Notice that the dimension of $\mathscr{S}$, which obviously determines the computational effort, is not a good measure of the quality of the calculation. In fact, all selection procedures lead to large differences in $N_{\mathscr{S}}$ according to state and geometry; in particular, when equal σ values are required, i.e. with *aimed* selection, the generated $\mathscr{S}$ spaces are much smaller at dissociation than at the equilibrium geometry.

The dissociation of $CH_3N{=}NCH_3$ into a methyl and a methyldiazenyl radical produces the same kind of

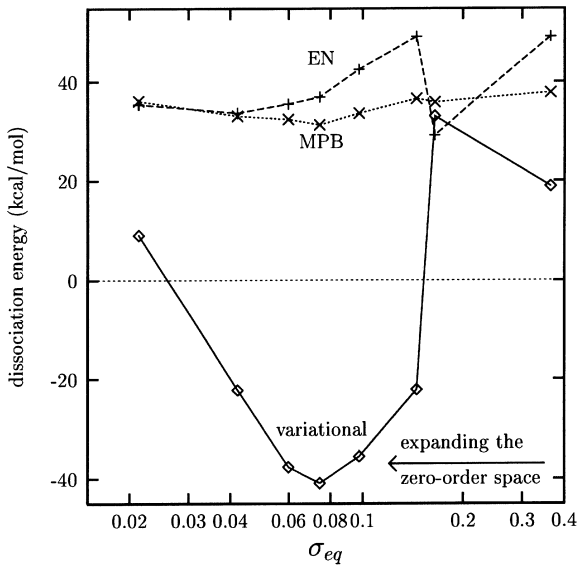


Fig. 4. Dissociation energies of CH_3NO, obtained with the Ψ-selection procedure, versus the EN square norm of the first-order correction to the wave function for the undissociated molecule (σ_{eq})

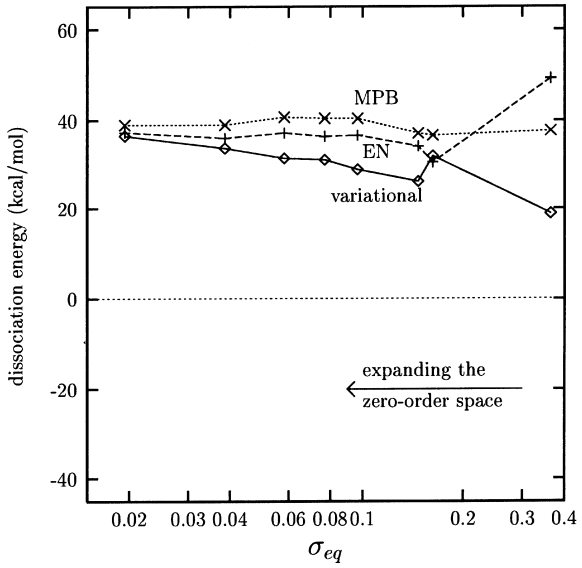


Fig. 5. As for Fig. 4, but with the *aimed* selection procedure (option 3, see text)

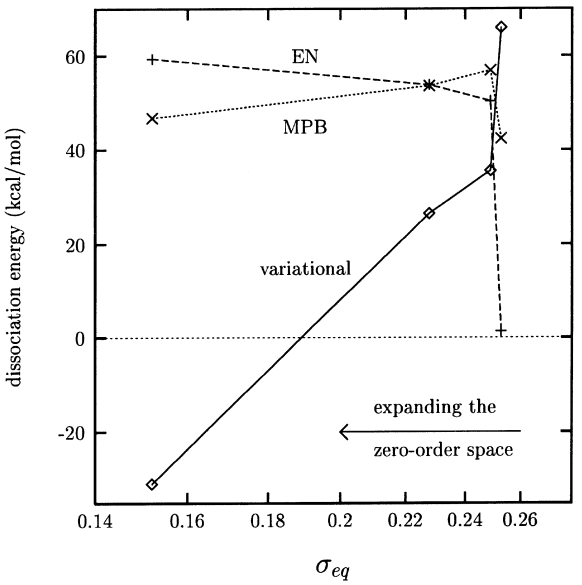


Fig. 6. Dissociation energies of $CH_3N{=}NCH_3$, obtained with the Ψ-selection procedure, versus the EN square norm of the first-order correction to the wave function for the undissociated molecule (σ_{eq})

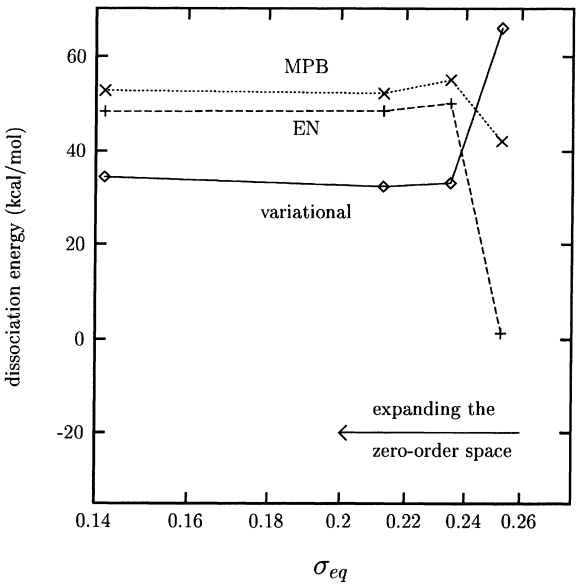


Fig. 7. As for Fig. 6, but with the *aimed* selection procedure (option 1, see text)

problems. Figures 6 and 7 show the convergence of variational, EN and MPB dissociation energies, with aimed and Ψ-selection, respectively. The decreasing trend of the variational dissociation energy with Ψ-selection, and the better stability obtained with *aimed* selection, parallel those of CH_3NO. The computed σ values for $CH_3N{=}NCH_3$, in analogy with those of CH_3NO, also show the unbalance between equilibrium geometry and dissociation which is characteristic of Ψ-selection. For the best calculations, we get $\sigma_{eq} = 0.1520$ and $\sigma_{diss} = 0.0964$, while with *aimed* selection $\sigma_{eq} = 0.1418$ and $\sigma_{diss} = 0.1421$. Because more electrons are correlated in $CH_3N{=}NCH_3$ than in CH_3NO, generally larger values of σ are obtained, and it would not be practical to increase further the size of $\mathscr{S}$ when computing potential energy surfaces [21]. This is the reason why the results shown in Figs. 6 and 7 are less close to convergence than those of CH_3NO.

3.2 Double-bond twisting and dissociation: $CH_2 = CH_2$

We have considered two ways of breaking the ethylene double bond: by torsion around the C-C axis, or by complete dissociation, leading to both CH_2 fragments in the same electronic state, either $^3B_1$ or $^1A_1$. As we have not

optimized the geometries of transition state and dissociated fragments, we cannot expect a good agreement with the experimental data. For all calculations we have adopted the same internal coordinates [7, 24] (except for the torsion angle at the twisted geometry and R_{CC} at dissociation): $R_{CC} = 1.330$ Å, $R_{CH} = 1.076$ Å, $\angle CCH = 121.7°$.

The split-valence plus polarization basis set of Saxe et al. [24] was used. The MOs were obtained by CAS calculations, with two electrons and two active MOs for the equilibrium and twisted geometries, and with four electrons and four MOs for dissociation. The corresponding CI spaces were taken as starting $\mathscr{S}$ spaces for the CIPSI calculations.

The results are summarized in Tables 2 and 3 and the energy convergence with increasing $N_{\mathscr{S}}$ is shown in Figs. 8 and 9. There are several analogies with the single-bond dissociations of the previous section. The Ψ-selection procedure generates $\mathscr{S}$ spaces which are variationally better for some geometries than for others, as can be seen by comparing the σ values and the zero-order energies. The smallest σ's and the steepest lowering of the variational energy are obtained for the two CH_2 fragments, with a good balance between the triplet and the singlet states. This is followed by the equilibrium geometry, while the largest σ's are obtained for the twisted structure. Therefore, at zero-order, the dissociation energies are too small, and the activation energy required for torsion is too large. These drawbacks disappear when we make use of *aimed* selection.

The differences are not dramatic as in the case of CH_3NO; therefore, the EN and MPB results obtained with Ψ-selection are only slightly less stable than with *aimed* selection (see Figs. 8 and 9). Also in this case we see that the number of determinants is not a good guide to assess the quality of the $\mathscr{S}$ space and to set up a balanced treatment of different geometries.

Finally, we notice that the excitation energy of CH_2, $\Delta E(^1A_1 - {}^3B_1)$, can be computed as half of the difference between the two dissociation limits. Because Ψ-selection, as well as *aimed* selection, yields a balanced treatment of the two states (see the σ values in Table 3), all

Table 2. CIPSI results for double-bond twisting in $CH_2{=}CH_2$. The headings *eq* and *tw* refer to the ground-state equilibrium geometry and the 90° twisted geometry, respectively. The heading *var* refers to variational, i.e. zero-order results. Energy differences ΔE are in kcal/mol

Ψ- selection

η	$N_{\mathscr{S}}$		σ, EN		ΔE		
	eq	tw	eq	tw	var	EN	MPB
–	2	2	0.1393	0.1339	75.8	80.7	72.6
0.0200	41	24	0.1085	0.1180	84.2	75.1	72.4
0.0150	78	84	0.0943	0.1066	86.3	73.7	72.0
0.0120	131	186	0.0829	0.0957	87.3	73.5	71.9
0.0100	214	368	0.0717	0.0817	86.3	74.7	72.8
0.0080	361	712	0.0597	0.0666	84.1	75.7	73.4
0.0050	945	1904	0.0394	0.0445	84.0	76.3	74.3
0.0040	1389	2852	0.0317	0.0368	84.8	76.1	74.4

Aimed selection

$\sigma^{(t)}$	$N_{\mathscr{S}}$		σ, EN		ΔE		
	eq	tw	eq	tw	var	EN	MPB
–	2	2	0.1393	0.0000	75.8	80.7	72.6
0.1200	29	24	0.1155	0.1180	79.2	76.9	72.1
0.1100	41	78	0.1080	0.1080	77.3	77.1	71.8
0.1000	68	160	0.0973	0.0994	77.0	75.6	70.8
0.0800	171	468	0.0783	0.0773	73.9	76.7	71.5
0.0700	261	732	0.0676	0.0664	75.0	76.8	72.4
0.0500	624	1681	0.0482	0.0473	75.2	76.3	72.8
0.0400	982	2660	0.0390	0.0383	75.1	76.2	73.2

Table 3. CIPSI results for the dissociation of $CH_2{=}CH_2$ into two CH_2 fragments, both in the $^3B_1$ or in the $^1A_1$ state. The headings *eq*, *trip* and *sing* refer to the ground-state equilibrium geometry, to two dissociated triplets and to two dissociated singlets, respectively. The heading *var* refers to variational, i.e. zero-order results. Dissociation energies ΔE(trip) and ΔE(sing) are in kcal/mol.

Ψ- selection

η	$N_{\mathscr{S}}$			σ, EN			ΔE(trip)			ΔE(sing)		
	eq	trip	sing	eq	trip	sing	var	EN	MPB	var	EN	MPB
–	2	6	20	0.1393	0.0951	0.0855	144.8	194.8	187.2	172.9	229.8	218.7
0.0110	213	495	162	0.0724	0.0345	0.0311	123.4	185.5	177.0	150.4	215.0	205.4
0.0075	396	864	272	0.0576	0.0257	0.0229	126.3	182.3	175.8	151.4	210.4	202.6
0.0050	920	1616	482	0.0398	0.0172	0.0162	134.9	180.6	176.4	162.3	208.7	203.9
0.0037	1553	2448	742	0.0296	0.0127	0.0124	142.2	179.7	176.7	170.2	207.8	204.4
0.0020	2945	4918	1589	0.0193	0.0074	0.0072	149.6	178.7	176.7	177.8	206.5	204.3

Aimed selection

$\sigma^{(t)}$	$N_{\mathscr{S}}$			σ, EN			ΔE(trip)			ΔE(sing)		
	eq	trip	sing	eq	trip	sing	var	EN	MPB	var	EN	MPB
–	2	6	20	0.1393	0.0951	0.0855	144.8	194.8	187.2	172.9	229.8	218.7
0.0800	211	34	24	0.0742	0.0775	0.0785	177.5	181.3	186.4	215.2	213.2	219.3
0.0600	376	106	39	0.0591	0.0590	0.0587	173.8	185.1	185.1	211.4	214.5	217.7
0.0400	941	362	100	0.0395	0.0401	0.0399	176.0	182.3	184.4	210.3	214.5	214.5
0.0300	1545	644	174	0.0299	0.0302	0.0303	176.6	183.2	184.0	209.3	213.2	213.7
0.0200	2762	1236	352	0.0204	0.0209	0.0205	178.6	182.4	182.7	209.9	212.5	212.5

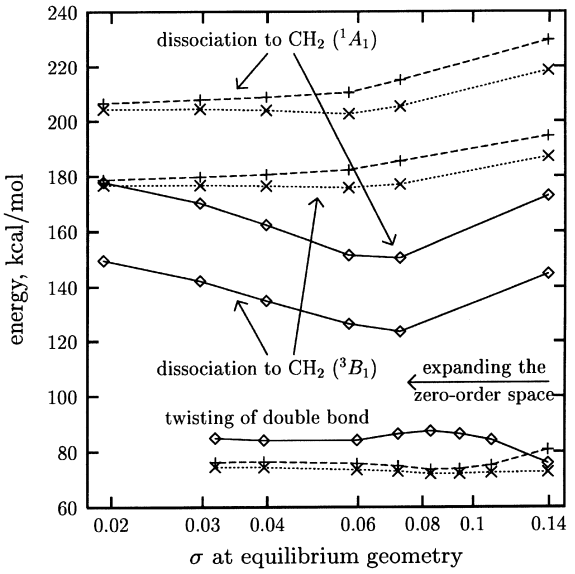


Fig. 8. Dissociation energy and barrier for the twisting of the $CH_2{=}CH_2$ double bond (unoptimized geometries, see text), versus the EN square norm of the first-order correction to the wave function using the Ψ-selection procedure

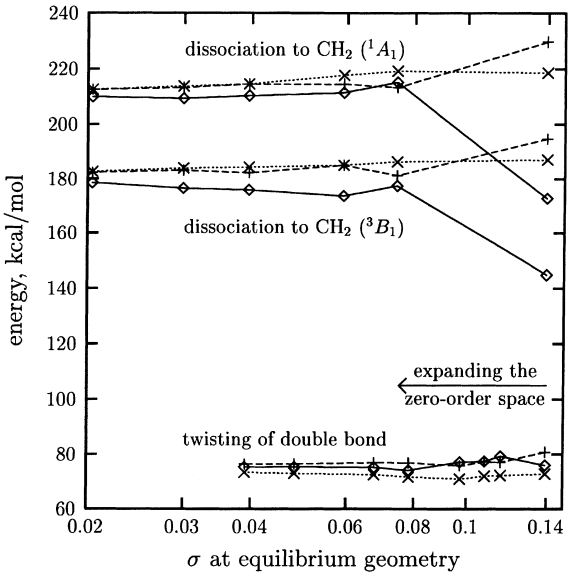


Fig. 9. As for Fig. 8, but with the *aimed* selection procedure (option 3, see text)

series of $\Delta E(^1A_1 - {}^3B_1)$ values are rapidly stabilized within 1 kcal/mol.

3.3 Triple-bond dissociation: $2N_2$

The N_2 system offers an example of triple-bond dissociation. The basis set and MOs are the same as in the first paper of this series [11]. They were chosen because full CI results [25] are available for comparison. However, the full CI was limited to six electrons, while the 1*s* and 2*s* orbitals were frozen; with this restriction, both the Ψ-selection and *aimed* selection results were very close to the full CI ones, even in the first steps of the iterative calculation.

In order to perform a more demanding test, we have included the 2*s* orbitals in the CI, and we have run two series of calculations: one for the $N_2 \rightarrow 2N$ reaction and one for a system composed of two non-interacting N_2 molecules (or four N atoms): $2N_2 \rightarrow 4N$. The degenerate orbitals were localized on each N_2 molecule or N atom as appropriate. Of course, the double N_2 energies should be twice the corresponding ones for single N_2, so in tables and figures we give the double N_2 values divided by two. The single N_2 system is still quite manageable, so that we can estimate the full CI limit from the EN results obtained for small σ (see Fig. 10). We get $E_{FCI} = -109.2765$ a.u. for N_2 ($R_{NN} = 2.10$ bohr) and $E_{FCI} = -108.9657$ a.u. for two N atoms. The latter result is confirmed by an extensive calculation on the single N atom, yielding $2E_{FCI} = -108.965688$ a.u. (the horizontal line in Fig. 10). The full CI-dissociation energy is $D_e = 195.0$ kcal/mol.

Table 4. CIPSI results for triple-bond dissociation. D_e is computed as half of the difference between the energy of four N atoms and that of two non-interacting N_2 molecules. ΔD_e errors with respect to the full CI result (195.0 kcal/mol) are given

Ψ- selection					*aimed* selection				
$N_{\mathscr{S}}$ $2N_2/4N$	σ, EN $2N_2/4N$	ΔD_e			$N_{\mathscr{S}}$ $2N_2/4N$	σ, EN $2N_2/4N$	ΔD_e		
		var	EN	MPB			var	EN	MPB
17/1	0.1821/0.1736	−36.8	−2.6	9.0	17/1	0.1821/0.1736	−36.8	−2.6	9.0
31/61	0.1311/0.0471	−79.5	2.7	−1.8	27/10	0.1396/0.1426	−30.4	−8.0	−9.1
					173/46	0.0757/0.0674	−33.3	−4.3	−7.2
229/149	0.0665/0.0142	−66.1	−4.3	−5.4					
					271/51	0.0601/0.0595	−24.8	−4.2	−5.9
381/193	0.0500/0.0089	−55.6	−5.0	−4.7					
					457/63	0.0452/0.0454	−18.9	−3.8	−5.1
615/197	0.0386/0.0087	−43.4	−4.2	−7.1					
2418/369	0.0170/0.0063	−19.5	−2.3	−3.9	2376/142	0.0173/0.0154	−9.8	−1.7	−2.8
4542/1651	0.0129/0.0041	−17.2	−1.9	−3.5	3970/168	0.0137/0.0116	−9.3	−1.6	−2.8
6266/2307	0.0114/0.0033	−16.2	−1.9	−3.2	5572/185	0.0120/0.0097	−9.0	−1.6	−2.5
9806/4803	0.0097/0.0020	−15.7	−1.8	−3.1	7728/208	0.0106/0.0083	−8.4	−1.5	−2.3
15198/6651	0.0083/0.0015	−14.4	−1.7	−2.9					
					49726/1503	0.0055/0.0043	−4.5	−0.9	−1.3

Table 4 shows the convergence of the CIPSI computed D_e to the full CI value, obtained with the Ψ-selection and *aimed* selection procedures, for the double N_2 system. These are several pairs of CIPSI calculations, belonging to either selection sequence, with approximately the same dimension of $\mathscr{S}$ and the same σ for the $2N_2$ system. These results appear in the same line of the table, in order to facilitate the comparison. We find that the *aimed* selection EN and MPB results are only slightly better than Ψ-selection ones; however, the variational energy differences show, also in this case, an unbalanced treatment of the molecule versus the separated atoms when Ψ-selection is employed. Notice that, in order to obtain approximately the same σ values, the dimension of $\mathscr{S}$ in the atomic calculations must be kept up to 40 times smaller than in the molecular ones. At dissociation, Ψ-selection yields σ values much smaller than at the equilibrium distance.

Figure 10 shows the convergence of the single and double N_2 total energies versus expansion of the $\mathscr{S}$ space by the *aimed* selection procedure. This comparison illustrates the important question of size consistency. Of course, the individual selection of determinants, no matter which criterion is employed, cannot produce fully size-consistent zero-order spaces, i.e. direct products of determinantal subspaces defined for the separated subsystems.

Ideally the square norm of the perturbative correction should increase linearly with the number of subsystems. Notice that (approximately) the same σ value may be obtained with different $\mathscr{S}$ spaces, depending on the ranking of determinants in the selection procedure. However, a meaningful comparison of the two series of calculations is obtained by plotting half of the double N_2 energies versus half of the corresponding σ values. Both variational and perturbative results show that this criterion is quite adequate, not only near to the full CI limit, but also in the first steps of the selection procedure. In other words, when the appropriate target values of σ are chosen, the *aimed* selection procedure can ensure an approximate size consistency at all levels. It may be noted that the MPB results are apparently closer to perfect size consistency than the EN ones [9], while the latter show a faster overall convergence [7].

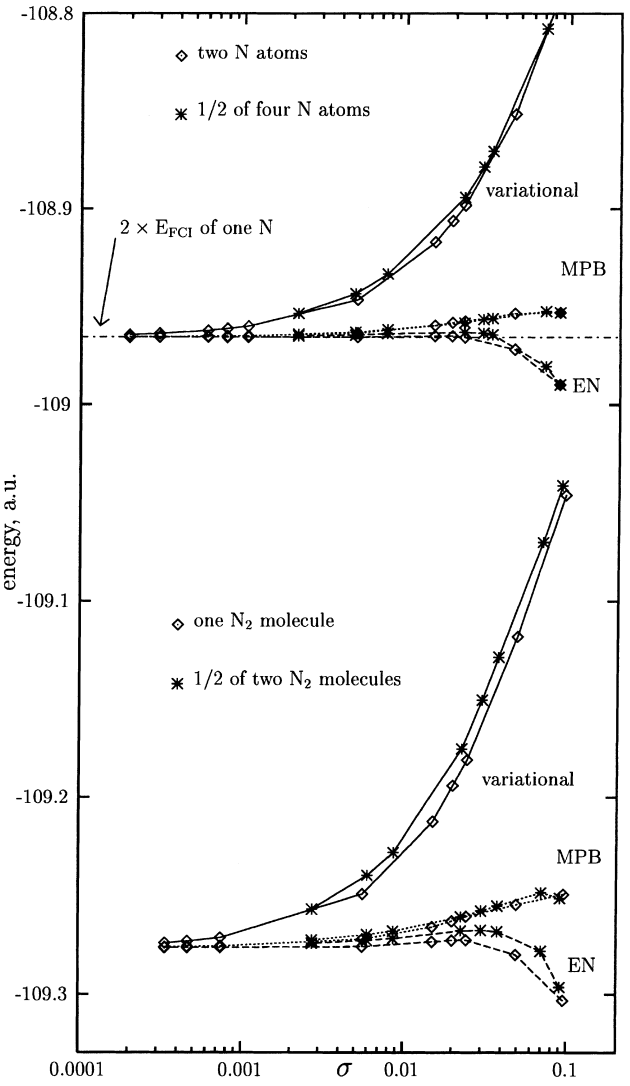


Fig. 10. Convergence of the CIPSI results to the full CI energy of N_2 and 2N. *Rhombs* Total energy of one N_2 molecule and two N atoms (*horizontal axis* EN square norm of the first-order correction to the wave function, σ). *Stars* Half of the total energy of two non-interacting N_2 molecules and of four N atoms (*horizontal axis*, $\sigma/2$, see text)

3.4 Excitation energy: Cd atom

We have considered the ground state ($5s^1S$) and the first excited state ($5p^3P$) of the cadmium atom. The computed transition energy is quite sensitive to the quality of the CI treatment, because the correlation energy of the excited state is smaller than that of the ground state [26]. We have employed a [$4s4p4d2f$] basis called E, which we partially optimized in our previous work [26]. In order to improve the CI convergence, we have used a state-specific set of orbitals obtained from CASSCF calculations (two electrons in eight orbitals, two *s* and two *p* shells).

The results obtained with Ψ-selection and *aimed* selection are shown in Figs. 11 and 12, respectively. It is again clear that the convergence of the EN and MPB transition energies is slightly more stable with *aimed* selection. On the other hand, the variational results obtained with the two selection procedures are strikingly different: Ψ-selection tends to favour the excited state, by selecting too many determinants for it, in comparison with the ground state. This bias in favour of the open-shell state bears some analogy with the behaviour found in the preceding examples of bond dissociation.

4 Conclusions

In this paper we have addressed the problem of the selection of Slater determinants belonging to the zero-order space in a multi-reference perturbation CI calculation. In particular, we have shown the importance of defining zero-order spaces of uniform quality for different electronic states and nuclear geometries. Our

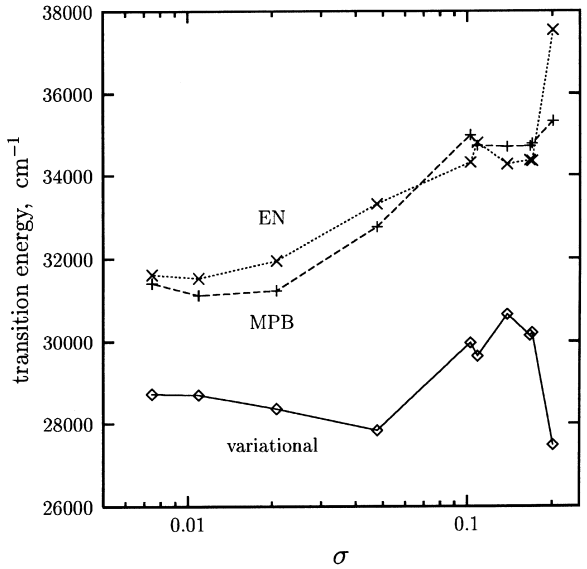
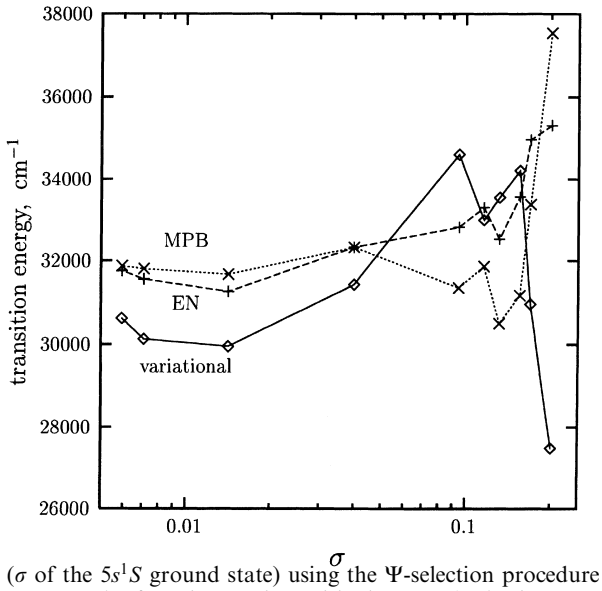


Fig. 11. $5s^1S \rightarrow 5p^3P$ excitation energy for the Cd atom, versus the EN square norm of the first-order correction to the wave function



(σ of the $5s^1S$ ground state) using the Ψ-selection procedure
Fig. 12. As for Fig. 11, but with the *aimed* selection procedure (option 3, see text)

study focusses on the CIPSI algorithm, which allows stepwise expansion of the zero-order space by individual selection of the determinants, up to convergence of the computed properties.

We have chosen the square norm of the first-order correction of the wave function, σ, as a numerical parameter expressing the quality of the zero-order space $\mathscr{S}$. We have then set up an expert system for multi-reference CI, which selects determinants according to the results obtained in a previous calculation and expands the $\mathscr{S}$ space to be used in the next one. The quality of $\mathscr{S}$ is controlled, so as to conform to a preset standard in the form of a target σ value, common to all geometries and states.

The new procedure leads to a much faster convergence of computed energy differences, with respect to the selection algorithm adopted so far. The most dramatic improvements concern the dissociation energies determined at the variational level (zero-order). In fact, we show that an individual selection not driven by a criterion of uniformity may lead to very unbalanced compositions of the $\mathscr{S}$ spaces for different nuclear geometries and, hence, to unacceptable biases in the computed energies. Adding the second-order perturbative contributions usually yields reasonable results, but the importance of a good starting point (zero-order wave functions and energies) cannot be neglected.

Notice that most multi-reference CI methods adopt more rigid definitions of the zero-order space: $\mathscr{S} \equiv$CAS, $\mathscr{S} \equiv$ single and double excitations, etc. Such restrictions in the choice of $\mathscr{S}$ are introduced to enable very efficient implementations rather than to satisfy a uniformity requirement. In fact, an $\mathscr{S}$ space which contains formally the same configurations at different geometries is not necessarily well balanced: our test calculations show that the appropriate dimensions of $\mathscr{S}$ may change by an order of magnitude according to the nuclear geometry. The issue we have brought out in this paper is therefore quite general, in that it concerns all multi-reference methods. Our approach could be profitably conjugated with recent proposals in the field of multi-reference CI [3, 27, 28], for instance in the selection of hierachically ordered subspaces, to be treated with different algorithms of increasing accuracy.

Acknowledgements. We are very grateful to J.-P. Malrieu for enlightening discussions and to P. Cattaneo for performing the $CH_3N{=}NCH_3$ calculations. This work has been supported by the Italian MURST within its 40% and 60% research funds.

References

1. Malrieu J-P, Heully J-L, Zaitsevskii A (1995) Theor Chim Acta 90: 167
2. Malrieu J-P, Durand P, Daudey J-P (1985) J Phys A 18: 809
3. Malrieu J-P, Nebot-Gil I, Sanchez-Marin J (1994) J Chem Phys 100: 1440
4. Huron B, Malrieu J-P, Rancurel P (1973) J Chem Phys 58: 5745
5. Andersson K, Malmqvist P-Å, Roos BO (1992) J Chem Phys 96: 1218
6. Spiegelmann F, Malrieu J-P (1984) J Phys B 17: 1235
7. Cimiraglia R, Persico M (1987) J Comp Chem 8: 39
8. Maynau D, Heully J-L (1993) Chem Phys Lett 211: 625
9. Malrieu J-P, Spiegelmann F (1979) Theor Chim Acta 52: 55
10. Duch W, Diercksen GHF (1994) J Chem Phys 101: 3018
11. Angeli C, Cimiraglia R, Persico M, Toniolo A (1997) Theor Chem Acc 98: 57–63
12. Cimiraglia R (1985) J Chem Phys 83: 1746
13. Cimiraglia R (1996) Int J Quantum Chem 60: 167
14. Cimiraglia R, Persico M, Tomasi J (1985) J Am Chem Soc 107: 1617
15. Persico M, Cacelli I, Ferretti A (1991) J Chem Phys 94: 5508
16. Granucci G, Persico M (1992) Chem Phys 167: 121
17. Angeli C, Cimiraglia R, Hofmann H-J (1996) Chem Phys Lett 259: 276
18. Buenker RJ, Peyerimhoff SD (1974) Theor Chim Acta 35: 33
19. Buenker RJ, Phillips RA (1989) J Mol Struct 123: 291
20. Hariharan PC, Pople JA (1973) Theor Chim Acta 28: 213

21. Cattaneo P, Persico M (1997) Chem Phys 214: 49
22. Jodkowski JT, Ratajczak E, Sillesen A, Pagsberg P (1993) Chem Phys Lett 203: 409
23. Wolff T, Wagner HG (1988) Ber Bunsenges Phys Chem 92: 678
24. Saxe P, Fox DJ, Schaefer III HF (1982) J Chem Phys 77: 5584
25. Bauschlicher CW, Langhoff SR (1987) J Chem Phys 86: 5595
26. Angeli C, Persico M (1996) Chem Phys 204: 57
27. Nebot-Gil I, Sanchez-Marin J, Malrieu J-P, Heully J-L, Maynau D (1995) J Chem Phys 103: 2576
28. Meller J, Malrieu J-P, Caballol R (1996) J Chem Phys 104: 4068