Partitioning in Multiconfiguration Perturbation Theory

P. R. Surján*, Z. Rolik, Á. Szabados, and D. Köhalmi
Department of Theoretical Chemistry, Eötvös University, 1518 Budapest 112, POB 32, Hungary

Received 4 July 2003, accepted 30 January 2004 by G. Röpke

Key words Perturbation theory, multi-configurational reference states, partitioning, optimization
PACS 31.15.-p, 31.15.Ar, 31.15.Md

A simple and effective formulation of multi-configuration perturbation theory is reviewed emphasizing the various possibilities for partitioning the Hamiltonian. We study different principles how traditional partitionings (like that of Møller and Plesset or Epstein and Nesbet) can be generalized to the multi-configurational case. Level shift parameters are introduced with the purpose of fine-tuning the partitioning. Variational conditions for optimizing level shift parameters in multi-configuration theory are investigated, both for \(N\)-particle levels and for quasiparticle (one-electron) energies.

1 Introduction

Perturbation theory (PT) is a powerful tool to account for small interactions. In spite of the long-standing existence of this theory, several important questions concerning its applicability have remained unanswered. Applicability of perturbation theory to the many-body problem depends largely upon finding the most appropriate splitting of the Hamiltonian into a zero order part and a perturbation. This problem of partitioning is enhanced if the zero order state is of multi-configurational character, due to the diverse formulations of multi-reference perturbation theory and the many ways how well known traditional partitionings, like that of Møller and Plesset[1] or of Epstein and Nesbet[2, 3], can be generalized to the multi-configurational case. In this work, we shall concentrate on the problem of partitioning from the point of view of molecular physics, in particular, we focus on describing correlation effects in many-electron systems.

The so called single-reference PT that starts from a single determinantal zero order state, has been proved to be useful only at around equilibrium geometries where the wave function is dominated by a single electronic configuration. This motivated an extensive research for developing multi-reference perturbation theories[4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 6]. In a recent paper, we have described a simple, still effective, version termed multi-configurational perturbation theory[25]. We summarize the basic features of this formalism in Section 2. It will be seen that the usefulness of this theory, just as of any other PT in general, depends crucially on the success of splitting the Hamiltonian (\(\hat{H}\)) into a zero order part (\(\hat{H}^0\)) and a remainder, or perturbation (\(\hat{W}\), i.e., the partitioning. Diverse possibilities for partitioning will be reviewed in Section 3. The guiding principle for generating different partitionings is the introduction of level shift parameters. These can fine-tune the partitioning without affecting the zero order states but have significant impact on finite-order sums and on convergence properties. Investigations on optimizing level-shift parameters in the framework of multi-configuration perturbation theory will be reported in Section 3.4. In Section 4 a few numerical examples are reported to illustrate the effect of various choices for \(\hat{H}^0\).

* Corresponding author: e-mail: surjan@chem.elte.hu, Phone: +36 1 209 0555 1632, Fax: +36 1 209 0602

adp header will be provided by the publisher
2 Multi-configurational perturbation theory

Consider a normalized, multiconfigurational reference state $|0\rangle$ and the associated projector $\hat{O} = |0\rangle \langle 0|$. The projector to the orthogonal complement space is $\hat{P} = 1 - \hat{O}$. Let us further introduce a set of determinants, denoted by $|k\rangle$, generated by applying single, double, etc. excitation operators on a Hartee-Fock-like determinant $|HF\rangle$ chosen as the 'principal' component of $|0\rangle$:

$$|0\rangle = d_0 |HF\rangle + \sum_{k=1}^{N} d_k |k\rangle.$$  \hspace{1cm} (1)

It is to be emphasized that $d_0$ should not be necessarily close to 1, merely the singular case $d_0 \sim 0$ is excluded.

We choose the set of vectors $|0\rangle$ and $|k\rangle$ ($k = 1, 2, \ldots$) as a basis in the full $M$-dimensional vector space. Since vectors $|k\rangle$ form an orthonormal basis in the $M - 1$ dimensional subspace, and the reference function does not lie in this subspace for $d_0 \neq 0$, $|0\rangle$ and $|k\rangle$-s together span the full space.

One finds immediately that $|0\rangle$ and $|k\rangle$ overlap. We construct therefore projected determinants $|k'\rangle = \hat{P}|k\rangle = (1 - \hat{O})|k\rangle$ which are orthogonal to the reference state so that $\hat{O}|k'\rangle = 0$, but overlap among themselves. The block structure of the full overlap matrix reads

$$S = \begin{bmatrix} \langle 0|0 \rangle & \langle 0|\hat{P}|l \rangle \\ \langle k|\hat{P}|0 \rangle & \langle k|\hat{P}|l \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & S \end{bmatrix}$$

where the overlap matrix $S$ of the projected excited determinants is given by the elements

$$S_{kl} = \langle k'|l' \rangle = \langle k|\hat{P}|l \rangle = \delta_{kt} - \langle k|0\rangle \langle 0|l \rangle = \delta_{kl} - d_k d_l.$$  \hspace{1cm} (2)

Here the second equality holds because $\hat{P}$ is Hermitian and idempotent.

A key point in this theory is that the inverse of matrix $S$ can be given analytically as

$$(S^{-1})_{kl} = \delta_{kl} + \frac{d_k d_l}{1 - \sum_{j=1}^{N} d_j^2} = \delta_{kl} + \frac{d_k d_l}{d_0^2},$$  \hspace{1cm} (2)

which facilitates an explicit bi-orthogonal formulation through the reciprocal vectors $\langle \tilde{k}'\rangle$:

$$\langle \tilde{k}'\rangle = \sum_{j=1}^{N} (S^{-1})_{kj} \langle j'\rangle = \langle k'\rangle - \frac{d_k}{d_0^2} \sum_{j=1}^{N} d_j \langle j'\rangle.$$  \hspace{1cm} (2)

and the definition of a non-Hermitian zero order Hamiltonian:

$$\hat{H}^0 = E_0 \hat{O} + \sum_{k=1}^{N} E_k |k'\rangle \langle \tilde{k}'|,$$  \hspace{1cm} (3)

where $E_0 = \langle 0|\hat{H}|0 \rangle$ is chosen as the energy of the reference state with $\hat{H}$ being the full Hamiltonian. The excited state energies $E_k$ are the parameters of the theory. They are in principle arbitrary quantities that define $\hat{H}^0$, i.e., the partitioning.

This definition of $\hat{H}^0$ possesses the properties

$$\hat{H}^0|0\rangle = E_0|0\rangle$$

and

$$\hat{H}^0|k'\rangle = E_k|k'\rangle.$$
One can also see that the left eigenvectors are the reciprocal projected determinants $\tilde{\langle k'|}$, and the bra reference state $\langle 0|$. The lowest order energy corrections, following standard bi-orthogonal perturbation theory, can be given as

$$E^1 = \langle 0|\hat{W}|0\rangle = 0$$

with the perturbation operator $\hat{W} = \hat{H} - \hat{H}^0$,

$$E^2 = -\langle 0|\hat{W}\hat{Q}\hat{W}|0\rangle = -\sum_{k=1} \frac{\langle 0|\hat{W}|k\rangle \langle k'|\hat{W}|0\rangle}{E_k - E_0}$$

$$E^3 = \langle 0|\hat{W}\hat{Q}(\hat{W} - \langle \hat{W} \rangle)\hat{Q}\hat{W}|0\rangle = \sum_{k,l=1} \frac{\langle 0|\hat{W}|k\rangle \langle k'|\hat{W}|l\rangle \langle l'|\hat{W}|0\rangle}{(E_k - E_0)(E_l - E_0)}$$

e tc.

3 Possibilities for choosing parameters $E_k$

In the formulae presented above, the zero order excited state energies $E_k$ are yet unspecified. The accuracy of low-order approximations as well as the convergence properties of the PT clearly depend on how we choose them. Here we mention several possibilities for choosing $E_k$-s and give numerical examples for some of these ideas. These parameters physically correspond to “excited state energies” of the unperturbed system, but a special feature of multi-configuration or multi-reference PT is that no natural choice is at our hands for $E_k$-s.

3.1 Constructing $E_k$-s from orbital energies: Møller-Plesset and Davidson-Kapuy partitionings

Following a Møller-Plesset-like philosophy one may specify $E_k$-s as

$$E_k = E_0 + \Delta \varepsilon_k \quad k = 1, 2, 3, \ldots$$

(4)

where $\Delta \varepsilon_k$ are differences of suitably chosen one-particle energies:

$$\Delta \varepsilon_k = \begin{cases} 0 & \text{for singles, doubles, etc. One-electron energies may suitably be obtained as eigenvalues (or diagonal elements) of a one-electron operator called multiconfiguration Fockian, which can be defined in two alternative forms}[26]:

F^1_{rs} = \langle 0|a^+_s \left[a_r, \hat{H}\right]_- |0\rangle \quad (5)

and

F^2_{rs} = \langle 0| a^+_s \left[a_r, \hat{H}\right]_+ |0\rangle \quad (6)

where $a^+_i (a_i)$ creates (annihilates) an electron on spin-orbital $i$, and the subscripts $\pm$ indicate commutator ($-$) or anticommutator ($+$). The two expressions are not equivalent in general, since expanding the (anti)commutators we get

$$F^1_{rs} = \sum_t h_{rt} P_{ts} + \sum_{tuv} [rt|uv] \Gamma_{uv,st} \quad (7)$$

where
and

\[ F_{rs}^2 = h_{rs} + \sum_{uv} P_{uv} [ru \mid sv] \]  

(8)

with \( P \) and \( \Gamma \) denoting the first- and second-order density matrices, and with the usual notations for the 
one-electron integrals \( h_{rs} \) and the (antisymmetrized) two-electron integrals in the \([12\mid 12] \) convention.

Both \( F^1 \) and \( F^2 \) have their own significance. Matrix \( F^1 \), called also the (negative) Koopmans matrix[26],
is non-symmetric in general, its antisymmetric part giving rise to orbital gradients in MC-SCF theory[26, 27]. Matrix \( F^2 \) is always symmetric and can be considered as the generalization of the usual Fockian built up by the actual (correlated) density matrix \( \hat{P} \). The eigenvalues of both matrices can be interpreted as approximate ionization potentials[26].

While the usual Fockian is diagonal in the MO basis set in HF theory, neither \( F^1 \) nor \( F^2 \) are diagonal
in the multi-configurational case. These two matrices therefore provide us with four possible definitions
for effective one-particle energies in the multi-configurational framework: we can use either the diagonal
elements or the eigenvalues of both matrices. If using the eigenvalues, we may speak about a generalized
Møller-Plesset (MP) partitioning. Defining just the diagonal elements as one-particle energies, one arrives
at a quite different partitioning. This, in connection with using non-canonical MOs in MBPT, was proposed
by Davidson[28, 29] and extensively used by Kapuy[30, 31, 32] (for a review, see [33]). To distinguish
this second possibility from MP, we shall refer to it as the generalized Davidson-Kapuy (DK) partitioning.
Although the DK partitioning was reported to provide numerical results inferior to those of MP in the
single-reference case[34], it appears to be important to investigate whether this conclusion holds for multi-
configurational reference states as well. The advantage of the DK partitioning from the computational
point of view in the MC case is obvious. Namely, if using the eigenvalues of the Fockian (7) or (8), one
has either to rewrite the reference state to the one-particle basis which diagonalizes these matrices, or deal
with a non-diagonal \( \hat{H}_0 \), both of which is quite impractical.

3.2 Optimized quasiparticle partitioning

The arbitrariness of orbital energies to construct energy denominators raises the question whether \( \varepsilon_i \)-s could be determined in an optimal way. The answer to this question is affirmative, and has so far been elaborated in connection with single-reference theory[35]. The basic idea is to make the third order energy stationary with respect to the variation of \( \varepsilon_i \)-s. Since \([E^0 + E^1] \) is invariant to such a level shift[35], this
leads to the variational condition:

\[ \frac{\partial [E^2 + E^3]}{\partial \varepsilon_i} = 0 \]  

(9)

for all (occupied and virtual) \( i \)-s. Solution of the above nonlinear equation for \( \varepsilon_i \)-s is possible numerically;
we have used here the BFGS technique with analytical gradients\(^1\).

3.3 Generalized Epstein-Nesbet partitioning

Instead of following a MP-type philosophy, one may also specify \( E_k \) in the Epstein-Nesbet spirit[2, 3].
Then, an alternative to Eq.(4) is

\[ E_k = \langle k \mid \hat{H} \mid k \rangle, \]  

(10)

which can be characterized as the generalized EN partitioning, also frequently used in MR-PT[19, 36].

\(^1\) in the original work[35] numerical first derivatives were applied thus the data presented there may suffer from numerical errors
3.4 Optimized partitioning in MCPT

Finally, if following neither MP nor EN concepts, the \( E_k \) energies themselves can be handled as variational parameters. The functional to be varied may be the same as in Section 3.2, i.e., the sum of PT energy corrections up to the third order. We call this ‘optimal’ (OPT) partitioning\[37, 38, 39\]. Several other ways of optimizing the partitioning have been considered previously\[18, 40, 41\]. Recently, Witek et al.\[42, 43\] investigated the optimal partitioning in the multi-reference framework.

In the single-reference\[38\], as well as the multi-reference version of optimized partitioning\[42, 43\] there are two equivalent ways to derive the working equations. One either solves the equations similar to (9)

\[
\frac{\partial (E^2 + E^3)}{\partial E_k} = 0,
\]

or forces the second order wave function correction to vanish:

\[
|\psi_k^{[2]} [E_l]| = 0.
\]

This latter condition, of course, involves that \( E^3 = 0 \), which also holds in the repartition by scaling the Hamiltonian studied by Feenberg, Goldhammer and Amos\[44, 45, 46\].

Both conditions (11) and (12) lead to the same system of linear equations for \( E_k \)-s in the single reference theory. In the MCPT version we use, however, these two conditions result in different system of equations due to the bi-orthogonal formulation (cf. Section 2). As we have checked, the differences are usually minor. The calculations reported below were obtained by using condition (12).

In the multi-reference PT by Witek et al.\[42, 43\], the second order energy in the optimal partitioning was shown to be equivalent to the multi-reference linearized coupled cluster energy\[47\]. This is a generalization of the resummation theorem\[38\] to the multi-reference case. In the present formulation of MCPT this equivalence does not exist, although the optimized second order results clearly correspond to a resummation of certain diagrams even in this case.

4 Numerical results

4.1 Role of full optimization

We present some results on the Be atom and the \( \text{H}_2\text{O} \) and \( \text{F}_2 \) molecules to show the performance of MCPT and the role of optimization of energy denominators.

Second and third order energies of the Be atom in a polarized triple-\( \zeta \) basis set in various partitionings are collected in Table 1. The reference state is an antisymmetrized product of strongly orthogonal geminals (APSG) \[48\], with the core geminal frozen and a valence geminal extended to the subspace of 2s and 2p orbitals. This reference state exhibits only a minor energy improvement from the HF approximation (the HF energy is -14.57187 a.u.), though it contains the principal \( (1s)^2(2s)^2 \) configuration and the three low-lying \( (1s)^2(2p)^2 \) excited configurations. The results of the EN partitioning are not too good: the second order represents some improvement but the third order energy is worse than the second order one. However, the DK partitioning (i.e., using the diagonals of the generalized Fockian (8) as one-particle energies) shows a good second order energy and a pretty accurate third order estimation. The results in the optimized partitioning have an error merely of 1 mH; here \( E^2 = E^3 \), since we have used Eq.(12) to find the optimal excited energies.

Table 2 presents results for the water molecule. One can observe that the behaviour of the Epstein-Nesbet partitioning is quite erratic in this case. EN2 overshoots; EN3 tends to correct this but the resulting energy becomes too high. The Davidson-Kapuy partitioning provides reasonable estimates, although the third order correction appears to be too small. However, optimizing the partitioning by Eq.(12) we obtain a highly accurate second order energy, its error being again in the order of a mH.
Table 1 MCPT energy results [a.u.] for the Be atom in various partitionings. APSG reference state, 6-311G** basis set.

<table>
<thead>
<tr>
<th>order of MCPT</th>
<th>partitioning</th>
<th>Epsten-Nesbet (EN)</th>
<th>Davidson-Kapuy (DK)</th>
<th>Optimized</th>
<th>partitioning-independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>-14.60565</td>
<td>-14.62993</td>
<td>-14.63458</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-14.60565</td>
<td>-14.62993</td>
<td>-14.63458</td>
<td></td>
</tr>
<tr>
<td>∞ (FCI)</td>
<td></td>
<td>-14.63337</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 The same as in Table 1, for the water molecule with GVB reference state, 6-31G basis.

<table>
<thead>
<tr>
<th>order of MCPT</th>
<th>partitioning</th>
<th>Epsten-Nesbet (EN)</th>
<th>Davidson-Kapuy (DK)</th>
<th>Optimized</th>
<th>partitioning-independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>-76.137702</td>
<td>-76.110592</td>
<td>-76.120697</td>
<td>-76.0279471</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>-76.103427</td>
<td>-76.112469</td>
<td>-76.120697</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-76.103427</td>
<td>-76.112469</td>
<td>-76.120697</td>
<td></td>
</tr>
<tr>
<td>∞ (FCI)</td>
<td></td>
<td>-76.120810</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results for the F$_2$ molecule computed in triple-ζ polarized basis (frozen core) are shown in Table 3. Since we were not able to run full CI for this system, the results of some other standard quantum chemical methods are reported for comparison. Two of them, CCSD and CCSD(T), are single-reference methods, while with the multi-reference averaged quartic coupled-cluster (MR-AQCC) calculations[49, 50] we used a full valence CAS reference function. The reference function for the MCPT calculations was a simple 2x2 CAS function which is equivalent to a GVB in which only one pair is correlated. For this system, the behaviour of the EN partitioning is similar to the previous cases: EN2 is erratically deep, EN3 is too high. The DK partitioning seems to do a fairly good job for this system, while the results of the optimized partitioning are again appear to be quite accurate, lying close to the CCSD(T) numbers.

Table 3 GVB-MCPT energies of the F$_2$ molecule in comparison to other methods at two different bond lengths R (6-311G** basis).

<table>
<thead>
<tr>
<th>method</th>
<th>R=1.2 Å</th>
<th>R=1.6 Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>GVB</td>
<td>-198.738251</td>
<td>-198.790696</td>
</tr>
<tr>
<td>MCPT-EN2</td>
<td>-199.187718</td>
<td>-199.202279</td>
</tr>
<tr>
<td>MCPT-EN3</td>
<td>-199.099669</td>
<td>-199.147977</td>
</tr>
<tr>
<td>MCPT-DK2</td>
<td>-199.112614</td>
<td>-199.142027</td>
</tr>
<tr>
<td>MCPT-DK3</td>
<td>-199.116706</td>
<td>-199.155061</td>
</tr>
<tr>
<td>MCPT-OPT2</td>
<td>-199.128580</td>
<td>-199.171839</td>
</tr>
<tr>
<td>CCSD</td>
<td>-199.120593</td>
<td>-199.150046</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>-199.129049</td>
<td>-199.167933</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>-199.122600</td>
<td>-199.163215</td>
</tr>
</tbody>
</table>

4.2 Role of optimized quasiparticle energies

To illustrate the effect of optimizing the quasiparticle energies, we have performed a few control calculations of which the results on the H$_2$ molecule are plotted in Fig.1. The figure compares the convergence of MCPT energies obtained with canonical quasiparticle energies (MP-MCPT), with optimized quasiparticle energies (ε-OPT-MCPT) and with the full-CI limit.

It is our general experience that, even if the MP-MCPT values converge monotonically, the ε-OPT-MCPT energies tend to show an oscillatory convergence. This has the advantage that it offers a better
possibility for extrapolation. For instance, taking simply the average energies of even and odd orders, one obtains a smooth, fast converging series denoted by DAMPED-OPT-MCPT in the figure.

As compared to the results of full optimization reported in the Tables, Figure 1 indicates that the $\epsilon$-optimization has on a moderate effect. Performing test calculations in larger systems we have found that the energy gain due to optimizing the quasiparticle energies tends to decrease with increasing number of electrons and/or basis orbitals. This is understandable since the number of optimizable parameters is the same as the number of basis functions ($M$) in the case of $\epsilon$-optimization, while the number of optimized parameters grows as $\binom{M}{N}$ in the fully optimized case, $N$ being the number of electrons in the molecule. Therefore, in spite of the computational advantages of $\epsilon$-optimization, it cannot be suggested for general use. The results obtained in the fully optimized partitioning, however, a rather accurate; here the problem is that solution of Eqs.(11) or (12) can be very time-consuming. Finding an efficient way to solve these equations for larger molecules, and finding acceptable approximations to these solutions is a challenging task; work in this line is in progress.
Acknowledgements  The authors are indebted to Péter Szalay (Budapest, Hungary) for his kind help in performing the MR-AQCC computations by the COLUMBUS program system[51, 52, 53, 54]. The APSG (GVB) reference functions were generated by the Budapest version [55] of the MUNGAUSS program package[56]. This work has been partly supported by the grants OTKA T-35094-43685-M45294-D45983 and TET-211/6/02. We are indebted to the project NIIF for computational facilities.
References