Multi-reference averaged quadratic coupled-cluster method: a size-extensive modification of multi-reference CI

Péter G. Szalay

Eötvös Loránd University, Department of Theoretical Chemistry, P.O. Box 32, 1518 Budapest, Hungary

and

Rodney J. Bartlett

Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, FL 32611, USA

Received 27 July 1993; in final form 7 September 1993

An extension of the earlier multi-reference linearized coupled-cluster method to include quadratic EPV terms in an averaged way is presented. The resulting functional is conceptually similar to the averaged coupled pair functional but it offers superior performance particularly with small reference spaces. This is demonstrated on the ozone molecule using a two electron–two orbital GVB-type reference function.

1. Introduction

The development of ab initio quantum chemical methods in the last few years demonstrates the interplay between CI methods and many-body methods like many-body perturbation theory and coupled-cluster theory. On one hand it is now appreciated that the (size) extensivity error [1] in CI can be large and therefore correction is required. As extensivity forms the rationale for all methods entitled to be called “many-body” methods, the derivation of the corrections to CI are mostly based upon many-body arguments [2–7]. On the other hand the usage of multi-reference states as a reference in many-body calculations is also important in some cases when the single reference function is not a qualitatively correct description of the system (see, e.g., refs. [8–10]). Although many multi-reference extensions of CC/MBPT have been presented [11–15], multi-reference (MR-CI) methods are frequently easier to employ because MR-CI is fundamentally a single state approach while most multi-reference coupled-cluster (MR-CC) methods are inherently effective Hamiltonian, multi-state theories [11–15].

This Letter deals with (size) extensivity corrected CI-type methods. Such a correction can be made in two ways. First, corrections can be made a posteriori following a CI calculation. These are the so-called Davidson-type corrections [2–7]. Another group of methods corrects the CI functional itself [7–9,16–19]. This type of a priori correction has the advantage that analytical energy gradients can be calculated and in this way (nearly) extensive properties can be obtained.

The rigorous connection between MR-CC and MR-CI is difficult, but much of the approximate correspondence can be observed via linearization of the Hilbert space MR-CC formation. Such an approach was considered by Paldus [7], while Bartlett and co-workers [8,9], further developed the linearized equations and added a transformation of the effective Hamiltonian matrix to introduce an MCSCF reference CC. In this way, a single state was approximately decoupled from all other states to provide a single but multi-configurational reference, linearized coupled-cluster method (MR-LCCM).

MR-LCCM was implemented into a MR-CI program and its performance tested on potential energy

0009-2614/93/$ 06.00 © 1993 Elsevier Science Publishers B.V. All rights reserved.
curves for N₂, F₂, [9], Be inserting into H₂ and bond stretching in H₂O [8]. MR-LCCM, augmented by inclusion of the other orthogonal combinations in the reference space, has been considered by Gdanitz and Ahlrichs [16] under the name MR-CEPA(0) and by Hoffman and Simons [17] under the name UCEPA (Unitary CEPA) (note, these later two methods are identical). The most successful variant, due to Gdanitz and Ahlrichs [16], is the averaged coupled pair functional (MR-ACPF) method which, in addition considers some EPV contributions to MR-LCCM. This has been implemented into many MR-CI programs and routinely used as an approximately extensive alternative to MR-CI. A more detailed comparison of these and other related methods [18,19] can be found in our subsequent paper [20].

Interestingly (to our knowledge) no example is reported where MR-ACPF fails badly, although it does regularly with small reference spaces. The absence of reported failures occurs partly because the original prescription suggests using as large a reference space as necessary, and increasing the size of the reference space will certainly remove all problems in the limit.

Increasing the reference space has several disadvantages, however. Besides the obvious increase in cost, the selection of the reference space can become a purely numerical experiment, and, of course, chemical insight will frequently be compromised. Since CC methods are extremely powerful, and they can frequently overcome even a very poor single reference description [21,22], it is not requisite to have a large reference space in CC. In fact, we prefer the minimum qualitatively correct reference, such as the two-determinant (TD)-CCSD [23] for open-shell singlets and automatic, well defined methods such as Pulay's scheme [24] to construct simple multi-reference functions as used for the UNO-CAS method [25]. However, as will be shown on the ozone example, the reference space predicted with this scheme [24] is not adequate to obtain good results from MR-ACPF. Also, it appears that MR-CI itself, frequently does not require as large a reference space as would be necessary for MR-ACPF. Often it is not even possible to increase the reference space as required by these methods. For example, for ozone we are not able to do multi-reference calculations with a reference space where all the problems of the MR-ACPF mentioned later would disappear, yet, calculations with two reference functions are widely applicable.

The main motivation of this work, therefore, is to derive a generalization of the MR-LCCM or MR-ACPF method which will work with smaller reference spaces. Analogous to CEPA approximations to single reference CC [26], we will start from a coupled-cluster Ansatz and approximate only those quadratic (EPV) terms which are necessary to get nearly extensive results. This multi-reference averaged quadratic CC (MR-AQCC) method thereby avoids the complications inherent to a fully quadratic multi-reference CC approach, retaining the simplicity of the MR-LCCM or MR-ACPF scheme. This is a step toward our ongoing effort to develop a fully quadratic, state-selective MR-QCC method.

The last section presents results demonstrating that the new MR-AQCC method is superior to MR-ACPF or MR-CISD for the ozone molecule.

2. Theory

Let us consider a reference space of nₙ reference functions \{Φᵣ, r=1...nᵣ\}. Our zeroth-order reference is the solution of the Schrödinger equation in this space: \(Ψ₀ = \sum cᵣΦᵣ\). The correlated wavefunction is formally written as [27]

\[Ψ = \sum_r \exp(T^r) P_r Ψ₀ + Ψ_p\],

(1)

\(P_r = |Φᵣ⟩⟨Φᵣ|\) and \(\exp(T^r)\) is defined relative to each \(Φᵣ\) to give the \(T^r\) cluster operator. \(Ψ_p\) corresponds to the part of the wavefunction composed of the other orthogonal combinations of \{Φᵣ\} in the \(P\) spaces. This exponential parameterization ensures extensivity if \(T^r\) is additively separable and if the reference function has this property. The single and doubly excited part of the wavefunction will be denoted by

\[Ψ₀ = \sum_r \left( \sum_{ia} cᵣ^iaΨᵣ^a(r) + \sum_{ijab} cᵣ^ibΨᵣ^b(r) \right)\]

(2)

with \(Ψ₀^a(r)\) and \(Ψ₀^b(r)\) denoting single and double substitution with respect to the reference space function \(Φᵣ\). The full correction to \(Ψ₀\) is \(Ψ_p = Ψ_p = Ψ₀\).

First, we discuss the single reference case where the operator in eq. (1) reduces to a single \(\exp(T)\). Consider the expectation value for the energy with
the exponentially parameterized wavefunction [28,29],
\[
F(\Delta E) = \frac{\langle e^T \Phi_0 | \hat{H} - E_0 | e^T \Phi_0 \rangle}{\langle \Phi_0 | e^T e^T | \Phi_0 \rangle} = \langle e^T \Phi_0 | \hat{H} - E_0 | e^T \Phi_0 \rangle_c
\]
where subscript \(c\) means that only connected terms are included. This expression is infinite due to EPV (exclusion principle violating) terms [28–30]. Now we assume that we can find a reference function \(\Phi_0\) for which the connected higher excitations can be neglected. Furthermore we select out the EPV terms and write them in a disconnected way [30] to obtain
\[
F(\Delta E) = \langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle - \sum_{ijab} (c_{ij}^{ab})^2 \langle \Psi_{ij}^{ab} | \Psi_{ij}^{ab} \rangle \\
\times \left( \sum_{klcd} \langle \Psi_0 | \hat{H} - E_0 | \Psi_{klcd}^{cd} \rangle c_{klcd}^{cd} \right). \tag{3}
\]
Here \(\sum_{klcd}\) means that at least one of the indices \(klcd\) has to match one of the letters \(ijab\). Using the CISD energy \(\Delta E_{\text{SD}} = \langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle / (1 + \langle \Psi_c | \Psi_c \rangle)\) one can obtain
\[
F(\Delta E) = \Delta E_{\text{SD}} + \langle \Psi_c | \Psi_c \rangle \Delta E_{\text{SD}} - \sum_{ijab} (c_{ij}^{ab})^2 \langle \Psi_{ij}^{ab} | \Psi_{ij}^{ab} \rangle \\
\times \left( \sum_{klcd} \langle \Psi_0 | \hat{H} - E_0 | \Psi_{klcd}^{cd} \rangle c_{klcd}^{cd} \right) = \Delta E_{\text{SD}} + \sum_{ijab} (c_{ij}^{ab})^2 \langle \Psi_{ij}^{ab} | \Psi_{ij}^{ab} \rangle \\
\times \sum_{klcd} \langle \Psi_0 | \hat{H} - E_0 | \Psi_{klcd}^{cd} \rangle c_{klcd}^{cd}. \tag{4}
\]
\(\sum\) means now that one of the \(klcd\) can match any of the \(ijab\). We introduce the usual pair energy,
\[
\epsilon_{ij} = \sum_{ab} \langle \Psi_0 | \hat{H} - E_0 | \Psi_{ij}^{ab} \rangle c_{ij}^{ab}.
\]
With this the second term of eq. (4) can be approximated as
\[
\sum_{ij} (c_{ij}^{ab})^2 \sum_{kl,ijkl} \epsilon_{kl},
\]
with \(c_{ij}^{ab} = \sum_{ab} c_{ij}^{ab}\). This approximation neglects EPV terms due to virtual labels.

Now we follow the arguments of Meissner [6] used to discuss a Davidson-type correction for CI. Recognizing that there are \(\binom{n_e}{2}\) different pair energies, and assuming that all contribute with the same amount to the correlation energy, we can define an averaged pair energy as
\[
\bar{\epsilon}_{ij} = \Delta E_{\text{SD}} \left[ \frac{n_e}{2} - \binom{n_e}{2} \right].
\]
With this the second term of eq. (4) becomes
\[
\sum_{ij} (c_{ij}^{ab})^2 \sum_{kl,ijkl} \overline{\epsilon}_{kl} = \sum_{ij} c_{ij}^{ab} \left( \frac{n_e - 2}{2} \right) \Delta E_{\text{SD}} \left[ \frac{n_e}{2} \right] = \langle \Psi_c | \Psi_c \rangle \left( \frac{n_e - 2}{2} \right) \frac{(n_e - 3)}{n_e(n_e - 1)} \Delta E_{\text{SD}}.
\]
Notice this vanishes for 2 electrons, since CISD would be exact and there are no unlinked diagram contributions to remove. Similarly, it should vanish for 3 electrons because there are no EPV contributions [6]. Therefore the correlations energy is
\[
F(\Delta E) \approx \Delta E_{\text{SD}} + \langle \Psi_c | \Psi_c \rangle \left( \frac{n_e - 2}{2} \right) \frac{(n_e - 3)}{n_e(n_e - 1)} \Delta E_{\text{SD}}.
\]
Neglecting terms which are higher than fourth order in \(\Psi_c\) [16] we get the final form of our functional
\[
F_{\text{AQCC}}(\Delta E) = \frac{\langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle}{1 + (1 - (n_e - 3)(n_e - 2)/[n_e(n_e - 1)])} \langle \Psi_c | \Psi_c \rangle \left( \frac{n_e - 2}{2} \right) \frac{(n_e - 3)}{n_e(n_e - 1)} \Delta E_{\text{SD}}. \tag{5}
\]
This is the functional form of Meissner's modified Davidson correction [6].

Although in the above derivation we used a single reference state, generalization to multi-reference states is possible. \(\Psi_0\) now plays a role, and we assume it contributes with unit value as argued in ref. [16]. Furthermore, one needs to neglect some linear terms the same way as was done for MR-LCCM [8,9]. As argued in that paper, those terms are probably less important than the neglect of the quadratic terms which are now approximated in our new method. Therefore, we choose to use eq. (5) with \(\Psi_0\) now representing a multi-reference function to define a new multi-reference averaged quadratic coupled-cluster (MR-AQCC) method. The name arises.
from the fact that compared to the linearized methods, we include EPV terms originating in the quadratic part of the exact functional but in an average way. The current method is viewed as a practical and potentially useful approximation in the evolution toward a fully quadratic, MR-QCC method.

Comparing with other methods we can easily see that MR-LCCM [8,9] neglects all EPV terms, i.e. the second term in eq. (3). In this sense the new method is an extension of it. The original MR-LCCM did not include \( \Psi_p \) in \( \Psi_c \). MR-CEPA(0) [16] (and the exactly equivalent UCEPA [17]) include this component. For an already approximate method, this slight constraint was recommended as it facilitated obtaining improved numerical results in MR-LCCM without the convergence difficulties encountered in MR-CEPA(0). On the other hand, using configuration-based methods, there is no reason why one should not include them and in this way account for the relaxation of the reference function. Therefore, we define MR-AQCC with the orthogonal complement of the reference function in the reference space. Whereas MR-CEPA(0) has a much poorer convergence behavior than MR-LCCM due to the presence of this orthogonal complement, for our MR-AQCC method, any such convergence problems seem to be alleviated.

Of course, the MR-AQCC functional is very similar to ACPF, as it can be derived simply by replacing the \( g=2/n_e \) factor [16] by \( g=1-(n_e-3)(n_e-2)/(n_e(n_e-1)) \). However, the latter seems to offer a superior estimate for the number of electron pairs.

In their derivation of the ACPF method Gdanitz and Ahlrichs [16] divided the electrons of the molecule into \( \frac{1}{2}n_e \) non-interacting electron pairs. Therefore the second term of eq. (4) can be written as

\[
\sum_m (c_m^*)^2 \sum_{m \neq m'} \epsilon_{m'}^*,
\]

where \( m \) labels the electron pairs. Assuming that the electron pairs are identical we can approximate the pair energy,

\[
\epsilon_{m'}^* = \frac{\Delta E_{SC}}{\frac{1}{2}n_e}
\]

and therefore the second term of eq. (4) becomes

\[
\sum_m (c_m^*)^2 \sum_{m \neq m'} \epsilon_{m'}^* = \sum_m (c_m^*)^2 (\frac{1}{2}n_e - 1) \frac{\Delta E_{SD}}{\frac{1}{2}n_e}
\]

\[
= \langle \Psi_c | \Psi_c \rangle (1 - 2/n_e) \Delta E_{SD}.
\]

Substitution into the energy functional,

\[
F(\Delta E) = \Delta E_{SD} + \langle \Psi_c | \Psi_c \rangle (1 - 2/n_e) \Delta E_{SD}
\]

we have the Pople et al. factor \( (1 - 2/n_e) \) [4]. Neglecting higher-order terms [16] we arrive at the ACPF functional,

\[
F_{ACPF}(\Delta E) = \frac{\langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle}{1 + 2/n_e \langle \Psi_c | \Psi_c \rangle}.
\]

This above comparison shows the basic difference of MR-AQCC and MR-ACPF, with MR-ACPF approximating the EPV terms through the non-interacting electron pairs while MR-AQCC includes all EPV terms in an averaged way. This means that MR-ACPF accounts for a smaller portion of this contribution and is, therefore, closer to the linearized versions. This fact also explains why MR-ACPF tends to overestimate higher excitations for small reference spaces (see section 3).

Gdanitz and Ahlrichs justify their choice of the \( g \)-factor by requiring that the functional be exact for non-interacting electron pairs. Our correction does not satisfy this condition but it rather vanishes for three electrons where EPV contributions do not exist [6]. Furthermore, we distribute the correlation energy among a larger number of "electron pairs" (all electrons are involved in several "pairs") and it is therefore potentially more justified than distributing the correlation energy solely among the non-interacting pairs. The difference between the two methods (and between MR-LCCM) will vanish when the correlation norm \( \langle \Phi_e | \Phi_e \rangle \) vanishes, i.e. when the reference function becomes a better approximation to the total wavefunction. Therefore we can assume a different behavior for the two methods especially for small reference spaces. Certainly, any such comparison ultimately depends upon numerical tests to decide which method is superior (see section 3). Although guided by EPV arguments, we can equally well consider that the averaging procedure we use might be shown to better account for all quadratic terms in an MR-QCC method.

MR-AQCC may be readily implemented into any
MR-CI program, and, in particular, it is trivial to modify a MR-ACPF program. No additional effort is needed to derive analytical energy derivatives either. Calculation of the effective density and the gradient itself can be performed as described in ref. [31] and other properties can be calculated as the energy derivatives [32]. Implementation into the COLUMBUS MR-CI program system [33] readily provided all those features.

3. Application to ozone

Calculations on $O_3$ are performed with the same DZP basis as in the comparative CCSDT study by Watts et al. [34]. The basis consists of the sp set of Huzinaga and Dunning [35,36] augmented with our usual correlation optimized polarization functions [37]. We retain all six Cartesian components of d functions. Because of the frozen core orbitals $n_e = 18$ was used in both ACPF and AQCC functionals. All singly and doubly substituted configurations out of the reference functions (see below) are included in the correlated wavefunctions with the constraint that the three lowest occupied orbitals ($1a_1, 2a_1, 1b_2$) be kept doubly occupied in all configurations. No frozen virtuals are used. Orbitals are optimized at the MCSCF level using the configurations in the reference space. All calculations are performed using an interfaced version of the ACES II [38] and COLUMBUS [33] program systems.

A qualitatively correct description of ozone requires the two configurations:

$$\phi_1 = \ldots 6a_1^2 4b_2^2 1b_1^2 1a_1^2 ,$$
$$\phi_2 = \ldots 6a_1^2 4b_2^2 1b_1^2 2b_1^2 .$$

This is in line with chemical intuition and with the prediction of Pulay's reference selection schema [24]. Here $1a_2$ and $2b_1$ represent $\pi$ orbitals with one and two nodes, respectively. The two-dimensional reference space corresponds to the GVB approximation and will be denoted by "2R" below.

From the viewpoint of a CAS wavefunction, a third function could be constructed,

$$\phi_3 = \ldots 6a_1^2 4b_2^2 1b_1^2 1a_1^2 2b_1^2 ,$$

but is has different symmetry than the ground state and will not contribute to the wavefunction at $C_{2v}$ geometries. There are, however, single and double excitations out of this reference function which have the right symmetry. These excited configurations do not belong to the first-order interacting space, therefore their effect should be comparatively small. At asymmetrically distorted geometries of the molecule both $a_1$ and $b_1$ orbitals will have $a''$ symmetry and consequently $\phi_3$ will contribute to the ground state wavefunction. In fact, these contributions have been suggested as being responsible for the extreme difficulty in obtaining the correct asymmetric stretch frequency in $O_3$ [39]. To obtain a smooth potential energy surface for the asymmetric distortion (also for calculations of the corresponding vibrational frequencies) $\phi_3$ and the single and double excitations out of this reference need to be included in the wavefunction. We will refer to this wavefunction as "3R".

Using the "2R" wavefunction we have calculated the equilibrium geometry and vibrational frequencies of ozone with several different methods. The results are shown in table 1. Also included for comparison are the results of CCSDT calculations by Watts et al. [34]. They should be viewed as a very good approximation of the exact results for the given basis. From the results of the table we can conclude that MR-ACPF overestimates the effect of quadruple corrections to MR-CI giving unreasonably long bonds and too low frequencies. Although $\omega_3$ is greater than $\omega_2$ contrary to experiment, the 2R-CI result is better, suggesting that the reference space is qualitatively reasonable. The same conclusion can be reached from the GVB-PT2 calculations of Wolinsky and Pulay [42], where exactly the same reference function is used and reasonable geometries are obtained. Therefore, it seems that it is not the correlation problem itself but the quadruple or, equivalently, extensivity correction, that is overestimated. This can be seen from the total energies, too, as MR-ACPF gives lower energies than CCSDT.

The results of the MR-AQCC calculations lie between the corresponding MR-CI and MR-ACPF results and are very close to CCSDT. Especially encouraging is the good agreement for the asymmetric stretching frequency, although $\omega_3 = \omega_2$. As was shown in ref. [34] it requires higher excitations in the single reference CC methods to correctly obtain this frequency. The total energy of MR-AQCC is higher than
Table 1
Ground state properties of ozone using 2 × 2 CAS reference space  

<table>
<thead>
<tr>
<th>Method</th>
<th>2R-CI</th>
<th>2R-LCCM</th>
<th>2R-ACP</th>
<th>2R-AQCC</th>
<th>CCSDT  c)</th>
<th>exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{OO} )</td>
<td>1.261</td>
<td>1.326</td>
<td>1.315</td>
<td>1.292</td>
<td>1.286</td>
<td>1.272</td>
</tr>
<tr>
<td>( \angle_{OOO} )</td>
<td>116.5</td>
<td>115.7</td>
<td>115.8</td>
<td>116.1</td>
<td>116.7</td>
<td>116.8</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>1235</td>
<td>796</td>
<td>930</td>
<td>1070</td>
<td>1141</td>
<td>1135</td>
</tr>
<tr>
<td>( \omega_2 )</td>
<td>761</td>
<td>547</td>
<td>640</td>
<td>694</td>
<td>705</td>
<td>716</td>
</tr>
<tr>
<td>( \omega_3 )</td>
<td>1338</td>
<td>478</td>
<td>898</td>
<td>1070</td>
<td>1077</td>
<td>1089</td>
</tr>
<tr>
<td>( E ) c)</td>
<td>-0.871296</td>
<td>-0.944178</td>
<td>-0.942051</td>
<td>-0.928506</td>
<td>-0.941197</td>
<td></td>
</tr>
</tbody>
</table>

\* Bond lengths in Å, bond angle in deg, harmonic frequencies in cm\(^{-1}\).
\b Total energy = -224 hartree.
\c Three highest virtual orbitals are frozen.

Fig. 1. Potential energy curves for simultaneous stretching of \( O-O \) bonds of the ozone molecule calculated with different methods (\( \angle_{OOO} \) is fixed at 116.48°).

that of CCSDT and MR-ACPF but still considerably lower than CI. This fact suggests that this new averaging procedure does not overestimate the effect of higher excitations in agreement with our goal. Although the energy is not especially close to the nearly exact values, more reasonably falling between MR-CI and CCSDT, the potential energy surface is rather parallel to the exact one as can be seen from the good agreement of the CCSDT and MR-AQCC geometry and vibrational frequencies.

The above results show that in the case of this rather small reference space MR-ACPF considerably overestimates the effect of higher excitations. By extending to the "3R" wavefunctions, which also includes more configurations outside the reference space, the effect is even pushed to the extreme. We
have calculated the potential energy surface for stretching both O0 bonds with fixed bond angle (116.48°) not too far from the equilibrium distance. The results are shown in fig. 1 and compared to the corresponding curves obtained with the “2R” wavefunction. As expected, the two MR-CI curves run nicely together, i.e. the effect of these non-interacting configurations is very small. This is not the case for MR-ACPF. With the “3R” wavefunction we do not even find a minimum in this range of values. The longer the O0 distance is, the smaller the contribution of the reference function to the total wavefunction and consequently the MR-ACPF correction becomes larger and larger. This effect is similar to that for MR-LCCM, where the missing points (1.50 and 1.55) could not be converged (see fig. 1). Much better behavior can be observed for the MR-AQCC method, especially around equilibrium.

In conclusion we can state that MR-AQCC gives excellent results for ozone with a small reference space and eliminates the failure of the MR-ACPF method.

We have also performed calculations on other systems like N2, Be2, H2O, C2, etc. with similarly positive results. These results will be published separately [20]. We conclude therefore that MR-AQCC provides reasonable results with small reference spaces even when MR-ACPF fails. In cases when MR-ACPF is a good approximation, MR-AQCC provides very similar results.

Acknowledgement

This Letter benefitted significantly from conversations with Professor I. Shavitt, Dr. L. Meissner, Dr. J. Stanton, and Dr. R. Gdanitz. This work was supported by the US Air Force Office of Scientific Research, Grant No. F49620-93-1-0127 and by the Hungarian Scientific Research Foundation (No. 354-0713).

References