I. INTRODUCTION

Configuration interaction (CI) and coupled-cluster (CC) methods would be exact, if the excitation operator were not truncated. This is, however, not realistic computationally and truncation is requisite. The most natural truncation is after double excitation (SD, singles and doubles model), which can be justified by perturbation theory provided the reference function \( \Psi_0 \) is a reasonable approximation to the exact wave function. Both CISD and CCSD are widely used in practice, however, the methods are fundamentally different. CISD truncates a linear excitation operator or configuration space to double excitations, while CCSD truncates the excitation operator in an exponential to double excitations, meaning that all nonvanishing products of the double with double and double and a single excitation operators remain in the wave function. The latter property is responsible for extensivity (correct scaling with number of electrons and the absence of unlinked diagrams) despite truncation of the wave function. In CI the lack of extensivity is a very serious deficiency for even a few electrons and the error grows fast with the size of the system. The removal of unlinked diagrams is only possible by mixing different categories of CI excitation spaces, manifesting itself only by explicit cancellation at the molecular integral level. With the wealth of development of CI technology, though, it is often operationally much easier to work in the underlying configuration space.

CI also has advantages in its variational property and therefore its energy functional includes only the wave function parameters. The latter is important since it allows the use of a generalized Hellmann–Feynman theorem to calculate energy derivatives and properties. Since the SD approximation requires \( \Psi_0 \) to be a good approximation, it often must consist of a multideterminantal reference function, leading to a MR-CISD description. There are effectively no essential complications in generalizing CI for multideterminantal reference functions.

Though truncated CC methods are extensive due to the exponential form of the wave function, their generalization to multideterminantal reference functions is not as straightforward. There are several multistate MRCC methods (Fock space CC, Hilbert space CC) and also some state selective versions (two-determinant TDCC, SS-CCSD(TQ), OSA-SS-CC), for specific reference spaces. Moreover, although a functional satisfying stationary principles and an associated generalized Hellmann–Feynman theorem can be defined for the CC energy and properties, this functional depends on additional parameters that have to be introduced. The determination of these extra parameters costs about the same as an energy calculation for single state methods, though it could be even more expensive for multistate methods, therefore the calculation of energy derivatives is somewhat more expensive for CC than for CI.

The ease of using a configuration space in CI calculations recommends attention at finding modified versions of CI which are, at least, approximately extensive. These methods should also provide the use of a stationary principle to facilitate property calculations. Such approaches are derived from variants of CEPA (coupled electron pair approximation) methods, having a functional form such as CPF (coupled
pair functional)\(^{23}\) and CEPA-VAR.\(^{24}\) Here, our interests lie in the multireference variants.

We have recently proposed a MR-AQCC (multireference averaged quadratic coupled-cluster) method\(^{25}\) which, as a generalization of MR-LCCM (multireference linearized CC method),\(^{26,27}\) is a step toward obtaining a general single state multireference CC theory. Beside MR-AQCC, several other approximately extensive modifications of MR-CI have been published in recent years.\(^{25–39}\) They are closely related (sometimes even exactly equivalent), but their derivation has been performed using different starting points and formalism, and only a few incomplete comparisons\(^{40,41}\) exist.

The aim of this paper is to provide a theoretical and numerical comparison of all such methods. In Sec. II, we present a theoretical comparison starting from the full CI equations. In Sec. III, we present several numerical examples chosen to test the capability of MR-AQCC compared to the other methods.

II. THEORY

We partition the full CI space according to three categories.\(^{42}\) The \(P\) space includes the reference functions \((\Phi^R_0)\), the \(Q\) space includes all unique single and double excitations out of the functions in \(P(\Phi^R_0)\), while \(R\) includes all higher excitations \((\Phi^R)\). The projection to these spaces are denoted by \(\hat{P}, \hat{Q}\), and \(\hat{R}\), respectively. The expectation value of the full CI correlation energy can be written as

\[
\langle \Psi_p + \Psi_Q + \Psi_R | \hat{H} - E_0 | \Psi_p + \Psi_Q + \Psi_R \rangle = E_c
\]

with \(\Psi_p = \sum c_i^p \Phi_i^p\), \(\Psi_Q = \sum c_i^Q \Phi_i^Q\) and \(\Psi_R = \sum c_i^R \Phi_i^R\) and the individual \(\Phi_i\)'s form an orthonormal set. \(E_0\) is the reference energy to be specified later. Varying the coefficients of the individual functions of any of these three sets gives the following CI equations, for all \(i\):

\[
\langle \Phi_i | \hat{H} - E | \Psi_p + \Psi_Q + \Psi_R \rangle = 0
\]

with \(\Phi_i \in \{P, Q, R\}\), and \(E = E_0 + E_c\). In practical applications the full CI problem is not solvable and therefore we consider the above equation with \(\Phi_i \in \{P, Q\}\) only, and try to express the coefficients of the higher excitations \(c_i^R\) by the coefficients \(c_i^p\) and \(c_i^Q\). Defining a constant \(K_i\) as

\[
K_i = \frac{\langle \Phi_i | \hat{H} | \Psi_R \rangle}{\langle \Phi_i | \Psi_p + \Psi_Q + \Psi_R \rangle} = \frac{\langle \Phi_i | \hat{H} | \Psi_R \rangle}{c_i}
\]

we can write the above equations as

\[
\langle \Phi_i | \hat{H} - E | \Psi_p + \Psi_Q \rangle + K_i \langle \Phi_i | \Psi_p + \Psi_Q \rangle = 0
\]

or

\[
\langle \Phi_i | \hat{H} - E + K_i | \Psi_p + \Psi_Q \rangle = 0.
\]

If we would know the coefficients of the functions in the \(R\) space, the above equation would still be exact, as this step simply corresponds to repartitioning the CI matrix.\(^{43}\) If the reference space is chosen such that it represents a good description of the final wave function, one would hope that the higher excitations would be less important in the wave function. On the other hand, we cannot neglect them completely (this would lead us to the CISD method) since a part of their contribution [second term in Eq. (2)] cancels the unlinked part of the first term \((E|\Psi_Q\rangle\)). This cancellation is essential to get extensive results. Therefore, corrected CI methods try to approximate \(K_i\).

Before analyzing this, we would like to comment on the so-called EPV (exclusion principle violating) terms. Configuration based methods, like CI, satisfy the exclusion principle, while many-body methods are interpreted such that some higher excitation diagrams include EPV contributions.\(^{14}\) This is not an unphysical effect, however, and its origin can be understood by analyzing the CI equations (2). The unlinked part of these equations include unrestricted summation over the excitations in \(E\) and \(|\Psi_Q\rangle\), therefore in \(E|\Psi_Q\rangle\) there are products of excitations including the same orbitals. Those are not called EPV terms since they are unlinked, but it can be shown that they equal the negative of the corresponding linked EPV terms.\(^{45}\) Since the higher excitation part of the CI equations do not include these linked EPV terms, EPV terms remain after complete cancellation of the unlinked terms.

Now we return to the expression for \(K_i\) and try to eliminate the higher excitation coefficients in it. We first consider the single reference case with \(\Phi_0\) being the reference function (generalizations to a multireference \(P\) space will be discussed later)

\[
K_i = \frac{\langle \Phi_i | \hat{H} | \Psi_R \rangle}{c_i} = \frac{\sum c_i^R \langle \Phi_i | \hat{H} | \Phi_i^R \rangle}{c_i}.
\]

These matrix elements are not zero if \(\Phi_i^R\) is a single or double excitation out of \(\Phi_0\)

\[
\Phi_i^R = \hat{E}_i \Phi_i = \hat{E}_i \hat{E}_i \Phi_0,
\]

where the single and double excitation operators are defined as

\[
\Phi_i = \hat{E}_i \Phi_0.
\]

Therefore, we can write

\[
K_i = \frac{\sum_k c_i^R \langle \Phi_0 | \hat{H} | \Phi_k \rangle}{c_i},
\]

where \(j\) is defined by \(k\) according to Eq. (4). One can easily recognize that (a) \(\Phi_j \in Q\) since all single and double excitations are included in \(Q\), (b) there is a restricted summation denoted by \((k)\) since not all \(\hat{E}_i\) acting on \(\Phi_i\) produce a function in the \(R\) space. Considering only double excitations, only those \(\hat{E}_i\)'s are excluded, which have common orbital indices with the excitation \(\hat{E}_i\) (i.e., \(\hat{E}_i \hat{E}_i\) cannot be an EPV excitation). Equivalently, included are those which have no common index (disconnected) with the excitation \(\hat{E}_i\). For single excitations additional requirements exist.\(^{35}\) We will use the notation of Ref. 35 and will denote this subset as \(\mathcal{D}(i)\).

If we do not want to deal with the higher excitation effects explicitly, we need to express the coefficient \(c_i^R\) by coefficients of the \(Q\) space functions. To that end the cluster condition can be used which is generally a good approximation
which is simply the sum of the energy contribution of the individual configurations over virtual orbitals. The sum of the pair energies gives the correlation energy  

$$
\sum_{ij} \epsilon_{ij} = E_c.
$$

In the multireference case Eq. (4) has the form  

$$
\Phi_j^R = \hat{E}_i \Phi_j^Q = \hat{E}_i \hat{E}_j \Phi_j^P.
$$

There are two complications compared to the single reference case: (a) in general there is no unique choice of $\Phi_j^R$ (i.e., a function in the reference space) to produce $\Phi_j^Q$; (b) not all $k \in \mathcal{D}(i)$ will produce a function $\Phi_k \in \mathcal{R}$. The right subset [say $\mathcal{D}'(i)$] cannot be given generally, and will depend on the choice of reference space. The detailed analysis of these problems are given in Refs. 35 and 37. Besides this, the validity of the cluster condition (6) is not completely clear for a multireference function.

As discussed in Sec. I, the knowledge of a functional of the energy is very important, particularly if one is interested in energy derivatives. Naturally, this functional will be an approximation of the full CI functional [Eq. (1)]. It can be shown that the functional corresponding to the stationary equation (2) exists if $K_i$ is independent of the coefficients $c_i$ and in this case it has the form

$$
K_i = \sum_{k \in \mathcal{D}(i)} c_k \langle \Phi_0 | \hat{H} | \Phi_i^Q \rangle.
$$

(7)

The correlation energy for the intermediate normalized full CI wave function can be expressed as  

$$
E_c = \sum_i c_i \langle \Phi_0 | \hat{H} | \Phi_i^Q \rangle.
$$

Then, we can rewrite $K_i$ as  

$$
K_i = E_c - \sum_{k \in \mathcal{D}(i)} c_k \langle \Phi_0 | \hat{H} | \Phi_k \rangle.
$$

(8)

Note, that in $K_i | \Psi_Q \rangle$ the last term represents the EPV contribution.

In the discussion of some of the methods below, it is useful to introduce the concept of pair energy defined by  

$$
\epsilon_{ij} = \sum_{ab} c_{ij} \langle \Psi_0 | \hat{H} | \Phi_{ij}^{ab} \rangle
$$

(9)

with $\hat{K}$ defined as  

$$
\hat{K} = \sum_{i \in P} K_i | \Phi_i \rangle \langle \Phi_i |.
$$

Making this functional stationary according to the coefficients in the wave function we clearly get Eq. (2). If $K_i$ does depend on $c_i$ then Eq. (2) is not linear in the parameters, and therefore the functional cannot be constructed. More precisely it can be constructed by the introduction of new parameters (Lagrangian multipliers) which are related to the explicit solution of the first derivative equations as in CC theory. To calculate energy derivatives one must also determine these new parameters, which make the calculation more expensive. If the functional (9) exists, however, the calculation of the gradient can be performed as for the MR-CI wave function, using slightly modified density matrices. Therefore we consider the existence of functional (9) essential and we prefer methods based on this functional.

It is usual to define the wave function as  

$$
\Psi = \Psi_0 + \Psi_c,
$$

where $\Psi_0$ is a function in the $P$ space and $\Psi_c$ is the correlation correction. There are two ways to define both $\Psi_0$ and $\Psi_c$. 

First, $\Psi_0$ can be the eigenfunction of the Hamilton in the reference space $P$  

$$
\hat{P} \hat{H} \hat{P} \Psi_0 = E_0 \Psi_0.
$$

(10)

These correspond to a reference function used in MCSCF-CI type applications. The second choice already includes the relaxation in the reference space due to the presence of the $Q$ space, and $\Psi_0$ is the solution of an effective Hamiltonian equation  

$$
(\hat{P} \hat{H} \hat{P} + \hat{P} \hat{H} \hat{Q} (E_c - \hat{H} \hat{Q}^{-1} \hat{Q}) \hat{H} \hat{P}) \Psi_0 = E \Psi_0
$$

(11)

with $E_c$ being a constant. This is simply the partitioned form of CI, if $E_c = E$, but other choices are possible, which are often used in many-body applications such as MR-MBPT. Note, that CI based on Eq. (11) is equivalent with a CI based on Eq. (10).

The definition of $\Psi_c$ depends on the choice of $\Psi_0$. Since the latter choice of $\Psi_0$ already benefits from relaxation in the reference space, one would have  

$$
\Psi_c = \Psi_Q,
$$

(12)

i.e., only the $Q$ space contributes to the correction. The other choice is  

$$
\Psi_c = \Psi_Q + \Psi'_P = \Psi_Q + (1 - \Psi_0 | \Psi_0 \rangle) \Psi_P
$$

(13)
so that $\Psi_c$ also includes the contribution of the orthogonal complement of the reference space $P$ with respect to $\Psi_0$.

Below we will discuss the actual equations for the different methods. The equations for all have the form of Eq. (2). They will differ in

- choice of $K_i$;
- choice of $\Phi_0$ and $E_0$;
- choice of the correlation correction of the wave function;
- existence of a functional.

### A. MR-CI

CI methods neglect all contributions from higher excitations and therefore the MR-CISD equations are simply obtained by setting $K_i = 0$. Doing this we introduce an extensivity error since the unlinked parts of the first term of Eq. (2) are not cancelled. Usually the orthogonal complement is included in the wave function [Eq. (13)]. The functional exists in the form of Eq. (9).

### B. MR-LCCM (Refs. 26 and 27)

The simplest approximation to $K_i$ [from Eq. (8)] is

$$K_i = E_c$$

for $i \in Q$ and $$K_i = 0$$ otherwise. In the single reference case this approximation was introduced by Čiček as the linearized coupled-pair many-electron theory (L-CPMET). It is equivalent to the sum of doubles excitation diagrams in many-body perturbation theory to infinite order [D-MBPT ($\infty$)], and is also called LCCD (linearized CC doubles) and CEPA(0). As Eq. (7) shows, it means that we neglect the restriction on $k$ in the summation, i.e., we include EPV terms in $K_i|\Psi_0\rangle$. This approximation seems to be justified if $\Phi_0$ is a good wave function, i.e., the $c_k$’s are small. However, for many cases the norm of the vector of these coefficients necessarily grows, and CEPA(0) becomes a poorer approximation. To make $\Phi_0$ better, we might consider a multireference description of the reference function. In this case $K_i$ in the form of Eq. (14), besides generating the EPV terms, will also produce excitations which are already present in the $Q$ part of the wave function according to the discussion above. Despite these obvious shortcomings, this approximation is valid through third-order perturbation theory and it offers a simple way to account for some effect of higher excitations.

The first MR method with this approximation was proposed by Bartlett and co-workers in their MR-LCCM (multireference linearized coupled-cluster method). The derivation is based on linearization of the multireference coupled-cluster (MR-CC) equations. For the wave function, this method chose to exclude the orthogonal complement of the reference space as in Eq. (12). This choice facilitates the CC-type derivation, while also improving the convergence behavior of the equations. The actual equations and the functional are

$$\langle \Phi_i | \hat{H} - E_c | \Psi_0 + \Psi_c \rangle = \langle \Phi_i | \hat{H} - E_0 | \Psi_0 \rangle = 0$$

$$F(\Psi_c) = \langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle$$

The papers discuss that the $P'$ space could logically be included. This slight addition leads us to the next group of methods.

### C. MR-CEPA(0), (Ref. 29) UCEPA, (Ref. 30) and the VPT (Ref. 31) methods

Comparing the equations of MR-CEPA(0), unitary CEPA (UCEPA), and the variational perturbation theory (VPT) methods, it is easy to see that these methods are equivalent. They only differ from MR-LCCM, as the $K_i$ approximation is the same, by the inclusion of the $P'$ complement, into the correlation correction $\Psi_c$ of the wave function.

Later in this paper we will refer to these three methods as MR-CEPA(0) since among these three methods MR-CEPA(0) was the first published.

### D. QDVPT method (Ref. 32)

The quasidegenerate variational perturbation theory (QDVPT) method of Cave and Davidson also uses a linearized approximation [Eq. (11)]. To avoid the convergence problem of MR-CEPA(0) but still account for relaxation in the $P$ space, the effective Hamiltonian equation (11) is used with $E_c = E_0$ to define $\Phi_0$ and $E_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle$. Since (a) the neglect of the relaxation in the reference space is normally less important than the CEPA(0) approximation itself, and (b) Eqs. (10) and (11) are equivalent for CI, it is expected that QDVPT will give very similar results as MR-CEPA(0) (or MR-LCCM). This was also found in calculations. Different behavior can be expected, however, for quasidegenerate cases or for cases where the reference space is not a good zeroth-order space. In this latter instance, however, the CEPA(0) approximation [Eq. (14)] is not valid and neither method should be used.

From our point of view there is another disadvantage: since $E_0$ depends on the coefficients $c_i$ through Eq. (11), no functional can be easily defined to facilitate analytical gradient calculations, making them more expensive for QDVPT than for MR-CEPA(0) or MR-LCCM.

### E. MR-CEPA method (Ref. 33)

The problems arising from the multireference nature of $\Psi_0$ have been discussed above. The MR-CEPA method of Ruttink et al. tries to account for this overcounting of certain excitations within the CEPA(0) approximation [Eq. (14)]. In this method excitations are assigned to classes based on the number of holes in the doubly occupied orbitals and the number of particles in the virtual space. Their choice of $K_i$ is

$$K_i = K_{\alpha(i)} = \sum_{(\beta)} e_{\beta}$$
where $\alpha(i)$ is the excitation class of $i$ (for example doubly occupied-virtual excitations, etc.). The summation goes only over those classes of excitations ($\beta$) which acting on $\Phi_i \in \alpha(i)$ lead to the $R$ space and

$$e_\beta = \sum_{k \in \beta} c_k \langle \Psi_0 | \hat{H} | \Phi_k \rangle. \quad (18)$$

By this choice, the overcounting of some excitations due to the MR nature of $P$ (as in any of the other methods above) is avoided. On the other hand, it neglects the effect of single excitations in $K_i$ and, due to the unrestricted summation in Eq. (18), some of the EPV terms are still present in $K_i$. In fact, for the single reference case (only one excitation class exists) this method is equivalent to CEPA(0)=LCCD.

Since $K_i$ now depends explicitly on $c_k$, no functional can be defined and the gradient calculations are again much more expensive than in CI. In addition, the dependence of $K_{\beta}$ on $c_k$ causes computational difficulties in Davidson-type direct methods.

**F. MR-CPA methods (Ref. 38)**

The MR-CPA (coupled pair approximation) methods of Tanaka et al.\textsuperscript{38} are based on a different approach than the other methods discussed above. The final equations of this method can, however, easily be compared to them.

Like MR-LCCM, the starting point for this method is the Hilbert space multireference CC method\textsuperscript{12} which is a “perturb then diagonalize” method. That is, the individual functions of the reference space $\Psi_{\beta} \in P$ are treated separately and for all of them an equation similar to Eq. (2) is solved

$$\langle \Phi_i | \hat{H} - E(\mu) + K_i | \Phi_{\beta} + \Psi_{\mu} \rangle(\mu) = 0 \quad (19)$$

for all $\mu=1, \ldots n_{\text{ref}}$. Therefore, we get as many different $\Psi_{\beta}(\mu)$ correlation corrections as the dimension of the reference space. Note, that the cost of solving such equations is $n_{\text{ref}}$ times higher than in the other approaches. The energy and wave function of the states are obtained by diagonalizing the effective Hamiltonian as Eq. (11). By this we will get $n_{\text{ref}}$ roots and the extra cost pays off since we also get excited states. An important property is that these simplified Hilbert space equations (unlike the full ones) are not coupled. By prediagonalizing at an earlier stage, MR-LCCM further eliminates the need for additional $\Psi_{\beta}(\mu)$, already the full Hilbert space CC method has the disadvantage that only states can be described which are similar. Here, due to the missed coupling of the equations, this effect is supposed to be an even more pronounced problem. On the other hand, the missing coupling helps the so-called “intruder state” problem.

Tanaka et al.\textsuperscript{38} proposed two choices of $K_i$

$$K_i = E(\mu) - \langle \Phi_{\mu} | \hat{H} | \Phi_{\mu} \rangle = E_{\alpha}(\mu) \quad (20)$$

for the MR-CPA(0) and

$$K_i = E_{\alpha}(\mu) - \langle \Phi_{\mu} | \hat{H} | \Phi_{\mu} \rangle c_i^{(1)}(\mu) \quad (21)$$

for the MR-CPA(2) method. Here, $c_i^{(1)}(\mu)$ is the coefficient of the function $\Phi_i$ from a lower order equation [not Eq. (19)]. For more detail see Ref. 38. Note that approximation (20) is essentially the CEPA(0) approximation [Eq. (14)] and in the limit of a single reference it is equivalent to LCCD or CEPA(0) method. The other approximation called MR-CPA(2) [Eq. (21)] already has an EPV correction but only the diagonal EPV term from the summation in Eq. (8) is included. We do not see any justification to include only this one term, although it is certainly in the summation. Most probably none of the two versions will give better performance than the other methods based on the approximation Eq. (14).

The advantage of this method is that it is claimed to be size consistent, for some special systems\textsuperscript{38} (the proof in Ref. 38 does not include cases where single excitations on the separated subsystems are allowed). The disadvantage is that no functional can be defined without the introduction of additional parameters.

**G. MR-ACPF method (Ref. 29)**

The essential approximation of the above discussed methods is that in the LCCD [or CEPA(0)] approximation [Eq. (14)] the second term in Eq. (8) (i.e., the EPV contribution) is neglected. The MR-averaged coupled pair functional (MR-ACPF) method of Gdanitz and Ahlrichs\textsuperscript{39} accounts for these EPV terms in an averaged way. If we assume that the $(n/2)$ electron pairs of the molecules ($n$ being the number of correlated electrons)\textsuperscript{51} are not interacting and they contribute by an equal amount to the correlation energy, the pair energy is given by

$$\bar{e} = \frac{E_c}{n/2}.$$

Now we can account for the restricted summation in Eq. (7) by excluding one of the pairs

$$K_i = (\frac{1}{2}n - 1) \frac{E_c}{n/2} = \frac{n-2}{n} E_c \quad (22)$$

for all $i \in Q$, while $K_i=0$ for $i \in P$. With this choice the MR-ACPF equation and the functional are\textsuperscript{29}

$$\langle \Phi_i | \hat{H} - E | \Psi_0 + \Psi_c \rangle \frac{n-2}{n} E_c \langle \Phi_i | \Psi_0 + \Psi_c \rangle = 0 \quad (23)$$

$$F(\Psi_c) = \langle \Psi_0 + \Psi_c | \hat{H} - E_0 | \Psi_0 + \Psi_c \rangle \frac{1}{n} \langle \Psi_0 | \Psi_0 \rangle + (2/n) \langle \Psi_c | \Psi_c \rangle. \quad (24)$$

The MR-ACPF method represents an improvement over the LCCD [or CEPA(0)] approximation of all other methods discussed above, since the EPV terms are approximately deleted from $K_i$. Although the method is not strictly size extensive, this choice ensures that the method is exact for non-interacting electron pairs. Note, that the rather crude approximation of noninteractive electron pairs is applied to account for the EPV effects and not for the full treatment of electron correlation.
H. MR-AQCC method (Ref. 25)

Obviously, the most satisfactory approach of this type that we could follow would be the rigorous inclusion of all quadratic terms. A step toward this solution, is to offer a better approximation for the quadratic terms. Hence, we recently proposed a slightly modified version of MR-ACPF. To approximate the EPV terms in the multireference averaged quadratic CC method, 35 the correlation energy is again distributed among equivalent electron pairs, but now all possible electron pairs (\( \binom{n}{2} \)) are considered. 52 The pair energy is given as

\[
\tilde{e}^i = \frac{E_c}{n} \binom{n-2}{2}
\]

and in \( K_i \) we can approximately exclude EPV terms by summing only for the disconnected electron pairs

\[
K_i = \frac{(n-2)(n-3)}{2(n-1)} E_c.
\]

The actual equation and the functional of the MR-AQCC method is then,

\[
\langle \Phi | \hat{H} - E | \Phi \rangle + \frac{(n-2)(n-3)}{n(n-1)} E_c \langle \Phi | \Phi \rangle = 0.
\]

(25)

\[
F(\Phi) = -\frac{\langle \Psi_0 + \Phi | \hat{H} - E | \Psi_0 + \Phi \rangle}{\langle \Psi_0 | \Psi_0 \rangle + \frac{(n-3)(n-2)}{n(n-1)} \langle \Phi | \Phi \rangle}.
\]

(26)

Our motivation in deriving this method is to improve upon the approximation of noninteracting electron pairs used in MR-ACPF but to keep its very convenient functional formulation. By distributing the correlation energy among all possible electron pairs, we account for the interaction of the electron pairs in an averaged way. As in MR-ACPF, this method is not strictly size extensive, but the size extensivity error of the method is still small. Unlike MR-ACPF, the method is not exact for noninteracting electron pairs, but the correction in the functional disappears for three electrons 52 as it should. Finally, initial applications on ozone using a small reference space 35 shows the good performance of the method for a difficult problem. In this paper we will present further examples.

I. QDVPT-APC method (Ref. 34)

Recently Murray et al. 34 extended the QDVPT method discussed earlier by using \( K_i \) of the MR-ACPF method [Eq. (22)] instead of \( K_i \) of the CEPA(0) approximation [Eq. (14)] used in the original QDVPT. 32 This means simply to use

\[
E_x = E_0 + \frac{2}{n} E_c
\]

in Eq. (11) This method is called QDVPT-APC (QDVPT with averaged pair correction) and it can also be viewed as an extension of MR-ACPF via the effective Hamiltonian Eq. (11). Since MR-ACPF does not show convergence problems as MR-CEPA(0) does, the relaxation in the reference space can be readily handled by the form of the correlation wave function Eq. (13) as in the original form of the MR-ACPF method. 35 The QDVPT-APC method is expected to perform better than MR-ACPF if quasidegeneracy is present, but should always be a better approximation than QDVPT, since the latter neglects EPV terms in \( K_i | \Phi \rangle | \Phi \rangle \) completely.

Finally, we would like to note that the QDVPT generalization of MR-AQCC can be trivially made by using

\[
E_x = E_0 + \left( 1 - \frac{(n-3)(n-2)}{n(n-1)} \right) E_c.
\]

J. MCCEPA method (Ref. 39)

The multiconfiguration CEPA method of Fink and Staemmler 39 is another effort to go beyond the CEPA(0) approximation. \( K_i \) is given by

\[
K_i = \sum_k \frac{1}{1 + \delta_k} \langle \Phi_0 | \hat{H} | \Phi_k \rangle
\]

(27)

which is essentially a multireference generalization of the scheme of Kelly and Sessler. 33 This approximation means that the EPV matrix elements are approximated by

\[
\langle \Phi_1 | \hat{H} | \Phi_2 \rangle = \langle \Phi_0 | \hat{H} \hat{\Phi}_0 | \Phi_0 \rangle \langle \hat{\Phi}_0 | \hat{\Phi}_0 | \hat{\Phi}_0 \rangle.
\]

It is a theoretically more justified approximation than that of MR-ACPF and MR-AQCC for the EPV terms, since the equivalence of electron pairs is not assumed. However, it is still an approximation, and its quality can only be determined from actual calculations. To do this comparison, only a few numbers calculated by MCCEPA are available for us in this paper. In addition, in practical application MCCEPA is used together with the PNO (pair natural orbitals) approach 52 to generate the multiconfigurational wave function (for details see the original paper). 39 Using MCCEPA results in our comparisons, we have to remember this additional approximation.

Since \( K_i \) depends on \( c_k \), there is no functional related to this method.

K. (SC) 2 SCI method (Ref. 35)

As has been shown by Daudet et al. 35 for the single reference case, it is possible to calculate \( K_i \) in the form of Eq. (8) without further approximation. This, however, requires determinantal instead of configuration based wave functions to generate the elements of the set \( \mathcal{Q}(i) \). Two possible extensions to multireference wave functions is proposed. The first variant chooses an element of the \( P \) space (say \( \Phi^D_0 \)) and treats the multireference problem as if it is a single reference one. Note that by this construction the \( Q \) space includes functions which are higher than double excitations with respect to \( \Phi^D_0 \). Therefore, a generalization of their single reference SD method to completely general, truncated, configurational space is required. 35 The second
variant now treats all functions in $P$ equivalently. Since functions in the $Q$ space may be produced by single and double excitations from more than one reference function, its contribution to $K_i$ is weighted. For more details we refer to the original paper. No application of this latter method has yet been published. The advantage of this method is that it is size extensive if certain conditions are fulfilled, while retaining the configuration space (instead of many-body) description. The disadvantage seems to be that no functional can be defined. Although the energy calculation does not cost much more than CI (in direct algorithms similar problems may occur as in the case of MR-CEPA), gradient calculations would seem to be much more expensive. It also lacks the unitary invariance of the rigorous theory.

III. NUMERICAL TESTS

In this section we present numerical tests of the methods discussed in the previous section. Most of the tests were chosen from the literature to allow comparison of all the methods to MR-AQCC. In our calculations the COLUMBUS program system was used.

A. Bond stretching in H$_2$O

Our first example is the simultaneous stretching of both OH bonds in the water molecule. In order to describe this motion, the reference space contains the OH bonds and the corresponding antibonding orbitals, i.e., a 4 orbital/4 electron CAS

$$1a_1^22a_1^1b_1^1(3a_1^4a_1^1b_2^12b_2)^4$$

This reference space can be considered to be large compared to the number of electrons.

There are two full CI calculations for this system which can be used in comparisons. The first one by Harrison and Handy uses a DZ basis, the other one by Bauschlicher and Taylor uses a DZP basis. Since some of the methods investigated in this paper are used with DZ, others with DZP, both bases will be discussed in this paper. For both bases the lowest occupied $(1a_1)$ orbitals are frozen.

Table I compares the DZ results and Table II the DZP results by listing the energy difference from full CI for the different methods. Since the numbers in the two tables are rather different for the same methods, we conclude that the DZ basis is not good enough to describe electron correlation, and, particularly the higher excitation effects.

However, since the MR-CPA methods have been tested on the DZ basis only, we are forced to make some conclusions in this case, too. There is no overestimation of the higher excitation effects by the MR-CPA methods, but the parallelism of the potential energy surface to the full CI [especially for the less approximate MR-CPA(2)] is not satisfactory.

For the DZP basis, we find that the CEPA(0) type methods overestimate the effect of higher excitations, but the absolute value of the error is small. MR-ACPF gives the smallest error, but MR-AQCC is the only method which does not overcompensate the effect of higher excitations. This latter behavior is very important, since, as the theory part shows, all methods treated in this paper approximate the disconnected higher excitations only, and therefore should not overshoot full CI. The absolute error of MR-AQCC is large, but not as large as in MR-CI. However, since in chemistry we are primarily interested in getting surfaces parallel to the exact one, the absolute error is less important. For this reason we plot the numbers of Table II on Fig. 1 to see how parallel the potential energy surfaces are to full CI. We have shifted all surfaces to the same origin at $R_e$ in order to show the quality of their parallelism to full CI (the best method would have a surface of $F=0$ at all distances). We again can see that the best performing methods are MR-ACPF and MR-AQCC; MR-ACPF is closer to the full CI, while the error of MR-AQCC has the same sign at all geometries. MR-CEPA(0) and MR-LCCM are worse, while MR-CEPA seems to perform better showing that the correction of the CEPA(0) approximation according to the MR space might be important for this large reference space.

B. Potential energy curve for Be$_2$

The potential energy curve for Be$_2$ is a challenging test for ab initio methods. It has a very shallow minimum and there might be a second minimum on the dissociation plateau, although the most recent work doubts that.

Two benchmarks are available: full CI by Harrison and Handy, for a limited number of point near the minimum, and CCSDT by Sosa et al. for a larger interval. Both cal-

<table>
<thead>
<tr>
<th>Geometry</th>
<th align="right">$R_e$</th>
<th align="right">$1.5R_e$</th>
<th align="right">$2R_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full CI (hartree)</td>
<td align="right">Ref. 56</td>
<td align="right">-75.157 866</td>
<td align="right">-75.014 521</td>
</tr>
<tr>
<td>MR-CI (Ref. 26)</td>
<td align="right">0.06</td>
<td align="right">0.25</td>
<td align="right">0.50</td>
</tr>
<tr>
<td>MR-CEPA (0)</td>
<td align="right">Ref. 59</td>
<td align="right">-0.78</td>
<td align="right">-0.83</td>
</tr>
<tr>
<td>MR-CEPA (0)</td>
<td align="right">Ref. 38</td>
<td align="right">-0.35</td>
<td align="right">0.75</td>
</tr>
<tr>
<td>MR-LCCM</td>
<td align="right">Ref. 37</td>
<td align="right">0.7</td>
<td align="right">1.0</td>
</tr>
<tr>
<td>MR-CPA(0)</td>
<td align="right">Ref. 38</td>
<td align="right">0.1</td>
<td align="right">0.5</td>
</tr>
<tr>
<td>MR-CPA(2)</td>
<td align="right">Ref. 38</td>
<td align="right">0.1</td>
<td align="right">0.8</td>
</tr>
<tr>
<td>MR-ACPF</td>
<td align="right"></td>
<td align="right">0.02</td>
<td align="right">0.19</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td align="right"></td>
<td align="right">0.62</td>
<td align="right">0.80</td>
</tr>
</tbody>
</table>

*Energies relative to full CI in mhartree.

*No interacting space restriction.

*First order interacting space restriction.
calculations are performed using a $7s3p1d$ basis with Cartesian $d$ functions, and the two lowest occupied ($1\sigma_g,1\sigma_u$) orbitals are frozen in the correlated calculations.

In this study our goal is to compare with benchmark calculations, therefore, we need to use the same basis as in the former calculations. For this reason the quality of our calculations, therefore, we need to use the same basis as in the correlated calculations.

To get a qualitatively correct description again, a $2\times2$ CAS reference space is required

$$\Phi_1 = 1\sigma_g^21\sigma_u^22\sigma_g^22\sigma_u^2,$$

$$\Phi_2 = 1\sigma_g^21\sigma_u^22\sigma_g^23\sigma_u^2.$$

The results for MR-CI, MR-ACPF, and MR-AQCC are compared with CCSDT and full CI on Fig. 2. Here, we have shifted all surfaces to the same origin at 5 a.u. to get a better comparison. From this we can see

- full CI and CCSDT are very close for all points where full CI is available;
- MR-CI overestimates the binding energy;
- MR-ACPF curve is dissociative;
- MR-AQCC is very close to the CCSDT surface (there might be a slight maximum at large distance; but this is likely an artifact of the basis).

In addition to our MR-ACPF and MR-AQCC results, there is a calculation of Daudey et al.\textsuperscript{55} on Be$_2$ with their (SC)$^2$SCI method using the same reference space. Their value for the binding energy (difference between the energy at 5 and 8 a.u.) is 0.461 mhartree which compares excellently with our MR-AQCC value of 0.461 mhartree (the CCSDT value is 0.494 mhartree\textsuperscript{58}). Unfortunately, there is no other example available to compare MR-AQCC to (SC)$^2$SCI with a multireference function. Thus although the above mentioned good agreement of the approximate MR-AQCC with the theoretically more justified (SC)$^2$SCI is very promising, general conclusions cannot be drawn.

C. Equilibrium properties of C$_2$

The ground state of the C$_2$ molecule requires a two-determinantal description in zeroth order. Therefore, it is again a good example to test the performance of MR-AQCC with a small reference space.

In this calculation we have used a $2\times2$ CAS reference function defined by the two determinants

$$\Phi_1 = 1\sigma_g^21\sigma_u^22\sigma_g^21\pi_u^2\sigma_g^2,$$

$$\Phi_2 = 1\sigma_g^21\sigma_u^22\sigma_g^21\pi_u^4\sigma_g^2.$$ 

We have used a DZP basis\textsuperscript{62–64} with Cartesian $d$ functions, and the two lowest occupied ($1\sigma_g,1\sigma_u$) orbitals have been frozen in the CI calculations. Using these we can compare our results with accurate CCSDT results.\textsuperscript{65}

We have calculated the equilibrium geometry and the harmonic vibrational frequency. The results are compared in Table III. We conclude:

- MR-CI gives a much shorter bond distance and larger frequency than CCSDT showing the importance of higher excitation effects;
- The MR-CEPA(0) approximation overestimates these effects by a large amount;
- MR-ACPF gives better results but still overestimates the effect especially for the frequency. Note that the MR-ACPF energy is slightly below the CCSDT energy;
- MR-AQCC gives results very close to CCSDT and it does not overestimate the effect of higher excitations.

The potential energy surface is given in Fig. 3. We have again shifted the curves to the same origin at 1.24 Å to show
how the curvature calculated with different methods compare. We again see how MR-ACPF overestimates the effect of higher excitations giving a curve that falls below CCSDT.

D. N₂ molecule

To investigate how the MR-AQCC performs when more than two reference configurations are required, we perform calculations on N₂. Here, we use three different reference functions

- Hartree–Fock wave function
  \[ 1 \sigma^2_2 \sigma^2_2 2 \sigma^2_2 3 \sigma^2_2 4 \pi^4. \]
- \(4 \times 4\) CAS (8 references)
  \[ 1 \sigma^3_2 1 \sigma^2_2 2 \sigma^2_2 3 \sigma^2_2 3(1 \pi_u 1 \pi_g). \]
- \(6 \times 6\) CAS (32 references)
  \[ 1 \sigma^2_1 2 \sigma^2_2 3 \sigma^2_2 4 \sigma^2_1 (3 \sigma^2_3 3 \sigma_u 1 \pi_u 1 g \pi_g)^6. \]

This allows us to investigate the dependence of the reference space size on the equilibrium bond distance and harmonic vibrational frequency.

To be able to compare with CCSDT results,⁶⁶ we used a DZP basis⁶²,⁶³,⁶⁶ with Cartesian \(d\) functions and freeze the two lowest occupied \((1 \sigma^2_1 1 \sigma_u)\) orbitals.

The dependence of the geometry on the reference size is given in Table IV, and that of the frequency in Table V. We see that

- MR-AQCC and MR-ACPF results using 8 references are very close, being even closer in the 32 reference case.

E. The Be/H₂ system

Insertion of Be into H₂ is a very often used model to test the performance of both single reference and multi-reference methods.⁶⁷ Using this example we can compare most of the methods treated in this paper.

There are two important configurations \(1 a^2_1 2 a^2_1 3 a^1_2\) and \(1 a^2_2 2 a^1_1 1 b^2_1\) for this system and these define the reference space in the calculations. Normally three geometries of the potential energy surface are considered. In geometry 1 the second configuration is dominating, while at geometry 3 the first one is the most important. Geometry 2 represents an intermediate between these two situations. Therefore, this model tests how good the methods perform with a strongly changing reference function.

A \(3s, 1p,\) and \(2s\) basis was used for Be and H, respectively.⁶⁸ The full CI results are given in Ref. 67.

Table VI collects the results. Inspection of the numbers leads to the following conclusions:

- MCCEPA, MR-ACPF (QDVPT+ACP) and MR-AQCC methods come closest to full CI, while all other

<table>
<thead>
<tr>
<th>No. of ref. functions</th>
<th>1</th>
<th>8</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCSDT (Ref. 66)</td>
<td>1.127</td>
<td>1.125</td>
<td>1.128</td>
</tr>
<tr>
<td>MR-CI</td>
<td>1.112</td>
<td>1.125</td>
<td>1.128</td>
</tr>
<tr>
<td>MR-ACPF</td>
<td>1.123</td>
<td>1.129</td>
<td>1.129</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>1.120</td>
<td>1.128</td>
<td>1.127</td>
</tr>
</tbody>
</table>

This method is a good approximation.

- It seems that all three schemes converge to the same geometry and frequency which seems to differ from the CCSD values.
- The size extensivity error diminishes with the larger reference space.
- For both MR-AQCC and MR-ACPF, the 8 reference wave function is a good approximation.

### Table IV. Calculated bond length (Å) of N₂ with different reference spaces.

<table>
<thead>
<tr>
<th>No. of ref. functions</th>
<th>1</th>
<th>8</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCSDT (Ref. 66)</td>
<td>1.127</td>
<td>1.125</td>
<td>1.128</td>
</tr>
<tr>
<td>MR-CI</td>
<td>1.112</td>
<td>1.125</td>
<td>1.128</td>
</tr>
<tr>
<td>MR-ACPF</td>
<td>1.123</td>
<td>1.129</td>
<td>1.129</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>1.120</td>
<td>1.128</td>
<td>1.127</td>
</tr>
</tbody>
</table>

### Table V. Calculated frequency (cm⁻¹) of N₂ with different reference spaces.

<table>
<thead>
<tr>
<th>No. of ref. functions</th>
<th>1</th>
<th>8</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCSDT (Ref. 66)</td>
<td>2294</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR-CI</td>
<td>2444</td>
<td>2303</td>
<td>2274</td>
</tr>
<tr>
<td>MR-ACPF</td>
<td>2320</td>
<td>2264</td>
<td>2264</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>2361</td>
<td>2275</td>
<td>2266</td>
</tr>
</tbody>
</table>

### Table VI. Relative energies for the Be/H₂ system.

<table>
<thead>
<tr>
<th>(r(\text{BeH}_3))</th>
<th>(r(\text{H}_3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full CI (hartree)²</td>
<td>-15.622 88</td>
</tr>
<tr>
<td>MCSCT³</td>
<td>53.31</td>
</tr>
<tr>
<td>MR-CPF</td>
<td>0.84</td>
</tr>
<tr>
<td>MR-LCCM³</td>
<td>-2.62</td>
</tr>
<tr>
<td>MR-CEPA(0)⁹</td>
<td>-3.28</td>
</tr>
<tr>
<td>QDVPT⁴</td>
<td>-2.9</td>
</tr>
<tr>
<td>MR-CEPA²</td>
<td>-1.65</td>
</tr>
<tr>
<td>MR-CPA(2)⁸</td>
<td>-1.2</td>
</tr>
<tr>
<td>MR-ACP²</td>
<td>-0.90</td>
</tr>
<tr>
<td>QDVPT+ACP⁹</td>
<td>-0.91</td>
</tr>
<tr>
<td>MCCEPA</td>
<td>0.20</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>0.29</td>
</tr>
</tbody>
</table>

²Energies relative to full CI in mhartree.
³Reference 32.
⁴Reference 29.
⁵Reference 46.
⁶Reference 33.
⁷Reference 38.
⁸Reference 34.
⁹Reference 29.
methods give a larger error of similar magnitude. Since only the former methods attempt to correct for EPV terms, while the latter methods use a CEPA(0) approximation, we can conclude that going beyond CEPA(0) is essential.

- Since the performance of CEPA(0) type methods are very similar, it seems that the differences among them [e.g., different description of the relaxation in the reference space as in QDVP or correction of the CEPA(0) approximation for the multireference case as in MR-CEPA] are not very important compared to the error of the common CEPA(0) approximation.

- Comparing MR-CEPA(0) and QDVP or MR-ACP shows that the different ways of considering relaxation in the reference space is not important compared to other errors.

- All methods but MR-AQCC and MCCEPA overestimate the error of MR-CI giving a lower energy than full CI. In fact this overestimation is very large for methods using just the CEPA(0) approximation, and the absolute value of the error is even larger than in plain MR-CI. MR-ACP also overestimates the effect of EPV terms, but the absolute error is rather small.

- The error of the CEPA(0) methods are larger than for water. This can be explained by the poorer reference space in this case, which gives a less satisfactory $\Psi_0$.

- The relative error of MR-AQCC especially for geometry 3 is large, although smaller than with CI or CEPA(0) methods. In addition, the surface of these three points is not as parallel to the full CI surface as it was for the other examples above. Analysis of the wave function shows that the weight of the reference functions ($\Sigma_i x_i P_i c_i^2$) is different for the three geometries, while in other examples such as the H$_2$O surface, Be$_2$ surface between 5 and 8 a.u., C$_2$ and O$_3$ surfaces about the equilibrium, for the A and B states of CH$_2$ with the 4×4 CAS reference space (see below) is about unchanged. Therefore we conclude, that MR-AQCC needs a reference space which gives a uniform description of the system for the whole range of interest.

F. $^1A_1 - ^3B_1$ excitation energy of CH$_2$ radical

To test the performance of the methods for excitation energies, we have chosen the CH$_2$ radical, because earlier calculations show that at least the A state requires a multireference description. To find a balanced description of both states is therefore not straightforward.

To be able to compare with the full CI results of Bauchlicher and Taylor, we use a DZP basis with polarization $d$ function on carbon with exponent 0.51 and 0.74 for the $^1A_1$ and $^3B_1$ states, respectively, and a $p$ function of exponent 1.0 on the hydrogen. The geometry is as in Ref. 69.

Table VII lists the energy difference between the two states calculated by the different methods. We conclude

- In the single reference case, none of the methods give satisfactory results, mostly because the SR description for the $^1A_1$ state is not good.

- In the 2×2 CAS case all methods are very good, especially the CEPA(0) methods.

- In the 4×4 CAS case, CI is perfect, and the modified methods give similarly satisfactory results.

These conclusions are not completely consistent with our previous observations. Therefore, in Table VIII we list the error of the energy for the two states calculated by the different methods.

Here, we see that the CEPA(0) methods overcompensate the extensivity error, as does MR-ACP with the larger reference space. The good performance of the CEPA(0) methods for the energy difference in Table VII is definitely a result of error compensation. For the total energies MR-ACP and MR-AQCC give results of similar quality for the 4×4 CAS case, while for the smaller cases the total energy calculated by MR-ACP is better. However, the excitation energies are very similar both for the 2×2 CAS and 4×4 CAS cases showing that the error of the total energy with MR-AQCC is systematic. Its very poor performance for the SR case can be explained again by the fact that this reference describes the A state poorer than the B state, i.e., the weight of the reference changes. Note that, although the MR-ACP excitation energy is better in this case, its error is also too large for practical purposes.

G. Extensivity

Since most of the methods discussed above are not rigorously extensive, only numerical tests can quantify the extensivity (or unlinked diagram) error. Unfortunately, the extensivity error can only be easily calculated for noninteracting subsystems, although it affects the results of interacting systems (molecules) as well.

<table>
<thead>
<tr>
<th>Method</th>
<th>SR</th>
<th>2×2 CAS</th>
<th>4×4 CAS</th>
<th>2×2 CAS</th>
<th>4×4 CAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCSCF</td>
<td>140.89</td>
<td>119.52</td>
<td>91.83</td>
<td>118.31</td>
<td>91.97</td>
</tr>
<tr>
<td>MR-CI</td>
<td>8.53</td>
<td>5.03</td>
<td>1.38</td>
<td>4.66</td>
<td>1.39</td>
</tr>
<tr>
<td>MR-LCCM</td>
<td>-2.95</td>
<td>-1.27</td>
<td>-0.55</td>
<td>-1.33</td>
<td>-0.40</td>
</tr>
<tr>
<td>MR-CEPA(0)</td>
<td>-2.95</td>
<td>-1.27</td>
<td>-1.59</td>
<td>-1.33</td>
<td>-1.26</td>
</tr>
<tr>
<td>MR-ACPF</td>
<td>1.87</td>
<td>1.01</td>
<td>-0.50</td>
<td>0.83</td>
<td>-0.28</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>4.96</td>
<td>2.96</td>
<td>0.28</td>
<td>2.44</td>
<td>0.41</td>
</tr>
</tbody>
</table>

*Relative to full CI in mhartree.
TABLE IX. Extensivity test for diatomic molecules.

<table>
<thead>
<tr>
<th>Method</th>
<th>Monomer</th>
<th>Dimer</th>
<th>∆</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCSCF</td>
<td>−54.398 748</td>
<td>−108.797 497</td>
<td>0.0</td>
</tr>
<tr>
<td>MR-CI</td>
<td>−54.504 190</td>
<td>−109.001 611</td>
<td>6.77e−3</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>−54.505 350</td>
<td>−109.010 536</td>
<td>1.64e−4</td>
</tr>
<tr>
<td>MR-ACPF</td>
<td>−54.506 569</td>
<td>−109.013 429</td>
<td>−2.92e−4</td>
</tr>
</tbody>
</table>

TABLE X. Dependence of extensivity error on reference space (N₂O–HF system).

<table>
<thead>
<tr>
<th>Method</th>
<th>SR</th>
<th>4×4 CAS</th>
<th>SR</th>
<th>6-31+G(d,p)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DZP</td>
<td>6-31+G(d,p)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCSCF</td>
<td>0.000 879</td>
<td>−0.031 518</td>
<td>0.037 088</td>
<td>0.032 695</td>
</tr>
<tr>
<td>MR-CI</td>
<td>0.000 045</td>
<td>0.000 184</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR-LCCM</td>
<td>−0.000 086 055</td>
<td>−0.000 44 842</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR-CEPA(0)</td>
<td>−0.000 551</td>
<td>−0.000 100</td>
<td>−0.000 755</td>
<td>−0.000 327</td>
</tr>
<tr>
<td>MR-AQCC</td>
<td>−0.000 376</td>
<td>−0.000 247</td>
<td>−0.000 662</td>
<td>0.000 002</td>
</tr>
</tbody>
</table>

For the noninteracting limit, we have chosen three systems for the extensivity tests: dissociation of N₂ and O₂, and the infinite separation of N₂O and HF molecules.

Correct dissociation of N₂ requires a 6×6 CAS, which was already used in an earlier subsection. As Table IX shows, neither MR-ACPF nor MR-AQCC shows excellent extensivity, although both improve upon CI by an order of magnitude. The overcompensation of higher excitation effects in MR-ACPF is well demonstrated by the negative sign of the error, while MR-AQCC has a positive error.

To be able to compare with the extensivity test by Ruttink et al. for the MR-CEPA method, our second example was O₂ with a DZ basis. The correct dissociation requires a 4×4 CAS. The results in Table IX show a much better behavior than for N₂. The extensivity error of all methods is very small and similar, only the error of MR-AQCC is two orders of magnitude larger. Note, however, that (a) its error is still an order of magnitude smaller than that of CI and therefore acceptable, and (b) the error is also smaller than in the case of N₂.

The third example is taken from the thorough study of N₂O complexes by Del Bene et al. Here, an excellent extensivity property was found for ACPF, therefore we test AQCC as well. We use the 6-31+G(d,p) basis used in Ref. 70 to test basis dependence, also a DZP basis. The reference space was SR and a 4×4 CAS (on N₂O). The results are summarized in Table X. The excellent extensivity property of ACPF is reproduced by AQCC, and for the 6-31+G(d,p) basis an even smaller error has been found. The extensivity error changes somewhat with the size of the basis, and in most of the cases the error is larger with the bigger 6-31+G(d,p) basis. This finding suggests that the size of basis might be included in the correction procedure. Work along this line is in progress in our groups. The effect of basis is even more pronounced for the CEPA(0) methods. The extensivity of MR-LCCM is excellent for both bases, but we had convergence problems for the dimer with the MR-CEPA(0) method.

Summarizing our findings we say that MR-AQCC and MR-ACPF performs very similarly in terms of extensivity. The fact that MR-ACPF is exact for noninteracting electron pairs (MR-AQCC does not have this property), does not seem to be particularly relevant regarding the extensivity of the method for real systems. In fact, the fluctuation of the extensivity error with the quality of basis or the size of reference space is much larger than the difference in the extensivity error of MR-AQCC and MR-ACPF. Extensivity of the QDVPt method is very similar to that of MR-CEPA(0). 32

IV. CONCLUSION

We have investigated several approximately extensive modifications to multireference CI. Theoretical analysis and numerical tests lead us to the following conclusions:

• MR-LCCM and MR-CEPA(0) approximations are usually not good enough and methods going beyond it [MR-ACPF, MR-AQCC, MCCEPA, (SC)²CI] are preferred.

• The error of the CEPA(0) approximation is much larger than that coming from the description of the relaxation in the reference space or from the correction of the CEPA(0) approximation for the multireference case. Therefore, CEPA(0) methods, those with and without this feature, perform similarly.

• The effective Hamiltonian approach (QDVPt32,34) [both the original version containing the CEPA(0) approximation32 and the extension with coupled pairs (QDVPt-CPA34)] fails to give any improvement over the nonquasidegenerate variants for the systems investigated here. However, for real quasidegenerate cases such a treatment might become essential, and we have proposed a QDVPt style modification to MR-AQCC.

• There are systems for which MR-ACPF fails with small reference spaces. In these cases a substantial increase of the size of the reference space is required, which is expensive and often inconvenient. In contrast, MR-AQCC seems to work well even in these cases with comparatively small reference spaces.

• In other cases MR-ACPF and MR-AQCC perform similarly.

• MR-AQCC needs a reference space which gives a uniform description of the system over the whole range (geometry, or different states, etc.) of interest. In these cases we can expect energy differences very close to full CI. If this requirement is not fulfilled, it still improves upon a CI or CEPA(0) approximation.

• The potential theoretical advantage of MR-ACPF over MR-AQCC is that the former is exact for noninteracting electron pairs, suggesting a better extensivity prop-
no effort is made in the current work to introduce any triples other than from those that arise from doubles or singles out of the reference space configurations. Their systematic addition should be considered in future developments. While MR-ACPF energies are often closer to full CI than MR-AQCC, in cases with small reference spaces MR-ACPF tends to fail qualitatively. No such example has yet been found for MR-AQCC (although could certainly happen), yet its use would appear to be safer.

ACKNOWLEDGMENTS

This paper benefited from conversations with Mr. L. Füsti Molnár. We thank Professor I. Shavitt for his unpublished results using the MR-AQCC method. This work was supported by the Hungarian Scientific Research Foundation (No. F-013962) and by the United States Air Force Office of Scientific Research Grant No. AFOSR-F49620-93-1-0127.

34. I. Shavitt (private communication).
43. To avoid confusion: inclusion of EPV terms in $K[\Psi_F]$ means that we cancel unlinked diagrams in the equation, but do not add all the EPV terms which would be necessary in the case of exact cancellation.
44. Because of this approximation, core electrons should not be included in the averaging.
64. Recently Jansen (Ref. 72) has found much larger extensivity error for SR-ACPF than for SR-AQCC in the case of the Ar–CO system.