

## Alternative definition of excitation amplitudes in multi-reference state-specific coupled cluster

Yann Garniron, Emmanuel Giner, Jean-Paul Malrieu, and Anthony Scemama

Citation: *The Journal of Chemical Physics* **146**, 154107 (2017); doi: 10.1063/1.4980034

View online: <http://dx.doi.org/10.1063/1.4980034>

View Table of Contents: <http://aip.scitation.org/toc/jcp/146/15>

Published by the [American Institute of Physics](#)

---

### Articles you may be interested in

[Equation-of-motion coupled-cluster methods for atoms and molecules in strong magnetic fields](#)

*The Journal of Chemical Physics* **146**, 154105 (2017); 10.1063/1.4979624

[Incremental full configuration interaction](#)

*The Journal of Chemical Physics* **146**, 104102 (2017); 10.1063/1.4977727

[Excited states from modified coupled cluster methods: Are they any better than EOM CCSD?](#)

*The Journal of Chemical Physics* **146**, 144104 (2017); 10.1063/1.4979078

[A new near-linear scaling, efficient and accurate, open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory](#)

*The Journal of Chemical Physics* **146**, 164105 (2017); 10.1063/1.4981521

[Correlated natural transition orbitals for core excitation energies in multilevel coupled cluster models](#)

*The Journal of Chemical Physics* **146**, 144109 (2017); 10.1063/1.4979908

[Driven similarity renormalization group: Third-order multireference perturbation theory](#)

*The Journal of Chemical Physics* **146**, 124132 (2017); 10.1063/1.4979016

---



**COMPLETELY  
REDESIGNED!**

*Physics Today* Buyer's Guide  
Search with a purpose.

# Alternative definition of excitation amplitudes in multi-reference state-specific coupled cluster

Yann Garniron,<sup>1</sup> Emmanuel Giner,<sup>2,3</sup> Jean-Paul Malrieu,<sup>1</sup> and Anthony Scemama<sup>1,a)</sup>

<sup>1</sup>Laboratoire de Chimie et Physique Quantiques, CNRS 5626, IRSAMC, Université Paul Sabatier, Toulouse, France

<sup>2</sup>Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Ferrara, Via Fossato di Mortara 17, I-44121 Ferrara, Italy

<sup>3</sup>Max Planck Institut for Solid State Research, Heisenbergstraße 1, 70569 Stuttgart, Germany

(Received 17 January 2017; accepted 30 March 2017; published online 19 April 2017)

A central difficulty of state-specific Multi-Reference Coupled Cluster (MR-CC) in the multi-exponential Jeziorski-Monkhorst formalism concerns the definition of the amplitudes of the single and double excitation operators appearing in the exponential wave operators. If the reference space is a complete active space (CAS), the number of these amplitudes is larger than the number of singly and doubly excited determinants on which one may project the eigenequation, and one must impose additional conditions. The present work first defines a state-specific reference-independent operator  $\hat{T}^m$  which acting on the CAS component of the wave function  $|\Psi_0^m\rangle$  maximizes the overlap between  $(1 + \hat{T}^m)|\Psi_0^m\rangle$  and the eigenvector of the CAS-SD (Singles and Doubles) Configuration Interaction (CI) matrix  $|\Psi_{\text{CAS-SD}}^m\rangle$ . This operator may be used to generate approximate coefficients of the triples and quadruples, and a dressing of the CAS-SD CI matrix, according to the intermediate Hamiltonian formalism. The process may be iterated to convergence. As a refinement towards a strict coupled cluster formalism, one may exploit reference-independent amplitudes provided by  $(1 + \hat{T}^m)|\Psi_0^m\rangle$  to define a reference-dependent operator  $\hat{T}^m$  by fitting the eigenvector of the (dressed) CAS-SD CI matrix. The two variants, which are internally uncontracted, give rather similar results. The new MR-CC version has been tested on the ground state potential energy curves of 6 molecules (up to triple-bond breaking) and two excited states. The non-parallelism error with respect to the full-CI curves is of the order of 1 mE<sub>h</sub>. Published by AIP Publishing. [<http://dx.doi.org/10.1063/1.4980034>]

## I. INTRODUCTION

The single-reference Coupled Cluster (CC) formalism<sup>1-4</sup> is the standard technique in the study of the ground state of closed-shell molecules, i.e., those for which a mean-field treatment provides a reasonable zeroth-order single-determinant wave-function  $\Phi_0$ . This method incorporates the leading contributions to the correlation energy in a given basis set; it is based on the linked-cluster theorem<sup>5</sup> and is size-extensive since it is free from unlinked contributions. The method generates an approximate wave function under the action of a wave operator  $\hat{\Omega}$  acting on the single-determinant reference  $\Phi_0$ , and assumes an exponential character to the wave operator

$$\Psi = \hat{\Omega}\Phi_0 = e^{\hat{T}}\Phi_0. \quad (1)$$

The most popular version only introduces single and double excitation operators in  $\hat{T}$ , and is known as the Coupled Cluster Singles and Doubles (CCSD) approximation. It incorporates the fourth-order correction of the quadruply excited determinants. The lacking fourth-order contribution concerns the triply excited determinants, which may be added in a perturbative manner. The CC equations, obtained by projecting the eigenequation on each of the Singles and Doubles (SD), lead to coupled quartic equations. In practice, guess values of

the amplitudes of the  $\hat{T}_{0 \rightarrow i}$  operators appearing in the  $\hat{T}$  operator may be taken as the coefficients of the singles and doubles  $|i\rangle$  in the intermediate normalization of the SD Configuration Interaction (CI) vector. The solution of the CC equations may be obtained by treating the effect of the triples and quadruples as an iterative dressing of the SD CI matrix,<sup>6</sup> according to the Intermediate Effective Hamiltonian (IEH) theory.<sup>7,8</sup> The field of application of this method, which satisfies formal requirements and is numerically efficient, is however limited to the systems and the situations where a single-determinant zeroth-order description is relevant. This is no longer the case when chemical bonds are broken, creating open shells, as occurs in most of the chemical reactions. The magnetic systems generally present several open shells, and the low spin-multiplicity states are inherently of multiple-determinant character. Due to near degeneracies, most of the excited states are not only of multi-determinantal but of multi-configurational character. The conception of a multi-reference (MR) counterpart of the CCSD formalism is highly desirable, and has been the subject of intense research. The most comprehensive review has been given by Bartlett and his colleagues.<sup>9</sup> For formal reasons and in particular to treat correctly the breaking of bonds, the reference space, or model space, is usually taken as a Complete Active Space (CAS), i.e., the Full-CI (FCI) of a well-defined number of electrons (the active electrons) in a well-defined set of orbitals (the active MOs). The other MOs are called

<sup>a)</sup>Electronic mail: scemama@irsamc.ups-tlse.fr

inactive. Let us label  $|I\rangle, |J\rangle, \dots$  the reference determinants. The determinants  $|i\rangle, |j\rangle, \dots$  which interact with the reference space are obtained under purely inactive or semi-active single and double excitations; they generate the CAS-SD CI space, the diagonalization of which provides a size-inconsistent energy  $E_{\text{CAS-SD}}^m$  and the corresponding eigenvector,

$$\begin{aligned} |\Psi_{\text{CAS-SD}}^m\rangle &= |\Psi_0^m\rangle + |\Psi_{\text{SD}}^m\rangle \\ &= \sum_{I \in \text{CAS}} C_I^m |I\rangle + \sum_{i \notin \text{CAS}} c_i^m |i\rangle \end{aligned} \quad (2)$$

with  $\langle \Psi_{\text{CAS-SD}}^m | \Psi_{\text{CAS-SD}}^m \rangle = 1$ .

One strategy, which is not very aesthetic since it breaks the symmetry between degenerate reference determinants, but which has given rather satisfactory results, consists in selecting (eventually in an arbitrary manner) a specific single reference and in introducing in the wave operator the multiple excitations which generate the other references (the other determinants of the model space).<sup>10</sup> A similar procedure was proposed by Li and Paldus which uses specific three and four body amplitudes issued from a MR-CISD function.<sup>11</sup> The other strategies consider all the references on an equal footing, and are really multi-reference. Let us call  $N$  the number of references, and  $n$  the number of SD determinants. If the treatment pretended to provide  $N$  eigenvectors simultaneously, one might define the  $N \times n$  amplitudes sending from the references to the outer-space determinants, in a unique manner but this state-universal approach is not practicable when the model space is a CAS.

Most of the proposed formalisms are state-specific. In this case one faces the famous multi-parentage problem. This problem is recalled in Section II A. Sufficiency conditions have to be imposed.<sup>12</sup> One solution was proposed by Mukherjee and co-workers, and has been widely tested.<sup>13–15</sup> Another one had been proposed earlier by one of us (JPM) and co-workers.<sup>16</sup> It consists, for a given outer-space determinant, in scaling the amplitudes of the various excitation operators  $\hat{T}_{I \rightarrow i}$  on the interaction between the outer-space determinant and its parents. A recent work has implemented this second solution of the state-specific Multi-Reference Coupled Cluster (MR-CC) problem and has tested its accuracy and robustness on a series of molecular benchmarks, comparing its results to the full-CI energies.<sup>17</sup> In the text, we will refer to this method as  $\lambda$ -MR-CCSD. The present work proposes an alternative process to define the amplitudes of the excitation operators, and this new method will be called  $\mu$ -MR-CCSD.

The state-specific MR-CC formalisms are usually based on the Jeziorski-Monkhorst<sup>18</sup> splitting of the wave operator into a sum of operators acting individually on the various references

$$\hat{T}^m = \sum_I \hat{T}_I^m |I\rangle \langle I|. \quad (3)$$

We shall leave in a first time this assumption and define in Section II B a reference-independent operator  $\hat{T}$  which acting on the component of the desired state in the model space,  $|\Psi_0^m\rangle$ , provides a vector as close as possible to the CAS-SD eigenvector. This solution, defining reference-independent amplitudes of the excitations, may be exploited directly to generate approximate values of the coefficients of the triply and quadruply excited determinants, according to the exponential

structure of the wave operator. From these coefficients, one may dress the CAS-SD CI matrix, redefine amplitudes, and iterate the process to convergence. This solution, presented in Section II C, is not an MR-CC technique; one may call it an exponential dressing of the CAS-SD CI matrix. Section II D redefines reference-dependent excitation amplitudes from the reference-independent amplitudes by a fitting of the previous amplitudes on the coefficients of the singles and doubles of the (dressed) CAS-SD CI eigenvector. This represents an alternative solution to multi-parentage problem and opens the way to a strict MR-CC formalism. Section III presents a series of numerical tests on the bond breaking of single, double, and triple bonds in ground states of molecules as well as a few tests on excited states. The results are compared to our previous proposal and with Full Configuration Interaction (FCI) results.

## II. FORMALISMS

In this section, all the presented formalisms are state-specific. To simplify the notations we will consider that the state superscript  $m$  is implicit for the wave functions ( $\Psi^m \rightarrow \Psi$ ) and for the excitation operators ( $\hat{T}^m \rightarrow \hat{T}$ ).

### A. The multi-parentage problem in the Jeziorski-Monkhorst approach

Since one wants to produce a MR-CCSD method, one may start from a preliminary CAS-SD CI calculation which will help to fix guess values of the amplitudes of the excitation operators. Let us call  $|I\rangle, |J\rangle, \dots$  the determinants of the CAS, i.e., the so-called reference vectors, and  $|i\rangle, |j\rangle, \dots$  the singles and doubles which do not belong to the CAS and interact with them. The resulting approximate wave function of the targeted state  $|\Psi\rangle$  is written as

$$|\Psi_{\text{CAS-SD}}\rangle = \sum_I C_I |I\rangle + \sum_i c_i |i\rangle. \quad (4)$$

Although this function is not size consistent, one may note that the coefficients on the CAS determinants are no longer those of the CAS-CI: they incorporate the effect of the dynamical correlation on the composition of the CAS component of the wave function.

In CC formalisms the wave operator  $\hat{\Omega}$  is assumed to take an exponential form

$$\hat{\Omega} = \exp(\hat{T}) \quad (5)$$

and in our previous MR-CC formalism<sup>17</sup> the Jeziorski-Monkhorst multi-exponential structure of the wave operator was adopted, introducing reference-specific wave operators acting specifically on each reference vector (Eq. (3)). One may exploit the knowledge of the CAS-SD CI eigenvector to determine guess operators  $\hat{T}_I$  defined in such a manner that

$$|\Psi_{\text{CAS-SD}}\rangle = \sum_I C_I \hat{T}_I |I\rangle. \quad (6)$$

Each of the  $\hat{T}_I$  operators is a sum of single and double excitations  $\hat{T}_{I \rightarrow i}$  possible on  $|I\rangle$ , multiplied by an amplitude  $t_{I \rightarrow i}$ ,

$$\hat{T}_I = \sum_i t_{I \rightarrow i} \hat{T}_{I \rightarrow i}. \quad (7)$$

In the single-reference CC, the amplitudes of the excitation operators are obtained by projecting the eigenequation on the singly and doubly excited determinants; the number of unknowns is equal to the number of equations. This is no longer the case in the MR context: projecting the eigenequation on each of the singly or doubly excited vectors  $|i\rangle$  is not sufficient to define the amplitudes  $t_{I\rightarrow i}$  since for many classes of excitation, an outer-space determinant interacts with several references,  $|i\rangle = \hat{T}_{I\rightarrow i}|I\rangle = \hat{T}_{J\rightarrow i}|J\rangle$ . The condition

$$C_i = \sum_I t_{I\rightarrow i} C_I \quad (8)$$

is not sufficient to define the amplitudes, even if one restricts the excitation operators to single and double excitations. Additional constraints have to be introduced to fix the amplitudes, and this is the famous *multi-parentage problem*. The number of amplitudes is larger than the number of outer-space determinants so that one cannot determine directly the guess values of the amplitudes from Eq. (6). Different additional constraints have been proposed. One of them consists in scaling the amplitudes on the Hamiltonian interactions between the references and the outer space determinants,

$$\frac{t_{I\rightarrow i}}{t_{J\rightarrow i}} = \frac{\langle i|\hat{H}|I\rangle}{\langle i|\hat{H}|J\rangle}. \quad (9)$$

This constraint is expressed as

$$t_{I\rightarrow i} = \lambda_i \langle i|\hat{H}|I\rangle, \quad (10)$$

where

$$\lambda_i = \frac{c_i}{\langle i|\hat{H}|\Psi_0\rangle}. \quad (11)$$

This solution has been recently implemented<sup>17</sup> and shown to provide excellent agreements with full-CI results on a series of molecular problems. From now on, we will refer to this method as the  $\lambda$ -MR-CCSD.

When the term  $\langle i|\hat{H}|\Psi_0\rangle$  is small, the  $\lambda$ -MR-CCSD presents minor stability problems which may introduce some jitter in the potential energy surfaces. A more important problem of the  $\lambda$ -MRCC is illustrated by considering the case of a non-interacting  $A \cdots B$  system with localized MOs on  $A$  and  $B$ . The 2-hole 2-particle inactive double excitations of the type  $\hat{T}_{i_A j_B \rightarrow r_A s_B}$  have zero amplitude in the  $\lambda$ -MRCC formalism since the integral  $\langle i_A j_B | r_A s_B \rangle = 0$ . The coefficient of the determinant  $\hat{T}_{i_A j_B \rightarrow r_A s_B} |I\rangle$  is not zero but it is equal to the product  $t_{i_A \rightarrow r_A} t_{j_B \rightarrow s_B} c_I$ . In Sec. II B, we propose an alternative solution to the multi-parentage problem to define amplitudes which do not suffer from this pathological behavior.

## B. Introduction of reference-independent amplitudes

The present method differs from the  $\lambda$ -MR-CCSD in the definition of the amplitudes, introduced in this section. The formalism will leave in the first step the Jeziorski-Monkhorst formulation of the wave operator and will consider the possibility to define a unique state-specific reference-independent operator  $\hat{T}$ , written as a sum of single and double excitation operators,

$$\hat{T} = \sum_{mnpq} t_{mn\rightarrow pq} a_p^\dagger a_q^\dagger a_n a_m + \sum_{mp} t_{m\rightarrow p} a_p^\dagger a_m, \quad (12a)$$

$$= \sum_{mnpq} t_{mn\rightarrow pq} \hat{T}_{mn\rightarrow pq} + \sum_{mp} t_{m\rightarrow p} \hat{T}_{m\rightarrow p}, \quad (12b)$$

where the indices  $p$  and  $q$  run on the virtual and active MOs and the indices  $m$  and  $n$  run on the inactive occupied and active MOs, excluding the possible occurrence of 4 active MOs. An operator of this kind (but keeping only the linearly independent combinations of the elementary operators) is used in the internally contracted MR-CC method (ic-MRCC) by Evangelista and Gauss,<sup>19</sup> and by Hanauer and Köhn.<sup>20</sup> A similar and more compact formulation was already suggested by Mahapatra *et al.*<sup>21</sup> Our formalism differs by both the determination of the amplitudes and by the way we use them, as will appear later. The ic-MRCC method determines the amplitudes of the excitations by solving the projected coupled cluster equations, where the amplitudes appear up to quartic terms. Hereafter we exploit the knowledge of the CAS-SD CI eigenvector to determine the guess values of the reference-independent amplitudes. These excitation amplitudes will be used later on to estimate the coefficients of the triples and quadruples, and perform an iterative dressing of the CAS-SD CI matrix introducing the coupling between the singles and doubles with the triples and quadruples.

We propose a criterion to fix the amplitudes  $t = \{t_{mn\rightarrow pq}, t_{m\rightarrow p}\}$ . Given the fact that we have at our disposal the CAS-SD wave function, a natural way to solve this overdetermined problem is to minimize the distance between the CAS-SD vector and the vector obtained by applying the  $(1 + \hat{T})$  operator on the CAS wave function

$$\begin{aligned} \arg \min_t \|(1 + \hat{T})|\Psi_0\rangle - |\Psi_{\text{CAS-SD}}\rangle\| \\ = \arg \min_t \|\hat{T}|\Psi_0\rangle - |\Psi_{\text{SD}}\rangle\|, \end{aligned} \quad (13)$$

$\hat{T}|\Psi_0\rangle$  being normalized such that  $\|\hat{T}|\Psi_0\rangle\| = \|\Psi_{\text{SD}}\|$ .

To perform the minimization, we build the  $N_{\text{SD}} \times N_t$  transformation matrix  $A_{i, mn\rightarrow pq} = \langle i|\hat{T}_{mn\rightarrow pq}|\Psi_0\rangle$  which maps from the outer space of determinants  $\{|i\rangle\}$  to the space of excited wave functions  $\{\hat{T}_{mn\rightarrow pq}|\Psi_0\rangle\}$ , and we search for the vector of amplitudes  $\mathbf{t}$  which minimizes  $\|\mathbf{A} \cdot \mathbf{t} - \mathbf{c}\|$  by solving the normal equations,

$$(\mathbf{A}^\dagger \mathbf{A}) \mathbf{t} = \mathbf{A}^\dagger \mathbf{c}. \quad (14)$$

Note that in the single-reference case,  $\mathbf{A}$  is a permutation matrix and the CAS-SD wave function is exactly recovered.

The matrix  $\mathbf{A}$  is usually so large that the use of standard singular value decomposition (SVD) routines to obtain the least squares solution is prohibitive.

Let us first consider the most numerous 2-hole-2-particle inactive double excitations  $\hat{T}_{j_k \rightarrow r_s}$ . These excitations consist in creating two holes in the doubly occupied orbitals and two particles in the unoccupied orbitals. For each excitation of this kind, as all the involved orbitals are outside of the active space, the number of determinants originating from this process is equal to the number of determinants in the reference. Moreover, each one of these excited determinants is doubly excited with respect to only one determinant  $|I\rangle$  of the reference, and the excitation degree with respect to all other reference determinants is necessarily higher than two. Therefore, all excited determinants created by such a 2-hole-2-particle process have only one parent in the reference, and

the corresponding rows of  $\mathbf{A}$  contain only one non-zero element located in the  $jk \rightarrow rs$  column with value  $A_{i,jk \rightarrow rs} = C_I$ . The condition fixing the amplitude  $t_{jk \rightarrow rs}$  is given by

$$\arg \min_{t_{jk \rightarrow rs}} \|\hat{T}_{jk \rightarrow rs} |\Psi_0\rangle t_{jk \rightarrow rs} - |\Psi_{SD}\rangle\| \quad (15)$$

which is obtained by minimizing

$$\min_{t_{jk \rightarrow rs}} \left( \sum_I \left( C_I t_{jk \rightarrow rs} - \sum_i c_i \langle i | \hat{T}_{jk \rightarrow rs} | I \rangle \right) \right)^2 \quad (16)$$

using Eq. (14), and this condition turns out to be satisfied using only one non-zero coefficient  $c_i$  with

$$t_{jk \rightarrow rs} = \frac{\sum_I C_I c_i}{\sum_I C_I^2}. \quad (17)$$

One may notice that this is the weighted average of the ratios between the coefficients of the doubly excited determinants  $|i\rangle$  and the coefficient of their unique reference generator,

$$t_{jk \rightarrow rs} = \frac{1}{\sum_I C_I^2} \left( \sum_I C_I^2 \left( \frac{c_i}{C_I} \right) \right). \quad (18)$$

The maximum number of non-zero elements per row of  $\mathbf{A}$  is equal to the number of reference determinants since each excitation operator applied on a reference produces no more than one excited determinant. Hence, for all the remaining active excitations,  $\mathbf{A}$  remains sparse and we solve Eq. (14) using Richardson's iterative procedure<sup>22</sup>

$$\begin{cases} \mathbf{t}_0 = \mathbf{A}^\dagger \mathbf{c}, \\ \mathbf{t}_{n+1} = \mathbf{A}^\dagger \mathbf{c} + (\mathbf{I} - \mathbf{A}^\dagger \mathbf{A}) \mathbf{t}_n, \end{cases} \quad (19)$$

which may be implemented very efficiently using sparse matrix products.

There are cases where multiple amplitudes applied to different references lead to same determinant:  $\hat{T}_{jk \rightarrow rs} |I\rangle = \hat{T}_{lm \rightarrow tv} |J\rangle = |i\rangle$ . The linear system is underdetermined, so there are infinitely many possible amplitudes verifying the equations. Among the infinity of possibilities, the SVD picks one particular solution given by  $\mathbf{A}^+ \mathbf{c}$ , where  $\mathbf{A}^+$  is the pseudo-inverse of  $\mathbf{A}$ . As this solution minimizes the norm of the amplitude vector,<sup>23</sup> the arbitrariness brought by the null space of  $\mathbf{A}$  is minimized and one obtains the most sensible solution.

### C. Evaluation of the coefficients of triples and quadruples and iterative dressing of the CAS-SD CI matrix

This section recalls the procedure described in our previous work.<sup>17</sup> The so-determined excitation operator  $\hat{T}$  may be used to generate the approximate values of the coefficients of the triples and quadruples as obtained by the action of  $\frac{1}{2} \hat{T}^2$ . Actually one may assume, in the spirit of the internally contracted MR-CC methods, that the wave operator  $\hat{\Omega}$  generating the correlated wave function  $\Psi$  from  $\Psi_0$ ,

$$\Psi = \hat{\Omega} \Psi_0, \quad (20)$$

has an exponential structure,

$$\hat{\Omega} = \exp(\hat{T}). \quad (21)$$

But this form will be simply used to estimate the coefficients of the triply and quadruply excited determinants  $\{|\alpha\rangle\}$ , leaving the internally contracted structure of the outer-space. The coefficients of these determinants are estimated as

$$c_\alpha = \frac{1}{2} \langle \alpha | \hat{T}^2 | \Psi_0 \rangle. \quad (22)$$

All the determinants  $\{|\alpha\rangle\}$  are generated by applying all the single and double substitutions on the singles and doubles, and filtering out the determinants which are already in the wave function. For each  $|\alpha\rangle$  one searches for the reference determinants  $\{|I\rangle\}_\alpha$  which differ by no more than 4 orbital substitutions from  $|\alpha\rangle$  (its grand-parents). One then identifies the set of all possible complementary excitations as the products of excitations  $\hat{T}_p$  and  $\hat{T}_q$ , which generate  $|\alpha\rangle$  from every member  $|I\rangle$  of the set  $\{|I\rangle\}_\alpha$ , i.e.,

$$\mathcal{S}_\alpha = \left\{ (p, q, I) : \forall |I\rangle \in \{|I\rangle\}_\alpha, (\hat{T}_p \hat{T}_q |I\rangle = |\alpha\rangle) \right\}. \quad (23)$$

It now straightforward to find the set of singles and doubles with which  $|\alpha\rangle$  interacts through the matrix elements  $\langle i | \hat{H} | \alpha \rangle$ , namely,

$$\{|i\rangle\}_\alpha = \left\{ |i\rangle : \forall (p, q, \cdot) \in \mathcal{S}_\alpha, (|i\rangle = \hat{T}_p^\dagger |\alpha\rangle) \right\}. \quad (24)$$

For each  $|i\rangle$ , in the eigenequation

$$\begin{aligned} (\langle i | \hat{H} | i \rangle - E) c_i + \sum_J \langle i | \hat{H} | J \rangle C_J + \sum_{j \neq i} \langle i | \hat{H} | j \rangle c_j \\ + \sum_\alpha \langle i | \hat{H} | \alpha \rangle c_\alpha = 0, \end{aligned} \quad (25)$$

the coefficient  $c_\alpha$  is given by the genealogy of  $|\alpha\rangle$ ,

$$c_\alpha = \sum_{(p,q,I) \in \mathcal{S}_\alpha} (-1)^{n(I \rightarrow \alpha)} t_p t_q C_I, \quad (26)$$

$n(I \rightarrow \alpha)$  being given by the number of permutations needed to go from  $|I\rangle$  to  $|\alpha\rangle$ . One may replace the sum over the  $\alpha$  by a dressing of the matrix elements between the determinant  $|i\rangle$  and the references which are grand-parents of  $|\alpha\rangle$ ,

$$\langle i | \hat{\Delta} | I \rangle = \sum_\alpha \langle i | \hat{H} | \alpha \rangle \left( \sum_{(p,q,J) \in \mathcal{S}_\alpha: (J=I)} (-1)^{n(I \rightarrow \alpha)} t_p t_q \right) \quad (27)$$

since

$$\sum_I \langle i | \hat{\Delta} | I \rangle C_I = \sum_\alpha \langle i | \hat{H} | \alpha \rangle c_\alpha. \quad (28)$$

The effect of the triples and quadruples is thus incorporated as a change of the columns of the CAS-SD CI matrix corresponding to the interaction between the references and the singles and doubles,

$$(\langle i | \hat{H} | i \rangle - E) c_i + \sum_J \langle i | (\hat{H} + \hat{\Delta}) | J \rangle C_J + \sum_j \langle i | \hat{H} | j \rangle c_j. \quad (29)$$

This type of dressing was already employed in our previous MR-CC implementation.<sup>17</sup> One will find in the same reference the practical procedure to make the dressed matrix Hermitian without any loss of information. Of course the whole process may be iterated. The diagonalization of the dressed CAS-SD CI matrix provides new values of the coefficients, not only of the singles and doubles which no longer suffer from the truncation but also those of the references: *the method is*

fully decontracted. From the new wave function new amplitudes are obtained, a new dressing is defined and the process is repeated till convergence, which is usually rapidly obtained (3-4 iterations).

As opposed to the  $\lambda$ -MR-CCSD method which uses reference-specific amplitudes, the amplitudes introduced in Sec. II B are reference-independent. As a consequence, the formalism is not a strict MR-CC method since we exploit the CAS-SD CI function which slightly differs from the vector resulting from the action of  $\hat{T}$  on the vector. Although the distance between these two vectors has been minimized, they are not identical,  $(1 + \hat{T})|\Psi_0\rangle \neq |\Psi_{\text{CAS-SD}}\rangle$ .

Once the  $\hat{T}$  operator has been obtained, one might imagine a contracted exponential formalism calculating  $\hat{T}^2|\Psi_0\rangle$  and the interaction between  $\hat{T}|\Psi_0\rangle$  and  $\hat{T}^2|\Psi_0\rangle$ , but this calculation requires to return to the determinants. This formalism would remain internally contracted and would be less accurate than the decontracted procedure we propose. Actually in this version, the deviations of the approximate reference-independent amplitudes from optimal ones, those which would generate the exact coefficients of the singles and doubles, only affect the evaluation of the coefficients of the triples and quadruples, and these deviations represent a minor source of error in the correction restoring the size extensivity. This reliability will be illustrated in the numerical tests.

#### D. State-specific MR-CC variant

In order to return to MR-CC formalism, one may simply exploit the reference-independent amplitudes as an initial guess to define reference-dependent amplitudes. Currently the determinant  $|i\rangle$  belonging to the singles and doubles has a coefficient  $\tilde{c}_i$  in  $\hat{T}|\Psi_0\rangle$ ,

$$\tilde{c}_i = \langle i|\hat{T}|\Psi_0\rangle = \sum_{\{(p,I):(\hat{T}_p|I)=|i\rangle\}} t_p C_I, \quad (30)$$

which differs from the coefficient  $c_i$  in  $|\Psi_{\text{SD}}\rangle$ . One can define a parameter  $\mu_i$ , specific of the determinant  $|i\rangle$ ,

$$\mu_i = \frac{c_i}{\tilde{c}_i}, \quad (31)$$

which multiplying with  $\tilde{c}_i$  will produce the exact coefficient  $c_i$  of  $|i\rangle$  in the (dressed) CAS-SD CI eigenvector. So the previous reference-independent amplitudes have now become reference-dependent. The excitation  $\hat{T}_p$  which excites  $|I\rangle$  to  $|i\rangle$  ( $|i\rangle = \hat{T}_p|I\rangle$ ) receives a reference-dependent amplitude

$$t_{I \rightarrow i} = t_{p,I} = \mu_i t_p. \quad (32)$$

The same excitation will receive a somewhat different amplitude when it acts on another reference  $t_{p,J} \neq t_{p,I}$ . This version returns to the Jeziorski-Monkhorst formalism as the wave operator again is a sum of reference-specific operators. The so-obtained amplitudes may be exploited to generate the coefficients of the triples and quadruples, and one may follow the same strategy as in our previous formalism, with an iterative column dressing of the interactions between the singles and doubles and the references. In what follows, we will refer to this method as  $\mu$ -MR-CCSD as it involves the  $\mu_i$  (Eq. (31)).

As the overlap between  $(1 + \hat{T})|\Psi_0\rangle$  and  $|\Psi_{\text{CAS-SD}}\rangle$  has been maximized, the coefficients  $\tilde{c}_i$  and  $c_i$  are expected to

be very close in particular if  $c_i$  is large, and the parameter  $\mu_i$  should be close to 1, at least for the determinants which contribute significantly to the wave function. In practice we observe this tendency, but the smallest coefficients are sacrificed during the maximization of the overlap and their  $\mu_i$  can be very far from 1. This introduces some instabilities in the iterations, so we chose to limit the values of  $\mu_i$  in the  $[-\mu_i^{\text{max}}, \mu_i^{\text{max}}]$  range, with

$$\mu_i^{\text{max}} = 2 + 100 \times \exp\left(-20 \frac{|c_i|}{\max_j |c_j|}\right). \quad (33)$$

In this way, when  $|c_i|$  is large  $\mu_i$  is constrained in the  $[-2,2]$  range, and when  $|c_i|$  is small,  $\mu_i$  is constrained in  $[-102,102]$ . The effect on the stability of the iterations is significant, and the effect on the energy differences is not noticeable, as seen in Sec. III.

Our procedure makes use of an (non-compulsory but convenient) approximation, namely, the fact that we have not subtracted the product of the single excitations (the  $\hat{T}_1^2$  contributions) from coefficients of the doubles to fix the  $\hat{T}_2$  amplitudes. From a perturbation expansion, one sees that this neglect only introduces fifth-order errors on the energy, which are responsible for small deviations from strict additivity of the energies. Notice that a correct treatment of the  $\hat{T}_1^n$  operations, although tedious, is perfectly conceivable in our formalism and would insure a perfect MR-CC character.

### III. NUMERICAL TESTS

In this section, we first numerically evaluate the errors made by the different approximations. Then, we compare the here-proposed dressed CAS-SD and  $\mu$ -MR-CCSD to the  $\lambda$ -MR-CCSD presented in Ref. 17 on standard benchmark systems.<sup>13,15,19,20,24-33</sup>

The basis set used is Dunning's cc-pVDZ,<sup>34</sup> and the molecular orbitals were obtained using the CAS-SCF code present in GAMESS.<sup>35</sup> All the following calculations were made using the Quantum Package,<sup>36</sup> an open-source program developed in our group. Full-CI energies were obtained using the CIPSI algorithm.<sup>37-39</sup> In all the calculations (full-CI, CAS-SD, and MR-CC), only the valence electrons are correlated (frozen core approximation).

#### A. Approximations

##### 1. $\hat{T}_1^2 \times \hat{T}_2$ , $\hat{T}_1^3$ , and $\hat{T}_1^4$

To estimate the errors due to the approximate treatment of the  $\hat{T}_1^2 \times \hat{T}_2$ ,  $\hat{T}_1^3$ , and  $\hat{T}_1^4$  operators, we chose a single-reference example in which the single excitations are important at the CISD level. In the single reference case, all the excitations are of the 2 hole-2 particle type, so the normal equations (Eq. (14)) are solved exactly and all the values of  $\mu$  are equal to 1. The only difference to standard CCSD is the approximate treatment of the  $\hat{T}_1^3$  and  $\hat{T}_1^4$ , as explained in Section II D.

We have calculated the energy of the FH molecule at a distance of 1.2 Å with the single-reference CCSD programs of GAMESS<sup>40</sup> and Gaussian 09,<sup>41</sup> and our  $\mu$ -MR-CCSD implementation using the Hartree-Fock determinant as reference. The results are presented in Table I. The Hartree-Fock and

TABLE I. Comparison of the single-reference energies obtained with Gaussian 09, GAMESS, and the Quantum Package for FH at 1.2 Å, cc-pVDZ. All energies are converged below  $10^{-10}$  a.u.

	Hartree-Fock	CISD	CCSD
Gaussian 09	-99.959 526 039	-100.170 216 059	-100.178 425 609
GAMESS	-99.959 526 065	-100.170 216 086	-100.178 425 629
Quantum package	-99.959 526 065	-100.170 216 097	-100.178 426 538

CISD energies agree up to  $10^{-8}$  a.u., but our implementation differs from the CCSD by almost  $\sim -10^{-6}$  a.u. We attribute this difference to the approximation in the  $\hat{T}_1^2 \times \hat{T}_2$ ,  $\hat{T}_1^3$ , and  $\hat{T}_1^4$  operators, and it represents a relative error of  $4.4 \times 10^{-6}$  on the correlation energy.

To measure the effect of this approximation on the size consistency, we have calculated the energy of the CH<sub>3</sub> radical, with C–H bond lengths of 1.103 Å and H–C–H bond angles of 107.69° in the 6-31G basis set. We have also calculated the energy of the dimer with an intermolecular distance of 100 Å. The active space of the monomer contains only the singly occupied orbital, and the dimer is an open shell singlet with a CAS(2,2) wave function. For the  $\mu$ -MR-CCSD calculation, we have constrained the  $\mu_i$  as in Eq. (33) or we have let it unconstrained. The results are given in Table II and show that the deviation to additivity of the energy is reduced by an order of magnitude going from the CAS-SD to the dressed CAS-SD, and by two orders of magnitude when including the  $\mu_i$  factors. The constraint on the  $\mu_i$  introduces a small error which is below  $10^{-4}$  a.u.

## B. Bond breaking

For all the applications we compare the dressed CAS-SD and  $\mu$ -MR-CCSD with the  $\lambda$ -MR-CCSD and the CAS-SD values. Results are also given using the reference-independent dressing of the CAS-SD CI matrix. All the applications are presented as energy differences with respect to the full-CI energy estimated by a CIPSI calculation with a second-order perturbative correction. The smallest and largest values of the CIPSI perturbative corrections along the curves are given in Table III. We empirically estimate the error to the FCI energy to be in the order of 10% of the largest contribution. For ethane and twisted ethylene, which have the largest perturbative corrections, we have performed a larger CIPSI calculation at the point with the largest PT2 contribution. For ethylene the PT2 contribution was reduced to  $-5.6 mE_h$ , but the total energy changed by only  $0.08 mE_h$ . In this case, we can consider that the CIPSI energy of ethylene is converged. In the case of ethylene, the

TABLE III. Second-order perturbative correction in the CIPSI calculations. Minimum and maximum values among all the points of the potential energy curve.

	$E_{PT2} (E_h)$	
	Smallest	Largest
C <sub>2</sub> H <sub>6</sub>	$-13.4 \times 10^{-3}$	$-17.8 \times 10^{-3}$
C <sub>2</sub> H <sub>4</sub> twisted	$-3.37 \times 10^{-3}$	$-9.86 \times 10^{-3}$
C <sub>2</sub> H <sub>4</sub>	$-2.41 \times 10^{-3}$	$-6.85 \times 10^{-3}$
F <sub>2</sub> $^3\Sigma_u^+$	$-0.31 \times 10^{-3}$	$-1.42 \times 10^{-3}$
F <sub>2</sub>	$-0.13 \times 10^{-3}$	$-0.47 \times 10^{-3}$
N <sub>2</sub>	$-61.1 \times 10^{-6}$	$-0.41 \times 10^{-3}$
BeH <sub>2</sub>	$-12.3 \times 10^{-6}$	$-35.4 \times 10^{-6}$
H <sub>2</sub> O	$-1.59 \times 10^{-6}$	$-69.1 \times 10^{-6}$
FH	$-0.23 \times 10^{-6}$	$-55.1 \times 10^{-6}$
LiF	$-0.17 \times 10^{-6}$	$-12.3 \times 10^{-6}$

largest calculation dropped the PT2 value to  $-6.7 mE_h$  and the total energies differed by  $1.4 mE_h$ . So we estimate that the CIPSI curve of ethane has an accuracy less than  $2 mE_h$  in total energy, and all the other curves have an accuracy below the  $mE_h$ . The non-parallelism errors (NPEs) of all the CIPSI curves are estimated with an error below the  $mE_h$ .

For the full series of compounds, Figure 1 shows the energy difference with respect to the full-CI along the reaction coordinate. Table IV summarizes the non-parallelism errors (NPEs) and the maximum of the error obtained along the curve. The MR-CC treatment reduces the average and maximum error of the CAS-SD with respect to the full-CI by a factor close to 4. The correction is larger when the system involves an important number of inactive electrons (F<sub>2</sub>, C<sub>2</sub>H<sub>6</sub>) than when this number is small (BeH<sub>2</sub>, N<sub>2</sub>). One actually knows that the size-consistency error of the CAS-SD treatment increases with the number of inactive electrons; this error disappears in the MRCC treatment, which essentially misses some fourth-order connected effects of the triples.

### 1. Single-bond breaking

We present here the single bond breaking of the  $\sigma$  bonds of the C<sub>2</sub>H<sub>6</sub> and F<sub>2</sub> molecules and of the  $\pi$  bond of ethylene. The active spaces were chosen with two electrons in two MOs, the minimal wavefunctions to describe properly the dissociation of the molecules. In the case of ethane, the NPE of the CAS-SD is  $5.1 mE_h$ , and is reduced to  $3.5 mE_h$  with the  $\mu$ -MR-CCSD. The curve of the dressed CAS-SD has the lowest NPE ( $1.3 mE_h$ ). The curves obtained by both MR-CCSD methods give equivalent results, with NPEs of  $3.5$  and  $3.6 mE_h$ .

TABLE II. Evaluation of the size-consistency error in the dressed CAS-SD and the  $\mu$ -MR-CCSD, 6-31G basis set.

	CAS-SCF	CAS-SD	Dressed CAS-SD	$\mu$ -MR-CCSD $\mu \in \text{Eq. (33)}$	$\mu$ -MR-CCSD unconstrained $\mu$
·CH <sub>3</sub>	-39.528 586 5	-39.622 437 9	-39.625 570 2	-39.625 570 2	-39.625 570 2
·CH <sub>3</sub> × 2	-79.057 173 0	-79.244 875 8	-79.251 140 3	-79.251 140 3	-79.251 140 3
H <sub>3</sub> C · · · CH <sub>3</sub>	-79.057 173 0	-79.237 098 2	-79.250 695 7	-79.251 039 3	-79.251 107 3
Error		$7.78 \times 10^{-3}$	$4.44 \times 10^{-4}$	$1.01 \times 10^{-4}$	$3.29 \times 10^{-5}$

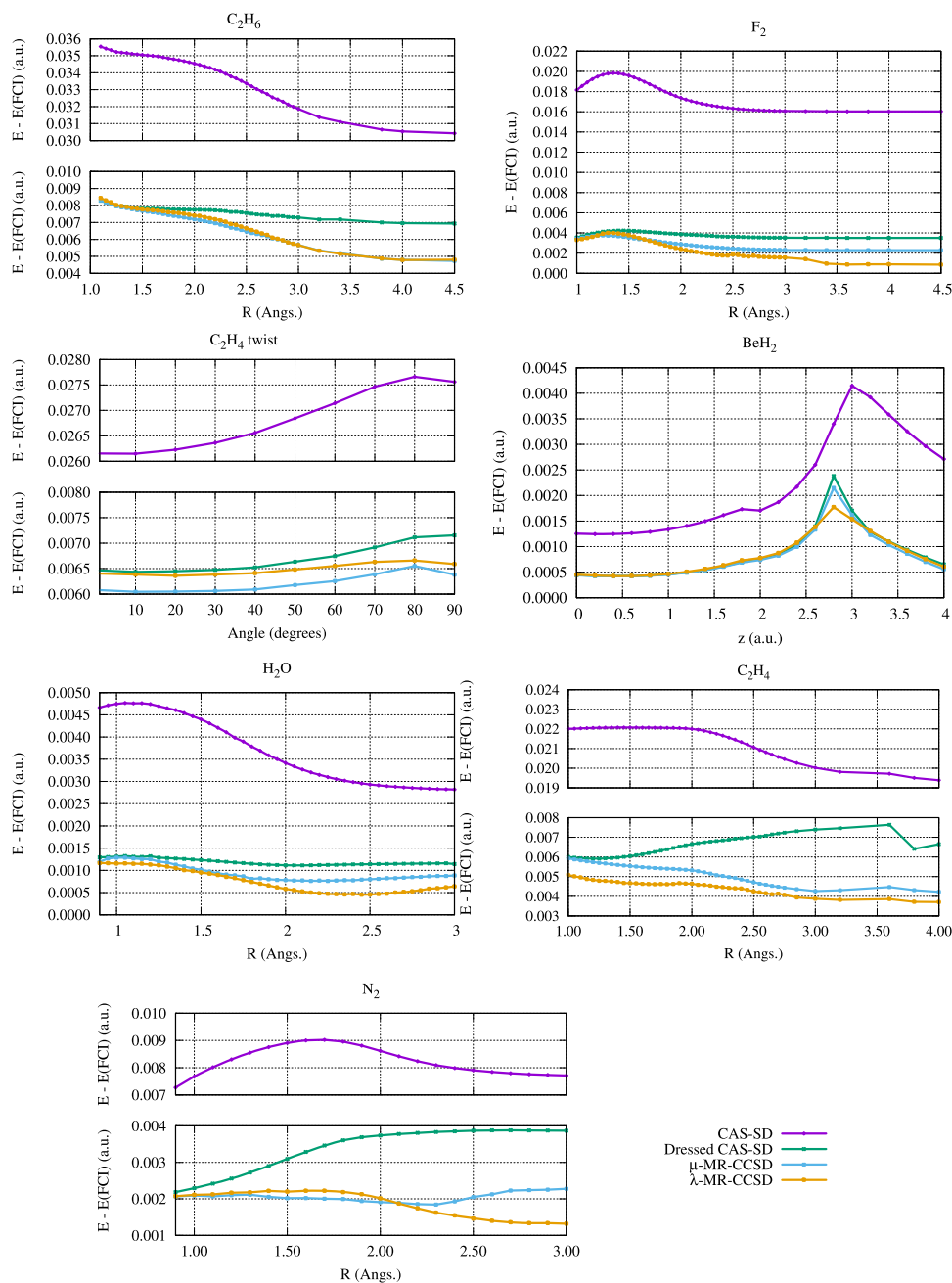


FIG. 1. Dissociation curves. Difference with respect to the full-CI energy using the MR-CCSD method presented in Ref. 17 and with the MR-CCSD method proposed in this work, as well as the CAS-SD and the dressed CAS-SD.

In the case of  $F_2$  the NPE of the dressed CAS-SD is  $0.9 mE_h$  and the NPE of the  $\mu$ -MR-CCSD is  $1.5 mE_h$ ; both are better than the NPE of the  $\lambda$ -MR-CCSD which is  $3.1 mE_h$ . Also, one can remark here some numerical instabilities in the  $\lambda$ -MR-CCSD where the curve is not smooth.

In the next example, the  $\pi$  bond of ethylene is broken by the rotation of the  $CH_2$  fragments. The CAS-SD has an NPE of  $1.5 mE_h$ , and using the dressed CAS-SD reduces the NPE to  $0.7 mE_h$ . The  $\mu$ -MR-CCSD gives an NPE of  $0.5 mE_h$ , and the NPE obtained with the  $\lambda$ -MR-CCSD is slightly better with a value of  $0.3 mE_h$ .

## 2. Insertion of Be in $H_2$

We present the results obtained by the insertion of a beryllium atom into the  $H_2$  molecule, which is a popular benchmark for MR-CC methods. The reference is still a CAS(2,2) for comparison with the literature, even though this choice of

reference is not the most appropriate for a correct description of the reaction. The geometries are given by the relation

$$z = 2.54 - 0.46x \quad (\text{a.u.}), \quad (34)$$

where the beryllium atom is at the origin and the hydrogen atoms are at the coordinates  $(x, 0, \pm z)$ . In this particular case, the  $\mu$ -MR-CCSD gives an NPE of  $1.7 mE_h$  which is larger than the NPE of  $1.3 mE_h$  obtained by the  $\lambda$ -MR-CCSD. This is due to only one point of the curve, the maximum which is higher by  $0.3 mE_h$ ; all the other points being very close by less than  $0.1 mE_h$ . Here, the dressed CAS-SD and the  $\mu$ -MR-CCSD are equivalent.

## 3. Two bond breaking

For breaking two bonds we have used CAS(4,4) wave functions as the reference space. The first example is the simultaneous breaking of the two O-H bonds of the water



TABLE IV. Non-parallelism errors (NPEs) and maximum errors with respect to the full-CI potential energy surface ( $mE_h$ ).

	CAS-SD		$\lambda$ -MR-CCSD		Dressed CAS-SD		$\mu$ -MR-CCSD	
	NPE	Max error	NPE	Max error	NPE	Max error	NPE	Max error
C <sub>2</sub> H <sub>6</sub>	5.1	35.5	3.6	8.4	1.3	8.3	3.5	8.3
F <sub>2</sub>	3.8	19.8	3.1	4.0	0.9	4.2	1.5	3.8
C <sub>2</sub> H <sub>4</sub> twist	1.5	27.7	0.3	6.7	0.7	7.1	0.5	6.5
BeH <sub>2</sub>	2.9	4.1	1.3	1.8	2.0	2.4	1.7	2.1
H <sub>2</sub> O	1.9	4.8	0.7	1.2	0.2	1.3	0.5	1.3
C <sub>2</sub> H <sub>4</sub> stretch	2.7	20.0	1.6	5.2	1.7	6.2	1.8	6.0
N <sub>2</sub>	1.7	9.0	0.9	2.2	1.7	3.8	0.3	2.3
F <sub>2</sub> $^3\Sigma_u^+(m_s = 1)$	2.6	18.6	1.3	3.3	1.2	3.5	1.2	3.3
F <sub>2</sub> $^3\Sigma_u^+(m_s = 0)$	2.6	18.6	1.2	1.8	1.3	3.5	1.1	3.3
FH (ground state)	2.6	14.6	1.8	4.0	2.1	4.5	1.8	4.0
FH (excited state)	3.3	20.9	8.8	8.5	10.5	10.1	7.1	8.3
F <sub>2</sub> (local)	3.8	19.8	1.2	3.2	1.5	3.5	0.9	3.0
N <sub>2</sub> (local)	1.7	9.0	3.6	5.0	1.1	3.5	0.4	1.8

molecule by stretching. Here, the CAS-SD exhibits an NPE of  $1.9 mE_h$  which is significantly improved to  $0.2 mE_h$  with the dressed CAS-SD. The  $\mu$ -MR-CCSD, with an NPE of  $0.5 mE_h$ , is slightly more parallel to the full-CI curve than the  $\lambda$ -MR-CCSD which has an NPE of  $0.7 mE_h$ .

The second example is the double-bond breaking of ethylene by stretching. One should first clarify that the energy differences in the figure do not match those of the torsion along the bond because in the former example the reference was a CAS(2,2), and here it is a CAS(4,4). Dressing the CAS-SD reduces the NPE from  $2.7 mE_h$  to  $1.7 mE_h$ . One can remark a discontinuity in the curve at large distances. The  $\mu$ -MR-CCSD and  $\lambda$ -MR-CCSD slightly improve the NPE to values of  $1.7 mE_h$  and  $1.8 mE_h$ .

#### 4. Triple-bond breaking

N<sub>2</sub> is the typical benchmark for breaking a triple bond. Here, we have used a CAS(6,6) reference wave function. At the CAS-SD level, the NPE is  $1.7 mE_h$ , and the dressed CAS-SD does not reduce the NPE. Here, it is necessary to use reference-dependant amplitudes to recover a low NPE:  $0.9 mE_h$  with the  $\lambda$ -MR-CCSD, and  $0.3 mE_h$  with the  $\mu$ -MR-CCSD.

### C. Excited states

#### 1. Triplet state of F<sub>2</sub>

We report here calculations on the triplet state  $^3\Sigma_u^+$  of F<sub>2</sub>. The reference wave function was prepared in two different ways, both using restricted open-shell Hartree-Fock molecular orbitals. The first reference wave function labeled  $m_s = 1$  is a single open-shell determinant, and the second wave function is the triplet  $m_s = 0$ , made of two determinants  $1/\sqrt{2}(\alpha\beta - \beta\alpha)$ .

To ensure that the CAS-SD is a strict eigenfunction of the  $\hat{S}^2$  operator, we have included in  $\Psi_{SD}$  all the determinants with the same space part as the singles and doubles with respect to the CAS. These determinants are treated in the same way as singles and doubles and are treated variationally in the diagonalizations. Of course, those which are triples or quadruples

with respect to  $\Psi_{ref}$  are excluded from the set of the  $\{\alpha\}$  and have no effect in the dressing.

To reduce the computational cost, the triples and quadruples were not augmented with all the determinants with the same space part. The absence of some determinants gives rise to a slight deviation ( $<10^{-6}$  a.u.) of  $\langle\hat{S}^2\rangle$  from the desired eigenvalue, and it is expected to have some impact on the iterative dressing. It is worth checking the effect of this deviation from the exact spin multiplicity. The first test concerns the comparison of the  $m_s = 0$  and  $m_s = 1$  components of a triplet state.

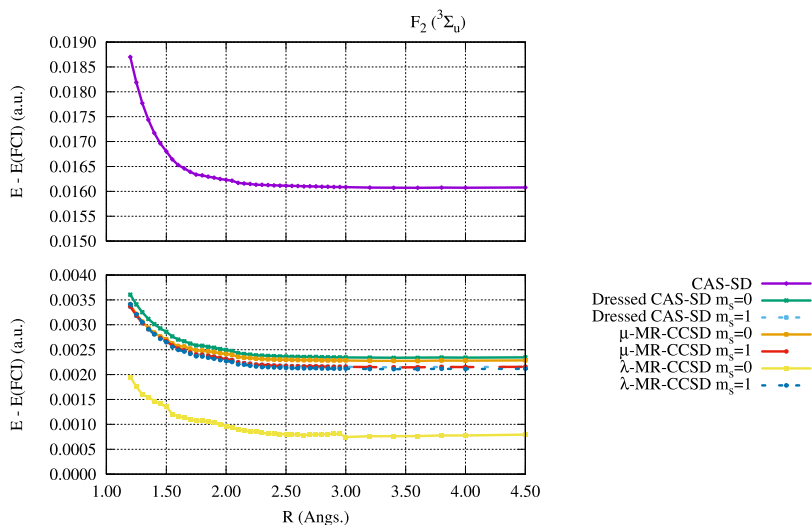
According to Figure 2, in all the cases, the NPE of the CAS-SD ( $2.6 mE_h$ ) is improved to a value of  $1.1$ – $1.3 mE_h$ . As expected the dressed CAS-SD and the  $\mu$ -MR-CCSD are strictly equivalent for  $m_s = 1$ . Indeed, for both variants, the usual single-reference amplitudes  $c_i/c_0$  are recovered. The amplitudes of the  $\lambda$ -MR-CCSD lower the curve by  $1 mE_h$  when going from  $m_s = 1$  to  $m_s = 0$ . The dressed CAS-SD gives a slightly higher energy by only  $0.3 mE_h$ , and introducing the reference-dependence via the  $\mu_i$  reduces the difference to  $0.2 mE_h$ .

If one considers the error on the singlet-triplet gap with respect to the full-CI reference, it appears clearly that the  $\mu$ -MR-CCSD with  $m_s = 0$  gives the most accurate results, with errors lying between 0 and  $0.9 mE_h$  along the curve.

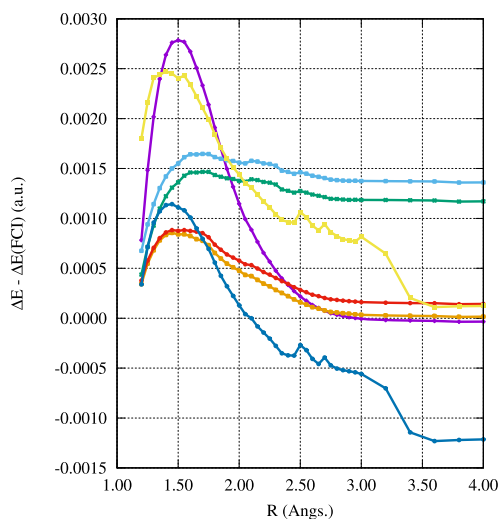
#### 2. Avoided crossing in FH and LiF

We have calculated the potential energy surfaces of the two lowest  $^1\Sigma^+$  states of FH, using as reference wave function the CAS(2,2) with state-averaged CAS-SCF molecular orbitals in the aug-cc-pVDZ basis set. Figure 3 shows the NPEs of the ground and excited states. In the ground state, the NPE is  $1.8 mE_h$  for both MR-CCSD variants, but the  $\lambda$ -MR-CCSD shows some numerical instabilities, as opposed to the  $\mu$ -MR-CCSD which gives a very smooth curve.

In the excited state, the situation is different: surprisingly the best NPE is obtained by the CAS-SD. The reason is the



CAS-SD —  
 Dressed CAS-SD  $m_s=0$  —  
 Dressed CAS-SD  $m_s=1$  - -  
 $\mu$ -MR-CCSD  $m_s=0$  —  
 $\mu$ -MR-CCSD  $m_s=1$  —  
 $\lambda$ -MR-CCSD  $m_s=0$  —  
 $\lambda$ -MR-CCSD  $m_s=1$  - -



CAS-SD —  
 Dressed CAS-SD  $m_s=0$  —  
 Dressed CAS-SD  $m_s=1$  - -  
 $\mu$ -MR-CCSD  $m_s=0$  —  
 $\mu$ -MR-CCSD  $m_s=1$  —  
 $\lambda$ -MR-CCSD  $m_s=0$  —  
 $\lambda$ -MR-CCSD  $m_s=1$  - -

FIG. 2.  $F_2 \ ^3\Sigma_u^+$ . Difference with respect to the full-CI energy for the  $m_s = 0$  and  $m_s = 1$  wave functions (top), and error on the singlet-triplet gap  $\Delta E = E(^3\Sigma_u^+) - E(^1\Sigma_g^+)$  (bottom). On both graphics, the two curves of the dressed CAS-SD coincide.

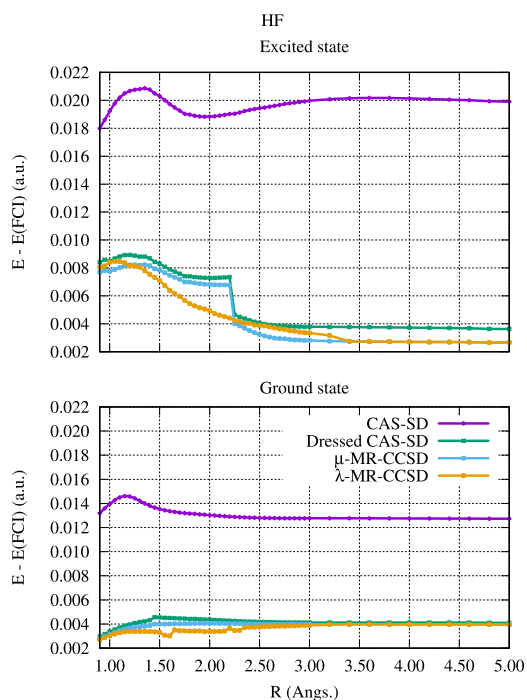


FIG. 3. Difference with respect to the full-CI energy for the two lowest  $^1\Sigma^+$  states of FH.

following. At large distances, the CAS-SD description is correct, but the size-consistency error raises the energy. At short distances, the CAS-SD description of the excited state is not as accurate as for large distances since there are determinants with a large coefficient which are not in the active space: the CAS contributes by 0.85 to the norm of the FCI wave function, but in the case of the CAS-SD its contribution is 0.91. This bad description of the CAS-SD raises also the energy at short distances, and the errors compensate along the curve. All the other methods correct the size-consistency error, so the long range errors are corrected but not the error due to the incompleteness of the CAS at short range. This explains why the deviations to the FCI decrease with the distance, and why the NPEs are so large. The two variants of the MR-CCSD agree at short and long distances, but they differ significantly between 1.5 and 2.5 Å; the region of the avoided crossing. To understand these differences, we have plotted in Figure 4 the two eigenvalues of the two state-specific Hamiltonians—one dressed for the ground state and one dressed for the excited state. It appears that the state of interest is very well described, but the dressing for other root has a much lower quality. This strong state-specific character is due to the fitting procedure which is implicitly weighted by the state of interest. The large coefficients have a higher quality in the amplitudes, but the

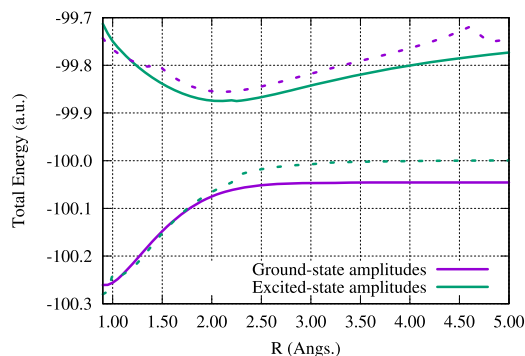


FIG. 4. Potential energy surfaces of the two lowest  $1\Sigma^+$  states of FH with the  $\mu$ -MR-CCSD method. The energy of the state corresponding to the dressing is plotted in plain curves, and the energy of the other state is plotted in dashed curves.

important determinants of the second root are usually not the same as in the state of interest. The  $\lambda$ -MR-CCSD has amplitudes which depend less on the wave function, so the quality is comparable on both states, and the choice of these amplitudes is better suited for calculating excited states within the same symmetry, as will be confirmed by the next example.

In Figure 5 we have represented the avoided crossing of LiF, also calculated with the aug-cc-pVDZ basis set. The physical situation is similar to FH, but the energy difference between the ground and the excited states is much smaller. A striking result is that the  $\lambda$ -MR-CCSD, although being state-specific, is able to reproduce very well the whole potential energy surfaces of both states. The position of the avoided crossing is very well reproduced by the three methods: the CAS-SD crosses at 6.3 Å, the full-CI crosses at 6.8 Å, and the dressed CAS-SD and the two MR-CCSD variants cross at 6.9 Å. The  $\mu$ -MR-CCSD and  $\lambda$ -MR-CCSD coincide in the short-range ( $\leq 5$  Å) and in the long range ( $\geq 7.2$  Å), but when the two states become very close in energy in the region of the crossing the dressed CAS-SD and the  $\mu$ -MR-CCSD are unable to give reasonable values. This disappointing result motivates a future work on a multi-state  $\mu$ -MR-CCSD.

#### D. Sensitivity to the choice molecular orbitals

The  $\mu$ -MR-CCSD algorithm we propose is in the Jeziorski-Monkhorst framework, so it is not invariant with respect to the choice of molecular orbitals. In this section, we checked its sensitivity to the choice of the MO set by

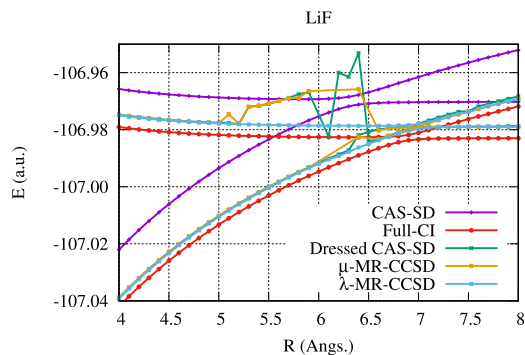


FIG. 5. Potential energy surfaces of the two lowest  $1\Sigma^+$  states of LiF.

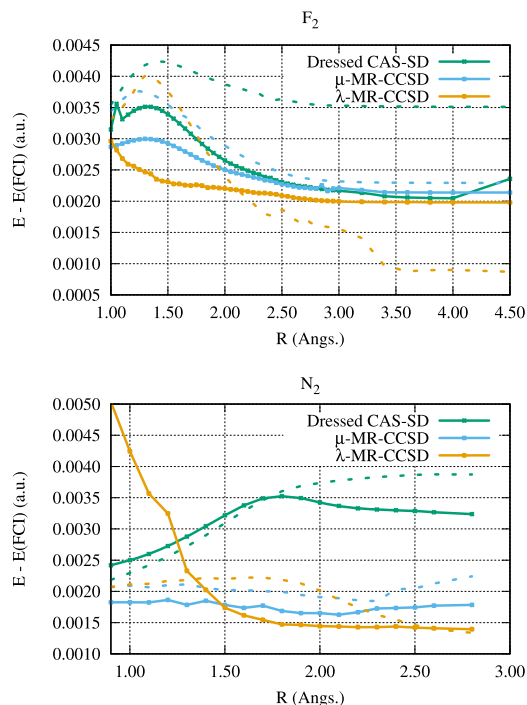


FIG. 6. Comparison between pseudo-canonical (dashed curves) and localized (plain curves) MOs in  $F_2$  and  $N_2$ . Difference with respect to the full-CI energy.

comparing results obtained with pseudo-canonical CAS-SCF orbitals and with localized MOs in the  $F_2$  and  $N_2$  molecules (Figure 6).

In the  $F_2$  molecule, using localized MOs is a better choice than the pseudo-canonical MOs. The best NPE is obtained by the  $\mu$ -MR-CCSD method with a value of 0.9  $mE_h$ . In the case of  $N_2$ , the situation is different: the NPE of the  $\lambda$ -MR-CCSD goes from 0.9  $mE_h$  to 3.6  $mE_h$ , and the NPE of the  $\mu$ -MR-CCSD is relatively stable around 0.3–0.4  $mE_h$ . On the other hand, the dressed CAS-SD gives a better NPE with local orbitals, going from 1.7  $mE_h$  to 1.1  $mE_h$ .

The fact that the  $\mu$ -MR-CCSD is less sensitive to the MO set than the  $\lambda$ -MR-CCSD can be understood. By changing the MO set, a single excitation rotates into a combination of single and double excitations. In the  $\lambda$ -MR-CCSD method, the amplitudes are calculated by taking into account the matrix elements of the Hamiltonian, which are of different nature depending on the degree of excitation, so the amplitudes are expected to change significantly. In the  $\mu$ -MR-CCSD variant, the amplitudes are adjusted in such a way that they fit the CAS-SD wave function, which is invariant by rotation of the MOs. Therefore, it is expected to be more robust with respect to the change of MO set.

#### IV. CONCLUSIONS

We have proposed a method to determine reference-independent amplitudes by fitting the CAS-SD CI vector. These amplitudes may be used to perform a state-specific iterative dressing of the CAS-SD Hamiltonian in order to take into account the effect of the triples and quadruples in the spirit of the coupled cluster formalism. Alternatively, these amplitudes may be rescaled to reproduce the exact coefficients of the singles and doubles to introduce a reference-dependent character.

In that case, the CAS-SD CI vector is recovered by the application of  $(1 + \hat{T})$  on the reference wave function, so we reach here the Jeziorski-Monkhorst coupled cluster formalism.

The CAS-SD dressed with reference-independent amplitudes gives excellent results for single-bond breaking ( $F_2$  and ethane) and the simultaneous breaking of the two O–H bonds of water, with a non-parallelism error lower than the milli-Hartree. When the active space becomes larger, it is necessary to go to the reference-dependent MR-CCSD introducing the  $\mu$  factors in Eq. (31). In the case of ethylene and  $N_2$ , this keeps the NPE to a value close to the milli-Hartree.

We have shown numerically that the here-proposed amplitudes are not very sensitive to the value of  $m_s$  for open-shell systems, and to the choice of the molecular orbitals. This is clearly an improvement compared the amplitudes proposed earlier.<sup>17</sup> But we have also shown that the former amplitudes are a better choice when computing excited states of the same symmetry because the here-proposed amplitudes have a much more pronounced state-specific character which may be disadvantageous if the states are too close in energy. This problem can be cured by leaving the state-specific formalism for a multi-state formalism,<sup>42</sup> and this will be the object of a future work.

## ACKNOWLEDGMENTS

This work has been made through generous computational support from CALMIP (Toulouse) under the Allocation No. 2015–0510, and GENCI under the Allocation No. x2015081738. We are also very grateful to the two anonymous reviewers for their comments, which helped us improve the manuscript significantly.

- <sup>1</sup>F. Coester, *Nucl. Phys.* **7**, 421 (1958).
- <sup>2</sup>F. Coester and H. Kümmel, *Nucl. Phys.* **17**, 477 (1960).
- <sup>3</sup>J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966).
- <sup>4</sup>R. J. Bartlett, J. Watts, S. Kucharski, and J. Noga, *Chem. Phys. Lett.* **165**, 513 (1990).
- <sup>5</sup>J. Goldstone, *Proc. R. Soc. A* **239**, 267 (1957).
- <sup>6</sup>I. Nebot-Gil, J. Sánchez-Mariñ, J. P. Malrieu, J. L. Heully, and D. Maynau, *J. Chem. Phys.* **103**, 2576 (1995).
- <sup>7</sup>J. P. Malrieu, P. Durand, and J. P. Daudey, *J. Phys. A: Math. Gen.* **18**, 809 (1985).
- <sup>8</sup>B. Kirtman, *J. Chem. Phys.* **75**, 798 (1981).
- <sup>9</sup>D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, *Chem. Rev.* **112**, 182 (2012).
- <sup>10</sup>N. Oliphant and L. Adamowicz, *J. Chem. Phys.* **96**, 3739 (1992).
- <sup>11</sup>X. Li and J. Paldus, *J. Chem. Phys.* **108**, 637 (1998).
- <sup>12</sup>U. S. Mahapatra, B. Datta, and D. Mukherjee, *Mol. Phys.* **94**, 157 (1998).
- <sup>13</sup>S. Das, D. Mukherjee, and M. Kállay, *J. Chem. Phys.* **132**, 074103 (2010).
- <sup>14</sup>A. Szabados, *J. Chem. Phys.* **134**, 174113 (2011).
- <sup>15</sup>U. S. Mahapatra, B. Datta, and D. Mukherjee, *J. Chem. Phys.* **110**, 6171 (1999).
- <sup>16</sup>J. Meller, J. P. Malrieu, and R. Caballol, *J. Chem. Phys.* **104**, 4068 (1996).
- <sup>17</sup>E. Giner, G. David, A. Scemama, and J. P. Malrieu, *J. Chem. Phys.* **144**, 064101 (2016).
- <sup>18</sup>B. Jeziorski and H. J. Monkhorst, *Phys. Rev. A* **24**, 1668 (1981).
- <sup>19</sup>F. A. Evangelista and J. Gauss, *J. Chem. Phys.* **134**, 114102 (2011).
- <sup>20</sup>M. Hanauer and A. Köhn, *J. Chem. Phys.* **136**, 204107 (2012).
- <sup>21</sup>U. S. Mahapatra, B. Datta, B. Bandyopadhyay, and D. Mukherjee, *Advances in Quantum Chemistry* (Elsevier BV, 1998), pp. 163–193.
- <sup>22</sup>L. F. Richardson, *Philos. Trans. R. Soc., A* **210**, 307 (1911).
- <sup>23</sup>G. Golub and W. Kahan, *J. Soc. Ind. Appl. Math. Ser. B Numer. Anal.* **2**, 205 (1965).
- <sup>24</sup>A. Engels-Putzka and M. Hanrath, *J. Mol. Struct.: THEOCHEM* **902**, 59 (2009).
- <sup>25</sup>G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- <sup>26</sup>G. D. Purvis, R. Shepard, F. B. Brown, and R. J. Bartlett, *Int. J. Quantum Chem.* **23**, 835 (1983).
- <sup>27</sup>W. D. Laidig and R. J. Bartlett, *Chem. Phys. Lett.* **104**, 424 (1984).
- <sup>28</sup>M. Kállay, P. G. Szalay, and P. R. Surján, *J. Chem. Phys.* **117**, 980 (2002).
- <sup>29</sup>M. Hanrath, *J. Chem. Phys.* **123**, 084102 (2005).
- <sup>30</sup>F. A. Evangelista, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **125**, 154113 (2006).
- <sup>31</sup>F. A. Evangelista, E. Prochnow, J. Gauss, and H. F. Schaefer, *J. Chem. Phys.* **132**, 074107 (2010).
- <sup>32</sup>M. Hanauer and A. Köhn, *J. Chem. Phys.* **134**, 204111 (2011).
- <sup>33</sup>Y. A. Aoto and A. Köhn, *J. Chem. Phys.* **144**, 074103 (2016).
- <sup>34</sup>T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>35</sup>M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.* **14**, 1347 (1993).
- <sup>36</sup>A. Scemama, T. Applencourt, Y. Garniron, E. Giner, G. David, and M. Caffarel, Quantum package v1.0, 2016, <https://github.com/LCPQ/quantum.package>.
- <sup>37</sup>B. Huron, P. Rancurel, and J. Malrieu, *J. Chem. Phys.* **58**, 5745 (1973).
- <sup>38</sup>S. Evangelisti, J. Daudey, and J. Malrieu, *Chem. Phys.* **75**, 91 (1983).
- <sup>39</sup>E. Giner, A. Scemama, and M. Caffarel, *Can. J. Chem.* **91**, 879 (2013).
- <sup>40</sup>P. Piecuch, S. A. Kucharski, K. Kowalski, and M. Musiał, *Comput. Phys. Commun.* **149**, 71 (2002).
- <sup>41</sup>M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, GAUSSIAN 09, Revision D.01, Gaussian, Inc., 2009.
- <sup>42</sup>J.-P. Malrieu, *Mol. Phys.* **111**, 2451 (2013).