Spin-component-scaling second-order Møller–Plesset theory and its variants for economical correlation energies: Unified theoretical interpretation and use for quartet N_3

A. J. C. Varandas[a)]

Departamento de Química, Universidade de Coimbra, 3004-535 Coimbra, Portugal

The spin-component-scaling second-order Møller–Plesset theory proposed by Grimme, the scaled opposite-spin variant of Head-Gordon and co-workers, and other variants of the theory to treat the electron correlation energy are examined. A refinement of scaled opposite-spin theory for strong chemical interactions is suggested where the scaled correlation contribution is chosen such as to mimic closely the one obtained by more sophisticated methods of the coupled cluster type. With the scaling factor chosen to vary in a simple statistical manner with the number of opposite-spin electron pairs of the system, the parameters have been calibrated from standard coupled cluster type calculations for a chosen *ab initio* test data set. The new approach, termed as variable-scaling opposite spin, aims to be applicable at any regions of the molecule configuration space where second-order Møller–Plesset perturbation theory converges. It thus benefits of all advantages inherent to the original theory, which makes it an attractive approach on a computational cost basis. Because the method in one of its formats fails size-extensivity, the consequences and remedies of this are analyzed. Illustrations are presented for many molecules utilizing Dunning-type basis sets, in particular, for a detailed analysis of N_3 in its lowest quartet state, which does not belong to the test set. Extrapolations of the calculated raw energies to the complete one-electron basis set limit are also reported, giving the most reliable estimates available thus far of the energetics for the $N(^4S)$ $+N_2$ exchange reaction. All spin-component-scaling schemes are known to show difficulties in dealing with weak interactions of the van der Waals type, which has justified the design of specific variants of the theory according to the property and regime of interactions. Several variants of the theory are then examined using a second test set of molecules, and shown to be linked via a coordinate that evolves gradually between two known extreme regimes. It is further shown that such a coordinate can be specified via a constrained Feenberg-type scaling approach, a theory whose merits are also explored. © *2010 American Institute of Physics*. [doi:10.1063/1.3465551]

I. INTRODUCTION

Electronic structure calculations play a key role on the prediction of the structure and reactivity of molecules and materials. Indeed, no doubt exists that having such calculations at an accurate, yet affordable, level is ubiquitous and a challenge to modern computational chemistry since the benefits can be enormous ranging from gas phase chemistry to biochemistry to materials science. In this regard, density functional theory[1,2] (DFT) has become the popular electronic structure tool for application to systems with large numbers of electrons. However, DFT is known to lack an *a priori* treatment of the dispersion interactions and suffer from spurious electron self-interaction. In fact, DFT methods are known to be somewhat suspect for reaction barriers as standard functionals tend to underestimate activation energies,[3] largely as a consequence of the self-interaction error.[4,5] Moreover, a surprising shortcoming to correctly predict bond dissociation energies has been reported,[6] in addition to other difficult to correct errors.[7]

The simplest molecular orbital electronic structure method alternative to DFT that can correctly treat dispersion and hydrogen-bonding interactions is second-order Møller–Plesset[8] (MP2) theory. Being size-consistent,[9] it can accurately treat long-range dispersion interactions, as well as the dispersion, polarization, and covalency effects associated with hydrogen bonding.[10,11] Moreover, recent progress has been made that enhances the success of MP2 computations by reducing further its low cost relative to the other standard MO algorithms. Regarding the steep cost increase of the MP2 calculations with molecular size, some well recognized developments are (a) methods that reduce the prefactor without changing the underlying scaling, such as "resolution-of-the-identity" methods[12,13] or the pseudospectral approach,[14] and others;[4] (b) methods that attempt to exploit "underlying locality" in the MP2 problem;[15,16] (c) efficient molecular orbital localization schemes;[12,13] (d) methods that exploit the locality of electronic structure by ansatz;[17–19] (e) combination of the above methods[20,21] such as resolution-of-the-identity/density fitting ones. Of course, there is still the need for using large atomic orbital basis sets in order to obtain good results,[22] which can further reduce the upper limit on system size and originate a diminished performance for open-shell systems.[23,24] Such a situation has been itself sub-

a)Electronic mail: varandas@qtvs1.qui.uc.pt.

stantially ameliorated in recent years by the appearance of extrapolation methods that can reliably predict the molecular energy at the complete basis set (CBS) limit. Indeed, it is now well established that MP2 correlation energies (or any wave function based correlation energy) can be systematically extrapolated toward the CBS limit by using the X^{-3} behavior of the basis set error with respect to the hierarchic number X of the Dunning[25–27] correlation-consistent polarized valence (cc-pVXZ or VXZ) basis sets, which are often augmented with diffused functions (aug-cc-pVXZ or AVXZ); X is commonly referred as the cardinal number. In fact, a refinement of this approach suggests[28,29] separate scaling of the same-spin (abbreviated as SS or T from "triplet" or numerically by spin 1) and opposite-spin (OS:S:0) correlation energies. Recall that the former actually converges as X^{-5} while the latter, numerically far larger, converges as X^{-3}. It should be added that we have suggested[30] a scheme, and later a generalized variant of it,[30,31] whereby such contributions are uniformly treated. The approach, known as the uniform singlet- and triplet-pair extrapolation (USTE) scheme,[30] has shown a high reliability being employed later to CBS extrapolate some of the raw energies reported in the present work.

The distinct X^{-n} dependence of the two spin cases noted in the previous paragraph may have inspired Grimme[32–34] to suggest that MP2 energies can be systematically improved by separate scaling of the OS and SS components of the MP2 correlation energy. This led to the "spin-component scaled" MP2 theory, or simply SCS-MP2. Grimme's idea was further explored by Jung *et al.*[35] who suggested a simplified variant of SCS-MP2 theory. By noting that the damping of the SS contribution is large (1/3), they proposed that results of comparable quality can be obtained by scaling just the OS component (i.e., ignoring the SS component). This offers the possibility of a significantly reduced computational cost for larger systems as it is possible to formulate SOS-MP2 as a fourth-order (rather than fifth-order as for MP2) method. It also offers other desirable practical implications for the efficiency of implementation since many of the algorithmic complications that arise in fast MP2 methods are associated with the exchange contribution to the SS correlation. Both SCS-MP2 and SOS-MP2 theories have shown impressive accuracy when calculating reaction energies, atomization energies, dissociation energies, molecular geometries, and barrier heights.[6,32,35] Unfortunately, the values of the optimal scaling parameters in SCS-MP2 and SOS-MP2 theories have been shown to vary drastically with the property of interest and the atomic orbital basis set that is employed. For example, they are found to vary drastically when stepping from strong to intermolecular interactions.[36–38] This led Head-Gordon and co-workers to the proposition of a modified-OS MP2 theory[36] in an attempt to correct the long-range domain of SOS-MP2 theory, and to develop an efficient linear scaling local SOS-MP2 algorithm that were utilized to compute the intermolecular interaction of fullerene and porphyrin.[37] In fact, both such developments have shown the ability to accurately describe nonbonding intermolecular interactions by the single-parameter scaling SOS-MP2 formalism. Other recent analyzes by Distasio and Head-Gordon,[38] and Hill and

Plats[39] also stressed that the set of optimal parameters for the above theories vary drastically with the physical property to be modeled. For some, SCS-MP2 performs best, for others is SOS-MP2, and sometimes even scaling the same-spin component (SSS-MP2) turns out to perform well.

In addition to low computer cost, a method in computational chemistry should ideally show size-extensivity and size-consistency, be applicable to both closed- and open-shell systems, and hopefully show uniform accuracy over the molecule configuration space. Although such prerequisites will be further examined below, suffice it to recall here that size-extensivity requires that the total energy of N infinitely separated systems is N times the energy of each system. Although such a property is known to apply to MP2 (and is retained[34] by SCS-MP2 and SOS-MP2), we will examine later how it can be preserved in the variable-scaling opposite-spin second-order Møller–Plesset (VOS-MP2) variant here suggested. Thus, the correct energy behavior of the system (size-consistency) can be guaranteed when the interaction between closed-shell molecules is nullified (such as when the intermolecular distance increases to its asymptotic value) since the Hartree–Fock (HF) wave function (also size-extensive) will then be a good starting point. This is especially important for large systems and weakly interacting fragments although, as noted above, SCS-MP2 theory and its variants are known to perform relatively poorer for such interactions. On the other hand, for strong interactions involving bond-breaking/formation, the HF wave function is often far from dominant and single-reference post-Hartree–Fock methods known to fail frequently to produce converged results. Thus, it appears that no strong reason exists to impose strict size-extensivity into the method even though a fair amount of configuration space may be expected to be sampled. Instead, its predictive value might seem to provide a fairer criterion for adoption. Despite this, the new theory will be shown to assume, if desired, a fully size-extensive format.

The aim of this work is twofold. First, an improvement of SOS-MP2 theory is suggested whereby a parametrization of scaling that varies with system size is envisaged. Specifically, a scaling factor is suggested that accounts for the dependence on the number of OS electron pairs in a simple statistical manner. The calibration is made such as to reproduce an *ab initio* test set of energies calculated with CC type methods, namely, CCSD(T) including singles and doubles electronic excitations with a perturbation estimate of connected triples. Alternatively, the quadratic configuration interaction $[QCISD(T)]$ method may be employed, which gives similar results.[10,11] As note above, the new variant of SOS-MP2 theory will then be referred to by the acronym VOS-MP2. Second, a theoretical interpretation of the various spin-component-scaling theories is presented using a generalization of the approach suggested by Feenberg[40] and adopted more recently by Szabados.[41]

The structure of the paper is as follows. Section II provides a brief review of MP2 theory and spin-component-scaling methods. A description of the novel VOS-MP2 theory is then given in Sec. III. Section IV presents the results obtained for the ATS1 *ab initio* test data set. A rationalization of the existing spin-component-scaling theories using

a variant of the Feenberg perturbation theory approach follows in Sec. IV, jointly with illustrative results for another *ab initio* test data set (ATS2) which is part of the one employed in Ref. 41. The major conclusions are gathered in Sec. V.

II. MP2 THEORY AND SPIN-COMPONENT ENERGY SCALING

Let the zeroth-order Hamiltonian be written as

$$H^{(0)} = E_0|0\rangle\langle 0| + \sum_{K\neq 0} E_K|K\rangle\langle K|, \qquad (1)$$

with the perturbation being defined in a common notation by $W = H - H^{(0)}$, and E_K and $|K\rangle$ denoting the eigenvalues and eigenfunctions of $H^{(0)}$. Within standard MP2 theory, the second-order energy assumes the form

$$E_{MP}^{(2)} = -\sum_{K\neq 0} \frac{|\langle 0|H|K\rangle|^2}{\Delta_K}, \qquad (2)$$

where $\Delta_K = E_K - E_0$. Thus, provided that the HF problem is solved, only doubly excited determinants will contribute to the second-order energy. These can either involve parallel spin excitations of the type $|T_K\rangle = b_\sigma^+ a_\sigma^+ i_\sigma^- j_\sigma^-|0\rangle$ or antiparallel-spin excitations $|T_K\rangle = b_\sigma^+ a_{\bar\sigma}^+ i_\sigma^- j_{\bar\sigma}^-|0\rangle$, with σ and $\bar\sigma$ specifying orthogonal spin functions, $i, j, \ldots$ referring to occupied indices, and $a, b, \ldots$ to virtual ones. For closed-shell systems, the excitation energy denominators are independent of the spin labels, being given by $\Delta_K = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j$, where the ϵ's stand for the canonical HF molecular orbital energies. The total correlation energy can therefore be split into SS [or ($\alpha\alpha + \beta\beta$) or triplet] electron-pair contributions and OS ($\alpha\beta$ or singlet) contributions. Thus,

$$E_{MP}^{(2)} = -\sum_{T_K} \frac{|\langle 0|H|T_K\rangle|^2}{\Delta_K} - \sum_{S_K} \frac{|\langle 0|H|S_K\rangle|^2}{\Delta_K}, \qquad (3)$$

where, for the MP partitioning, $|0\rangle$ denotes the HF wave function and $|K\rangle$ refers to excited determinants. In terms of two-electron integrals, the result is

$$\begin{aligned} E_{MP}^{(2)} &= E_{MP2}^T + E_{MP2}^S \\ &= -\sum_{ijab} \left(\frac{\langle ij|ab\rangle\langle ij\|ab\rangle}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} + \frac{\langle ij|ab\rangle^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \right), \end{aligned} \qquad (4)$$

where the usual notation $\langle ij\|ab\rangle = \langle ij|ab\rangle - \langle ij|ba\rangle$ has been utilized.

At the MP1 (or HF) level, the Fermi correlation between spin-parallel electron pairs is already considered while the Coulomb correlation between spin antiparallel pairs is completely neglected, a biased description that continues at higher levels of perturbation theory. This leads at second-order (MP2) to an unbalanced description of dynamic (short-range) versus static (long-range) electron correlation effects. The scaling procedure aims therefore, in particular, to reduce the often overestimated correlation of the SS contribution. This involves mainly distant electrons, which represent a significant part of the static electron correlation. Grimme's scaling approach assumes then the form

$$E_{SCS-MP2} = -p_T\sum_{T_K} \frac{|\langle 0|H|T_K\rangle|^2}{\Delta_K} - p_S\sum_{S_K} \frac{|\langle 0|H|S_K\rangle|^2}{\Delta_K}, \qquad (5)$$

hence leading as noted above to SCS-MP2 theory. Note that p_T and p_S denote scaling factors of the modified correlation energy, which can help to effectively reduce the outliers in standard MP2 calculations. Clearly, this immediately reminds one that any excessive reduction of the SS interactions may have critical implications when dealing with properties dominated by subtle long-range interactions. The scaling parameters used by Grimme are $p_S = 6/5$ for the OS correlation and $p_T = 1/3$ for SS. For a given (large) basis, Grimme's approach showed clear statistical improvements in the quality of geometries (of diatomics), and a wide range of relative energies of reactions, and atomization energies.[32–34] Efficient implementations of orbital-optimized MP2 theory within the RI and their impact on the SCS-MP2 method have also been explored.[42] In a similar fashion, improvements of coupled cluster singles and doubles (CCSD) theory by scaling the SS and OS components of the double excitation correlation energy were reported by Sherrill and co-workers.[43]

Grimme's idea was further explored by Head-Gordon and co-workers[35] who, noting that the damping of the SS contribution is large ($p_T = 1/3$), suggested that results of comparable quality could be obtained by scaling just the OS component (i.e., ignoring the SS components). As noted above, this has desirable practical implications for efficiency of implementation since many of the algorithmic complications that arise in fast MP2 methods are associated with the exchange contribution to the SS correlation. In SOS-MP2 theory, the correlation energy assumes then the form

$$E_{SOS-MP2} = -p_S\sum_{S_K} \frac{|\langle 0|H|S_K\rangle|^2}{\Delta_K}, \qquad (6)$$

where $p_S = 1.2$ or 1.3 [we utilize the latter as it gives a somewhat smaller root mean squared deviation (rmsd) in the comparisons to be presented later]. The magnitude of this scaling coefficient has been fixed from the observation that the ratio of $S:T$ correlation is typically 3 or 4 to 1, and thus the OS scaling factor would requires to be increased by about $1/(3\times 3)$ to mimic the absence of explicit SS correlations. Note that scaled OS second-order perturbation corrections to single excitation configuration interaction have also been suggested to treat quasidegeneracies between excited states.[44–46]

A theoretical rationalization of Grimme's approach was given by Szabados[41] using a generalization of the Feenberg[40] scaling approach. By noting that the partitioning of the Hamiltonian into a zeroth-order term and a perturbation leaves some arbitrariness in defining the latter, Feenberg suggested to introduce one parameter to scale the zeroth-order Hamiltonian which could then be chosen by some appropriate criterion. Szabados[41] generalized the idea by introducing two spin-dependent excitation energies in the Hamiltonian. This assumes now the form

$$H_{\text{SCS-MP}}^{(0)} = E_0 |0\rangle\langle 0| + \sum_{T_K} E_K^T |T_K\rangle\langle T_K| + \sum_{S_K} E_K^S |S_K\rangle\langle S_K|$$

$$+ \sum_K^{\text{other}} E_K |K\rangle\langle K|, \tag{7}$$

where "other" in the last summation implies singly, triply, and other excited terms. Thus, Eq. (7) provides a generalization of Eq. (1) with the additional relations,

$$E_K^T = \frac{\Delta_K}{p_T} + E_0, \tag{8}$$

$$E_K^S = \frac{\Delta_K}{p_S} + E_0. \tag{9}$$

Following Feenberg, Szabados[41] determined the optimum scaling parameters from the condition that the sum of the scaled second- and third-order energies is stationary with respect to p_T and p_S. The latter assumes the form

$$E_{\text{SCS-MP}}^{(3)} = p_T^2 \sum_{T_K T_L} \frac{\langle 0|H|T_K\rangle\{\langle T_K|H|T_L\rangle - \delta_{T_K T_L}[(\Delta_K/p_T) + E^{(0)} + E^{(1)}]\}\langle T_L|H|0\rangle}{\Delta_K \Delta_L}$$

$$+ p_S^2 \sum_{S_K S_L} \frac{\langle 0|H|S_K\rangle\{\langle S_K|H|S_L\rangle - \delta_{S_K S_L}[(\Delta_K/p_S) + E^{(0)} + E^{(1)}]\}\langle S_L|H|0\rangle}{\Delta_K \Delta_L} + 2p_T p_S \sum_{T_K S_L} \frac{\langle 0|H|T_K\rangle\langle T_K|H|S_L\rangle\langle S_L|H|0\rangle}{\Delta_K \Delta_L}, \tag{10}$$

with the sum of the second- and third-order energies yielding

$$E_{\text{SCS-MP}}^{(2)} + E_{\text{SCS-MP}}^{(3)} = p_T^2 A_{TT} - 2p_T E_{\text{MP}}^{(2),T} + p_S^2 A_{SS} - 2p_S E_{\text{MP}}^{(2),S} + 2p_T p_S A_{ST}. \tag{11}$$

In Eq. (11), E_{MP2}^S and E_{MP2}^S are the second-order contributions in Eq. (4),

$$A_{TT} = \sum_{T_K T_L} \frac{\langle 0|H|T_K\rangle\{\langle T_K|H|T_L\rangle - \delta_{T_K T_L}[E^{(0)} + E^{(1)}]\}\langle T_L|H|0\rangle}{\Delta_K \Delta_L}, \tag{12}$$

$$A_{ST} = \sum_{T_K S_L} \frac{\langle 0|H|T_K\rangle\langle T_K|H|S_L\rangle\langle S_L|H|0\rangle}{\Delta_K \Delta_L}, \tag{13}$$

with a corresponding expression applying to A_{SS}. Explicitly, the above terms assume the forms[41]

$$A_{SS} = 2\sum_{ijk}\sum_{abc} \frac{\langle ij|ba\rangle\langle jc||ak\rangle\langle bc|ik\rangle}{\Delta_{ij}^{ab}\Delta_{ik}^{bc}}$$

$$- 2\sum_{ijk}\sum_{abc} \frac{\langle ij|ab\rangle\langle jc|ka\rangle\langle bc|ki\rangle}{\Delta_{ij}^{ab}\Delta_{ik}^{bc}}$$

$$+ \sum_{ijkl}\sum_{ab} \frac{\langle ij|ab\rangle\langle ij|kl\rangle\langle ab|kl\rangle}{\Delta_{ij}^{ab}\Delta_{kl}^{ab}}$$

$$+ \sum_{ij}\sum_{abcd} \frac{\langle ij|ab\rangle\langle cd|ab\rangle\langle cd|ij\rangle}{\Delta_{ij}^{ab}\Delta_{ij}^{cd}} + \sum_{ij}\sum_{ab} \frac{\langle ij|ab\rangle^2}{\Delta_{ij}^{ab}}, \tag{14}$$

$$A_{TT} = 2\sum_{ijk}\sum_{abc} \frac{\langle ij||ab\rangle\langle ci||ak\rangle\langle cb||jk\rangle}{\Delta_{ij}^{ab}\Delta_{jk}^{bc}}$$

$$+ \sum_{ijkl}\sum_{ab} \frac{\langle ij||ab\rangle\langle ij|kl\rangle\langle ab|kl\rangle}{\Delta_{ij}^{ab}\Delta_{kl}^{ab}}$$

$$+ \sum_{ij}\sum_{abcd} \frac{\langle ij||ab\rangle\langle cd|ab\rangle\langle cd|ij\rangle}{\Delta_{ij}^{ab}\Delta_{ij}^{cd}}$$

$$+ \sum_{ij}\sum_{ab} \frac{\langle ij|ab\rangle\langle ij||ab\rangle}{\Delta_{ij}^{ab}}, \tag{15}$$

$$A_{ST} = 2\sum_{ijk}\sum_{abc} \frac{\langle ij|ab\rangle\langle ci|ka\rangle\langle cb|kj\rangle}{\Delta_{ij}^{ab}\Delta_{jk}^{bc}}, \tag{16}$$

where Δ_{ij}^{ab} is defined as above. Note that in stepping from MP2 to MP3 no new repartitioning parameters appear. Thus, the approach differs from Grimme's[47] third-order one where a further scaling of the MP3 contribution is done. Szabados[41] suggested to obtain the optimum scaling parameters p_T and p_S by the above condition of stationary, and applied the method to various systems. Although the results were occasionally poorer than those from the standard MP2 approach, they have shown in some cases an improvement over the ones obtained via Grimme's empirical method.

Consider now the constraint $p_F = p_S = p_T$. Application of the above stationarity condition leads then straightforwardly to the Feenberg[40] minimal condition

$$p_F = \frac{E_{\text{MP}}^{(2)}}{E_{\text{MP}}^{(2)} - E_{\text{MP}}^{(3)}}, \tag{17}$$

where all terms have the meaning assigned above. The same result is obtained for the single spin-component case. For example, if the assumption is made of $p_T = 0$, the stationarity condition yields

$$p_F^S = -\frac{E_{\mathrm{MP}}^{(2),S}}{A_{SS}} = \frac{E_{\mathrm{MP}}^{(2),S}}{E_{\mathrm{MP}}^{(2),S} - E_{\mathrm{MP}}^{(3),S}}, \tag{18}$$

which can easily be shown to correspond to the root of the third-order perturbation energy. Moreover, from Eqs. (6) and (17), the following result is obtained:

$$E_{\mathrm{SOS\text{-}MP2}}^F = \frac{[E_{\mathrm{MP}}^{(2)}]^2}{E_{\mathrm{MP}}^{(2)} - E_{\mathrm{MP}}^{(3)}}, \tag{19}$$

where the superscript F emphasizes that the scaling factor has been obtained from Feenberg's single-parameter perturbation analysis. Note the appearance of the square of $E_{\mathrm{MP2}}^{(2)}$ in Eq. (19), which is responsible for the energy at second-order of Feenberg's perturbation development to obey size-extensivity only when N identical particles are involved, but not otherwise.[48]

III. IMPROVED SOS-MP2 THEORY

In attempting to obtain an improved spin-component-scaling theory, a few issues should desirably be taken into consideration. Two are general and appear critical when aiming at a description of covalent interactions: (a) the theory should reproduce any energies from the *ab initio* test set that are known exactly, such as for the prototypical covalent bond of H_2 which has a single OS pair of electrons; (b) it should rely only on the OS spin component of the MP2 energy such as to benefit of the important cost-effective advantages pointed out by Head-Gordon and co-workers. Aiming at a wider applicability for the theory, it should also be applicable to contracted/stretched bonds and, hopefully, van der Waals (vdW) interactions. Although this will not be explicitly attempted at this occasion, some guiding remarks will be pointed out later. In fact, this can hardly be achieved with SOS-MP2 theory since T excitations are formally discarded although they are known to play a critical role in vdW interactions as discussed in detail in several recent publications[38,39,49–52] devoted to this theme.

Consider then H_2 in its singlet ground state. Let the energy be expressed in terms of the reduced bond distance

$$\rho = \frac{R - \eta R_e}{\eta R_e}, \tag{20}$$

where R_e is the equilibrium geometry, η is the parameter, and the reduced bond distance satisfies $-1 \leq \rho < \infty$. For a chosen value of the cardinal number X, the following condition should be verified:

$$\lim_{\substack{n_S \to 1 \\ \rho \to \rho_e}} p_S(\rho, n_S, X) = \lim_{\substack{n_S \to 1 \\ \rho \to \rho_e}} \frac{E_{\mathrm{FCI}}}{E_{\mathrm{MP2}}} = \left. \frac{\epsilon_{\mathrm{FCI}}}{\epsilon_{\mathrm{MP2}}} \right|_{H_2}, \tag{21}$$

where n_S is the number of OS electron pairs, and the acronym FCI stands for full-configuration interaction. For very short distances, the H_2 molecule will collapse to its united atom,
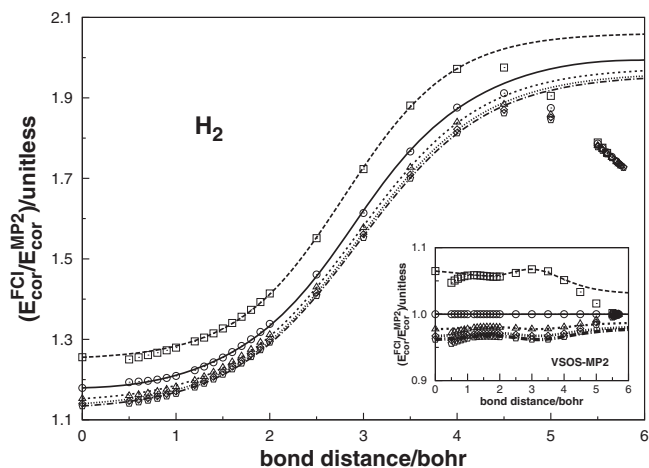

FIG. 1. Calibration of the OS weight in Eq. (24) by fitting the ratio between the exact (nonrelativistic) and MP2 correlation energies for the ground electronic state of the hydrogen molecule.

$$\lim_{\substack{n_S \to 1 \\ \rho \to -1}} p_S(\rho, n_S, X) = \lim_{\substack{n_S \to 1 \\ \rho \to -1}} \frac{E_{\mathrm{FCI}}}{E_{\mathrm{MP2}}} = \left. \frac{\epsilon_{\mathrm{CCSD}}}{\epsilon_{\mathrm{MP2}}} \right|_{He}, \tag{22}$$

with the correlation energies ϵ_{MP2} and ϵ_{CCSD} (or FCI) in Eq. (22) referring now to the helium atom in the ground electronic state.

The dependence of p_S is unknown and its form can hardly be anticipated. For simplicity, the drastic assumption is made here that it can be split into a shape term that varies with molecular geometry, $\mathcal{G}_S(\rho, X)$, and a size-term $\mathcal{S}(n_S, X)$ that accounts for the number of electrons of the system. The latter should then vary with the number of OS electron pairs. It is further argued that the theory should be applicable whenever the wave function of the system can be reliably approximated by a single reference. Thus, it is specially designed for regions of configuration space in the vicinity of the equilibrium geometry, as single-reference methods are generally bound to fail when chemical bonds are broken or formed. To summarize,

$$p_S(\rho, n_S, X) = \mathcal{G}_S(\rho, X)\mathcal{S}(n_S, X) \tag{23}$$

with an attempt being made here to suggest forms from simple physical arguments. For H_2 (this should also hold qualitatively for other bonds), the geometry-dependence can be well approximated by the form

$$\mathcal{G}_S(\rho, X) = g_0(X) + g_1 \tanh(g_2 \rho), \tag{24}$$

where g_i ($i = 0 - 2$) are parameters. Note that for a given X, Eq. (24) has enough flexibility to reproduce the value of p_S for the hydrogen molecule at equilibrium ($n_S = 1, \rho = \rho_e$), and the corresponding value for atomic helium ($n_S = 1, \rho = -1$). In fact, Fig. 1 shows that the calculated FCI energies for He and H_2 are accurately modeled by Eq. (24) for all basis sets; the fitted parameters are gathered in Table I. Obviously, the CCSD or QCISD curves coincide in this case with the FCI one whenever the former methods converge, which occurs for distances up to $3a_0$ or so. In turn, the inset of Fig. 1 shows that the curves for distinct values of X are nearly parallel to each other, thus justifying our ignoring of the

TABLE I. Coefficients in Eqs. (20) and (24) that model the correlation energy recovered by MP2 relative to FCI for H_2 and He in their ground electronic states.

Parameter	$X=D$	T	Q	5	6
VXZ					
g_0	1.655 310	1.585 550	1.559 044	1.546 542	1.540 626
g_1	0.406 116	0.415 079	0.416 458	0.416 552	0.417 076
g_2	2.432 888	2.269 377	2.208 711	2.183 740	2.167 738
η [a]	1.996 167	2.071 458	2.091 474	2.099 637	2.106 001
$\mathcal{G}_S(\rho_e,X)$	1.315 022	1.242 904	1.218 119	1.206 781	1.201 175
AVXZ					
$\mathcal{G}_S(\rho_e,X)$	1.312 562	1.238 138	1.214 751	1.204 520	1.189 232

[a]Reference geometry R_e is $1.40 a_0$.

X-dependence in the parameters g_1 and g_2. Besides the excellent fit to the converged data points, the notable feature from Fig. 1 is the systematic extrapolation to values of $\mathcal{G}_S \sim 2$ even at large distances where both the MP2 and CCSD single-reference methods show no longer a proper convergence. This is to be expected since the S and T contributions to the ground state curve are expected to be equal at the asymptote. In fact, a three-parameter fit based on Eq. (24) but fixing $g_1 = 2 - g_0$ yields essentially indistinguishable results. It is further observed that the calculated value of $p_S(\rho_e,X) = 1.238$ for H_2 with $X=T$ differs by about 5% from the corresponding value for He. In turn, Table I shows that the values of $\mathcal{G}_S(\rho_e,X)$ are identical within a few percent for the VXZ and AVXZ basis set families, when the cardinal number varies over the whole series that is commonly utilized. The small difference in the predicted correlation energy for the single OS pair in the molecule and corresponding united atom can be attributed at least in part to the different performance of the atomic and molecular basis sets even if they refer to the same cardinal number and basis set family. Ignoring such a difference (i.e., by viewing the atom as the limit of the molecule), the near parallel behavior of the various curves in the inset of Fig. 1 suggests that the model is expected to be universal in the sense of being valid irrespective of the basis set that is utilized.

To account for system size, the following expression has been utilized:

$$\mathcal{S}(n,X) = 1 + \lambda(X)\xi_{OS}, \qquad (25)$$

where $\xi_{OS} = (n-2)/(n-1)$ is the fraction of OS pairs that can be formed from n electrons. Assuming maximum OS pairing, the OS weight assumes then the form

$$p_S(n,\rho_e,X) = \mathcal{G}_S^{H_2}(\rho_e,X)\left[1 + \lambda(X)\frac{n-2}{n-1}\right], \qquad (26)$$

where the factor $\lambda(X)$ is to be calibrated from the *ab initio* test set (ATS1). As noted above, the dependence of p_S on the cardinal number is mild for H_2, and can be neglected to good accuracy. However, it will be given here for completeness,

$$p_0(X) = a + b\exp(-cX), \qquad (27)$$

where $a = 1.19875$, $b = 0.79746$, and $c = 0.96445$ are parameters. No such a fit has been done for $\lambda(X)$ in Eq. (26) as the data for the test set has only been obtained at the QCISD(T)/VTZ level of theory. Thus, λ will be assumed

here to be X-independent, with its value being specified below. It should be emphasized that the inclusion of any geometry-dependence in $\mathcal{G}_S$ should be done cautiously as it may question the size-extensivity (consistency) of the method besides making the interface with available electronic structure codes more difficult, in particular, MOLPRO (Ref. 53) where the SCS-MP2 theory is implemented. Thus, it has been fixed at $\mathcal{G}_S^{H_2}(\rho_e,X)$. In fact, both the MP2 and target CCSD(T) methods are valid over restricted regions of the molecule configuration space (where the single-reference approximation is valid), and hence interest is focused on such regions. Noting further the above mild-dependence on X, the shape term will be reduced to a constant $\mathcal{G}_S(\rho_e)$, and identified with the corresponding value for H_2. Only λ remains therefore to be specified. As noted above, it is chosen from a fit to ATS1, the *ab initio* test set adopted by Head-Gordon and co-workers.[35] First, an estimate of λ has been obtained by minimizing the sum of squared relative deviations in a fit of the ratio SS/OS of MP2 energies versus the number of electrons. This led to the curve with $\lambda = 0.173\,51$ in Fig. 2. Note that n has been chosen for simplicity as the total number of electrons, irrespective of using frozen-core or all-electron correlated calculations (however, see the next paragraph). As expected, the fitted ratio overshoots on average by $\sim 2\%$ the recovered correlation energy with respect to
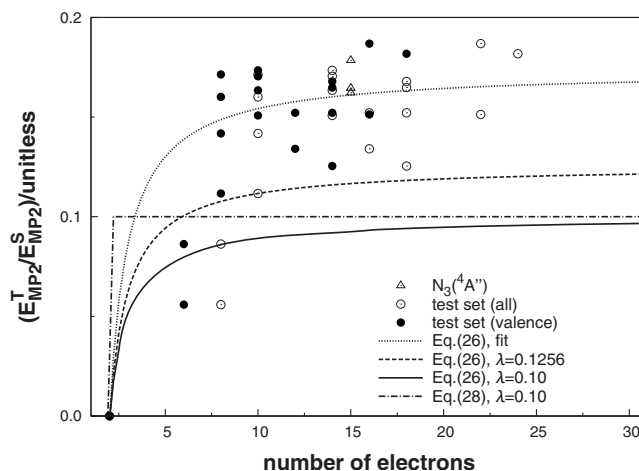

FIG. 2. Calibration of the size-term in VOS-MP2 theory using the *ab initio* test set ATS1, and Eqs. (26) and (28). The labels "valence" and "all" are meant to imply that in the x-axis the number of electrons excludes and includes the core, respectively. See the text.

QCISD(T). A value that leads to just one exception (the overshooting is of 2% only for BH_3) is $\lambda = 0.1256$. This can itself be prevented by a further reduction to $\lambda = 0.10$. In fact, this value may have the advantage of working almost as reliably also in the case of all-electron correlated calculations (solid line in Fig. 2). Thus, it is perhaps more general than $\lambda = 0.1256$, although this (or a nearby) value may still be recommendable for valence-only calculations by far the ones for which the theory is most required.

In spite of the above observations, size-extensivity has not been guaranteed yet. Consider the dissociation process of a weakly bound cluster, $A_xB_y \rightarrow xA + yB$, where A and B are monomers. If the cluster is homonuclear, size-extensivity will be warranted since the value of the scaling parameter p_S is equal both for the cluster and monomers. Yet, the correlation energies will be system-specific in the sense that may vary slightly with cluster size. This can be avoided if p_S is defined instead from the monomers, which should be realistic since no bonds are broken or formed and the monomers keep essentially unaltered their equilibrium geometries. However, if A and B are distinct, size-extensivity can only be kept in a system-specific sense provided that the same scaling is retained for both fragments. All the above subtleties can, however, be avoided by noting that Eq. (26) can be advantageously replaced by the form

$$p_S(n, \rho_e, X) = \mathcal{G}_S^{H_2}(\rho_e, X)\{1 + \lambda(X) H[n-2]\}, \quad (28)$$

where $H[n-2]$ is the Heaviside or unit step function defined here as follows: 0 for $(n-2) \leq 0$; 1 for $(n-2) > 0$. Although being a discontinuous function, this has no implications whatsoever. In fact, many analytic approximations exist for $H[n-2]$, one being Eq. (28) itself for $(n-2) \geq 0$. Thus, for all but two-electron systems where $p_S = 1.2429$, a scaling factor of $p_S = 1.3672$ is obtained ($p_S = 1.3990$, if $\lambda = 0.1256$). Clearly, the above substantiates in practice into a small difference from SOS-MP2 theory which utilizes a single value of p_S irrespectively of n. However, the designation of VOS-MP2 is justified for two reasons. First, it highlights the n-dependence of the model. Second, it opens the possibility for further flexibility if size-extensivity is kept within a system-specific sense. A comparison of how both approaches perform is given later. As it will be shown, the differences between the results of the size-extensive [based on Eq. (28)] and non-size-extensive [Eq. (26)] formulations are minor. In fact, illustrative cuts along bond-breaking processes have been analyzed (not reported) for a few extra molecules, and no exception to the trends here presented has been encountered.

A. Analysis of performance relative to the *ab initio* test set

We start by observing that any new improvement of the VOS-MP2 theory should satisfy a few requirements to be considered successful. First, as the scaling is free of cost compared with the calculations needed for obtaining the S and T components of the energy, one should not require a very drastic reduction of the error. However, the results should be perceptively improved by the new approach when compared with its original form in SOS-MP2. It would also be desirable that the new theory reproduces as accurately as possible the energies for which exact or nearly exact results are available, with H_2 being a paradigmatic example among molecules. Of course, the correlation energies of the test set are far from exact in most cases and there is no variational sense in the method that prevents its predicted results of exceeding the target values. Thus, it should not be surprising that the predicted correlation energies can be overestimated, although any large overestimation may signal a warning concerning the performance of the new theory. For practical purposes, it would then be fair enough to get better results than one often obtains using other schemes from the same family. This is clearly the case, as shown in Table II, which compares the correlation energy recovered by MP2, SCS-MP2, SOS-MP2, and VOS-MP2 relative to CCSD(T). As shown, the VOS-MP2 method with $\lambda = 0.10$ recovers $(97.0 \pm 1.4)\%$ of the CCSD(T) correlation energy, while an even better performance of $(99.0 \pm 1.2)\%$ is achieved with $\lambda = 0.1256$. Note that the average is taken here of the unsigned values such as to give a meaningful result. Such values compare with the more modest SCS-MP2 and SOS-MP2 results for the same *ab initio* test set (ATS1) of (92.8 ± 1.8) and $(93.4 \pm 2.9)\%$, respectively. Resulting from a calibration, it cannot be excluded that in some cases the predictions may slightly exceed 100% as indeed described later. In fact, this has been observed in a few cases, although one should point out that the correlation energies available for comparison are accurate but far from exact. Moreover, the largest error here reported is small when compared with the one of SOS-MP2 for the H_2 system (4.7%). It is also relevant to stress at this junction that the enhanced agreement in absolute value between the VOS-MP2 and accurate correlation energies is desirable but not essential in practice. In fact, it often suffices that the energies relative to some chosen reference agree, say relative to the separated atoms in a dissociative process.

B. Hydrogen fluoride: A prototypical chemical bond

The hydrogen fluoride molecule has been extensively studied by generations of experimentalists and theoreticians, resulting in an extensive body of information available for critical comparisons.[54] Although now amenable to rather high-accuracy type calculations, this ten-electron system will serve as a prototype of a bound diatomic molecule on which the schemes here discussed will be tested. Figure 3 compares the correlation energy calculated at the various MP2 levels of theory with the results obtained from CCSD(T)/VTZ calculations. In all cases, the VTZ basis set of Dunning was employed.[55] Both calculations with core correlation and without it have been performed, although only the former are discussed here. Since the interest is only to compare the calculated correlation energies, no attempt has been made to compensate for the basis set superposition error (BSSE), either by using the popular counterpoise[56] (CP) technique or via extrapolation to the CBS limit.[28,30,31] For illustration, the results obtained with $\lambda = 0.10$ [using both Eqs. (26) and (28)], and 0.1256 are presented. Moreover, the results obtained with other variants of SCS-MP2 theory which derive from applications of the Feenberg[40,41] theoretical approach are in-

TABLE II. The correlation energy recovered by MP2, SCS-MP2, SOS-MP2, and VOS-MP2 relative to CCSD(T) (energies in kcal mol^{-1}).

System[b]	CCSD(T)	MP2				SCS-MP2		SOS-MP2		VOS-MP2[a]			
		OS	SS	%CE[c]	%[d]	%CE[c]	%[d]	%CE[c]	%[d]	%CE[c,e]	%[d,e]	%CE[c,f]	%[d,f]
$^1CH_2$	−169.1	−118.2	−10.2	82.0	−18.0	87.9	−12.1	90.9	−9.1	94.3	−5.7	96.2	−3.8
C_2H_2	−340.4	−239.4	−36.1	91.5	−8.5	91.5	−8.5	91.4	−8.6	95.5	−4.5	97.5	−2.5
C_2H_4	−375.4	−264.7	−35.5	89.4	−10.6	90.9	−9.1	91.7	−8.3	95.8	−4.2	97.9	−2.1
C_2H_6	−414.9	−295.8	−37.1	89.2	−10.8	91.5	−8.5	92.7	−7.3	97.0	−3.0	99.1	−0.9
CH_4	−224.9	−162.1	−18.1	88.2	−11.8	91.9	−8.1	93.7	−6.3	97.5	−2.5	99.6	−0.4
CO	−379.4	−267.5	−45.6	94.5	−5.5	92.6	−7.4	91.7	−8.3	95.7	−4.3	97.8	−2.2
Cyclopropene	−533.7	−374.1	−56.6	91.3	−8.7	91.2	−8.8	91.1	−8.9	95.4	−4.6	97.5	−2.5
F_2	−545.6	−391.9	−65.8	95.9	−4.1	94.2	−5.8	93.4	−6.6	97.7	−2.3	99.8	−0.2
H_2O	−276.1	−198.6	−31.8	95.0	−5.0	94.0	−6.0	93.5	−6.5	97.3	−2.7	99.4	−0.6
H_2O_2	−525.6	−375.5	−61.9	95.0	−5.0	93.6	−6.4	92.9	−7.1	97.2	−2.8	99.3	−0.7
HCN	−371.8	−263.7	−43.1	94.1	−5.9	92.8	−7.2	92.2	−7.8	96.3	−3.7	98.4	−1.6
HF	−281.3	−203.0	−34.8	96.9	−3.1	94.8	−5.2	93.8	−6.2	97.7	−2.3	99.7	−0.3
H_2	−39.4	−31.7	0.0	80.5	−19.5	96.6	−3.4	104.7	4.7	100.1	0.1	100.1	0.1
N_2	−398.5	−284.2	−49.3	96.1	−3.9	93.8	−6.2	92.7	−7.3	96.8	−3.2	98.9	−1.1
N_2H_2	−475.4	−337.8	−51.4	92.7	−7.3	92.5	−7.5	92.4	−7.6	96.6	−3.4	98.7	−1.3
N_2O	−672.1	−479.4	−89.6	98.0	−2.0	94.5	−5.5	92.7	−7.3	97.1	−2.9	99.3	−0.7
NH_3	−255.8	−183.4	−26.0	92.0	−8.0	92.8	−7.2	93.2	−6.8	97.0	−3.0	99.1	−0.9
O_3	−805.5	−579.2	−105.3	98.1	−1.9	95.0	−5.0	93.5	−6.5	97.9	−2.1	100.1	0.1
N_2H_4	−475.4	−337.8	−51.4	92.7	−7.3	92.5	−7.5	92.4	−7.6	96.6	−3.4	98.8	−1.2
BH_3	−137.6	−102.0	−5.7	82.4	−17.6	91.7	−8.3	96.4	−3.6	100.0	0.0	102.1	2.1
⟨%CE⟩[g]				91.9	9.7	92.8	7.4	93.4	7.2	97.0	3.3	99.0	1.6
rmsd[h]				±5.1		±1.8		±2.9		±1.4		±1.2	

[a]The results for the size-extensive version of VOS-MP2 theory with $p_S(n.\rho_e, X) = 1.3672$ (two bottom entries of columns 11 and 12) are 97.7 ± 1.4 and 2.7%. With $p_S(n.\rho_e, X) = 1.3990$, the corresponding results (columns 12 and 14) are 99.8 ± 1.3 and 1.3%; see text.
[b]In all cases the results refer to valence-only VTZ calculations performed at (or near) the equilibrium geometry.
[c]Percentage of recovered correlation energy relative to QCISD(T) values.
[d]Percentage of missing correlation.
[e]Equation (26), with λ=0.10.
[f]With λ=0.1256.
[g]Average unsigned value.
[h]Defined as $[\Sigma_{i=1}^{N}(\%CE_i - \langle\%CE\rangle)^2/N]^{1/2}$.

cluded; these are indicated by carrying the letter "F" as a subscript, and will be addressed in detail in Sec. IV. The notable feature is the rather close agreement of the VOS-MP2 and CCSD(T) curves over the whole range of internuclear distances where the latter method is stable. This is especially true when λ=0.1256, although the general trend in a comparison with the curves calculated using the standard SCS-MP2 and SOS-MP2 theories is good overall. In fact, although showing visible differences in absolute value when compared with VOS-MP2, the curves referring to other spin-component-scaling variants of MP2 theory behave pretty much parallel and close to each other (except for the VOS$_F$-MP2-II case curve that predicts the smallest absolute value) and hence are not expected to yield drastic differences for the energetics. It should also be noted that similar trends (in the sense of the agreement reported above) are obtained when core correlation effects are neglected.

C. The lowest quartet state of trinitrogen

The potential energy surface for the $N_3(^4A'')$ system has been the subject of much theoretical and experimental work,[57–60] mostly due to its interest on the design of space aircrafts. It is characterized by a metastable minimum which is replicated in a threefold way due to the permutation symmetry of the nuclei. Each minimum is separated from the two

reactive atom-diatom asymptotes by equivalent saddle points of C_s symmetry, in a total of six equivalent saddle points. The calculations reported here will be carried out pointwise at the geometries optimized at the CCSD(T)/AVTZ level of theory in Ref. 60, with the restricted-Hartree–Fock wave function being taken as the unperturbed wave function for the post-Hartree–Fock RMP2 and RCCSD(T) calculations (the R will be omitted from these acronyms for brevity). Since the open shell works only with spin orbitals rather than singlet and triplet pairs, the $\alpha\beta$ and $(\alpha\alpha + \beta\beta)$ energies have been scaled instead. A detailed analysis of the above three stationary points has first been carried out at the valence-only level with the VXZ and AVXZ basis sets;[55] $X = D, T, \ldots, 6$. The results, which are collected in Table I of ESI (Ref. 61) carrying the designation C6 to indicate the number of core electrons, show that the VOS-MP2 method with λ=0.1256 recovers on average more than 98% of the CCSD(T) correlation energy in a performance similar to what has been found for ATS1. Only a slightly poorer performance is found with λ=0.10, namely, (97 ± 1.7)%. In both cases, the fluctuations over the potential energy surface appear to be rather small: rmsd= ± 1.7%. Note that the average performance of Eq. (28) is rather close or even slightly enhanced relatively to Eq. (26), and hence only detailed results obtained from the latter are here reported.
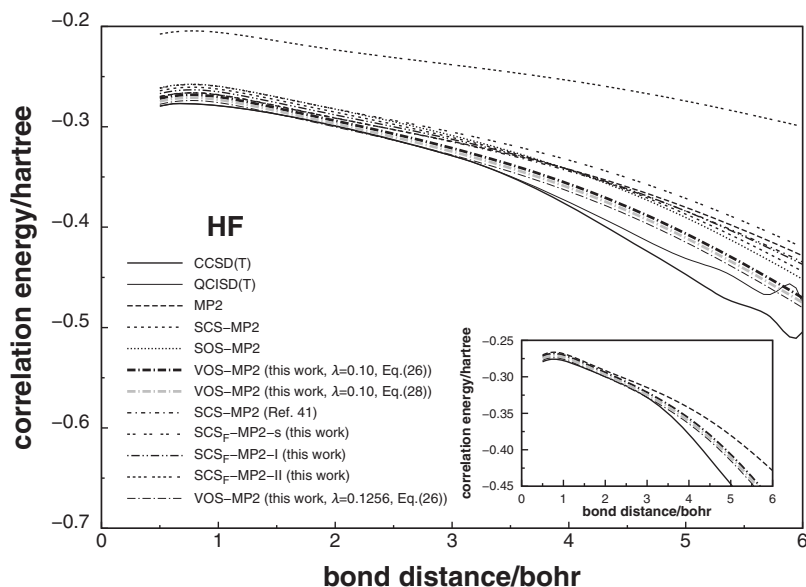
FIG. 3. Recovered correlation energy for the ground electronic state of hydrogen fluoride from the various methods studied in the present work.

Calculations with all electrons correlated (C0) have also been performed by employing correlation-consistent core-valence basis sets[27,62] (cc-pCVXZ and aug-cc-pCVXZ or CVXZ and ACVXZ), thus allowing an estimate of the core correlation effects. A comparison with the CCSD(T) results [Table II of ESI (Ref. 61)] shows that the correlation recovered by the VOS-MP2 method with $\lambda=0.1256$ overshoots the CCSD(T) value by 3%–4%, although this reduces to $\approx1\%$ when $\lambda=0.10$. This should not be surprising as the ATS1 calibration set does not include core correlation effects. Besides, the exact energy is unknown in such cases. Note, however, that any small overestimation should have a minor impact on the relative energies (cf. results for $\lambda=0.1256$ and 0.10).

The reliability of the single-reference CCSD(T) method for the title open-shell species has been examined by looking at some diagnostics. For the (all-electron) CCSD calculations, the T_1-diagnostic[63] computed from single-substitution amplitudes are ≈0.025 at the saddle point geometry and ≈0.022 at the minimum, thus meeting the usual requirement ($\lesssim0.02$) for the single-reference CCSD(T) method (the chemist's golden rule) to be trusted. Yet, the T1 diagnostic cannot by itself be a sign that CCSD(T) is reliable. Since a doubles diagnostic has been reported only in nuclear structure theory,[64] another [D_1 (Ref. 65)] from single-substitution amplitudes will be utilized. This assumes values of $D_1 \approx0.071$ and 0.082 which, used in tandem,[63,65,66] yield ratios of $T_1/D_1 \approx0.345$ and 0.267, in the above order. While the values of D_1 slightly exceed 0.025, signaling some inadequacy of the CCSD approach, the ratios T_1/D_1 less than $1/\sqrt{2}$ point to some inhomogeneity in the molecular electronic structure, i.e., some problem areas in the molecule and others where the CC approach performs better.[66] The largest T1 and T2 amplitudes may also provide an additional judgment: In all cases the print threshold value of 0.05 has been exceeded by only a few amplitudes, typically fewer (1 or 2) at the saddle point than at the minimum. It is then argued that CCSD(T) gives a solution as close as one might ambition with a single-reference method to the FCI limit, particularly at the saddle point geometry.

Of special relevance for dynamics is the barrier height (V_b) for exchange of a nitrogen atom. The results [Tables III and IV of ESI (Ref. 61)] show that the VOS-MP2 values at valence-only correlated level agree within 1 kcal mol^{-1} with the CCSD(T) ones, irrespectively of the value of λ and basis set. Larger differences are, however, observed for the metastability. This attains values at CCSD(T) that are smaller than at any other level of theory here considered, with SCS-MP2 being generally largest (~7.5 kcal mol^{-1}). Note that MP2 predicts already a value of 3–4 kcal mol^{-1} larger than CCSD(T). The scenario is not changed drastically when correlating all electrons, although there is a tendency for a slight increase of the barrier height [Table V of ESI (Ref. 61)]. The best VOS-MP2 [CCSD(T)] raw estimates are now 45.8(46.2) and 45.9(46.2) kcal mol^{-1} (with CV6Z and ACV5Z basis) when $\lambda=0.10$, with the VOS-MP2 values changing slightly when $\lambda=0.1256$ (45.4 and 45.5 kcal mol^{-1} in the same order). They can be compared with the value utilized on a recent parametrization[60] of a potential energy surface of the double-many-body expansion type for N$_3$($^4A''$), namely, 45.9 kcal mol^{-1} as predicted by the correlation scaling[70]/uniform singlet- and triplet-pair extrapolation[30] [CS/USTE (Ref. 71)] method employing AVDZ and AVTZ basis sets and a single pivotal geometry calculated with AVQZ. Clearly, there is a remarkable agreement with the best estimate here reported of (45.6 ± 0.8) kcal mol^{-1} [which embraces all CBS/SCS-MP2 and CBS/CCSD(T) estimates; see later].

As noted above, the scaled spin-component theories tend to predict a well depth larger by a few kcal mol^{-1} than the CCSD(T) value, irrespectively of the basis set. This should not come entirely as a surprise. In fact, the standard MP2/VTZ theory already predicts a barrier height relative to the minimum of 6.1 kcal mol^{-1} while the corresponding CCSD(T) value is 2.3 kcal mol^{-1}, with a similar trend being observed for other basis sets. The CCSD(T) and MP2 potential energy surfaces show, therefore, a nonparallelism that can hardly be accounted for prior to including some geometry-dependence on the correlation scaling factor.

TABLE III. Energies for $N(^4S)+N_2$ reaction from various single-reference methods (absolute total energies have been added 163, with geometries as in Ref. 60) at the CBS extrapolation level.

Method	(X_1,X_2) [a]	Reactants Total	Minimum Total	ϵ [b]	ϵ [c]	Saddle point Total	V_b [d]	V_b [e]	V_b [f]	V_b [g]
HF/VXZ	(5,6)	−0.392 562	−0.259 029	83.8		−0.269 712	77.1		−6.7	
HF/AVXZ	(5,6)	−0.392 650	−0.259 214	83.7		−0.269 906	77.0		−6.7	
HF/CVXZ	(5,6)	−0.392 639	−0.259 155	83.8		−0.269 840	77.1		−6.7	
HF/ACVXZ	(5,6)	−0.392 747	−0.259 266	83.8		−0.269 966	77.1		−6.7	
CCSD(T)-C6/VXZ	(5,6)	−0.951 632	−0.883 732	42.6		−0.878 687	45.8		3.2	
CCSD(T)-C6/AVXZ	(5,6)	−0.951 497	−0.883 643	42.6		−0.878 607	45.7		3.1	
MP2-C6/VXZ	(5,6)	−0.928 394	−0.856 502	45.1		−0.845 296	52.1		7.0	
MP2-C6/AVXZ	(5,6)	−0.928 272	−0.856 337	45.1		−0.845 132	52.2		7.0	
SCS-MP2-C6/VXZ	(5,6)	−0.922 041	−0.873 310	35.4		−0.852 743	46.1		10.7	
SCS-MP2-C6/AVXZ	(5,6)	−0.921 048	−0.864 681	35.4		−0.847 549	46.1		10.7	
SOS-MP2-C6/VXZ	(5,6)	−0.916 381	−0.859 198	35.9		−0.842 986	46.1		10.2	
SOS-MP2-C6/AVXZ	(5,6)	−0.916 368	−0.859 176	35.9		−0.842 963	46.1		10.2	
VOS-MP2-C6/VXZ	(5,6)	−0.940 726	−0.885 501	34.7	34.0	−0.868 756	45.2	44.7	10.5	10.7
VOS-MP2-C6/AVXZ	(5,6)	−0.940 715	−0.885 488	34.7	34.0	−0.868 742	45.2	44.7	10.5	10.7
CCSD(T)-C0/CVXZ	(5,6)	−1.129 663	−1.061 702	42.6		−1.056 416	46.0		3.4	
CCSD(T)-C0/ACVXZ	(Q,5)	−1.130 764	−1.062 706	42.7		−1.057 476	46.0		3.3	
MP2-C0/CVXZ	(5,6)	−1.101 962	−1.030 085	45.1		−1.018 663	52.3		7.2	
MP2-C0/ACVXZ	(Q,5)	−1.108 197	−1.031 435	48.2		−1.020 095	55.3		7.1	
SCS-MP2-C0/CVXZ	(5,6)	−1.114 085	−1.057 582	35.5		−1.037 899	47.8		10.3	
SCS-MP2-C0/ACVXZ	(Q,5)	−1.115 428	−1.058 673	35.6		−1.041 523	46.4		10.8	
SOS-MP2-C0/CVXZ	(5,6)	−1.119 005	−1.061 592	36.0		−1.045 323	46.2		10.2	
SOS-MP2-C0/ACVXZ	(Q,5)	−1.120 214	−1.062 548	36.2		−1.046 373	46.3		10.1	
VOS-MP2-C0/CVXZ	(5,6)	−1.152 840	−1.097 379	34.8	34.2	−1.080 575	45.4	44.9	10.4	10.7
VOS-MP2-C0/ACVXZ	(Q,5)	−1.154 050	−1.098 328	35.0	34.4	−1.081 620	45.5	45.0	10.5	10.6

[a]From fit to Hartree–Fock energies for $X=X_1$ and X_2 with (Ref. 67) $E_X^{HF}=E_\infty^{HF}+B\exp(-9\sqrt{X})$; see Table VI of ESI (Ref. 61). Similarly for correlation energy but using USTE (Refs. 30, 68, and 69): $E_\infty^{cor}=E_{X_2}^{cor}-A_5^\circ(X_2+\alpha)^{-5}+[E_{X_1}^{cor}-E_{X_2}^{cor}+A_5^\circ(Y_2^{-5}-Y_1^{-5})]/[c(Y_2^{-5}-Y_1^{-5})+Y_2^{-3}-Y_1^{-3}](Y_2^{-3}+cY_2^{-5})$, where $Y_i=X_i+\alpha$ and $i=1,2$, and $\alpha=-3/8$. A_5° and c are "universal" coefficients: 0.096 066 8 E_h and −1.582 009 for MP2-type methods; $A_5^\circ=0.166\,069\,9$ E_h and $c=-1.422\,251\,2$ for CCSD(T).
[b]Energy relative to reactants, in kcal mol^{-1} with $\lambda=0.1$.
[c]Energy relative to reactants, in kcal mol^{-1} with $\lambda=0.1256$.
[d]Energy relative to reactants, in kcal mol^{-1} with $\lambda=0.10$.
[e]Energy relative to reactants, in kcal mol^{-1} with $\lambda=0.1256$.
[f]Energy relative to minimum, in kcal mol^{-1} with $\lambda=0.10$.
[g]Energy relative to minimum, in kcal mol^{-1} with $\lambda=0.1256$.

Clearly, the unknown form of such a factor and the high dimensionality of the configuration space call for simplicity. A pragmatic approximate way of accounting for such an unbalance consists of scaling the predicted barrier height for metastability by the *a priori* ratio of the CCSD(T) and MP2 results as obtained with the cheapest basis set (VDZ). Calculations show such a scaling factor to have a value of 1.595/5.185=0.308. If utilized, the corrected values for the metastability fall essentially over the CCSD(T) results, with a most likely value of (3.1 ± 0.2) kcal mol^{-1}. Slightly larger values are obtained if the VTZ ratio is utilized. Such considerations apply also to the CBS results discussed next.

For enhanced accuracy, the raw energies discussed above have been extrapolated to the CBS limit. Since the energy is already split into its HF and correlation parts, each component just requires to be CBS extrapolated. This has been done using the energies associated to the highest pair of cardinal numbers (X_1,X_2) that can be considered. For the Hartree–Fock energy, the two-point extrapolation formula recommended by Karton and Martin[67] has been utilized. As expected, the convergence of the HF energy with basis set size is fast. In turn, the two-point USTE (Ref. 30) scheme

has been employed to CBS extrapolate the correlation energy. The data is gathered in Table III. As shown, the general trend observed in the previous paragraph is maintained, with a recommended best estimate for the barrier height of (45.6 ± 0.8) kcal mol^{-1}, thus in excellent agreement with the CS/USTE(T,Q) value of 45.9 kcal mol^{-1} reported in Ref. 60. Although use has been made of Eq. (26) [rather than Eq. (28)], any potential error due to this is expected to be much smaller than the reported uncertainty, and hence ignorable. It should be noted that both SCS-MP2 and SOS-MP2 have a similar performance to VOS-MP2 theory for the quartet state of trinitrogen. Thus, the above results open the possibility for a slight increase of the rate constant predicted in Ref. 72, and so to a further slight improvement on the agreement with the available experimental data.[73–75] Of course, an estimate of relativistic and nonadiabatic effects will be required for a better judgment, although this falls outside the scope of the present work.

D. The helium-neon van der Waals molecule

The spin-component-scaling theories studied here are now applied to a prototypical vdW interaction that does not

TABLE IV. Optimum values of coefficients in Eqs. (29) and (30) (with $\lambda = 0.10$).

System	Variational[a]			Equation (35)[b]			Feenberg[c]	Variational[d]	
	ξ	p_S	p_T	ξ	p_S	p_T	p_S	p_S	p_T
H$_2$	−0.027 291	1.215 810		−0.027 291	1.215 810		1.215 810	1.213 85	
He$_2$	−0.021 411	1.175 393	0.516 642	−0.021 441	1.175 615	0.516 134	1.175 692		
				0.224 067	0.000 104	1.899 935			
NH$_3$	−0.006 945	1.037 768	0.799 521	−0.019 884	1.163 512	0.543 490	1.202 830	1.089 57	0.911 73
				0.042 984	0.328 976	1.668 941			
H$_2$O	−0.005 985	1.026 601	0.820 038	−0.015 538	1.126 381	0.624 241	1.179 969	1.062 17	0.882 99
				0.035 404	0.423 691	1.588 355			
F$_2$	−0.006 053	1.027 390	0.818 598	−0.012 696	1.099 332	0.680 224	1.161 439	1.049 91	0.856 21
				0.030 548	0.491 607	1.525 718			
CO	−0.006 738	1.035 380	0.803 936	−0.011 935	1.091 718	0.695 573	1.134 318	1.051 41	0.832 79
				0.031 162	0.482 745	1.534 140			
HF (VDZ)	−0.006 053	1.027 396	0.818 587	−0.018 360	1.151 063	0.571 086	1.194 742	1.058 27	0.826 78
				0.030 930	0.486 083	1.530 976			
HF (VTZ)	−0.005 008	1.014 967	0.841 067	−0.012 909	1.101 440	0.675 943	1.183 003	1.058 27	0.826 78
				0.030 857	0.487 146	1.529 966			
N$_2$	−0.007 318	1.042 035	0.791 595	−0.010 189	1.073 628	0.731 350	1.109 737	1.050 18	0.808 21
				0.029 843	0.501 864	1.515 874			

[a]This work, from the condition of a stationary sum of second plus third-order (correlation) energies, Eq. (11).
[b]This work, from the condition of a vanishing third-order energy. Whenever two entries are reported, they refer to the two roots that are encountered (see text).
[c]This work, from Feenberg's (Ref. 41) condition in Eq. (17), for the cc-pVTZ basis set.
[d]Reference 41, by treating p_S and p_T as independent parameters.

belong to ATS1, namely, HeNe; see Fig. 4. All calculations have employed the VTZ basis set, with only the valence electrons correlated. Despite being a rather poor basis set for such a dimer, no attempt has been made to correct for BSSE via CP, an issue that will be addressed in Sec. IV for the even more fragile helium dimer. Two results deserve a comment. First, both SCS-MP2 and SOS-MP2 theories tend to underestimate the well depth, a result that is often interpreted as positive, since somehow it accounts for the fact that CP tends to weaken the attraction. In this regard, VOS-MP2 sits in between such curves, being the results almost indistinguishable when the theory utilizes the weights in Eqs. (26) and (28). Second, and quite interestingly, SCS-MP2 calculations with $p_S = p_T = 0.95$ yield a curve in rather good agreement with the CCSD(T) interaction potential. It turns out though that a calculation with $p_S = 0$ and $p_T = 1.90$ also yields a quite fair account of the intermolecular interaction, especially at long-range distances. Note, however, that the unsigned (unscaled) MP2 correlation energy at the asymptote is about 165.8 mE_h larger than the value obtained with $p_S = 0$ and $p_T = 1.90$, which is in turn smaller by 150.9 mE_h than the SCS-MP2 result obtained with $p_S = p_T = 0.95$. Although improved agreement could possibly be obtained using a somewhat larger value of p_T, such an analysis will not be pursued. The above findings provide supporting evidence for the role of SS correlation, and will be interpreted theoretically in Sec. IV using a constrained Feenberg[40,41] approach.

IV. FROM SOS TO SSS: RATIONALIZATION VIA FEENBERG SCALING

The results of Sec. III D suggest that one may evolve from SOS to scaling of SSS theory via some coordinate that measures the contribution of S and T excitations. In fact, similar findings have been gathered elsewhere.[38,39] This is illustrated in Fig. 5 where the solid symbols indicate the set of spin-component-scaling parameters that have been employed in SCS-MP2,[32] SOS-MP2,[32] and SSS-MP2[38] (these carry the initials MI in the original paper to stress that they have been optimized for intermolecular binding energies) calculations at the VTZ level of theory. Also shown are the values utilized in the present work, as well as the value reported above for the HeNe system, which is indicated by the open circle. As seen, such data are well modeled analytically by the functions

$$p_S(\xi) = \frac{A}{2}\{1.0 - \tanh[g_S(\xi - \xi_0)]\}, \tag{29}$$

$$p_T(\xi) = \frac{B}{2}[1.0 + \tanh(g_T\xi)], \tag{30}$$

where $A = 1.243 + \lambda(n-2)/(n-1)$ and $B = 1.90$; a value of $\lambda = 0.1$ has been utilized in plotting Fig. 3, together with a large value of n (i.e., $A = 1.3672$). In turn, ξ is the chosen independent coordinate, whose optimum value will be formally cho-

TABLE V. Correlation energies calculated at equilibrium geometry.

System[a]	MP2	SCS-MP2	SOS-MP2	VOS-MP2[b]	MP2-SCS$_F$ [c]	MP2-SCS$_F$-v[d]	MP2-SCS$_F$-I[e]	MP2-SCS$_F$-II[f]	CCSD(T)
H$_2$	−0.031 679	−0.038 015	−0.041 183	−0.039 374	−0.038 454	−0.038 516	−0.038 516		−0.039 374
He$_2$	−0.066 294	−0.079 545	−0.086 170	−0.090 625	−0.077 931	−0.077 912	−0.077 930	−0.000 024	−0.078 178
NH$_3$	−0.235 120	−0.237 085	−0.238 067	−0.250 373	−0.246 934	−0.228 765	−0.239 230	−0.146 214	−0.255 305
	−0.249 986	−0.252 503	−0.253 762	−0.266 880	−0.262 634	−0.243 347	−0.254 697	−0.154 795	−0.270 428
H$_2$O	−0.261 508	−0.258 764	−0.257 393	−0.270 698	−0.266 385	−0.252 262	−0.260 516	−0.183 294	−0.275 068
	−0.275 118	−0.272 452	−0.271 119	−0.285 134	−0.280 295	−0.265 444	−0.274 223	−0.192 536	−0.288 641
F$_2$	−0.522 477	−0.513 032	−0.508 309	−0.534 585	−0.523 088	−0.503 215	−0.515 633	−0.388 745	−0.544 060
	−0.547 266	−0.537 352	−0.532 395	−0.559 916	−0.547 902	−0.527 085	−0.540 085	−0.407 213	−0.568 472
CO	−0.366 530	−0.360 153	−0.356 964	−0.375 416	−0.365 273	−0.353 896	−0.361 521	−0.270 728	−0.386 854
	−0.379 860	−0.373 574	−0.370 432	−0.389 580	−0.378 638	−0.366 854	−0.374 854	−0.280 176	−0.400 117
HF (VTZ)	−0.271 802	−0.266 030	−0.263 144	−0.276 747	−0.271 579	−0.260 666	−0.267 729	−0.202 755	−0.280 372
	−0.284 990	−0.278 945	−0.275 922	−0.290 185	−0.284 758	−0.273 315	−0.280 723	−0.212 585	−0.293 366
N$_2$	−0.373 677	−0.365 401	−0.361 263	−0.379 937	−0.369 252	−0.360 967	−0.366 609	−0.281 329	−0.390 365
	−0.399 426	−0.392 059	−0.388 376	−0.408 452	−0.395 109	−0.386 274	−0.392 647	−0.298 953	−0.416 402

[a]Where two entries exist, the first refers to valence-only correlated calculations, the other to all-electron calculations.
[b]Using Eq. (28).
[c]Using p_S and p_T values from Ref. 41 (columns 9 and 10 of Table IV).
[d]Using p_S and p_T values obtained variationally from the constrained Feenberg approach utilized in this work.
[e]Using p_S and p_T values obtained from the first root of Eq. (35).
[f]Using p_S and p_T values obtained from the second root of Eq. (35).

sen via the Feenberg scaling. Assuming a pure sigmoid form that reaches the ceiling at the average value of[38] $p_T = 1.90$, one predicts the half-weight to be $p_T = 0.95$ at the origin ($\xi = 0$), in striking accord with the result obtained for HeNe. In turn, the data from Table III of Ref. 38 (solid symbols) are seen to follow the above sigmoid pattern with a range parameter of $g_T = 23$ or so. Since an inverse sigmoid variation pattern is expected for $p_S(\xi)$ as a function of ξ, g_S is fixed equal to g_T and ξ_0 varied until $p_T(\xi)$ attains the value of 0.95 at $\xi = 0$. Such a procedure leads to $\xi_0 = 0.018$, with the data fitted well enough having in mind its coarse statistical nature.[38]

Following the traditional approach,[40,41] the optimum value of ξ (and hence p_T and p_S) may now be obtained by imposing stationarity in Eq. (11),

$$\frac{d[E^{(2)}_{\text{SCS-MP}}(\xi) + E^{(3)}_{\text{SCS-MP}}(\xi)]}{d\xi} = 0, \tag{31}$$

or equivalently,

$$\frac{dp_T(\xi)}{d\xi}[2p_T(\xi)A_{TT} - 2E^T_{\text{MP2}} + 2p_S(\xi)A_{ST}]$$
$$+ \frac{dp_S(\xi)}{d\xi}[2p_S(\xi)A_{SS} - 2E^S_{\text{MP2}} + 2p_T(\xi)A_{ST}] = 0. \tag{32}$$

Although the stationary condition usually implies a vanishing third-order perturbation energy, this may not necessarily be the case now due to the constraints in Eqs. (29) and (30). However, a vanishing third-order perturbation energy is warranted provided that a single weight is kept (a single parameter is determined from one condition). For example, if the
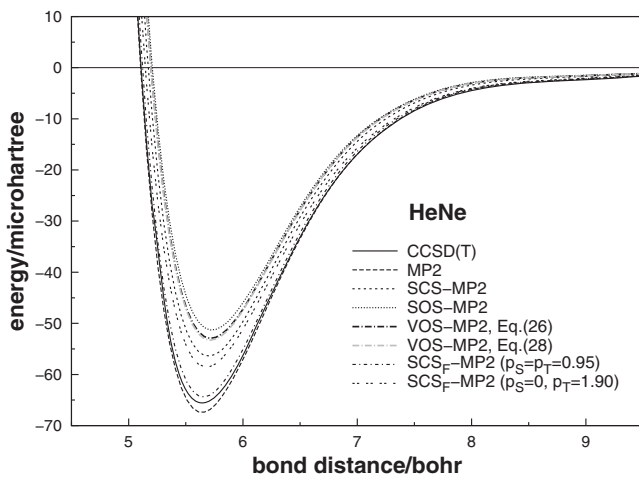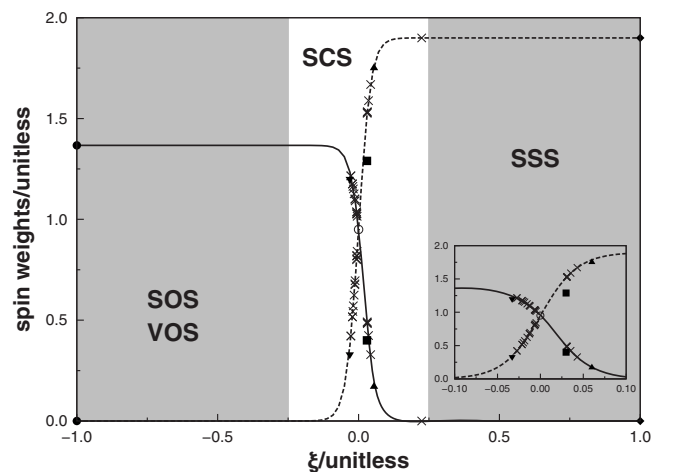

FIG. 4. Interaction potentials for the HeNe dimer as obtained from the methods studied in the present work. The same notation is employed as in Fig. 3. Note that the interaction curve from the non-size-extensive VOS-MP2 theory (shown by the black dashed-dotted line) based on Eq. (26) can hardly be distinguished from the size-extensive one (same style, in gray) that employs Eq. (28).

FIG. 5. Diagram showing the interrelation between SOS, VOS, SCS, and SSS theories. The sigmoidal weight-functions used in the SCS$_F$-MP2 theory are indicated by the solid and dashed lines. Indicated by the solid symbols are results from the literature[32,38] while the crosses refer to values obtained by solving Eq. (35) for ATS2.
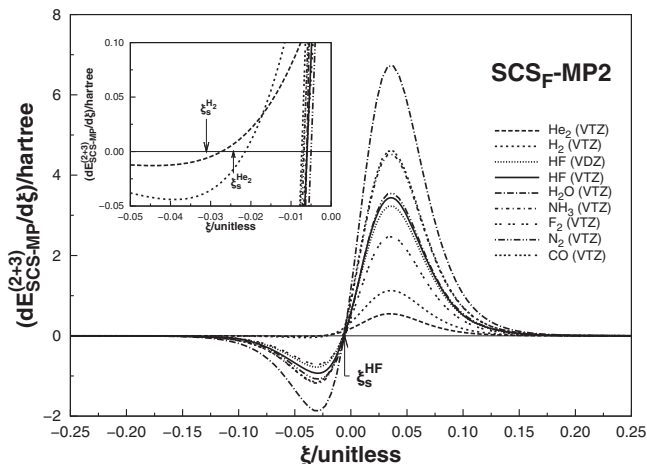
FIG. 6. Roots as obtained from the stationary condition in Eq. (32), and hence indicated as ξ_s^{HF} in the case of hydrogen fluoride. Although less visible in the main plot, a root also exists for He_2 and H_2 as illustrated in the insert.
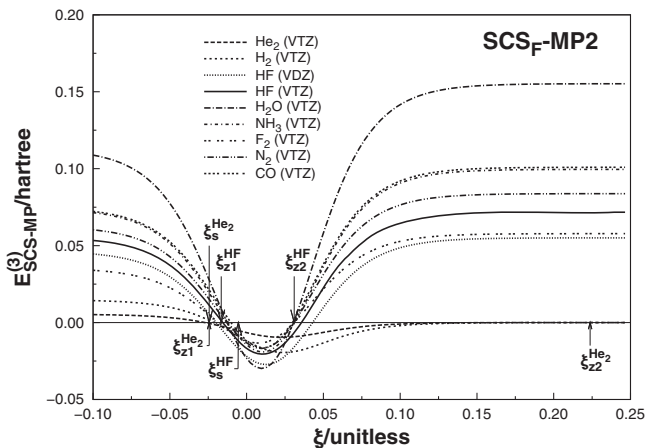

FIG. 7. Roots as obtained from the condition of zeroing the third-order energy. In all cases but H_2 a double root is predicted. They are indicated by the arrows for hydrogen fluoride and the helium dimer. They are labeled for the former as ξ_{z1}^{HF} and ξ_{z2}^{HF}, with a corresponding notation for He_2. Also indicated is the root obtained from the condition of stationarity for HF, ξ_s^{HF}, yielding a non-vanishing third-order energy (correspondingly for He_2 but giving nearly vanishing third-order energy).

contribution to the total correlation energy due to T excitations can be ignored, the terms involving $p_T(\xi)$ in Eq. (32) disappear, yielding

$$\frac{dp_S(\xi)}{d\xi}[2p_S(\xi)A_{SS} - 2E_{MP2}^S] = 0, \tag{33}$$

whose solution can be approximated by

$$\xi_S \to \approx g_S^{-1} \tanh^{-1}\left(1 - \frac{0.001\,61 E_{MP2}^S}{A_{SS}}\right). \tag{34}$$

A corresponding expression is obtained if T excitations are instead ignorable. For mixed regimes, where both p_s and p_T are relevant, the solution must be found by solving Eq. (32). This is the approach followed here, with the results so obtained for a second *ab initio* test set (ATS2) data being given in columns 2–4 of Table IV, and shown in Fig. 6. Of course, a more flexible approach is to treat p_S and p_T as unrelated, with the stationary condition then leading to two independent equations from which the two parameters are determined. This approach has already been utilized,[41] and hence will not be pursued here.

To circumvent the problem raised in the previous paragraph, one may instead impose the condition of vanishing the third-order energy in Eq. (10). This approach has also been utilized in the present work by solving numerically

$$E_{SCS-MP}^{(3)}(\xi) = 0. \tag{35}$$

The roots and weights obtained in this way are collected in columns 5–7 of Table IV, and shown in Fig. 7. In all cases but H_2 (where only a single root exists since $p_T=0$), two roots are obtained for all other systems in the *ab initio* test set (ATS2) that has been utilized in the present work, and which is largely coincident with the one adopted by Szabados.[41] Also given for completeness in column 8 of Table IV are the results obtained from Feenberg's one-parameter approach in Eq. (17). Moreover, columns 9 and 10 gather the results[41] obtained by treating p_S and p_T as unrelated.

Before analyzing the results in Table IV, a few technicalities are in order. The first is to note that all calculations of p_S and p_T here reported have been carried out with the VTZ basis set (except for HF, where the VDZ basis set has also been utilized) but correlating all electrons. Second, Eqs. (4)–(16) have been programed in double precision using the two-center molecular integrals and orbital energies calculated with MOLPRO.[53] The results have been checked by calculating the MP2 and MP3 energies, and found to agree typically to ten decimal places with the output from MOLPRO. Similarly, the agreement observed between the values of p_S and p_T obtained by the variational approach utilized in the present work and by Szabados[41] is quite satisfactory. Note that a single root has been obtained in all cases; Fig. 5. Note further (see Fig. 6) that the solution of the stationary condition in Eq. (32) does not correspond to a vanishing third-order perturbation energy. The correlation energies calculated using such optimized values are collected in Table V, where the results from both valence-only and all-electron calculations are given. We observe that the unsigned correlation energies obtained from the variational procedure utilized in the present work (column 7 of Table V) tend to be somewhat smaller than those of Szabados[41] (column 6), which may be attributed (from a variational point of view) to the fact that p_S and p_T are free to vary in the case of Ref. 41. A similar explanation may be given when comparing the results obtained by choosing the roots that impose a vanishing third-order energy, although the difference is now even smaller in the case of root I (the one leading to a large p_S and a small p_T). In fact, the values so obtained compare well with the values reported in Ref. 41. Recall that the variational procedure yields a vanishing third-order perturbation energy in the case of independent weights, as does in the case of a single root. However, given the restriction imposed by Eqs. (29) and (30), one is bound to expect somewhat smaller unsigned correlation energies as indeed observed. Moreover, due to the opposing nature of the two sigmoids, the set of
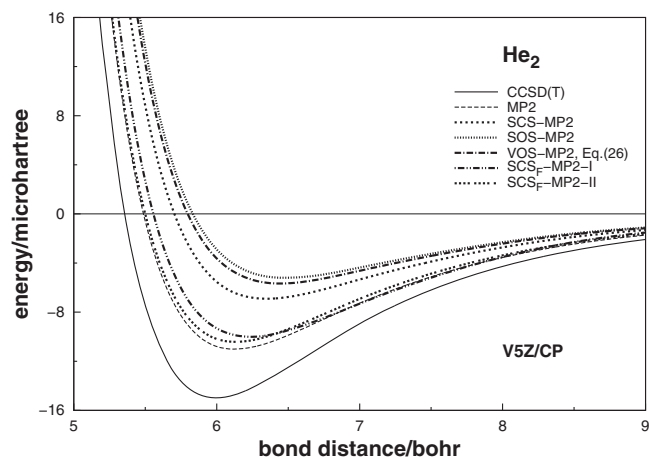
FIG. 8. Counterpoise corrected interaction potentials for the helium dimer as obtained from the methods studied in the present work.

unsigned correlation energies obtained with root II (large p_T and small p_S) is rather small which, as discussed above, does not imply at all poor interaction correlation energies. Root I, yielding a large singlet weight, should therefore be preferred when dealing with strong interactions, while root II should describe better vdW interactions as it favors the triplet electron-pair excitations. To test this, we have done calculations on He$_2$ with the VXZ ($X=T,5$) basis sets. Although much larger basis sets and advanced techniques (often accompanied by correlation extrapolation to the CBS and FCI limits) are known to be required for high accuracy[76,77] (the list of references is by no means extensive; see also references therein), the results here reported may serve to illustrate the role of BSSE which is instead enhanced.

Two sets of calculations have been done. While all results are shown in Fig. 1 of the ESI[61] [without counterpoise correction (NCP) on the left-hand-side panels, CP corrected ones on the right hand side], only the V5Z/CP ones are illustrated in Fig. 8. In both cases, the VOS-MP2 size-extensive formalism based on Eq. (28) has been employed [the differences in the interaction energy are negligible with respect to using Eq. (26) instead]. The notable observation is that the interaction correlation energies obtained with the second root (SCS$_F$-MP2-II) yield quite satisfactory interaction energies for He$_2$ when compared with the corresponding CCSD(T) results. Moreover, when corrected for BSSE, such results lie closest to the MP2 and CCSD(T) curves. Additionally, root I is seen to yield satisfactory results both within NCP and CP approaches. As expected, the VOS-MP2 and SOS-MP2 curves lie close to each other being the ones that recover less of the vdW dispersion interaction as a result of having formally ignored the triplet-pair excitations. Yet, both VOS-MP2 and VOS$_F$-MP2 (this for both roots) theories look quite promising and successful overall. In fact, it is fair to say that the Feenberg-type approach here utilized has the merit of covering the whole range of interactions, from chemically bound- to vdW-molecules, in a fully consistent way.

A final remark to note that the energy of two He atoms at the largest calculated distance ($1680a_0$) coincides with the energy of the dimer to at least eight significant figures, as one would expect from a size-extensive (consistent) method.

However, as noted above, the energy at second-order of perturbation theory may not preserve size-extensivity in Feenberg's approach.[41,48] Nevertheless, a workaround can still be used to exploit the method potentialities: Use it at a single geometry (say, equilibrium), with the scaling parameters so generated being then utilized as input for the SCS-MP2 (SCS-MP2$_F$) method (recall size-extensivity considerations for VOS-MP2). This has actually been the procedure utilized in the present work (Figs. 3, 4, and 8).

V. CONCLUDING REMARKS

The SOS-MP2 theory suggested by Head-Gordon and co-workers has been refined hoping to capture an as large as possible fraction of the correlation energy that is affordable only by more expensive post-Hartree–Fock methodologies such as CCSD(T) or QCISD(T). The corrected MP2 energy assumes the general form

$$E_{\text{VOS-MP2}} = 1.2429\left(1 + 0.1\frac{n-2}{n-1}\right)E_{\text{MP2}}^{\text{S}}$$

or, in a more convenient size-extensive format,

$$E_{\text{VOS-MP2}} = 1.2429(1 + 0.1H[n-2])E_{\text{MP2}}^{\text{S}},$$

where n is the number of electrons, $H[n-2]$ is the Heaviside step function distribution [equal to 0 for $(n-2) \leq 0$; 1 for $(n-2) > 0$], and $E_{\text{MP2}}^{\text{OS}}$ is the calculated raw OS MP2 energy. The new formulation, referred to as VOS-MP2, is applicable to any regions of the molecule configuration space where MP2 converges (preferably near the equilibrium geometry), shows a rather consistent performance in the quality of its predictions regarding the basis set type and correlation treatment that are employed, and benefits of all advantages inherent to SOS-MP2 theory. Thus, it offers an economical, yet reliable, approach even when compared with MP2 itself. In particular, for small systems such as N$_3$ in its lowest quartet state, the reduction in time relatively to CCSD(T) can exceed two orders of magnitude when high-hierarchy basis sets are utilized. Indeed, using various families of Dunning basis sets, CBS extrapolations of the calculated raw energies have allowed to present the most reliable estimates available thus far of the energetics of the N(^4S)+N$_2$ reaction. It is then reasonable to expect that VOS-MP2 theory may also be valuable for studying large systems as more sophisticated approaches will then become prohibitive.

VOS-MP2 as well as other variants of SCS-MP2 theory is known, however, to have difficulties in dealing with interactions of the vdW type such as the weakly bound rare gas dimers, where the attraction is dominated by dispersion forces. To investigate this issue, conventional wisdom has been followed by examining several variants of SCS-MP2 theory designed for handling separately the two regimes of interactions: chemical versus vdW species. In this regard, we have shown that all such variants can be interpreted theoretically and linked smoothly through a coordinate that measures the proportion of S and T electron-pair excitations. Furthermore, a procedure has been proposed to specify such a coordinate via a constrained Feenberg-type approach. Unfortunately, Feenberg's second-order perturbation energy fails

size-extensivity. In spite of this, the scaling parameters emerging from such a treatment at a chosen geometry (say equilibrium) can still provide a good guess for use in the simpler SCS-MP2$_F$ theory. Naturally, although promising, the results reported here are exploratory and demand further investigation. In particular, applications to large systems are most desirable since any possible benefits can surely be regarded as most welcome.

ACKNOWLEDGMENTS

[1] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[2] W. Kohn, A. D. Becke, and R. G. Parr, J. Phys. Chem. **100**, 12974 (1996).

[3] B. J. Lynch, P. I. Fast, M. Harrix, and D. G. Truhlar, J. Phys. Chem. A **104**, 4811 (2000).

[4] B. G. Johnson, C. A. Gonzalez, P. M. W. Gill, and J. A. Pople, Chem. Phys. Lett. **221**, 100 (1994).

[5] J. Gräfenstein and D. Cremer, Theor. Chem. Acc. **123**, 171 (2009).

[6] E. I. Izgorodina, M. L. Coote, and L. Radom, J. Phys. Chem. A **109**, 7558 (2005).

[7] R. J. Bartlett, Chem. Phys. Lett. **484**, 1 (2009).

[8] Chr. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[9] J. S. Binkley and J. A. Pople, Int. J. Quantum Chem. **9**, 229 (1975).

[10] C. J. Cramer, *Essentials of Computational Chemistry*, 2nd ed. (Wiley, Hoboken, 2004).

[11] F. Jensen, *Introduction to Computational Chemistry*, 2nd ed. (Wiley, Chichester, 2006).

[12] M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).

[13] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).

[14] R. A. Friesner, R. B. Murphy, M. D. Beachy, M. N. Rignalda, W. T. Pollard, B. D. Dunietz, and Y. Cao, J. Phys. Chem. A **103**, 1913 (1999).

[15] P. Y. Ayala and G. E. Scuseria, J. Chem. Phys. **110**, 3660 (1999).

[16] P. Y. Ayala and K. N. K. G. E. Scuseria, J. Chem. Phys. **115**, 9698 (2001).

[17] S. Saebø and P. Pulay, Annu. Rev. Phys. Chem. **44**, 213 (1993).

[18] P. E. Maslen and M. Head-Gordon, Chem. Phys. Lett. **283**, 102 (1998).

[19] M. S. Lee, P. E. Maslen, and M. Head-Gordon, J. Chem. Phys. **112**, 3592 (2000).

[20] H. Werner, F. R. Manby, and P. J. Knowles, J. Chem. Phys. **118**, 8149 (2003).

[21] R. A. DiStasio, Jr., Y. S. Yung, and M. Head-Gordon, J. Chem. Theory Comput. **1**, 862 (2005).

[22] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[23] T. Helgaker, J. Gauss, P. Jørgensen, and J. Olsen, J. Chem. Phys. **106**, 6430 (1997).

[24] E. F. C. Byrd, C. D. Sherrill, and M. Head-Gordon, J. Phys. Chem. A **105**, 9736 (2001).

[25] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[26] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[27] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **103**, 4572 (1995).

[28] W. Klopper, Mol. Phys. **99**, 481 (2001).

[29] W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).

[30] A. J. C. Varandas, J. Chem. Phys. **126**, 244105 (2007).

[31] A. J. C. Varandas, J. Phys. Chem. A **112**, 1841 (2008).

[32] S. Grimme, J. Chem. Phys. **118**, 9095 (2003).

[33] M. Gerenkamp and S. Grimme, Chem. Phys. Lett. **392**, 229 (2004).

[34] T. Schwabe and S. Grimme, Acc. Chem. Res. **41**, 569 (2008).

[35] Y. Jung, R. C. Lochan, A. D. Dutoi, and M. Head-Gordon, J. Chem. Phys. **121**, 9793 (2004).

[36] R. C. Lochan, Y. S. Yung, and M. Head-Gordon, J. Phys. Chem. A **109**, 7598 (2005).

[37] Y. Jung and M. Head-Gordon, Phys. Chem. Chem. Phys. **8**, 2831 (2006).

[38] R. A. Distasio, Jr. and M. Head-Gordon, Mol. Phys. **105**, 1073 (2007).

[39] J. G. Hill and J. A. Platts, J. Chem. Theory Comput. **3**, 80 (2007).

[40] E. Feenberg, Phys. Rev. **103**, 1116 (1956).

[41] Á. Szabados, J. Chem. Phys. **125**, 214105 (2006).

[42] F. Neese, T. Schwabe, S. Kossmann, B. Schirmer, and S. Grimme, J. Chem. Theory Comput. **5**, 3060 (2009).

[43] T. Takatani, E. G. Hohenstein, and C. D. Sherrill, J. Chem. Phys. **128**, 124111 (2008).

[44] Y. M. Rhee and M. Head-Gordon, J. Phys. Chem. A **111**, 5314 (2007).

[45] D. Casanova, Y. M. Ree, and M. Head-Gordon, J. Chem. Phys. **128**, 164106 (2008).

[46] Y. M. Rhee, D. Casanova, and M. Head-Gordon, J. Phys. Chem. A **113**, 10564 (2009).

[47] S. Grimme, J. Comput. Chem. **24**, 1529 (2003).

[48] A. Szabados and P. R. Surján, Int. J. Quantum Chem. **101**, 287 (2005).

[49] J. Antony and S. Grimme, J. Phys. Chem. A **111**, 4862 (2007).

[50] T. Takatani and C. D. Sherrill, Phys. Chem. Chem. Phys. **9**, 6106 (2007).

[51] A. Tkatchenko, R. A. DiStasio, Jr., M. Head-Gordon, and M. Scheffler, J. Chem. Phys. **131**, 094106 (2009).

[52] R. P. Steele, R. A. DiStasio, Jr., and M. Head-Gordon, J. Chem. Theory Comput. **5**, 1560 (2009).

[53] MOLPRO, a package of *ab initio* programs, version 2008.3, H.-J. Werner, P. J. Knowles, R. Lindh *et al.*

[54] D. Feller and K. A. Peterson, J. Mol. Struct.: THEOCHEM **400**, 69 (1997).

[55] Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 02/25/04, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multiprogram laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract No. DE-AC06-76RLO 1830. Contact K. Schuchardt for further information.

[56] F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).

[57] W. Wang, W. M. Huo, C. E. Dateo, D. W. Schwenke, and J. R. Stallcop, J. Phys. Chem. A **120**, 132 (2003).

[58] P. Zhang, K. Morokuma, and A. M. Wodtke, J. Chem. Phys. **122**, 014106 (2005).

[59] E. Garcia, A. Saracibar, S. Gomez-Carrasco, and A. Laganá, J. Phys. Chem. A **10**, 2552 (2008).

[60] B. R. L. Galvão and A. J. C. Varandas, J. Phys. Chem. A **113**, 14424 (2009).

[61] See supplementary material at http://dx.doi.org/10.1063/1.3465551 for the whole set of results obtained for trinitrogen in its lowest quartet state, and for the helium dimer.

[62] A. K. Wilson and T. Dunning, Jr. (to be published), cited in Ref. 55.

[63] T. J. Lee and P. R. Taylor, Int. J. Quantum Chem. **S23**, 199 (1989).

[64] M. Horoi, J. R. Gour, M. Włoch, M. D. Lodriguito, B. A. Brown, and P. Piecuch, Phys. Rev. Lett. **98**, 112501 (2007).

[65] M. L. Leininger, I. M. B. Nielsen, T. D. Crawford, and C. L. Janssen, Chem. Phys. Lett. **328**, 431 (2000).

[66] T. J. Lee, Chem. Phys. Lett. **372**, 362 (2003).

[67] A. Karton and J. M. L. Martin, Theor. Chem. Acc. **115**, 330 (2006).

[68] A. J. C. Varandas, Chem. Phys. Lett. **463**, 225 (2008).

[69] A. J. C. Varandas, J. Comput. Chem. **30**, 379 (2009).

[70] A. J. C. Varandas and P. Piecuch, Chem. Phys. Lett. **430**, 448 (2006).

[71] A. J. C. Varandas, J. Chem. Phys. **127**, 114316 (2007).

[72] P. J. S. B. Caridade, B. R. L. Galvão, and A. J. C. Varandas, J. Phys. Chem. A **114**, 6063 (2010).

[73] R. A. Back and J. Y. P. Mui, J. Chem. Phys. **66**, 1362 (1962).

[74] A. Bar-Nun and A. Lifshitz, J. Chem. Phys. **47**, 2878 (1967).

[75] R. K. Lyon, Can. J. Chem. **50**, 1437 (1972).

[76] A. J. C. Varandas, "Extrapolation to the complete basis set limit without counterpoise: The pair potential of helium revisited," J. Phys. Chem. A (to be published), doi:10.1021/jp908835v.

[77] A. J. C. Varandas, Theor. Chem. Acc. **119**, 511 (2008).