Spin Component Scaling in Multiconfiguration Perturbation Theory

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY A · FEBRUARY 2011

Impact Factor: 2.69 · DOI: 10.1021/jp108575a · Source: PubMed

2 AUTHORS, INCLUDING:

Peter Nagy
Eötvös Loránd University
9 PUBLICATIONS **96** CITATIONS

SEE PROFILE

Spin Component Scaling in Multiconfiguration Perturbation Theory

Ágnes Szabados* and Péter Nagy

Laboratory of Theoretical Chemistry, Institute of Chemistry, Loránd Eötvös University, H-1518 Budapest, POB 32, Hungary

Received: September 8, 2010; Revised Manuscript Received: November 20, 2010

We investigate a term-by-term scaling of the second-order energy correction obtained by perturbation theory (PT) starting from a multiconfiguration wave function. The total second-order correction is decomposed into several terms, based on the level and the spin pattern of the excitations. To define individual terms, we extend the same spin/different spin categorization of spin component scaling in various ways. When needed, identification of the excitation level is facilitated by the pivot determinant underlying the multiconfiguration PT framework. Scaling factors are determined from the stationary condition of the total energy calculated up to order 3. The decomposition schemes are tested numerically on the example of bond dissociation profiles and energy differences. We conclude that Grimme's parameters determined for single-reference Møller−Plesset theory may give a modest error reduction along the entire potential surface, if adopting a multireference based PT formulation. Scaling factors obtained from the stationary condition show relatively large variation with molecular geometry, at the same time they are more efficient in reducing the error when following a bond dissociation process.

I. Introduction

Second-order Møller−Plesset perturbation theory (MP2)[1] is a highly successful tool of computational chemistry used for describing ground state electronic structure and properties of molecules. Ample use of MP2 over the years led to the accumulation of a broad experience on the general performance of the theory. Extensive developments in the line of reducing the calculation cost have helped to extend its applicability to systems of considerable size.[2−16] Apart from calculation time reduction, improvement of the accuracy of MP2 has also been attempted by various means.[17,18] Among these techniques, spin component scaled (SCS) MP2 proposed by Grimme[19] is outstanding in its spread and success. It has lived many applications[20−22] and initiated further developments toward a reduced cost approximate method called scaled opposite spin MP2,[23,24] reparametrizations for intermolecular interactions,[25−27] and excited state calculations.[28−30] Spin-contamination in the underlying perturbed functions has been recently pinpointed by Fink,[31] who also suggested an alternative two-parameter scaling which conserves the spin. Introduction of Grimme-type scaling factors was also investigated within the coupled-cluster framework.[32] An extension of spin-component-scaling to near equilibrium geometries was recently suggested by Varandas, who assigned geometry and particle number dependence to the scaling factors.[33] Popularity of SCS-MP2 lies with the fact that it is easy to implement, outperforms MP2 in numerical terms, and does not bring any cost increase or loss of the important characteristics of MP2, like size extensivity and invariance to degenerate orbital rotations.

Since the Hartree−Fock (HF) wave function is the zero-order behind MP2, its applicability is limited to systems for which the HF approximation is meaningful. Being derived from MP2, the same holds for SCS-MP2. To correctly describe molecular systems with a significant amount of static correlation—e.g., bond fission, biradicals, or transition metal complexes of open shell character—one may turn to approaches based on a multireference (MR) initial function, like MR PT. In general MR PTs are less costly but also less accurate than more elaborated MR coupled-cluster or MR configuration interaction methods. For this reason it would be desirable to improve the performance of MR PTs by a scaling similar to Grimme's. This would mean the application of a few parameters, that reduce the error without any increase in calculation time.

It is however not straightforward to generalize the scaling by Grimme for the MR case. The reason, on one hand, is the question of categorization. As there are several determinants building the reference function, identification of the excitation pattern is ambiguous, i.e., an excited determinant may arise by parallel spin excitation with respect to one but by antiparallel spin excitation with respect to another. For this reason Robinson and McDouall scaled only those components which can be classified unambiguously, when experimenting with Grimme's factors in their MRMP study.[34] A second difficulty is the determination of generally applicable values for the parameters. The statistical approach adopted by Grimme—which was confined for molecules at equilibrium geometry—would be much harder to do for molecular systems of any geometry.

Our intention in the present study is to explore various scaling possibilities in the framework of multiconfiguration (MC) PT.[35,36] The pivot determinant underlying MCPT makes the theory well suited for spin-component scaling, since the problem of excitation pattern identification is not present. To determine the value of the scaling parameters, we turn to the approach suggested by Feenberg[37,38] which means a minimization condition imposed on the energy written up to order 3. This way extensive statistics for obtaining the scaling parameters are avoided. Generalized for the case of two parameters, Feenberg's procedure has been shown to give energies and parameters similar to those obtained by Grimme.[39] At the same time, Feenberg's scaling necessitates the computation of the third-order energy, which brings a considerable increase in calculation time as compared to the second order. Obviously, the ultimate

* To whom correspondence should be addressed, szabados@chem.elte.hu.

10.1021/jp108575a
Published on Web 12/29/2010

goal would be the deduction of universal, system-independent values for the parameters, in order to keep the computational cost relatively low.

Apart from its advantages, the presence of a pivot determinant in the theory also has a downside. Following a potential energy curve, it may happen that the pivot determinant has to be changed as commanded by the weights of the determinants in the reference function. This may produce a defective discontinuity at the point of the change. Redefinition of the pivot determinant is problematic also because it represents the origin of excitation level identification. To overcome this difficulty we investigate the application of averaging the wave function and energy corrections over all determinants present in the reference function.[40] The averaging technique is combined with an appropriate definition for one-particle energies following Malrieu,[41] to avoid quasi-degeneracies when picking a determinant of small weight in the reference as pivot determinant.

In order to exclude any specific effect that may arise from the MCPT formalism, we also present test calculations within the complete active space (CAS) reference function based multireference PT (MRPT) first advocated by Davidson.[42] The two theories—MRPT and MCPT—are related, the main difference being the treatment of overlap among zero-order functions of the many-electron space. While in MCPT a biorthogonal approach is adopted, MRPT considers all roots of the CAS problem and determinants orthogonal to the CAS space as zero order functions.

When working with just a handful of scaling parameters, it is essential to collect those determinants behind a common scaling factor, which have in some sense a similar contribution. Grimme's intuition on distinguishing parallel and antiparallel spin double excitations has proved appropriate in this respect. There are several possibilities for transplanting this idea to the MR case. One may focus, e.g., on the excitation level. In this case a decision is called for on the scaling strategy of excitations different from doubles. This is obvious, since singly excited determinants enter the energy at order 2, while excitations up to quadruples may figure at order 3 in MCPT. Concerning doubly excited determinants, we keep the singlet−triplet type separation advocated by Grimme.

An alternative way of generalizing spin-component scaling may be based on the sum of s_z eigenvalues. In the single reference case these are 0 in both the occupied and virtual subspaces for the different-spin excitation, while we get ± 1 in the two subspaces for the same-spin category. Focusing on this property, one may take the distinct orbital subspaces underlying the reference function as a starting point. The excitations attainable at the second- and third-order energy formula can be grouped into categories, based on the sum of s_z eigenvalues in the different subspaces.

Considering either the excitation level based categorization or the grouping based on the sum of s_z eigenvalues, we explore various possibilities of merging different groups and study the variation of scaled energies and scaling parameters.

II. Theory

A. Multiconfiguration Perturbation Theory (MCPT). A brief overview of MCPT necessary to develop the formulation of the scaled theory is presented below.

Let us suppose that $\Psi^{(0)}$ is a normalized wave function of multireference nature, which can be expanded in terms of determinants as

$$|\Psi^{(0)}\rangle = d_1|1\rangle + \sum_{k=2}^{M} d_k|k\rangle \tag{1}$$

In the above expression M refers to the number of determinants with nonzero coefficient d_k in $\Psi^{(0)}$. One may regard the set of these determinants as an M-dimensional reference space, which will be referred to as $\mathscr{R}$. Determinant $|1\rangle$ is written separately in eq 1 to indicate its pivot role. Instead of the normalized function $\Psi^{(0)}$, we are going to consider

$$\Psi^{(0)}_{N1} = \Psi^{(0)}/d_1$$

as the zero-order function (cf. intermediate normalization), and we will keep the pivot coefficient fixed to 1 in the PT treatment.

When constructing a correction to the zero-order function by Rayleigh−Schrödinger (RS) PT, one has to make a choice for the excited state functions and energies. In one variant of MCPT[36] excited zero-order functions are the determinants, excluding the pivot determinant

$$|k\rangle, k = 2, ..., N$$

where N is the dimension of the many-electron space. Energies corresponding to the excited determinants are defined in such a way that zero-order excitation energies take a form similar to single-reference MP theory

$$\Delta_k = E^{(0)}_k - E^{(0)} = \varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j$$

if, e.g., $|k\rangle = a_b^+ a_a^+ a_i a_j|1\rangle$. The reference energy $E^{(0)}$ and one-particle energies ε_i are to be specified later. Perturbation corrections are developed by supposing a zero-order Hamiltonian diagonal on the basis of vectors $|\Psi^{(0)}_{N1}\rangle$ and $|k\rangle$ ($k > 1$), with the associated zero-order energies $E^{(0)}$ and $E^{(0)}_k$

$$\hat{H}^{(0)} = E^{(0)}|\Psi^{(0)}_{N1}\rangle\langle 1| + \sum_{k=2}^{N} E^{(0)}_k|k\rangle\langle\tilde{k}| \tag{2}$$

Note, that the zero-order Hamiltonian is non-Hermitean. This is a consequence of the biorthogonal treatment of the overlap of basis vectors $|\Psi^{(0)}_{N1}\rangle$ and $|k\rangle$. Definition of the reciprocal functions, which appear in the bra vectors is

$$\langle\tilde{k}| = \langle k| - \frac{d_k}{d_1}\langle 1| \tag{3}$$

Function $\langle\tilde{k}|$ is different from $\langle k|$ only if $k \in \mathscr{R}$. The reciprocal function corresponding to $\Psi^{(0)}_{N1}$ is the pivot determinant, $\langle 1|$. For more details we refer to previous reports on the subject.[35,36,43]

A definition of the zero-order energy suited for the biorthogonal vector sets is the nonsymmetrical expression

$$E^{(0)} = \langle 1|\hat{H}|\Psi^{(0)}_{N1}\rangle$$

Based on the zero-order Hamiltonian eq 2, biorthogonal RS theory gives the following second- and third-order energy corrections

$$E^{(2)} = -\sum_{k=2}^{N} \frac{\langle 1|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|\Psi_{N1}^{(0)}\rangle}{\Delta_k} \tag{4}$$

and

$$E^{(3)} = \sum_{k,l=2}^{N} \frac{\langle 1|\hat{H}|k\rangle\langle\tilde{k}|\hat{H} - \hat{H}^{(0)}|l\rangle\langle\tilde{l}|\hat{H}|\Psi_{N1}^{(0)}\rangle}{\Delta_k\Delta_l} \tag{5}$$

An analysis of excitation levels contributing to the above expressions reveals that $|k\rangle$ may be singly or doubly excited with respect to $|1\rangle$ in eq 4. The same holds for $|k\rangle$ in eq 5, while $|l\rangle$ can be excited between 1× and 4×. On the basis of eq 3 we consider the excitation level of vector $\langle\tilde{k}|$ to agree with that of $|k\rangle$.

1. One-Particle Energies. It is the specification of one-particle energies ε_i that completes the definition of the zero-order eq 2 and, consequently, the partitioning. In the present study we calculate the energy of an occupied level as generalized ionization potential (IP)

$$\varepsilon_i^{\oplus} = \langle\Psi^{(0)}|\hat{H}|\Psi^{(0)}\rangle/\langle\Psi^{(0)}|\Psi^{(0)}\rangle - \langle\Psi^{(0)}|a_i^{+}\hat{H}a_i|\Psi^{(0)}\rangle/\langle\Psi^{(0)}|a_i^{+}a_i|\Psi^{(0)}\rangle \tag{6}$$

while the energy of a virtual level is obtained as generalized electron affinity (EA):

$$\varepsilon_i^{\ominus} = \langle\Psi^{(0)}|a_i\hat{H}a_i^{+}|\Psi^{(0)}\rangle/\langle\Psi^{(0)}|a_ia_i^{+}|\Psi^{(0)}\rangle - \langle\Psi^{(0)}|\hat{H}|\Psi^{(0)}\rangle/\langle\Psi^{(0)}|\Psi^{(0)}\rangle \tag{7}$$

Index i refers to spin orbitals, and occupied-virtual classification of orbitals is based on the pivot determinant. The advantage of definitions (6) and (7) as orbital energies—referred to as IP/EA partitioning—over Fockian diagonals applied earlier[35,36,43] is that occupied and virtual levels are less likely to get close to each other.[41] One-particle energies computed as Fockian diagonals may become close to zero especially for occupied orbitals of small occupation number or virtual orbitals of large occupation number. A demonstration for such an ill effect showing at order 3 is given in Figure 1 of the Supporting Information. Generalization of MP partitioning as in eqs 6 and 7 or by taking the diagonals of the Koopmans matrix and its EA analogue[43] puts more stress on the IP and EA interpretation of orbital energies. Numerical experience shows that this can successfully increase the gap in difficult situations.[41]

The zero-order excitation energy in the IP/EA partitioning takes the form

$$\Delta_k^{\mathrm{IP/EA}} = E_k^{(0)} - E^{(0)} = \varepsilon_a^{\ominus} + \varepsilon_b^{\ominus} - \varepsilon_i^{\oplus} - \varepsilon_j^{\oplus}$$

for, e.g., determinant $|k\rangle = a_b^{+}a_a^{+}a_ia_j|1\rangle$.

2. Fermi-Vacuum Invariance (MCPT-avg). Pivot dependence of a theory may cause ill effects when following a potential energy surface. To avoid these problems an averaging procedure may be applied.[40] This involves computation of energy corrections by all possible choices for the pivot determinant, followed by the construction of weighted sums. Weight factors are given by the squared coefficients of the normalized reference function. For example the zero-order energy takes the form

$$E_{\mathrm{avg}}^{(0)} = \sum_{i\in\mathscr{R}} d_i^2\langle i|\hat{H}|\Psi_{Ni}^{(0)}\rangle = \langle\Psi^{(0)}|\hat{H}|\Psi^{(0)}\rangle$$

The first-order energy is zero in the averaged formulation, similarly to the unaveraged version of the theory. The second- and third-order corrections become

$$E_{\mathrm{avg}}^{(2)} = -\sum_{i\in\mathscr{R}} d_i^2 \sum_{\substack{k=1\\k\neq i}}^{N} \frac{\langle i|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|\Psi_{Ni}^{(0)}\rangle}{\Delta_k} \tag{8}$$

and

$$E_{\mathrm{avg}}^{(3)} = \sum_{i\in\mathscr{R}} d_i^2 \sum_{\substack{k,l=1\\k,l\neq i}}^{N} \frac{\langle i|\hat{H}|k\rangle\langle\tilde{k}|\hat{H} - \hat{H}^{(0)}|l\rangle\langle\tilde{l}|\hat{H}|\Psi_{Ni}^{(0)}\rangle}{\Delta_k\Delta_l} \tag{9}$$

Note, that the sum for determinants i of the model space $\mathscr{R}$ cannot be interchanged with the sum for k and l in eqs 8 and 9. The reason is not only that k and l must not coincide with i but also that reciprocal vectors $\langle\tilde{k}|$ and $\langle\tilde{l}|$, denominators Δ_k and Δ_l, and the zero-order operator $\hat{H}^{(0)}$ depend on index i.

B. Scaling. In analogy with SCS MP2, we introduce scaling parameters, associated with groups of excited determinants. Several grouping schemes are going to be investigated, as detailed in section C. Whatever the idea is behind the definition of excitation groups, the scaling involves a redefinition of the zero-order Hamiltonian in the form

$$\hat{H}^{(0)}(\mathbf{p}) = E^{(0)}|\Psi_{N1}^{(0)}\rangle\langle 1| + \sum_{I=1}^{g}\sum_{k\in I} E_k^{(0)}(\mathrm{p}_I)|k\rangle\langle\tilde{k}|$$

$$E_k^{(0)}(\mathrm{p}_I) = E_k^{(0)}\frac{1}{\mathrm{p}_I} + E^{(0)}\frac{\mathrm{p}_I - 1}{\mathrm{p}_I} \tag{10}$$

$$I = 1, ..., g$$

g is the number of groups and vector $\mathbf{p}(p_1, ..., p_g)$ collects the scaling parameters. (We note at this point, that a scaling in the generalized MP partitioning of MCPT[44] is also possible to formulate, e.g. by using the zero-order Hamiltonian given by Fink.[31]) The reason for introducing parameters in the peculiar form of eq 10 is to get terms of the energy correction scaled by p_I's. On the basis of the zero-order Hamiltonian eq 10, PT corrections in the scaled partitioning can be expressed as

$$E^{(2)}(\mathbf{p}) = -\sum_{I=1}^{g} B_I p_I \tag{11}$$

$$E^{(3)}(\mathbf{p}) = \sum_{I,J=1}^{g} p_I A_{IJ} p_J - \sum_{I=1}^{g} B_I p_I \tag{12}$$

where vector $\mathbf{B}$ and matrix $\mathbf{A}$ collect terms of the unscaled second- and third-order energy, respectively. For example

$$B_I = \sum_{k\in I} \frac{\langle 1|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|\Psi_{N1}^{(0)}\rangle}{\Delta_k} \tag{13}$$

and

$$A_{IJ} = \sum_{k\in I}\sum_{l\in J} \frac{\langle 1|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|l\rangle\langle\tilde{l}|\hat{H}|\Psi_{N1}^{(0)}\rangle}{\Delta_k\Delta_l}$$

Other elements B_I and A_{IJ} can be constructed analogously. Due to the non-Hermitean character of the zero-order Hamiltonian, in general $A_{IJ} \neq A_{JI}$.

To determine the value of the scaling parameters, we follow the idea first proposed by Feenberg and Goldhammer.[37,38] Let us consider the energy as a function of the scaling parameters

$$E = \sum_{i=0}^{3} E^{(i)}(\mathbf{p}) + \mathrm{O}^{(4)}(\mathbf{p})$$

and require that[45,46]

$$\frac{\partial(E^{(2)}(\mathbf{p}) + E^{(3)}(\mathbf{p}))}{\partial \mathbf{p}} = 0 \tag{14}$$

Evaluating the derivatives, condition (14) leads to the linear system of equations

$$\boldsymbol{\alpha}\mathbf{p} = \boldsymbol{\beta} \tag{15}$$

with elements of matrix $\boldsymbol{\alpha}$ being

$$\alpha_{IJ} = \frac{A_{IJ} + A_{JI}}{2}$$

and vector $\boldsymbol{\beta}$ agreeing with vector $\mathbf{B}$

$$\beta_I = B_I$$

Having determined parameter vector $\mathbf{p}$ via inversion of matrix $\boldsymbol{\alpha}$, the scaled energies can be computed by substitution into eqs 11 and 12 to get

$$E^{(2)}_{\mathrm{scaled}} = -\mathbf{B}^T\boldsymbol{\alpha}^{-1}\mathbf{B} \tag{16}$$

and

$$E^{(3)}_{\mathrm{scaled}} = \mathbf{B}^T\boldsymbol{\alpha}^{-1}(\mathbf{A} - \boldsymbol{\alpha})\boldsymbol{\alpha}^{-1}\mathbf{B} \tag{17}$$

We are going to explore scaling strategies where some terms of the second- and third-order energy remain unscaled. In such circumstances, matrix $\boldsymbol{\alpha}$ is truncated to involve only those groups for which a parameter is to be computed. Vector $\boldsymbol{\beta}$ is changed along with, according to

$$\beta_I = B_I - \sum_{J}^{\text{unscaled}} \frac{A_{IJ} + A_{JI}}{2}$$

with I denoting groups to be scaled. When computing the energies of eqs 11 and 12 $p_J = 1$ is applied if group J is unscaled, and the full matrix $\boldsymbol{\alpha}$ is used.

Rewriting eq 12 in the form

$$E^{(3)}(\mathbf{p}) = \mathbf{p}^T(\boldsymbol{\alpha}\mathbf{p} - \mathbf{B})$$

one may realize that the third-order energy vanishes in the scaled partitioning since the term in round brackets on the right-hand side is set to zero when solving eq 15. Obviously, $E^{(3)}(\boldsymbol{p})$ is zero only if all parameters are determined from eq 15. If $p_I = 1$ for some I, there is a nonzero third order-contribution to the energy.

A scaling based on condition eq 15 is obviously the time-determining step of the present theory. Computation of the elements of matrix $\boldsymbol{\alpha}$ agrees with the cost of the third-order energy. This is proportional to $n_v^4 n_o^2$ in the case of MCPT and increases to $Mn_v^4 n_o^2$ for MCPT-avg, n_o (n_v) denoting the number of occupied (virtual) spin orbitals in the pivot determinant. These time requirements apply for a code which loops for orbitals, based on many-body expressions or diagrams, along the lines suggested, e.g., by Hose and Kaldor.[47]

C. Grouping Strategies. When collecting excited determinants behind a common scaling factor, we either focus on the s_z pattern of the excitation or on the excitation level.

In the former case the absolute values of the sum of s_z eigenvalues in distinct orbital subspaces are computed for each excited determinant. The orbital subspaces are naturally defined as doubly occupied, active, and virtual in the case of a CAS reference function. For an antisymmetrized product of strongly orthogonal geminals (APSG) function, geminal orbital subspaces seem to be appropriate in addition to the doubly occupied and the virtual subspace. If considering three orbital subspaces, the determinants entering the second- and third-order energy formula may give rise to the $|\sum_i' s_{zi}|$ values indicated in Table 1. To keep the number of scaling parameters relatively low, we always combine the different cases into at most 5 groups, as indicated in the table.

A different grouping strategy may keep the excitation level in primary focus. Here we rely on the pivot determinant to determine the excitation level and define Grimme's categories for doubles: S for antiparallel-spin and T for parallel-spin.[19,39] (Note, that the antiparallel and parallel spin double excitations are not eigenfunctions of the spin-squared operator, hence they are neither pure triplets nor singlets.) In addition to these, we introduce one parameter assigned to single excitations (R) and another associated with determinants of excitation of level 3 or higher (U). There are at maximum four parameters in this scheme. A special feature of this categorization is that $A_{UI} = 0$, $I = R, S, T$ as well as $B_U = A_{UU} = 0$. (We prefer labeling of excitation groups by letters instead of numbers, to help the identification when presenting the numerical results.)

Both the s_z pattern and the excitation level based grouping are related to Grimme's original proposition. The former is a straightforward extension of Grimme's idea, while in the latter

TABLE 1: Absolute Value of the Sum of s_z Eigenvalues in Distinct Orbital Subspaces for Determinants Entering the Second- and Third-Order Energy Correction in Either MRPT or MCPT[a]

		T_1			T_{2a}			T_{2b}		
	S	case I	case II	case III	case I	case II	case III	case I	case II	T_{2c}
$\lvert\sum_i^{\text{doubly occ.}} s_{zi}\rvert$	0	0	0.5	0.5	0	1	1	0.5	1	0.5
$\lvert\sum_i^{\text{active}} s_{zi}\rvert$	0	0.5	0	0.5	1	0	1	0.5	0.5	1
$\lvert\sum_i^{\text{virtual}} s_{zi}\rvert$	0	0.5	0.5	0	1	1	0	1	0.5	0.5

[a] Label of the given excitation group used in section III is given in the table heading. Orbital categorization is based on their occupation in the CAS space determinants.

we strictly keep the S and T type doubly excited determinants. Note, that there is no one-to-one correspondence between groups in the s_z pattern and the excitation level based categorization. For example, determinants of excitation level based group S may belong to group S, T_1, or T_{2c} in the s_z pattern based scheme. Group R forms a subspace of group T_1.

An important difference between the two grouping strategies is that the s_z pattern based scheme is characteristic for the excited determinant itself and lacks any reference to the pivot determinant. At contrast with this, excitation level based scheme relies on the pivot determinant. As a consequence classification of the determinants is changed when changing the pivot determinant. This feature is undesirable in Fermi-vacuum invariant MCPT. For this reason, when combining the excitation level based categorization with MCPT-avg, we consider the whole model space when assigning determinants to groups. To obtain the weight of determinant $|k\rangle$ in group R for example, we calculate

$$v(R,k) = \sum_{\substack{l=1 \\ k \text{ is 1}\times\text{ex. wrt } l}}^{M} d_l^2$$

and similarly for categories S, T, and U. Note, that case $k = l$ does not fall into any of the groups; hence the sum $\sum_{I=R,S,T,U} v(I,k)$ is not equal to 1. For this reason the normalized weight factors are obtained as

$$w(I,k) = \frac{v(I,k)}{\sum_{J=R,S,T,U} v(J,k)}, \quad I = R,S,T,U$$

The expression of vector **B** and matrix **A** in the scaled and averaged treatment take the form

$$B_I = \sum_{i\in\mathscr{R}} d_i^2 \sum_{\substack{k=1 \\ k\neq i}}^{N} w(I,k) \frac{\langle i|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|\Psi_{Ni}^{(0)}\rangle}{\Delta_k}$$

and

$$A_{IJ} = \sum_{i\in\mathscr{R}} d_i^2 \sum_{\substack{k=1 \\ k\neq i}}^{N} w(I,k) \left(\frac{\langle i|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|\Psi_{Ni}^{(0)}\rangle}{\Delta_k^2} + \sum_{\substack{l=1 \\ l\neq i,k}}^{N} w(J,l) \frac{\langle i|\hat{H}|k\rangle\langle\tilde{k}|\hat{H}|l\rangle\langle\tilde{l}|\hat{H}|\Psi_{Ni}^{(0)}\rangle}{\Delta_k\Delta_l} \right)$$

III. Examples

A numerical test of scaling strategies applied in different PT frameworks is presented below on covalent bond dissociation energy profiles, the singlet−triplet gap of the CH_2 molecule, and the barrier of HCN to HNC isomerization. We use rather small basis sets so that the full configuration interaction (FCI) results are attainable as benchmark. The calculations are perfomed using a transparent but inefficient code that loops on configurations instead of orbitals. Since this makes the computation time longer than the ideal, no timings are reported for these pilot studies.

Scaling strategies are denoted by giving the labels of the groups that obtain a parameter determined from condition (15). Underscores are used to separate labels that form individual groups. For example notation S_T refers to a two-parameter scaling, with all other scaling parameters set to 1. Notation S_TRU also stands for a two-parameter scaling, but in this case one parameter is applied for antiparallel spin doubles and one for all other excitations. In many cases we will display results obtained by applying Grimme's two parameters $p_S = 1.2$ and $p_T = 1/3$, under the acronym SCS-MCPT2. In this scheme $p_R = p_U = 1$.

To get an impression on the role of the theoretical formalism, we present some numerical results within the multireference PT framework originally proposed by Davidson[42] and later pursued by, e.g., Mitrushenkov.[48] The formalism is also closely related to the MRMP approach of Hirao,[49,50] which was the method of choice by McDouall et al. when testing SCS-MP in the multireference case.[34] In this approach higher roots of the CAS problem are taken as zero-order excited functions. In the space orthogonal to the model space, determinants are considered, similarly to MCPT. The main difference between MRPT and MCPT is that the former operates with a set of orthogonal functions and does not rely on any procedures of treating an overlap matrix. Due to this fact, its applicability is restricted to functions that arise as the solution of the linear variation problem within a subspace of the many-electron space, like a CAS function for example. Regarding, e.g., the first-order wave function, the difference between MRPT and MCPT manifests in the coefficients of the determinants of the CAS space. These are zero in MRPT while MCPT gives a relaxation stemming from the treatment of overlap. For MRPT we apply either the excitation level based categorization averaged over all determinants in the CAS space as discussed in section C or study the grouping based on the s_z pattern.

With some of the scaling strategies it may occur that the components of the third-order energy are distributed unevenly in matrix **A**. This mainly affects the excitation level based grouping where, e.g., singly excited determinants typically give negligible contribution as compared with other groups. If the corresponding eigenvalue of matrix $\boldsymbol{\alpha}$ is close to zero, it may even cause numerical instabilities when computing $\boldsymbol{\alpha}^{-1}$. To deal

TABLE 2: Upper and Lower Bounds to the Feenberg Scaled Second-Order PT Term Based on the Sign and Relation of the Original Second- and Third-Order Energy[a]

	$E^{(2)}_{FE} = (E^{(2)})^2/(E^{(2)} - E^{(3)})$		
	case I	case II	case III
if	$E^{(3)} > 0$	$E^{(3)} < 0$ and $E^{(2)} < (r/(r-1))\cdot E^{(3)},\ r \in \mathbb{R},\ 1 < r$	$E^{(3)} < 0$ and $E^{(3)} < E^{(2)}$
then	$E^{(2)} < E^{(2)}_{FE} < 0$	$r\cdot E^{(2)} < E^{(2)}_{FE} < E^{(2)} + E^{(3)}$	$0 < E^{(2)}_{FE}$

[a] It is assumed that $E^{(2)} < 0$.

with these cases, we apply a damping when computing the inverse of $\boldsymbol{\alpha}$ according to

$$\boldsymbol{\alpha}^{-1'} = \sum_{\mu} \frac{a_\mu}{a_\mu^{\,2} + \omega^2} \mathbf{c}_\mu \otimes \mathbf{c}_\mu \qquad (18)$$

where ω is a small damping parameter and a_μ and $\mathbf{c}_\mu$ are the eigenvalues and eigenvectors of $\boldsymbol{\alpha}$, respectively. Expression 18 is a regularization according to Tikhonov that was successfully applied recently, when solving linearized coupled-cluster (LCC) equations.[51] A common characteristic of the present problem and LCC is that both may give an unrealistic result in cases where the inverse of the coefficient matrix is not yet ill-defined, but there are orders of magnitude differences among its eigenvalues. We note here that there is also a formal relation between scaled PT and LCC. If starting from MP2, assigning each doubly excited determinant a separate parameter, and determining them from Feenberg's condition, one arrives at the linearized coupled-cluster doubles (LCCD) approximation as the first-order wave function and second-order energy.[46]

When analyzing numerical results, we find it useful to compare multiparameter scaling with the single-parameter Feenberg scheme. This latter approach has its advantages and pitfalls depending on the relation of the initial second- and third-order energy terms. It is possible to judge the behavior of Feenberg scaling by giving upper and lower bounds to the Feenberg-scaled second order, as collected in Table 2. From the table it is apparent that one-parameter scaling with a parameter value by Feenberg's condition is obviously successful in case I, while it is a complete failure in case III. The performance in case II is good, as long as $E^{(0)} + E^{(1)} + rE^{(2)}$ does not seriously overshoot the exact result.

A. BH Molecule. On the example of the BH molecule, we demonstrate the typical behavior of scaled methods in a situation where the third-order energy represents a considerable improvement over order two and both are upper bounds to the exact energy. Single parameter Feenberg scaling is expected to perform well in these circumstances, case II of Table 2 applies with $r = 2$. As basis we take Pople's 6-31G** set. A CAS(2,2) function serves as reference for the MCPT treatment.

In Figure 1a the error of the second- and third-order energy is plotted as a function of the internuclear distance. Since there are merely two determinants contributing to the reference function, pivot dependent MCPT and MCPT-avg give similar results. Curves by MRPT deviate from MCPT at around 2 Å and give somewhat larger errors in the broken bond limit.

Excitation level based scalings are presented in parts b and c of Figure 1. For comparison we include the excitation level based scaling computed with Grimme's parameters in the pivot-dependent MCPT framework. This curve, labeled SCS MCPT2, successfully reduces the error of the original second order by a

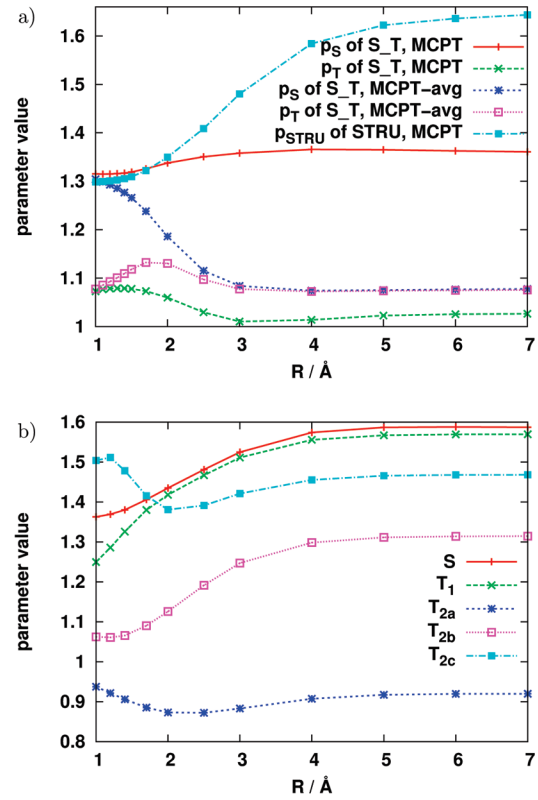

Figure 1. Energy difference curves (with respect to FCI) of the BH dissociation computed by MCPT methods, in Pople's 6-31G** basis. Reference function underlying the PT approximations is a CAS(2,2) function. Acronym "SCS" refers to Grimme's two parameter scaling with $p_S = 1.2$ and $p_T = 1/3$, within the pivot-dependent MCPT framework. Scalings S, S_T, and STRU belong to the excitation level based categorization, S_T_1_T_{2a}_T_{2b}_T_{2c} refers to an s_z pattern based categorization. (a) Unscaled second- and third-order results. (b) Scaled results in the MCPT framework. (c) Scaled results in the MCPT-avg framework.

rough 30% and improves parallelity. Regarding scalings obtained from eqs 15 and 16, it is apparent that the results of S and S_T schemes are very close. They definitely represent an improvement over the unscaled results in pivot-dependent MCPT. In MCPT-avg this holds only at order two. Errors are better by single parameter Feenberg scaling (STRU) but the parallelity of the curve is deteriorated. Schemes S_T_R_U, S_T_RU, and S_TRU are not shown in the figures, as their results are indistinguishable from STRU on the scale of Figure 1c. The same scalings give far too large corrections—even if applying the damping of eq 18—in the pivot-dependent MCPT framework. These are omitted from Figure 1b for clarity but included in the Supporting Information with a demonstrative example of the effect of the damping parameter ω.

A five-parameter scaling based on the s_z pattern is presented in Figure 1c, computed with MRPT. Second and third orders agree in this case, since all parameters were determined from Feenberg's condition. Performance of S_T_1_T_{2a}_T_{2b}_T_{2c} is comparable to the third order of excitation level based S or S_T scaling in MCPT-avg.

Inspecting the variation of scaling parameters with geometry shown in Figure 2, we see that parameters of the excitation level based scaling S_T are fairly constant in the pivot dependent framework. (Parameter p_S is practically the same in schemes S_T and S; it is therefore not indicated twice.) When the averaged MCPT framework or including group U is applied in the parameter determination, p_S or p_{STRU} varies in a broader

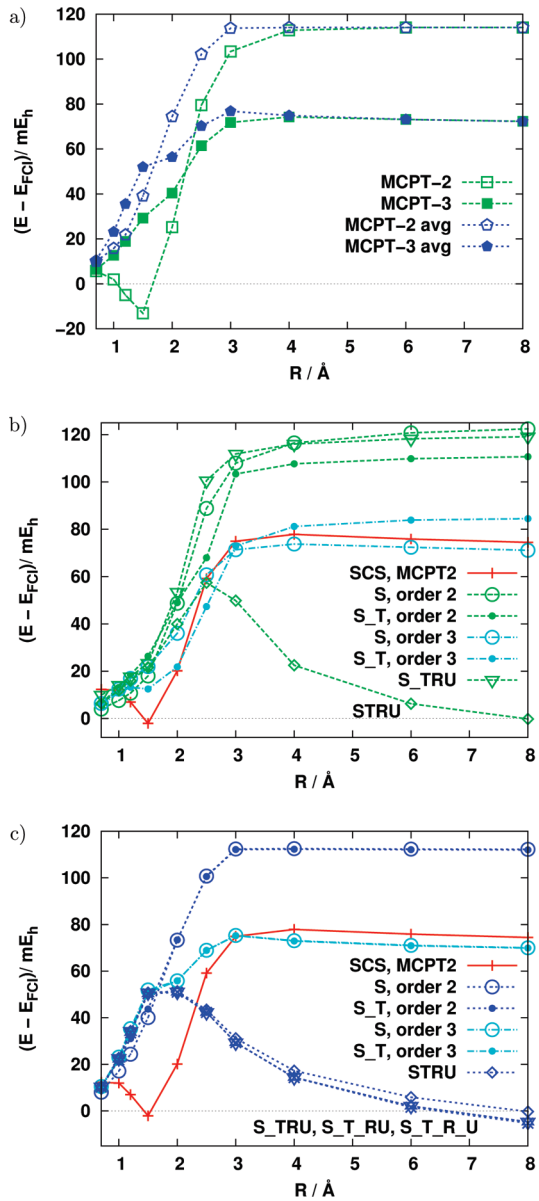

Figure 2. Selected scaling parameters for the BH molecule as a function of the internuclear distance. Basis set and reference function agree with Figure 1. (a) Parameters of excitation level based categorization, MCPT and MCPT-avg framework. (b) Parameters of s_z pattern based categorization, within the MRPT formulation.

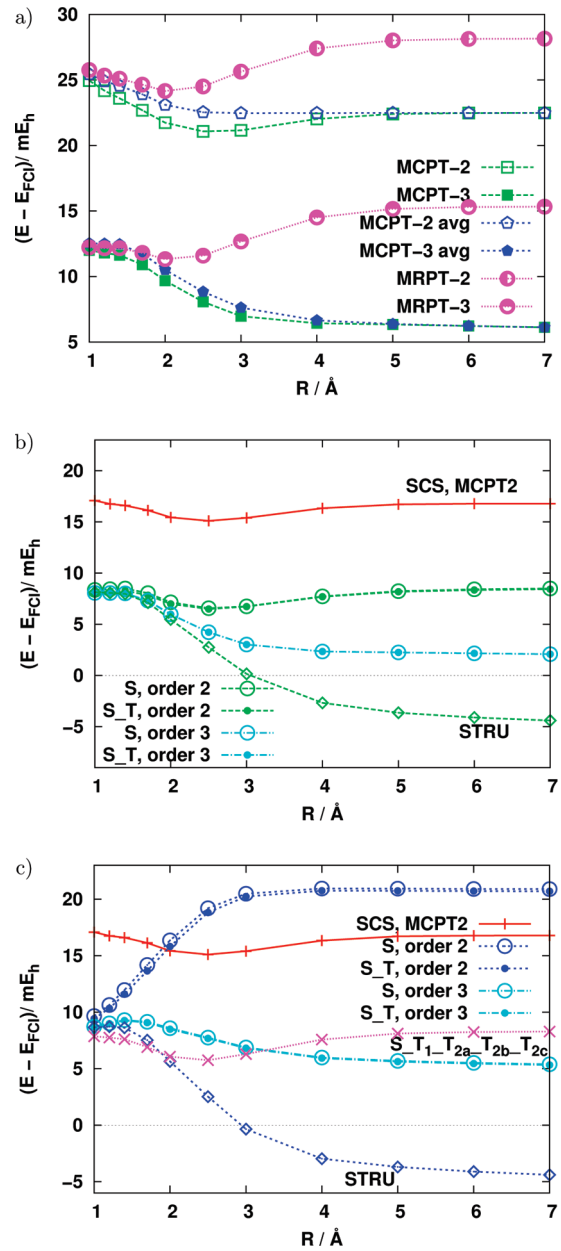

Figure 3. Energy difference curves (with respect to FCI) of the N_2 dissociation computed by MCPT methods, in valence double-ζ basis set. Reference function underlying the PT approximations is an APSG function, with two orbitals assigned to each bonding electron pair. Acronym SCS refers to Grimme's two parameter scaling with $p_S = 1.2$ and $p_T = 1/3$, within the pivot-dependent MCPT framework. (a)Unscaled second- and third-order results. (b) Scaled results in the MCPT framework. (c) Scaled results in the MCPT-avg framework.

range. Three parameters of the five occurring in s_z pattern based categorization show a variation similar to p_{STRU}; i.e., they are increased monotonically by about 0.2−0.3 from their value at around equilibrium. Curves of the remaining two parameters are more flat; they change roughly by 0.1.

B. N_2 Molecule. On the example of the triple bond dissociation of the nitrogen molecule, we study the effect of basis set and reference function.

We applied the cc-pVDZ basis set of Dunning.[52] The polarization function was excluded in a series of small basis set calculations while it was kept in a second set of calculations. Cores were frozen in every case.

1. APSG Reference. As an example for a small model space an antisymmetrized product of strongly orthogonal geminals (APSG) function[53] was computed, with one virtual function assigned to each dissociating bond. This is equivalent to a generalized valence bond treatment of the system. The expansion of the reference function in terms of determinants built with natural orbitals is composed of merely eight terms.

In Figure 3a. we compare errors of the second- and third-order energies computed by different perturbation theories. We observe tens of millihartrees difference between MCPT-avg and pivot-dependent MCPT, the latter typically giving smaller errors at around equilibrium geometry.

In Figure 3b, we present the energy errors of various scaling schemes for MCPT. For comparison, second-order MCPT scaled with the original parameters of Grimme (labeled SCS) is shown. All other curves were obtained by the application of eqs 15 and 16. Third orders computed by formula 17 are shown when nonzero. Figure 3b indicates that Grimme's two-parameter scaling successfully reduces the error along the whole potential

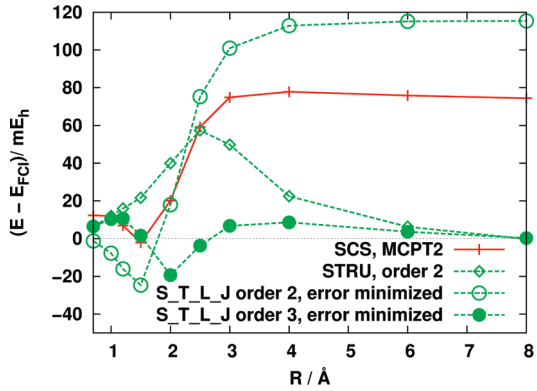

Figure 4. Energy difference curves (with respect to FCI) of the N_2 dissociation computed by pivot-dependent MCPT. Basis set and reference agree with Figure 3. Acronym SCS refers to Grimme's two parameter scaling with $p_S = 1.2$ and $p_T = 1/3$. Error minimized S_T_L_J operates with $p_S = 1.032$, $p_T = 1.147$, $p_R = 0.260$, and $p_U = 2.056$. See text for more details.

curve to become the quality of MCPT3. With the application of eq 15 it is not possible to achieve a considerable error reduction if adopting the excitation level based S or S_T scaling. On the other hand scaling STRU is very effective in the long distance range. This holds also for the scalings in the MCPT-avg which include group U in parameter determination. This is notable, since determinants classified exclusively as U in the pivot dependent theory belong to category S or T to a certain degree, when averaging is applied. Still, scaling of the full contribution of these determinants is needed to reduce the error efficiently, as demonstrated by Figure 3c.

Schemes S_T_R_U, S_T_RU are again not shown in Figure 3b. since they are unsuccessful. In the MCPT-avg framework, these are reliable but give results closely matching those of STRU. (For completeness, errors and nonparallelity errors of all scalings computed are collected in the Supporting Information.) This experience indicates that there is no point in using more than a single parameter within the excitation level based classification scheme.

It is notable that scaling has a minor effect in the short bond distance range in the MCPT-avg framework. This behavior is the direct consequence of $E^{(3)}$ being almost zero at short distance as seen in Figure 3a. Since Feenberg-type scaling is that unique partitioning which sets the third-order correction zero, it has hardly any role if the third order is negligible at start.

To examine the performance of constant scaling factors, with values fitted for the problem under investigation, we also performed numerical minimization of the function $\Sigma_\mu(E_\mu^{(2)}(\mathbf{p}) + E_\mu^{(3)}(\mathbf{p}) - E_{\mathrm{FCI}}^\mu)^2$, where μ runs over the geometry points sampled. We present the results in Figure 4 for excitation level based scaling in the pivot-dependent MCPT framework. Error minimization gave parameters $p_S = 1.032$, $p_T = 1.147$, $p_R = 0.260$ and $p_U = 2.056$ in the interval $(0, 3)$ for the nitrogen dissociation process studied here. As apparent in Figure 4, the second-order energy by these parameters is worse than either SCS MCPT2 or the best scaling based on eq 15, i.e., STRU, while the third-order energy is superior. The above parameter values however are not generally applicable; they give worse results than MCPT2 if applied for different basis set or different reference function. We also observe a relatively large variation in scaling parameters that minimize the errors if taking only the short or the long distance regime. Largest variation with geometry is shown by parameters p_S and p_U. A graphical example for this is given in Figure 4 of the Supporting Information.[7]

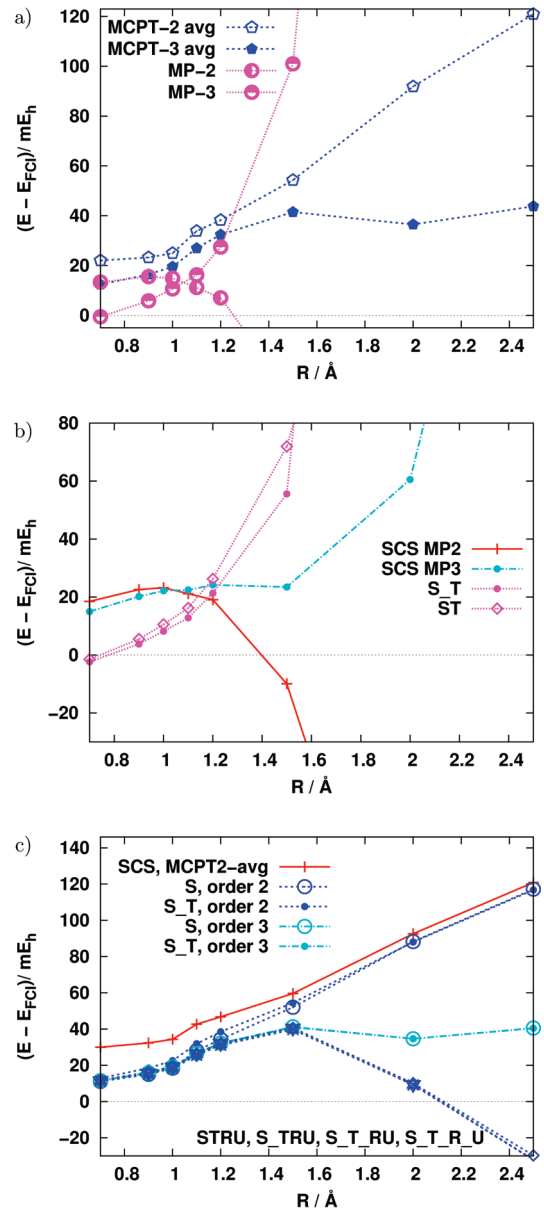

Figure 5. The same as Figure 3 in cc-pVDZ basis. (a) Unscaled second and third order results. (b) Scaled results in single reference MP theory. (c) Scaled results in the MCPT-avg framework.

Presenting the results obtained in polarized basis set, we focus on the comparison of scalings in single reference MP theory and averaged MCPT. Single reference MP becomes unreliable over 1.2 Å as indicated by Figure 5a. Errors of averaged MCPT are slightly worse than single reference MP at equilibrium geometry (around 1.1 Å), but remains more acceptable at stretched geometries.

Feenberg scaling may be efficient below the point where the third-order MP correction becomes zero. Even in this bond distance range neither the single parameter (ST) nor the two-parameter (S_T) schemes get considerably better than MP3. Interestingly, scaling by the parameters of Grimme increases the errors slightly around equilibrium but improves the parallelity within this region.

Scalings based on averaged MCPT on the same example show hardly noticeable deviation from the third order around equilibrium, as seen in Figure 5c. This is acceptable regarding the small absolute value of the third-order energy term. As stretching the bond, excitation level based S and S_T are similar and run

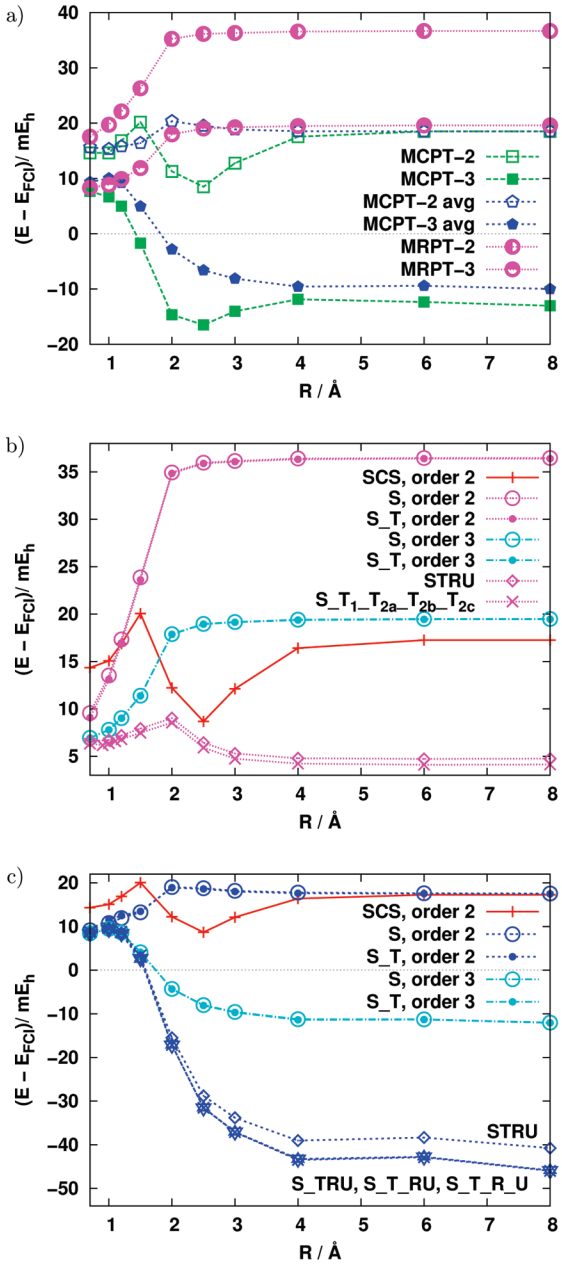

Figure 6. The same as Figure 3 with a CAS(6,6) reference function. (a) Unscaled second- and third-order results. Acronym MRPT refers to the theory originally proposed by Davidson.[42] (b) Scaled results in the MRPT framework. (c) Scaled results in the MCPT-avg framework.

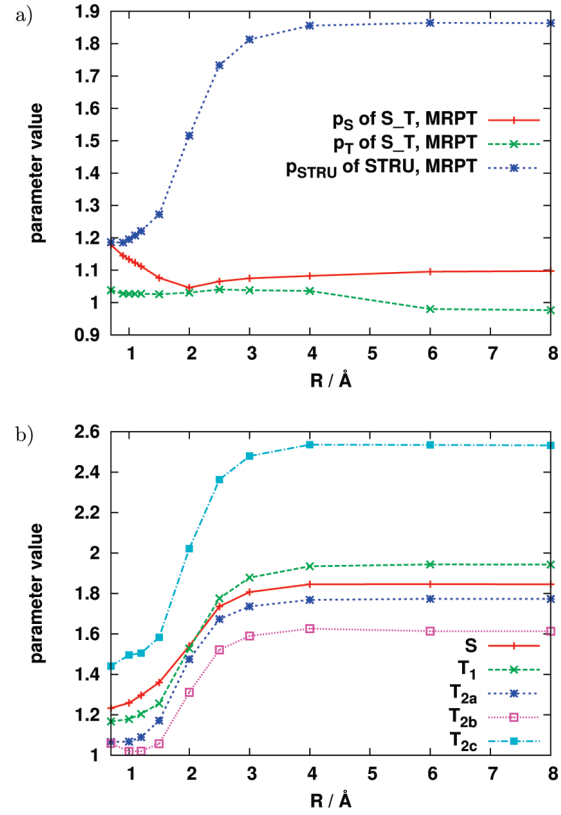

Figure 7. Selected scaling parameters for the N_2 molecule as a function of the internuclear distance. The MRPT framework is applied. Basis set and reference function agree with Figure 6. (a) Parameters of excitation level based categorization. (b) Parameters of s_z pattern based categorization.

rather close to the curve of the respective order. Schemes that involve a scaling of quadruples also tend to agree with each other and give a considerable overcorrection upon dissociation.

2. CAS(6,6) Reference. For the N_2 dissociation a complete active space (CAS) function was also computed as a reference, with six active electrons distributed over six active orbitals. This function is a better approximation than the previous APSG, its determinantal expansion including 56 determinants when spatial symmetry is exploited.

Second- and third-order results based on the CAS(6,6) reference are presented in Figure 6a. Zero orders are not shown in the figure; their errors vary between 73 and 94 mE_h. An important qualitative difference between MRPT and MCPT in this example is that the latter overshoots at the third order while MRPT3 remains an upper bound. This makes MRPT better fitted for Feenberg-scaling than MCPT (cf. Table 2). We again see that averaging in MCPT improves the parallelity at the price of getting the errors worse at some points.

Similar to previous examples we see the close agreement of excitation level based S and S_T scalings. They however do not bring improvement neither in MRPT nor in MCPT-avg. Grimme's constant parameter values are not successful in this example: MCPT2 remains practically unaffected. Determining all parameters from eq 15 brings a dual benefit in MRPT: errors are diminished and parallelity is better. There is hardly any difference between the results of excitation level based and s_z pattern based categorization. The slight difference visible in Figure 6b is due to the separate treatment of groups T_{2b} and T_{2c}, i.e., by the merger of these categories the two curves are overlapped. As before, no significant difference is revealed among scaling variants STRU, S_TRU, S_T_RU, and S_T_R_U, the latter three are therefore omitted from Figure 6b. (Errors and nonparallelity errors of all scalings are given in the Supporting Information for this example also.)

In MCPT-avg Feenberg scaling decreases the error of the third order and therefore necessarily means a deterioration.

Parameters determined from condition 15 show a picture somewhat different from the case of the BH system. Inspecting excitation level based grouping in Figure 7a, we see that both parameters p_S and p_T vary in a rather narrow range. It is notable that the value of p_T is larger than 1 until about 5 Å bond distance. The value of the single parameter scaling scheme shows drastic increase upon bond length elongation tending from about 1.2 to 2. A similar behavior is shown by all parameters in the s_z pattern based scaling in Figure 7b. It is the range of variation that is either larger—T_{2c}—or smaller—T_{2b}. This experience can be taken as a contraindication to the intent of extracting

TABLE 3: Parameters of the Function (19) Fitted for the Data Points of p_S Obtained by Feenberg Scaling for Various Multireference Pt Methods, Cathegorization Schemes, and Molecules

method	g_0	g_1	g_2	λ	η	rms[a]
S_T, MCPT[b]	0.899073	0.0178359	2.73441	0.608619	1.63006	1.98813×10^{-3}
S_T, MCPT-avg[b]	0.821558	–0.0831031	2.87162	0.570482	1.55156	1.24969×10^{-3}
$S\_T_{1}\_T_{2a}\_T_{2b}\_T_{2c}$,[b] MRPT	0.96955	0.0863499	1.77392	0.63079	1.81366	1.90874×10^{-3}
$S\_T_{1}\_T_{2a}\_T_{2b}\_T_{2c}$,[c] MRPT	0.955365	0.197598	2.60358	0.652064	1.63267	6.38666×10^{-3}
S_T, MCPT-avg[c]	0.793131	0.0971636	2.63428	0.569324	1.82812	1.4129×10^{-3}

[a] Root mean squared error. [b] BH molecule, 6-311G** basis, CAS(2,2) reference. [c] N_2 molecule, VDZ basis, CAS(6,6) reference.

geometry-independent scaling parameters stemming from eq 15 in the s_z pattern based categorization. Regarding excitation level based categorization, parameters p_S and p_T give more hope for transferability, but their optimized values are not really effective without including quadruples. Inclusion of group U however gets the parameters to vary within a broader range.

C. Parameter Dependence on Geometry and Size. Various scaling shemes applied to the description of covalent bond dissociation show a diverse picture and an expressed dependence on system, geometry, and method. These results make the impression that the aim of finding a universal set of scaling parameters is unrealistic. This conclusion complements a recent study by Varandas, where geometry dependence of the scaling factor was investigated.[33] In particular, Varandas supposed the following form of the p_S parameter of single-reference SOS-MP2, as a function of geometry(R), number of electrons(n_s), and basis set cardinal number (X)[33]

$$p_S = \left(g_0(X) + g_1 \tanh\left(g_2 \frac{R - \eta R_e}{\eta R_e}\right)\right)\left(1 + \lambda \frac{n_S - 2}{n_S - 1}\right) \tag{19}$$

where R_e characterizes the equilibrium geometry. He has shown that by using this functional form of p_S, applicability of SOS-MP2 can be extended to near equilibrium geometries without losing the performance observed at the equilibrium. In parallel with his findings, we observe that most p_S parameter curves plotted in Figures 2 and 7 exhibit a sigmoid shape. If fitting the above function to the p_S parameters of some well-performing multireference scaling examples, we indeed obtain small root-mean-square errors (rms), as shown in Table 3. Comparing the values of the parameters in Table 3 with those reported by Varandas for the H_2 molecule (Table I of ref 33), g_2 and η are in acceptable agreement (the average values of g_2 are within 12%, those of η are within 23%). Parameters g_0 and g_1 of the shape term show larger rms, g_0 being systematically smaller in our case. The g_1 parameters collected in Table 3 vary strongly with method and with system. It is out of the scope of the present study to further explore the possibility of geometry-dependent scaling in multireference PT. The reason for showing this unrepresentative set of examples is merely to demonstrate that a systematization of the dissapointingly diverse scaling results may be possible to do if, e.g., finding a common functional form of the parameters.

D. Singlet−Triplet Splitting of the CH_2 Molecule. In Table 4 we collect the results obtained for the CH_2 molecule in the valence double-ζ basis of Dunning and Hay,[54] with one set of d functions applied for the carbon atom. For both states a CAS(6,6) function gives the zero-order approximation. The reference function is dominated by a single determinant (with an expansion coefficient 0.969) while the triplet state is built of a pair of open shell determinants predominantly. For this reason no averaging is applied for the singlet state. The two dominant determinants were averaged in the triplet case.

TABLE 4: Energy Errors ($E - E_{FCI}$) and Singlet−Triplet Splittings ($E_{singlet} - E_{triplet}$) of the CH_2 Molecule Computed by Different Scaled MCPT and MRPT in the Polarized VDZ Basis of Dunning and Hay[54] [c]

	MCPT				MRPT		
	triplet error, mE_h	singlet error, mE_h	S-T splitting, mE_h		triplet error, mE_h	singlet error, mE_h	S-T splitting, smE_h
MCPT0	69.28	68.35	21.76	MRPT0	69.28	68.35	21.76
MCPT2	11.68	11.63	22.64	MRPT2	19.26	15.37	18.80
MCPT3	1.87	3.15	23.96	MRPT3	7.41	5.63	20.90
SCS MCPT2	11.49	10.53	21.72				
S order 2	10.81	3.80	15.68	5 parameters[b]	3.65	3.39	22.42
S order 3	1.70	2.02	23.00	4 parameters[b]	3.65	3.39	22.42
S_T order 2	10.60	2.96	15.05	1 parameter[b]	3.74	3.43	22.38
S_T order 3	1.67	1.95	22.96				
S_T_RU[a]	−0.15	4.19	27.03				
S_T_R_U[a]	−0.15	4.08	26.92				
S_TRU	−0.15	1.64	24.47				
STRU	−0.14	1.65	24.48				
FCI			22.69				22.69

[a] Damping with $\omega = 5 \times 10^{-4}$ E_h for the singlet state. [b] 5 parameters: $S\_T_{1}\_T_{2a}\_T_{2b}\_T_{2c}$, 4 parameters: $S\_T_{1}\_T_{2a}\_T_{2b}$ T_{2c}, 1 parameter: ST_1 T_{2a} T_{2b} T_{2c}. [c] Zero-order reference function is obtained by a CAS(6,6) calculation. Averaging in MCPT is performed over the two leading determinants in the triplet case. No averaging is applied for the singlet state. Acronym SCS refers to Grimme's $p_S = 1.2$ and $p_T = 1/3$. Left panel of the table includes excitation level based categorization, in the right panel results by s_z pattern based categorization are collected. Geometry is $r_{CH} = 1.0780$ Å, $\angle(HCH) = 132.9°$ for the triplet state and $r_{CH} = 1.1086$ Å, $\angle(HCH) = 102.0°$ for the singlet state.

Unscaled PT results are better by MCPT than by MRPT in this example. This is understandable, considering that the contribution of CAS space determinants to the dynamical correlation is included in MCPT but not in MRPT at the orders computed. Regarding excitation level based scalings, S and S_T schemes prove to be similar again to each other and decrease the error only lightly, with the exception of the second order for the triplet state. Determination of all scaling parameters via condition eq 15 slightly overshoots for the triplet in the excitation level based scheme. This effect is not seen in the s_z pattern based categorization, where the errors are diminished by scaling also showing a balanced treatment of the two states. This has the effect that the S-T splitting is well captured by any of the s_z pattern-based scalings shown.

Splitting by scaling based on the excitation level categorization is also correct considering the order of magnitude. Slight imbalance of the errors committed in the singlet and the triplet case lead to relatively larger deviation from the FCI value for S and S_T at the second order. The second order of MCPT gives a notably good splitting, which is worse by MCPT3, even if the error of individual states is diminished by an order of magnitude. Within these circumstances, a scaling may be declared successful if the splitting remains in the range of

TABLE 5: Energy Errors ($\varepsilon = E - E_{FCI}$), Forward ($\Delta E_F$), and Reverse ($\Delta E_R$) Barriers of the HCN → HNC Reaction in mE_h Computed by Single Reference MP and MCPT, Unscaled and Scaled[b]

	ε_{HCN}	ε_{TS}	ε_{HNC}	ΔE_F	ΔE_R
HF	221.48	221.40	207.29	98.54	85.17
MP2	7.48	16.33	13.64	107.46	73.76
MP3	18.08	22.46	17.47	103.00	76.06
SCS MP2	12.78	21.86	17.51	107.70	75.42
SCS MP3	15.43	23.40	18.46	106.58	76.00
S_T MP2	13.38	19.06	14.19	104.30	75.94
ST MP2	17.58	22.28	17.39	103.32	75.96
CAS(10,9)	62.54	67.17	60.82	103.25	77.42
MCPT 2[a]	10.74	12.23	14.98	100.11	68.32
MCPT 3[a]	3.58	3.83	6.92	98.88	67.98
S_T MCPT 2[a]	8.11	8.46	10.44	98.97	69.10
S_T MCPT 3[a]	3.08	3.11	6.23	98.65	67.95
S MCPT 2[a]	6.98	7.26	9.91	98.90	68.41
S MCPT 3[a]	3.25	3.31	6.27	98.68	68.11
S_TRU MCPT 2[a]	1.69	1.52	4.55	98.45	68.05
STRU MCPT 2[a]	2.43	2.32	5.20	98.51	68.19
FCI[a]				98.62	71.07

[a] Frozen core approximation. [b] Dunning's cc-VDZ basis is applied. Zero-order reference function underlying MCPT is obtained by a CAS(10,9) calculation. Geometries are taken from the database of Truhlar et al.[55] Scaled results are computed with the excitation level based categorization.

MCPT2 and MCPT3. In this respect schemes S_TRU and STRU as well as third order of S and S_T are worth mentioning.

E. Barrier of HCN to HNC Isomerization. In Table 5 we collect barrier heights of the forward and reverse reaction HCN → HNC as well energy errors for the reactant, product and the transition state (TS). The calculation was performed in double-ζ basis (cc-pVDZ, polarization function omitted).[52] A full valence CAS reference function serves as zero order, with 10 active electrons distributed over nine active orbitals. Geometries were taken from the database of Truhlar et al.[55]

For all the three species there is a dominant, closed shell determinant in the CI expansion of the wave function. For this reason single reference MP results and their scaled variants are also indicated. Examining errors in the single-reference framework, we see that the third order is worse than MP2. This explains why MP2 scaled by parameters determined from eq 15 give poorer approximation, than MP2 itself (cf. Table 2). Apparently SCS MP2 and SCS MP3 (computed according to ref 56) also exhibit larger errors than MP2. Barriers show a more scattered picture. The forward barrier is captured exceedingly well by Hartree−Fock, while there is a ca. 10 mE_h error in the MP2 value, which is approximately halved by MP3. At difference with this, the HF value of the reverse barrier is nicely improved by MP2 and worsened by MP3. Barriers of scaled methods vary between MP2 and MP3 for the forward barrier while they remain close to the MP3 value for the reverse process.

If inspecting the numbers obtained in the multiconfiguration framework, we observe that MCPT2 is comparable in quality with MP2 despite the much improved reference function. At the same time, MCPT3 represents a considerable improvement for all three species, at difference with the MP expansion. As a consequence, parameters determined from eq 15 nicely diminish the errors. Second orders of excitation level based S and S_T schemes are somewhat better than MCPT2, and third orders are comparable to MCPT3. Barriers show improvement (or only slight deterioration) in the order of CAS(10,9), MCPT2, and MCPT3. Scaled variants give barriers rather close to the MCPT3 value.

IV. Summary

We have explored the applicability of Feenberg-type scaling in PT built upon a multiconfiguration reference function. Various strategies were tested for assigning excited determinants to categories scaled by a common parameter, the guiding rule being either the level and/or the spin pattern of the excitation. The number of parameters was varied between 1 and 5.

A stationary condition for the total energy calculated up to order 3 was applied to determine the value of the scaling parameters. In a few attempts we also deduced parameters in a statistical manner, based on a comparison with the exact solution. These latter parameters however failed to be system independent.

We have diagnosed the failure of Feenberg's condition for giving reliable parameters in cases where the third order overshoots.

In successful scaling examples scaling just S is practically equivalent to scaling S and T. We have found that determining all parameters from Feenberg's condition is superior to leaving some parameters unscaled. Distinction of categories S and T among double excitations in the excitation level based grouping is not uniformly fruitful along a potential curve of covalent bond breaking. Neither could we find a spectacular effect of using five parameters in s_z pattern based categorization as compared to single parameter scaling.

We saw a considerable increase in the value of the parameters when stretching covalent bonds. This mainly affects category U in the excitation level based grouping, while all parameters are affected in the s_z pattern based categorization. Parameters of excitation level based categories S and T vary in a narrower range. They however seem to be ineffective in themselves in achieving a notable error reduction.

Though the variation of scaling parameters may be strong with geometry and system, the present results are in agreement with a previous observation of a hyperbolic tangent type geometry dependence of parameter p_S. On the basis of this, further studies for determining geometry-dependent parameters may be warranted.

Acknowledgment. We thank professor P. Surján (Budapest) for inspiring discussions and for reading the manuscript. This work has been supported by the Hungarian National Research Fund (OTKA), Grant Number K-81590. The Project was also supported by the European Union and cofinanced by the European Social Fund (Grant Agreement No. TAMOP 4.2.1./B-09/1/KMR-2010-0003). The APSG (GVB) reference functions were generated by the Budapest version of the MUNGAUSS program package.[57] The CAS functions with more than two electrons in the active space were computed by the GAMESS program suite.[58]

Supporting Information Available: Figures and tables of energy errors. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(2) Pulay, P.; Saebø, S. *Theor. Chim. Acta* **1986**, *69*, 357.
(3) Saebø, S.; Pulay, P. *J. Chem. Phys.* **1987**, *86*, 914.
(4) Maslen, P. E.; Head-Gordon, M. *Chem. Phys. Lett.* **1998**, *283*, 102–108.
(5) Schütz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691–5705.
(6) Nakao, Y.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 6375–6380.
(7) Hetzer, Georg; Schütz, Martin; Stoll, Hermann; Werner., Hans-Joachim *J. Chem. Phys.* **2000**, *113*, 9443.

(8) Koch, H.; Sánchez de Merás, A.; Bondo Pedersen, T. *J. Chem. Phys.* **2003**, *118*, 9481.
(9) Häser, M.; Almlöf, J. *J. Chem. Phys.* **1992**, *96*, 489–494.
(10) Ayala, P. Y.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 3660–3671.
(11) Surján, P. R. *Chem. Phys. Lett.* **2005**, *406*, 318–320.
(12) Kobayashi, M.; Nakai, H. *Chem. Phys. Lett.* **2006**, *420*, 250–255.
(13) Kobayashi, M.; Akama, T.; Nakai, H. *J. Chem. Phys.* **2006**, *125*, 204106.
(14) Doser, B.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2009**, *130*, 064107.
(15) Schweizer, S.; Doser, B.; Ochsenfeld, C. *J. Chem. Phys.* **2009**, *128*, 154101.
(16) Weijo, V.; Manninen, P.; Jørgensen, P.; Christiansesn, O.; Olsen, J. *J. Chem. Phys.* **2007**, *127*, 074106.
(17) He, Z.; Cremer, D. *Int. J. Quantum Chem.* **1996**, *59*, 71.
(18) Goodson, D. Z. *J. Chem. Phys.* **2000**, *112*, 4901.
(19) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.
(20) Goumans, T. P. M.; Ehlers, A. W.; Lammertsma, K.; Würthwein, E.-U.; Grimme, S. *Chem.—Eur. J.* **2004**, *10*, 6468.
(21) Grimme., S. *J. Phys. Chem.* **2005**, *A 109*, 3067.
(22) Grimme, S.; Diedrich, C.; Korth, M. *Angew. Chem., Int. Ed.* **2006**, *45*, 625.
(23) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793.
(24) Aquilante, F.; Bondo Pedersen, T. *Chem. Phys. Lett.* **2007**, *449*, 354.
(25) Hill, J. G.; Platts, J. A. *J. Chem. Theory Comput.* **2007**, *3*, 80–85.
(26) Distasio, R. A., Jr.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073–1083.
(27) King, R. A. *Mol. Phys.* **2009**, *107*, 789–795.
(28) Grimme, S.; Izgorodina, E. I. *Chem. Phys.* **2004**, *305*, 223.
(29) Casanova, D.; Rhee, Y. M.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 164106.
(30) Hellweg, A.; Grün, S. A.; Hättig, C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4119–4127.
(31) Fink, R. *J. Chem. Phys.* **2010**, *133*, 174113.
(32) Takatani, T.; Hohenstein, E. G.; Sherill, C. D. *J. Chem. Phys.* **2008**, *128*, 124111.
(33) Varandas, A. J. C. *J. Chem. Phys.* **2010**, *133*, 064104.
(34) Robinson, D.; McDouall, J. W. *J. Phys. Chem. A* **2007**, *111*, 9815.
(35) Rolik, Z.; Szabados, Á.; Surján, P. R. *J. Chem. Phys.* **2003**, *119*, 1922.
(36) Szabados, A.; Rolik, Z.; Tóth, G.; Surján, P. R. *J. Chem. Phys.* **2005**, *122*, 114104.
(37) Goldhammer, P.; Feenberg, E. *Phys. Rev.* **1955**, *101*, 1233.
(38) Feenberg, E. *Phys. Rev.* **1956**, *103*, 1116.
(39) Szabados, Á. *J. Chem. Phys.* **2006**, *125*, 214105.
(40) Szabados, Á.; Surjan., P. R. *Progress in Theoretical Chemistry and Physics*; Springer: Dordrecht, 2009; pp 257−269.
(41) Zaitevskii, A.; Malrieu, J.-P. *Chem. Phys. Lett.* **1996**, *250*, 366.
(42) Davidson, E. R.; Bender, C. F. *Chem. Phys. Lett.* **1978**, *59*, 369–374.
(43) Surján, P. R.; Rolik, Z.; Szabados, Á.; Kohalmi, D. *Ann. Phys. (Leipzig)* **2004**, *13*, 223–231.
(44) Nakai, H.; Kobayashi, M.; Szabados, Á.; Surjan, P. R. *J. Chem. Theory Comput.* **2010**, *6*, 2024–2033.
(45) Szabados, Á.; Surjan, P. R. *Chem. Phys. Lett.* **1999**, *308*, 303.
(46) Surján, P. R.; Szabados, Á. *J. Chem. Phys.* **2000**, *112*, 4438–4446.
(47) Hose, G.; Kaldor, U. *J. Phys. B* **1979**, *12*, 3827.
(48) Mitrushenkov, A. O. *J. Chem. Phys.* **1996**, *105*, 10487.
(49) Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.
(50) Hirao, K. *Int. J. Quantum Chem.* **1992**, *S26*, 517.
(51) Taube, A. G.; Bartlett, R. J. *J. Chem. Phys.* **2009**, *130*, 144112.
(52) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
(53) Surjan, P. R. *Top. Curr. Chem.* **1999**, *203*, 63–88.
(54) Dunning, T. H., Jr.; Hay, P. J. In *Modern Theoretical Chemistry, The Methods of Electronic Structure Theory*; Plenum: New York, 1977; pp 1−27.
(55) Zhao, Y.; Gonzáles-García, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.
(56) Grimme, S. *J. Comput. Chem.* **2003**, *24*, 1529.
(57) Surjan, P. R. *Program BP-MUNGAUSS*; Department of Theoretical Chemistry, Eötvös University, Budapest, 2002.
(58) Gordon, M. S.; Schmidt, M. W. In *Theory and Applications of Computational Chemistry, the first forty years*; Elsevier: Amsterdam, 2005; pp 1167−1189.

JP108575A