# Sensitivity analysis of state-specific multireference perturbation theory

Ágnes Szabados[a)]
*Laboratory of Theoretical Chemistry, Loránd Eötvös University, H-1518 Budapest, POB 32, Hungary*

State-specific multireference perturbation theory (SS-MRPT) developed by Mukherjee *et al.* [Int. J. Mol. Sci. **3**, 733 (2002)] is examined focusing on the dependence of the perturbed energy on the initial model space coefficients. It has been observed earlier, that non-physical kinks may appear on the potential energy surface obtained by SS-MRPT while related coupled-cluster methods may face convergence difficulties. Though exclusion or damping of the division by small coefficients may alleviate the problem, it is demonstrated here that the effect does not originate in an ill-defined division. It is shown that non-negligible model space coefficients may also be linked with the problem. Sensitivity analysis is suggested as a tool for detecting the coefficient responsible. By monitoring the singular values of sensitivity matrices, orders of magnitude increase is found in the largest value, in the vicinity of the problematic geometry point on the potential energy surface. The drastic increase of coefficient sensitivities is found to be linked with a degeneracy of the target root of the effective Hamiltonian. The nature of the one-electron orbitals has a profound influence on the picture: a rotation among active orbitals may screen or worsen the effect. © *2011 American Institute of Physics.* [doi:10.1063/1.3585604]

## I. INTRODUCTION

Molecular systems that are intriguing for chemists are often characterized by a wavefunction which can not be well approximated by a single Hartree-Fock determinant. Single bond dissociation, as the simplest example, requires two determinants when working with restricted orbitals, to get a qualitatively correct characterization. It has been well established to describe the electronic structure of such systems in two steps, i.e., by solving a linear variation problem in a preferably small determinantal space collecting the most important configurations, and correct this wavefunction in a second step by a properly formulated, multireference function based correlation method.

Among various perturbation approaches that rely on a multireference zero-order function (MRPT) the state-specific (SS) MRPT developed by Mukherjee and co-workers[1–4] is one which possesses the desirable features of size-extensivity and intruder free character. Similarly to most MRPT's, SS-MRPT does not conserve orbital invariance of the underlying complete active space (CAS) wavefunction neither in its original form, nor in a reformulation suggested recently.[5] Nevertheless, SS-MRPT is an appealing and competitive method[6,7] and has been successfully applied for various problems.[8,9]

In spite of the manifest intruder-free character of the theory, unphysical kinks have been recently observed on the order of one to ten $mE_h$ when computing potential energy surfaces by SS-MRPT. Similar experience has been reported in the parent coupled-cluster (CC) theory[10] as well as in a related multireference CC theory developed by Hanrath.[11] The unexpected behavior has been attributed to small coefficients in the CAS wavefunction, which appear in the denominator of the amplitude equations. To circumvent the problem a damping procedure has been suggested.[10]

In the present study we apply a sensitivity analysis of the SS-MRPT equations with the aim of unfolding the problem. We show that sensitivity analysis may be used to point to the problematic CAS space coefficients. The investigation reveals that the coefficients responsible for the effect are not necessarily the smallest in absolute value. It is found that the appearance of the kinks depends severely on the nature of the one particle basis, showing orders of magnitude difference in the singular values of coefficient sensitivity matrices, for different orbital sets. Knowing that SS-MRPT energies are orbital dependent, we argue that orbital sets giving small sensitivities are preferable to others, since these produce small variation of SS-MRPT results upon infinitesimal orbital rotations.

In what follows, the formulation of SS-MRPT is briefly recollected, followed by the presentation of sensitivity analysis and a numerical illustration section demonstrating the utility of the suggested analysis.

## II. THEORY

### A. SS-MRPT

Derived from the parent CC theory,[12] SS-MRPT operates with the Jeziorski-Monkhorst parametrization of the wavefunction[13]

$$\Psi = \sum_\mu e^{T_\mu} |\phi_\mu\rangle \, c_\mu, \tag{1}$$

and assumes the existence of a CAS wavefunction written as

$$\Phi = \sum_\mu |\phi_\mu\rangle \, c_\mu^{(0)}, \tag{2}$$

[a)]Electronic mail: szabados@chem.elte.hu.

**134**, 174113-1

the associated CAS energy being $E_{CAS}$. Introducing $P$ for the projector of the CAS space

$$P = \sum_\mu |\phi_\mu\rangle\langle\phi_\mu|, \qquad (3)$$

and $Q$ for the orthogonal complement

$$Q = 1 - P = \sum_l |\chi_l\rangle\langle\chi_l|, \qquad (4)$$

upon substitution of the Ansatz Eq. (1) into the Schrödinger equation one may arrive at the following pair of equations[12]

$$\sum_\mu \widetilde{H}_{\mu\nu} c_\nu = E c_\mu, \qquad (5)$$

defining coefficients $c_\mu$ and

$$\sum_\mu \left( e^{T_\mu} Q \overline{H}_\mu |\phi_\mu\rangle c_\mu + \sum_\nu e^{T_\mu} Q e^{-T_\mu} e^{T_\nu} |\phi_\mu\rangle \widetilde{H}_{\mu\nu} c_\nu \right) = 0, \qquad (6)$$

determining amplitudes in excitation operators $T_\mu$. In the above $\overline{H}_\mu = e^{-T_\mu} H e^{T_\mu}$ is the similarity transformed Hamiltonian and $\widetilde{H}_{\mu\nu} = \langle\phi_\mu|\overline{H}_\nu|\phi_\nu\rangle$ denotes its matrix elements taken with CAS space wavefunctions.

Obviously, projection of Eq. (6) by $Q$-space functions does not provide sufficient number of equations to determine the amplitudes in all operators $T_\mu$. This well known redundancy problem is surpassed in SS-MRPT by requiring that Eq. (6) holds separately for each index $\mu$. Omitting the sum in Eq. (6) and projecting by function $\langle\chi_l|$ one obtains the amplitude equations

$$\langle\chi_l|\overline{H}_\mu|\phi_\mu\rangle c_\mu + \sum_\nu \langle\chi_l|e^{-T_\mu}e^{T_\nu}|\phi_\mu\rangle \widetilde{H}_{\mu\nu} c_\nu = 0, \qquad (7)$$

that should hold for each $\mu$ and $l$. Equations (7) and (5) are the basic equations of SS-MRCC that—with the appropriate truncation introduced in operators $T_\mu$—are iterated till self-consistency.

The Rayleigh-Schrödinger variant of SS-MRPT proposed by Mukherjee et al.[4] stems by a linearization of the amplitude equations giving

$$\langle\chi_l|H|\phi_\mu\rangle c_\mu + \langle\chi_l|[H, T_\mu]|\phi_\mu\rangle c_\mu$$
$$+ \sum_\nu \langle\chi_l|T_\nu - T_\mu|\phi_\mu\rangle H_{\mu\nu} c_\nu = 0. \qquad (8)$$

Introducing a partitioning of the total Hamiltonian as

$$H = H^{(0)} + V, \qquad (9)$$

with the zero-order operator diagonal on the basis of functions $\phi_\mu$ and $\chi_l$

$$H^{(0)} = \sum_\mu E_\mu^{(0)} |\phi_\mu\rangle\langle\phi_\mu| + \sum_l E_l^{(0)} |\chi_l\rangle\langle\chi_l|, \qquad (10)$$

the equation determining the first order amplitudes $t_{l\mu}^{(1)\nu} = \langle\chi_l|T_\nu^{(1)}|\phi_\mu\rangle$ takes the form

$$\sum_\nu \left[ H_{\mu\nu} + \left( E_l^{(0)} - E_\mu^{(0)} - E_{CAS} \right) \delta_{\mu\nu} \right] c_\nu^{(0)} t_{l\mu}^{(1)\nu}$$
$$= -H_{l\mu} c_\mu^{(0)}, \qquad (11)$$

with $H_{\mu\nu} = \langle\phi_\mu|H|\phi_\nu\rangle$ and $H_{l\mu} = \langle\chi_l|H|\phi_\mu\rangle$. Zero-order energies $E_\mu^{(0)}$ and $E_l^{(0)}$ are subject to choice. See Sec. III for the specification of these quantities. Substituting the first order amplitudes into Eq. (5), the energy correct up to order two is obtained as an eigenvalue

$$\sum_\mu H_{\nu\mu}^{[2]} c_\mu = E^{[2]} c_\nu, \qquad (12)$$

of the non-symmetric effective Hamiltonian

$$H_{\nu\mu}^{[2]} = \langle\phi_\nu|H + \left(HT_\mu^{(1)}\right)_c|\phi_\mu\rangle. \qquad (13)$$

We remark here that the derivation of the above equations is not based on a rigorous order-by-order analysis of the SS-MRCC equations. Note, for example, that the CAS wavefunction Eq. (2) is not an eigenfunction of the zero-order Hamiltonian Eq. (10). A strict perturbation approach based on the Taylor-expansion of the wavefunction and the energy has been given recently[5] and it results in working equations different from the above. An important point of dissimilarity is the expression of the second order correction: in Eq. (12) it appears as an eigenvalue of $H_{\nu\mu}^{[2]}$ while it is simply an expectation value of $H_{\nu\mu}^{[2]}$ in Ref. 5.

Introducing the composite index $I$ for excitation $\mu \to l$, working equation [Eq. (11)] of SS-MRPT can be solved separately for each $I$ in the spectral resolution form

$$t_I^{(1)\nu} = -\frac{1}{c_\nu^{(0)}} \sum_\mu{}' A_{\nu\mu}^{-1}(I) c_\mu^{(0)} H_{I\mu,\mu}, \qquad (14)$$

with

$$A_{\mu\nu}(I) = H_{\mu\nu} + \left( E_l^{(0)} - E_\mu^{(0)} - E_{CAS} \right)\delta_{\mu\nu}, \qquad (15)$$

and $H_{I\mu,\mu} = H_{l\mu}$. Once $t_I^{(1)\nu}$-s are computed for a given $I$, their contribution is accumulated into $H_{\nu\mu}^{[2]}$ by Eq. (13) and subsequently they are dropped. When all excitations, $I$, have been taken into account, $H_{\nu\mu}^{[2]}$ is diagonalized to obtain the second order SS-MRPT energy according to Eq. (12). It is to be noted, that the dimension of the linear system of equations [Eq. (14)] is equal or less than the dimension of the model space, depending on the creation/annihilation indices which carry out the excitation $\mu \to l$. Whenever these indices are inconsistent with the occupied/virtual classification of a model space determinant $\phi_\nu$, the corresponding amplitude $t_I^{(1)\nu}$ is zero, and $\nu$ is omitted when constructing $A_{\mu\nu}(I)$ and evaluating the sum on the right hand side of Eq. (14). A prime on the sum in Eq. (14) indicates this restriction.

Although the amplitude equation in the form of Eq. (11) shows similarity with the SS-MRCEPA(0) equations,[14] they markedly differ in the treatment of the Hamiltonian in the space of the virtual functions, $\chi_l$. In particular, term $H_{lm}$ in matrix **A** of the SS-MRCEPA(0) equations is approximated as $\delta_{lm} E_l^{(0)}$ in Eq. (15), facilitating the decoupling of $t_I^{(1)\mu}$ and $t_J^{(1)\mu}$ for $I \neq J$.

As demonstrated in Sec. III, the second order SS-MRPT energy may show sudden, non-physical kinks on a potential energy surface. Similar effects are observable in Figs. 12 and 14 of a recent study by Mahapatra et al.[15] reporting applications of SS-MRPT. Divergence of PT denominators—intruder

in matrix $\mathbf{A}(I)$, in other terms—could be an obvious reason. This however cannot be claimed for the effect, which may be reasoned by a rearrangement of the diagonals of matrix $\mathbf{A}(I)$ of Eq. (15) in the form[4]

$$A_{\mu\mu}(I) = H_{\mu\mu} - E_{\mu}^{(0)} + E_{l}^{(0)} - E_{\mathrm{CAS}}. \qquad (16)$$

The main contribution to $A_{\mu\mu}(I)$ comes from the last two terms on the right hand side in the form of an excitation energy. Since this is a difference of an uncorrelated excited state energy and a correlated ground state energy, $A_{\mu\mu}(I)$ is never expected to vanish if treating the ground state.

Another plausible explanation for the kinks would be the close to zero value of CAS coefficient $c_{\nu}^{(0)}$ which would render $t_{I}^{(1)\nu}$ of Eq. (14) unphysically large in absolute value and non-computable in the limit $c_{\nu}^{(0)} \to 0$. For this reason, a damping of $1/c_{\nu}^{(0)}$ attributed to Tikhonov[16] has been applied in the parent SS-MRCC[10] in the form

$$\frac{c_{\nu}^{(0)}}{\left(c_{\nu}^{(0)}\right)^2 + \omega^2}, \qquad (17)$$

while Hanrath suggested to drop every $c_{\nu}^{(0)}$ falling below a threshold in magnitude and zero the corresponding $t_{I}^{\nu}$-s.[11]

When applied to SS-MRPT, both of the above techniques may be used to eliminate the kinks. However, the magnitude of coefficients which need to be eliminated or damped may be rather large. To complete the picture, Tikhonov damping of $\mathbf{A}^{-1}(I)$ in the form

$$\sum_{\mu} \frac{\alpha_{\mu}}{\alpha_{\mu}^2 + \omega^2} \, \mathbf{u}_{\mu} \otimes \mathbf{u}_{\mu} \qquad (18)$$

has also been tested, while keeping $1/c_{\nu}^{(0)}$ undamped. Here $\alpha_{\mu}$ are the eigenvalues and $\mathbf{u}_{\mu}$ are the eigenvectors of matrix $\mathbf{A}(I)$. Remarkably, in spite of matrix $\mathbf{A}(I)$ being regular, this procedure may also eliminate the kinks with a sufficiently large damping parameter. Both of the above experiences indicate that the problem is not purely of numerical nature.

A further watchword of effective or intermediate operator based MRPT's is the (quasi) degeneracy in the spectrum of the effective matrix.[17–20] To distinguish this effect from quasi-degeneracy of matrix $\mathbf{A}(I)$ we refer to this problem as an intruder in the effective Hamiltonian. Being a state-selective theory, SS-MRPT is free from this pitfall since only one root of $\mathbf{H}^{[2]}$ of Eq. (13) is meaningful. To select the desired root, one may monitor symmetry of the vectors as well as the overlap with the starting CAS wavefunction. Interestingly, however, we have observed that the problem of kinks is accompanied by a degeneracy building up in the spectrum of $\mathbf{H}^{[2]}$, involving the target root. Though selection of the target root becomes problematic only in the exactly degenerate point, the performance of the theory drops already in the quasi-degenerate region, regarding either of the roots. From this point of view, the success of a damping as described above, lies in the modification of $\mathbf{H}^{[2]}$ in a way that the degeneracy is lifted.

## B. Sensitivity analysis

To explore the problem of kinks in more detail, a technique called sensitivity analysis may be useful. Sensitivity analysis is a tool for investigating the effect of parameter change on the solution of mathematical models, that was largely developed in connection with differential equations appearing in reaction kinetics.[21] The method is however not confined to chemical kinetics, it may be applied whenever the relation between the parameters and the results of a mathematical model is to be explored. In the present case we are dealing with linear equations of a non-dynamical model, which renders sensitivity analysis rather simple.

The unknowns of SS-MRPT are the cluster amplitudes and the linear expansion coefficients of the Jeziorski-Monkhorst Ansatz Eq. (1). The quantities are determined in two steps:

(i) Equation (14) is solved first for the amplitudes $t_{I}^{(1)\nu}$, the CAS coefficients $c_{\nu}^{(0)}$ playing the role of parameters;

(ii) Equation (12) is solved next for the coefficients $c_{\nu}$, the amplitudes $t_{I}^{(1)\nu}$ playing the role of parameters.

In principle, the one- and two-electron integrals defining the Hamiltonian at the given geometry and basis set may also be regarded as parameters of the equations. For the sake of clarity, they will not be indicated among the parameters presently, since it is neither geometry nor basis set sensitivity that we are interested in. We will also omit upper index (1) of the first order amplitudes from here on, to simplify the notations.

### 1. Coefficient sensitivity of the amplitudes

Let us first study the effect of a change in $c_{\nu}^{(0)}$ on the amplitudes. This is reflected by the partial derivatives $\partial t_{I}^{\mu}/\partial c_{\nu}^{(0)}$

$$\frac{\partial t_{I}^{\mu}}{\partial c_{\nu}^{(0)}} = -\frac{1}{c_{\nu}^{(0)}} \left(\delta_{\mu\nu} t_{I}^{\mu} + A_{\mu\nu}^{-1}(I) H_{I\nu,\nu}\right). \qquad (19)$$

The Taylor-expansion of amplitude $t_{I}^{\mu}$ around a set of reference parameter values $c_{1}^{(0)}, c_{2}^{(0)}, \ldots$ collected in column vector $\mathbf{c}^{(0)}$ reads as

$$t_{I}^{\mu}(\mathbf{c}^{(0)} + \Delta\mathbf{c}) = t_{I}^{\mu}(\mathbf{c}^{(0)}) + \sum_{\nu} \frac{\partial t_{I}^{\mu}}{\partial c_{\nu}^{(0)}} \Delta c_{\nu} + \mathcal{O}(2), \qquad (20)$$

if truncated at the quadratic term. A collective measure of amplitude perturbation produced by the alteration of the coefficients can be given by an objective function, e.g. in the form

$$e = \sum_{\mu} \left(\frac{t_{I}^{\mu}(\mathbf{c}^{(0)} + \Delta\mathbf{c}) - t_{I}^{\mu}(\mathbf{c}^{(0)})}{t_{I}^{\mu}(\mathbf{c}^{(0)})}\right)^2. \qquad (21)$$

Substituting the linear expansion (20) into the above expression one gets

$$e = \mathbf{d}^T \mathbf{S}^T \mathbf{S} \mathbf{d}, \qquad (22)$$

where

$$S_{\mu\nu} = \frac{c_\nu^{(0)}}{t_I^\mu} \frac{\partial t_I^\mu}{\partial c_\nu^{(0)}} = \frac{\partial \ln t_I^\mu}{\partial \ln c_\nu^{(0)}} \qquad (23)$$

is the normalized sensitivity matrix, and elements of vector **d** give the relative parameter change

$$d_\nu = \frac{\Delta c_\nu^{(0)}}{c_\nu^{(0)}} . \qquad (24)$$

By definition, the sensitivity matrix (23) is non-symmetric. Moreover it is a non-square matrix, since excitation $\mu \to l$ acting on model space function $\phi_\nu$ may give zero. In such a case the corresponding $t_I^\nu$ is zero by definition and it is eliminated from the problem, while $c_\nu^{(0)}$ (if nonzero) still counts as a parameter.

To grab which are the parameters having the largest influence on the solution when changed, it is useful to perform the singular value decomposition (SVD) of the sensitivity matrix. This allows to express **S** as

$$\mathbf{S} = \mathbf{U}\,\sigma\,\mathbf{V}^T, \qquad (25)$$

with **U** and **V** built on the normalized eigenvectors of $\mathbf{S}\mathbf{S}^T$ and $\mathbf{S}^T\mathbf{S}$, respectively, and the diagonal matrix $\sigma$ collecting the singular values $\sigma_i$. Substituting this form into the objective function one gets

$$e = \mathbf{d}^T \mathbf{V}\,\sigma^T\,\sigma\,\mathbf{V}^T \mathbf{d} = \sum_i \sigma_i^2 |\delta_i|^2, \qquad (26)$$

where $\delta_i$ stands for the elements of the relative parameter-change vector transformed by $\mathbf{V}^T$

$$\delta = \mathbf{V}^T \mathbf{d} . \qquad (27)$$

The right hand side of Eq. (26) gives the canonical form of the objective function $e$. This form can be used to identify the most influential parameters, since an exceedingly large $\sigma_i$ points to the coefficients which have strong perturbing effect on the amplitudes, if changed. The corresponding column of matrix **V** reveals what combination of relative changes in $c_\nu^{(0)}$-s is the most effective in achieving a large change in the amplitudes.

### 2. Amplitude sensitivities: energy derivatives

Having obtained the amplitudes, the second step of SS-MRPT is to compute the energy and the relaxed coefficients according to Eq. (12). Here one may wish to examine the effect of an alteration in the amplitudes on the eigenvectors and eigenvalues of $\mathbf{H}^{[2]}$.

Let us consider the eigenvalue first. Computing the derivative of $E^{[2]}$ with respect to an amplitude $t_I^\mu$ one may make use of the Hellman–Feynman theorem

$$\frac{\partial E^{[2]}}{\partial t_I^\lambda} = \langle \tilde{\mathbf{c}} | \frac{\partial \mathbf{H}^{[2]}}{\partial t_I^\lambda} \mathbf{c} \rangle, \qquad (28)$$

with $\tilde{\mathbf{c}}$ denoting the left eigenvector of $\mathbf{H}^{[2]}$ corresponding to $E^{[2]}$, satisfying $\langle \tilde{\mathbf{c}} | \mathbf{c} \rangle = 1$. Elements of the second order effective Hamiltonian are linear functions of the amplitudes

$$H_{\nu\mu}^{[2]} = H_{\nu\mu} + \sum_l H_{\nu l} t_{l\mu}^\mu, \qquad (29)$$

leading to the simple derivative expression

$$\frac{\partial H_{\nu\mu}^{[2]}}{\partial t_I^\lambda} = \delta_{\lambda\mu} H_{\nu,I\mu}, \qquad (30)$$

where $H_{\nu,I\mu} = H_{\nu l}$, composite index $I$ being associated with the transition $\mu \to l$. Substituting this into the energy derivative one obtains

$$\frac{\partial E^{[2]}}{\partial t_I^\lambda} = \sum_\nu \tilde{c}_\nu^* H_{\nu,I\lambda} c_\lambda, \qquad (31)$$

which may be used to build the sensitivity matrix as

$$S_{I\lambda} = \frac{t_I^\lambda}{E^{[2]}} \frac{\partial E^{[2]}}{\partial t_I^\lambda} , \qquad (32)$$

and compute its SVD to reveal whether any amplitudes have an influence on $E^{[2]}$ that is out of proportion as compared to the others. Note, that in this case matrix **S** is a row vector, since it is the single quantity $E^{[2]}$ whose amplitude sensitivity is examined. As a consequence, the only nonzero singular value is simply the norm of vector **S**.

### 3. Amplitude sensitivities: coefficient derivatives

Turning to the amplitude sensitivity of the eigenvectors of $\mathbf{H}^{[2]}$, let us take the derivative of the eigenvalue equation [Eq. (12)] with respect to $t_I^\lambda$. Denoting operation $\partial/\partial t_I^\lambda$ by prime for brevity, one obtains

$$\mathbf{H}^{[2]\prime} \mathbf{c} + \mathbf{H}^{[2]} \mathbf{c}' = E^{[2]\prime} \mathbf{c} + E^{[2]} \mathbf{c}'. \qquad (33)$$

Rearranging this equation one gets

$$(\mathbf{H}^{[2]} - E^{[2]})\mathbf{c}' = -(\mathbf{H}^{[2]\prime} - E^{[2]\prime})\mathbf{c}, \qquad (34)$$

which would yield $\mathbf{c}'$ if multiplied by the inverse of $\mathbf{H}^{[2]} - E^{[2]}$ from the left. However, this matrix is singular, $E^{[2]}$ being an eigenvector of $\mathbf{H}^{[2]}$. For this reason the reduced resolvent is defined as[27,28]

$$\mathbf{G}(\mathbf{H}^{[2]} - E^{[2]}) = \mathbf{1} - |\mathbf{c}\rangle\langle\tilde{\mathbf{c}}|. \qquad (35)$$

Left multiplication of Eq. (34) by **G** yields

$$|\mathbf{c}'\rangle - |\mathbf{c}\rangle\langle\tilde{\mathbf{c}}|\mathbf{c}'\rangle = -\mathbf{G}(\mathbf{H}^{[2]\prime} - E^{[2]\prime})|\mathbf{c}\rangle . \qquad (36)$$

Projection of Eq. (36) by $\langle \mathbf{c}|$ gives the scalar product of $\langle\tilde{\mathbf{c}}|$ and $|\mathbf{c}'\rangle$:

$$\langle\tilde{\mathbf{c}}|\mathbf{c}'\rangle = \langle\mathbf{c}|\mathbf{G}(\mathbf{H}^{[2]\prime} - E^{[2]\prime})|\mathbf{c}\rangle, \qquad (37)$$

having used that $\langle\mathbf{c}|\mathbf{c}'\rangle = 0$, a consequence of the normalization condition of the eigenvector, $\langle\mathbf{c}|\mathbf{c}\rangle = 1$. Combining Eqs. (37) and (36) the derivative $\mathbf{c}'$ can be expressed as

$$|\mathbf{c}'\rangle = -(1 - |\mathbf{c}\rangle\langle\mathbf{c}|)\,\mathbf{G}(\mathbf{H}^{[2]\prime} - E^{[2]\prime})|\mathbf{c}\rangle . \qquad (38)$$

Substituting expressions for $\mathbf{H}^{[2]\prime}$ and $E^{[2]\prime}$ derived previously, the derivative of the coefficients with respect to the amplitudes can be finally written as

$$\frac{\partial c_\mu}{\partial t_I^\lambda} = -\sum_{\nu\sigma\tau} L_{\mu\nu}\,G_{\nu\sigma}\,K_{\sigma\tau}\,H_{\tau,I\lambda}\,c_\lambda, \qquad (39)$$

and can be utilized to construct the sensitivity matrix

$$S_{\mu,I\lambda} = \frac{t_I^\lambda}{c_\mu} \frac{\partial c_\mu}{\partial t_I^\lambda}. \tag{40}$$

Matrices **L** and **K** in the above [Eq. (39)] are the representation of projector $1 - |\mathbf{c}\rangle\langle\mathbf{c}|$ and $1 - |\mathbf{c}\rangle\langle\tilde{\mathbf{c}}|$, respectively, with elements $L_{\mu\nu} = \delta_{\mu\nu} - c_\mu c_\nu^*$ and $K_{\sigma\tau} = \delta_{\sigma\tau} - c_\sigma \tilde{c}_\tau^*$. Computation of matrix **G** is straightforward, once the spectral form of $\mathbf{H}^{[2]}$ is available.

### 4. Eventual coefficient sensitivities: derivatives of the energy

It may be useful to link the above derivatives via the chain rule and study the eventual effect of CAS coefficient variation on the final SS-MRPT quantities. Regarding the energy, the derivative can be expressed as

$$\frac{\partial E^{[2]}}{\partial c_\mu^{(0)}} = \sum_{I\lambda} \frac{\partial E^{[2]}}{\partial t_I^\lambda} \frac{\partial t_I^\lambda}{\partial c_\mu^{(0)}} = -\frac{1}{c_\mu^{(0)}} \sum_{\nu\lambda} \tilde{c}_\nu^* F_{\nu\lambda}^\mu c_\lambda, \tag{41}$$

the elements of matrices $\mathbf{F}^\mu$ being

$$F_{\nu\lambda}^\mu = \sum_I H_{\nu,I\lambda} \left( \delta_{\lambda\mu} t_I^\mu + A_{\lambda\mu}^{-1}(I) H_{I\mu,\mu} \right). \tag{42}$$

For each $I$, terms of matrices $\mathbf{F}^\mu$ can be computed and accumulated and then dropped to spare the storage of the numerous derivatives. The sensitivity matrix can finally be expressed as

$$S_\mu = \frac{c_\mu^{(0)}}{E^{[2]}} \frac{\partial E^{[2]}}{\partial c_\mu^{(0)}}, \tag{43}$$

and subjected to SVD.

### 5. Eventual coefficient sensitivities: derivatives of the coefficients

Finally, we give the expressions for studying the CAS coefficient sensitivity of the relaxed coefficients. For this end the partial derivatives

$$\frac{\partial c_\mu}{\partial c_\nu^{(0)}} = \sum_{I\lambda} \frac{\partial c_\mu}{\partial t_I^\lambda} \frac{\partial t_I^\lambda}{\partial c_\nu^{(0)}} = \frac{1}{c_\nu^{(0)}} \sum_{\kappa\lambda\sigma\tau} L_{\mu\kappa} G_{\kappa\sigma} K_{\sigma\tau} F_{\tau\lambda}^\nu c_\lambda \tag{44}$$

are to be constructed and substituted into

$$S_{\mu\nu} = \frac{c_\nu^{(0)}}{c_\mu} \frac{\partial c_\mu}{\partial c_\nu^{(0)}}, \tag{45}$$

which is the sensitivity matrix to be analyzed by performing its SVD.

We note at this point that instead of the relative change in the parameters, one may alternatively monitor absolute changes. Derivation of this analysis is analogous to the above and leads to the SVD of un-normalized sensitivity matrices as shown in Appendix A. Numerical experience indicates that normalized sensitivities show better correlation with kinks on energy curves than un-normalized quantities, we therefore present only normalized sensitivities in the numerical applications.

## III. NUMERICAL ILLUSTRATIONS

In the examples presented below singular values of the eventual coefficient sensitivity matrices [Eqs. (43) and (45)] are reported. To avoid numerical ill-effects, removal of amplitudes or coefficients below threshold in magnitude is performed, whenever a division is needed in the final expression of the sensitivity matrix. For example $c_\mu$ is checked in the case of Eq. (45). The dropout threshold is typically $10^{-8}$.

In the case of Eq. (43) the only nonzero singular value, which is the norm of vector **S**, is examined. In the case of Eq. (45) we are focusing on the few largest singular values. We note here, that if an exact degeneracy affects root **c** of $\mathbf{H}^{[2]}$, the reduced resolvent **G** of Eq. (35) is still ill-defined, since the projector of the entire degenerate subspace should be subtracted from unity on the right hand side. This occurs just at a few geometry points on the potential surface, these are simply omitted when performing sensitivity analysis.

The sensitivity diagnostic tool is illustrated on the example of the bond dissociation process of the LiH molecule and the insertion of Be among two H atoms. Relatively small basis sets are applied, so that full configuration interaction (full CI or FCI) is attainable as benchmark.

Of the possible formulations of SS-MRPT, the spin-free variant is applied here.[22,30] Specification of the zero-order Hamiltonian—i.e., the partitioning—requires the definition of $E_\mu^{(0)}$ and $E_I^{(0)}$ in Eq. (10). We investigate two choices for constructing the zero-order energies: diagonals of Fockians are used in Møller–Plesset (MP) partitioning, while expectation values of the full Hamiltonian are computed with functions $\phi_\mu$ and $\chi_I$ in the Epstein–Nesbet (EN) partitioning. Size-extensivity requirement is satisfied both by MP and EN partitionings.[4]

Based on Eq. (14), our equations for the single excitations' amplitudes read

$$t_{ia}^\nu = -\frac{1}{c_\nu^{(0)}} \sum_\mu{}' A_{\nu\mu}^{-1}(ia) c_\mu^{(0)} \tilde{f}_{ia}^\mu, \tag{46}$$

the double excitations' amplitudes are given by

$$t_{ijab}^\nu = -\frac{1}{c_\nu^{(0)}} \sum_\mu{}' A_{\nu\mu}^{-1}(ijab) c_\mu^{(0)} \langle ij|ab\rangle. \tag{47}$$

Here, the composite excitation index—previously referred to as $I$—is resolved as orbital index pair $ia$ or the quartet $ijab$, $i$, $j$ labeling orbitals singly or doubly occupied in $\phi_\nu$, $a$, $b$ labeling orbitals unoccupied or singly occupied in $\phi_\nu$. For the two-electron integral the $\langle 12|12\rangle$ convention is applied. Note that $H_{I\mu,\mu}$ in the inhomogeneous term is substituted by $\tilde{f}_{ia}$ for singles and by $\langle ij|ab\rangle$ for doubles. Matrix elements of the Fockian are defined as

$$\tilde{f}_{pq}^\mu = f_{pq}^0(\mu) + \sum_{u_s \in \phi_\mu}{}' \langle pu_s|qu_s\rangle, \tag{48}$$

where indices $p$, $q$ are generic, $u_s$ refers to orbitals singly occupied in $\phi_\mu$ and $f_{pq}^0(\mu)$ are elements of the Fockian built with the doubly occupied part of $\phi_\mu$. The prime on the sum above indicates a restriction: $\langle pu_s|qu_s\rangle$ is excluded for $u_s = p = q$, with the aim of omitting self-interaction. Note, that elements of the Fockian show a dependence on the reference

function $\phi_\mu$. This means that not one, but several zero-order Hamiltonians are used, a treatment reminiscent of the multi-partitioning theory introduced by Malrieu and co-workers.[23]

Matrix $\mathbf{A}(I)$ of Eq. (15) is the final ingredient to be specified. This is essentially the model space block of $(\mathbf{H} - E_{CAS})$ to which the zero-order excitation energy

$$X(I) = E_l^{(0)} - E_\mu^{(0)} \tag{49}$$

is added in the diagonals. In MP partitioning the excitation energy is

$$X(ia) = -\tilde{f}_{ii}^\mu + \overline{f}_{aa}^\mu , \tag{50}$$

if single excitation $E_i^a$ connects functions $\phi_\mu$ and $\chi_l$, and

$$X(ij, ab) = -\tilde{f}_{ii}^\mu - \tilde{f}_{jj}^\mu + \overline{f}_{aa}^\mu + \overline{f}_{bb}^\mu$$
$$- \delta_{ib}\langle ib|ib\rangle - \delta_{ja}\langle ja|ja\rangle, \tag{51}$$

if double excitation $E_j^b E_i^a$ takes from $\phi_\mu$ to $\chi_l$. Note, that $\overline{f}$ is used for particle labels, where self-interaction is retained in Eq. (48) (i.e., no restriction on the sum). Two-electron integrals in $X(ij, ab)$ take care of converting $\overline{f}$ to $\tilde{f}$ so that spectator excitations do not have any contribution.

Formulae of zero-order excitation energies in EN partitioning are somewhat lengthier, these expressions are given in Appendix B.

Orbital dependence of the PT results is investigated by using two choices for the active set: pseudo-canonical or natural. In the former case the active block of the generalized Fockian built with the density matrix of the target root is diagonal, while in the latter case the active block of the density matrix is diagonal. The computation of derivatives necessitates well converged CAS orbitals, which was achieved by requiring that asymmetry of the Lagrangian got smaller than $2 \times 10^{-10}$ $E_h$ upon convergence.

## A. LiH

The single bond dissociation of the LiH molecule is computed in Dunning's double zeta plus polarization (DZP) set.[24] The CAS reference function is produced by distributing two active electrons on five active orbitals, with symmetry labels $2a_1, 3a_1, 4a_1, 1b_1, 1b_2$ classified according to $C_{2v}$.

In Fig. 1 we present the energy error of second order SS-MRPT results in both partitionings, using pseudo-canonical orbitals. A kink of ca. 0.5 $mE_h$ amplitude is apparent in MP partitioning around 6 Å internuclear distance. A larger, ca. 1.5 $mE_h$ amplitude kink is produced around 3 Å internuclear distance in EN partitioning. If applying the damping of Eq. (17) when computing the effective Hamiltonian, the kinks get smoothed, as shown by the hollow symbols in Fig. 1.

An illustration focusing on the problematic geometry range is given by Figs. 2–4 in MP and by Figs. 5–7 in EN partitioning.

According to Figs. 2 and 5 a degeneracy of the target root of $\mathbf{H}^{[2]}$ is built up at the problematic geometries. Passing the degenerate region of the potential energy curve, one of the roots of the effective Hamiltonian falls below the target root. The fact, that the ground state of the system does not correspond to the lowest root of $\mathbf{H}^{[2]}$ does not represent a problem
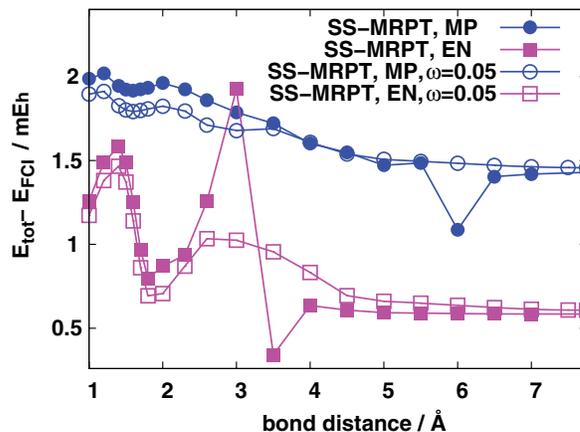


FIG. 1. Errors of second order SS-MRPT energies for the ground state of the LiH molecule in Dunning's DZP basis (Ref. 24). Reference function is CAS(2,5), active orbitals are pseudo-canonicals. Full CI values are subtracted from total SS-MRPT energies. Partitioning is either MP or EN. Filled symbols correspond to undamped equations, results obtained by Tikhonov damping according to Eq. (17) are shown by hollow symbols. The damping parameter is denoted by $\omega$.

in itself: the kink appears in the quasi-degenerate region but not at larger bond distances. Note, that the degeneracy in the spectrum of $\mathbf{H}^{[2]}$ is a pure artifact, no such effect is observed in the full CI solution. Either by applying Tikhonov damping [Eq. (17)] or by turning to natural orbitals in the active space, the degeneracy of the effective Hamiltonian is lifted. The second (and unphysical) root of SS-MRPT—not shown in the Figures — lies by a rough 0.2 $E_h$ above the ground state in MP and by 0.05–0.2 $E_h$ (depending on the actual treatment) in EN partitioning in the respective geometry region.

Analysis of the CAS coefficient sensitivity of the relaxed coefficients presented in Figs. 3 and 6 show a drastic increase in the largest singular value for the method producing the kink (undamped SS-MRPT with pseudo-canonical orbitals) in the
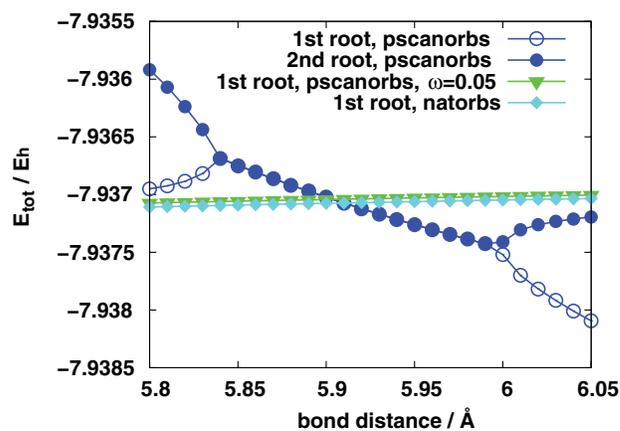


FIG. 2. Total energy of the ground state of the LiH molecule as a function of the internuclear distance, obtained by second order SS-MRPT in MP partitioning. Basis set and reference function agrees with that of Fig. 1. The nature of orbitals (pscanorb: pseudo-canonicals; natorbs: naturals) is varied. Whenever Tikhonov damping is used [cf. Eq. (17)] symbol $\omega$ is indicated and shows the value of the damping parameter. Both roots of the second-order Hamiltonian $\mathbf{H}^{[2]}$ are shown for the undamped equations, displayed by circles.
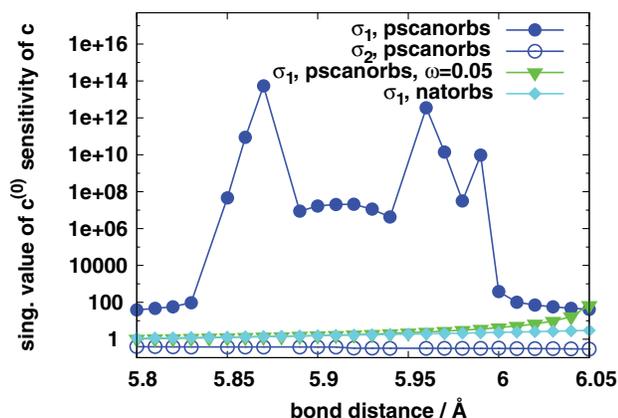
FIG. 3. Singular values of the coefficient sensitivity matrix Eq. (45) around 6 Å internuclear distance, for the ground state of the LiH molecule in MP partitioning. The largest singular value is denoted by $\sigma_1$, the second largest is $\sigma_2$. Basis set and reference function agrees with that of Fig. 1. The nature of orbitals (pscanorb: pseudo-canonicals; natorbs: naturals) is varied. Symbol $\omega$ refers to Tikhonov damping (c.f. Eq. (17)) and gives the value of the damping parameter.

geometry regions under study. The singular value shoots up more than 10 orders of magnitude in either partitionings. This occurs only for the largest singular value, the second largest is unaffected and remains similar to $\sigma_1$ of the techniques giving a smooth potential curve. This latter value remains $\sigma_1 < 10^2$ if applying damping with pseudo-canonical orbitals or $\sigma_1 < 10^3$ if using natural orbitals, as manifested by Figs. 3 and 6.

Regarding CAS coefficient sensitivity of the SS-MRPT energy, Figs. 4 and 7 also show peaks for the ill-behaving method in the quasi-degenerate regime. The increase in the singular value reflected by Figs. 4 and 7 is less drastic, than what is observed with derivatives of the coefficients: it mounts to about two orders of magnitude. The peaks, however, can be clearly identified. Interestingly, the shoot up in the coefficient sensitivity of the energy is smaller for EN than for MP, though the kink is larger for EN in Fig. 1. Similarly to the observations made above, singular values of the well-
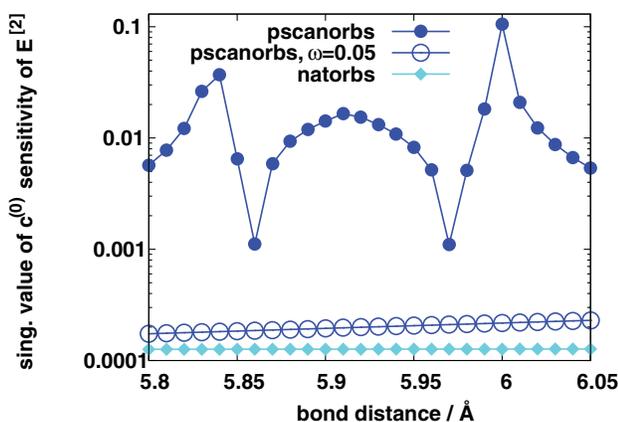


FIG. 4. Nonzero singular value (i.e. norm) of the coefficient sensitivity matrix Eq. (43), around 6 Å internuclear distance, for the ground state of the LiH molecule in MP partitioning. Basis set and reference function agrees with that of Fig. 1. The nature of orbitals (pscanorb: pseudo-canonicals; natorbs: naturals) is varied. Symbol $\omega$ refers to Tikhonov damping (c.f. Eq. (17)), and gives the value of the damping parameter.
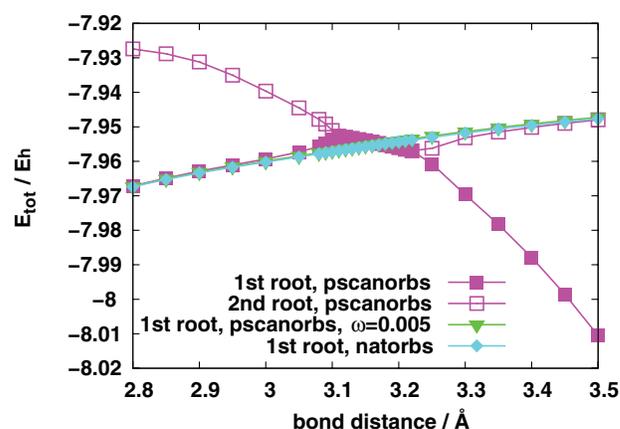


FIG. 5. Same as Fig. 2 in EN partitioning, around 3 Å bond distance.

behaving methods in Figs. 4 and 7 remain flat and small, on the order of $10^{-4}$.

## B. BeH$_2$

The classical test case of the BeH$_2$ system with the geometry points defined by Purvis and Bartlett[25] is computed in a valence double zeta basis. Dunning's DZ set[24] is taken for the hydrogen atoms. For beryllium the basis of Purvis *et al.*[25] is used with the *p* function decontracted leaving the most compact primitive (exponent 5.693880) alone and contracting the remaining two into a second *p* function (exponents 1.555630, 0.171855 and coefficients 0.144045, 0.949692, respectively). Coordinates of the hydrogen atoms in atomic units are $(0, \pm2.54, 0)$, $(0, \pm2.08, 1.0)$, $(0, \pm1.62, 2.0)$, $(0, \pm1.39, 2.5)$, $(0, \pm1.275, 2.75)$, $(0, \pm1.16, 3.0)$ $(0, \pm0.93, 3.5)$, $(0, \pm0.70, 4.0)$, $(0, \pm0.70, 6.0)$, respectively, at points A, B, C, D, E, F, G, H, and I. The beryllium lies in the origin of the coordinate system. A CAS(4,4) function is computed as reference, with two active orbitals of $a_1$ symmetry, and the other two being $b_1$ as classified in $C_{2v}$.
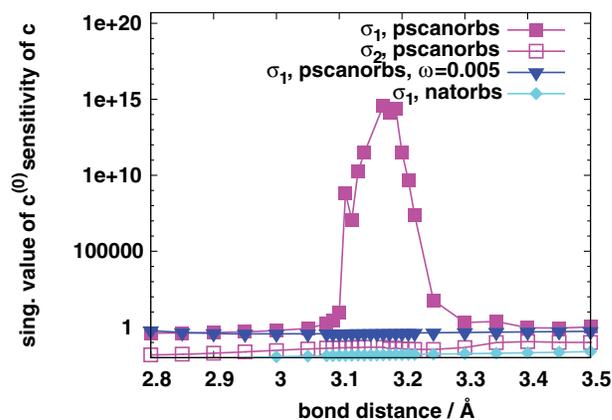


FIG. 6. Same as Fig. 3 in EN partitioning, around 3 Å bond distance.
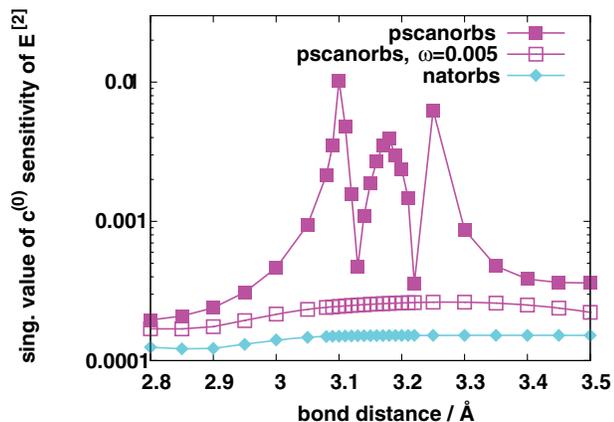
FIG. 7. Same as Fig. 4 in EN partitioning, around 3 Å bond distance.



FIG. 9. Nonzero singular value (i.e. norm) of the coefficient sensitivity matrix Eq. (43) (upper plot) and the two largest singular values of the of the coefficient sensitivity matrix Eq. (45) (bottom plot) for the ground state of the BeH$_2$ system in MP partitioning, using pseudo-canonical active orbitals. Basis set and reference function agrees with that of Fig. 8. Symbol $\omega$ refers to Tikhonov damping (c.f. Eq. (17)) and gives the value of the damping parameter.

In Fig. 8 energy errors are displayed in both partitionings. The curves obtained with pseudo-canonical active orbitals show kinks at geometry point E in MP and at points B and E in EN partitioning. Similarly to the previous experience, a damping according to Eq. (17) gets the error curve smoother, as apparent in Fig. 8. Instead of damping the division by $c_\mu^{(0)}$, a damping of inversion of matrix $\mathbf{A}(I)$ according to Eq. (18) is also tested in this case, and gives results indistinguishable from the hollow symbols on the scale of Fig. 8. However, a damping of $\mathbf{A}(I)^{-1}$ may require larger damping factor than the damping of $1/c_\mu^{(0)}$: in particular, $\omega = 0.17$ is needed for point E and $\omega = 0.005$ suffices at point B. It is also observed that the optimal factor $\omega$ needed for damping $\mathbf{A}(I)^{-1}$ varies more with geometry than the factor needed for damping according to Eq. (17).

Error curves of SS-MRPT when using natural active orbitals are also shown in Fig. 8 and they again prove to be smoother than the curves by pseudo-canonical orbitals. A kink of ca. 5 m$E_h$ is apparent in the geometry region E-F in EN partitioning. This however is unaffected by damping: a

roughly constant shift of the error curve is all that is achievable by a sufficiently large factor $\omega$.

In Figs. 9 and 10 an insight offered by sensitivities is presented for the case of the pseudo-canonical active orbitals. Derivatives of the SS-MRPT energy are always informative, as reflected in the upper plots of the figures. We see a rough order of magnitude increase in the singular value of Eq. (43) at the problematic geometries, i.e., E in the case of MP and B and E in the case of EN partitioning. Coefficient sensitivity of the relaxed coefficients in MP partitioning, shown at the bottom plot of Fig. 9 is unindicative. Taking into account that this sensitivity is seen to shoot up several orders of magnitude at the problematic points, the rough order of magnitude increase cannot be regarded as significant. The fact, that the peak at point D is insignificant is underlined by the observation that the singular value at point D *increases* upon damping. From this we deduce that the values shown at the bottom of Fig. 9 are insignificantly small. In the case of EN partitioning on the other hand we see a considerable effect in coefficient sensitivities of the coefficients: a peak of ca. 12 orders of magnitude at point E at the bottom of Fig. 10.



FIG. 8. Errors of second order SS-MRPT energies for the ground state of the BeH$_2$ system in DZ basis (see text for more detail). Reference function is CAS(4,4), active orbitals are either pseudo-canonicals (pscanorb) or naturals (natorbs). Full CI values are subtracted from total SS-MRPT energies. Partitioning is either MP or EN. Filled symbols correspond to undamped equations, results obtained by Tikhonov damping according to Eq. (17) are shown by hollow symbols. The damping parameter is denoted by $\omega$.



FIG. 10. Same as Fig. 9 in EN partitioning.

TABLE I. Right singular vectors of the CAS coefficient sensitivity of the relaxed coefficients, Eq. (45) for the ground state of the $BeH_2$ system in EN partitioning, with pseudo-canonical active orbitals. Basis set and reference function agrees with that of Fig. 8. The largest singular value is given by $\sigma_1$.

| geometry B $\sigma_1 = 461$ | | geometry E $\sigma_1 = 1.17 \cdot 10^{14}$ | |
| --- | --- | --- | --- |
| CAS coeff. | right sing. vector | CAS coeff. | right sing. vector |
| 0.99109292 | 0.116 | −0.80937851 | −0.050 |
| −0.00420961 | −0.316 | −0.07705062 | 0.009 |
| −0.00151338 | 0.892 | 0.11577611 | −0.031 |
| −0.06926363 | 0.021 | 0.51498539 | 0.939 |
| −0.08774899 | 0.111 | 0.19855833 | −0.059 |
| −0.00464325 | −0.256 | 0.09262726 | 0.332 |
| −0.04218947 | 0.073 | 0.06034612 | 0.014 |
| 0.00259208 | 0.084 | 0.00308458 | −0.004 |
| −0.05028701 | −0.004 | −0.00097524 | 0.010 |
| 0.00144061 | −0.001 | −0.04022365 | 0.008 |
| −0.02854409 | 0.007 | 0.06490194 | −0.006 |
| 0.00797646 | 0.001 | −0.05355391 | 0.003 |

The case of natural orbitals is not displayed. An inspection of singular values along the reaction path reveals a mild peak in sensitivities at point B in MP partitioning. This means an increase from ca. $10^{-4}$ to ca. $10^{-3}$ in the singular value of Eq. (43) and from ca. $10^2$ to ca. $10^5$ in the singular value of Eq. (45). Since no pronounced kink is visible at point B in Fig. 8 with natural orbitals, one may deduce that sensitivities represent a finer tool of problem-detection than monitoring kinks on error curves.

Apart from singular values, it is instructive to examine singular vectors as well. Right singular vector corresponding to the largest singular value of Eq. (45), for example, carries information on which CAS coefficient the effect of kink originates, while the corresponding left singular vector reflects which are the most affected coefficients of the relaxed eigen-



FIG. 11. Analysis of the unrelaxed second order SS-MRPT energy, $E^{[2]} = \langle \Phi | H^{[2]} | \Phi \rangle$ for the ground state of the $BeH_2$ system in DZ basis. Reference function is CAS(4,4), active orbitals are pseudo-canonicals. Filled symbols correspond to undamped equations, results obtained by Tikhonov damping according to Eq. (17) are shown by hollow symbols. The damping parameter is denoted by $\omega$. Energy errors are shown in the upper plot. Nonzero singular value (i.e. norm) of the coefficient sensitivity of the unrelaxed $E^{[2]}$ is shown in the bottom plot.

vector of $H^{[2]}$. The two geometries shown in Table I represent examples of two typical cases. At point B these are indeed the small coefficients—on the order of $10^{-3}$—for which the elements of the right singular vector are significant. In this case, the corresponding reference functions may be omitted from the SS-MRPT calculation and the corresponding amplitudes may be zeroed. The SS-MRPT energy may get remarkably better by this procedure, e.g., omitting reference function No. 3 at point B of the $BeH_2$ molecule reduces the error from 10.4 m$E_h$ to 0.2 m$E_h$. At geometry E on the other hand, the most significant component of the right singular vector is the second most important component of the CAS vector. In this case dropping the responsible element of the CAS wavefunction does not result in a reasonable SS-MRPT energy. Omitting reference function No.4, for example, reduces the original error of 26.6 m$E_h$ at point E only to 11.7 m$E_h$. This is still out of proportion if compared to the errors at point D and F, 0.1 m$E_h$ and −1.2 m$E_h$, respectively.

Finally, we present energy and sensitivity results obtained with the unrelaxed second order energy expression [Eq. (30)] of the study by Evangelista et al.[5] As the upper plot of Fig. 11 shows, no kinks are produced on the errors curves by this method. There is only a slight increase of the absolute error at point E in EN partitioning and it is apparently unaffected by damping. In accordance with this, sensitivities shown in the bottom part of the plot are unindicative. The peak apparent at point E is negligible, since it occurs in the $10^{-3}$ range while the value lies in the $10^0$ range.

## IV. CONCLUSION

To summarize, it has been demonstrated that sensitivity analysis represents a diagnostic tool for the unexpected ill-behavior of SS-MRPT potential energy curves at certain geometries. The analysis does not provide any explanation for why the effect occurs. It rather offers an insight on what quantities of the theory are affected and why damping of the equations is helpful. By the inspection of right singular vectors of coefficient sensitivity matrices, the analysis also reveals why a cutoff of CAS coefficients is helpful in some cases and ineffective in others.

It is apparent from the examples shown, that analyzing singular values of sensitivity matrices is not a clear-cut issue. It is hard to give a threshold value, above which a singular value would point to a problem. It is rather an expressed increase in sensitivity than its actual value, which can be regarded as an indicator. As reference value, either sensitivities of neighboring geometry points or those lying below the largest one (if there is any) may be useful. In our examples, we have regarded a singular value of Eq. (43) indicative if it stood out of the others by at least one order of magnitude, and an increase of at least 2–3 orders of magnitude has been called significant for Eq. (45). A sensitivity out of proportion of the others often appears in the form of a peak, i.e., it builds up in a relatively narrow geometry range. Two types of coefficient sensitivities were studied: that of the SS-MRPT energy and the relaxed coefficients. Sensitivities getting large implies a large change in the respective quantity upon orbital rotation.

This effect is clearly undesirable, giving the motive of considering it a problem.

Our numerical experience indicates that the kinks produced on energy error curves vary a lot with orbitals: they rather appear if using pseudo-canonical active orbitals and are less likely to be observed with natural active orbitals. Changing the definition of the Fockian of Eq. (48) and the expression of the inhomogeneous term in Eq. (14) is also expected to affect the appearance of the kinks. The problem is manifestly more severe in EN partitioning: the amplitude and the number of kinks is often larger in EN than in MP.

A question not addressed in the present study is the sensitivity of molecular properties as obtained by SS-MRPT. It would be also desirable to arrive at a modified formulation of SS-MRPT equations, which avoids the shoot up of sensitivities. According to our experience, the second order expression of Ref. 5 is free from this effect, it however leaves the coefficients unrelaxed. The cure used for the relaxed formulation—Tikhonov damping—represents a numerical workaround with no physical motivation. The ultimate goal would clearly be to take only physically well-grounded approximation steps.

## ACKNOWLEDGMENTS

## APPENDIX A: ANALYSIS WITH UN-NORMALIZED SENSITIVITY MATRIX

Taking the example of coefficient sensitivity of the amplitudes, the objective function for absolute changes would look

$$f = \sum_{\mu} \left( t_I^{\mu}(\mathbf{c}^{(0)} + \Delta \mathbf{c}) - t_I^{\mu}(\mathbf{c}^{(0)}) \right)^2, \qquad (A1)$$

which can be expressed with the un-normalized sensitivity matrix[29]

$$R_{\mu\nu}^I = \frac{\partial t_I^{\mu}}{\partial c_{\nu}^{(0)}}, \qquad (A2)$$

in the form

$$f = \Delta \mathbf{c}^T \mathbf{R}^T \mathbf{R} \Delta \mathbf{c}. \qquad (A3)$$

Calculating the SVD of $\mathbf{R}$ written as

$$\mathbf{R} = \mathbf{X} \rho \mathbf{Y}^T, \qquad (A4)$$

the canonical form of the objective function becomes

$$f = \sum_i \rho_i^2 |\gamma_i|^2, \qquad (A5)$$

with $\gamma = \mathbf{Y}^T \Delta \mathbf{c}$. An analysis based on $\mathbf{R}$ instead of the normalized matrix $\mathbf{S}$ may be advantageous if one would run into an ill-defined division, e.g., an amplitude being almost zero in this example. On the other hand, the normalized coefficient sensitivity matrix Eq. (23) offers the advantage of canceling the leading $1/c_\nu^{(0)}$ factor of Eq. (19). This prompts to prefer normalized sensitivities to un-normalized ones.

## APPENDIX B: ZERO-ORDER EXCITATION ENERGIES IN EN PARTITIONING

Labeling convention is different from the text. Here $i$, $j$ is restricted to doubly occupied orbitals in function $\phi_\mu$ and $u_s$, $v_s$ is introduced for singly occupied actives. (In the text $i$, $j$ refers to both categories.) Hence

$i$, $j$    doubly occupied, active or inactive,
$u_s$, $v_s$ active, singly occupied,
$a$, $b$    virtual or singly occupied, active or inactive.

Single excitations

$$X(ia) = -\widetilde{f}_{ii}(\mu) + \overline{f}_{aa}(\mu) - \langle ia|ia \rangle + 2\langle ia|ai \rangle, \qquad (B1)$$

$$X(u_s a) = -\widetilde{f}_{u_s u_s}(\mu) + \overline{f}_{aa}(\mu) - \langle u_s a|u_s a \rangle. \qquad (B2)$$

Double excitations

$$\begin{aligned}
X(ij, ab) = {} & -\widetilde{f}_{ii}(\mu) - \widetilde{f}_{jj}(\mu) + \overline{f}_{aa}(\mu) + \overline{f}_{bb}(\mu) - \langle ia|ia \rangle \\
& - \langle ib|ib \rangle - \langle ja|ja \rangle - \langle jb|jb \rangle \\
& + \langle ab|ab \rangle + \langle ij|ij \rangle + 2\langle ia|ai \rangle + 2\langle jb|bj \rangle \\
& + \delta_{ij}(1 - \delta_{ab})\langle ab|ba \rangle + \delta_{ab}(1 - \delta_{ij})\langle ij|ji \rangle \\
& - [\delta_{ij}(1 - \delta_{ab}) + \delta_{ab}(1 - \delta_{ij}) + \delta_{ab}\delta_{ij}] \\
& \times [\langle ia|ai \rangle + \langle jb|bj \rangle], \qquad (B3)
\end{aligned}$$

$$\begin{aligned}
X(u_s j, ab) = {} & -\widetilde{f}_{u_s u_s}(\mu) - \widetilde{f}_{jj}(\mu) + \overline{f}_{aa}(\mu) + \overline{f}_{bb}(\mu) \\
& - \langle u_s a|u_s a \rangle - \langle u_s b|u_s b \rangle - \langle ja|ja \rangle - \langle jb|jb \rangle \\
& + \langle ab|ab \rangle + \langle u_s j|u_s j \rangle + 2\langle jb|bj \rangle \\
& - \delta_{ab}\langle jb|bj \rangle, \qquad (B4)
\end{aligned}$$

$$\begin{aligned}
X(u_s v_s, ab) = {} & -\widetilde{f}_{u_s u_s}(\mu) - \widetilde{f}_{v_s v_s}(\mu) + \overline{f}_{aa}(\mu) + \overline{f}_{bb}(\mu) \\
& - \langle u_s a|u_s a \rangle - \langle v_s a|v_s a \rangle - \langle u_s b|u_s b \rangle \\
& - \langle v_s b|v_s b \rangle + \langle ab|ab \rangle + \langle u_s v_s|u_s v_s \rangle \\
& - \delta_{ab}\langle u_s v_s|v_s u_s \rangle. \qquad (B5)
\end{aligned}$$

[1] U. S. Mahapatra, B. Datta, and D. Mukherjee, J. Phys. Chem. A **103**, 1822 (1999).

[2] U. S. Mahapartra, B. Datta, and D. Mukherjee, Chem. Phys. Lett. **299**, 42 (1999).

[3] U. S. Mahapatra, B. Datta, and D. Mukherjee, Chem. Phys. Lett. **301**, 206 (1999).

[4] P. Ghosh, S. Chattopadhyay, D. Jana, and D. Mukherjee, Int. J. Mol. Sci. **3**, 733 (2002).

[5] F. A. Evangelista, A. C. Simmonett, H. F. Schaefer III, D. Mukherjee, and W. D. Allen, Phys. Chem. Chem. Phys. **11**, 4728 (2009).

[6] R. K. Chaudhuri, K. F. Freed, G. Hose, P. Piecuch, K. Kowalski, M. Wloch, S. Chattopadhyay, D. Mukherjee, Z. Rolik, Á. Szabados, G. Tóth, and P. R. Surján, J. Chem. Phys. **122**, 134105 (2005).

[7] M. R. Hoffmann, D. Datta, S. Das, D. Mukherjee, Á. Szabados, Z. Rolik, and P. R. Surján, J. Chem. Phys. **131**, 204104 (2009).

[8] S. Chattopadhyay, U. S. Mahapatra, and D. Mukherjee, J. Chem. Phys. **111**, 3820 (1999).

[9] U. S. Mahapatra, S. Chattopadhyay, and R. K. Chaudhuri, J. Chem. Phys. **130**, 014101 (2009).

[10] S. Das, D. Mukherjee, and M. Kállay, J. Chem. Phys. **132**, 074103 (2010).

[11] A. Engels-Putzka and M. Hanrath, Mol. Phys. **107**, 143 (2009).

[12] U. S. Mahapatra, B. Datta, and D. Mukherjee, J. Chem. Phys. **110**, 6171–6188 (1999).

[13] B. Jeziorski and H. J. Monkhorst, Phys. Rev. A **24**, 1668 (1981).

[14] S. Chattopadhyay, D. Pahari, D. Mukherjee, and U. S. Mahapatra, J. Chem. Phys. **120**, 5968 (2004).

[15] U. S. Mahapatra, S. Chattopadhyay, and R. K. Chaudhuri, J. Chem. Phys. **129**, 024108 (2008).

[16] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems* (Dordrecht, Kluwer, 1996).

[17] Ph. Durand and J.-P. Malrieu, Adv. Chem. Phys. **67**, 1 (1987).

[18] S. Evangelisti, J.-P. Daudey, and J.-P. Malrieu, Phys. Rev. A **35**, 4930 (1987).

[19] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, Phys. Rev. A **47**, 2738 (1993).

[20] K. Kowalski and P. Piecuch, Phys. Rev. A **61**, 052506 (2000).

[21] T. Turányi, J. Math. Chem. **5**, 203 (1990).

[22] D. Pahari, S. K. Chattopadhyay, S. Das, D. Mukherjee, and U. S. Mahapatra, in *Theory and Applications of Computational Chemistry: The First 40 years of Quantum Chemistry*, edited by C. E. Dykstra, K. S. Kim, G. Frenking, and G. E. Scuseria (Elsevier, Amsterdam, 2005), pp. 581–633.

[23] A. Zaitevskii and J.-P. Malrieu, Chem. Phys. Lett. **233**, 597 (1995).

[24] T. H. Dunning Jr, J. Chem. Phys. **53**, 2829 (1970).

[25] G. D. Purvis, R. Shepard, F. B. Brown, and R. J. Bartlett, Int. J. Quantum Chem. **23**, 835 (1983).

[26] M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry: The First 40 years of Quantum Chemistry*, edited by C. E. Dykstra, K. S. Kim, G. Frenking, and G. E. Scuseria (Elsevier, Amsterdam, 2005), pp. 1167–1189.

[27] P.-O. Löwdin, J. Mol. Spectrosc. **10**, 12 (1963).

[28] Note, that operator $\mathbf{G}$ is not fully defined by Eq. (35). Following Löwdin (Ref. 27), a proper formal definition can be given as $\mathbf{G} = \Pi \, [\eta\Omega + \Pi(\mathbf{H}^{[2]} - E^{[2]})\Pi]^{-1} \, \Pi$, with $\eta$ being an arbitrary but nonzero scalar, $\Omega = |\mathbf{c}\rangle\langle\tilde{\mathbf{c}}|$ and $\Pi = 1 - |\mathbf{c}\rangle\langle\tilde{\mathbf{c}}|$.

[29] In sensitivity analysis term "normalization" refers to setting the sensitivity matrix dimensionless.

[30] D. Datta and D. Mukherjee, J. Chem. Phys. **134**, 054122 (2011).