

Spin-adaptation and redundancy in state-specific multireference perturbation theory

Péter Jeszenszki, Péter Surján and Ágnes Szabados

Laboratory of Theoretical Chemistry, Loránd Eötvös University,

H-1518 Budapest, POB 32, Hungary

Abstract

Spin-adaptation of virtual functions in state-specific multireference perturbation theory (SS-MRPT) is examined. Redundancy occurring among virtual functions generated by unitary group based excitation operators on a model-space function, is handled by canonical orthogonalization. The treatment is found to remove non-physical kinks observed earlier on potential energy surfaces. Sensitivity analysis of the new approach confirms the elimination of the drastic increase in singular values of sensitivity matrices, reported earlier.

I. INTRODUCTION

Molecular systems often exhibit multireference character, which persists to be a challenge for theoreticians to describe. Systems of fundamental chemical interest may belong to this class, like stretched covalent bonds, transition metal compounds with open shells or excited electronic states of molecules. From the methodological point of view, the difficulty lies with the fact, that a single determinant does not give a qualitatively correct wavefunction even at the zero order. Constructing a reasonably good multideterminantal wavefunction may not represent a problem in itself. It is however not straightforward to describe correlation effects, based on such a reference function. Development of standard electron correlation approaches for the multireference (MR) case continues to be an active field. For surveys on techniques in MR configuration interaction (CI), MR coupled-cluster (CC) and MR perturbation theory (PT) see e.g. Ref.'s¹⁻³, respectively.

There exists an especially great variety of multireference perturbation theories (MRPT's). To help orientation, it is usual to categorize MRPT techniques as "diagonalize then perturb" or "perturb then diagonalize". Our present study focuses on an approach belonging to the latter class. Methods of this type rely of an effective Hamiltonian, often but not exclusively based on the Bloch-equation⁴. Effective Hamiltonian theories have the potential to target more than one state at a time, though intruder effect may hamper the calculation of numerous states simultaneously. One solution to the intruder problem is to target less states than the dimension of the model space, leading to intermediate Hamiltonian theories⁵. The extreme case of a single target state was considered by Mukherjee and co-workers when developing state-specific (SS) MRPT⁶⁻⁸. Besides manifest intruder avoidance, an attractive feature of SS-MRPT is the extensive nature of its energy corrections. The method of SS-MRPT has been successfully applied for determining potential energy surfaces and spectroscopical data⁹⁻¹¹. It has also been demonstrated, that SS-MRPT gives competitive results, compared to various alternative MRPT methods^{12,13}.

In addition to extensivity, the question of spin-symmetry breaking may come as a further problematic point of multireference correlation methods. Following the pioneering studies on spin symmetry conservation within the coupled-cluster methodology¹⁴⁻¹⁶ two paradigms appear to be persistent in MR CC theories: i) application of spin-free unitary generators for expressing the cluster-operator¹⁷⁻²¹, or ii) consideration of the eigenvalue equation of

operator S^2 in addition to that of $H^{22,23}$. The first approach has the drawback of taking the exponential of non-commuting operators, resulting in contractions among different terms of the cluster-operator. In contrast to this, spin-adaptation based on S^2 -equations works with "pure" excitation-operators, but it relies on a single determinantal reference state.

In the context of the parent CC theory of SS-MRPT, called SS-MRCC, the unitary group approach (UGA) has been used in various manners to arrive at a spin-pure formulation^{24–26}. First attempts in this line used exponentials of normal ordered spin-free operators²⁴. More recently, normal-ordering of the entire exponential, $\{\exp(T^\mu)\}$ has been advocated to avoid the long expansion including contractions among various cluster-operators²⁶.

The UGA type spin-adaptation of SS-MRPT has been reported and applied in several studies^{13,24}. Mao and co-workers suggested a partitioning which suppresses dependence of the results on the spin-coupling scheme²⁷. Explicit electron correlation correction to spin-adapted SS-MRPT been also implemented and tested recently²⁸.

In spite of the manifest intruder-free character of SS-MRPT, unphysical kinks have been observed on the order of 1-10 mE_h when computing potential energy surfaces. Similar experience has been reported in the parent coupled-cluster theory²⁹ as well as in a related multireference CC theory, developed by Hanrath³⁰. The unexpected behavior has been attributed to small coefficients in the CAS wavefunction, which appear in the denominator of the amplitude equations. Later, Hanrath hinted that reference specific excitation spaces used for projection of the Schrödinger equation in SS-MRCC may be responsible for the numerical inaccuracy³¹. To circumvent the problem a damping procedure has been applied both in SS-MRCC²⁹ and SS-MRPT^{11,32}. Our study of the problem has revealed, that the unphysical kinks are accompanied by orders of magnitude increase in the coefficient sensitivity of SS-MRPT energies and relaxed coefficients³². It could however provide no explanation on the source of the phenomenon.

In the present work we study a redundancy among virtual functions, which arises due to the application of spin-free unitary generators. An orthogonalization procedure is introduced to remove the redundancy stemming from spin-adaptation. Numerical applications of the theory demonstrate that the problem of kinks gets cured by screening redundancy. Sensitivity analysis confirms the finding, showing no notable increase in sensitivities in the previously problematic regions of the potential energy surface.

II. THEORY

1. SS-MRPT

Essential features of SS-MRPT are briefly summarized below, focusing on spin-adaptation applying unitary group generators. For detailed derivation we refer to previous works on the determinantal version^{6,8,33–35} and the unitary group approach (UGA) based spin-adaptation of the theory^{24,26,36}.

State-specific MRPT relies on a model space, the corresponding projector being

$$P = \sum_{\mu} |\phi_{\mu}\rangle \langle \phi_{\mu}| .$$

Functions ϕ_{μ} are presently considered to form a complete active space (CAS). This restriction is not necessary to impose: incomplete model space formulation of SS-MRCC has been also worked out³⁷.

The initial target function expanded in terms of model-space functions is written as

$$|\Phi\rangle = \sum_{\mu} |\phi_{\mu}\rangle c_{\mu}^{(0)} , \quad (1)$$

the associated CAS energy denoted by E_{CAS} . Parametrization of the target function

$$|\Psi\rangle = \sum_{\mu} \exp(T^{\mu}) |\phi_{\mu}\rangle c_{\mu} , \quad (2)$$

involves a separate cluster-operator T^{μ} , associated to each model-space function³⁸.

Equations determining the amplitudes in T^{μ} are derived from the Schrödinger-equation, assuming a set of redundancy conditions characteristic for SS-MRCC. Upon linearization, the amplitude equations become⁸

$$\langle \chi_l | H | \phi_{\mu} \rangle c_{\mu}^{(0)} + \langle \chi_l | [H, T^{\mu}] | \phi_{\mu} \rangle c_{\mu}^{(0)} + \sum_{\nu} \langle \chi_l | T^{\nu} - T^{\mu} | \phi_{\mu} \rangle H_{\mu\nu} c_{\nu}^{(0)} = 0 . \quad (3)$$

Functions χ_l are elements of the outer space Q , complementary to P

$$Q = 1 - P = \sum_l |\chi_l\rangle \langle \chi_l| . \quad (4)$$

Since Eq.(3) holds for each μ and l , the number of equations is appropriate if using as many outer-space functions χ_l for a given ϕ_{μ} , as the number of parameters in T^{μ} .

Invoking a perturbative splitting of the Hamiltonian

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

with the zero-order operator diagonal on the basis of functions ϕ_μ and χ_l

$$H^{(0)} = \sum_{\mu} E_{\mu}^{(0)} |\phi_{\mu}\rangle \langle \phi_{\mu}| + \sum_l E_l^{(0)} |\chi_l\rangle \langle \chi_l| , \quad (5)$$

equations determining the amplitudes at the first order take the form⁸

$$\sum_{\nu} \left[H_{\mu\nu} + \left(E_l^{(0)} - E_{\mu}^{(0)} - E_{\text{CAS}} \right) \delta_{\mu\nu} \right] c_{\nu}^{(0)} t_{l\mu}^{\nu} = -H_{l\mu} c_{\mu}^{(0)} , \quad (6)$$

with $H_{\mu\nu} = \langle \phi_{\mu} | H | \phi_{\nu} \rangle$, $H_{l\mu} = \langle \chi_l | H | \phi_{\mu} \rangle$ and $t_{l\mu}^{\nu} = \langle \chi_l | T^{\nu} | \phi_{\mu} \rangle$.

Having determined the amplitudes, the non-symmetric effective Hamiltonian is constructed up to order two as

$$H_{\nu\mu}^{[2]} = \langle \phi_{\nu} | H + (HT^{\mu})_c | \phi_{\mu} \rangle .$$

The eigenvalue-equation of $H_{\nu\mu}^{[2]}$ yields the energy and the relaxed coefficients of SS-MRPT at order two:

$$\sum_{\mu} H_{\nu\mu}^{[2]} c_{\mu} = E^{[2]} c_{\nu} .$$

2. Spin-adaptation

When aiming a correct treatment of spin-symmetry, it is practical to consider functions ϕ_{μ} as configuration state functions (CSF). Conserving spin quantum number of the reference functions is possible if expressing the cluster operator in terms of generators of the unitary group

$$E_i^a = a_{a\alpha}^+ a_{i\alpha} + a_{a\beta}^+ a_{i\beta} .$$

Such a parametrization – ensuring orthogonality of $T^\mu \phi_\mu$ to the model space – can be given as

$$\begin{aligned}
T^\mu &= \sum_{ia} t_i^a \{E_i^a\}_c + \sum_{iu_s} t_i^{u_s} \{E_i^{u_s}\}_c + \sum_{au_s} t_{u_s}^a \{E_{u_s}^a\}_c \\
&+ \frac{1}{2} \sum_{ij} \sum_{ab} (1 + \delta_{ij} \delta_{ab}) t_{ij}^{ab} \{E_{ij}^{ab}\}_c + \sum_{ij} \sum_{au_s} t_{ij}^{au_s} \{E_{ij}^{au_s}\}_c + \sum_{iu_s} \sum_{ab} t_{iu_s}^{ab} \{E_{iu_s}^{ab}\}_c \\
&+ \frac{1}{2} \sum_{ij} \sum_{u_s \neq v_s} t_{ij}^{u_s v_s} \{E_{ij}^{u_s v_s}\}_c + \frac{1}{2} \sum_{\substack{u_s v_s \\ u_s \neq v_s}} \sum_{ab} t_{u_s v_s}^{ab} \{E_{u_s v_s}^{ab}\}_c \\
&+ \sum_{iu_s} \sum_{av_s} t_{iu_s}^{av_s} \{E_{iu_s}^{av_s}\}_c + \sum_{iu_s} \sum_{av_s} t_{iu_s}^{v_s a} \{E_{iu_s}^{v_s a}\}_c \\
&+ \sum_{iu_s} \sum_{v_s \neq w_s} t_{iu_s}^{v_s w_s} \{E_{iu_s}^{v_s w_s}\}_c + \sum_{\substack{u_s v_s \\ u_s \neq v_s}} \sum_{aw_s} t_{u_s v_s}^{aw_s} \{E_{u_s v_s}^{aw_s}\}_c , \tag{7}
\end{aligned}$$

with $\{E_{ij}^{ab}\}_c = \{E_i^a E_j^b\}_c$. The above expression of T^μ is essentially in agreement with Ref.'s^{24,27}, with the exception that normal ordering with respect to the common core of all model space functions, ϕ_c is used, denoted by $\{\cdot\}_c$. As it was argued before²⁷, in case of the operators appearing in Eq.(7), normal ordering with respect to ϕ_c is the same as normal ordering with respect to the doubly occupied part of ϕ_μ . Indices i, j, \dots stand for orbitals doubly occupied in ϕ_μ (i.e. they can be either holes or particles as defined by ϕ_c). Labels a, b, \dots denote orbitals unoccupied in ϕ_μ (either inactive or active). Active indices are labeled by u, v, \dots . Occupancy of active indices is indicated in subscript when needed: u_d refers to doubly occupied, u_s to singly occupied and u_z to an unoccupied active orbital. Note the dichotomy of the notation for active indices: a doubly occupied active may be labeled by i or by u_d , an unoccupied active may be denoted by either u_z or by a . Inactive particles, also called virtuals, are labeled by a, b, \dots only. When needed, indices p, q, \dots are used to refer to generic orbitals.

For open shell CSF's, a matching between upper and lower indices of unitary generators can occur for orbitals singly occupied in ϕ_μ . Upper-lower matching affects the so-called spectator excitations, where active labels appear both as creation and annihilation indices. Any matching between upper and lower indices are excluded among single excitations but they can appear among doubles in T^μ . When matching orbitals are necessarily of the same spin, we speak of direct spectator excitations, these are of the form $\{E_{iu_s}^{u_s}\}_c$. In contrast to

this, excitations of the type $\{E_{u_s}^{u_s}\}_c$ are called exchange spectators.

Excitation types admissible in T^μ are inherently connected with the question of redundancy. State-specific theories relying on a Jeziorski-Monkhorst type parametrization³⁸ encounter a redundancy stemming from the fact that a given excited function χ_l is generated by T^μ acting on ϕ_μ as well as by T^ν acting on ϕ_ν , $\mu \neq \nu$. Redundancy of this sort is handled by introducing sufficiency conditions³³, which are exploited in Eq.(3), the linearized form of the SS-MRCC amplitude equations. In spin-orbital based formulations of SS-MRPT, Eq.(6) written for all μ, l , determines all the amplitudes considered.

In a spin-adapted treatment further complications arise from the fact that acting on ϕ_μ with the spin-free operators in Eq.(7) of the cluster operator, a linearly dependent set is generated, if ϕ_μ is open-shell. This additional redundancy has been taken care of in different manners in various spin-adapted MRCC theories. Paldus and coworkers have been using spin functions as χ_l , constructed according to the Gel'fand coupling scheme^{18,19}. Sen et al. also considered linearly independent spin-adapted functions in their UGA-based state-universal MRCC³⁹. In the context of state-specific MRCC, Mukherjee and coworkers choose to take the overlapping and redundant set of functions χ_l , as resulted by acting with unitary generators on model space functions²⁴. When following the latter approach, the use of additional sufficiency conditions, analogous those of the the spin-orbital based theory, have been reported²⁵⁻²⁷. These are necessary, since the number of amplitudes in a given truncation scheme of T^μ exceeds the dimension of the space spanned by functions χ_l , generated from ϕ_μ .

To give a specific example, consider a case of two active electrons and take ϕ_μ as an open-shell singlet

$$|\phi_\mu\rangle = \frac{1}{\sqrt{2}} (v_\beta^+ u_\alpha^+ + u_\beta^+ v_\alpha^+) |\phi_c\rangle, \quad u \neq v. \quad (8)$$

Let us examine two double excitations, $\{E_{iu_s}^{v_s a}\}_c$ and $\{E_{iu_s}^{a v_s}\}_c$, acting on ϕ_μ :

$$\begin{aligned} \frac{1}{2} \{E_{iu_s}^{a v_s}\}_c |\phi_\mu\rangle &= |\chi_{iu_s}^{a v_s}\rangle \\ \{E_{iu_s}^{v_s a}\}_c |\phi_\mu\rangle &= |\chi_{iu_s}^{v_s a}\rangle, \end{aligned}$$

with factor 1/2 introduced for normalization. It is easy to show, that the above functions differ only in sign

$$\chi_{iu_s}^{a v_s} = - \chi_{iu_s}^{v_s a}.$$

Taking $E_l^{(0)} = \langle \chi_l | H | \chi_l \rangle$ for zero order excited energies (c.f. Epstein-Nesbet (EN) partitioning in PT), it is clear from Eq.(6) that the equations obtained by projection with $\chi_{iu_s}^{av_s}$ and $\chi_{iu_s}^{v_s a}$ are the same. This would prohibit the inversion of the coefficient matrix of Eq.(6), if written for all μ and l indices. The amplitudes of SS-MRPT are however not obtained by direct inversion of the coefficient matrix of Eq.(6). They are decoupled instead, c.f. Ref.²⁴ or the discussion in Section II.4. For each excitation pattern – in our example for $\{E_{iu_s}^{v_s a}\}_c$ and $\{E_{iu_s}^{av_s}\}_c$ – separate equations are solved, which are of the dimension of the model space. In a spin-orbital based approach, where no overlap occurs among various χ_l 's generated from the same ϕ_μ , the decoupling follows from the structure of Eq.(6) and does not represent any approximation⁸. In contrast to this, coupling does occur in Eq.(6) in a spin-adapted formulation, affecting excitations which generate overlapping functions χ_l , when acting on the same ϕ_μ . Neglecting this coupling solves the additional redundancy problem of spin-adapted theory, since the number of equations gets augmented, to match the number of unknowns.

When dealing with a spin-adapted multireference formulation, one has to be aware of an even further complexity: spin-functions of a given open-shell structure may correspond to non-uniform excitation levels when generated from a particular, open-shell ϕ_μ . To give an example again, consider excitation $\{E_{ii}^{ab}\}_c$, $a \neq b$, acting on ϕ_μ of Eq.(8). The resulting function

$$\frac{1}{\sqrt{2}}\{E_{ii}^{ab}\}_c|\phi_\mu\rangle = |\chi_{ii}^{ab}\rangle$$

has four singly occupied orbitals: u, v, a, b . With four open shells, there exists a second singlet function, related to ϕ_μ by a triple excitation, according to¹⁸

$$\{E_{i u_s i}^{u_s a b}\}_c|\phi_\mu\rangle \sim |\chi_{i u_s i}^{u_s a b}\rangle.$$

If truncating T^μ based on excitation level, χ_{ii}^{ab} is used for projecting the equations, while $\chi_{i u_s i}^{u_s a b}$ not. Since these functions mix, when rotating e.g. active orbitals, it is necessary to consider both of them, if unitary invariance is to be kept¹⁸. In any approximate theory it is a matter of decision, whether to include all spin components of open-shell excited functions, or not. In the present approach we wish to investigate the role of redundancy, and keep all other aspects of the theory unchanged. For this reason we conserve the treatment of our previous studies^{13,32}, and stick with truncation of T^μ based on excitation level. We note in passing, that choice Eq.(5) of the zero-order Hamiltonian – i.e. diagonal on the basis applied – prohibits invariance of SS-MRPT to a unitary transformation anyhow.

3. Removal of redundancy in T^μ

Redundancy in T^μ stemming from spin-adaptation has been so far treated by applying additional sufficiency conditions in SS-MRPT^{24,27,32}. Here we examine a variant of spin-adapted SS-MRPT, using orthogonal excited functions. The basic idea is to identify overlapping sets of χ_l -s obtained by acting with the unitary generators in T^μ on ϕ_μ . The overlapping functions are orthogonalized by the canonical scheme⁴⁰. Eigenvectors corresponding to zero eigenvalues of the overlap matrix are dropped to eliminate redundancy. The treatment is similar in spirit to the orthogonalization applied in internally contracted MR approaches, both in the framework of PT^{41,42} and CC⁴³⁻⁴⁶. A notable difference with internally contracted theories is that overlap arises only due to spin-adaptation in the present case. This means that the open-shell structure of nonorthogonal functions must be the same, overlapping sets are consequently rather small. In the examples studied here they are three dimensional at most, lifting the need to perform any numerical procedure. In parallel with filtering redundancy, the number of amplitudes is reduced to match the number orthogonal virtual functions. The orthogonalization procedure yields a modified cluster operator

$$\tilde{T}^\mu = \sum_{\mathcal{I}, \mathcal{A}} \sum_g \tilde{t}_{\mathcal{I}}^{\mathcal{A}} \{g \tilde{E}_{\mathcal{I}}^{\mathcal{A}}\}_c, \quad (9)$$

where upper case calligraphy letters $\mathcal{I}, \mathcal{J}, \mathcal{A}, \mathcal{B}$ may be composite indices: they refer either to a single orbital, or to an orbital pair. The new cluster operator, \tilde{T}^μ is used for writing amplitude equations in the next Section.

The very first step in eliminating overcompleteness was taken by Mukherjee and coworkers^{24,27,47}, when combining direct spectator excitations with the corresponding single excitations in T^μ . This is an obvious redundancy screening, since $\{E_{u_s}^{u_s}\}_c$ acts as an operator of unity on ϕ_μ , where u_s is singly occupied. For this reason the result of a double excitation, involving direct spectators, matches the effect of the corresponding single substitution, spectator indices omitted.

In the present approach we choose to omit direct spectator excitations entirely. Even after excluding direct spectators, there remain redundancies among spin-free operator generated excited functions. Overlapping sets of excited functions and their orthogonalized counterparts are presented in Table I for single excitations and in Table III for doubles. Excitation operators appearing in \tilde{T}^μ , defined based on the orthogonalized functions are collected in

the same Tables. Relation between amplitudes in Eq.(7) and those appearing in Eq.(9). A detailed derivation of the results collected in Tables I-IV is given in the Supplementary Material⁴⁸.

When categorizing excitations, designations core, docc, active, active(1), empty and virtual are used. Terminology core refers to orbitals doubly occupied in all model functions, active designates orbitals with variable occupancy in the model functions, while virtual refers to orbitals unoccupied in any model function. Classifications docc, active(1) and empty are defined with respect to a particular model function, and stand for a doubly occupied, singly occupied or unoccupied orbital, respectively.

For the sake of simplicity we restrict the analysis to two active electrons. In such a situation singlet reference functions are either closed-shell determinants, or two-determinantal open-shell functions. Tables III and IV do not include the case of 2 docc→2 empty excitation performed on a closed-shell reference functions. (As discussed in the Supplementary Material⁴⁸, this situation is completely parallel with the case of 2 core→2 empty excitation exercised upon an open-shell reference function.) Apart from this case, and those considered in Tables I and III, parametrization of T^μ agrees with that of \tilde{T}^μ , with the difference that direct spectators are excluded from the latter.

We wish to stress here that from the purely theoretical point of view, the use of orthogonal functions is by no means superior to overlapping functions. If invariance to a linear transformation of the basis holds, the two approaches produce the same result. When dealing with an overlapping basis however, care has to be taken to achieve a proper – possibly equivalent – formulation. Deriving equations of SS-MRPT, orthogonality of the virtual functions was assumed in the form of projector Q (c.f. Eq.(4)) and the zero-order Hamiltonian (c.f. Eq.(5)). Formulation of the theory with overlapping functions χ_l , requires either the overlapping form of the above quantities, e.g.

$$Q = \sum_{lm} |\chi_l\rangle (S^{-1})_{lm} \langle \chi_m| , \quad (10)$$

or orthogonalization of functions χ_l (with $S_{lm} = \langle \chi_l | \chi_m \rangle$ giving the elements of the overlap matrix). It is the latter route we explore in this study.

Use of orthogonal or overlapping functions have their specific advantages as well as shortcomings. Orthogonalization of functions χ_l brings a considerable complication in the present study, which discouraged us to step beyond two active electrons. Keeping the overlapping

functions and introducing their overlap in the equations may offer a simpler alternative. Overcompleteness is still to be removed if relying on the spectral form of projector Q , since \mathbf{S}^{-1} is singular otherwise.

A known caveat in connection with orthogonalization is the possibility for spoiling size-extensivity^{45,49,50}. To examine this property, one may start from the original, spin-adapted formulation of the theory, which is size-extensive²⁴. Consider now a separation of the total system for two non-interacting parts, characterized by $H = H_A + H_B$. Assume that orbitals are localized to subsystems, e.g. i_A is localized to system A and j_B to system B. In this situation, amplitudes $t(\mu)$ may also be given a subsystem label A or B. Amplitude assignment is based on the Hamiltonian which appears in the interaction matrix element, e.g. $t_{i_A}^{a_A}(\mu)$ is given label A, since only H_A figures in

$$\langle \phi_\mu | \{E_{i_A}^{a_A}\}_c^\dagger (H_A + H_B) | \phi_\mu \rangle = \langle \phi_\mu | \{E_{i_A}^{a_A}\}_c^\dagger H_A | \phi_\mu \rangle .$$

Size-extensivity is conserved upon orthogonalization if the quantities mixed by the transformation are assigned the same subsystem label. As Tables Eq.(I) and Eq.(III) reflect, this holds true for every overlapping subspace affected by the present transformation. The case of double excitations is trivial, since unitary generators combined to form $\{\tilde{E}_{\mathcal{I}}^A\}_c$ have the very same indices. Regarding single excitations, $\{E_{i_A u_B}^{u_B a_A}\}_c$ may mix with $\{E_{i_A}^{a_A}\}_c$. This however presents no problem, since the corresponding interaction matrix element involves only H_A :

$$\langle \phi_\mu | \{E_{i_A u_B}^{u_B a_A}\}_c^\dagger (H_A + H_B) | \phi_\mu \rangle = \langle \phi_\mu | \{E_{i_A u_B}^{u_B a_A}\}_c^\dagger H_A | \phi_\mu \rangle ,$$

hence $t_{i_A u_B}^{u_B a_A}$ is assigned to system A. Separability of the theory is therefore expected with two active electrons.

Overlapping spaces may get larger when increasing the number of active electrons. Size-extensivity of the theory is expected to still hold, since the overlapping functions exhibit the same open-shell structure. Upon separation, this is expected to lead to a subsystem assignment uniform for the overlapping set.

4. Coupling of amplitudes within \tilde{T}^μ

A compact form of amplitude equations (6) can be written in the spin-adapted case as

$$\sum_{\nu} \sum_{\mathcal{I}, \mathcal{B}} H_{\mu\nu} c_{\nu}^{(0)} C(\mu, \mathcal{I} \rightarrow \mathcal{A}; \nu, \mathcal{J} \rightarrow \mathcal{B}) t_{\mathcal{J}}^{\mathcal{B}}(\nu) = - \langle \chi_{\mathcal{I}}^{\mathcal{A}}(\mu) | H | \phi_{\mu} \rangle c_{\mu}^{(0)}. \quad (11)$$

Connection with Eq.(6) is set by recognizing, that

$$\langle \chi_{\mathcal{I}}^{\mathcal{A}}(\mu) | = \frac{1}{\mathcal{N}_{\mathcal{I}}^{\mathcal{A}}(\mu)} \langle \phi_{\mu} | \{ E_{\mathcal{I}}^{\mathcal{A}} \}_c^\dagger$$

stands for $\langle \chi_l |$ of Eq.(6), factor $\mathcal{N}_{\mathcal{I}}^{\mathcal{A}}(\mu)$ ensuring normalization. For $\mu \neq \nu$, amplitude t_{μ}^{ν} of Eq.(6) is replaced by

$$\langle \chi_{\mathcal{I}}^{\mathcal{A}}(\mu) | T^{\nu} | \phi_{\mu} \rangle = \sum_{\mathcal{J}, \mathcal{B}} C(\mu, \mathcal{I} \rightarrow \mathcal{A}; \nu, \mathcal{J} \rightarrow \mathcal{B}) t_{\mathcal{J}}^{\mathcal{B}}(\nu), \quad \mu \neq \nu$$

in Eq.(11). For $\mu = \nu$, the product $C(\mu, \mathcal{I} \rightarrow \mathcal{A}; \mu, \mathcal{I} \rightarrow \mathcal{A}) t_{\mathcal{I}}^{\mathcal{A}}(\nu)$ in Eq.(11) incorporates term $(E_l^{(0)} - E_{\mu}^{(0)} - E_{\text{CAS}}) t_{l_{\mu}}^{\mu}$ on the left hand side of Eq.(6) as well. Reference index ν of the amplitude is written in round braces to avoid confusion with excitation indices. Index μ of the reference function is also indicated in round braces in the virtual function, e.g. $\chi_{\mathcal{I}}^{\mathcal{A}}(\mu)$, when its parentage is not obvious.

The additional conditions applied for the redundancy stemming from spin adaptation, can now be given expression-wise, by the following decoupled form of matrix C :

$$C(\mu, \mathcal{I} \rightarrow \mathcal{A}; \nu, \mathcal{J} \rightarrow \mathcal{B}) = \delta_{\mathcal{I}\mathcal{J}} \delta_{\mathcal{A}\mathcal{B}} \mathcal{N}_{\mathcal{I}}^{\mathcal{A}}(\mu) (1 + \delta_{\mu\nu} (X_{\mathcal{I}}^{\mathcal{A}}(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}), \quad (12)$$

with the shorthand

$$X_{\mathcal{I}}^{\mathcal{A}}(\mu) = \langle \chi_{\mathcal{I}}^{\mathcal{A}}(\mu) | H^{(0)} | \chi_{\mathcal{I}}^{\mathcal{A}}(\mu) \rangle - \langle \phi_{\mu} | H^{(0)} | \phi_{\mu} \rangle \quad (13)$$

for zero order excitation energy. Note, that $X_{\mathcal{I}}^{\mathcal{A}}(\mu)$ is shifted by $H_{\mu\mu} - E_{\text{CAS}}$ in the energy denominator of SS-MRPT, a reason why the theory is manifestly robust against intruders⁸.

Inspecting matrix C of Eq.(12), factors $\delta_{\mathcal{I}\mathcal{J}} \delta_{\mathcal{A}\mathcal{B}}$ apparently exclude any coupling between excitation patterns, Eq.(11) can therefore be solved for each $\mathcal{I} \rightarrow \mathcal{A}$ pattern separately.

Turning now to the equations determining amplitudes of cluster operator \tilde{T}^μ , the compact form of Eq.(6) becomes

$$\sum_{\nu} \sum_{\mathcal{J}, \mathcal{B}} \sum_g H_{\mu\nu} c_{\nu}^{(0)} \tilde{C}(\mu, \mathcal{I} \rightarrow \mathcal{A}, f; \nu, \mathcal{J} \rightarrow \mathcal{B}, g) {}_g \tilde{t}_{\mathcal{J}}^{\mathcal{B}}(\nu) = - \langle {}_f \tilde{\chi}_{\mathcal{I}}^{\mathcal{A}}(\mu) | H | \phi_{\mu} \rangle c_{\mu}^{(0)} \quad (14)$$

Compared to Eq.(11), an additional sum is introduced for lower left index g , since

$$\langle {}_f\tilde{\chi}_{\mathcal{I}}^A(\mu)|\tilde{T}^\nu|\phi_\mu\rangle = \sum_{\mathcal{J},\mathcal{B}} \sum_g \tilde{C}(\mu, \mathcal{I} \rightarrow \mathcal{A}, f; \nu, \mathcal{J} \rightarrow \mathcal{B}, g) {}_g\tilde{t}_{\mathcal{J}}^{\mathcal{B}}(\nu), \quad \mu \neq \nu \quad (15)$$

holds in this case. Indices f and g can take values 1 and 2, when the case requires (c.f. Tables I, III). Due to the orthogonality of functions ${}_f\tilde{\chi}_{\mathcal{I}}^A(\mu)$, we can allow couplings to appear in Eq.(14). Coupling affects only those excitations which are combined when orthogonalizing virtual functions, consequently the factor $\delta_{\mathcal{I}\mathcal{J}}\delta_{\mathcal{A}\mathcal{B}}$ appears in \tilde{C} as well. Equation (14) can therefore also be solved separately for excitation patterns $\mathcal{I} \rightarrow \mathcal{A}$, but not separately for indices f .

It is a rather tedious task to derive the matrix \tilde{C} , since its complexity increases with the number of active electrons. Matrix elements for two active electrons are collected in Tables V, VI and VII for single excitations. Double excitations are uncoupled in index f in most of the cases. The corresponding values for \tilde{C} are collected in Table VIII. Couplings among doubles' amplitudes occur for two excitation types: core, active \rightarrow 2 active and core, active \rightarrow active, virtual. Matrix form of \tilde{C} for these excitations is shown in Table IX. Type core, active \rightarrow active, virtual is not tabulated, as it can be derived from matrix \tilde{C} of the core, active \rightarrow 2 active case, by substituting index a for v_z , and omitting the last row and column.

Orthogonalization results in a slight increase in computational cost as compared to the decoupled variant of the redundant, spin-adapted formulation. Based on Eq.(11), the latter shows a rough $(N_{\text{core}} + N_{\text{active}})^2(N_{\text{active}} + N_{\text{virtual}})^2 N_{\text{CAS}}^2$ scaling, while the sum for g introduced in Eq.(14) brings a factor of two. This scaling is obviously more favourable than the requirement of the parent CC theory, nonlinear terms bringing a considerable increase in computational time.

5. Construction of the effective Hamiltonian

Once an amplitude $\tilde{t}_{l\mu}^\mu = \langle \tilde{\chi}_l(\mu)|\tilde{T}^\mu|\phi_\mu\rangle$ is determined, its contribution to the effective Hamiltonian can be readily calculated by multiplication with the matrix element $\langle \phi_\nu|H|\tilde{\chi}_l(\mu)\rangle$, involving Hamiltonian integrals. Accumulating these products yields the ef-

fective Hamiltonian

$$\begin{aligned} H_{\nu\mu}^{[2]} &= H_{\nu\mu} + \langle \phi_\nu | H \tilde{T}^\mu | \phi_\mu \rangle \\ &= H_{\nu\mu} + \sum_l \langle \phi_\nu | H | \tilde{\chi}_l(\mu) \rangle \tilde{t}_{l\mu}^\mu . \end{aligned} \quad (16)$$

In Eq.(16) above, the projection operator Q in the form

$$Q = \sum_l |\tilde{\chi}_l(\mu)\rangle \langle \tilde{\chi}_l(\mu)| ,$$

was inserted.

When working with amplitudes $t_{\mathcal{I}}^A(\mu)$ of T^μ , instead of the non-redundant $\tilde{t}_{l\mu}^\mu$ parameters of \tilde{T}^μ , it is practical to express the effective Hamiltonian as

$$H_{\nu\mu}^{[2]} = H_{\nu\mu} + \sum_{\mathcal{IA}} \langle \phi_\nu | H | \chi_{\mathcal{I}}^A(\mu) \rangle t_{\mathcal{I}}^A(\mu) \mathcal{N}_{\mathcal{I}}^A(\mu) . \quad (17)$$

Note, that expression of $H_{\nu\mu}^{[2]}$ in Eq.(17) is not equivalent to inserting operator Q' between H and T^μ in the form $Q' = \sum_\mu |\chi_l(\mu)\rangle \langle \chi_l(\mu)|$. There are two reasons for this: i) Q' as written here is not a proper projector, ii) $t_{\mathcal{I}}^A(\mu)$ does not arise as the matrix element $\langle \chi_{\mathcal{I}}^A(\mu) | T^\mu | \phi_\mu \rangle$.

In the applications of Section III, Eq.(17) is used for building the effective Hamiltonian. When using Eq.(17) with the orthogonalized parametrization of \tilde{T}^μ , amplitudes \tilde{t} are first converted to amplitudes t , governed by the formulae of the Section II.3. Alternatively, amplitudes \tilde{t} of the orthonormal parametrization could be readily used for constructing the effective Hamiltonian of Eq.(16). The two approaches lead to the same result.

6. Sensitivity analysis

The main motivation of this work is to explore the origin of kinks which occasionally appear on potential energy surfaces computed by spin-adapted SS-MRPT^{11,13}. A diagnostic tool for this purpose is offered by sensitivity analysis, a technique for characterizing the change in the solution of a mathematical model when an input parameter is altered. Sensitivity analysis finds widespread applications when solving systems of differential equations by numerical means, e.g. in reaction kinetics⁵¹. Sensitivities are presently used for comparing the nonredundant parametrization of the cluster operator with the redundant variant.

The essence of sensitivity analysis as applied to SS-MRPT is briefly recapitulated here. We refer to a previous report³² for more detailed presentation.

The key quantities of coefficient sensitivity of SS-MRPT are derivatives of the energy, $\partial E^{[2]}/\partial c_\mu^{(0)}$, and the final coefficients, $\partial c_\mu/\partial c_\mu^{(0)}$. The related sensitivity matrices are built of

$$\mathcal{S}_\mu = \frac{c_\mu^{(0)}}{E^{[2]}} \frac{\partial E^{[2]}}{\partial c_\mu^{(0)}}, \quad (18)$$

and

$$\mathcal{S}_{\mu\nu} = \frac{c_\nu^{(0)}}{c_\mu} \frac{\partial c_\mu}{\partial c_\nu^{(0)}}, \quad (19)$$

respectively. Sensitivity matrices are subjected to singular value decomposition (SVD), allowing to write e.g. $\mathcal{S}_{\mu\nu}$ as

$$\mathcal{S}_{\mu\nu} = \sum_i U_{\mu i} \sigma_i V_{\nu i},$$

where $V_{\nu i}$ and $U_{\mu i}$ are components of the right and left singular vector, corresponding to singular value σ_i .

When altering initial coefficients $c_\nu^{(0)}$ by $\Delta c_\nu^{(0)}$, the relative change of the quantity studied – e.g. relaxed coefficients, c_μ – can be expressed as the squared sum of relative deviations

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

Using the singular values and singular vectors, function e takes the simple, so called canonical form:

$$e = \sum_i \sigma_i^2 |\delta_i|^2 + \mathcal{O}(3), \quad (20)$$

with

$$\delta_i = \sum_\nu \frac{\Delta c_\nu^{(0)}}{c_\nu^{(0)}} V_{\nu i}.$$

Singular values σ_i are called sensitivities. If one, or a few sensitivities are exceedingly large compared to the others, they give the dominant term(s) of function e , and point to the initial coefficients which have the largest influence on \mathbf{c} , when changed. The corresponding

column of matrix \mathbf{V} reveals what combination of relative changes in $c_\nu^{(0)}$'s is the most effective in achieving a large relaxation effect in \mathbf{c} .

When studying coefficient sensitivity of the SS-MRPT energy, the corresponding sensitivity matrix is a row vector, leading to the only nonzero singular value $\sigma = (\sum_\mu \mathcal{S}_\mu^2)^{1/2}$. There is no point in examining the right singular vector in this case, since the first column of \mathbf{V} is uniformly built of numbers 1, apart from normalization.

To express elements of sensitivity matrices, derivatives of the working equations are to be taken. Here we compute sensitivities based either on the amplitude equations of the redundant parametrization of T^μ , c.f. Eqs.(11)-(12), or the orthogonalized parametrization of \tilde{T}^μ , c.f. Eq.(14). The effective Hamiltonian is built according to Eq.(17) in both cases. Sensitivities based on the effective Hamiltonian of Eq.(16) are not examined in this work.

There are three basic quantities necessary for constructing sensitivity matrices:

- i) derivative of the amplitudes with respect to the initial coefficients;

Regarding T^μ , derivatives of amplitudes $t_{\mathcal{I}}^A(\mu)$ are readily expressed from Eqs.(11)-(12) as¹

$$\frac{\partial t_{\mathcal{I}}^A(\mu)}{\partial c_\nu^{(0)}} = - \frac{1}{c_\mu^{(0)}} \left(\delta_{\mu\nu} t_{\mathcal{I}}^A(\mu) + A_{\mu\nu}^{-1}(\mathcal{I} \rightarrow \mathcal{A}) \langle \phi_\nu | \{E_{\mathcal{I}}^A\}_c^\dagger H | \phi_\nu \rangle \right) ,$$

with

$$A_{\mu\nu}(\mathcal{I} \rightarrow \mathcal{A}) = H_{\mu\nu} + \delta_{\mu\nu} (X_{\mathcal{I}}^A(\mu) - E_{\text{CAS}}) .$$

If considering the orthogonalized parametrization, \tilde{T}^μ , Eq.(14) leads to the derivative

$$\frac{\partial_f \tilde{t}_{\mathcal{I}}^A(\mu)}{\partial c_\nu^{(0)}} = - \frac{1}{c_\mu^{(0)}} \left(\delta_{\mu\nu} \tilde{t}_{\mathcal{I}}^A(\mu) + \sum_g A_{\mu f, \nu g}^{-1}(\mathcal{I} \rightarrow \mathcal{A}) \langle {}_g \tilde{\chi}_{\mathcal{I}}^A(\nu) | H | \phi_\nu \rangle \right) ,$$

with

$$A_{\mu f, \nu g}(\mathcal{I} \rightarrow \mathcal{A}) = H_{\mu\nu} \tilde{C}(\mu, \mathcal{I} \rightarrow \mathcal{A}, f; \nu, \mathcal{I} \rightarrow \mathcal{A}, g) .$$

- ii) derivative of the energy with respect to the amplitudes

$$\frac{\partial E^{[2]}}{\partial t_{\mathcal{I}}^A(\mu)} = \sum_\nu \tilde{c}_\nu^* \langle \phi_\nu | H | \chi_{\mathcal{I}}^A(\mu) \rangle c_\mu ;$$

¹ Note the typographical error in Eq.(19) of Ref.³².

iii) derivative of the relaxed coefficients with respect to the amplitudes

$$\frac{\partial c_\mu}{\partial t_{\mathcal{I}}^A(\mu)} = - \sum_{\nu\sigma\tau} L_{\mu\nu} G_{\nu\sigma} K_{\sigma\tau} \langle \phi_\tau | H | \chi_{\mathcal{I}}^A(\mu) \rangle c_\mu .$$

Notation \tilde{c}_ν^* stands for elements of the left eigenvector of the effective Hamiltonian, $\mathbf{H}^{[2]}$, corresponding to eigenvalue $E^{[2]}$. Elements of matrices \mathbf{L} and \mathbf{K} above are $L_{\mu\nu} = \delta_{\mu\nu} - c_\mu c_\nu^*$ and $K_{\sigma\tau} = \delta_{\sigma\tau} - c_\sigma \tilde{c}_\tau^*$. Quantities $G_{\nu\sigma}$ are elements of the reduced resolvent of the effective Hamiltonian, defined as

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

With the above ingredients, derivatives necessary for the sensitivity matrices are expressed via the chain rule, e.g.

$$\frac{\partial E^{[2]}}{\partial c_\nu^{(0)}} = \sum_{\mathcal{I},\mathcal{A}} \sum_{\mu} \frac{\partial E^{[2]}}{\partial t_{\mathcal{I}}^A(\mu)} \frac{\partial t_{\mathcal{I}}^A(\mu)}{\partial c_\nu^{(0)}}$$

if using amplitude equations of Eqs.(11)-(12). When amplitudes are determined from Eq.(14), derivative of $E^{[2]}$ is written as

$$\frac{\partial E^{[2]}}{\partial c_\nu^{(0)}} = \sum_{\mathcal{I},\mathcal{A}} \sum_{\mu} \sum_{\mathcal{J},\mathcal{B}} \sum_g \frac{\partial E^{[2]}}{\partial t_{\mathcal{I}}^A(\mu)} \frac{\partial t_{\mathcal{I}}^A(\mu)}{\partial_g \tilde{t}_{\mathcal{J}}^B(\mu)} \frac{\partial_g \tilde{t}_{\mathcal{J}}^B(\mu)}{\partial c_\nu^{(0)}} ,$$

with $\partial t_{\mathcal{I}}^A(\mu)/\partial_g \tilde{t}_{\mathcal{J}}^B(\mu)$ computed based on Tables II and IV. Analogous expressions apply for the derivatives of relaxed coefficients, c_μ .

III. NUMERICAL ILLUSTRATIONS

Two single bond dissociation processes are selected as test cases: that of the HF and the LiH molecule. The HF molecule is computed in Dunning's double zeta correlation consistent (cc-pVDZ) set⁵². The CAS reference function is generated by distributing two active electrons on two active orbitals, with symmetry labels $3a_1$ and $4a_1$, classified according to C_{2v} . The bond dissociation curve of the LiH molecule is computed in Dunning's double zeta plus polarization (DZP) set⁵³. The CAS wavefunction is constructed with two active electrons and five active orbitals of symmetry $2a_1$, $3a_1$, $4a_1$, $1b_1$, $1b_2$ in the C_{2v} molecular point group.

The basis sets being relatively small, full configuration interaction (FCI) results are feasible, and serve for computing errors.

Møller-Plesset (MP) and Epstein-Nesbet (EN) partitionings are applied within SS-MRPT. In the MP case the partitioning of the Hamiltonian depends on the reference CSF, in analogy with the multipartitioning applied in many-body PT schemes⁵⁴. Fockian matrix elements are constructed as

$$\begin{aligned}\tilde{f}_{pq}^-(\mu) &= f_{pq}^0(\mu) + \sum_{u_s \in \phi_\mu} [(pq|u_s u_s) - \delta_{qu_s}(pu_s|u_s u_s)] \\ \tilde{f}_{pq}^+(\mu) &= f_{pq}^0(\mu) + \sum_{u_s \in \phi_\mu} (pq|u_s u_s)\end{aligned}$$

adopting the (11|22) integral convention,

$$f_{pq}^0(\mu) = h_{pq} + \sum_{i \in \phi_\mu} [2(pq|ii) - (pi|i q)]$$

standing for matrix elements of the Fockian corresponding to the core of CSF ϕ_μ and h_{pq} being the one-electron integral of the kinetic energy operator plus nuclear-electron attraction. Note, that $\tilde{f}_{pq}^-(\mu)$ and $\tilde{f}_{pq}^+(\mu)$ differ only if index q is singly occupied.

Quantity $X_{\mathcal{I}}^A(\mu)$ defined in Eq.(13), is expressed in MP partitioning with the Fockian matrix elements as

$$X_q^p = \tilde{f}_{pp}^+(\mu) - \tilde{f}_{qq}^-(\mu)$$

for single excitations, and as

$$X_{rs}^{pq} = \tilde{f}_{pp}^+(\mu) + \tilde{f}_{qq}^+(\mu) - \tilde{f}_{rr}^-(\mu) - \tilde{f}_{ss}^-(\mu)$$

for double excitations.

In EN partitioning the diagonal part of the Hamiltonian is regarded as zero-order, accordingly H can be substituted for $H^{(0)}$ in Eq.(13). Many-body expressions for $X_{\mathcal{I}}^A(\mu)$ in the EN case are given both in Ref.²⁷ and Ref.³², they are however not in complete agreement. To eliminate any ambiguity, a numerical code, based on Wick's theorem was used for constructing EN excitation energies in the newly developed, redundancy-filtered formulation. In calculations relying on the redundant parametrization of T^μ , expressions of Ref.³² are used for EN excitation energies.

One-electron orbitals used for generating the PT results are either 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.

In the following we report the results for the various aspects of SS-MRPT, studied. Let us first examine the kinks in more detail. As emphasized earlier, if plotting the difference between SS-MRPT second order energy and FCI, kinks on the order of 1-10 mE_h may appear on potential energy surfaces.

1. Kinks due to small coefficients

An example for the kinks was reported by Mao et al.¹¹ on the case of the HF molecule, using natural orbitals. To alleviate the problem, Mao and coworkers applied Tikhonov regularization⁵⁵ (damping).

The kinks observed by Mao are also apparent in Fig.1. As Fig.1 illustrates, the effect is considerably larger than a few mE_h and appears at the same geometry for MP and EN partitioning, at around 2 Å. Both curves are smoothed, if setting a numerical threshold of 10^{-8} for model space coefficients. This means dropping CSF-s from the initial function with coefficients smaller than 10^{-8} in absolute value. Omitting small coefficients – either in form of Tikhonov damping or a numerical threshold – is necessary for SS-MRPT since a division by $c_\mu^{(0)}$ has to be carried out at some point to obtain amplitudes.

Kinks which disappear due to this treatment are not interesting from our present point of view. In the followings, we focus on those effects, which appear even if the appropriate numerical threshold is set for $c_\mu^{(0)}$'s.

In the results reported below, a numerical threshold of 10^{-8} is applied for CAS coefficients when working with natural orbitals. Such treatment did not have any significant effect with pseudo-canonicals, hence it was not applied.

2. Kinks due to redundancy

Examples for kinks that show up even if setting proper numerical threshold for small $c_\mu^{(0)}$'s are given by the curves labeled 'MP, T ' and 'EN, T ' in Figs. 2 and 3. These calculations were

carried out with the redundant parametrization of T^μ , including direct spectator excitations and using the decoupled form of the amplitude equations, c.f. Eqs. (11)-(13). Omitting direct spectators but keeping all other features unchanged we get the curves labeled 'T, no dir spec'. As Figs. 2 and 3 demonstrate, the error curves get smoothed by the exclusion of direct spectators. Redundancy in T^μ is however not eliminated completely by omitting direct spectators. If working with the non redundant parametrization of \tilde{T}^μ , amplitude equations Eq.(14), and the effective Hamiltonian of Eq.(17), the curves labeled 'MP, \tilde{T} ' and 'EN, \tilde{T} ' are obtained. Apparently, orthogonalization of virtual functions has only a minor numerical effect in MP partitioning, if compared with the redundant parametrization of T^μ , without direct spectators. In EN partitioning the curve can get shifted even by cca. 10 mE_h for the HF molecule. The larger effect be attributed to the fact that besides orthogonalization, the expression of $X_{\tilde{T}}^A(\mu)$ has also been altered in EN partitioning.

3. Sensitivity analysis

The largest singular value of coefficient sensitivity matrices, c.f. Eqs. (18) and (19), are presented for the HF molecule in Fig.4 and for the LiH molecule in Fig.5. Sensitivity curves of both the second order energy and the relaxed coefficients show a smoothing when stepping from the redundant parametrization of T^μ including direct spectators ('MP or EN, T ') to the method where direct spectators are excluded ('T, no dir spec'). Sudden jumps on orders of magnitude in σ_i visible on 'MP or EN, T ' curves are all extinguished in 'T, no dir spec' curves. Sensitivities are further diminished when applying the orthogonal parametrization of \tilde{T}^μ .

4. Determinantal versus spin-adapted formulation

Spin-adapted results are compared with the determinantal formulation^{29,56} in the MP partitioning in Figs. 6 and 7. Curves by the redundant parametrization of T^μ without direct spectators as well as those by the redundancy filtered method show good correspondence with the determinantal formulation, for both molecules. The largest difference occurs for intermediate bond distances but remains rather insignificant. Of the two molecules, HF shows the worse picture, but even here the difference does not exceed 1 mE_h .

5. Effect of redundancy treatment

It is important to stress that the SS-MRPT energy is not invariant to the orthogonalization used for screening redundancy in T^μ . Results of canonical orthogonalization are compared with those obtained by alternative procedures in Fig.6. One of the alternative schemes omits direct spectators and follows a different scheme for constructing functions $\tilde{\chi}_l(\mu)$ in two cases. Formulae of this orthogonalization are summarized in Table I of the Supplementary Material⁴⁸. Results produced by this approach are labeled ' \tilde{T} , alternative ort' in Fig.6. As Fig.6 shows, the non-parallelity error of the curve ' \tilde{T} , alternative ort' is significantly closer to that of the determinantal curve, if compared with the error produced by canonical orthogonalization. We have also tested the effect of including direct spectators when constructing orthogonalized functions $\tilde{\chi}_l(\mu)$, in the spirit of Mukherjee et al^{24,27,47}. This curve is not plotted in Fig.6, as it runs on the top of the curve ' \tilde{T} , alternative ort'. Similarly, ' \tilde{T} , alternative ort' and the results obtained by inclusion of direct spectators is not shown for the LiH molecule, as they could not be distinguished from ' \tilde{T} , canonical ort' on the scale of the figure.

Figs. 6 and 7 show one more method, where redundancy was achieved by dropping certain excitations instead of constructing $\{\tilde{E}_T^A\}_c$. In particular exchange spectators were dropped for core \rightarrow active(1) and active(1) \rightarrow empty excitations, $\{E_i^a\}_c$ was ignored for core \rightarrow empty and $\{E_{i_u s}^{av_s}\}_c$ was omitted for core, active(1) \rightarrow active(1), empty. The curve labeled ' T , drop out' shows these results, where the decoupled amplitude equations (11) and (12) were used. Apparently, this method suffers from the largest non-parallelity error, indicating that coupling among amplitudes within T^μ can be important.

IV. CONCLUSION

The results presented above indicate that a redundancy among excited functions, generated by spin-free excitations acting on a given model function is responsible for occasional kinks on potential energy surfaces, obtained by spin-adapted SS-MRPT, reported previously. Filtering this redundancy, potential energy curves get smoothened. In parallel with this, the orders of magnitude increase in coefficient sensitivities of SS-MRPT is diminished.

Redundancy due to spin-adaptation is either completely removed or just partially filtered.

Numerical examples show that the dominant part of the effect is caused by direct spectator excitations. Direct spectators can be handled either by ignoring them entirely or combining them with the corresponding single excitation. Numerical comparison of the two treatments show insignificant difference.

There are in principle two ways to handle redundancy arising from the application of unitary generators in spin-adapted SS-MRCC and related theories. One can either introduce additional sufficiency conditions to determine the redundant set of parameters, or carry out an orthogonalization accompanied by a reduction in the number of parameters. Mukherjee et al. have been following the first route^{25,26}, with the actual additional conditions for determining the redundant parameters in T^μ often implicit. When making use of sufficiency conditions, any step that relies on the orthogonal spectral form of the projector of the virtual space, Eq.(4) has to be considered approximative. The overlapping form of Eq.(10) can not be constructed in this case, due to matrix \mathbf{S}^{-1} being singular.

To our knowledge, the technique of orthogonalizing virtual functions in spin-adapted SS-MRPT has been explored for the first time here. Our development is restricted to two active electrons. Increasing the number of active electrons necessitates the derivation of new expressions for the orthogonalized functions and matrix elements, which are expected to become more complicated. It would still be desirable to check how the number of active electrons influences the picture gained here. Examination of redundancy treatment by orthogonalization in SS-MRCEPA theories in another issue of interest, which we plan to perform in the future.

It is important to stress that the findings presented here concern the spin-adapted formulation of SS-MRPT. Though numerical problems have been reported with the determinantal formulation of the parent coupled-cluster theory²⁹, as well as in a related multireference CC theory³⁰, the orthogonalization proposed above is irrelevant in this context, as the redundancy examined here does not show up in determinant based theories.

V. ACKNOWLEDGMENTS

This work has been supported by the Hungarian National Research Fund (OTKA), grant numbers K-81588 and K-81590. The European Union and the European Social Fund have also provided financial support to the project under the grant agreement TÁMOP 4.2.2./B-

10/1-2010-0030.

The authors wish to express their thanks to professor D. Mukherjee (Kolkata) and R. Maitra (Kolkata) for valuable discussions. The authors are also thankful to S. Das for providing her code⁵⁶ which was used to generate determinantal SS-MRPT results. The GAMESS (US) program package⁵⁷ was used for computing the CAS wavefunctions and integral lists.

-
- ¹ P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard. *Chem. Rev.*, **112**, 108, (2012).
- ² I. Shavitt and R. J. Bartlett. *Many-Body Methods in Chemistry and Physics*. Cambridge University Press, New York, 2009.
- ³ E. R. Davidson and A. A. Jarzecki. in: *Recent Advances in Multireference Methods*, volume ., pages 31–63. World Scientific, Singapore, 1999.
- ⁴ I. Lindgren and J. Morrison. *Atomic Many-Body Theory*. Springer, Berlin, 1986.
- ⁵ S. Evangelisti, J-P. Daudey, and J-P. Malrieu. *Phys. Rev. A*, **35**, 4930, (1987).
- ⁶ U. S. Mahapatra, B. Datta, and D. Mukherjee. *J. Phys. Chem. A*, **103**, 1822–1830, (1999).
- ⁷ U. S. Mahapatra, B. Datta, and D. Mukherjee. *Chem. Phys. Letters*, **299**, 42–50, (1999).
- ⁸ P. Ghosh, S. Chattopadhyay, D. Jana, and D. Mukherjee. *Int. J. Mol. Sci.*, **3**, 733–754, (2002).
- ⁹ S. Chattopadhyay, U. S. Mahapatra, and R. K. Chaudhuri. *Theor. Chem. Acc.*, **131**, 1213, (2012).
- ¹⁰ F.A. Evangelista and A.C. Simmonett and H.F. Schaefer III and D. Mukherjee and W.D. Allen. *Phys. Chem. Chem. Phys.*, **11**, 4728, (2009).
- ¹¹ S. Mao, L. Cheng, W. Liu, and D. Mukherjee. *J. Chem. Phys.*, **136**, 024106, (2012).
- ¹² 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).
- ¹³ M. R. Hoffmann, D. Datta, S. Das, D. Mukherjee, Á. Szabados, Z. Rolik, and P. R. Surján. *J. Chem. Phys.*, **131**, 204104, (2009).
 - ¹⁴ P. Piecuch and J. Paldus. *Int. J. Quantum Chem.*, **36**, 429, (1989).
 - ¹⁵ C. L. Janssen and H. F. Schaefer. *Theor. Chim. Acta*, **79**, 1, (1991).
 - ¹⁶ P. Neogrady, M. Urban, and I. Hubac. *J. Chem. Phys.*, **100**, 3706, (1994).
 - ¹⁷ P. Piecuch and J. Paldus. *J. Chem. Phys.*, **101**, 5875, (1994).
 - ¹⁸ X. Li and J. Paldus. *J. Chem. Phys.*, **101**, 8812, (1994).
 - ¹⁹ X. Li and J. Paldus. in: *Recent Advances in Coupled-Cluster Methods*, volume 3 of *Recent Advances in Computational Chemistry*, page 183. World Scientific, Singapore, 1997.
 - ²⁰ M. Nooijen and R. J. Bartlett. *J. Chem. Phys.*, **104**, 2652, (1996).
 - ²¹ D. Jana, U. S. Mahapatra, and D. Mukherjee. *Chem. Phys. Letters*, **353**, 100, (2002).
 - ²² P. G. Szalay and J. Gauss. *J. Chem. Phys.*, **107**, 9028, (1997).
 - ²³ M. Heckert, O. Heun, J. Gauss, and P. G. Szalay. *J. Chem. Phys.*, **124**, 124105, (2006).
 - ²⁴ 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*, 2005.
 - ²⁵ D. Datta and D. Mukherjee. *J. Chem. Phys.*, **134**, 054122, (2011).
 - ²⁶ R. Maitra, D. Sinha, and D. Mukherjee. *J. Chem. Phys.*, **137**, 024105, (2012).
 - ²⁷ S. Mao, L. Cheng, W. Liu, and D. Mukherjee. *J. Chem. Phys.*, **136**, 024105, (2012).
 - ²⁸ R. Haunschild, S. Mao, D. Mukherjee, and W. Klopper. *Chem. Phys. Letters*, **531**, 247, (2012).
 - ²⁹ S. Das, D. Mukherjee, and M. Kallay. *J. Chem. Phys.*, **132**, 074103, (2010).
 - ³⁰ A. Engels-Putzka and M. Hanrath. *Mol. Phys.*, **107**, 143, (2009).
 - ³¹ M. Hanrath. *J. Chem. Phys.*, **128**, 154118, (2008).

- ³² Á. Szabados. *J. Chem. Phys.*, **134**, 174113, (2011).
- ³³ U.S. Mahapatra, B. Datta, and D. Mukherjee. *J. Chem. Phys.*, **110**, 6171–6188, (1999).
- ³⁴ U.S. Mahapatra, B. Datta, and D. Mukherjee. *Chem. Phys. Letters*, **299**, 42–50, (1999).
- ³⁵ U.S. Mahapatra, B. Datta, and D. Mukherjee. *Chem. Phys. Letters*, **301**, 206, (1999).
- ³⁶ D. Sinha, R. Maitra, and D. Mukherjee. *J. Chem. Phys.*, **137**, 094104, (2012).
- ³⁷ D. Pahari, S. Chattopadhyay, S. Das, and D. Mukherjee. *Chem. Phys. Letters*, **381**, 223, (2003).
- ³⁸ B. Jeziorski and H. J. Monkhorst. *Phys. Rev. A*, **24**, 1668, (1981).
- ³⁹ S. Sen, A. Shee, and D. Mukherjee. *J. Chem. Phys.*, **137**, 074104, (2012).
- ⁴⁰ P.-O. Löwdin. *Adv. Phys.*, **5**, 1, (1966).
- ⁴¹ P. Celani and H. Werner. *J. Chem. Phys.*, **112**, 5546, (2000).
- ⁴² K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. *J. Phys. Chem.*, **94**, 5483, (1990).
- ⁴³ T. Yanai and G. K-L. Chan. *J. Chem. Phys.*, **124**, 194106, (2006).
- ⁴⁴ F. A. Evangelista and J. Gauss. *J. Chem. Phys.*, **134**, 114102, (2011).
- ⁴⁵ M. Hanauer and A. Köhn. *J. Chem. Phys.*, **134**, 204111, (2011).
- ⁴⁶ U. S. Mahapatra, B. Datta, B. Bandyopadhyay, and D. Mukherjee. *Adv. Quantum Chem.*, **30**, 163–193, (1998).
- ⁴⁷ D. Pahari, P. Ghosh, D. Mukherjee, and S. Chattopadhyay. *Theor. Chim. Acta*, **116**, 621, (2006).
- ⁴⁸ See Supplementary Material Document No.

for a detailed derivation of the orthogonalization procedure yielding terms of the redundancy-filtered cluster-operator. For information on Supplementary Material, see

<http://www.aip.org/pubservs/epaps.html> .

- ⁴⁹ F. Mertins and J. Schirmer. *Phys. Rev. A*, **53**, 2140, (1996).
- ⁵⁰ M. Hanauer and A. Köhn. *J. Chem. Phys.*, **137**, 131103, (2012).
- ⁵¹ T. Turányi. *J. Math. Chem.*, **5**, 203–248, (1990).
- ⁵² T. H. Dunning, Jr. *J. Chem. Phys.*, **90**, 1007, (1989).
- ⁵³ T. H. Dunning Jr. *J. Chem. Phys.*, **53**, 2829, (1970).
- ⁵⁴ A. Zaitevskii and J-P. Malrieu. *Chem. Phys. Letters*, **233**, 597, (1995).
- ⁵⁵ H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Dordrecht, Kluwer, 1996.
- ⁵⁶ S. Das, M. Kállay, and D. Mukherjee. *Chem. Phys.*, **392**, 83, (2012).
- ⁵⁷ M.S. Gordon and M.W. Schmidt. in: *theory and applications of computational chemistry: The first 40 years of quantum chemistry*, 2005.

TABLE I: Overlapping sets of normalized excited functions and their orthonormalized counterparts for single excitations. Excitation operators based on the orthogonalized functions and the respective amplitudes of cluster operator \tilde{T}^μ are also tabulated. Model function ϕ_μ is two-determinantal open-shell as given by Eq.(8), $u < v$ is assumed. See text for labeling convention.

overlapping functions	orthonormal functions	corresponding term in \tilde{T}^μ	
		$\left\{ {}_f \tilde{E}_{\mathcal{I}}^A \right\}_c$	${}_f \tilde{t}_{\mathcal{I}}^A$
core→empty			
$\chi_i^a = \frac{1}{\sqrt{2}} \{E_i^a\}_c \phi_\mu$	${}_1 \tilde{\chi}_i^a = -\frac{2}{3} \chi_i^a$	$\left\{ {}_1 \tilde{E}_i^a \right\}_c = -\{E_i^a\}_c$	${}_1 \tilde{t}_i^a$
$\chi_{i u_s}^{u_s a} = \frac{1}{\sqrt{2}} \{E_{i u_s}^{u_s a}\}_c \phi_\mu$	$+$ $\frac{1}{3} (\chi_{i u_s}^{u_s a} + \chi_{i v_s}^{v_s a})$	$+$ $\frac{1}{2} (\{E_{i u_s}^{u_s a}\}_c + \{E_{i v_s}^{v_s a}\}_c)$	
$\chi_{i v_s}^{v_s a} = \frac{1}{\sqrt{2}} \{E_{i v_s}^{v_s a}\}_c \phi_\mu$	${}_2 \tilde{\chi}_i^a = \frac{1}{\sqrt{3}} (\chi_{i u_s}^{u_s a} - \chi_{i v_s}^{v_s a})$	$\left\{ {}_2 \tilde{E}_i^a \right\}_c = \{E_{i u_s}^{u_s a}\}_c - \{E_{i v_s}^{v_s a}\}_c$	${}_2 \tilde{t}_i^a$
core→active(1)			
$\chi_i^{u_s} = \{E_i^{u_s}\}_c \phi_\mu$	$\tilde{\chi}_i^{u_s} = \frac{1}{2} (\chi_i^{u_s} - \chi_{i v_s}^{v_s u_s})$	$\left\{ \tilde{E}_i^{u_s} \right\}_c = \{E_i^{u_s}\}_c - \frac{1}{2} \{E_{i v_s}^{v_s u_s}\}_c$	$\tilde{t}_i^{u_s}$
$\chi_{i v_s}^{v_s u_s} = \frac{1}{2} \{E_{i v_s}^{v_s u_s}\}_c \phi_\mu$			
active(1)→empty			
$\chi_{u_s}^a = \{E_{u_s}^a\}_c \phi_\mu$	$\tilde{\chi}_{u_s}^a = \frac{1}{2} (\chi_{u_s}^a + \chi_{u_s v_s}^{v_s a})$	$\left\{ \tilde{E}_{u_s}^a \right\}_c = \{E_{u_s}^a\}_c + \{E_{u_s v_s}^{v_s a}\}_c$	$\tilde{t}_{u_s}^a$
$\chi_{u_s v_s}^{v_s a} = \{E_{u_s v_s}^{v_s a}\}_c \phi_\mu$			

TABLE II: Relation between cluster amplitudes in the redundant parametrization Eq.(7) of T^μ and the orthogonal parametrization Eq.(9) of \tilde{T}^μ . Case of single excitations. Model function ϕ_μ is two-determinantal open-shell, as given by Eq.(8). See text for labeling convention.

core→empty

$$t_i^a = -{}_1\tilde{t}_i^a$$

$$t_{i u_s}^{u_s a} = \frac{1}{2} {}_1\tilde{t}_i^a + {}_2\tilde{t}_i^a$$

$$t_{i v_s}^{v_s a} = \frac{1}{2} {}_1\tilde{t}_i^a - {}_2\tilde{t}_i^a$$

core→active(1)

$$t_i^{u_s} = \tilde{t}_i^{u_s}$$

$$t_{i v_s}^{v_s u_s} = -\frac{1}{2} \tilde{t}_i^{u_s}$$

active(1)→empty

$$t_{u_s}^a = \tilde{t}_{u_s}^a$$

$$t_{v_s u_s}^{a v_s} = \tilde{t}_{u_s}^a$$

TABLE III: Overlapping sets of normalized excited functions and their orthonormalized counterparts for double excitations. Excitation operators based on the orthogonalized functions and the respective amplitudes of cluster operator \tilde{T}^μ are also tabulated. Model function ϕ_μ is two-determinantal open-shell, according to Eq.(8). Index ordering $i < j$, $a < b$ and $u < v$ is assumed. See text for labeling convention. Notation $p(ua)$ stands for the parity of the permutation ordering the pair (ua) .

overlapping functions	orthonormal functions	corresponding term in \tilde{T}^μ	
		$\left\{ {}_f \tilde{E}_{\mathcal{I}}^A \right\}_c$	${}_f \tilde{t}_{\mathcal{I}}^A$
2 core→2 empty			
$\chi_{ij}^{ab} = \frac{1}{2} \left\{ E_{ij}^{ab} \right\}_c \phi_\mu$	${}_1 \tilde{\chi}_{ij}^{ab} = \chi_{ij}^{ab} + \chi_{ij}^{ba}$	$\left\{ {}_1 \tilde{E}_{ij}^{ab} \right\}_c = \left\{ E_{ij}^{ab} \right\}_c + \left\{ E_{ij}^{ba} \right\}_c$	${}_1 \tilde{t}_{ij}^{ab}$
$\chi_{ij}^{ba} = \frac{1}{2} \left\{ E_{ij}^{ba} \right\}_c \phi_\mu$	${}_2 \tilde{\chi}_{ij}^{ab} = \frac{1}{\sqrt{3}} \left(\chi_{ij}^{ab} - \chi_{ij}^{ba} \right)$	$\left\{ {}_2 \tilde{E}_{ij}^{ab} \right\}_c = \left\{ E_{ij}^{ab} \right\}_c - \left\{ E_{ij}^{ba} \right\}_c$	${}_2 \tilde{t}_{ij}^{ab}$
2 core→active(1), empty			
$\chi_{ij}^{u_s a} = \frac{1}{\sqrt{2}} \left\{ E_{ij}^{u_s a} \right\}_c \phi_\mu$	${}_1 \tilde{\chi}_{ij}^{u_s a} = \chi_{ij}^{u_s a} + \chi_{ij}^{a u_s}$	$\left\{ {}_1 \tilde{E}_{ij}^{u_s a} \right\}_c = \left\{ E_{ij}^{u_s a} \right\}_c + \left\{ E_{ij}^{a u_s} \right\}_c$	${}_1 \tilde{t}_{ij}^{u_s a}$
$\chi_{ij}^{a u_s} = \frac{1}{\sqrt{2}} \left\{ E_{ij}^{a u_s} \right\}_c \phi_\mu$	${}_2 \tilde{\chi}_{ij}^{u_s a} = (-1)^{p(u_s a)} \cdot \frac{1}{\sqrt{3}} \left(\chi_{ij}^{u_s a} - \chi_{ij}^{a u_s} \right)$	$\left\{ {}_2 \tilde{E}_{ij}^{u_s a} \right\}_c = (-1)^{p(u_s a)} \cdot \left(\left\{ E_{ij}^{u_s a} \right\}_c - \left\{ E_{ij}^{a u_s} \right\}_c \right)$	${}_2 \tilde{t}_{ij}^{u_s a}$
core, active(1)→2 empty			
$\chi_{iu_s}^{ab} = \frac{1}{\sqrt{2}} \left\{ E_{iu_s}^{ab} \right\}_c \phi_\mu$	${}_1 \tilde{\chi}_{iu_s}^{ab} = \chi_{iu_s}^{ab} + \chi_{iu_s}^{ba}$	$\left\{ {}_1 \tilde{E}_{iu_s}^{ab} \right\}_c = \left\{ E_{iu_s}^{ab} \right\}_c + \left\{ E_{iu_s}^{ba} \right\}_c$	${}_1 \tilde{t}_{iu_s}^{ab}$
$\chi_{iu_s}^{ba} = \frac{1}{\sqrt{2}} \left\{ E_{iu_s}^{ba} \right\}_c \phi_\mu$	${}_2 \tilde{\chi}_{iu_s}^{ab} = \frac{1}{\sqrt{3}} \left(\chi_{iu_s}^{ab} - \chi_{iu_s}^{ba} \right)$	$\left\{ {}_2 \tilde{E}_{iu_s}^{ab} \right\}_c = \left\{ E_{iu_s}^{ab} \right\}_c - \left\{ E_{iu_s}^{ba} \right\}_c$	${}_2 \tilde{t}_{iu_s}^{ab}$

TABLE III: Cont'd

overlapping functions	orthonormal functions	corresponding term in \tilde{T}^μ	$\tilde{t}_{\mathcal{I}}^A$
2 core→2 active(1)			
$\chi_{ij}^{u_s v_s} = \{E_{ij}^{u_s v_s}\}_c \phi_\mu$ $\chi_{ij}^{v_s u_s} = \{E_{ij}^{v_s u_s}\}_c \phi_\mu$	$\tilde{\chi}_{ij}^{u_s v_s} = \frac{1}{2} (\tilde{\chi}_{ij}^{u_s v_s} + \tilde{\chi}_{ij}^{v_s u_s})$	$\{\tilde{E}_{ij}^{u_s v_s}\}_c = \{E_{ij}^{u_s v_s}\}_c + \{E_{ij}^{v_s u_s}\}_c$	$\tilde{t}_{ij}^{u_s v_s}$
2 active(1)→2 empty			
$\chi_{u_s v_s}^{ab} = \{E_{u_s v_s}^{ab}\}_c \phi_\mu$ $\chi_{u_s v_s}^{ba} = \{E_{u_s v_s}^{ba}\}_c \phi_\mu$	$\tilde{\chi}_{u_s v_s}^{ab} = \frac{1}{2} (\tilde{\chi}_{u_s v_s}^{ab} + \tilde{\chi}_{u_s v_s}^{ba})$	$\{\tilde{E}_{u_s v_s}^{ab}\}_c = \{E_{u_s v_s}^{ab}\}_c + \{E_{u_s v_s}^{ba}\}_c$	$\tilde{t}_{u_s v_s}^{ab}$
core, active(1)→active(1), empty			
$\chi_{iu_s}^{v_s a} = \{E_{iu_s}^{v_s a}\}_c \phi_\mu$ $\chi_{iu_s}^{a v_s} = \frac{1}{2} \{E_{iu_s}^{a v_s}\}_c \phi_\mu$	$\tilde{\chi}_{iu_s}^{v_s a} = (-1)^{p(v_s a)} \cdot$ $\frac{1}{2} (\tilde{\chi}_{iu_s}^{v_s a} - \tilde{\chi}_{iu_s}^{a v_s})$	$\{\tilde{E}_{iu_s}^{v_s a}\}_c = (-1)^{p(v_s a)} \cdot$ $\left(\{E_{iu_s}^{v_s a}\}_c - \frac{1}{2} \{E_{iu_s}^{a v_s}\}_c \right)$	$\tilde{t}_{iu_s}^{v_s a}$

TABLE IV: Relation between cluster amplitudes in the redundant parametrization Eq.(7) of T^μ and the orthogonal parametrization Eq.(9) of \tilde{T}^μ . Case of double excitations. Model function ϕ_μ is two-determinantal open-shell, as given by Eq.(8). See text for labeling convention. Index ordering $i < j$, $a < b$ and $u < v$ is assumed. Notation $p(ua)$ stands for the parity of the permutation ordering the pair (ua) .

2 core→2 empty
$t_{ij}^{ab} = {}_1\tilde{t}_{ij}^{ab} + {}_2\tilde{t}_{ij}^{ab}$ $t_{ij}^{ba} = {}_1\tilde{t}_{ij}^{ab} - {}_2\tilde{t}_{ij}^{ab}$
2 core→active(1), empty
$t_{ij}^{u_s a} = {}_1\tilde{t}_{ij}^{u_s a} + (-1)^{p(u_s a)} {}_2\tilde{t}_{ij}^{u_s a}$ $t_{ij}^{a u_s} = {}_1\tilde{t}_{ij}^{u_s a} - (-1)^{p(u_s a)} {}_2\tilde{t}_{ij}^{u_s a}$
core, active(1)→2 empty
$t_{iu_s}^{ab} = {}_1\tilde{t}_{iu_s}^{ab} + {}_2\tilde{t}_{iu_s}^{ab}$ $t_{iu_s}^{ba} = {}_1\tilde{t}_{iu_s}^{ab} - {}_2\tilde{t}_{iu_s}^{ab}$
2 core→2 active(1)
$t_{ij}^{u_s v_s} = \tilde{t}_{ij}^{u_s v_s}$ $t_{ij}^{v_s u_s} = \tilde{t}_{ij}^{u_s v_s}$
2 active(1)→2 empty
$t_{u_s v_s}^{ab} = \tilde{t}_{u_s v_s}^{ab}$ $t_{u_s v_s}^{ba} = \tilde{t}_{u_s v_s}^{ab}$
core, active(1)→active(1), empty
$t_{iu_s}^{a v_s} = (-1)^{p(v_s a)} \frac{1}{2} \tilde{t}_{iu_s}^{v_s a}$ $t_{iu_s}^{v_s a} = (-1)^{p(v_s a)} \tilde{t}_{iu_s}^{v_s a}$

TABLE V: Elements \tilde{C} of Eq.(15), for core \rightarrow virtual excitation, $i \rightarrow a$. Abbreviations: ‘cs’ – closed-shell, ‘os’ – open-shell. Indices f and g are left blank, when not applicable. Virtual function ${}_f\tilde{\chi}_i^a(\mu)$ is assumed normalized. See text for labeling convention and the definition of ${}_fX_i^a(\mu)$.

μ	f	ν	g	$\tilde{C}(\mu, i \rightarrow a, f; \nu, i \rightarrow a, g)$
cs		cs		$\sqrt{2} + \sqrt{2} \delta_{\mu\nu} (X_i^a(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$
cs, x_d		os, $u_s < v_s$	1	$-\sqrt{2} (1 + \frac{1}{2}(\delta_{xu} + \delta_{xv}))$
cs, x_d		os, $u_s < v_s$	2	$-\sqrt{2}(\delta_{xu} - \delta_{xv})$
os, $x_s < y_s$	1	cs		$-\sqrt{2}$
os, $x_s < y_s$	1	os, $u_s < v_s$	1	$\sqrt{2} (1 + \frac{1}{4}(\delta_{xu} + \delta_{yu} + \delta_{xv} + \delta_{yv})) + \frac{3}{\sqrt{2}} \delta_{\mu\nu} ({}_1X_i^a(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$
os, $x_s < y_s$	1	os, $u_s < v_s$	2	$\frac{1}{\sqrt{2}} (\delta_{xu} - \delta_{xv} + \delta_{yu} - \delta_{yv})$
os, $x_s < y_s$	2	cs		0
os, $x_s < y_s$	2	os, $u_s < v_s$	1	$\frac{\sqrt{3}}{2\sqrt{2}}(\delta_{xu} + \delta_{xv} - \delta_{yu} - \delta_{yv})$
os, $x_s < y_s$	2	os, $u_s < v_s$	2	$\frac{\sqrt{3}}{\sqrt{2}} (\delta_{xu} - \delta_{xv} - \delta_{yu} + \delta_{yv}) + \sqrt{6} \delta_{\mu\nu} ({}_2X_i^a(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$

TABLE VI: Elements \tilde{C} of Eq.(15), for core→active excitations, $i \rightarrow w$. Abbreviations: ‘cs’ – closed-shell, ‘os’ – open-shell. Indices f and g are left blank, when not applicable. Virtual function ${}_f\tilde{\chi}_i^w(\mu)$ is assumed normalized. See text for labeling convention and the definition of ${}_fX_i^w(\mu)$.

μ	f	ν	g	$\tilde{C}(\mu, i \rightarrow w, f; \nu, i \rightarrow w, g)$
CS, w_z		CS, w_z		$\sqrt{2} + \sqrt{2} \delta_{\mu\nu} (X_i^w(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$
CS, x_d, w_z		OS, $u_s < v_s, w_z$	1	$-\sqrt{2} (1 + \frac{1}{2}(\delta_{xu} + \delta_{xv}))$
CS, w_z		OS, $u_s < v_s, w_z$	2	$-\sqrt{2}(\delta_{xu} - \delta_{xv})$
CS, w_z		OS, u_s, w_s		$\sqrt{2}(1 + \frac{1}{2}\delta_{xu})$
OS, $x_s < y_s, w_z$	1	CS, w_z		$-\sqrt{2}$
OS, $x_s < y_s, w_z$	1	OS, $u_s < v_s, w_z$	1	$\sqrt{2} (1 + \frac{1}{4}(\delta_{xu} + \delta_{xv} + \delta_{yu} + \delta_{yv})) + \frac{3}{\sqrt{2}}\delta_{\mu\nu}({}_1X_i^w(\mu) - E_{\text{CAS}})H_{\mu\mu}^{-1}$
OS, $x_s < y_s, w_z$	1	OS, $u_s < v_s, w_z$	2	$\frac{1}{\sqrt{2}} (\delta_{xu} - \delta_{xv} + \delta_{yu} - \delta_{yv})$
OS, $x_s < y_s, w_z$	1	OS, u_s, w_s		$-\sqrt{2} (1 + \frac{1}{4} (\delta_{xu} + \delta_{yu}))$
OS, $x_s < y_s, w_z$	2	CS, w_z		0
OS, $x_s < y_s, w_z$	2	OS, $u_s < v_s, w_z$	1	$\frac{\sqrt{3}}{2\sqrt{2}} (\delta_{xu} + \delta_{xv} - \delta_{yu} - \delta_{yv})$
OS, $x_s < y_s, w_z$	2	OS, $u_s < v_s, w_z$	2	$\frac{\sqrt{3}}{\sqrt{2}} (\delta_{xu} - \delta_{xv} - \delta_{yu} + \delta_{yv}) + \sqrt{6} \delta_{\mu\nu} ({}_2X_i^w(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$
OS, $x_s < y_s, w_z$	2	OS, u_s, w_s		$\frac{\sqrt{3}}{2\sqrt{2}} (-\delta_{xu} + \delta_{yu})$
OS, x_s, w_s		CS, w_z		1
OS, x_s, w_s		OS, $u_s < v_s, w_z$	1	$-(1 + \delta_{xu} + \delta_{xv})$
OS, x_s, w_s		OS, $u_s < v_s, w_z$	2	$2(-\delta_{xu} + \delta_{xv})$
OS, x_s, w_s		OS, u_s, w_s		$(1 + \delta_{xu}) + 2\delta_{\mu\nu} (X_i^w(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$

TABLE VII: Elements \tilde{C} of Eq.(15), for active \rightarrow virtual excitations, i.e. $w \rightarrow a$. Abbreviations: ‘cs’ – closed-shell, ‘os’ – open-shell. Virtual functions $\tilde{\chi}_w^a(\mu)$ are assumed normalized. See text for labeling convention and the definition of $X_w^a(\mu)$.

μ	ν	$\tilde{C}(\mu, w \rightarrow a; \nu, w \rightarrow a)$
cs, w_d	cs, w_d	$\sqrt{2} + \sqrt{2} \delta_{\mu\nu} (X_w^a(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$
cs, w_d	os, u_s, w_s	$\sqrt{2}$
os, x_s, w_s	cs, w_d	1
os, x_s, w_s	os, u_s, w_s	$(1 + \delta_{xu}) + 2 \delta_{\mu\nu} (X_w^a(\mu) - E_{\text{CAS}}) H_{\mu\mu}^{-1}$

TABLE VIII: Elements \tilde{C} of Eq.(15), for non coupled double excitations, i.e. $f = g$. Values of \tilde{C} are collected for the possible combinations of $\mathcal{I} \rightarrow \mathcal{A}$ with reference functions, ϕ_μ . Admissible types for CSF ν agree with types listed for μ . Description of CSF μ is indicated in the rows, indices \mathcal{I}, \mathcal{A} and f are given in column headers. Abbreviations: ‘cs’ – closed-shell, ‘os’ – open shell, ‘n.a.’ – not applicable. Shorthand Y stands for ${}_f Y_{ij}^{ab}(\mu) = 1 + \delta_{\mu\nu}({}_f X_{ij}^{ab}(\mu) - E_{\text{CAS}})H_{\mu\mu}^{-1}$. See text for the definition of ${}_f X_{ij}^{ab}(\mu)$. Virtual functions ${}_f \tilde{\chi}_{ij}^{ab}(\mu)$ are assumed normalized.

excitation 2 core \rightarrow 2 virtual: $(i, j) \rightarrow (a, b)$

type of μ	$i < j$		$i = j$	$i < j$		$i = j$
	$a < b$	$a < b$	$a < b$	$a = b$	$a = b$	$a = b$
	$f = 1$	$f = 2$				
any	$2 {}_1 Y_{ij}^{ab}$	$2\sqrt{3} {}_2 Y_{ij}^{ab}$	$\sqrt{2} Y_{ii}^{ab}$	$\sqrt{2} Y_{ij}^{aa}$	$2 Y_{ii}^{aa}$	

excitation 2 core \rightarrow active, virtual: $(i, j) \rightarrow (u, a)$

type of μ	$i < j$		$i = j$
	$u < a$	$u < a$	$u < a$
	$f = 1$	$f = 2$	
cs or os, u_z	$2 {}_1 Y_{ij}^{ua}$	$2\sqrt{3} {}_2 Y_{ij}^{ua}$	$\sqrt{2} Y_{ii}^{ua}$
os, u_s	$\sqrt{2} {}_1 Y_{ij}^{ua}$	$\sqrt{6} {}_2 Y_{ij}^{ua}$	Y_{ii}^{ua}

excitation core, active \rightarrow 2 virtual: $(i, u) \rightarrow (a, b)$

type of μ	$i < u$		$i < u$
	$a < b$	$a < b$	$a = b$
	$f = 1$	$f = 2$	
cs or os, u_d	$2 {}_1 Y_{iu}^{ab}$	$2\sqrt{3} {}_2 Y_{iu}^{ab}$	$\sqrt{2} Y_{iu}^{aa}$
os, u_s	$\sqrt{2} {}_1 Y_{iu}^{ab}$	$\sqrt{6} {}_2 Y_{iu}^{ab}$	Y_{iu}^{aa}

TABLE VIII: Cont'd.

excitation 2 core→2 active: $(i, j) \rightarrow (u, v)$

type of μ	$i < j$		$i = j$	$i < j$	$i = j$
	$u < v$		$u < v$	$u = v$	$u = v$
	$f = 1$	$f = 2$			
cs or os, u_z, v_z	$2 \ 1 Y_{ij}^{uv}$	$2\sqrt{3} \ 2 Y_{ij}^{uv}$	$\sqrt{2} \ Y_{ii}^{uv}$	$\sqrt{2} \ Y_{ij}^{uu}$	$2 \ Y_{ii}^{uu}$
os, u_s, v_z or u_z, v_s	$\sqrt{2} \ 1 Y_{ij}^{uv}$	$\sqrt{6} \ 2 Y_{ij}^{uv}$	Y_{ii}^{uv}	n.a.	n.a.
os, u_s, v_s	$2 \ 1 Y_{ij}^{uv}$	0	$\sqrt{2} \ Y_{ii}^{uv}$	n.a.	n.a.

excitation 2 active→2 virtual: $(u, v) \rightarrow (a, b)$

type of μ	$u < v$	
	$a < b$	$a = b$
os, u_s, v_s	$2 \ Y_{uv}^{ab}$	$\sqrt{2} \ Y_{uv}^{aa}$

excitation 2 active→active, virtual: $(u, v) \rightarrow (w, a)$

type of μ	$u < v$
	$w < a$
os, u_s, v_s, w_z	$2 \ Y_{uv}^{wa}$

TABLE IX: Matrix \tilde{C} of Eq.(15), for core, active \rightarrow 2 active excitations, $(i, w) \rightarrow (u, v)$. Description of CSF μ is indicated in the rows, together with index f , when applicable. Characterization of CSF ν is given in column headers, together with index g , when applicable. Abbreviations: ‘cs’ – closed-shell, ‘os’ – open shell. Shorthand Y stands for ${}_f Y_{iw}^{uv}(\mu) = 1 + \delta_{\mu\nu}({}_f X_{iw}^{uv}(\mu) - E_{\text{CAS}})H_{\mu\mu}^{-1}$. See text for the definition of ${}_f X_{iw}^{uv}(\mu)$. Virtual functions ${}_f \tilde{\chi}_{iw}^{uv}(\mu)$ are assumed normalized. The table applies for the core, active \rightarrow active, virtual excitations also, with a substituted for v_z . Column and row referring to a v_s is not applicable in this case.

	f	$g = 1$		$g = 2$		os	os
		cs	os	cs	os		
		$w_d, u_z < v_z$	$w_s, u_z < v_z$	$w_d, u_z < v_z$	$w_s, u_z < v_z$		
cs, $w_d, u_z < v_z$	1	$2 {}_f Y_{iw}^{uv}(\mu)$	2	0	0	$\frac{1}{2}$	$-\frac{1}{2}$
os, $w_s, u_z < v_z$	1	$\sqrt{2}$	$\sqrt{2} {}_f Y_{iw}^{uv}(\mu)$	0	0	$\frac{1}{2\sqrt{2}}$	$-\frac{1}{2\sqrt{2}}$
cs, $w_d, u_z < v_z$	2	0	0	$2\sqrt{3} {}_f Y_{iw}^{uv}(\mu)$	$2\sqrt{3}$	$\frac{3\sqrt{3}}{2}$	$\frac{3\sqrt{3}}{2}$
os, $w_s, u_z < v_z$	2	0	0	$\sqrt{6}$	$\sqrt{6} {}_f Y_{iw}^{uv}(\mu)$	$\frac{3\sqrt{3}}{2\sqrt{2}}$	$\frac{3\sqrt{3}}{2\sqrt{2}}$
os, $w_s, u_s < v_z$		-1	-1	3	3	$2 Y_{iw}^{uv}(\mu)$	$\frac{5}{2}$
os, $w_s, u_z < v_s$		1	1	3	3	$\frac{5}{2}$	$2 Y_{iw}^{uv}(\mu)$

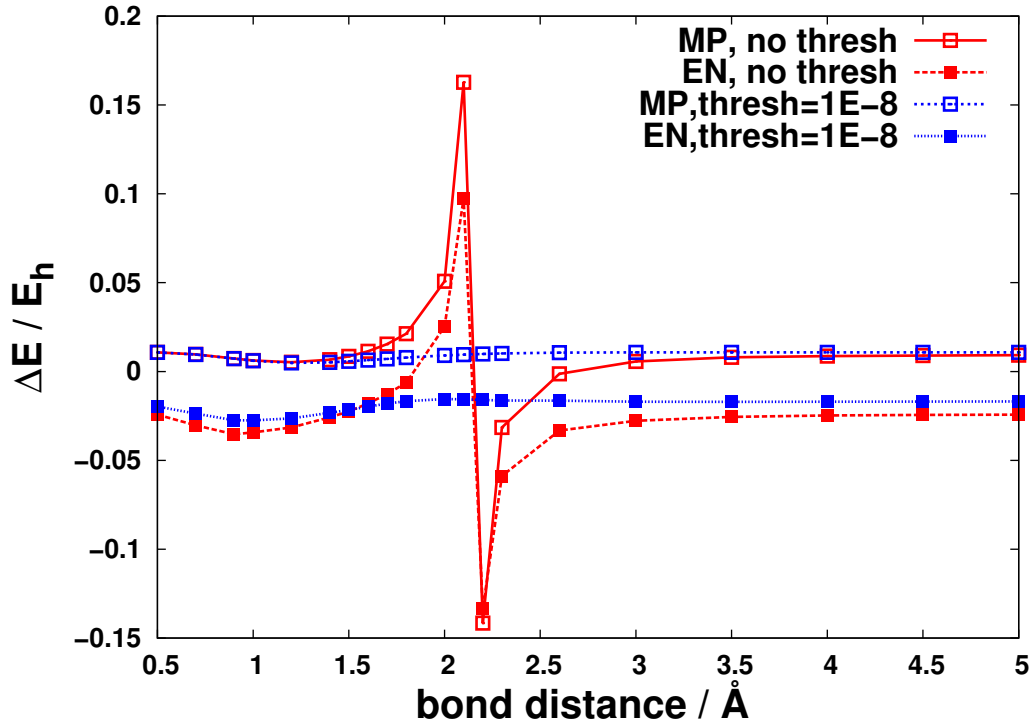


FIG. 1: Errors of second order SS-MRPT energies for the ground state of the HF system in cc-pVDZ basis⁵². Reference function is CAS(2,2), active orbitals are naturals. Full CI values are subtracted from total SS-MRPT energies. Partitioning is either MP or EN. Label 'thresh=1E-8' corresponds to omitting model space CSF's with coefficients smaller than 10^{-8} in absolute value. No such treatment is applied for label 'no thresh'.

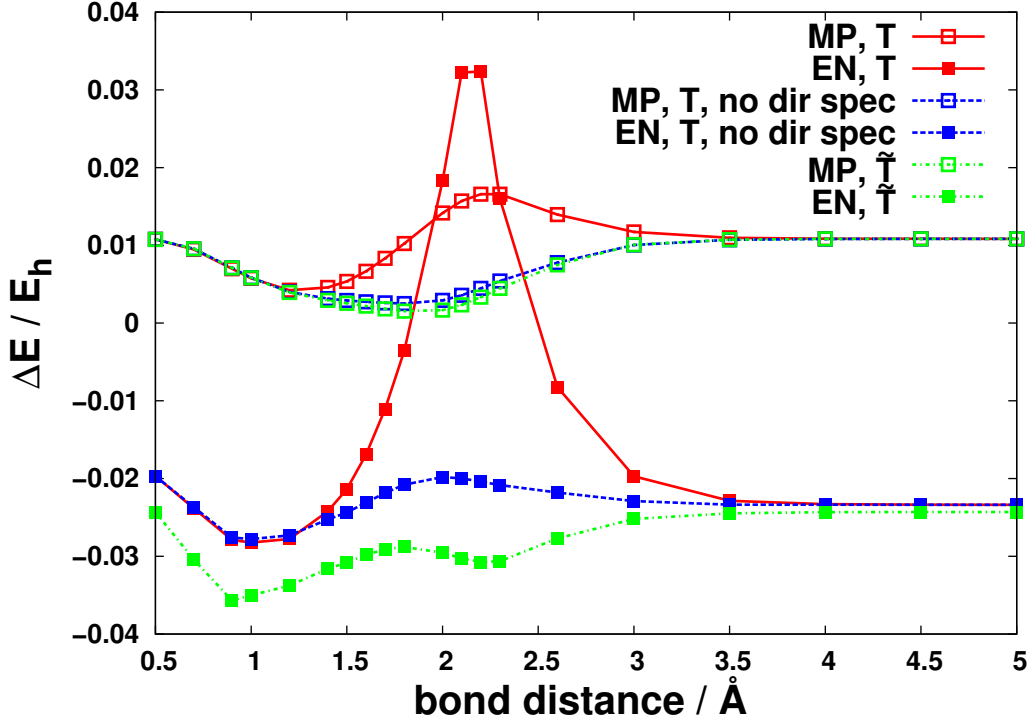


FIG. 2: Errors of second order SS-MRPT energies for the ground state of the HF system. Active orbitals are pseudo-canonicals. Full CI values are subtracted from total SS-MRPT energies. Basis set, reference function and partitionings agree with that of Fig.1. Key legends: 'T' refers to the redundant parametrization of T^μ with direct spectators included, 'T, no dir spec' applies the parametrization of Eq.(7) without direct spectators, ' \tilde{T} ' uses the non-redundant parametrization of \tilde{T}^μ .

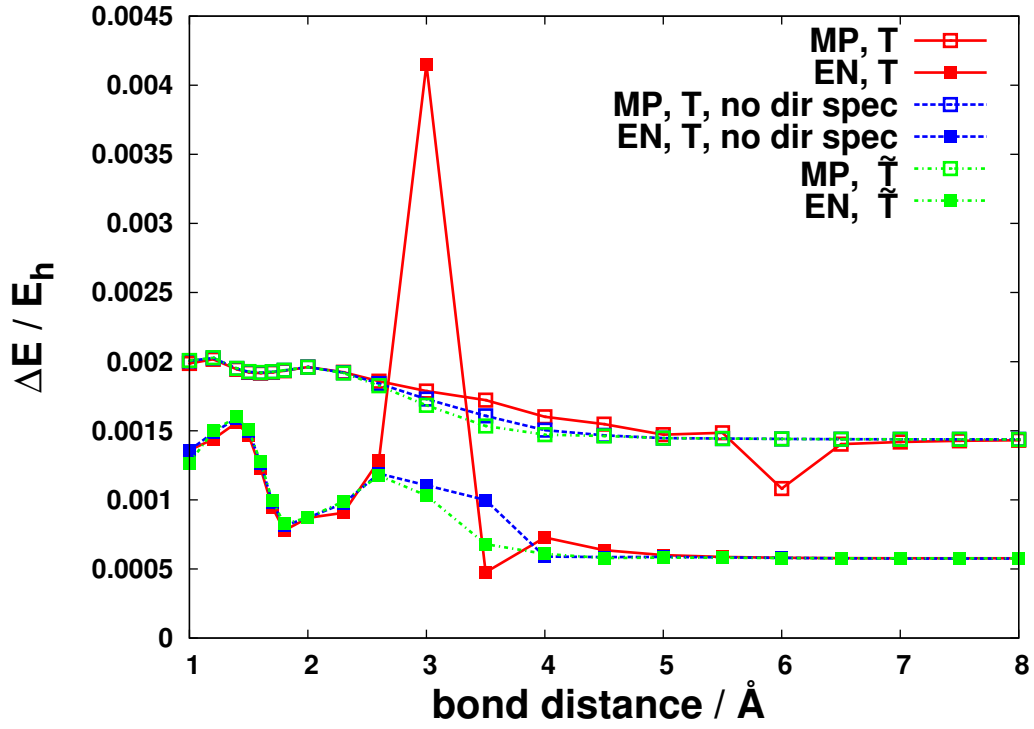


FIG. 3: Errors of second order SS-MRPT energies for the ground state of the LiH molecule in Dunning's DZP basis⁵³. Reference function is CAS(2,5), active orbitals are pseudo-canonicals. Full CI values are subtracted from total SS-MRPT energies. Partitioning is either MP or EN. See Fig.2 for key legends.

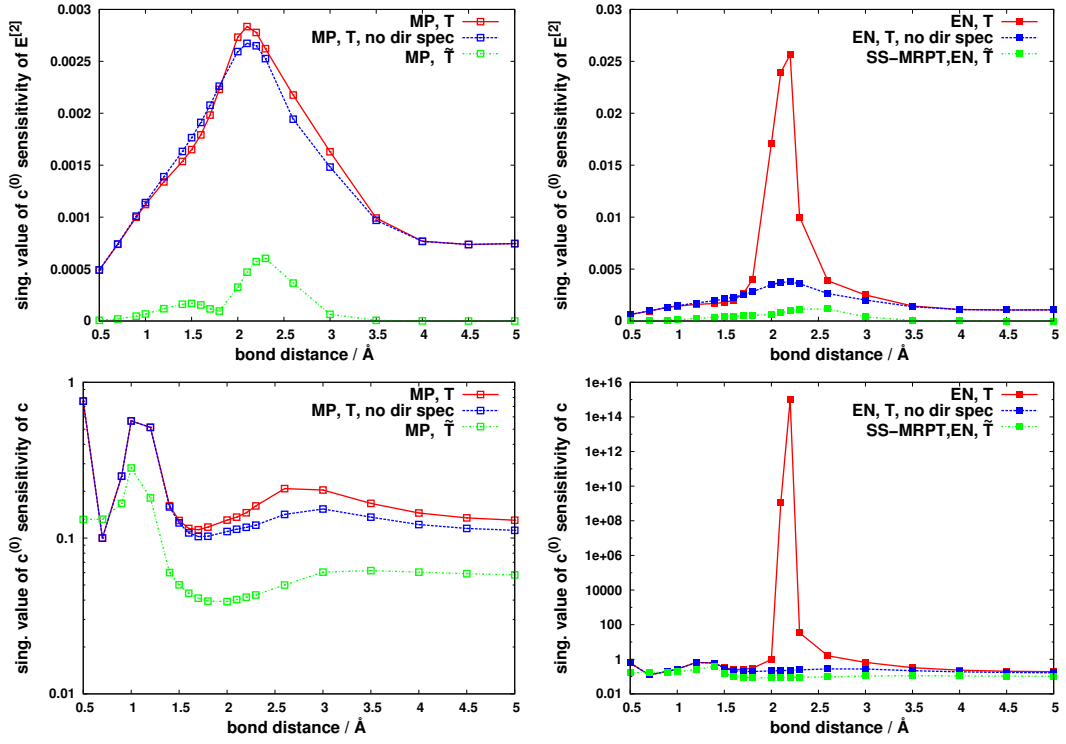


FIG. 4: The largest singular value of the coefficient sensitivity matrix of SS-MRPT energy (Eq.(18)) and relaxed coefficients (Eq.(19)) for the ground state of the HF molecule. Basis set, reference function and partitioning agrees with Fig.2. Key legends are also given at Fig.2.

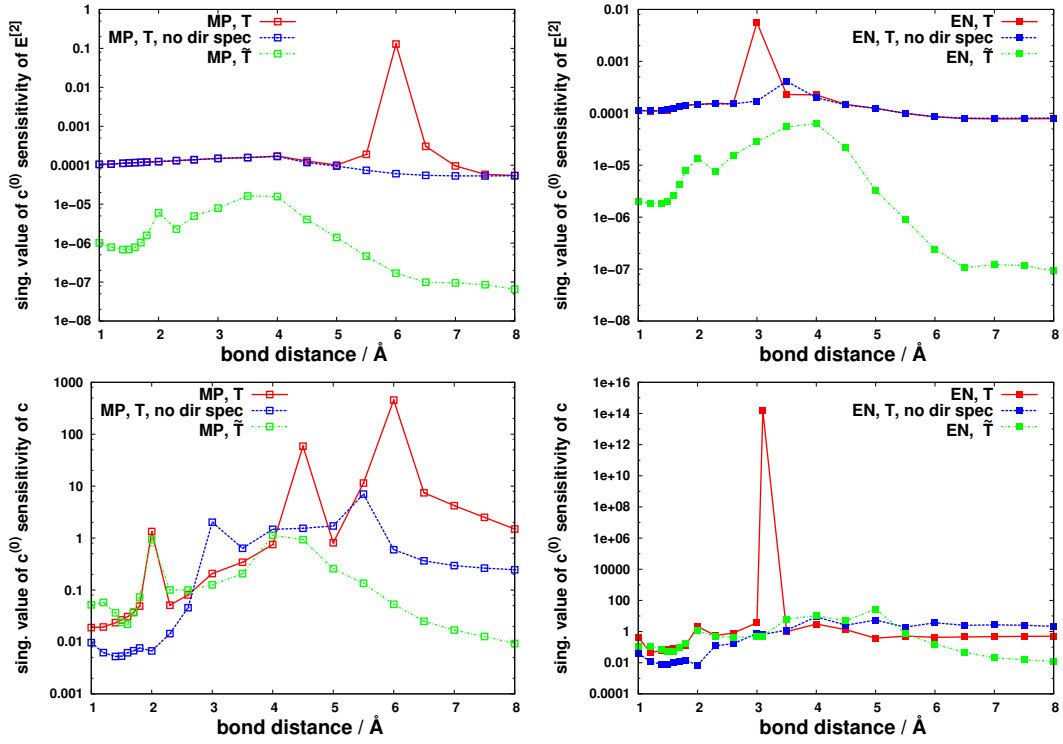


FIG. 5: The largest singular value of the coefficient sensitivity matrix of SS-MRPT energy (Eq.(18)) and relaxed coefficients (Eq.(19)) for the ground state of the LiH molecule. Basis set, reference function and partitioning agrees with Fig.3. Key legends are given at Fig.2.

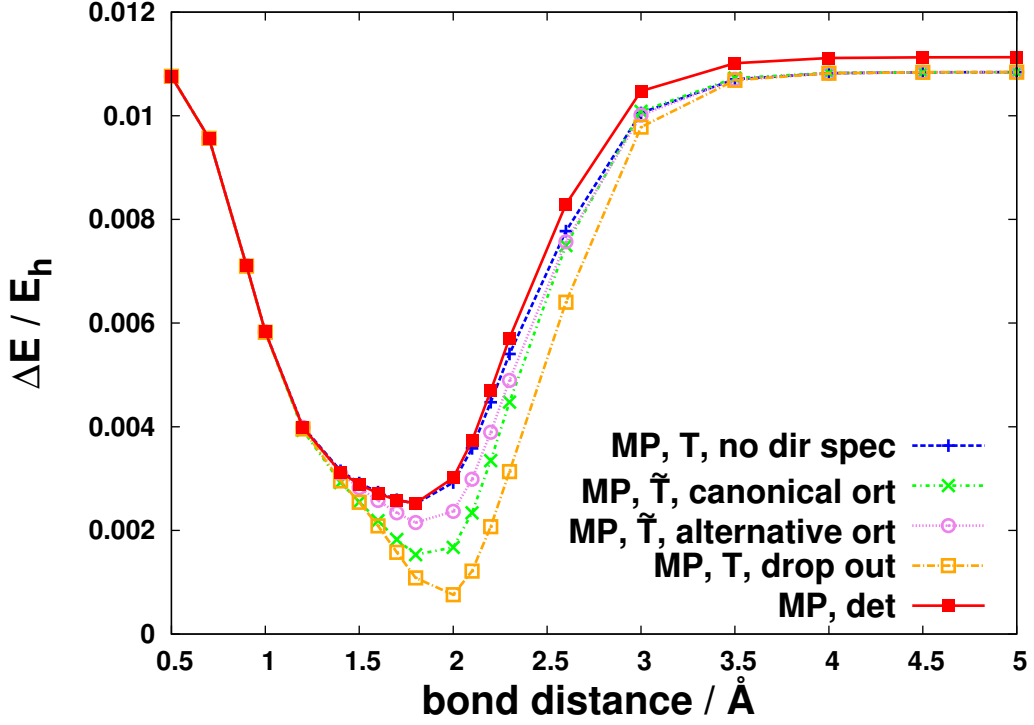


FIG. 6: Errors of second order SS-MRPT energies in MP partitioning for the ground state of the HF system. Full CI values are subtracted from total SS-MRPT energies. Basis set and reference function agrees with that of Fig.2. Key legends: 'T, no dir spec' applies the parametrization of Eq.(7) without direct spectators, ' \tilde{T} canonical ort' uses the non-redundant parametrization of \tilde{T}^μ with the functions of Tables I and III, ' \tilde{T} , alternative ort' uses the non-redundant parametrization of \tilde{T}^μ with the functions described in Table I of the Supplementary Material⁴⁸, 'T, drop out' refers to the method where terms of T^μ of Eq.(7) are dropped to set the parametrization non-redundant, 'det' refers to the determinantal approach.

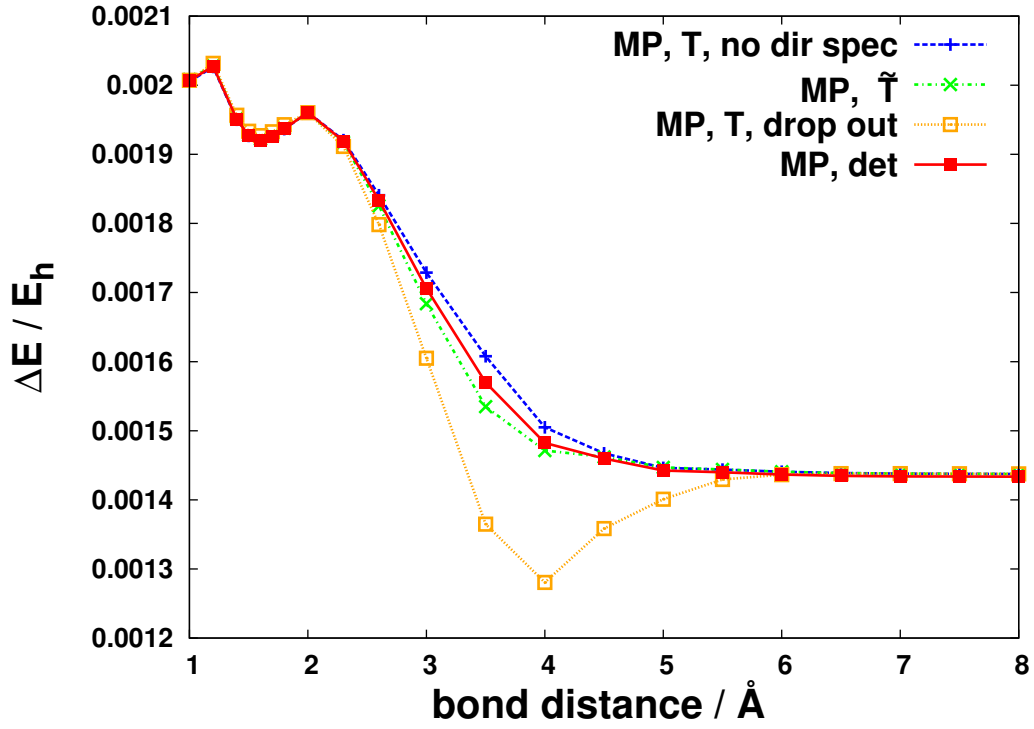


FIG. 7: Errors of second order SS-MRPT energies in MP partitioning for the ground state of the LiH system. Full CI values are subtracted from total SS-MRPT energies. Basis set and reference function agrees with that of Fig.3. Legend 'T-tilde' refers to the non-redundant parametrization of \tilde{T}^μ , see Fig.6 for other key legends.