

Equation-of-motion coupled cluster perturbation theory revisited

Janus J. Eriksen¹, Poul Jørgensen, Jeppe Olsen, and Jürgen Gauss

Citation: *The Journal of Chemical Physics* **140**, 174114 (2014); doi: 10.1063/1.4873138

View online: <http://dx.doi.org/10.1063/1.4873138>

View Table of Contents: <http://aip.scitation.org/toc/jcp/140/17>

Published by the *American Institute of Physics*

Articles you may be interested in

[Convergence of coupled cluster perturbation theory](#)

The Journal of Chemical Physics **145**, 224104 (2016); 10.1063/1.4971294

[Stochastic multi-reference perturbation theory with application to the linearized coupled cluster method](#)

The Journal of Chemical Physics **146**, 044107 (2017); 10.1063/1.4974177

COMPLETELY

REDESIGNED!



**PHYSICS
TODAY**

Physics Today Buyer's Guide
Search with a purpose.

Equation-of-motion coupled cluster perturbation theory revisited

Janus J. Eriksen,^{1,a)} Poul Jørgensen,¹ Jeppe Olsen,¹ and Jürgen Gauss²

¹*qLEAP Center for Theoretical Chemistry, Department of Chemistry, Aarhus University, Langelandsgade 140, DK-8000 Aarhus C, Denmark*

²*Institut für Physikalische Chemie, Johannes Gutenberg-Universität Mainz, Duesbergweg 10-14, D-55128 Mainz, Germany*

(Received 10 March 2014; accepted 14 April 2014; published online 7 May 2014)

The equation-of-motion coupled cluster (EOM-CC) framework has been used for deriving a novel series of perturbative corrections to the coupled cluster singles and doubles energy that formally converges towards the full configuration interaction energy limit. The series is based on a Møller-Plesset partitioning of the Hamiltonian and thus size extensive at any order in the perturbation, thereby remedying the major deficiency inherent to previous perturbation series based on the EOM-CC ansatz.

© 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4873138>]

I. INTRODUCTION

To obtain high accuracy coupled cluster¹ (CC) electronic structure models, the energy of the coupled cluster singles and doubles² (CCSD) model has to be augmented by corrections originating from higher-level excitations, in particular triple, but often also quadruple excitations. Different schemes have been proposed for this purpose; one acclaimed strategy has been to correct the CCSD energy by the leading-order triples contributions from many-body perturbation theory³ (MBPT). Using this approach, energy corrections are determined with respect to the Hartree-Fock (HF) reference state, but not directly with respect to the CCSD state. The most prominent of these models is the CCSD(T)^{4,5} non-iterative approximate triples model that has become the gold standard within the field of applied electronic structure theory. Related models, which differ by taking into account the asymmetric nature of CC theory, have also been proposed. Of these, the Λ -CCSD(T)⁶ model (also known as a-CCSD(T)⁷) is the CCSD(T) counterpart.

A different strategy for determining specific triples corrections to the CCSD energy revolves around equation-of-motion CC (EOM-CC) theory. Depending on the focus of the derivation and whether or not the reference energy and state(s) are expanded in terms of a perturbation, different series have been derived such as those of the CC(*m*)PT(*n*)^{8,9} hierarchy and those of the biorthogonal method of moments of CC equations (MMCC) theory.¹⁰ Using such approaches, corrections are determined directly to the CCSD energy.

In a recent paper,¹¹ we have used coupled cluster energy Lagrangians and a Møller-Plesset partitioning of the Hamiltonian to derive a new class of perturbative triple and higher-level excitation models where corrections are likewise determined directly to the CCSD energy. In particular, we have described the development of the CCSD(T-*n*) and CCSD(TQ-*n*) perturbation series that converge towards the coupled cluster singles, doubles, and triples¹² (CCSDT) or coupled cluster singles, doubles, triples, and quadruples¹³

(CCSDTQ) energy limits, respectively, depending on what target energy has been used in the derivation. We refer to Ref. 11 for an overview of previous triples (and quadruples) models and for a comparison of the characteristics of these with those of the CCSD(T-*n*) and CCSD(TQ-*n*) series.

In the present paper, we restrict ourselves to reconsider series of corrections to the CCSD energy that are obtained by applying perturbation theory within an EOM-CC framework. Despite the plethora of such models, none of the current perturbation series are size extensive through all orders. In particular, difficulties have been encountered whenever the EOM-CC framework has been used to (non-iteratively) describe relaxation effects in the singles and doubles spaces. We describe in the present paper how such problems may be overcome, and by using a Møller-Plesset partitioning of the Hamiltonian, we propose a novel series of perturbative energy corrections — denoted the EOM(2)PT(*n*) series — based on the EOM-CC ansatz. The devised corrections to the CCSD energy are size extensive to *any* order and converge formally towards the FCI energy limit. A direct comparison of the EOM(2)PT(*n*) and CC(2)PT(*n*) series is performed that reveals the origin of the deficiencies of the latter. The EOM(2)PT(*n*) series is also compared to our recently developed CCSD(TQ-*n*) series,¹¹ stressing that while both series are size extensive, their target energies differ, as this is always the FCI energy for the EOM(2)PT(*n*) series while it is the CCSDTQ energy for the CCSD(TQ-*n*) series.

In Sec. II, the EOM-CC energy and amplitude equations are derived, formulated such that a Møller-Plesset partitioning of the electronic Hamiltonian may be performed, which gives rise to the EOM(2)PT(*n*) series of corrections to the CCSD energy derived in Sec. III. In Sec. IV, the EOM(2)PT(*n*) series is viewed from a variational energy perspective where the left-hand EOM-CC state by construction is biorthonormal to the right-hand EOM-CC state. In Sec. V, a detailed comparison of the EOM(2)PT(*n*), CC(2)PT(*n*), and CCSD(TQ-*n*) series is carried out, while Sec. VI contains a summary and some conclusive remarks. The Appendix examines in detail why the CC(2)PT(*n*) series is not size extensive.

^{a)} Author to whom correspondence should be addressed. Electronic mail: janusje@chem.au.dk

II. EOM-CC ENERGY AND AMPLITUDE EQUATIONS

In EOM-CC theory,¹⁴ the FCI eigenvalue equation may be expressed as

$$\begin{aligned} \hat{H} \exp(*\hat{T}) \left(1 + \sum_{k=1} \hat{\mathcal{R}}_k \right) |\text{HF}\rangle \\ = E_{\text{FCI}} \exp(*\hat{T}) \left(1 + \sum_{k=1} \hat{\mathcal{R}}_k \right) |\text{HF}\rangle, \end{aligned} \quad (2.1)$$

where the parameters to be determined are the amplitudes of the intermediate normalized right-hand FCI state

$$|\mathcal{R}\rangle = \left(1 + \sum_{k=1} \hat{\mathcal{R}}_k \right) |\text{HF}\rangle. \quad (2.2)$$

$\hat{\mathcal{R}}_k$ is a linear operator that generates all excitations at excitation level k

$$\hat{\mathcal{R}}_k = \sum_{\mu_k} \mathcal{R}_{\mu_k} \hat{t}_{\mu_k}, \quad (2.3)$$

where the greek indices, $\{\mu_i, \nu_j, \eta_k, \dots\}$, are reserved for specific excitations within manifolds at excitation levels, i, j, k, \dots . The summation over index k in Eq. (2.1) runs over all excitation levels. Furthermore, Eq. (2.1) is expressed in terms of the CCSD cluster operator, $*\hat{T}$, which is build from the CCSD singles, $*t_{\mu_1}$, and doubles, $*t_{\mu_2}$, amplitudes

$$*\hat{T} = *\hat{T}_1 + *\hat{T}_2 = \sum_{\mu_1} *t_{\mu_1} \hat{t}_{\mu_1} + \sum_{\mu_2} *t_{\mu_2} \hat{t}_{\mu_2}. \quad (2.4)$$

These amplitudes satisfy the CCSD singles and doubles amplitude equations

$$\langle \mu_i | \hat{H}^{*\hat{T}} | \text{HF} \rangle = 0 \quad i = 1, 2 \quad (2.5)$$

and determine the CCSD energy

$$E_{\text{CCSD}} = \langle \text{HF} | \hat{H}^{*\hat{T}} | \text{HF} \rangle, \quad (2.6)$$

where $\hat{H}^{*\hat{T}}$ is the CCSD similarity-transformed Hamiltonian

$$\hat{H}^{*\hat{T}} = \exp(-*\hat{T}) \hat{H} \exp(*\hat{T}). \quad (2.7)$$

Note that the amplitudes in Eq. (2.4) carry an asterisk and that the asterisk notation will be used to denote the pre-determined CCSD parameters in the following.

From an EOM-CC point of view, the right-hand CCSD ground state becomes $|\mathcal{R}\rangle = |\text{HF}\rangle$, since the corresponding singles and doubles parameters vanish, i.e., $*\mathcal{R}_{\mu_1} = *\mathcal{R}_{\mu_2} = 0$, whereas the left-hand CCSD state

$$\langle *\mathcal{L} | = \sum_{j=1}^2 \langle *\mathcal{L}_j | = \sum_{j=1}^2 \sum_{\nu_j} *\mathcal{L}_{\nu_j} \langle \text{HF} | \hat{t}_{\nu_j}^\dagger \quad (2.8)$$

contains singles and doubles parameters, $\{*\mathcal{L}_{\nu_j}\}$, that satisfy the CCSD singles and doubles left-hand state equations

$$\langle \text{HF} | [\hat{H}^{*\hat{T}}, \hat{t}_{\mu_i}] | \text{HF} \rangle + \langle *\mathcal{L} | [\hat{H}^{*\hat{T}}, \hat{t}_{\mu_i}] | \text{HF} \rangle = 0 \quad (i = 1, 2), \quad (2.9)$$

The CCSD left-hand state parameters, $\{*\mathcal{L}_{\mu_1}, *\mathcal{L}_{\mu_2}\}$, are identical to the CCSD Lagrange multipliers, $\{*\tilde{t}_{\mu_1}, *\tilde{t}_{\mu_2}\}$, of Ref. 11.

Multiplying Eq. (2.1) by $\langle \text{HF} |$ (or equivalently by $\langle \text{HF} | \exp(-*\hat{T})$) gives a closed expression for the FCI energy

$$E_{\text{FCI}} = \langle \text{HF} | \hat{H}^{*\hat{T}} \left(1 + \sum_k \hat{\mathcal{R}}_k \right) | \text{HF} \rangle, \quad (2.10)$$

which — using Eq. (2.6) — becomes

$$E_{\text{FCI}} = E_{\text{CCSD}} + \sum_{k=1}^2 \langle \text{HF} | \hat{H}^{*\hat{T}} \hat{\mathcal{R}}_k | \text{HF} \rangle, \quad (2.11)$$

where the $\hat{\mathcal{R}}$ operator has been restricted to single and double excitations ($k \leq 2$), as the similarity-transformed Hamiltonian, $\hat{H}^{*\hat{T}}$, has the same maximal deexcitation rank as \hat{H} , i.e., two. Further using Eq. (2.9) and that $\langle \text{HF} | \hat{t}_{\mu_k} = 0$, we arrive at

$$E_{\text{FCI}} = E_{\text{CCSD}} - \sum_{k=1}^2 \langle *\mathcal{L} | [\hat{H}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle. \quad (2.12)$$

Equation (2.12) expresses the FCI energy as a sum of the CCSD energy and a correction, which involves the left- and right-hand CCSD states as well as the parameters of the right-hand FCI state.

The FCI parameters may be determined with the right-hand CCSD state as reference point by multiplying Eq. (2.1) by $\langle \mu_i | \exp(-*\hat{T})$ (where i may refer to any excitation level), which yields

$$\langle \mu_i | \hat{H}^{*\hat{T}} \left(1 + \sum_k \hat{\mathcal{R}}_k \right) | \text{HF} \rangle = E_{\text{FCI}} \mathcal{R}_{\mu_i}. \quad (2.13)$$

Substituting Eq. (2.10) into Eq. (2.13) gives

$$(\langle \mu_i | - \mathcal{R}_{\mu_i} \langle \text{HF} |) \hat{H}^{*\hat{T}} \left(1 + \sum_k \hat{\mathcal{R}}_k \right) | \text{HF} \rangle = 0 \quad (2.14)$$

which may be rewritten as

$$\begin{aligned} \langle \mu_i | \hat{H}^{*\hat{T}} | \text{HF} \rangle - \mathcal{R}_{\mu_i} \langle \text{HF} | \hat{H}^{*\hat{T}} | \text{HF} \rangle \\ + \sum_k \{ \langle \mu_i | [\hat{H}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle + \langle \mu_i | \hat{\mathcal{R}}_k \hat{H}^{*\hat{T}} | \text{HF} \rangle \\ - \mathcal{R}_{\mu_i} \langle \text{HF} | [\hat{H}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle \} = 0. \end{aligned} \quad (2.15)$$

Using the expansion

$$\begin{aligned} \sum_k \langle \mu_i | \hat{\mathcal{R}}_k \hat{H}^{*\hat{T}} | \text{HF} \rangle \\ = \sum_{k < i} \langle \mu_i | \hat{\mathcal{R}}_k \hat{H}^{*\hat{T}} | \text{HF} \rangle + \mathcal{R}_{\mu_i} \langle \text{HF} | \hat{H}^{*\hat{T}} | \text{HF} \rangle, \end{aligned} \quad (2.16)$$

the amplitude equations in Eq. (2.14) may be written in the compact form

$$\begin{aligned} \langle \mu_i | \hat{H}^{*\hat{T}} | \text{HF} \rangle + \sum_k \langle \mu_i | [\hat{H}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle \\ + \sum_{k < i} \langle \mu_i | \hat{\mathcal{R}}_k \hat{H}^{*\hat{T}} | \text{HF} \rangle - \mathcal{R}_{\mu_i} \sum_k \langle \text{HF} | [\hat{H}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle = 0. \end{aligned} \quad (2.17)$$

The FCI energy may thus be obtained by using right-hand state amplitudes obtained from Eq. (2.17).

III. PERTURBATION EXPANSION OF THE EOM-CC ENERGY AND AMPLITUDE EQUATIONS: THE EOM(2)PT(n) SERIES

We now describe how perturbation theory may be applied to the EOM-CC ansatz in order to derive a series of size-extensive corrections to the CCSD energy that formally converges towards the FCI energy. From the amplitude equation in Eq. (2.17), we will determine an order expansion of the FCI amplitudes, keeping the right-hand CCSD state as our reference point, while from Eq. (2.12), energy corrections are determined by using all of the CCSD parameters (the energy as well as the left- and right-hand states) of our reference point. We first perform a Møller-Plesset partitioning of the CCSD similarity-transformed Hamiltonian

$$\begin{aligned}\hat{H}^{*\hat{T}} &= \hat{f}^{*\hat{T}} + \hat{\Phi}^{*\hat{T}} \\ &= \hat{f} + \sum_{i=1}^2 \sum_{\mu_i}^* t_{\mu_i} \epsilon_{\mu_i} \hat{t}_{\mu_i} + \hat{\Phi}^{*\hat{T}},\end{aligned}\quad (3.1)$$

where \hat{f} is the Fock operator, $\hat{\Phi}$ the fluctuation potential, and ϵ_{μ_i} denotes the difference in orbital energies between the virtual and occupied spin orbitals of excitation μ_i . $\hat{f}^{*\hat{T}}$ is considered as the zeroth-order operator and $\hat{\Phi}^{*\hat{T}}$ as the perturbation. By introducing the partitioning scheme in Eq. (3.1) into Eq. (2.17), we arrive at an expression for the wave function correction

$$\begin{aligned}\epsilon_{\mu_i} \mathcal{R}_{\mu_i} &= -\langle \mu_i | \hat{\Phi}^{*\hat{T}} | \text{HF} \rangle (1 - \delta_{i1} - \delta_{i2}) \\ &\quad - \sum_{k=1}^{i+2} \langle \mu_i | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle - \sum_{k=1}^{i-3} \langle \mu_i | \hat{\mathcal{R}}_k \hat{\Phi}^{*\hat{T}} | \text{HF} \rangle \\ &\quad + \mathcal{R}_{\mu_i} \sum_{k=1}^2 \langle \text{HF} | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle.\end{aligned}\quad (3.2)$$

In the first summation (second term) of Eq. (3.2), k is restricted to be less than or equal to $i + 2$, as the operator $[\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k]$ must be of excitation rank i and the minimal excitation rank of $\hat{\Phi}^{*\hat{T}}$ is -2 . The upper limit on the second summation (third term) in Eq. (3.2) arises from noting that nonvanishing contributions to this term occur only when $i - k > 2$, due to the CCSD amplitude equation in Eq. (2.5). Furthermore, δ_{i1} and δ_{i2} in Eq. (3.2) are Kronecker deltas that take into account the CCSD amplitude equation for the first term.

The right-hand state parameters may now be expanded in orders of the fluctuation potential

$$\mathcal{R}_{\mu_i} = \mathcal{R}_{\mu_i}^{(0)} + \mathcal{R}_{\mu_i}^{(1)} + \mathcal{R}_{\mu_i}^{(2)} + \dots, \quad (3.3)$$

where, from Eq. (3.2), we note that

$$\mathcal{R}_{\mu_i}^{(0)} = 0. \quad (3.4)$$

To understand which excitation levels contribute to a given order in the perturbation, we note that the CCSD similarity-transformed Hamiltonian can have an excitation rank of at most six (hextuple excitations), originating from the termination of the Baker-Campbell-Hausdorff expansion of $\hat{H}^{*\hat{T}}$ after

five terms.¹⁵ Through second order, the amplitudes thus read

$$\mathcal{R}_{\mu_i}^{(1)} = -\epsilon_{\mu_i}^{-1} \langle \mu_i | \hat{\Phi}^{*\hat{T}} | \text{HF} \rangle \quad (i = 3, \dots, 6), \quad (3.5)$$

$$\begin{aligned}\mathcal{R}_{\mu_i}^{(2)} &= -\epsilon_{\mu_i}^{-1} \sum_{k=3}^6 \langle \mu_i | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k^{(1)}] | \text{HF} \rangle \quad (i = 1, \dots, 10) \\ &\quad - \epsilon_{\mu_i}^{-1} \sum_{k=3}^6 \langle \mu_i | \hat{\mathcal{R}}_k^{(1)} \hat{\Phi}^{*\hat{T}} | \text{HF} \rangle \quad (i = 6, \dots, 12),\end{aligned}\quad (3.6)$$

while the third-order amplitudes are given as

$$\begin{aligned}\mathcal{R}_{\mu_i}^{(3)} &= -\epsilon_{\mu_i}^{-1} \sum_{k=1}^{12} \langle \mu_i | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k^{(2)}] | \text{HF} \rangle \quad (i = 1, \dots, 16) \\ &\quad - \epsilon_{\mu_i}^{-1} \sum_{k=1}^{12} \langle \mu_i | \hat{\mathcal{R}}_k^{(2)} \hat{\Phi}^{*\hat{T}} | \text{HF} \rangle \quad (i = 4, \dots, 18).\end{aligned}\quad (3.7)$$

Equations (3.5) through (3.7) show how relaxation in the singles and doubles amplitude spaces is introduced beyond first-order amplitudes and how excitations of increasingly higher excitation levels enter the perturbative expansion of \mathcal{R}_{μ_i} . We also note that the first-order amplitudes, $\mathcal{R}_{\mu_i}^{(1)}$, only contain connected contributions, while the second-order amplitudes, $\mathcal{R}_{\mu_i}^{(2)}$, contain connected contributions only for $i < 6$. Disconnected contributions arise from the last term in Eq. (3.6) and will start to contribute for $i \geq 6$. Similarly, the connected-only contributions to the third-order amplitudes, $\mathcal{R}_{\mu_i}^{(3)}$, are limited to $i < 4$ due to the second term in Eq. (3.7).

In order to derive a perturbation series of corrections to the CCSD energy that formally converges to the FCI energy, we rewrite the last term in Eq. (2.12) such that only the fluctuation potential is referenced. First, Eq. (3.1) is inserted into the second term of Eq. (2.12), giving

$$E_{\text{FCI}} = E_{\text{CCSD}} - \sum_{k=1}^2 \langle * \mathcal{L} | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle - \sum_{j=1}^2 \sum_{v_j}^* \mathcal{L}_{v_j} \epsilon_{v_j} \mathcal{R}_{v_j}. \quad (3.8)$$

With the further use of the expression for $\epsilon_{v_j} \mathcal{R}_{v_j}$ from Eq. (3.2) (for $j = 1, 2$), the FCI energy becomes

$$\begin{aligned}E_{\text{FCI}} &= E_{\text{CCSD}} - \sum_{k=1}^2 \langle * \mathcal{L} | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle \\ &\quad + \sum_{k=1}^4 \langle * \mathcal{L} | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle \\ &\quad - \sum_{j=1}^2 \sum_{v_j}^* \mathcal{L}_{v_j} \mathcal{R}_{v_j} \sum_{k=1}^2 \langle \text{HF} | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle \\ &= E_{\text{CCSD}} + \sum_{k=3}^4 \langle * \mathcal{L} | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle \\ &\quad - \sum_{j=1}^2 \sum_{v_j}^* \mathcal{L}_{v_j} \mathcal{R}_{v_j} \sum_{k=1}^2 \langle \text{HF} | [\hat{\Phi}^{*\hat{T}}, \hat{\mathcal{R}}_k] | \text{HF} \rangle.\end{aligned}\quad (3.9)$$

Note that the first and third term of Eq. (3.2) do not contribute to Eq. (3.9) due to the restricted summation over single and double excitations. From the first- and second-order amplitudes in Eqs. (3.5) and (3.6), we furthermore recognize that only the second term in Eq. (3.9) will contribute to energy corrections up to fourth order, as the third term involves a product of doubles amplitudes and therefore first enters at fifth order.

In the following, we will consider only energy corrections up to fourth order. The energy corrections obtained from Eq. (3.9) lead to a perturbation series that obeys an $n + 1$ rule for the right-hand state amplitudes. The first-order correction vanishes

$$E^{(1)} = 0 \quad (3.10)$$

and the second-order energy correction becomes

$$E^{(2)} = \sum_{k=3}^4 \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{\mathcal{R}}_k^{(1)}] | \text{HF} \rangle, \quad (3.11)$$

where only triple and quadruple excitations contribute. In a similar way, we may write the third- and fourth-order energy corrections as

$$E^{(3)} = \sum_{k=3}^4 \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{\mathcal{R}}_k^{(2)}] | \text{HF} \rangle, \quad (3.12)$$

$$E^{(4)} = \sum_{k=3}^4 \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{\mathcal{R}}_k^{(3)}] | \text{HF} \rangle. \quad (3.13)$$

Whereas the third-order energy correction does not contain any singles and doubles contributions, it does include contributions from quintuples and hexuples in addition to triples and quadruples, as may be recognized from expanding Eq. (3.12), using the second-order amplitudes of Eq. (3.6)

$$E^{(3)} = - \sum_{k=3}^4 \sum_{\mu_k} \sum_{l=3}^6 \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{t}_{\mu_k}] | \text{HF} \rangle \epsilon_{\mu_k}^{-1} \langle \mu_k | [\hat{\Phi}^{*T}, \hat{\mathcal{R}}_l^{(1)}] | \text{HF} \rangle. \quad (3.14)$$

Note that the second term of Eq. (3.6) does not contribute to Eq. (3.14), as this term only contributes to second-order amplitudes of excitation levels six through twelve.

Unlike the lower-order corrections, the fourth-order correction does contain, in addition to contributions from all excitation levels ranging from triples to hexuples, contributions that arise from relaxation effects in the singles and doubles spaces. Noteworthy, the fourth-order correction is also the first level at which disconnected contributions start to appear. This may be realized by expanding Eq. (3.13), using the third-order amplitudes of Eq. (3.7)

$$E^{(4)} = - \sum_{k=3}^4 \sum_{\mu_k} \sum_{l=1}^6 \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{t}_{\mu_k}] | \text{HF} \rangle \epsilon_{\mu_k}^{-1} \langle \mu_k | [\hat{\Phi}^{*T}, \hat{\mathcal{R}}_l^{(2)}] | \text{HF} \rangle - \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{t}_{\mu_4}] | \text{HF} \rangle \epsilon_{\mu_4}^{-1} \langle \mu_4 | \hat{\mathcal{R}}_1^{(2)} \hat{\Phi}^{*T} | \text{HF} \rangle. \quad (3.15)$$

From the last term in Eq. (3.15), we recognize that disconnected contributions to $E^{(4)}$ arise from $\hat{\mathcal{R}}_4^{(3)}$. Further expanding Eq. (3.15) by using Eq. (3.6) gives

$$\begin{aligned} E^{(4)} = & \sum_{k=3}^4 \sum_{\mu_k} \sum_{l=1}^6 \sum_{v_l} \sum_{i=3}^6 \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{t}_{\mu_k}] | \text{HF} \rangle \epsilon_{\mu_k}^{-1} \langle \mu_k | [\hat{\Phi}^{*T}, \hat{t}_{v_l}] | \text{HF} \rangle \epsilon_{v_l}^{-1} \langle v_l | [\hat{\Phi}^{*T}, \hat{\mathcal{R}}_i^{(1)}] | \text{HF} \rangle \\ & + \sum_{\mu_4} \sum_{v_6} \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{t}_{\mu_4}] | \text{HF} \rangle \epsilon_{\mu_4}^{-1} \langle \mu_4 | [\hat{\Phi}^{*T}, \hat{t}_{v_6}] | \text{HF} \rangle \epsilon_{v_6}^{-1} \langle v_6 | \hat{\mathcal{R}}_3^{(1)} \hat{\Phi}^{*T} | \text{HF} \rangle \\ & + \sum_{\mu_4} \sum_{v_1} \langle * \mathcal{L} | [\hat{\Phi}^{*T}, \hat{t}_{\mu_4}] | \text{HF} \rangle \epsilon_{\mu_4}^{-1} \langle \mu_4 | \hat{t}_{v_1} \hat{\Phi}^{*T} | \text{HF} \rangle \epsilon_{v_1}^{-1} \langle v_1 | [\hat{\Phi}^{*T}, \hat{\mathcal{R}}_3^{(1)}] | \text{HF} \rangle. \end{aligned} \quad (3.16)$$

Obviously, the last two terms in Eq. (3.16) contain disconnected contributions. Similar terms will continue to appear when moving to higher orders in the perturbation series.

The energy corrections in Eqs. (3.11) through (3.13) constitute the lowest-order corrections to the CCSD energy of a perturbation series based on EOM-CC theory, which we will denote as the EOM(2)PT(n) series. As the EOM(2)PT(n) series is formulated in terms of an additively separable zeroth-order Hamiltonian — the similarity-transformed Fock operator (cf. Eq. (3.1)) — the series is size extensive through all orders in the perturbation. This size extensivity is not term-wise, though, since the EOM-CC framework does not provide a formulation in which all contributions are connected, and the statement about size extensivity therefore only holds in the limit where all terms through a given order are retained. In particular, this implies that if the excitations levels in the series are pragmatically restricted to the level of, e.g.,

quadruples, the size extensivity is in general lost. However, for $E^{(3)}$ in Eq. (3.14), such a truncation does not pose a problem with respect to size extensivity as the correction is expressed in a closed commutator form — thereby ensuring term-wise size extensivity¹⁵ — both prior to and after the restriction is made to the level of quadruple excitations. For $E^{(4)}$ in its form in Eq. (3.16), on the other hand, more careful considerations are called for; for example, we may obtain a size-extensive energy correction if the last two terms are neglected as the first term is expressed solely in terms of commutators of connected quantities. As continually more disconnected terms will appear at higher orders, it will become exceedingly more difficult to ensure size extensivity upon a limitation of the involved excitation levels in the series as certain terms will have to be cherry-picked from the resulting energy corrections. In addition, for such series it is unclear what the target energy becomes. In Sec. V, we will compare the EOM(2)PT(n) series to

the related CC(2)PT(n) series of Ref. 8 and the CCSD(TQ- n) series of Ref. 11.

IV. THE LAGRANGIAN EOM-CC ANSATZ

In the present section, we show that the EOM(2)PT(n) series may equivalently be derived from the EOM-CC variational energy expression. In particular, we show how the EOM-CC energy expression can be expressed in a form where the left-hand EOM-CC state per construction is biorthonormal to the right-hand EOM-CC state and use this to obtain a Lagrangian for the EOM-CC energy and amplitude equations of Sec. II.

To obtain a variational energy expression in the conventional EOM-CC picture, Eq. (2.1) is multiplied from the left by $\langle \text{HF} | (1 + \sum_{j=1} \hat{\mathcal{L}}_j) \exp(-\hat{T})$, giving

$$E_{\text{FCI}} = \frac{\langle \text{HF} | (1 + \sum_{j=1} \hat{\mathcal{L}}_j) \hat{H}^{*\hat{T}} (1 + \sum_{i=1} \hat{\mathcal{R}}_i) | \text{HF} \rangle}{\langle \text{HF} | (1 + \sum_{j=1} \hat{\mathcal{L}}_j) (1 + \sum_{i=1} \hat{\mathcal{R}}_i) | \text{HF} \rangle}. \quad (4.1)$$

For convenience, we have assumed that an intermediate normalization is also invoked for the left-hand EOM-CC state. In the same way as for $\hat{\mathcal{R}}_k$ in Eq. (2.3), the operator $\hat{\mathcal{L}}_j$ is a linear operator that contains all deexcitations within a manifold at excitation level j

$$\hat{\mathcal{L}}_j = \sum_{v_j} \mathcal{L}_{v_j} \hat{t}_{v_j}^\dagger. \quad (4.2)$$

The left- and right-hand states in Eq. (4.1), $\langle \mathcal{L}_{\text{EOM}} | = \langle \text{HF} | (1 + \sum_{j=1} \hat{\mathcal{L}}_j)$ and $|\mathcal{R}\rangle = (1 + \sum_{i=1} \hat{\mathcal{R}}_i) | \text{HF} \rangle$, respectively, satisfy the relation

$$\langle \mathcal{L}_{\text{EOM}} | \mathcal{R} \rangle = 1 + \sum_j \sum_{v_j} \mathcal{L}_{v_j} \mathcal{R}_{v_j}. \quad (4.3)$$

The left-hand state, $\langle \mathcal{L}_{\text{EOM}} |$, describes an intermediate normalized state expanded in the basis

$$\mu_{\text{EOM}} = \{ \langle \text{HF} |, \langle \mu_1 |, \langle \mu_2 |, \dots \}. \quad (4.4)$$

To simplify the variational expression in Eq. (4.1), we may choose to reparametrize the left-hand state such that it has a unit overlap with the right-hand state. We thus introduce the new left-hand state, $\langle \mathcal{L} |$

$$\langle \mathcal{L} | = \langle \text{HF} | + \sum_i \sum_{v_i} \mathcal{L}_{\mu_i} (\langle \mu_i | - \mathcal{R}_{\mu_i} \langle \text{HF} |) \quad (4.5)$$

which has the required unit overlap with $|\mathcal{R}\rangle$

$$\langle \mathcal{L} | \mathcal{R} \rangle = 1. \quad (4.6)$$

In order to replace $\langle \mathcal{L}_{\text{EOM}} |$ with $\langle \mathcal{L} |$, we recognize that $\langle \mathcal{L} |$ is expanded in the basis

$$\tilde{\mu}_{\text{EOM}} = \{ \langle \text{HF} |, \langle \mu_1 | - \mathcal{R}_{\mu_1} \langle \text{HF} |, \langle \mu_2 | - \mathcal{R}_{\mu_2} \langle \text{HF} |, \dots \}. \quad (4.7)$$

As the μ_{EOM} basis in Eq. (4.4) is connected to the $\tilde{\mu}_{\text{EOM}}$ basis in Eq. (4.7) via a non-singular transformation, the solution to the EOM-CC problem in Eq. (4.1) does not change upon the reparametrization of the left-hand state. By replacing $\langle \mathcal{L}_{\text{EOM}} |$ with $\langle \mathcal{L} |$, the EOM-CC variational energy expression

thus simplifies as

$$\begin{aligned} E_{\text{FCI}} &\equiv L(\mathcal{L}, \mathcal{R}) \\ &= \langle \text{HF} | \hat{H}^{*\hat{T}} \left(1 + \sum_k \hat{\mathcal{R}}_k \right) | \text{HF} \rangle \\ &+ \sum_i \sum_{\mu_i} \mathcal{L}_{\mu_i} (\langle \mu_i | - \mathcal{R}_{\mu_i} \langle \text{HF} |) \hat{H}^{*\hat{T}} \left(1 + \sum_k \hat{\mathcal{R}}_k \right) | \text{HF} \rangle. \end{aligned} \quad (4.8)$$

Equation (4.8) is the Lagrangian for the FCI energy in Eq. (2.10), hence the $L(\mathcal{L}, \mathcal{R})$ notation, where the amplitude equations in Eq. (2.14) are added via a set of undetermined multipliers, $\{\mathcal{L}_{\mu_i}\}$. Determining energy corrections from the Lagrangian in Eq. (4.8) is therefore equivalent to determining energy corrections from the energy and amplitude equations of Sec. II. However, by using the Lagrangian form, left-hand state parameters (i.e., the Lagrange multipliers) are introduced, allowing for the energy corrections of the EOM(2)PT(n) series to be derived by means of a $2n + 1$ rule for the right-hand state amplitudes and a $2n + 2$ rule for the left-hand state parameters. As the EOM(2)PT(n) series may be derived from the Lagrangian in Eq. (4.8), we note that the energy corrections of the series can be viewed as obtained from a perturbative solution to the EOM-CC variational problem subjected to biorthonormality as a constraint.

V. PERTURBATIVE CORRECTIONS TO THE CCSD ENERGY

To understand the similarities and differences between the present EOM(2)PT(n) series and the CC(2)PT(n)^{8,9} and CCSD(TQ- n)¹¹ series, we briefly summarize the derivations leading to the latter two series. In Secs. V A and V B, the basic equations of the CC(2)PT(n) and CCSD(TQ- n) series are summarized, respectively, followed by the comparison of all three series in Sec. V C.

A. Short summary of the CC(2)PT(n) series

In Ref. 8, Hirata *et al.* used the EOM-CC framework to derive the CC(2)PT(n) perturbation series. In deriving this series, Eq. (2.1) is multiplied by $\exp(-\hat{T})$ from the left, giving

$$\hat{H}^{*\hat{T}} \left(1 + \sum_{k=1} \hat{\mathcal{R}}_k \right) | \text{HF} \rangle = E_{\text{FCI}} \left(1 + \sum_{k=1} \hat{\mathcal{R}}_k \right) | \text{HF} \rangle, \quad (5.1)$$

which contains the CCSD similarity-transformed Hamiltonian, $\hat{H}^{*\hat{T}}$, of Eq. (2.7). By partitioning the CCSD similarity-transformed Hamiltonian into a zeroth-order part, $\hat{H}^{(0)}$, and a perturbation, $\hat{H}^{(1)}$, while treating the CCSD amplitudes as zeroth-order parameters, a Rayleigh-Schrödinger perturbation expansion (RSPT) is derived. Expanding the energy and the right state parameters in Eq. (2.2) in orders of the perturbation then gives the n th order equation

$$\hat{H}^{(0)} | \mathcal{R}^{(n)} \rangle + \hat{H}^{(1)} | \mathcal{R}^{(n-1)} \rangle = \sum_{p=0}^n E^{(p)} | \mathcal{R}^{(n-p)} \rangle. \quad (5.2)$$

By requiring that $\hat{H}^{(0)}|*\mathcal{R}\rangle = E_{\text{CCSD}}|*\mathcal{R}\rangle$ and imposing an intermediate normalization of the right-hand state with respect to the left-hand CCSD state in Eq. (2.8), we obtain

$$\langle *\mathcal{L}|\mathcal{R}^{(p)}\rangle = 0 \quad (5.3)$$

for $p \geq 1$. Multiplying Eq. (5.2) by $\langle *\mathcal{L}|$ gives

$$E^{(n)} = \langle *\mathcal{L}|\hat{H}^{(1)}|\mathcal{R}^{(n-1)}\rangle \quad n \geq 1, \quad (5.4)$$

where the wave function correction to the right-hand state is obtained from the recursion relation

$$(E_{\text{CCSD}} - \hat{H}^{(0)})|\mathcal{R}^{(p)}\rangle = \hat{H}^{(1)}|\mathcal{R}^{(p-1)}\rangle - \sum_{j=1}^p E^{(j)}|\mathcal{R}^{(p-j)}\rangle. \quad (5.5)$$

The RSPT-based corrections to the CCSD energy are thus obtained using an $n + 1$ rule for the right-hand state amplitudes, as in the derivation of the EOM(2)PT(n) series in Sec. III.

In the context of deriving corrections to the EOM-CCSD energy, the HF state and the space generated by all singly and doubly excited determinants from the HF reference is denoted the primary CCSD space, while the complementary space, made up from all triply, quadruply, etc., excited determinants, is denoted the secondary CCSD space. The primary and secondary spaces may be generated by projection operators \hat{P} and $\hat{Q} = \hat{I} - \hat{P}$, respectively. When the zeroth- and first-order Hamiltonians entering Eqs. (5.4) and (5.5) are defined as in Ref. 8

$$\hat{H}^{(0)} = \hat{P}\hat{H}^*\hat{T}\hat{P} + \hat{Q}\left(E_{\text{CCSD}} + \hat{f} - \sum_i^{\text{occ}} \epsilon_i\right)\hat{Q}, \quad (5.6a)$$

$$\hat{H}^{(1)} = \hat{H}^*\hat{T} - \hat{H}^{(0)}, \quad (5.6b)$$

the CC(2)PT(n) series is obtained.

B. Short summary of the CCSD(TQ- n) series

The CCSD(TQ- n) series¹¹ is formulated in terms of a CCSDTQ energy Lagrangian,¹⁵ which is parametrized using the CCSD energy as reference point

$$\begin{aligned} L_{\text{TQ}}(*\mathbf{t}, *\mathcal{L}, \delta\mathbf{t}, \delta\mathcal{L}) &= \langle \text{HF} | \exp(-\delta\hat{T})\hat{H}^*\hat{T} \exp(\delta\hat{T}) | \text{HF} \rangle \\ &+ \sum_{i=1}^2 \langle *\mathcal{L}_i + \delta\mathcal{L}_i | \exp(-\delta\hat{T})\hat{H}^*\hat{T} \exp(\delta\hat{T}) | \text{HF} \rangle \\ &+ \sum_{j=3}^4 \langle \delta\mathcal{L}_j | \exp(-\delta\hat{T})\hat{H}^*\hat{T} \exp(\delta\hat{T}) | \text{HF} \rangle, \quad (5.7) \end{aligned}$$

where we again stress the equivalence between the $\{*\mathcal{L}_\mu, \delta\mathcal{L}_\mu\}$ amplitudes of Eq. (5.7) and the Lagrange multipliers, $\{*\tilde{t}_\mu, \delta\tilde{t}_\mu\}$, of Ref. 11. The CCSD similarity-transformed Hamiltonian is partitioned using the Møller-Plesset partitioning scheme in Eq. (3.1), considering again the fluctuation potential as the perturbation. Perturbative corrections to the CCSD energy, which define the CCSD(TQ- n) series, may then be determined by expanding the Lagrangian in orders of the fluctuation potential with the CCSD energy as

expansion point, resulting in energy corrections that converge towards the CCSDTQ energy. In order to obtain these energy corrections, the amplitudes, $\{\delta t_\mu\}$, and multipliers, $\{\delta \mathcal{L}_\mu\}$, of Eq. (5.7) are expanded in terms of the fluctuation potential. By means of the CCSD amplitude and multiplier equations, the CCSDTQ Lagrangian can be simplified and corrections to the CCSD energy obtained in accordance with Wigner's $2n + 1$ and $2n + 2$ rules for the amplitudes and multipliers, respectively. All energy corrections of the CCSD(TQ- n) series are given in terms of closed commutator expressions and are thus term-wise size extensive.

C. Comparison of the EOM(2)PT(n), CC(2)PT(n), and CCSD(TQ- n) series

The three series, EOM(2)PT(n), CC(2)PT(n), and CCSD(TQ- n), all determine corrections to the CCSD energy by means of perturbation theory. In the EOM(2)PT(n) and CC(2)PT(n) series, the EOM-CC formalism is used to determine corrections that target the FCI energy. The linear parametrizations of the wave functions in EOM-CC theory imply that a truncation of the left- and right-hand states at any given excitation level, e.g., quadruples, does not lead to a target energy for the series that equals the CC energy of the given excitation level. A perturbation series that converges towards the CCSDTQ energy thus cannot be derived within the EOM-CC framework. For the CCSD(TQ- n) series, however, the corrections to the CCSD energy do indeed target the CCSDTQ energy, as the right-hand state is exponentially parametrized while a linear parametrization is invoked only for the left-hand state.

In order to obtain a size-extensive perturbation series, the zeroth-order Hamiltonian must be additively separable. For both the EOM(2)PT(n) and CCSD(TQ- n) series, this is clearly satisfied from the use of the CCSD similarity-transformed Fock operator, $\hat{f}^*\hat{T}$, as $\hat{H}^{(0)}$. Thus, both series are size extensive to any order. In fact, the CCSD(TQ- n) series is even term-wise size extensive as all terms in this series are expressed in terms of commutators that contain only connected quantities. On the contrary, as stated in Ref. 11, the CC(2)PT(n) zeroth-order Hamiltonian in Eq. (5.6a) does not separate additively due to the presence of projection operators in its $\hat{P}\hat{H}^*\hat{T}\hat{P}$ term.¹⁶ In the Appendix, we demonstrate explicitly how the involved projection operators make $\hat{H}^{(0)}$ non-separable, and the energy corrections of the CC(2)PT(n) series, which reference this particular $\hat{P}\hat{H}^*\hat{T}\hat{P}$ term, are therefore not size extensive.

Inspecting the energy corrections of the EOM(2)PT(n), CC(2)PT(n), and CCSD(TQ- n) series, we first note that the second-order corrections of all three series are identical, containing only contributions from triple and quadruple excitations. At third order, the energy corrections of the EOM(2)PT(n) and CC(2)PT(n) series are identical and differ from the CCSD(TQ-3) corrections by containing also contributions from quintuple and hextuple excitations. However, if the EOM(2)PT(n) and CC(2)PT(n) series are truncated at the level of quadruple excitations, all three series give identical third-order energy corrections (see the discussion at the end of Sec. III). At fourth order, contributions from singles

and doubles relaxation enter any of the series, and by inspecting the three series it is observed how the energy corrections differ, irrespective of whether the CC(2)PT(n) and EOM(2)PT(n) series are truncated at the level of quadruples or not. This difference has three distinct origins; first and foremost, the elements of the CCSD(TQ-4) model differ from the EOM(2)PT(4) and CC(2)PT(4) elements as a linear parametrization is used for both the left- and right-hand states in the EOM-CC theory-based series. Second, only excitation levels up to quadruples are considered in the CCSD(TQ-4) model. Finally, the EOM(2)PT(4) and CC(2)PT(4) models differ because different zeroth-order Hamiltonians are used in the singles and doubles spaces in the two models. In summary, while the fourth-order singles and doubles contributions in the CC(2)PT(n) series are not size extensive, those in the EOM(2)PT(n) series are, however only sum-wise and not term-wise as in the CCSD(TQ- n) series. At fifth and higher order, the energy corrections of the three series will continue to differ; while only the EOM(2)PT(n) and CCSD(TQ- n) series are size extensive, only the corrections of the CCSD(TQ- n) series converge towards the CCSDTQ energy. Finally, it should be noted that the corrections of a CCSD(TQ... - n) model, where TQ... imply all possible excitations, are identical to those of the EOM(2)PT(n) model at any given order, as both models target the FCI energy in this case. However, even in this limit, the distinction between the corrections of the two models in terms of sum-wise (EOM(2)PT(n)) and term-wise (CCSD(TQ... - n)) size extensivity remains.

VI. SUMMARY AND CONCLUSION

The EOM-CC framework has previously been used to derive series of perturbative corrections to the CCSD energy that converge towards the FCI energy limit. However, these series have had fundamental flaws as they have not been size extensive at higher orders in the perturbation. In the present paper, we have shown how it is possible to derive a size-extensive perturbation theory scheme starting from the EOM-CC ansatz and using a Møller-Plesset partitioning of the Hamiltonian. In particular, we have derived the size-extensive EOM(2)PT(n) perturbation series of corrections to the CCSD energy that converges towards the FCI energy. We have also demonstrated how the size extensivity problems that occur in previously derived series arise due to a partitioning of the Hamiltonian into a zeroth-order part that is non-separable, thereby violating a fundamental requirement for obtaining a size-extensive perturbation series.

The EOM(2)PT(n) series has also been compared to our recently developed CCSD(TQ- n) series. Both of these series use a Møller-Plesset partitioning of the Hamiltonian which guarantees size extensivity through all orders since an additively separable zeroth-order Hamiltonian is used. The major drawback of the EOM(2)PT(n) series and all other schemes based on EOM-CC theory is that they can only target the FCI energy limit, and thus not the energies of any truncated CC model. On the contrary, the coupled cluster Lagrangian-based CCSD(T- n) and CCSD(TQ- n) series will target the CCSDT and CCSDTQ energies, respectively. This is of no relevance

in lowest order, i.e., second order, but affects all of the higher-order energy corrections.

ACKNOWLEDGMENTS

J.J.E. and P.J. acknowledge support from The European Research Council under the European Union's (EU) Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. 291371 and J.J.E., P.J., and J.O. acknowledge support from The Danish Council for Independent Research – Natural Sciences. J.G. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG) (DFG GA 370/5-1).

APPENDIX: A PROJECTION OPERATOR-BASED ZERO-ORDER HAMILTONIAN

In the present appendix, we demonstrate explicitly that the zeroth-order Hamiltonian, $\hat{H}^{(0)}$, in Eq. (5.6a) is not additively separable. To do so, we consider two non-interacting systems, denoted subsystems A and B , and show how the sum of the zeroth-order Hamiltonians, $\hat{H}_A^{(0)} + \hat{H}_B^{(0)}$, for the two subsystems differs from the zeroth-order Hamiltonian, $\hat{H}_{AB}^{(0)}$, for the compound system AB . To simplify the proof, we restrict $\hat{H}^{(0)}$ to its primary space part, $\hat{P}\hat{H}^{(0)}\hat{P}$, noting that the total zeroth-order Hamiltonian will not separate additively if its primary space part does not.

For subsystem A , we may write the CCSD similarity-transformed Hamiltonian in Eq. (2.7) as

$$\hat{H}_A^{*\hat{T}_A} = \exp(-*\hat{T}_A)\hat{H}_A\exp(*\hat{T}_A), \quad (\text{A1})$$

where $*\hat{T}_A = *\hat{T}_{1,A} + *\hat{T}_{2,A}$. The EOM-CCSD eigenvalue equation for the i th state of subsystem A may be expressed as

$$\langle *\mathcal{L}_{i,A} | \hat{H}_A^{*\hat{T}_A} | *\mathcal{R}_{i,A} \rangle = *E_{i,A}, \quad (\text{A2})$$

where $*E_{i,A}$ is the i th eigenvalue and $\{*\mathcal{L}_{i,A}, |*\mathcal{R}_{i,A}\}$ are the corresponding i th left and right eigenvectors. For convenience, we assume that the left and right eigenvectors are biorthonormal

$$\langle *\mathcal{L}_{i,A} | *\mathcal{R}_{j,A} \rangle = \delta_{ij}. \quad (\text{A3})$$

In a matrix notation, Eq. (A3) reads

$$\langle *\mathcal{L}_A | *\mathcal{R}_A \rangle = \mathbf{I}_A. \quad (\text{A4})$$

The projector for the primary space of subsystem A can be expressed as

$$\hat{P}_A = \sum_{k=0}^{N_{\hat{P}_A}-1} |*\mathcal{R}_{k,A}\rangle\langle *\mathcal{L}_{k,A}|, \quad (\text{A5})$$

where the dimension of \hat{P}_A is $N_{\hat{P}_A} = 1 + N_{\mu_1,A} + N_{\mu_2,A}$ in terms of the number of single, $N_{\mu_1,A}$, and double, $N_{\mu_2,A}$, excitations from the HF reference state in subsystem A . The primary space part (p) of the zeroth-order Hamiltonian for subsystem A may thus be written as

$${}^p\hat{H}^{(0),A} = \hat{P}_A\hat{H}_A^{*\hat{T}_A}\hat{P}_A. \quad (\text{A6})$$

Similarly, for subsystem B we have

$${}^p\hat{H}^{(0),B} = \hat{P}_B \hat{H}_B^{*\hat{T}_B} \hat{P}_B, \quad (\text{A7})$$

where \hat{P}_B is defined in analogy with \hat{P}_A in Eq. (A5).

Considering a cluster operator for the compound system that is equal to the sum of the cluster operators for the individual subsystems

$${}^*\hat{T}_{AB} = {}^*\hat{T}_A + {}^*\hat{T}_B, \quad (\text{A8})$$

the CCSD similarity-transformed Hamiltonian for the compound system may be expressed as

$$\begin{aligned} \hat{H}_{AB}^{*\hat{T}_{AB}} &= \exp(-({}^*\hat{T}_A + {}^*\hat{T}_B))(\hat{H}_A + \hat{H}_B) \exp({}^*\hat{T}_A + {}^*\hat{T}_B) \\ &= \hat{H}_A^{*\hat{T}_A} \otimes \mathbf{I}_B + \mathbf{I}_A \otimes \hat{H}_B^{*\hat{T}_B} \end{aligned} \quad (\text{A9})$$

and is thus additively separable. Furthermore, the left and right eigenvectors of $\hat{H}_{AB}^{*\hat{T}_{AB}}$ may in the direct product basis be written as

$$|{}^*\mathcal{R}_{AB}\rangle = |{}^*\mathcal{R}_A\rangle \otimes \mathbf{I}_B + \mathbf{I}_A \otimes |{}^*\mathcal{R}_B\rangle, \quad (\text{A10a})$$

$$\langle{}^*\mathcal{L}_{AB}| = \langle{}^*\mathcal{L}_A| \otimes \mathbf{I}_B + \mathbf{I}_A \otimes \langle{}^*\mathcal{L}_B|, \quad (\text{A10b})$$

with the eigenvectors satisfying the biorthonormality condition

$$\langle{}^*\mathcal{L}_{AB}|{}^*\mathcal{R}_{AB}\rangle = \mathbf{I}_A \otimes \mathbf{I}_B. \quad (\text{A11})$$

The operator which projects onto the eigenvectors of $\hat{H}_{AB}^{*\hat{T}_{AB}}$ is given as

$$\begin{aligned} \hat{P}_{AB} &= \hat{P}_A + \hat{P}_B \\ &= \sum_{k=0}^{N_{\hat{P}_A}-1} |{}^*\mathcal{R}_{k,A}\rangle \langle{}^*\mathcal{L}_{k,A}| + \sum_{l=0}^{N_{\hat{P}_B}-1} |{}^*\mathcal{R}_{l,B}\rangle \langle{}^*\mathcal{L}_{l,B}|. \end{aligned} \quad (\text{A12})$$

For the compound system, the primary space part of the zeroth-order Hamiltonian may be written as

$$\begin{aligned} {}^p\hat{H}_{AB}^{(0)} &= \hat{P}_{AB} \hat{H}^{*\hat{T}_{AB}} \hat{P}_{AB} \\ &= \hat{P}_A \hat{H}_A^{*\hat{T}_A} \hat{P}_A + \hat{P}_B \hat{H}_B^{*\hat{T}_B} \hat{P}_B + \hat{H}_A^{*\hat{T}_A} \hat{P}_B + \hat{P}_A \hat{H}_B^{*\hat{T}_B} \end{aligned} \quad (\text{A13})$$

since \hat{P}_A and \hat{P}_B are idempotent and since

$$\begin{aligned} \hat{P}_A \hat{H}^{*\hat{T}_A} \hat{P}_B &= \sum_{k=0}^{N_{\hat{P}_A}-1} |{}^*\mathcal{R}_{k,A}\rangle \langle{}^*\mathcal{L}_{k,A}| (\hat{H}^{*\hat{T}_A} \otimes \mathbf{I}_B) \sum_{l=0}^{N_{\hat{P}_B}-1} |{}^*\mathcal{R}_{l,B}\rangle \langle{}^*\mathcal{L}_{l,B}| \\ &= \sum_{k=0}^{N_{\hat{P}_A}-1} {}^*E_{k,A} |{}^*\mathcal{R}_{k,A}\rangle \langle{}^*\mathcal{L}_{k,A}| (\mathbf{I}_A \otimes \mathbf{I}_B) \sum_{l=0}^{N_{\hat{P}_B}-1} |{}^*\mathcal{R}_{l,B}\rangle \langle{}^*\mathcal{L}_{l,B}| = 0, \end{aligned} \quad (\text{A14})$$

where Eq. (A11) has been used. Thus, for the primary part of the excitation space, ${}^p\hat{H}_A^{(0)} + {}^p\hat{H}_B^{(0)} \neq {}^p\hat{H}_{AB}^{(0)}$ and $\hat{H}^{(0)}$ as a whole is therefore non-separable.

In the energy corrections at second and third order, the primary space does not enter. As only the secondary space is referenced in the expressions for these corrections, the following zeroth-order Hamiltonian is effectively used

$${}^q\hat{H}^{(0)} = E_{\text{CCSD}} + \hat{f} - \sum_i^{\text{occ}} \epsilon_i. \quad (\text{A15})$$

Since E_{CCSD} , \hat{f} , and $\sum_i^{\text{occ}} \epsilon_i$ are all additively separable quantities, the resulting second- and third-order energy corrections are size extensive.

¹J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966); J. Paldus, J. Čížek, and I. Shavitt, *Phys. Rev. A* **5**, 50 (1972).

²G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).

³I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: Many-Body Perturbation Theory and Coupled-Cluster Theory* (Cambridge University Press, Cambridge, UK, 2009).

⁴K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).

⁵J. F. Stanton, *Chem. Phys. Lett.* **281**, 130 (1997).

⁶S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **108**, 5243 (1998); **108**, 9221 (1998); A. G. Taube and R. J. Bartlett, *ibid.* **128**, 044110 (2008); **128**, 044111 (2008).

⁷T. D. Crawford and J. F. Stanton, *Int. J. Quantum Chem.* **70**, 601 (1998).

⁸S. Hirata, M. Nooijen, I. Grabowski, and R. J. Bartlett, *J. Chem. Phys.* **114**, 3919 (2001); S. Hirata, P.-D. Fan, A. A. Auer, M. Nooijen, and P. Piecuch, *ibid.* **121**, 12197 (2004); T. Shiozaki, K. Hirao, and S. Hirata, *ibid.* **126**, 244106 (2007).

⁹S. R. Gwaltney and M. Head-Gordon, *Chem. Phys. Lett.* **323**, 21 (2000); *J. Chem. Phys.* **115**, 2014 (2001).

¹⁰K. Kowalski and P. Piecuch, *J. Chem. Phys.* **113**, 18 (2000); P. Piecuch and M. Włoch, *ibid.* **123**, 224105 (2005); P. Piecuch, M. Włoch, J. R. Gour, and A. Kinal, *Chem. Phys. Lett.* **418**, 467 (2006).

¹¹J. J. Eriksen, K. Kristensen, T. Kjærgaard, P. Jørgensen, and J. Gauss, *J. Chem. Phys.* **140**, 064108 (2014).

¹²J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987); G. E. Scuseria and H. F. Schaefer, *Chem. Phys. Lett.* **152**, 382 (1988).

¹³N. Oliphant and L. Adamowicz, *J. Chem. Phys.* **95**, 6645 (1991); S. A. Kucharski and R. J. Bartlett, *ibid.* **97**, 4282 (1992).

¹⁴J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.* **98**, 7029 (1993).

¹⁵T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*, 1st ed. (Wiley & Sons, Ltd., West Sussex, England, 2000).

¹⁶A similar issue is already known from CASPT(2) theory, cf. Ref. 15.