The divide-expand-consolidate family of coupled cluster methods: Numerical illustrations using second order Møller-Plesset perturbation theory

Ida-Marie Høyvik, Kasper Kristensen, Branislav Jansik, and Poul Jørgensen

Citation: The Journal of Chemical Physics **136**, 014105 (2012); doi: 10.1063/1.3667266
View online: http://dx.doi.org/10.1063/1.3667266
View Table of Contents: http://aip.scitation.org/toc/jcp/136/1
Published by the American Institute of Physics

The divide-expand-consolidate family of coupled cluster methods: Numerical illustrations using second order Møller-Plesset perturbation theory

Ida-Marie Høyvik,[a] Kasper Kristensen, Branislav Jansik, and Poul Jørgensen
Lundbeck Foundation Center for Theoretical Chemistry, Department of Chemistry, Aarhus University, Langelandsgade 140, DK-8000 Aarhus C, Denmark

(Received 4 October 2011; accepted 18 November 2011; published online 4 January 2012)

Previously, we have introduced the linear scaling coupled cluster (CC) divide-expand-consolidate (DEC) method, using an occupied space partitioning of the standard correlation energy. In this article, we show that the correlation energy may alternatively be expressed using a virtual space partitioning, and that the Lagrangian correlation energy may be partitioned using elements from both the occupied and virtual partitioning schemes. The partitionings of the correlation energy leads to atomic site and pair interaction energies which are term-wise invariant with respect to an orthogonal transformation among the occupied or the virtual orbitals. Evaluating the atomic site and pair interaction energies using local orbitals leads to a linear scaling algorithm and a distinction between Coulomb hole and dispersion energy contributions to the correlation energy. Further, a detailed error analysis is performed illustrating the error control imposed on all components of the energy by the chosen energy threshold. This error control is ultimately used to show how to reduce the computational cost for evaluating dispersion energy contributions in DEC. © *2012 American Institute of Physics*. [doi:10.1063/1.3667266]

I. INTRODUCTION

Standard implementations of coupled cluster (CC) wave function models using a set of canonical Hartree-Fock (HF) orbitals exhibit a scaling wall. This scaling wall prevents CC calculations from being carried out for any but the smallest molecular systems. It has been attempted to remove this scaling wall by expressing the CC wave function in a set of local HF orbitals. Local occupied HF orbitals have been obtained using various localization strategies,[1–7] and both local occupied and local virtual HF orbitals have recently been obtained by minimizing powers of the orbital variance.[8]

Local orbital correlation methods were pioneered by Pulay[9] and Pulay and Saebo,[10] and an early prominent contribution is the local coupled cluster (LCC) method of Werner and coworkers.[11–16] Many other local CC methods have been proposed, e.g., the atomic orbital-based CC,[17–19] the natural linear scaling approach,[20] the cluster-in-a-molecule approach[21–23] the divide-and-conquer approach,[24] the fragment molecular orbital approach,[25] and the incremental scheme.[26–28] Near-linear scaling for the local CC model using bump functions has been considered by Head-Gordon and coworkers.[29–31] For second order Møller-Plesset (MP2) theory linear scaling has been obtained using a Laplace transformation[32,33] of the energy denominators and using effective integral screening techniques.[34–37]

In the local CC methods, *ad hoc* approximations are introduced in the cluster amplitude equations and in the correlation energy, for example by assigning fixed virtual correlating orbital spaces to local occupied HF orbitals (e.g., using the completeness criterion of Boughton and Pulay,[38]) or by carrying out a physical fragmentation of the molecule (e.g., using dangling bonds). When *ad hoc* approximations are introduced, the precision of the correlation energy compared to a conventional calculation is in general unknown.

We have recently introduced the divide-expand-consolidate (DEC) local CC method.[39,40] In this method, a set of local HF orbitals is assigned to each atomic site. The summation over two occupied orbitals in the correlation energy is then divided into atomic site and pair interaction energy contributions. For local orbitals the individual site energies may be obtained by carrying out CC calculations on small fragments of the total orbital space, which include self-adaptive buffer spaces ensuring that the atomic site and pair interaction energies are determined to a preset energy tolerance – the fragment optimization threshold (FOT). The precision of a DEC calculation is defined by the FOT. The theoretical foundation for DEC given in Ref. 40 was based on a locality analysis of the CC correlation energy and amplitude equations.

The fundamental property of the DEC scheme is the partitioning of the correlation energy into atomic site and pair interaction energy contributions leading to a linear scaling algorithm for local orbitals. In this paper, we present a family of DEC schemes where such partitionings can be performed. In addition to the occupied space partitioning of the correlation energy previously reported in Refs. 39 and 40, we describe a virtual space partitioning of the correlation energy and a partitioning of the Lagrangian correlation energy expression. In the latter, errors in the atomic site and pair interaction energies are bilinear in the errors in the cluster amplitudes and multipliers. The Lagrangian partitioning scheme is also

[a]Electronic mail: idamh@chem.au.dk.

suitable for determining molecular properties. The various DEC approaches give different ways of evaluating the correlation energy and an internal consistency check to validate the calculated correlation energy.

The theoretical foundation for the DEC method is in this paper analyzed from a different perspective than in Refs. 39 and 40. An important characteristic of the coupled cluster energy is that it is invariant with respect to rotations among the occupied and/or among the virtual orbitals. A theoretical derivation demonstrating that the invariances of the correlation energy is partly preserved for the DEC atomic site and pair interaction energies will be presented here.

For local HF orbitals the atomic site and pair interaction energies describe Coulomb hole and dispersion energy contributions to the correlation energy, respectively. For local HF orbitals the pair interaction energies thus decay with the inverse pair distance to the sixth power. The distance decay of the pair interaction energy has previously been used in the DEC method to reduce the computational scaling from quadratic to linear by restricting pair interaction energies to distances where the dispersion energy is non-vanishing. In this paper, we describe how this distance dependence may also be used to dramatically reduce the computational cost for evaluating the pair interaction energies for distances where the dispersion energy contributions are non-vanishing.

The paper is organized as follows. In Sec. II, we derive the DEC family of methods for evaluating the standard and Lagrangian coupled cluster correlation energy, and describe how the correlation energy may be efficiently evaluated for a set of local HF orbitals. Section III contains numerical results to illustrate the foundation for the DEC method, and in Sec. IV we give some concluding remarks.

II. THE DEC FAMILY OF COUPLED CLUSTER METHODS

The DEC coupled cluster method was introduced in Refs. 39 and 40 as a method for carrying out coupled cluster calculations for large molecular systems using a linear scaling algorithm. In this section, we will introduce a family of DEC coupled cluster methods and use MP2 as an illustrative and simple example, recalling that MP2 constitutes the starting point for an extension to higher order CC methods.

In Sec. II A, the standard and Lagrangian correlation energies are given, and in Sec. II B the HF orbitals are assigned to atomic sites. In Sec. II C the DEC family of CC methods is derived by partitioning the correlation energy into atomic site and pair interaction energy contributions. Two different partitioning schemes are carried out for the standard correlation energy, while a single partitioning scheme is considered for the Lagrangian correlation energy. Numerical illustrations are given in Sec. II D showing that the atomic site energies may be associated with a description of the Coulomb holes in the wave function while the pair interaction energies describe dispersion energy contributions. The numerical results show that the site energies may be efficiently calculated in a local basis as described in Sec. II E. A physical interpretation of the atomic site and pair interaction energy contributions in terms of local orbitals is given in Sec. II F.

A. Standard and Lagrangian correlation energy

We consider a closed-shell molecule. The total MP2 energy, E_{MP2}, is obtained from the general coupled cluster energy in Eq. (A2) neglecting the singles contributions

$$E_{\mathrm{MP2}} = E_{\mathrm{HF}} + E_{\mathrm{corr}}, \tag{1}$$

where the correlation energy E_{corr} is given as

$$E_{\mathrm{corr}} = \sum_{ijab} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \tag{2}$$

t_{ij}^{ab} are the cluster doubles amplitudes, expressed in the molecular orbital (MO) basis. Indices i and j refer to occupied MOs, and a and b to virtual MOs. g_{iajb} is a two-electron integral in the HF orbital basis using Mulliken notation.

The MP2 cluster amplitude equation

$$g_{iajb} + \sum_c \left(t_{ij}^{cb} F_{ac} + t_{ij}^{ac} F_{bc}\right) - \sum_k \left(t_{kj}^{ab} F_{ki} + t_{ik}^{ab} F_{kj}\right) = 0 \tag{3}$$

is the simplification of the general cluster amplitude equation in Eq. (A8) where terms of second and higher orders in the fluctuation potential are neglected.

The Lagrangian expression for the MP2 correlation energy may be expressed as

$$\begin{aligned} L_{\mathrm{corr}} = &\sum_{ijab} t_{ij}^{ab}(2g_{iajb} - g_{ibja}) \\ &+ \frac{1}{2}\sum_{ijab} \bar{t}_{ij}^{ab}\left[g_{iajb} + \sum_c \left(t_{ij}^{cb} F_{ac} + t_{ij}^{ac} F_{bc}\right)\right. \\ &\left. - \sum_k \left(t_{kj}^{ab} F_{ki} + t_{ik}^{ab} F_{kj}\right)\right], \end{aligned} \tag{4}$$

where the cluster multiplier equation reads

$$\begin{aligned} &2(2g_{iajb} - g_{ibja}) + \sum_c \left(\bar{t}_{ij}^{cb} F_{ac} + \bar{t}_{ij}^{ac} F_{bc}\right) \\ &- \sum_k \left(\bar{t}_{kj}^{ab} F_{ki} + \bar{t}_{ik}^{ab} F_{kj}\right) = 0, \end{aligned} \tag{5}$$

which is the simplification of Eq. (A9) where the terms of second and higher orders in the fluctuation potential are neglected. From Eqs. (3) and (5) we see that

$$\bar{t}_{ij}^{ab} = 4t_{ij}^{ab} - 2t_{ij}^{ba}. \tag{6}$$

For MP2 the multipliers can thus be determined directly from the amplitudes.

B. Assignment of HF orbitals to atomic sites

Assuming we have carried out a HF calculation, the HF orbitals may be assigned to atomic sites, for example by associating a HF orbital to the atomic site where it has the largest Mulliken charge. An atomic site P thus has assigned a set of occupied $\underline{P}$ and virtual $\overline{P}$ orbitals, see Figure 1. To obtain

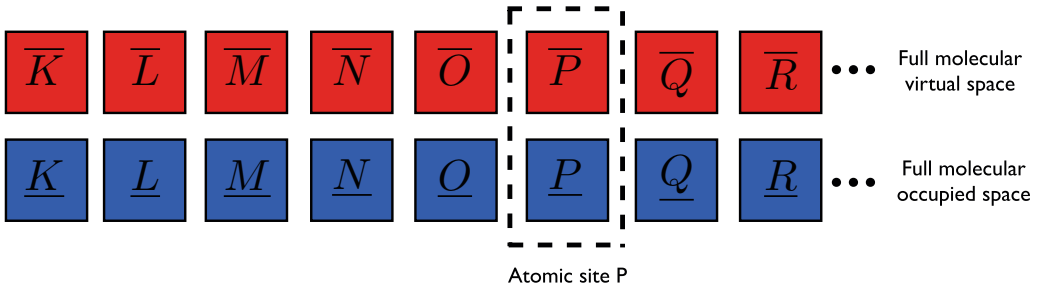

FIG. 1. An illustration showing how a one dimensional system is divided into atomic sites I, J, K.. which has each been assigned a set of occupied (red) $\underline{I}$, $\underline{J}$.. and virtual (blue) $\overline{I}$, $\overline{J}$.. HF orbitals.

a more homogeneous assignment of orbital spaces to atomic sites, and to reduce the number of atomic sites, HF orbitals originally assigned to hydrogenic sites are reassigned to the atomic site to which the hydrogen form a covalent bond.

For a set of delocalized HF orbitals there is no advantage in assigning HF orbitals to atomic sites when evaluating the correlation energy, but for local HF orbitals, simplifications occur that allows the correlation energy to be evaluated using a linear scaling algorithm. This will be detailed in Secs. II D and II E.

Having assigned a set of occupied and virtual HF orbitals to each atomic site, we may express the correlation energy in terms of a summation over atomic sites and atomic pair sites instead of a summation over the individual occupied and virtual HF orbitals.

C. A family of DEC partitionings of the correlation energy

For the standard and Lagrangian correlation energy given in Eqs. (2) and (4), respectively, we may introduce a family of DEC CC methods by replacing the summations in Eqs. (2) and (4) over two occupied or two virtual orbital indices by a summation over atomic sites and pairs of atomic sites. This leads to CC methods containing a quadratic number of pair interaction energy contributions. In Sec. II D, we demonstrate that only a linear number of pairs need to be considered when local orbitals are used.

For the correlation energy in Eq. (2) there exists two partitionings. Either an occupied space partitioning of the orbital space (superscript o) or a virtual space partitioning of the orbital space (superscript v). The third partitioning comes from a partitioning of the Lagrangian correlation energy in Eq. (4) and contains elements of both the occupied and virtual partitioning. The three partitionings are

$$E_{\mathrm{corr}} = \sum_P E_{\mathrm{P}}^{\mathrm{o}} + \sum_{P>Q} \Delta E_{\mathrm{PQ}}^{\mathrm{o}}, \tag{7}$$

$$E_{\mathrm{corr}} = \sum_P E_{\mathrm{P}}^{\mathrm{v}} + \sum_{P>Q} \Delta E_{\mathrm{PQ}}^{\mathrm{v}}, \tag{8}$$

$$L_{\mathrm{corr}} = \sum_P L_{\mathrm{P}} + \sum_{P>Q} \Delta L_{\mathrm{PQ}}. \tag{9}$$

Equation (7) defines the occupied space partitioning, Eq. (8) the virtual space partitioning, and Eq. (9) the Lagrangian partitioning. The occupied space partitioning was introduced in Ref. 39 while the virtual and Lagrangian partitioning have not been considered previously. The definitions of the atomic site energies are

$$E_{\mathrm{P}}^{\mathrm{o}} = \sum_{ij\in\underline{P}} \sum_{ab} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \tag{10}$$

$$E_{\mathrm{P}}^{\mathrm{v}} = \sum_{ab\in\overline{P}} \sum_{ij} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \tag{11}$$

$$\begin{aligned} L_{\mathrm{P}} = &\sum_{ij\in\underline{P}} \sum_{ab} \Bigg[t_{ij}^{ab}(2g_{iajb} - g_{ibja}) \\ &+ \frac{1}{2} \sum_c \bar{t}_{ij}^{ab}\left(t_{ij}^{cb} F_{ac} + t_{ij}^{ac} F_{bc}\right) \Bigg] \\ &+ \sum_{ab\in\overline{P}} \sum_{ij} \left[\frac{1}{2} \bar{t}_{ij}^{ab} g_{iajb} - \frac{1}{2} \sum_k \bar{t}_{ij}^{ab}\left(t_{kj}^{ab} F_{ki} + t_{ik}^{ab} F_{kj}\right) \right], \end{aligned} \tag{12}$$

while the pair interaction energies are given by

$$\begin{aligned} \Delta E_{\mathrm{PQ}}^{\mathrm{o}} = &\sum_{\substack{i\in\underline{P} \\ j\in\underline{Q}}} \sum_{ab} t_{ij}^{ab}(2g_{iajb} - g_{ibja}) \\ &+ \sum_{\substack{i\in\underline{Q} \\ j\in\underline{P}}} \sum_{ab} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \end{aligned} \tag{13}$$

$$\begin{aligned} \Delta E_{\mathrm{PQ}}^{\mathrm{v}} = &\sum_{\substack{a\in\overline{P} \\ b\in\overline{Q}}} \sum_{ij} t_{ij}^{ab}\left(2g_{iajb} - g_{ibja}\right) \\ &+ \sum_{\substack{a\in\overline{Q} \\ b\in\overline{P}}} \sum_{ij} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \end{aligned} \tag{14}$$

$$\begin{aligned} \Delta L_{\mathrm{PQ}} = &\left[\sum_{\substack{i\in\underline{P} \\ j\in\underline{Q}}} + \sum_{\substack{i\in\underline{Q} \\ j\in\underline{P}}} \right] \sum_{ab} \Bigg[t_{ij}^{ab}(2g_{iajb} - g_{ibja}) \\ &+ \frac{1}{2} \sum_c \bar{t}_{ij}^{ab}\left(t_{ij}^{cb} F_{ac} + t_{ij}^{ac} F_{bc}\right) \Bigg] \end{aligned}$$

$$+\left[\sum_{\substack{a\in\overline{P}\\ b\in\overline{Q}}}+\sum_{\substack{a\in\overline{Q}\\ b\in\overline{P}}}\right]\sum_{ij}\left[\frac{1}{2}\bar{t}_{ij}^{ab}g_{iajb}-\frac{1}{2}\sum_{k}\bar{t}_{ij}^{ab}\left(t_{kj}^{ab}F_{ki}+t_{ik}F_{kj}\right)\right]. \quad (15)$$

For the Lagrangian partitioning scheme we have collected the terms containing Fock matrix elements such that the individual energy contributions contain either summations over occupied or virtual orbitals.

Note that no approximations have been made in the correlation energy expressions in Eqs. (7)–(9) relative to Eqs. (2) and (4); the summations over two occupied or two virtual indices have just been split into sums over atomic sites and pair sites.

Using the derivation in Appendix B 1, it follows that the correlation energy is term-wise invariant with respect to an orthogonal transformation among the occupied and/or virtual HF orbitals. From these invariances it follows that some of the invariance properties are preserved for the different partitionings. The invariance with respect to rotations among virtual orbitals is preserved in $E_{\mathrm{P}}^{\mathrm{o}}$ and $\Delta E_{\mathrm{PQ}}^{\mathrm{o}}$, while $E_{\mathrm{P}}^{\mathrm{v}}$ and $\Delta E_{\mathrm{PQ}}^{\mathrm{v}}$ are invariant with respect to rotations among the occupied orbitals. Conversely, $E_{\mathrm{P}}^{\mathrm{o}}$ and $\Delta E_{\mathrm{PQ}}^{\mathrm{o}}$ are not invariant with respect to rotations among the occupied orbitals and $E_{\mathrm{P}}^{\mathrm{v}}$ and $\Delta E_{\mathrm{PQ}}^{\mathrm{v}}$ are not invariant with respect to rotations among the virtual orbitals. For L_{P} and ΔL_{PQ} the first two terms in Eqs. (12) and (15) are invariant with respect to rotations among virtual orbitals but not among occupied orbitals, whereas the last two terms are invariant with respect to rotations among occupied orbitals but not virtual orbitals. The invariances of the atomic site and pair interaction energies are used to select a set of local orbitals for their evaluation, reducing the computational scaling to become linear.

D. The correlation energy separated into Coulomb hole and dispersion energy contributions

We now examine the dependence of the occupied (virtual) site energies on the occupied (virtual) orbitals by carrying out calculations using occupied (virtual) orbitals of different locality, where the locality is measured by the maximum orbital spread (MOS) for the occupied (virtual) orbitals. An equivalent investigation will subsequently be carried out for the Lagrangian energy.

We report MP2/cc-pVDZ calculations on arachidic acid (saturated fatty acid with composition $C_{20}H_{40}O_2$) and on a polyalanine alpha-helix containing 8 alanine residues, referred to as alanine(8). We use geometries obtained from the MAESTRO program[42] where no further optimizations have been carried out. Arachidic acid is an approximately 24 Å long (semi-)one-dimensional system, and it thus constitutes an instructive example where a standard MP2 calculation can easily be carried out, and where the distance dependence of the pair interaction energies ΔE_{PQ} can be observed for large pair distances r_{PQ}. Alanine(8) is a more three-dimensional system, where the maximum distance between two atoms is about 13 Å.

We consider calculations using the localized HF orbitals of Ref. 43 with m = 2 and using nonlocal canonical orbitals. In addition, we consider calculations with two sets of HF orbitals, where the locality is in between the ones of the canonical HF orbitals and the localized HF orbitals. For arachidic acid we thus consider four calculations of the occupied (virtual) site energies labeled I_2^{aa}, I_5^{aa}, I_7^{aa}, and I_{19}^{aa}, where the subscript denotes the approximate MOS (in a.u.) for the occupied (virtual) orbitals – i.e., I_2^{aa} represents a calculation with MOS≈2 a.u. for the occupied (virtual) orbitals, and I_{19}^{aa} represents calculations where delocalized canonical orbitals with MOS≈19 a.u. are used. Similarly, for alanine(8) we consider four calculations of the occupied (virtual) fragment energies denoted I_2^{al8}, I_5^{al8}, I_7^{al8}, and I_{10}^{al8}.

A log-log plot for arachidic acid showing the atomic site (red crosses) and pair interaction energies (blue crosses) as a function of the pair distance are given in Figure 2, while the corresponding alanine(8) results are given in Figure 3. Atomic site energies are plotted at a hypothetical 0 on the abscissa. Results for the occupied partitioning scheme are given in the plots on the left, while the corresponding results for the virtual partitioning scheme are given in the plots on the right. The plots are presented in order of increasing orbital locality (decreasing MOS) from the top – i.e., the top panels contains canonical HF basis results, and the bottom panels contain results for the most local HF orbitals. The MOS for the occupied orbitals are given in the left panels, while the MOS for the virtual orbitals are given in the right panels recalling that the occupied site energies are invariant with respect to a rotation among the virtual orbitals –and vice versa for the virtual site energies.

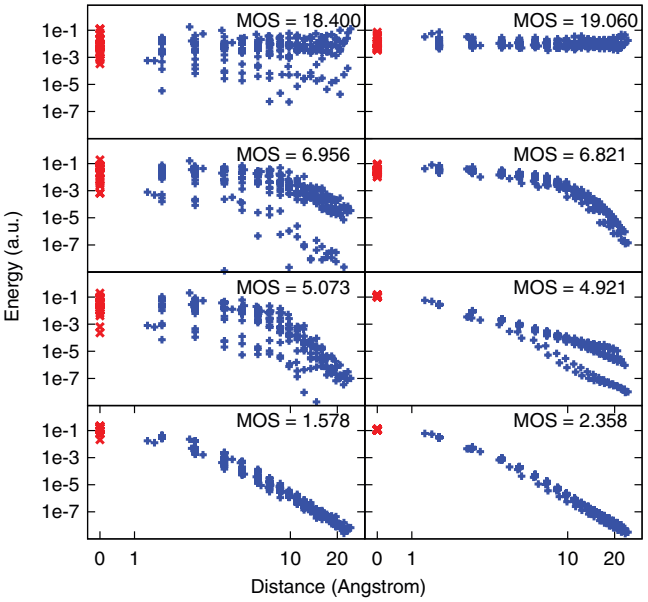

FIG. 2. Absolute atomic site energies (red crosses at hypothetical zero) and pair interaction (blue crosses) energies in a.u. for arachidic acid plotted as a function of the distance (Å) between atomic sites (zero for atomic sites) for orbitals of different locality as indicated by the maximum orbital spread (MOS, in a.u.), see the text. The left and right panels show results for the occupied and virtual partitioning schemes, respectively.

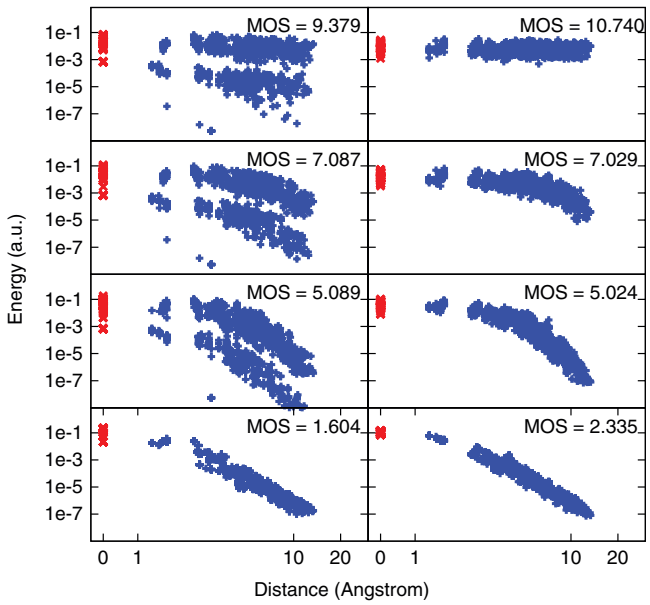

FIG. 3. Absolute atomic site energies (red crosses at hypothetical zero) and pair interaction (blue crosses) energies in a.u. for alanine(8) plotted as a function of the distance (Å) between atomic sites (zero for atomic sites) for orbitals of different locality as indicated by the maximum orbital spread (MOS, in a.u.), see the text. The left and right panels show results for the occupied and virtual partitioning schemes, respectively.

For the delocalized canonical HF orbital calculations the pair interaction energies are independent of r_{PQ}, while for the semi-local orbitals in the middle plots the fragment energies decay with r_{PQ} for larger r_{PQ}. The I_2 results show a clear r_{PQ}^{-n} dependence for both arachidic acid and alanine(8) and for both partitioning schemes. A least squares fit shows that $n \approx 6$ as expected for pair interaction energies describing dispersion effects (see Appendix C for the distance dependence of pair interaction energies for local orbitals). The r_{PQ}^{-6} decay makes it possible to reduce the computational scaling since pair interaction energies for large r_{PQ} are vanishing and can be neglected (Figures 2 and 3, bottom).

For the I_2 calculations, the fragment energies have thus been separated into a contribution (the pair interaction energy) describing the dispersion energy (the second term of Eqs. (7) and (8)) and a contribution at $r_{PQ} = 0$ (the atomic site energy) describing the Coulomb hole contribution (the first term of Eqs. (7) and (8)). Note that in the lower left (right) plots of Figures 2 and 3 we may still have a delocal description of the Coulomb holes and dispersion energies since the individual energy contributions are invariant with respect to an orthogonal transformation of the virtual (occupied) orbitals. For example, the virtual (occupied) orbitals in the lower left (right) plot of Figures 2 and 3 may be canonical orbitals.

For both the arachidic acid calculations I_5^{aa}, I_7^{aa} and I_{19}^{aa} and the alanine(8) I_5^{al8}, I_7^{al8} and I_{10}^{al8} calculations some of the site energy contributions are very small compared to the main trend. This occurs because in our assignment of orbitals to atomic sites some of the carbon atomic sites only gets assigned one occupied HF orbital (the 1s core orbital), and the dispersion energy where such an atomic site is involved is significantly smaller than the dispersion energy for atomic sites where valence HF orbitals have also been assigned.

In Figure 4, we have given a log-log plot of the Lagrangian pair interaction energies for arachidic acid (left) and alanine(8) (right) as a function of r_{PQ} using the local occupied and virtual orbitals described in connection with Figures 2 and 3, bottom. From Figure 4 we see that the Lagrangian pair interaction energies also have an r_{PQ}^{-6} decay when local occupied and virtual orbitals are used.

For local HF orbitals Figures 2–4 show that if atomic site energies are determined to a given tolerance, then beyond a certain distance the pair interaction energies fall below this tolerance value (as a result of the distance decay). Thus the overall precision of the correlation energy will not be compromised by neglecting pairs well beyond this distance, nor will the precision be made better by including them (due to the errors in the atomic site energies). Neglecting pairs beyond a certain cut-off distance reduces the number of pair fragments from quadratic to linear.

Furthermore, for local HF orbitals Figures 2–4 show that computing the correlation energy to a precision of 10^{-2} a.u. is equivalent to neglecting dispersion effects. This is

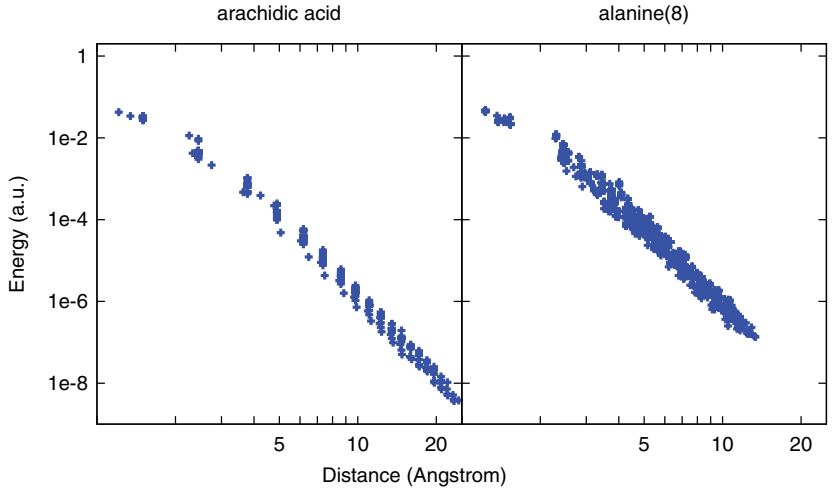

FIG. 4. Absolute values (a.u) of the Lagrangian pair interaction energies for arachidic acid (left) and alanine(8) (right) plotted as a function of distance (Å) between atomic sites.

the case since all pair interaction energies beyond ≈2 Å fall below 10^{-2} a.u. A treatment where dispersion effects are neglected is in general insufficient for accurate chemical applications.

E. Evaluation of DEC atomic site energies and pair interaction energies

In this section, we discuss how occupied, virtual and Lagrangian atomic site and pair interaction energies may be evaluated if both the occupied and virtual orbitals are local.

Consider initially the occupied atomic site energy in Eq. (10) and assume that both the occupied and virtual orbitals are local. For $ij \in \underline{P}$ the g_{iajb} integral is vanishing if one (or both) of the charge distributions $\phi_i^*\phi_a$ and $\phi_b^*\phi_j$ is vanishing. For a non-vanishing g_{iajb} integral it thus follows that $ab \in [\overline{P}]$, where $[\overline{P}]$ denotes the set of virtual orbitals assigned to the atomic sites local to atomic site P (including P itself). Similar arguments may be used for $\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$, and the summation over the virtual orbital indices in the occupied atomic site and pair interaction energies may thus – based on vanishing integrals – be restricted as

$$E^{\mathrm{o}}_{\mathrm{P}} = \sum_{ij\in\underline{P}} \sum_{ab\in[\overline{P}]} \left(t_{ij}^{ab}\right)(2g_{iajb} - g_{ibja}), \tag{16}$$

$$\Delta E^{\mathrm{o}}_{PQ} = \sum_{\substack{i\in\underline{P}\\ j\in\underline{Q}}} \sum_{ab\in[\overline{P}]\cup[\overline{Q}]} \left(t_{ij}^{ab}\right)(2g_{iajb} - g_{ibja}) + \sum_{\substack{i\in\underline{Q}\\ j\in\underline{P}}} \sum_{ab\in[\overline{P}]\cup[\overline{Q}]} \left(t_{ij}^{ab}\right)(2g_{iajb} - g_{ibja}). \tag{17}$$

The summations in Eq. (16) define the energy orbital space (EOS) for $E^{\mathrm{o}}_{\mathrm{P}}$ ($\underline{P} \cup [\overline{P}]$), and the EOS for $\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$ is the union of EOS for $E^{\mathrm{o}}_{\mathrm{P}}$ and $E^{\mathrm{o}}_{\mathrm{Q}}$ ($\underline{P} \cup \underline{Q} \cup [\overline{P}] \cup [\overline{Q}]$).

For the virtual and Lagrangian atomic site and pair interaction energies we may similarly use the locality of the integrals to express these energies as

$$E^{\mathrm{v}}_{\mathrm{P}} = \sum_{ab\in\overline{P}} \sum_{ij\in[\underline{P}]} \left(t_{ij}^{ab}\right)(2g_{iajb} - g_{ibja}), \tag{18}$$

$$\Delta E^{\mathrm{v}}_{\mathrm{PQ}} = \sum_{\substack{a\in\underline{P}\\ b\in\underline{Q}}} \sum_{ij\in[\underline{P}]\cup[\underline{Q}]} \left(t_{ij}^{ab}\right)(2g_{iajb} - g_{ibja}) + \sum_{\substack{a\in\underline{Q}\\ b\in\underline{P}}} \sum_{ij\in[\underline{P}]\cup[\underline{Q}]} \left(t_{ij}^{ab}\right)(2g_{iajb} - g_{ibja}), \tag{19}$$

$$L_{\mathrm{P}} = \sum_{ij\in\underline{P}} \sum_{ab\in[\overline{P}]} \left[t_{ij}^{ab}(2g_{iajb} - g_{ibja}) + \frac{1}{2}\sum_{c\in[\overline{P}]} \bar{t}_{ij}^{ab}\left(t_{ij}^{cb}F_{ac} + t_{ij}^{ac}F_{bc}\right)\right] + \sum_{ab\in\overline{P}} \sum_{ij\in[\underline{P}]} \left[\frac{1}{2}\bar{t}_{ij}^{ab}g_{iajb} - \frac{1}{2}\sum_{k\in[\underline{P}]} \bar{t}_{ij}^{ab}(t_{kj}^{ab}F_{ki} + t_{ik}^{ab}F_{kj})\right], \tag{20}$$

$$\Delta L_{\mathrm{PQ}} = \left[\sum_{\substack{i\in\underline{P}\\ j\in\underline{Q}}} + \sum_{\substack{i\in\underline{Q}\\ j\in\underline{P}}}\right] \sum_{ab\in[\overline{P}]} \left[t_{ij}^{ab}(2g_{iajb} - g_{ibja}) + \frac{1}{2}\sum_{c\in[\overline{P}]} \bar{t}_{ij}^{ab}\left(t_{ij}^{cb}F_{ac} + t_{ij}^{ac}F_{bc}\right)\right] + \left[\sum_{\substack{a\in\overline{P}\\ b\in\overline{Q}}} + \sum_{\substack{a\in\overline{Q}\\ b\in\overline{P}}}\right] \sum_{ij\in[\underline{P}]} \left[\frac{1}{2}\bar{t}_{ij}^{ab}g_{iajb} - \frac{1}{2}\sum_{k\in[\underline{P}]} \bar{t}_{ij}^{ab}\left(t_{kj}^{ab}F_{ki} + t_{ik}F_{kj}\right)\right]. \tag{21}$$

The EOSs for the Lagrangian and virtual partitionings are defined by the atomic site summation, in the same manner as for the occupied partitioning.

To elaborate on the choice of partitioning for the standard correlation energy, we note that only an occupied or virtual partitioning can be exploited to give a linear scaling algorithm. A mixed partitioning gives the atomic site energies,

$$E^{\mathrm{ov}}_{\mathrm{P}} = \sum_{i\in\underline{P}} \sum_{a\in\overline{P}} \sum_{jb} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \tag{22}$$

which are nonlocal because the summation indices jb cannot be restricted to be local to atomic site P due to vanishing charge distributions as can be done for the occupied and virtual partitioning schemes in Eqs. (16) and (18).

The conventional correlation energy contribution may be evaluated using either an occupied or a virtual orbital space partitioning. For the Lagrangian, we have chosen an approach where one of the standard correlation energy contributions are evaluated using an occupied space partitioning and the other using a virtual space partitioning. This puts equal weight on the occupied and virtual contributions to the correlation energy.

In Ref. 40 it was shown that the CC amplitudes entering the energy expression for an atomic site energy may be determined by solving CC amplitude equations in a local orbital fragment space, denoted the amplitude orbital space (AOS). A locality analysis of the cluster amplitude equation shows that the AOS in addition to containing the EOS, must contain buffer spaces for the occupied and virtual orbitals, see Figure 5. In particular, the orbital fragment for the occupied partitioning scheme in Figure 5(a) contains an occupied buffer space (light-blue). Conceptually, it also contains a virtual buffer space, but in practice the virtual buffer space is absorbed in the virtual EOS (dark-red). Similarly, the

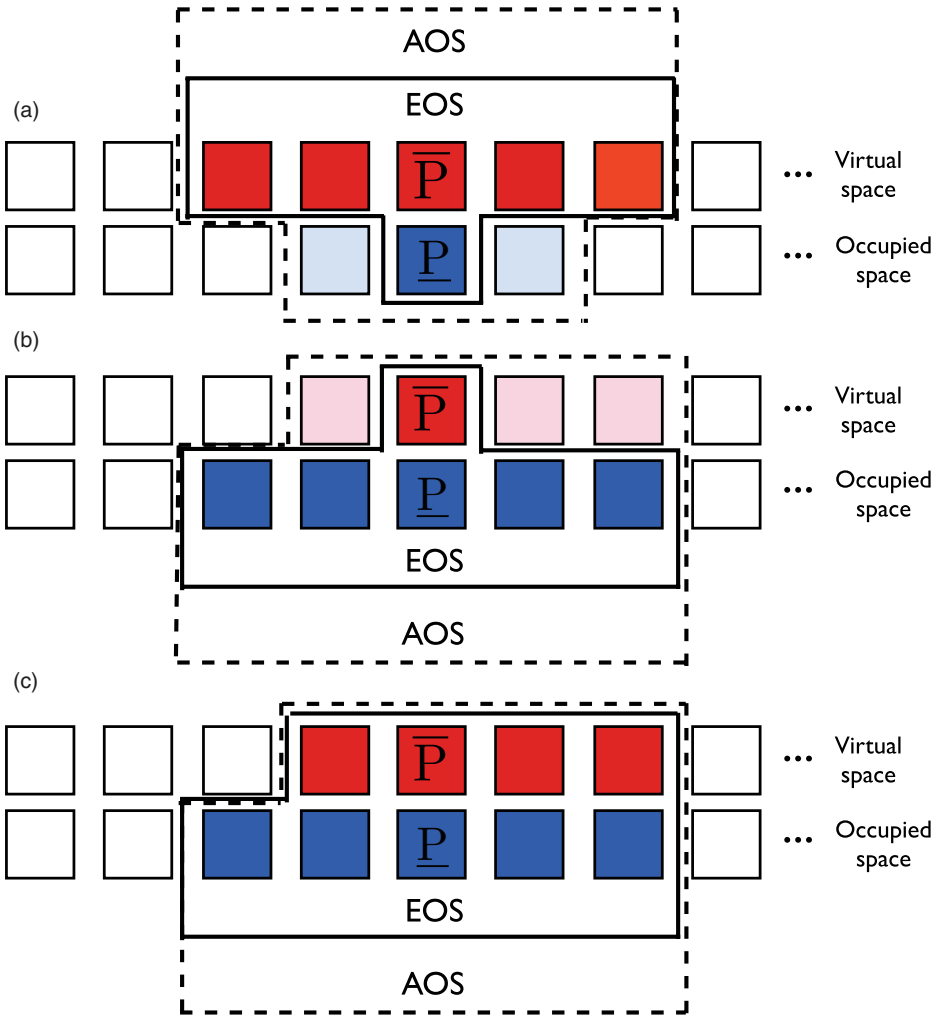

FIG. 5. Illustration of a single atomic fragment for atomic site P using occupied (a), virtual (b) and Lagrangian (c) partitionings of the orbital space. Each square represents the set of occupied or virtual orbitals assigned to a particular atom. For each partitioning scheme the atomic site energy is evaluated using the energy orbital space (EOS) (dark-red and dark-blue), whereas the CC amplitude equations are solved in the amplitude orbital space (AOS) which is the union of the EOS and the buffer space (light-blue and pink).

atomic fragment entering the virtual partitioning scheme in Figure 5(b) contains a virtual buffer space (pink), and an occupied buffer space, which, in practice, is absorbed in the occupied EOS (dark-blue). The Lagrangian atomic fragments Figure 5(c) similarly also contains both an occupied and a virtual buffer space which both may be absorbed into the EOS. The EOS and AOS are thus coinciding in the Lagrangian scheme. This gives a more uniform treatment of the occupied and virtual orbital space where, e.g., the self adaptive determination of orbital spaces is simplified and may be done more rigorously.

The use of buffer spaces allows the EOS amplitudes to adjust to the surrounding environment to obtain amplitudes which, to a high precision, are the same as those of a full CC calculation.

The AOS for determining CC amplitudes for pair interaction energies is the union of the AOSs for the atomic site calculations E_P and E_Q. In other words, the locality of a DEC calculation is determined by the locality of the atomic site calculations, in which locality is determined in a system-adaptive black-box manner. The black box determination of the sizes of orbital spaces (as defined by the FOT) ultimately ensures control of the error in the total correlation energy compared to a full molecular CC calculation.[40]

When the DEC correlation energy is evaluated the errors in the amplitudes and multipliers may be divided into two contributions:

- the amplitudes and multipliers are determined using a restricted AOS where weak interactions with amplitudes outside the AOS have been neglected
- the EOSs have been truncated and the amplitudes outside the truncated EOS have been neglected.

When the correlation energy is determined using the DEC Lagrangian framework both the amplitude and multiplier errors are approximately proportional to the square root of the preset energy threshold, FOT, because the energy errors of the standard CC Lagrangian are bilinear in the amplitude errors (see Appendix A).

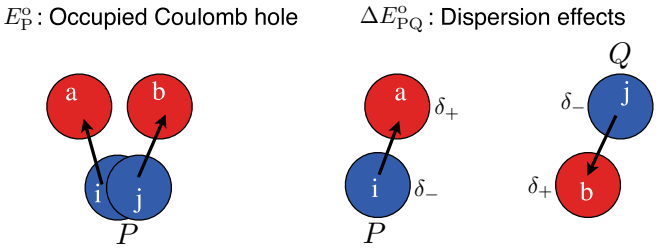

FIG. 6. Illustrations of excitations from local occupied orbitals ϕ_i and ϕ_j (blue) into local virtual orbitals ϕ_a and ϕ_b (red). Left: Short-range electron-electron repulsion described by the atomic site energies $E_{\mathrm{P}}^{\mathrm{o}}$ (Coulomb hole). Right: Dispersion effects described by occupied pair interaction energies $\Delta E_{\mathrm{PQ}}^{\mathrm{o}}$.

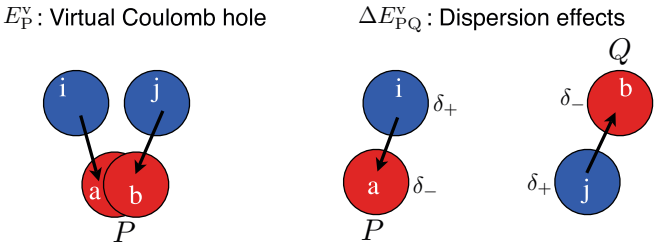

FIG. 7. Illustrations of excitations from local occupied orbitals ϕ_i and ϕ_j (blue) into local virtual orbitals ϕ_a and ϕ_b (red). Left: Short-range electron-electron repulsion described by the atomic site energies $E_{\mathrm{P}}^{\mathrm{v}}$ (Coulomb hole). Right: Dispersion effects described by virtual pair interaction energies $\Delta E_{\mathrm{PQ}}^{\mathrm{v}}$.

F. Physical interpretation in terms of local orbitals

In Sec. II D, the physical interpretation of the fragment energies was discussed for full molecular calculations using HF orbitals of different locality. Having established that the site energies can be calculated using small local orbital spaces when both occupied and virtual orbitals are local, we are now in a position to give a more explicit intuitive interpretation of the site energies.

The contributions to the atomic site energies may be expressed in terms of orbital pair contributions as

$$E_{\mathrm{P}}^{\mathrm{o}} = \sum_{ij \in \underline{P}} E_{\mathrm{P}_{\mathrm{ij}}}^{\mathrm{o}} \tag{23}$$

$$E_{\mathrm{P}}^{\mathrm{v}} = \sum_{ab \in \overline{P}} E_{\mathrm{P}_{\mathrm{ab}}}^{\mathrm{v}} \tag{24}$$

where

$$E_{\mathrm{P}_{\mathrm{ij}}}^{\mathrm{o}} = \sum_{ab} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \quad ij \in \underline{P} \tag{25}$$

$$E_{\mathrm{P}_{\mathrm{ab}}}^{\mathrm{v}} = \sum_{ij} t_{ij}^{ab}(2g_{iajb} - g_{ibja}), \quad ab \in \overline{P} \tag{26}$$

$E_{\mathrm{P}_{\mathrm{ij}}}^{\mathrm{o}}$ arises predominantly from a description of the Coulomb hole resulting from occupying both ϕ_i and ϕ_j in the wave function, recalling that for local orbitals the orbitals assigned to the same atomic site occupy the same regions in space. In Figure 6 (left), a graphical representation is given for an individual contribution to $E_{\mathrm{P}_{\mathrm{ij}}}^{\mathrm{o}}$ of Eq. (25). Similarly $E_{\mathrm{P}_{\mathrm{ab}}}^{\mathrm{v}}$ arise predominantly from a description of the *virtual Coulomb hole* resulting from occupying both ϕ_a and ϕ_b in the wave function. A graphical representation is given for an individual contribution to $E_{\mathrm{P}_{\mathrm{ab}}}^{\mathrm{v}}$ in Figure 7 (left).

The dispersion interactions (also known as induced dipole-dipole interactions) described by $\Delta E_{\mathrm{PQ}}^{\mathrm{o}}$ (Figure 6, right) and $\Delta E_{\mathrm{PQ}}^{\mathrm{v}}$ (Figure 7, right) has an r_{PQ}^{-6} dependence (as discussed in connection with Figures 2 and 3). This is shown explicitly for local orbitals in Appendix C. The dispersion energies are in general attractive as indicated by the induced partial charges. It turns out to be crucial to exploit the distance dependence of dispersion energies for an efficient evaluation of the correlation energy as discussed in Sec. III D.

III. NUMERICAL EXAMPLES OF ATOMIC SITE AND PAIR SITE INTERACTION ENERGIES

We now compare DEC MP2 results evaluated to a given FOT against results from full DEC MP2 calculations for the arachidic acid and alanine(8) molecules of Sec. II D, where we use the local occupied and virtual orbitals of Ref. 43 with $m = 2$. Full DEC MP2 corresponds to Eqs. (7)–(15) and is equivalent to a full MP2 calculation (which in principle corresponds to FOT = 0). In particular, we compare atomic site energies and pair interaction energies obtained in DEC calculations to a given FOT with their counterparts from the full DEC MP2 calculations. The error analysis of the pair interaction energies will be used to demonstrate how computational savings may be obtained for evaluating the dispersion (pair interaction) contribution to the correlation energy.

A. Total errors in DEC MP2 energies

In Tables I–IV, we report the total errors in the atomic site energies, pair interaction energies, and the correlation energy for arachidic acid and alanine(8) respectively. Errors are computed for the FOTs = 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, and 10^{-6} a.u.

TABLE I. Sum of all errors (a.u.) in atomic site energies and pair interaction energies for arachidic acid.

	Occupied partitioning		Virtual partitioning		Lagrangian partitioning	
FOT	$\sum_P \delta E_{\mathrm{P}}^{\mathrm{o}}$	$\sum_{P>Q} \delta\Delta E_{\mathrm{PQ}}^{\mathrm{o}}$	$\sum_P \delta E_{\mathrm{P}}^{\mathrm{v}}$	$\sum_{P>Q} \delta\Delta E_{\mathrm{PQ}}^{\mathrm{v}}$	$\sum_P \delta E_{\mathrm{P}}^{\mathrm{v}}$	$\sum_{P>Q} \delta\Delta E_{\mathrm{PQ}}^{\mathrm{v}}$
10^{-2}	8.97×10^{-2}	4.82×10^{-2}	1.79×10^{-1}	7.22×10^{-2}	1.50×10^{-1}	4.99×10^{-2}
10^{-3}	1.79×10^{-2}	1.04×10^{-2}	2.09×10^{-2}	6.70×10^{-3}	1.38×10^{-2}	4.93×10^{-3}
10^{-4}	2.19×10^{-3}	1.01×10^{-3}	1.40×10^{-3}	4.69×10^{-4}	1.37×10^{-3}	5.80×10^{-4}
10^{-5}	1.78×10^{-4}	9.07×10^{-5}	1.97×10^{-4}	5.92×10^{-5}	1.25×10^{-4}	5.33×10^{-5}
10^{-6}	2.18×10^{-5}	1.07×10^{-5}	7.64×10^{-6}	3.22×10^{-6}	1.50×10^{-5}	5.09×10^{-6}

TABLE II. Sum of all errors (a.u.) in atomic site energies and pair interaction energies for alanine(8).

	Occupied partitioning		Virtual partitioning		Lagrangian partitioning	
FOT	$\sum_P \delta E^{\mathrm{o}}_{\mathrm{P}}$	$\sum_{P>Q} \delta\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$	$\sum_P \delta E^{\mathrm{v}}_{\mathrm{P}}$	$\sum_{P>Q} \delta\Delta E^{\mathrm{v}}_{\mathrm{PQ}}$	$\sum_P \delta E^{\mathrm{v}}_{\mathrm{P}}$	$\sum_{P>Q} \delta\Delta E^{\mathrm{v}}_{\mathrm{PQ}}$
10^{-2}	2.04×10^{-1}	1.14×10^{-1}	3.96×10^{-1}	2.26×10^{-1}	3.37×10^{-1}	1.37×10^{-1}
10^{-3}	5.18×10^{-2}	3.75×10^{-2}	5.44×10^{-2}	3.66×10^{-2}	2.72×10^{-2}	1.22×10^{-2}
10^{-4}	4.71×10^{-3}	2.11×10^{-3}	5.43×10^{-3}	3.99×10^{-3}	4.21×10^{-3}	2.17×10^{-3}
10^{-5}	4.45×10^{-4}	2.94×10^{-4}	4.98×10^{-4}	9.33×10^{-4}	4.44×10^{-4}	1.94×10^{-4}
10^{-6}	5.66×10^{-5}	9.01×10^{-5}	4.10×10^{-5}	5.09×10^{-5}	3.71×10^{-5}	1.53×10^{-5}

The results in Tables I–IV demonstrate that the total errors for both atomic site and pair interaction energies are proportional to the FOT. The errors in the atomic site energies appear to be slightly larger than the pair interaction energy error, but of the same order of magnitude. The error in the total correlation energy – being the sum of the atomic site and pair interaction energy errors – is also proportional to the FOT for all three partitioning schemes. Consequently, the occupied, virtual and Lagrangian partitionings schemes constitute three alternative strategies for calculating the correlation energy. The errors in the Lagrangian approach appear to be slightly more systematic than the errors for the occupied and virtual partitioning schemes, which is presumably due to the variational nature of the Lagrangian approach and the uniform treatment of the occupied and virtual orbital spaces.

B. Errors in DEC MP2 atomic site energies

The errors for each individual atomic site energy is visualized in Figure 8 for arachidic acid and in Figure 9 for alanine(8). Note that since no orbitals have been assigned to the hydrogen atoms, there are 22 atomic sites in arachidic acid and 41 in alanine(8).

Figures 8 and 9 illustrates the meaning of the FOT, namely that all atomic site errors are proportional to the FOT. As the FOT is tightened with an order of magnitude, the errors drop one order of magnitude, showing the error control that the FOT poses on the calculation. In other words, the sizes of the spaces used in the summations in Eqs. (16), (18), and (20) have been expanded in a black-box manner until the atomic site energy is converged with an error proportional to the FOT.

C. Orbital fragment sizes in DEC MP2 calculations

In Tables V and VI, the mean and maximum number of orbitals in the AOS are given for arachidic acid and alanine(8) for various FOTs. The number of orbitals in a fragment is an important measure, since it dictates the cost of both atomic site and pair interaction energy calculations. To make it easier to compare the three partitionings the number of orbitals are presented as X(Y), where X is the number of occupied orbitals and Y the number of virtual orbitals.

Tables V and VI show how the size of the fragments correlates with the FOT. As the FOT is tightened, both the occupied and virtual spaces increase in size. As the FOT is proportional to the error in the correlation energy, Tables V and VI thus show the sizes of the fragment orbital spaces required to obtain the correlation energy to a given precision.

We note that the AOS sizes are FOT dependent, but independent of molecular size. To illustrate this, the maximum and average size (Figure 10 left and right respectively) of the occupied and virtual orbitals in the AOS is plotted as a function of molecular size for the occupied partitioning scheme (similar results are obtained for the other partitioning schemes). The calculations were carried out for a series of polyalanine α-helices for FOT $= 10^{-4}$ a.u. for the occupied space partitioning in a cc-pVDZ basis. The polyalanine sizes ranges from 13 atoms for alanine(1) to 203 atoms for alanine(20). As seen from Figure 10 the orbital fragment sizes increase with molecular size for very small molecules, but converges to become independent of molecular size.

Due to overlapping fragments, there is a large overhead associated with fragment evaluation. To reduce this overhead several spatially close atomic fragments may be combined into a super fragment.

D. Errors in DEC MP2 pair interaction calculations

We now investigate the pair interaction energies for arachidic acid and alanine(8). We consider the errors in pair interaction energies obtained from DEC MP2, where the reference energies are the MP2 pair interaction energies from full MP2 calculations. The pair interaction energy error are plotted as a function of r_{PQ} for arachidic acid in Figure 11

TABLE III. Total errors (a.u.) in the correlation energy for arachidic acid.

FOT	$\delta E^{\mathrm{o}}_{\mathrm{corr}}$	$\delta E^{\mathrm{v}}_{\mathrm{corr}}$	δL_{corr}
10^{-2}	1.38×10^{-1}	2.51×10^{-1}	2.0×10^{-1}
10^{-3}	2.83×10^{-2}	2.76×10^{-2}	1.87×10^{-2}
10^{-4}	3.20×10^{-3}	1.87×10^{-3}	1.95×10^{-3}
10^{-5}	2.69×10^{-4}	2.56×10^{-4}	1.78×10^{-4}
10^{-6}	3.25×10^{-5}	1.09×10^{-5}	2.01×10^{-5}

TABLE IV. Total errors (a.u.) in the correlation energy for alanine(8).

FOT	$\delta E^{\mathrm{o}}_{\mathrm{corr}}$	$\delta E^{\mathrm{v}}_{\mathrm{corr}}$	δL_{corr}
10^{-2}	3.18×10^{-1}	6.22×10^{-1}	4.74×10^{-1}
10^{-3}	8.93×10^{-2}	9.10×10^{-2}	3.95×10^{-2}
10^{-4}	6.82×10^{-3}	9.42×10^{-3}	6.38×10^{-3}
10^{-5}	7.39×10^{-4}	1.43×10^{-3}	6.38×10^{-4}
10^{-6}	1.47×10^{-4}	9.90×10^{-5}	5.24×10^{-5}

TABLE V. The average and maximum number of orbitals in the occupied space (virtual space) for occupied, virtual and Lagrangian atomic sites for arachidic acid, determined to a given FOT in a DEC MP2 calculation.

	Occupied scheme		Virtual scheme		Lagrangian scheme	
FOT (a.u)	Avg	Max	Avg	Max	Avg	Max
10^{-2}	7.9(68.4)	9(100)	14.1(20.2)	19(35)	8.2(40.1)	12(75)
10^{-3}	7.9(93.5)	9(104)	19.5(91.5)	28(104)	17.1(93.9)	28(141)
10^{-4}	19.5(139.3)	21(185)	28.6(96.9)	36(115)	26.0(131.4)	33(185)
10^{-5}	25.5(191.0)	29(225)	37.5(143.9)	56(165)	36.0(186.3)	48(241)
10^{-6}	32.4(240.2)	40(310)	53.9(167.2)	77(245)	46.5(231.5)	60(285)

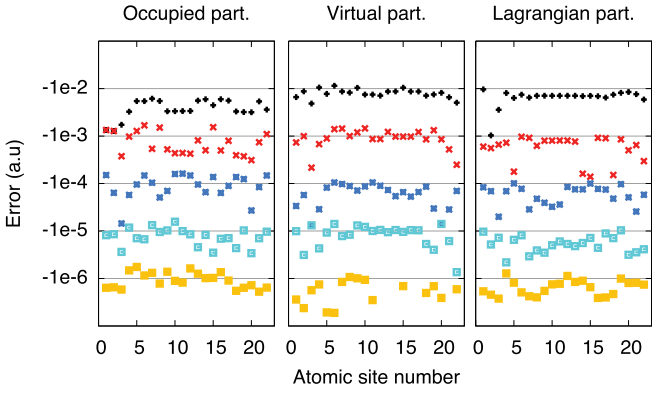

FIG. 8. Errors (a.u.) for arachidic acid atomic site energies plotted against site number compared to a full MP2 calculation using the occupied (left), virtual (middle) and Lagrangian (right) partitioning schemes for different FOTs.

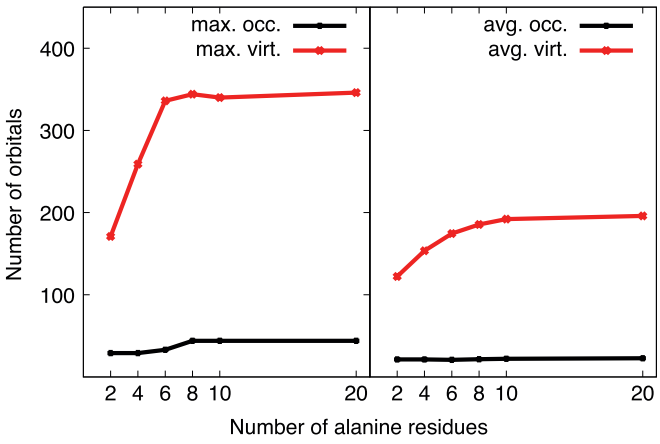

FIG. 10. The maximum (left) and average (right) number of occupied and virtual orbitals entering a DEC calculation for polyalanines of different sizes.

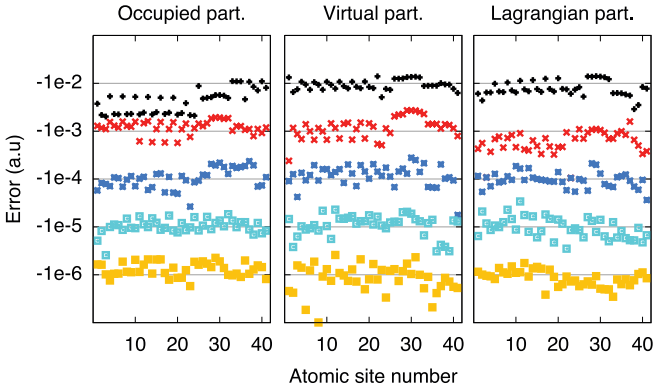

FIG. 9. Errors (a.u.) for alanine(8) atomic site energies plotted against fragment number compared to a full MP2 calculation using the occupied (left), virtual (middle) and Lagrangian (right) partitioning schemes for different FOTs.

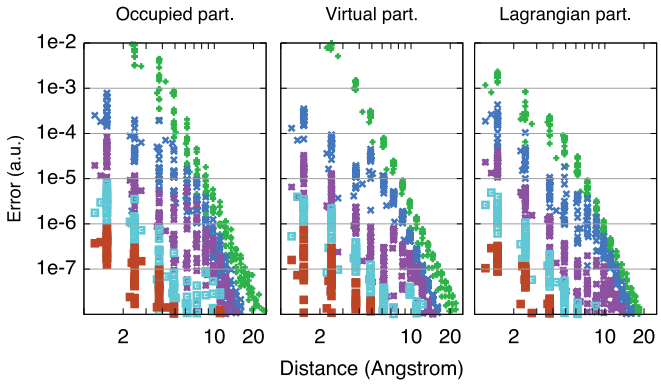

FIG. 11. Errors in $\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$ (left), $\Delta E^{\mathrm{v}}_{\mathrm{PQ}}$ (middle) and ΔL_{PQ} (right) for arachidic acid when computed for FOTs 10^{-2} (green pluses), 10^{-3} (blue crosses), 10^{-4} (purple stars), 10^{-5} (blue open squares) and 10^{-6} (brown filled squares). All FOTs are in a.u.

TABLE VI. The average and maximum number of orbitals in the occupied space (virtual space) for occupied, virtual and Lagrangian atomic fragments for alanine(8), determined to a given FOT in a DEC MP2 calculation.

	Occupied scheme		Virtual scheme		Lagrangian scheme	
FOT (a.u)	Avg	Max	Avg	Max	Avg	Max
10^{-2}	3.8(52.5)	5(117)	7.7(19.9)	12(36)	8.2(41.4)	12(104)
10^{-3}	6.0(108.8)	12(205)	20.2(67.1)	32(111)	25.7(119.9)	37(229)
10^{-4}	22.0(184.3)	44(344)	43.5(140.4)	67(229)	45.5(181.8)	68(330)
10^{-5}	35.2(312.2)	62(490)	77.7(214.4)	144(345)	69.0(317.1)	96(535)
10^{-6}	48.4(424.7)	103(602)	107.1(338.1)	146(514)	96.0(419.9)	128(582)

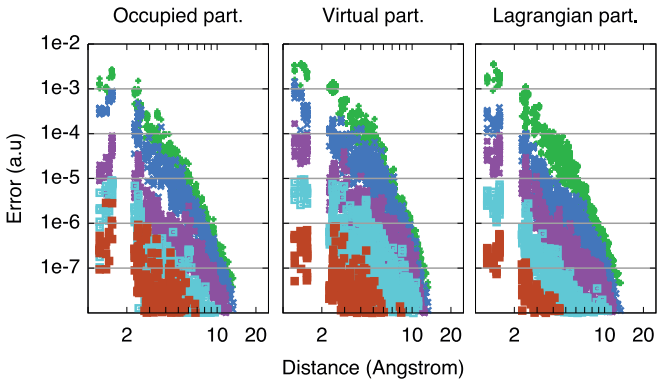

FIG. 12. Errors in $\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$ (left), $\Delta E^{\mathrm{v}}_{\mathrm{PQ}}$ (middle) and ΔL_{PQ} (right) for alanine(8) when computed for FOTs 10^{-2} (green pluses), 10^{-3} (blue crosses), 10^{-4} (purple stars), 10^{-5} (blue open squares) and 10^{-6} (brown filled squares). All FOTs are in a.u.

and for alanine(8) in Figure 12. The errors are given for FOTs $= 10^{-2}$, 10^{-3}, 10^{-4}, 10^{-5} and 10^{-6} a.u. for all three partitioning schemes.

The errors displayed in Figures 11 and 12 decay rapidly with pair distance as does the exact pair interaction energies (Figures 2 and 3 (bottom) and Figure 4). Further, the errors for different FOTs are separated substantiating the error control that implicitly is imposed on the pair interaction energies via the FOT. However, from Figures 11 and 12 we also see that the errors for different FOTs are most clearly separated for the Lagrangian partitioning. This is due to the variational nature of the Lagrangian correlation energy and the more uniform treatment of the occupied and virtual orbital spaces in the Lagrangian partitioning scheme. This makes the Lagrangian the best candidate for the cost reduction described below.

Having illustrated the predictable behavior of the DEC pair interaction energy errors, and considering the r_{PQ}^{-6} decay of the pair interaction energies (see Appendix C), we are now in position to reduce the cost for evaluating pair interaction energies that are not negligible.

Consider the case where we have optimized the atomic site energies to a FOT $= 10^{-4}$ a.u. The correlation energy error relative to a full MP2 calculation will then be proportional to 10^{-4} a.u. The pair fragments are constructed using the union of spaces from the corresponding atomic site calculations, and as the pair distance increases the errors in the pair interaction energies decrease rapidly. Thus, beyond a certain distance the precision of the pair interaction energies is unnecessarily high, since the error in the total correlation energy will be proportional to 10^{-4} a.u. From Figures 11 and 12 we see that beyond, e.g., $\approx$5 Å the errors of pair interaction energies corresponding to FOT $= 10^{-3}$ a.u. are well below 10^{-4} a.u. Hence, to reduce the computational cost we may consider computing pair interaction energies using the atomic site spaces corresponding to FOT $= 10^{-3}$ a.u., without affecting the precision of the total correlation energy. Furthermore, pairs separated by more than $\approx$10 Å may be completely neglected (as discussed in Sec. II D and in Ref. 39).

The cost reducing scheme described above yields three regions for the pair interaction energies;

- Region 1: Pair interaction energies are computed from original atomic site spaces (FOT $= 10^{-4}$ a.u., $r_{\mathrm{PQ}} < 5$ Å for the above example).
- Region 2: Pair interaction energies are computed from atomic site spaces obtained using a less tight threshold (FOT $= 10^{-3}$ a.u., 5 Å $< r_{\mathrm{PQ}} <$ 10 Å for the above example).
- Region 3: Pair interaction energies are vanishing and are neglected ($r_{\mathrm{PQ}} >$ 10 Å for the above example).

We note that the behavior of the rather different molecules arachidic acid and alanine(8) is very similar, since it reflects the nature of dispersion. Therefore, such cost effective schemes are expected to be general for non-metallic closed-shell molecules. The most important thing is to be conservative when choosing where to start making approximations.

IV. SUMMARY AND PERSPECTIVE

In this paper, we have introduced a family of DEC coupled cluster methods which – in addition to the occupied partitioning of the standard CC correlation energy in Refs. 39 and 40 – include a virtual partitioning of the standard correlation energy and a partitioning of the Lagrangian correlation energy. A detailed theoretical analysis and numerical illustrations are provided to demonstrate the performance of the different DEC partitionings.

We have shown that the atomic site and pair interaction energies entering the occupied (virtual) partitioning scheme for the correlation energy are invariant with respect to an orthogonal transformation among the virtual (occupied) orbitals. By selecting local orbitals we have provided a description for the practical evaluation of the atomic site and pair interaction energies in terms of small orbital fragment calculations. The orbital fragment CC equations of the occupied, virtual or Lagrangian DEC partitioning schemes are solved in an amplitude orbital space that includes buffer spaces, while the atomic site and pair interaction energies are evaluated using the subset energy orbital space. Importantly, the occupied, virtual and Lagrangian partitioning schemes constitute independent strategies for evaluating the correlation energy. The combined use of these schemes may be used to validate the calculated DEC correlation energy, also for large molecular systems where a standard reference calculation cannot be carried out.

There are two error sources in the atomic site and pair interaction energies: the error in amplitudes resulting from solving the CC equations in an orbital fragment space instead of the full space and the error arising from neglecting amplitudes outside the energy orbital space. We have shown numerically that these amplitude errors introduce errors in the atomic site and pair interaction energies that are proportional to the fragment optimization threshold (FOT). Consequently, the errors in the atomic site energies, in the pair interaction energies, and in the total correlation energy are proportional to the predefined FOT.

The DEC evaluation of the correlation energy is on par with a standard MP2 (canonical) evaluation of the correlation energy, as the error in the correlation energy and amplitudes

are fully controlled in both approaches. In a standard calculation, the precision of the amplitudes and the correlation energy is determined by a preset residual norm in the amplitude and multiplier equations, while in the DEC scheme the precision of the correlation energy is determined by the energy threshold (FOT) for the atomic site calculations which in turn determine the precision of the amplitudes and the multipliers.

Although all three schemes have the same error control and will yield reliable results, the Lagrangian scheme is seen to have some advantages compared to the occupied and virtual schemes;

- The variational nature of the Lagrangian scheme yields errors in amplitudes/multipliers that are roughly proportional to the square root of the FOT.
- The Lagrangian scheme has a more uniform treatment of the occupied and virtual space.
- With the Lagrangian scheme it is straightforward to extend the DEC scheme to calculate molecular properties.

For these reasons the Lagrangian partitioning is our preferred scheme.

Numerical investigations of the DEC errors show that the errors in pair interaction energies decay with r_{PQ}^{-6} in a local orbital basis. This result is of major importance, since it implies that distant pair sites may be evaluated at a much lower FOT (smaller orbital space) with no loss of precision in the correlation energy. Very distant pairs may be neglected completely, also without loss of precision. Pair interaction energies evaluated at a lower FOT contain much smaller amplitude orbital spaces, thus speeding up the DEC calculation dramatically.

We have shown that DEC CC and standard CC share the same solid theoretical foundation. The DEC is governed by a FOT threshold and – in a local orbital basis – tightening the FOT results in rapid reduction of errors in the DEC amplitudes and energy. The DEC CC method scales linearly with system size and provides a predictable accuracy of the results. The method is also embarassingly parallel (the orbital fragment calculations are independent), making it computationally efficient for calculations on large molecular systems on contemporary computer architectures.

ACKNOWLEDGMENTS

This work has been supported by the Lundbeck Foundation and the Danish Center for Scientific Computing (DCSC).

APPENDIX A: LAGRANGIAN COUPLED CLUSTER THEORY

The coupled cluster energy may be evaluated in terms of the CC Lagrangian

$$L_{CC} = \langle \mathrm{HF}|H\exp(T)|\mathrm{HF}\rangle + \sum_{\mu} \bar{t}_{\mu}\langle \mu|\exp(-T)H\exp(T)|\mathrm{HF}\rangle \tag{A1}$$

where $|\mathrm{HF}\rangle$ is the Hartree–Fock state. The first term is the standard coupled cluster energy

$$E_{\mathrm{CC}} = \langle \mathrm{HF}|H\exp(T)|\mathrm{HF}\rangle \tag{A2}$$

and the second term contains the cluster amplitude equations multiplied with the Lagrangian multipliers. H is the Hamiltonian and T is the cluster operator containing singles, doubles, etc. excitations,

$$T = T_1 + T_2 + \cdots \tag{A3}$$

where

$$T_1 = \sum_{AI} t_I^A a_A^\dagger a_I = \sum_{AI} t_I^A \tau_I^A, \tag{A4}$$

$$T_2 = \sum_{\substack{I>J\\A>B}} t_{IJ}^{AB} a_A^\dagger a_I a_B^\dagger a_J = \frac{1}{4}\sum_{IJAB} t_{IJ}^{AB}\tau_{IJ}^{AB}, \tag{A5}$$

and indices $I, J, \ldots$ $(A, B, \ldots)$ denote occupied (virtual) HF spin orbitals. $|\mu\rangle$ refers to the excitation manifold

$$|\mu\rangle = \tau_\mu|\mathrm{HF}\rangle \tag{A6}$$

where the excitation operator τ_μ refers to either singles τ_I^A, doubles τ_{IJ}^{AB} etc. excitation operators. Using a shorthand notation the cluster operator may be expressed as

$$T = \sum_{\mu} t_\mu \tau_\mu. \tag{A7}$$

The coupled cluster Lagrangian is variational with respect to the cluster amplitudes and multipliers and the equations for determining these thus becomes

$$\frac{\partial L_{CC}}{\partial \bar{t}_\nu} = \langle \nu|\exp(-T)H\exp(T)|\mathrm{HF}\rangle = 0, \tag{A8}$$

$$\frac{\partial L_{CC}}{\partial t_\nu} = \langle \mathrm{HF}|H\tau_\nu\exp(T)|\mathrm{HF}\rangle + \sum_{\mu}\bar{t}_\mu\langle\mu|\exp(-T)[H,\tau_\nu]\exp(T)|\mathrm{HF}\rangle = 0. \tag{A9}$$

The advantage of using the energy expression L_{CC} compared to E_{CC} is that when L_{CC} is used the error in the energy will be bilinear in the error in the amplitudes and multipliers. To see this assume that we have determined a set of approximate amplitudes t_μ^A and multipliers $\bar{t}_\mu^A$ and have expressed these in terms of the optimized amplitudes t_μ^* and multipliers $\bar{t}_\mu^*$ (which satisfy Eqs. (A8) and (A9)) and some correction terms

$$t_\mu^A = t_\mu^* + \delta t_\mu, \tag{A10}$$

$$\bar{t}_\mu^A = \bar{t}_\mu^* + \delta\bar{t}_\mu. \tag{A11}$$

Evaluating the energy for the approximate amplitudes and multipliers using Eq. (A1) gives

$$\begin{aligned} L_{CC}^A &= \langle \mathrm{HF}|H\exp(T^*+\delta T)|\mathrm{HF}\rangle + \sum_{\mu}(\bar{t}_\mu^* + \delta\bar{t}_\mu) \\ &\quad\times\langle\mu|\exp(-T^*-\delta T)H\exp(T^*+\delta T)|\mathrm{HF}\rangle \\ &= \langle \mathrm{HF}|H\exp(T^*)|\mathrm{HF}\rangle + \sum_{\mu}(\bar{t}_\mu^* + \delta\bar{t}_\mu) \\ &\quad\times\langle\mu|\exp(-T^*)H\exp(T^*)|\mathrm{HF}\rangle \end{aligned}$$

$$+\langle \mathrm{HF}|H\delta T\exp(T^*)|\mathrm{HF}\rangle$$
$$+\sum_\mu \bar{t}^*_\mu \langle \mu|\exp(-T^*)[H,\delta T]\exp(T^*)|\mathrm{HF}\rangle$$
$$+\mathcal{O}(\delta t^2,\delta\bar{t}^2,\delta t\delta\bar{t}), \tag{A12}$$

where T^* and δT are cluster operators that contain the optimized amplitudes and correction amplitudes respectively. We have used that T^* and δT commute and performed a Baker-Campbell-Hausdorff expansion of $\exp(-\delta T)H\exp(\delta T)$. The second term in Eq. (A12) is zero because it contains the amplitude equation in Eq. (A8) for optimized amplitudes. Using $\delta T = \sum_\nu \delta t_\nu \tau_\nu$ the third and fourth term in Eq. (A12) may be expressed in terms of the multiplier equation in Eq. (A9) and they therefore vanish. Equation (A12) may therefore be expressed as

$$L^A_{CC} = L^*_{CC} + \mathcal{O}(\delta t^2,\delta\bar{t}^2,\delta t\delta\bar{t}), \tag{A13}$$

where $L^*_{CC} = E^*_{CC}$ is the energy for the optimized amplitudes and multipliers. From Eq. (A13) we see that the CC Lagrangian has a bilinear precision in the errors in the amplitudes and multipliers. Thus, if the coupled cluster energy is known to have an error which is proportional to δ then the errors in the amplitudes and multipliers will be approximately proportional to $\delta^{\frac{1}{2}}$. We note that for MP2 the amplitude and multiplier equations are identical in the spin orbital basis.[41] The error in the Lagrangian correlation energy is therefore quadratic in the errors in the amplitudes for MP2.

APPENDIX B: ORBITAL INVARIANCE OF THE COUPLED CLUSTER ENERGY

1. Orbital invariance of the standard coupled cluster energy

In this section, we summarize the proof that the standard coupled cluster energy is invariant with respect to rotations among the occupied and/or virtual HF orbitals, and that it is termwise so. This proof is important for understanding orbital rotation invariance properties for atomic site and pair interaction energies entering the DEC scheme.

Consider initially the amplitude equations in a HF spin orbital basis where an orthogonal transformation has been carried out among the occupied and virtual HF orbitals. Operators and states in the rotated basis are denoted with a tilde, i.e., the creation and annihilation operators in the rotated basis are given by[41]

$$\tilde{a}^\dagger_A = \exp(-\kappa)a^\dagger_A\exp(\kappa) = \sum_C a^\dagger_C[\exp(-\boldsymbol{\kappa})]_{CA}, \tag{B1}$$

$$\tilde{a}_I = \exp(-\kappa)a_I\exp(\kappa) = \sum_K a_K[\exp(-\boldsymbol{\kappa})]_{KI}, \tag{B2}$$

where $\boldsymbol{\kappa}$ is a real anti-symmetric matrix and κ an anti-Hermitian orbital rotation operator

$$\kappa = \sum_{K>L}\kappa_{KL}\left(a^\dagger_K a_L - a^\dagger_L a_K\right) + \sum_{C>D}\kappa_{CD}\left(a^\dagger_C a_D - a^\dagger_D a_C\right). \tag{B3}$$

The rotated Hartree–Fock state and the rotated excitation manifold are connected to the non-rotated HF state and excitation manifold as

$$|\widetilde{\mathrm{HF}}\rangle = \exp(-\kappa)|\mathrm{HF}\rangle = |\mathrm{HF}\rangle, \tag{B4}$$

$$|\tilde{\mu}\rangle = \exp(-\kappa)|\mu\rangle, \tag{B5}$$

where we have used that $\kappa|\mathrm{HF}\rangle = 0$ to obtain Eq. (B4). The CC amplitude equations in the rotated basis read

$$\langle\tilde{\mu}|\exp(-\tilde{T})H\exp(\tilde{T})|\widetilde{\mathrm{HF}}\rangle = 0. \tag{B6}$$

Equation (B6) holds for any state $\langle\tilde{\mu}|$ in the rotated basis, and it therefore also holds for any linear combination of excited states in the rotated basis. In particular, since there is a non-singular transformation matrix $\boldsymbol{\kappa}$ connecting $|\tilde{\mu}\rangle$ and $|\mu\rangle$, Eq. (B6) is also satisfied for $|\mu\rangle$, i.e.,

$$\langle\mu|\exp(-\tilde{T})H\exp(\tilde{T})|\mathrm{HF}\rangle = 0, \tag{B7}$$

where we also have used Eq. (B4). The cluster operator in the transformed basis $\tilde{T}$ may be expressed in terms of operators in the untransformed basis. Consider initially the singles excitation operator. Using Eqs. (B1) and (B2), we obtain

$$\tilde{T}_1 = \sum_{CK}\tilde{t}^C_K\tilde{a}^\dagger_C\tilde{a}_K = \sum_{AICK}\tilde{t}^C_K a^\dagger_A[\exp(-\boldsymbol{\kappa})]_{AC}a_I[\exp(-\boldsymbol{\kappa})]_{IK}$$
$$= \sum_{AI}\left(\sum_{CK}\tilde{t}^C_K[\exp(-\boldsymbol{\kappa})]_{AC}[\exp(-\boldsymbol{\kappa})]_{IK}\right)a^\dagger_A a_I. \tag{B8}$$

The amplitudes entering the non-rotated basis in Eq. (A8) and in Eq. (B7) are both solutions to the amplitude equations, and therefore $\tilde{T}_1$ must be identical to T_1 term by term when the amplitudes are expressed in the same basis. By comparing Eqs. (A4) and (B8), it thus follows that

$$t^A_I = \sum_{CK}\tilde{t}^C_K[\exp(-\boldsymbol{\kappa})]_{AC}[\exp(-\boldsymbol{\kappa})]_{IK}. \tag{B9}$$

Equation (B9) shows how the cluster singles amplitudes in the rotated basis is connected to cluster amplitudes in the non-rotated basis. Likewise for the higher excitation operators, i.e., $T_n = \tilde{T}_n$ for $n = 1, 2, \ldots$, and therefore T and $\tilde{T}$ are identical,

$$T = \tilde{T}. \tag{B10}$$

Using Eqs. (B4) and (B10), it now follows that the coupled cluster energy in Eq. (A2) is invariant with respect to an orthogonal transformation among the occupied and among the virtual HF orbitals,

$$\tilde{E}_{\mathrm{CC}} = \langle\widetilde{\mathrm{HF}}|H\exp(\tilde{T})|\widetilde{\mathrm{HF}}\rangle = \langle\mathrm{HF}|H\exp(T)|\mathrm{HF}\rangle = E_{\mathrm{CC}}. \tag{B11}$$

The correlation energy in Eq. (A2) may be expressed as

$$E_{\mathrm{CC}} = \langle\mathrm{HF}|[H,T_2] + \frac{1}{2}\big[[H,T_1],T_1\big]|\mathrm{HF}\rangle. \tag{B12}$$

From Eq. (B11) it follows that the individual terms in the correlation energy are invariant with respect to a rotation between the occupied and/or virtual orbitals. MP2 is a subset of coupled cluster, and therefore the MP2 energy is also invariant with respect to rotations among occupied and among virtual orbitals.

The invariance of the CC energy is also valid when rotations are considered only among the occupied orbitals or only among the virtual orbitals. This is easily seen since rotations only among occupied orbitals is accomplished by using $[\exp(-\kappa)]_{AC} = \delta_{AC}$, and similarly $[\exp(-\kappa)]_{IK} = \delta_{IK}$ for rotations only among the virtual orbitals. This means that for a given subset of the occupied orbitals, one may still rotate freely among the full set of virtual orbitals, and vice versa. For example, the T_2 atomic site energy contribution

$$E_{\mathrm{P}}^{\mathrm{o}}(T_2) = \frac{1}{2}\sum_{IJ\in\underline{P}}\sum_{AB}\langle \mathrm{HF}|\left[H, a_A^\dagger a_I a_B^\dagger a_J\right]|\mathrm{HF}\rangle t_{IJ}^{AB} \tag{B13}$$

is invariant with respect to rotations among virtual orbitals. Likewise,

$$E_{\mathrm{P}}^{\mathrm{v}}(T_2) = \frac{1}{2}\sum_{IJ}\sum_{AB\in\overline{P}}\langle \mathrm{HF}|\left[H, a_A^\dagger a_I a_B^\dagger a_J\right]|\mathrm{HF}\rangle t_{IJ}^{AB} \tag{B14}$$

is invariant with respect to rotations among occupied orbitals.

2. Orbital invariances of the Lagrangian coupled cluster energy

In this section, we discuss orbital invariances of the Lagrangian coupled cluster energy. We consider initially the Lagrangian multiplier equation in Eq. (A9) in the rotated HF basis,

$$\langle\widetilde{\mathrm{HF}}|H\tilde{\tau}_\nu\exp(\tilde{T})|\widetilde{\mathrm{HF}}\rangle + \sum_{\tilde{\mu}}\bar{t}_{\tilde{\mu}}\langle\tilde{\mu}|\exp(\tilde{T})[H,\tilde{\tau}_\nu] \times\exp(\tilde{T})|\widetilde{\mathrm{HF}}\rangle = 0, \tag{B15}$$

where $\bar{t}_{\tilde{\mu}}$ denotes a multiplier in the rotated basis. Equation (B15) holds for any excitation operator $\tilde{\tau}_\nu$ and therefore also for τ_ν since there is a non-singular orbital transformation connecting $\tilde{\tau}_\nu$ and τ_ν. Using Eqs. (B4) and (B10), we may therefore express Eq. (B15) as

$$\langle\mathrm{HF}|H\tau_\nu\exp(T)|\mathrm{HF}\rangle + \sum_{\tilde{\mu}}\bar{t}_{\tilde{\mu}}\langle\tilde{\mu}|\exp(-T)[H,\tau_\nu] \times\exp(T)|\mathrm{HF}\rangle = 0. \tag{B16}$$

Comparing Eq. (B16) and Eq. (A9), it follows that

$$\sum_\mu \bar{t}_\mu\langle\mu| = \sum_{\tilde{\mu}}\bar{t}_{\tilde{\mu}}\langle\tilde{\mu}|. \tag{B17}$$

Using Eqs. (B4), (B10), and (17) we thus see that the Lagrangian energy is invariant with respect to an orthogonal transformation among the occupied and/or virtual orbitals

$$\begin{aligned}\tilde{L}_{\mathrm{CC}} &= \langle\widetilde{\mathrm{HF}}|H\exp(\tilde{T})|\widetilde{\mathrm{HF}}\rangle \\ &\quad+\sum_{\tilde{\mu}}\bar{t}_{\tilde{\mu}}\langle\tilde{\mu}|\exp(-\tilde{T})H\exp(\tilde{T})|\widetilde{\mathrm{HF}}\rangle \\ &= \langle\mathrm{HF}|H\exp(T)|\mathrm{HF}\rangle \\ &\quad+\sum_\mu\bar{t}_\mu\langle\mu|\exp(-T)H\exp(T)|\mathrm{HF}\rangle = L_{\mathrm{CC}}.\end{aligned} \tag{B18}$$

The coupled cluster Lagrangian energy may be expressed as

$$\begin{aligned}L_{\mathrm{CC}} &= \langle\mathrm{HF}|[H,T_2] + \frac{1}{2}[[H,T_1],T_1]|\mathrm{HF}\rangle \\ &\quad+\sum_\mu\bar{t}_\mu\langle\mathrm{HF}|\tau_\mu^\dagger\Bigg(H+[H,T]+\frac{1}{2}[[H,T],T] \\ &\quad+\frac{1}{6}[[[H,T],T],T]+\frac{1}{24}[[[[H,T],T],T],T]\Bigg)|\mathrm{HF}\rangle.\end{aligned} \tag{B19}$$

From Eq. (B18) it follows that the individual contributions in Eq. (B19) are invariant with respect to an orthogonal transformation among the occupied and/or virtual orbitals. Further if in Eq. (B19) H is replaced by the Fock operator each individual contribution will be invariant with respect to an orthogonal transformation among the occupied and/or virtual orbitals.

The term $\sum_\mu\bar{t}_\mu\langle\mathrm{HF}|\tau_\mu^\dagger[F,T]|\mathrm{HF}\rangle$ corresponds to the second and fourth terms in Eq. (12). Since F has an occupied-occupied and virtual-virtual block structure, an argument similar to the one leading to Eqs. (B13) and (B14) shows that the contributions from the second term in Eq. (12) is invariant with respect to rotation among the virtual orbitals, while the fourth term in Eq. (12) is invariant with respect to rotations among the occupied orbitals.

APPENDIX C: DISTANCE DEPENDENCE OF PAIR INTERACTION ENERGIES

In this appendix, we demonstrate that – for a local HF basis – the pair interaction energies decay with the inverse pair distance to the sixth power. For simplicity in the following analysis we now consider only the MP2 model and therefore the singles contributions are omitted. Eqs. (13) and (14) may then be written as

$$\begin{aligned}\Delta E_{\mathrm{PQ}}^{\mathrm{o}} &= \sum_{\substack{i\in P\\ j\in Q}}\sum_{ab}t_{ij}^{ab}(2g_{iajb}-g_{ibja}) \\ &\quad+\sum_{\substack{i\in Q\\ j\in P}}\sum_{ab}t_{ij}^{ab}(2g_{iajb}-g_{ibja})\end{aligned} \tag{C1}$$

and

$$\begin{aligned}\Delta E_{\mathrm{PQ}}^{\mathrm{v}} &= \sum_{\substack{a\in P\\ b\in Q}}\sum_{ij}t_{ij}^{ab}(2g_{iajb}-g_{ibja}) \\ &\quad+\sum_{\substack{a\in Q\\ b\in P}}\sum_{ij}t_{ij}^{ab}(2g_{iajb}-g_{ibja}),\end{aligned} \tag{C2}$$

where the MP2 amplitudes are determined from the MP2 amplitudes equations

$$\sum_c t_{ij}^{cb}F_{ac}+\sum_c t_{ij}^{ac}F_{bc}-\sum_k t_{kj}^{ab}F_{ki}-\sum_k t_{ik}^{ab}F_{kj} = -g_{iajb}, \tag{C3}$$

where $\boldsymbol{F}$ is the Fock matrix. The Fock matrix is block diagonal with vanishing occupied-virtual matrix elements due to the HF orbital optimization condition.

We now examine the decay of $\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$ for large distances between atomic sites P and Q, i.e., for large r_{PQ}. For a diagonally dominant Fock matrix the cluster amplitudes in Eq. (C3) may be approximated as

$$t_{ij}^{ab} \approx -g_{iajb}\left(\Delta F_{ij}^{ab}\right)^{-1}, \tag{C4}$$

where

$$\Delta F_{ij}^{ab} = F_{aa} + F_{bb} - F_{ii} - F_{jj} \tag{C5}$$

and the dominant contribution to $\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$ in Eq. (C1) arising from the first term is

$$M_{PQ} = -\sum_{\substack{i\in P\\ j\in Q}}\sum_{ab} g_{iajb}(2g_{iajb} - g_{ibja})\left(\Delta F_{ij}^{ab}\right)^{-1}. \tag{C6}$$

A similar contribution arises from the second term in Eq. (C1) where i resides on Q and j on P. For local HF orbitals the first integral g_{iajb} entering Eq. (C6) is only nonzero if ϕ_a is local to ϕ_i and ϕ_b is local to ϕ_j. For large r_{PQ} distances the charge distributions $\phi_i^*\phi_b$ and $\phi_j^*\phi_a$ will thus be vanishing since ϕ_b (ϕ_a) cannot be simultaneously local to both ϕ_i and ϕ_j. Hence, the contributions from g_{ibja} in Eq. (C6) can be neglected, and the dominant contribution in Eq. (C6) may be expressed as

$$M_{PQ} \approx -\sum_{\substack{i\in P\\ j\in Q}}\sum_{ab} 2g_{iajb}^2\left(\Delta F_{ij}^{ab}\right)^{-1}. \tag{C7}$$

We now investigate how M_{PQ} decays as a function of r_{PQ} by carrying out a multipole expansion of the integral g_{iajb}. Introducing the charge distributions $\Omega_{ia} = \phi_i^*\phi_a$ centered at X and $\Omega_{jb} = \phi_j^*\phi_b$ centered at Y and following Ref. 41 a multipole expansion of g_{iajb} can be expressed as

$$g_{iajb} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sum_{k=0}^{\infty}\sum_{n=-k}^{k} q_{lm}^{ia}(X)T_{lm,kn}(\boldsymbol{r}_{XY})q_{kn}^{jb}(Y). \tag{C8}$$

The multipole moments of the charge distribution Ω_{ia} are given as

$$q_{lm}^{ia}(X) = \int \Omega_{ia}(\boldsymbol{r})R_{lm}(\boldsymbol{r}_X)d\boldsymbol{r}, \tag{C9}$$

where the integration variable $\boldsymbol{r}$ refers to an electronic coordinate, and where

$$R_{lm}(r_X) = \frac{1}{\sqrt{(l-m)!(l+m)!}} r_X^l C_{lm}(\theta,\phi), \tag{C10}$$

and C_{lm} are complex solid harmonics in Racah's normalization. The multipole moment for the charge distribution Ω_{jb} is given in a similar way. The interaction matrix in Eq. (C8) is defined as

$$T_{lm,kn}(\boldsymbol{r}_{XY}) = (-1)^k I^*_{l+k,m+n}(\boldsymbol{r}_{XY}), \tag{C11}$$

where

$$I_{lm}(\boldsymbol{r}_{XY}) = \sqrt{(l-m)!(l+m)!}\,\boldsymbol{r}_{XY}^{-l-1}C_{lm}(\theta,\phi). \tag{C12}$$

Since $R_{00} = 1$ and the HF orbitals are orthonormal, it follows that

$$q_{00}^{ia}(X) = \int \Omega_{ia}(\boldsymbol{r})R_{00}(\boldsymbol{r}_X)d\boldsymbol{r} = \int \Omega_{ia}(\boldsymbol{r})d\boldsymbol{r}$$
$$= \int \phi_i^*(\boldsymbol{r})\phi_a(\boldsymbol{r})d\boldsymbol{r} = 0, \tag{C13}$$

and similarly $q_{00}^{jb}(Y) = 0$. The multipole expansion in Eq. (C8) thus starts out with $l = 1$ and $k = 1$, and the first non-vanishing contribution in the interaction matrix becomes $I^*_{l+k,m+n}(\boldsymbol{r}_{XY}) = I^*_{2,m+n}(\boldsymbol{r}_{XY})$. Consequently, it follows from Eq. (C12) that the lowest order contribution to g_{iajb} decays as r_{XY}^{-3}. The dominant contributions to the pair interaction energy in Eq. (C6) contain g_{iajb}^2 and will hence decay as r_{XY}^{-6}. For local HF orbitals, r_{XY} may be approximated by r_{PQ} and the pair interaction energies will thus decay as r_{PQ}^{-6} as expected for dispersion effects. The smaller non-diagonal contributions to $\Delta E^{\mathrm{o}}_{\mathrm{PQ}}$ may likewise be shown to decay as r_{PQ}^{-6}. Using a similar analysis it may similarly be shown that $\Delta E^{\mathrm{v}}_{\mathrm{PQ}}$ in Eq. (C2) also decays as r_{PQ}^{-6}. We note that the r_{XY}^{-3} decay of g_{iajb} was used by Hetzer, Pulay and Werner in their multipole approximation to distant pair energies in local MP2 calculations.[44]

[1] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).
[2] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).
[3] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).
[4] C. Edmiston and K. Ruedenberg, J. Chem. Phys. **43**, S97 (1965).
[5] J. Pipek, Int. J. Quantum Chem. **36**, 487 (1989).
[6] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
[7] M. Ziółkowski, B. Jansík, P. Jørgensen, and J. Olsen, J. Chem. Phys. **131**, 124112 (2009).
[8] B. Jansík, S. Høst, K. Kristensen, and P. Jørgensen, J. Chem. Phys. **134**, 194104 (2011).
[9] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[10] S. Saebø and P. Pulay, Ann. Rev. Phys. Chem. **44**, 213 (1993).
[11] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[12] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[13] G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, J. Chem. Phys. **113**, 9443 (2000).
[14] M. Schütz, J. Chem. Phys. **113**, 9986 (2000).
[15] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[16] R. A. Mata and H.-J. Werner, J. Chem. Phys. **125**, 184110 (2006).
[17] G. E. Scuseria and P. Y. Ayala, J. Chem. Phys. **111**, 8330 (1999).
[18] O. Christiansen, P. Manninen, P. Jørgensen, and J. Olsen, J. Chem. Phys. **124**, 084103 (2006).
[19] V. Weijo, P. Manninen, P. Jørgensen, O. Christiansen, and J. Olsen, J. Chem. Phys. **127**, 074106 (2007).
[20] N. Flocke and R. J. Bartlett, J. Chem. Phys. **121**, 10935 (2004).
[21] S. Li, J. Ma, and Y. Jiang, J. Comput. Chem. **23**, 237 (2002).
[22] W. Li, P. Piecuch, J. R. Gour, and S. Li, J. Chem. Phys. **131**, 114109 (2009).
[23] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 8644 (2010).
[24] M. Kobayashi and H. Nakai, J. Chem. Phys. **129**, 044103 (2008).
[25] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **123**, 134103 (2005).
[26] H. Stoll, Chem. Phys. Lett. **191**, 548 (1992).
[27] J. Friedrich, M. Hanrath, and M. Dolg, J. Chem. Phys. **126**, 154110 (2007).
[28] J. Friedrich, M. Hanrath, and M. Dolg, Z. Phys. Chem. **224**, 513 (2010).
[29] J. E. Subotnik and M. Head-Gordon, J. Chem. Phys. **123**, 064108 (2005).
[30] J. E. Subotnik, A. Sodt, and M. Head-Gordon, J. Chem. Phys. **125**, 074116 (2006).
[31] J. E. Subotnik, A. Sodt, and M. Head-Gordon, J. Chem. Phys. **128**, 034103 (2008).
[32] J. Almlöf, Chem. Phys. Lett. **181**, 319 (1991).
[33] M. Häser, Theor Chim Acta **87**, 147 (1993).
[34] P. Y. Ayala and G. E. Scuseria, J. Chem. Phys. **110**, 3660 (1999).

[35]D. S. Lambrecht, B. Doser, and C. Ochsenfeld, J. Chem. Phys. **123**, 184102 (2005).
[36]B. Doser, D. S. Lambrecht, and C. Ochsenfeld, Chem. Phys. Phys. Chem **10**, 3335 (2008).
[37]B. Doser, J. Zienau, L. Clin, D. S. Lambrech, and C. Ochsenfeld, Z. Phys. Chem. **224**, 397 (2010).
[38]J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).
[39]M. Ziółkowski, B. Jansík, T. Kjærgaard, and P. Jørgensen, J. Chem. Phys. **133**, 014107 (2010).
[40]K. Kristensen, M. Ziółkowski, B. Jansík, T. Kjærgaard, and P. Jørgensen, J. Chem. Theory Comput. **7**, 1677 (2011).
[41]T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory, First Edition* (Wiley, New York, 2000).
[42]Maestro v. 8.5; Schrödinger, LLC: Cambridge, MA (2008); http://www.schrodinger.com.
[43]B. Jansík, S. Høst, K. Kristensen, and P. Jørgensen, J. Chem. Phys. **134**, 194104 (2011).
[44]G. Hetzer, P. Pulay, and H.-J. Werner, Chem. Phys. Lett. **290**, 143 (1998).