Local energy decomposition of coupled-cluster interaction energies: Interpretation, benchmarks, and comparison with symmetry-adapted perturbation theory

Ahmet Altun | Róbert Izsák | Giovanni Bistoni

Department of Molecular Theory and Spectroscopy, Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, Germany

Correspondence
Giovanni Bistoni, Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany. Email: giovanni.bistoni@kofo.mpg.de

Funding information
Deutsche Forschungsgemeinschaft

Abstract
Local energy decomposition analysis provides a breakdown of the domain-based local pair natural orbital CCSD(T) [DLPNO-CCSD(T)] energy into additive contributions representing the interaction between pairs of user-defined fragments. Each of these fragment-pairwise components can be further decomposed into a sum of physically meaningful terms, such as electrostatics, dispersion, and exchange. In this study, the dependence of such energy terms on the basis set size, the approximations used for the two-electron integrals, and the localization scheme used for the virtual orbitals have been carefully evaluated on the interaction energies of the S66 benchmark set. A comparison with the energy components obtained at the SAPT2 + (3)δiMP2 level of Symmetry-Adapted Perturbation Theory is also provided.

KEYWORDS
correlation energy, DLPNO-CCSD(T), energy decomposition analysis, interaction energy, London dispersion, resolution of identity

1 | INTRODUCTION

Noncovalent interactions (NCIs) are essential in all fields of chemical research. Thus, many quantum mechanical methods have been developed over the years that allow for their accurate quantification and analysis. These are typically grouped into two families, that is, perturbative and supramolecular approaches.[1]

The most popular perturbative approach is symmetry-adapted perturbation theory (SAPT).[2] In this method, the Hamiltonian of the system is partitioned as the sum of that of the monomers plus the intermolecular interaction operator. The interaction energy is then expressed as a perturbation series in the Rayleigh-Schrödinger (RS) perturbation theory framework, plus a series of energy corrections arising from the antisymmetry requirement of the wave function. SAPT and its variants[2–6] provide accurate NCI energies, as well as their breakdown into exchange-repulsion, electrostatic, induction, and London dispersion energies.

Within a supramolecular approach, NCI energies are computed as the difference between the energy of a molecular adduct and those of the fragments. Additional insights can be obtained with energy decomposition analysis (EDA) schemes,[7,8] which provide a SAPT-like decomposition of NCI energies computed within a supramolecular approach.

Correlated wave function methods are often used for calculating accurate NCI energies within a supramolecular approach. In particular, the coupled-cluster method with single, double, and perturbative treatment of triple excitations, that is, CCSD(T),[9] is known to provide extremely accurate results with deviations from reference experimental data that often fall within the experimental uncertainty.[10] Moreover, linear scaling local variants of the CCSD(T) method have been developed that allow for the calculation of CCSD(T)-quality energies for systems of hundreds
2 | THEORY

2.1 | Local energy decomposition

Given a molecular complex of two (or more) weakly interacting monomers, a supermolecular approach can be used to compute their interaction energy ($\Delta E_{\text{int}}$). If a post-HF method such as DLPNO-CCSD(T) is used, the resulting interaction energy can be divided into the HF ($\Delta E_{\text{HF}}$) and correlation ($\Delta E_{\text{cor}}$) contributions:

$$\Delta E_{\text{int}} = \Delta E_{\text{HF}} + \Delta E_{\text{cor}}$$  \hspace{1cm} (1)

Each of these contributions is obtained as the corresponding energy of the adduct minus the sum of those of its constituting monomers frozen in their in-adduct geometry.

Within the LED scheme, $\Delta E_{\text{HF}}$ can be decomposed into a series of terms by exploiting the localization of the occupied orbitals in the DLPNO-CCSD(T) method:[28,33]

$$\Delta E_{\text{HF}} = \Delta E_{\text{HF, prep}} + E_{\text{elstat}} + E_{\text{exch}}$$  \hspace{1cm} (2)

in which $\Delta E_{\text{HF, prep}}$ is the energy required to distort the electronic structure of the isolated monomers from their ground state to the one that is optimal for the interaction, and $E_{\text{elstat}}$ and $E_{\text{exch}}$ are the electrostatic and exchange interactions, respectively, between the monomers.

This decomposition is achieved by assigning each orbital to the fragment in which it is dominantly localized, as detailed in Ref. [28]. As an example, here, we report the expressions of $E_{\text{elstat}}$ and $E_{\text{exch}}$ for an arbitrary pair of fragments X and Y:

$$E_{\text{elstat}} = P_{X-Y} \left[ \sum_{A_i \neq B_j} \frac{Z_X^{(X)} Z_Y^{(Y)}}{\| r_{A_i}^{(X)} - r_{B_j}^{(Y)} \|} - 2 \sum_{i_x \neq j_y} \left\langle i_x \left| \frac{Z_X^{(Y)}}{\| r_{B_j}^{(Y)} - r_{A_i}^{(X)} \|} \right| j_y \right\rangle + 4 \sum_{i_x > j_y} \left\langle i_x | j_y \right\rangle \right]$$  \hspace{1cm} (3)

$$E_{\text{exch}} = P_{X-Y} \left[ -2 \sum_{i_x > j_y} \left( i_x | j_y \right) \right]$$  \hspace{1cm} (4)
where $Z_A$ is the nuclear charge of atom A at position $R_A$; $i$ and $j$ label the localized occupied orbitals (the subscripts $X$ and $Y$ denote the fragment in which the orbital is mostly localized); and $\langle \mu \nu \alpha \beta \rangle$ and $\langle \mu \nu \alpha \beta \rangle$ are two-electron integrals in Mulliken notation. Typically, the calculation of such two-electron integrals is the most computationally demanding step in LED calculations. Various integral decomposition techniques can be used to increase the efficiency of the technique, as detailed in the following section.

It is worth mentioning here that the dependence of the LED terms on the localization scheme used for the internal orbital space was already investigated in the original LED paper.\[^{[28]}\] Although Foster-Boys (FB)\[^{[57]}\] and Pipek-Mezey (PM)\[^{[58]}\] localization schemes provide an analogous partitioning of HF energies, in some cases (e.g., certain delocalized $\pi$ systems), noticeable deviations might occur. Generally, FB localization is recommended as it typically provides more localized orbitals.

The DLPNO-CCSD(T) correlation energy can be expressed as a sum of double-excitation contributions from the localized occupied orbitals to the virtual space, which is spanned by a set of pair natural orbitals (PNOs) that is specific for each electron pair. In LED, PNOs are subjected to an additional localization procedure, such as FB and PM, and then assigned to the fragment in which they are dominantly localized.

By exploiting the localization of both occupied and virtual orbitals,\[^{[28]}\] the correlation energy can be expressed as the sum of double-excitation contributions from pairs of occupied orbitals to the localized PNOs. This allows one to identify double-excitation types as given in Scheme 1.

By exploiting this decomposition, $\Delta E_{\text{int}}$ can be written as

$$\Delta E_{\text{int}} = \Delta E_{\text{int}}^{\text{CCSD}} + \Delta E_{\text{int}}^{\text{C}^-(T)} = E_{\text{disp}}^{\text{CCSD}} + \Delta E_{\text{no-disp}}^{\text{CCSD}} + \Delta E_{\text{int}}^{\text{C}^-(T)}$$

where $E_{\text{disp}}^{\text{CCSD}}$ is the London dispersion attraction between the fragments, and $\Delta E_{\text{no-disp}}^{\text{CCSD}}$ is the nondispersive correlation, which, among other things, corrects for the permanent and induced electrostatics approximately described by the HF method. Finally, $\Delta E_{\text{int}}^{\text{C}^-(T)}$ is the perturbative triple contribution to the interaction energy. Note that $\Delta E_{\text{int}}^{\text{C}^-(T)}$ can be further decomposed into dispersive and nondispersive contributions.\[^{[34]}\] This allows us to define the overall DLPNO-CCSD(T) dispersion $E_{\text{disp}}^{\text{CCSD}(T)}$ and nondispersion $\Delta E_{\text{no-disp}}^{\text{CCSD}(T)}$ correlation terms. This approach is used in this study in the comparison of LED and SAPT terms (Section 4.4). A more detailed description of the terms just discussed can be found in Refs. [29, 31, 34].

**SCHEME 1** Double-excitation contributions to the correlation energy

| Dispersion Excitations | Non-dispersive Excitations |
|------------------------|---------------------------|
| $X$ $Y$                | $X$ $Y$                   |
| $X$ $Y$                | $X$ $Y$                   |
| $X$ $Y$                | $X$ $Y$                   |
| $X$ $Y$                | $X$ $Y$                   |
| $X$ $Y$                | $X$ $Y$                   |
| $X$ $Y$                | $X$ $Y$                   |
2.2 | LED/RIJCOSX vs LED/RIJK

We now turn to the evaluation of the two-electron integral contributions in Equations (3) and (4). Introducing the density belonging to a given fragment $X$ in terms of the molecular orbital (MO) coefficients $c_{i\mu}$,

$$D_{\mu\nu}^X = \sum_i c_{i\mu} c_{i\nu}$$  \hspace{1cm} (6)

we can write for the Coulomb part:

$$\sum_{\ell x} \langle \ell x | j_y | j_y \rangle = \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda) D_{\mu\kappa}^X D_{\nu\lambda}^Y = \sum_{\mu\nu\kappa\lambda} J_{\mu\nu} D_{\mu\nu}^X$$  \hspace{1cm} (7)

and for the exchange part:

$$\sum_{\ell x} \langle \ell x | l_x j_y l_x j_y \rangle = \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda) D_{\mu\kappa}^X D_{\nu\lambda}^X = \sum_{\mu\nu\kappa\lambda} K_{\mu\nu} D_{\mu\nu}^X$$  \hspace{1cm} (8)

in which we have implicitly defined the Coulomb and exchange matrixes $J_{\mu\nu}$ and $K_{\mu\nu}$, respectively. The Greek indices denote the atomic orbital (AO) basis. The evaluation of both terms in the AO basis scales as $O(N^4)$, where $N$ is the number of AO basis functions. The various integral decomposition schemes improve efficiency by accelerating the evaluation of these terms.

For the Coulomb term, the most successful approximation has been RIJ, which relies on the following decomposition pattern:

$$J_{\mu\nu} \approx \sum_{AB} (\mu\nu|A) V_{AB} d_{B}$$  \hspace{1cm} (9)

where $V_{AB}$ is the two-index Coulomb integral, and the indices $A$, $B$ label auxiliary basis functions. The advantage of this decomposition is that the density can be contracted with a three-index integral to yield the vector $d_{B}$ as

$$d_{B} = \sum_{\kappa\lambda} D_{\kappa\lambda}^Y (\kappa\lambda|B)$$  \hspace{1cm} (10)

The most expensive steps in the above equations scale formally as $O(LN^2)$, where $L$ is the number of auxiliary basis functions. The method can be implemented efficiently to approach quadratic scaling with a small prefactor.\[^{[45]}\] In typical calculations, Coulomb terms usually cost much less than exchange terms.

The RI for exchange approach can be combined with RIJ, but the resulting RIJK method does not reduce the formal scaling of the exchange term as it is not possible to fully contract $D_{\mu\nu}^X$ with only one of the three index integrals. In this case, it is more advantageous to use the molecular orbital coefficients as the number of occupied MOs $i$ is much smaller than the dimension of the AO basis

$$K_{\mu\nu} \approx \sum_{ABi} (\mu\nu|A) V_{AB} (B|i\nu)$$  \hspace{1cm} (11)

As a result, the method scales as $O(N_O LN^2)$, where $N_O$ is the number of occupied MOs. As this is usually small, the prefactor is expected to be much less than the number of occupied MOs $i$.

Alternatively, RIJ may be combined with the COS approximation for the exchange, yielding the RIJCOSX approximation. The exchange term is then approximated as

$$K_{\mu\nu} \approx \sum_{g} Q_{g} \sum_{i} A_{g}^i \sum_{\kappa\lambda} D_{\kappa\lambda}^X X_{g\kappa\lambda}$$  \hspace{1cm} (12)

in which $X_{g\kappa\lambda}$ is the value of the AO $\kappa$ at grid point $g$ multiplied by the square root of the grid weight. $Q_{g}$ can be obtained from $X_{g\kappa\lambda}$ by the overlap fitting process.\[^{[52]}\] that is, by requiring that the product of the two matrices is just the AO overlap matrix. Finally,

$$A_{g}^i = \int \frac{\psi(\mathbf{r})\lambda(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_g|} d\mathbf{r}$$  \hspace{1cm} (13)
The formal scaling of these steps is $O(MN^2)$, where $M$ is the number of grid points. In practice, the scaling may be reduced to one between linear and quadratic using proper prescreening techniques. As $M$ is much larger than the number of basis functions, the method has a lower formal scaling than RUK but a larger prefactor. Thus, it is expected that RUK will perform better for smaller molecules, and RIJCOSX will take over as the system size grows. Note that the summation over grid points can be divided into many independent batches, which makes RIJCOSX an ideal approach for parallelization.

3 | COMPUTATIONAL DETAILS

All DLPNO-CCSD(T)/LED calculations were performed with a development version of the ORCA 4.2 program package. The accuracy and efficiency of RIJK and RIJCOSX approximations in the decomposition of HF energies are assessed in this work. The localization of the occupied orbitals was carried out using the FB scheme. This method typically provides a satisfactorily local representation of the occupied orbital space, and it is thus chosen as the default localization scheme in LED applications.

DLPNO-CCSD(T) correlation energies were calculated using the default frozen core settings of ORCA. TightPNO settings were used. All correlated electron pairs were included in the CCSD treatment ($T_{\text{CutPairs}} = 0$). The perturbative triples contributions were calculated using the iterative ($T_1$) algorithm. In LED calculations, FB and PM schemes were tested for the localization of the PNOs in the correlation energy decomposition.

The aug-cc-pVnZ ($n = T$ and Q) correlation consistent basis sets were used in conjunction with their matching auxiliary counterparts. Interaction energies were corrected for the basis set superposition error (BSSE). As for COSX integration grids, the default GRIDXS2 and a larger and more accurate GRIDX7 were used in this study.

4 | RESULTS AND DISCUSSION

In this section, the dependence of the LED results by the technical parameters of the calculation is initially evaluated (Section 4.1: basis set; Section 4.2: localization scheme; Section 4.3: integral approximations). Then, LED and SAPT results are compared on the S66 benchmark set. For a detailed discussion of the accuracy of the DLPNO-CCSD(T) method in the calculation of relative energies, the reader is referred to Refs. [27–31].

4.1 | Basis set convergence

Figure 1 shows HF and DLPNO-CCSD(T) correlation interaction energies, that is, $\Delta E_{\text{HF}}$ and $\Delta E_{\text{CCSD}}$, respectively, for the 66 reactions of the S66 set computed with the aug-cc-pVTZ and aug-cc-pVQZ basis sets. As expected, HF interaction energies show only negligible differences with the two basis sets, while the convergence of correlation energies with respect to the basis set size is much slower. In particular, $\Delta E_{\text{HF}}$ increases up to 0.53 kcal mol$^{-1}$ by increasing the basis set size, with a mean absolute deviation (MAD) between aug-cc-pVTZ and aug-cc-pVQZ results of 0.146 kcal mol$^{-1}$.

Figure 2 shows the effect of the basis set size on the decomposed HF interaction energy terms, that is, $\Delta E_{\text{elprep}}$ and $\Delta E_{\text{elstat}}$, and $\Delta E_{\text{exch}}$. Consistent with the behavior of the overall $\Delta E_{\text{HF}}$ (MAD = 0.019 kcal mol$^{-1}$), the decomposed terms are not significantly affected by the basis set size (MAD = 0.013, 0.024, and 0.004 kcal mol$^{-1}$, respectively).

Figure 3 shows the basis set convergence behavior of the LED components of $\Delta E_{\text{HF}}$, that is, $\Delta E_{\text{elprep}}$ (Figure 3A), $\Delta E_{\text{elstat}}$ (Figure 3B), and $\Delta E_{\text{exch}}$ (Figure 3C). $\Delta E_{\text{elprep}}$ and $\Delta E_{\text{elstat}}$ values demonstrate the weakest dependence on the basis set size (MAD = 0.016 kcal mol$^{-1}$), which is comparable to that found for the reference energy (MAD = 0.019 kcal mol$^{-1}$). $\Delta E_{\text{exch}}$ values also show a weak dependence on the basis set size (0.073 kcal mol$^{-1}$). Note that, in our recent HFLD study, it was pointed out that the small aug-cc-pVDZ basis set already provides nearly converged $\Delta E_{\text{elstat}}$ values. Hence, the dominant part of the variation observed for $\Delta E_{\text{elstat}}$, going from aug-cc-pVTZ to aug-cc-pVQZ arises from the $\Delta E_{\text{elprep}}$ component (MAD = 0.139 kcal mol$^{-1}$). This behavior is somewhat consistent with the notion that short-range correlation effects are the most challenging to treat using wavefunction-based methods.

Finally, it is worth emphasizing that $\Delta E_{\text{elstat}}$ incorporates the effect of many different double-excitation classes, as shown in Scheme 1. A detailed analysis of the basis set dependence of these terms is given in the Figure S1.
4.2 The effect of localization scheme for PNOs

Figure 4 shows the impact of the localization scheme used for localizing the PNOs (FB vs PM) on $E_{\text{disp}}^{\text{CCSD}}$ and $\Delta E_{\text{no-disp}}^{\text{CC}}$. Both contributions are only weakly affected by the localization scheme adopted. Typically, FB provides slightly larger absolute values for the individual contributions. Importantly, the PNOs obtained from FB localization are significantly more local than those obtained with PM. For this reason, FB is chosen as the default for the localization of the virtual space in LED calculations, starting from ORCA 4.1. This aspect is further discussed in the Figure S2.

4.3 Accuracy and efficiency of integral approximations

In this section, the efficiency and accuracy of the RIJK and RIJCOSX approximations for the reference HF energy decomposition are discussed. For the S66 set, the RIJK and RIJCOSX errors in the calculation of the overall $\Delta E_{\text{HF}}^{\text{int}}$ are shown in Figure 5.

As expected, RIJK provides almost exact interaction energies when used in conjunction with very large auxiliary JK basis, as shown in Figure 5A. If the smaller aug-cc-pVTZ/JK basis is used, the error in individual reactions reaches up to 0.02 kcal mol$^{-1}$ (MAD = 0.006 kcal mol$^{-1}$, see Figure 5B).

If RIJCOSX is used instead of RIJK to approximate the two-electron integrals, a MAD of only 0.001 kcal mol$^{-1}$ is obtained with a conservative grid “GridX7”, as shown in Figure 5C. With the default grid “GridX” (Figure 5D), the maximum error increases to ~0.06 kcal mol$^{-1}$, with a MAD of 0.006 kcal mol$^{-1}$. These results confirm that these two approaches provide very robust results for relative energies, as already discussed in previous papers.[47-54,65]

It is now interesting to discuss how these approximations influence the individual LED energy components that constitute the reference interaction energy (Figure 6). All calculations were carried out at the HF/aug-cc-pVTZ level. Reference values were obtained using LED/RIJK results in conjunction with a large aug-cc-pV5Z/JK basis as this approach has close to zero MAD for the overall $\Delta E_{\text{HF}}^{\text{int}}$ (Figure 5A).
In all cases, the largest error was obtained for $E_{\text{elstat}}$ and $\Delta E^{\text{HF}}_{\text{el}-\text{prep}}$ if RIJCOSX is used in combination with the default COSX grid. This error originates from the fact that same density is used to compute the Coulomb and exchange contributions, and COSX seems to have a larger indirect effect on the Coulomb component rather than a direct one on the exchange. The error on $E_{\text{elstat}}$ and $\Delta E^{\text{HF}}_{\text{el}-\text{prep}}$ amounts to 0.07 kcal mol$^{-1}$, which is of the order of the inherent accuracy of the reference interaction energies in the S66 set.[66] Therefore, it can be concluded that all these approaches provide relatively small errors in all cases.

We now discuss the efficiency of RIJK and RIJCOSX approximations in LED calculations. The average computational time required for performing an LED calculation on the DLPNO-CCSD/aug-cc-pVTZ energy is given in Figure 7A for the dimers in the S66 set by using RIJK and RIJCOSX in the HF part. In addition, the time required for the LED calculations at the HFLD/TightPNO/aug-cc-pVDZ level is given in Figure 7B on the 3a (126 atoms) and 7b (153 atoms) molecules of S12L set when using RIJK and RIJCOSX/GridX7 in the HF part. On both benchmark sets, RIJCOSX is at least twice more efficient than RIJK in the LED calculations.
4.4 | Comparison with SAPT

Within the SAPT framework,[2] intermolecular interactions are treated as a perturbation to the energy of the isolated monomers. Many SAPT variants exist, differing for the level of theory used for description of the monomers and for how the perturbation series is truncated.[1–6]

Typically, the SAPT interaction energy $E^{\text{SAPT}}$ is expressed as the sum of a few physically meaningful energy terms, such as exchange-repulsion $E_{\text{exch-rep}}$, electrostatic interaction $E_{\text{elstat}}$, induction $E_{\text{ind}}$, and dispersion $E_{\text{disp}}$. For example, at the SAPT2 + (3)$\delta$MP2 level, the interaction energy $E^{\text{SAPT2+(3)$\delta$MP2}}$ reads:[1]

**FIGURE 3** Basis set dependence of the decomposed DLPNO-CCSD correlation interaction energy terms in terms of local energy decomposition on the S66 set. A, basis set dependence of the dispersion energy $E_{\text{disp}}^{\text{C-CCSD}}$. B, basis set dependence of the nondispersive interaction energy $\Delta E_{\text{no-disp}}^{\text{C-CCSD}}$. C, basis set dependence of the triples contribution to the interaction energy $\Delta E_{\text{int}}^{\text{C--(T)}}$. The FB localization scheme was used for both occupied orbitals and pair natural orbitals. RIJK approach with the standard matching auxiliary basis sets was utilized in the HF part. The diagonal line corresponds to the ideal relation.
FIGURE 4 The effect of the localization scheme used for pair natural orbitals on the decomposed DLPNO-CCSD correlation interaction energy terms in terms of local energy decomposition on the S66 set, that is, $A, E_{\text{disp}}^{\text{CCSD}}$ and $B, \Delta E_{\text{non-disp}}^{\text{CCSD}}$. The FB localization scheme was used for occupied orbitals. The aug-cc-pVTZ and its standard matching auxiliary basis sets (within RUJK approach in the HF part) were utilized.

![Graph showing the effect of localization schemes on interaction energy terms.](image)

FIGURE 5 Error associated to different RI approaches in HF/aug-cc-pVTZ interaction energy calculations on the S66 set. RIJK calculations were performed with the A, aug-cc-pV5Z/JK and B, aug-cc-pVTZ/JK auxiliary basis sets. RIJCOSX calculations were performed with the aug-cc-pVTZ/JK auxiliary basis set by using C, the large grid "GridX7" and D, the default grid "GridX". Reference values were obtained from HF/aug-cc-pVTZ calculations carried out without using any approximation for the two-electron integrals.

![Graphs showing error associated to different RI approaches.](image)
\[ E^{\text{SAPT2+3} \delta \text{MP2}} = E_{\text{exch-rep}} + E_{\text{elstat}} + E_{\text{ind}} + E_{\text{disp}} + \delta \text{HF} + \delta \text{MP2} \] (14)

in which

\[ E_{\text{exch-rep}} = E_{\text{exch}}^{(10)} + E_{\text{exch}}^{(11)} + E_{\text{exch}}^{(12)} \] (15)

\[ E_{\text{elstat}} = E_{\text{elstat}}^{(10)} + E_{\text{elstat}}^{(12)} + E_{\text{elstat}}^{(13)} \] (16)

\[ E_{\text{ind}} = E_{\text{ind}}^{(20)} + E_{\text{ind}}^{(20)} + E_{\text{ind}}^{(22)} + E_{\text{exch-ind}}^{(22)} \] (17)
\[ E_{\text{disp}} = E_{\text{disp}}^{20} + E_{\text{disp}}^{21} + E_{\text{disp}}^{22} + E_{\text{disp}}^{30} \]  \hspace{1cm} (18)

\[ \delta_{\text{HF}} = \Delta E_{\text{HF}}^{\text{int}} - \left( E_{\text{elstat}}^{10} + E_{\text{exch}}^{10} + E_{\text{ind,r}}^{10} \right) \]  \hspace{1cm} (19)

\[ \delta_{\text{MP2}} = \Delta E_{\text{HF}}^{\text{MP2}} - \left( E_{\text{elstat}}^{12} + E_{\text{exch}}^{12} + E_{\text{ind,r}}^{12} + E_{\text{disp}}^{20} + E_{\text{exch-disp}}^{20} \right) \]  \hspace{1cm} (20)

**FIGURE 8**  The correlation between the A, "exchange terms", B, "electrostatic terms", and C, dispersion energies in DLPNO-CCSD/LED and SAPT2 + (3)δMP2 calculations on the S66 set. The diagonal line corresponds to the correlation relation between the corresponding local energy decomposition (LED) and SAPT2 + (3)δMP2 terms, shown in each panel. The aug-cc-pVTZ basis set was used in all cases. In LED calculations, the FB localization scheme was used for both the occupied and the virtual orbitals. The RIJK approach was used in the LED decomposition of HF energies in conjunction with the aug-cc-pVTZ/JK auxiliary basis set.
where the notation $E_{\text{elstat}}^{\text{prep}}$ is used to indicate terms of order $\nu$ in the intermolecular potential and of order $w$ with respect to intramolecular correlation, as detailed in Ref. [6]; the superscript $t$ indicates that only a certain part of the corresponding term is actually computed; the subscript $r$ denotes that orbital relaxation is included; the $\text{exch-ind}$ and $\text{exch-disp}$ abbreviations denote the exchange correction to the induction and dispersion energies, respectively; $\Delta E_{\text{rep}}^{\text{MP2}}$ is the BSSE-corrected MP2 correlation interaction energy; and $\delta_{\text{HF}}$ and $\delta_{\text{MP2}}$ are correction terms and are often added to the $E_{\text{ind}}$ term. A more detailed discussion on the individual terms can be found in Refs. [1, 6].

Establishing a one-to-one correspondence between the energy terms of SAPT and LED is not possible. However, both approaches are expected to provide similar trends in their most comparable terms. Here, we establish a series of useful correlations between the energy terms found in LED and SAPT. In particular, the present LED/aug-cc-pVTZ results are compared with the previously published$^{67,68}$ SAPT2 + (3)$\text{iMP2}$/aug-cc-pVTZ energy terms on the S66 set (see Figure 8 and Tables S10 and S11).

As mentioned above, the LED scheme allows us to separate the repulsive electronic preparation energy $\Delta E_{\text{elstat}}$, the attractive Slater exchange energy $E_{\text{exch}}$, and the sum of the LED electrostatics and non-dispersive correlation, $E_{\text{elstat}} + \Delta E_{\text{elstat}}^{\text{CCSD(T)}}$, with the sum of SAPT2 + (3)$\text{iMP2}$ electrostatics, induction, and HF and MP2 correction terms, that is, $E_{\text{elstat}} + \delta_{\text{HF}} + \delta_{\text{MP2}}$ ($R^2 = 0.995$), as shown in Figure 8B. Note that, as described in Ref. [31], LED terms can, in principle, also be decomposed into their permanent and induced contributions. Finally, the SAPT2 + (3)$\text{iMP2}$ estimate of the dispersion energy is about 16% larger than that obtained from LED at the DLPNO-CCSD(T) level, with $R^2 = 0.986$ (Figure 8C).

5 | CONCLUSIONS

On the S66 benchmark set for NCIs, we studied the dependence of the individual LED energy terms on the size of the basis set, the localization scheme used for the virtual PNO space, and the approximations used in the calculation of the reference HF energy. Typically, we found a fast basis set convergence behavior of the individual terms and a weak dependence on the localization scheme used for the PNOs. We also found that, without significant loss in accuracy, the RUCCSD approximation generally increases the efficiency of the LED decomposition in comparison with RIJK. Finally, we have established a series of useful linear relationships between the energy terms found in LED and SAPT2 + (3)$\text{iMP2}$.

ACKNOWLEDGMENT

This work was supported by the Priority Program “Control of Dispersion Interactions in Molecular Chemistry” (SPP 1807) of the Deutsche Forschungsgemeinschaft.

AUTHOR CONTRIBUTIONS

Giovanni Bistoni: Conceptualization; formal analysis; investigation; methodology; supervision; validation; visualization; writing—original draft; writing-review and editing. Ahmet Altun: Data curation; formal analysis; investigation; methodology; validation; visualization; writing—original draft; writing-review and editing. Robert Izsák: Formal analysis; investigation; methodology; validation; visualization; writing—original draft; writing-review and editing.

ORCID

Giovanni Bistoni https://orcid.org/0000-0003-4849-1323

REFERENCES

[1] E. G. Hohenstein, C. D. Sherrill, Rev. Comput. Mol. Sci. 2012, 2, 304.
[2] B. Jeziorski, R. Moszynski, K. Szalewicz, Chem. Rev. 1994, 94, 1887.
[3] R. Podeszwa, R. Bukowski, K. Szalewicz, J. Chem. Theory Comput. 2006, 2, 400.
[4] A. Heßelmann, G. Jansen, M. Schütz, J. Chem. Phys. 2005, 122, 014103.
[5] K. U. Lao, J. M. Herbert, J. Chem. Theory Comput. 2018, 14, 2955.
[6] T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, C. D. Sherrill, J. Chem. Phys. 2014, 140, 094106.
[7] M. von Hopffgarten, G. Frenking, Rev. Comput. Mol. Sci. 2012, 2, 43.
[8] E. Pastorczak, C. Corminboeuf, J. Chem. Phys. 2017, 146, 120901.
[9] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, Chem. Phys. Lett. 1989, 157, 479.
[10] R. J. Bartlett, M. Musial, Rev. Mod. Phys. 2007, 79, 291.
[11] R. Zalesny, M. G. Papadopoulos, P. G. Mezey, J. Leszczynski Eds., Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications, Springer Netherlands, Dordrecht 2011.
[12] Q. Ma, M. Schwilk, C. Köppl, H.-J. Werner, J. Chem. Theory Comput. 2017, 13, 4871.
[13] C. Ripplinger, P. Pinski, U. Becker, E. F. Valeev, F. Neese, J. Chem. Phys. 2016, 144, 024109.
[14] F. Neese, A. Hansen, D. G. Liakos, J. Chem. Phys. 2009, 131, 064103.
