

RESEARCH ARTICLE | APRIL 14 2023

Assessment of random phase approximation and second-order Møller–Plesset perturbation theory for many-body interactions in solid ethane, ethylene, and acetylene

Khanh Ngoc Pham  ; Marcin Modrzejewski  ; Jiří Klimeš  



J. Chem. Phys. 158, 144119 (2023)

<https://doi.org/10.1063/5.0142348>



View
Online



Export
Citation

CrossMark



The Journal of Chemical Physics
Special Topic: Adhesion and Friction

Submit Today!

AIP
Publishing

Assessment of random phase approximation and second-order Møller-Plesset perturbation theory for many-body interactions in solid ethane, ethylene, and acetylene

Cite as: J. Chem. Phys. 158, 144119 (2023); doi: 10.1063/5.0142348

Submitted: 13 January 2023 • Accepted: 27 March 2023 •

Published Online: 14 April 2023



View Online



Export Citation



CrossMark

Khanh Ngoc Pham,¹ Marcin Modrzejewski,² and Jiří Klimeš^{1,a}

AFFILIATIONS

¹ Department of Chemical Physics and Optics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, CZ-12116 Prague 2, Czech Republic

² Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

^a Author to whom correspondence should be addressed: klimes@karlov.mff.cuni.cz

ABSTRACT

The relative energies of different phases or polymorphs of molecular solids can be small, less than a kilojoule/mol. A reliable description of such energy differences requires high-quality treatment of electron correlations, typically beyond that achievable by routinely applicable density functional theory (DFT) approximations. At the same time, high-level wave function theory is currently too computationally expensive. Methods employing an intermediate level of approximations, such as Møller-Plesset (MP) perturbation theory and the random phase approximation (RPA), are potentially useful. However, their development and application for molecular solids has been impeded by the scarcity of necessary benchmark data for these systems. In this work, we employ the coupled-cluster method with singles, doubles, and perturbative triples to obtain a reference-quality many-body expansion of the binding energy of four crystalline hydrocarbons with a varying π -electron character: ethane, ethene, and cubic and orthorhombic forms of acetylene. The binding energy is resolved into explicit dimer, trimer, and tetramer contributions, which facilitates the analysis of errors in the approximate approaches. With the newly generated benchmark data, we test the accuracy of MP2 and non-self-consistent RPA. We find that both of the methods poorly describe the non-additive many-body interactions in closely packed clusters. Using different DFT input states for RPA leads to similar total binding energies, but the many-body components strongly depend on the choice of the exchange-correlation functional.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0142348>

I. INTRODUCTION

Accurate prediction of structural and energetic properties of molecular solids is an important ingredient for their numerous applications, particularly in pharmaceutical and materials science fields.^{1,2} However, binding energies of polymorphs or phases of molecular solids can differ only by a fraction of kJ/mol, and this level of accuracy and precision is currently difficult to reach.^{3–7} For accurate methods, such as quantum Monte Carlo (QMC) or coupled clusters with singles, doubles, and perturbative triples [CCSD(T)], the issues are usually related to the computational time needed for the calculations or their non-trivial setup.^{8–11} More approximate

methods, such as density functional theory (DFT) approximations, are simpler to perform but are less accurate due to, e.g., errors related to self-interaction or the description of correlation.^{12,13} It is often difficult to predict the effect of these errors beforehand, and approximate DFT methods require validation and testing on reference datasets.^{14–16}

Simpler methods based on perturbation theory, such as the random phase approximation (RPA) or second-order Møller-Plesset (MP2) theory, offer a promising way for the calculation of binding energies of molecular solids.^{17–22} While less accurate than CCSD(T),¹⁰ their errors are more consistent or predictable than those of standard DFT approximations.^{22,23} For example, RPA with

the so-called singles corrections was shown to offer better and more consistent results than dispersion-corrected hybrids for a set of binding energies of molecular solids.^{15,22,24} Moreover, the central processing unit (CPU) time required to obtain the RPA energy non-self-consistently with a DFT input from generalized gradient approximation (GGA) functionals is similar to the time required for the hybrid calculations within periodic settings.^{25,26} It is important to identify systems for which the approximations in the simpler schemes lead to reduced accuracy to understand the limits of their applicability. For example, MP2 lacks any correlation above second order, which contributes to overestimated binding for systems with delocalized electrons and larger errors of binding energies of trimers and larger clusters.^{27–34} RPA energy expression accounts for a set of correlation terms up to infinite order so that it can describe systems both with localized and delocalized electrons.^{35–37} However, the accuracy of RPA is not satisfactory in some situations, as shown earlier by tests of atomization energies, interaction energies of molecular dimers, and others.^{38–42} The work to improve the accuracy has proceeded in several directions. For example, it was suggested to include additional terms, such as higher-order exchange terms or various single corrections, or include approximate exchange–correlation kernels.^{21,24,43–52} Moreover, several groups developed methods to perform self-consistent RPA (scRPA) calculations that reduce some of the issues of the non-scRPA approach.^{53–63} However, some of the aforementioned modifications substantially increase the computational demands of RPA or are not yet available within periodic boundary conditions (PBC). For the application to molecular solids, it is, therefore, useful to understand how the accuracy of the simple and affordable scheme, RPA with renormalized singles corrections (RSE),²⁴ is affected by the input DFT states.^{17,23,64}

For molecular solids, the accuracy of different methods is usually tested by comparing the binding energies to reference data.^{12,15,65} This, however, is partially a limitation as the binding energy is only a single number, and it is, thus, difficult to understand in detail the deviations of approximate methods from the reference. A much more detailed understanding can be obtained from the many-body expansion (MBE) of the binding energy.^{4,9} In MBE, the binding energy is divided into a set of dimer interaction energies and non-additive three- and higher-order contributions.^{66–68} This reduces the computational requirements of a single energy evaluation so that the use of the reference methods such as CCSD(T) is feasible.⁶⁹ The accuracy of the simpler scheme, such as RPA or MP2, can be then tested on each of the individual MBE contributions to obtain the errors of the individual *n*-body fragments. This elucidates the origin of the error and the extent of error cancellation and uncovers problems specific to the methods tested.^{4,9,23} Such an analysis is especially useful for the RPA binding energies of molecular solids where RPA with singles corrections was shown to give accurate results.^{10,22}

The convergence of MBE can be slow, especially for systems with important electrostatic contributions.^{7,70} One of the ways to reduce this issue is a subtractive embedding procedure in which a calculation with a simpler scheme is performed within periodic boundary conditions and the binding energy is corrected using MBE involving a more accurate method.^{71–74} For example, periodic Hartree–Fock calculations can be combined with MBE of the correlation energy.^{6,73,75,76} In another example, a fitted empirical force

field is used as the simpler method.^{71,77} In any case, if the subtractive embedding is to be efficient, the simpler scheme should be such that the number of individual MBE contributions that needs to be calculated explicitly is as small as possible. The MP2 and RPA approaches are possible choices for the simpler scheme and we have already shown for a methane clathrate cluster that four-body terms could be well approximated by RPA.²³ Compared to MP2, the benefit of RPA within PBC is its more favorable scaling with the number of *k*-points sampling the reciprocal space, as well as with the number of occupied and virtual states, and also a substantially lower cost of diagonalization if a GGA functional is used.^{26,78,79} The scaling with the system size is also more favorable for RPA [$O(N^4)$] than for MP2 [$O(N^5)$] when localized basis sets are used.

In this work, we obtain MBE reference *n*-body energies for a set of molecular solids and use the data to understand the origin of the low errors of RPA-based methods observed before,^{10,22} and to assess the suitability of MP2 and RPA for the subtractive embedding scheme. We use four molecular crystals of simple hydrocarbons to obtain reference CCSD(T) energies up to the fourth order of MBE. The distance cutoffs that we use for MBE are sufficient to obtain a diverse set of fragments, with molecules both in contact and separated for the two- and three-body contributions. We analyze the basis-set convergence of the different methods and its dependence on the fragment size and the spatial separation of the molecules of the fragment. This also allows us to identify contributions for which large basis-set sizes are not necessary. We use the reference to assess the predictions of MP2 and RPA with and without the RSE corrections. We use the Perdew–Burke–Ernzerhof (PBE)⁸⁰ and the strongly constrained and appropriately normed (SCAN) functionals to provide the input states for RPA as these are readily available within PBC settings with an affordable computational cost.

II. COMPUTATIONAL AND THEORETICAL DETAILS

We selected four crystals for our study: monoclinic ethane⁸² and ethylene⁸³ and cubic and orthorhombic forms of acetylene.⁸⁴ These molecules are small enough to allow for reference CCSD(T) calculations in high-quality basis sets. The importance of electrostatic contributions increases from ethane to acetylene, which we expect to have an effect on the relative importance of the different MBE terms or on their convergence with the number of fragments considered. To differentiate between the two forms of acetylene, we denote the cubic form as acetylene/c and the orthorhombic as acetylene/o.

The initial structures of the crystals were taken from the Cambridge Structural Database (CSD),⁸⁵ and Table S1 shows the CSD codes. The positions of atoms were subsequently optimized using the optB88-vdW functional.^{86–89} We kept the lattice parameters at their experimental values. The geometries of isolated molecules were extracted from the optimized crystal structures and used without further optimization to build the clusters for MBE. All the crystal structures and additional information are provided in the supplementary material and data repository.⁹⁰

The binding energy, E_b , of a molecular solid is

$$E_b = \frac{E_{\text{sol}}}{Z} - E_{\text{mol}}, \quad (1)$$

where E_{sol} and E_{mol} are the energies of the solid per unit cell and isolated molecule, respectively. Z is the number of molecules in the unit cell. There are two main ways to obtain E_b : a direct evaluation using periodic boundary conditions and MBE. We use the MBE approach in this work and discuss it as follows.

The basic idea of MBE is to decompose a calculation of a large (or infinite) system into many smaller subsystem (fragment) calculations. If all the molecules in a crystal are symmetry equivalent, we can select one of them as a reference molecule (ref). The binding energy of the solid, E_b , is then evaluated from interaction energies of dimers $\Delta^2 E$ and non-additive three-, four-, and higher-body energies $\Delta^3 E$, $\Delta^4 E$, ... as follows:

$$E_b = \frac{1}{2} \sum_j \Delta^2 E_{\text{ref},j} + \frac{1}{3} \sum_{j < k} \Delta^3 E_{\text{ref},j,k} + \frac{1}{4} \sum_{j < k < l} \Delta^4 E_{\text{ref},j,k,l} + \dots, \quad (2)$$

where i , j , and k are indices of molecules other than the reference one. The summations run over all the molecules in the crystal, in principle, but in practice, cutoffs are introduced. Here, we use cutoffs based on the distance between the molecules in the fragments. For dimers, we define the distance as the average Cartesian distance of all the pairs of atoms of the two molecules. For trimers and tetramers, the distance is the sum of the distances of all the dimers contained in the cluster. Note that, in general, the structure of the reference gas phase molecule E_{mol} differs from the one in solid, and in that case, monomer deformation energy should be also included in Eq. (2). However, our main aim here is to compare different theoretical methods; we keep the gas phase structure identical to the one in solid and, thus, the monomer term is zero.

The two-body interaction energies $\Delta^2 E_{\text{ref},j}$ are obtained from the dimer energies $E_{\text{ref},j}$ and monomer energies E_{ref} and E_j as

$$\Delta^2 E_{\text{ref},j} = E_{\text{ref},j} - E_{\text{ref}} - E_j. \quad (3)$$

The non-additive three-body contributions $\Delta^3 E_{\text{ref},j,k}$ are evaluated from the trimer energies $E_{\text{ref},j,k}$ using

$$\Delta^3 E_{\text{ref},j,k} = E_{\text{ref},j,k} - \Delta^2 E_{\text{ref},j} - \Delta^2 E_{\text{ref},k} - \Delta^2 E_{j,k} - E_{\text{ref}} - E_j - E_k. \quad (4)$$

Finally, the non-additive tetramer contributions $\Delta^4 E_{\text{ref},j,k,l}$ are obtained as

$$\begin{aligned} \Delta^4 E_{\text{ref},j,k,l} = & E_{\text{ref},j,k,l} - \Delta^3 E_{\text{ref},j,k} - \Delta^3 E_{\text{ref},j,l} - \Delta^3 E_{\text{ref},k,l} - \Delta^3 E_{j,k,l} \\ & - \Delta^2 E_{\text{ref},j} - \Delta^2 E_{\text{ref},k} - \Delta^2 E_{\text{ref},l} - \Delta^2 E_{j,k} - \Delta^2 E_{j,l} \\ & - \Delta^2 E_{k,l} - E_{\text{ref}} - E_j - E_k - E_l. \end{aligned} \quad (5)$$

We obtained MBE contributions up to the four-body term for all the considered systems. Higher-order terms are likely to contribute marginally,^{7,69} and their precise evaluation can be difficult due to numerical errors.³ The structures of the fragments were generated by an in-house library,⁷ and symmetry equivalent clusters were identified using the approach suggested in Ref. 91.

The MBE calculations were performed for CCSD(T), MP2, and RPA. The Molpro program⁹² was used for the MP2 and CCSD(T) calculations. The RPA calculations were performed non-self-consistently, as it is the current practice for molecular solids and other solid-state systems. An in-house code using a canonical-orbital variant of the algorithm described in Ref. 64 was used and the

input states were obtained by the PBE and SCAN functionals. All the correlation energies were obtained within the frozen-core approximation. Care was taken to obtain the many-body contributions with high precision. Specifically, the frequency integration grids are optimized separately for each interacting complex as described in Ref. 64. Moreover, it is known that the SCAN functional requires dense integration grids.^{93–95} Our tests show that this is also the case for RPA based on the SCAN input states, as shown in Table S2. To reduce the numerical errors related to the DFT integration grid, we used a dense molecular grid with 150 radial and 590 spherical points. These settings guarantee a precision of a few percent for the three-body interactions (Tables S2 and S3), which is sufficient for the tests presented here.

In all the calculations, Dunning's augmented correlation-consistent basis sets,⁹⁶ shortened as AVXZ (X = D, T, Q, and 5), were used. The energies needed to evaluate each of the individual n -body contributions, i.e., $\Delta^2 E_{\text{ref},j}$, $\Delta^3 E_{\text{ref},j,k}$, and $\Delta^4 E_{\text{ref},j,k,l}$, were obtained using the basis set of the whole n -body fragment. To reduce the basis-set incompleteness errors, we extrapolate the correlation component of the interaction energies to the complete basis-set (CBS) limit using the formula of Halkier *et al.*,⁹⁷

$$E_{\text{CBS}} = \frac{(X+1)^n E_{X+1} - X^n E_X}{(X+1)^n - X^n}, \quad (6)$$

where E_X is the energy in the AVXZ basis set. We set $n = 3$ for the canonical versions of CCSD(T) and MP2 and for the RPA calculations.⁹⁷

In the case of MP2 and CCSD, we also used the explicitly correlated (F12) versions of the methods to reduce the basis-set dependence of the correlation energies.^{98,99} The correlation energies obtained with the F12 methods have a smaller dependence on the basis-set size and, thus, allow an independent validation of the CBS limit. While they are often close to the CBS limit when the AVQZ basis set is used, we also extrapolated them using Eq. (6) with $n = 5$.^{7,100} The triples (T) contribution was scaled for the two-body term¹⁰¹ and unscaled for the three- and four-body contributions, similar to the approach used in Ref. 23. Finally, the complete auxiliary basis set singles (CABS) corrections^{102,103} to the HF energy were also calculated and included where appropriate.

III. RESULTS

A. Reference CCSD(T) binding energies

In this section, we discuss the setup used for the reference CCSD(T) energies. The cutoffs used to obtain the n -body terms are given in Table 1 together with the number of symmetry inequivalent fragments that are within the cutoff for each of the crystals. The assumed values of cutoff distances allow enough configurations to be sampled to reliably assess the accuracy of MP2 and RPA. While the n -body contributions are not completely converged with the finite cutoffs, increasing the cutoff distances would add to the numerical noise.^{3,7} The total n -body terms also depend on the basis-set size used for the calculations. In the following, we discuss each of the n -body terms separately, focusing first on the dependence on the basis-set size and then on the dependence on the cutoff distance. Note that in the tables and text, the energies are given to three

TABLE I. Cutoff distance (r_{cut} , in Å) and corresponding number (N) of symmetry inequivalent dimers, trimers, and tetramers within the selected cutoff distance for the MBE calculations.

Systems	Two-body		Three-body		Four-body	
	r_{cut}	N	r_{cut}	N	r_{cut}	N
Ethane	19.5	436	25.2	991	34.6	200
Ethylene	18.6	428	26.3	1672	33.1	202
Acetylene/c	24.4	1174	27.0	2875	31.8	282
Acetylene/o	24.8	1094	27.4	2655	32.1	164

decimal digits; this is primarily to be able to show also small changes between energies.

1. Two-body terms

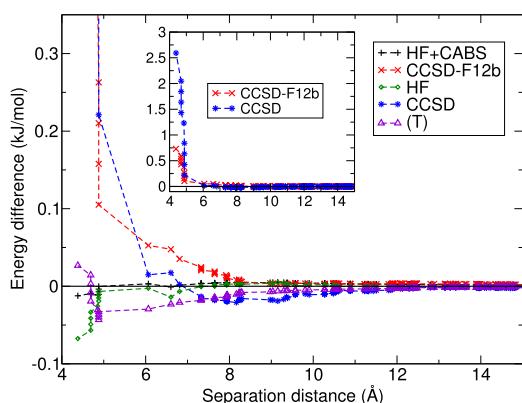
We now discuss the two-body terms, starting with their basis-set convergence. We expect that the convergence will be similar for all the systems. We thus use ethylene to analyze the convergence in detail. We then use the ethylene data to find reliable settings that we use to obtain the CCSD(T) reference energies for all the

systems. Specifically for ethylene, we obtained the two-body energies of CCSD(T) and its components using the AVDZ, AVTZ, and AVQZ basis sets. Moreover, we have also calculated the two-body MP2 energies using AVDZ, AVTZ, AVQZ, and AV5Z basis sets. This allows us to compare the MP2 basis-set convergence behavior to that of CCSD(T) and also test composite schemes that estimate the CCSD(T) basis-set incompleteness error based on the MP2 data.¹⁰⁴

The dimer interaction energies typically depend strongly on the basis-set size. However, one could argue that the importance of using a large basis set might be smaller for dimers where the molecules are far from each other. In general, for large separations, the interactions become smaller and also the perturbing potential of the other molecule becomes more homogeneous. We, therefore, first ask what the distance dependence of the basis-set error of the dimer energies is. To assess this, we consider the two-body energies obtained with AVDZ and AVQZ basis sets, taking AVQZ as the reference values. We then calculate the error that occurs when contributions above some distance, called the separation distance, are obtained with the less precise AVDZ basis set instead of the AVQZ basis. The resulting error is plotted in Fig. 1 for the different contributions to the CCSD(T) energy. One can see that the use of a large basis is, indeed, critical for the nearest neighbors, i.e., molecules within a distance smaller than ~5 Å. Using AVDZ for all the other dimers leads to errors well below 0.1 kJ/mol for each of the energy component.

As the basis-set convergence of the two-body energies depends on the intermolecular distance r , we divided the dimers into two groups: a proximate group ($r < 10$ Å) and a distant group ($10 < r < r_{\text{cut}}$). The separation distance of 10 Å is based on the data shown in Fig. 1. For ethylene, there are 64 dimers in the proximate group and 364 in the distant group. The basis-set convergence of different energy components for the proximate and distant dimers is shown in Tables II and III, respectively, and we discuss them in the following.

The two-body HF energy converges quickly with the basis-set size, and the value for proximate dimers obtained with the AVTZ basis set differs by less than 0.01 kJ/mol from the energy calculated with the AV5Z basis set (Table II). This small error is further reduced to around 0.002 kJ/mol when the CABS corrections are used. The distant dimers, separated by more than 10 Å, contribute by only 0.061 kJ/mol to the two-body energy of ethylene (Table III). The value changes only marginally, by 0.002 kJ/mol, when going from the AVDZ to the AV5Z basis set.

**FIG. 1.** Difference of the two-body CCSD(T) energy components of ethylene between the AVDZ and AVQZ basis sets. The data show the error made in the two-body energy when calculations for dimers above the separation distance are made using the AVDZ basis set, instead of the AVQZ basis set.**TABLE II.** Basis-set convergence of the two-body HF, MP2, CCSD, and (T) contributions for proximate dimers of ethylene in kJ/mol. We also show data obtained with the Δ MP2 procedure. The proximate dimers have an intermolecular distance smaller than 10 Å.

Methods	AVDZ	AVTZ	AVQZ	AV5Z	AVDZ/AVTZ	AVTZ/AVQZ	AVQZ/AV5Z
HF	9.253	9.332	9.323	9.325
HF+CABS	9.299	9.322	9.324	9.324
MP2	-30.445	-33.136	-33.873	-34.126	-34.270	-34.411	-34.390
MP2-F12	-33.570	-34.264	-34.349	-34.383	-34.370	-34.375	-34.400
CCSD	-24.066	-26.142	-26.670	...	-27.017	-27.055	...
CCSD-F12b	-26.232	-26.848	-26.960	...	-26.942	-26.994	...
CCSD+ Δ MP2	-27.890	-27.417	-27.186
CCSD-F12b+ Δ MP2-F12	-27.031	-26.959	-27.011
(T) _{scaled}	-5.281	-5.323	-5.312	...	-5.341	-5.304	...

TABLE III. Basis-set convergence of the two-body term of the HF, MP2, CCSD, and (T) energies for distant dimers of ethylene in kJ/mol. The intermolecular distance is between 10 Å and the total cutoff 18.6 Å for the distant dimers.

Methods	AVDZ	AVTZ	AVQZ	AV5Z	AVDZ/AVTZ	AVTZ/AVQZ	AVQZ/AV5Z
HF	0.063	0.062	0.061	0.061
HF+CABS	0.059	0.061	0.061	0.061
MP2	-0.625	-0.628	-0.628	-0.627	-0.629	-0.628	-0.627
MP2-F12	-0.621	-0.624	-0.628	-0.611	-0.625	-0.629	-0.603
CCSD	-0.478	-0.470	-0.466	...	-0.467	-0.465	...
CCSD-F12b	-0.463	-0.467	-0.465	...	-0.467	-0.466	...
(T) _{scaled}	-0.099	-0.097	-0.095	...	-0.095	-0.094	...

The components of the correlation energy depend more strongly on the basis-set size, as expected. As noted before, the dependence is larger for the proximate dimers than for the distant dimers. For example, for the proximate dimers, the two-body CCSD/AVDZ energy differs by almost 10% from the value obtained with the AVQZ basis. In the case of distant dimers, the difference is only around 2.5%. The difference is also much more significant in absolute numbers. The error is close to 2.6 kJ/mol for the proximate dimers and around 0.01 kJ/mol for the distant dimers. Clearly, small basis sets are sufficient to obtain the interaction energies of the distant dimers. Larger basis sets and extrapolations to the CBS limit are required for the proximate dimers, and we discuss our findings and the resulting setup in the following paragraphs.

As the two-body CCSD energy of the proximate dimers depends strongly on the basis-set size, it is more difficult to obtain the reference data at the CBS limit. There are several ways to obtain values close to the CBS limit and we compared extrapolation, use of the F12 corrections, and the so-called Δ MP2 correction where CCSD is combined with MP2 energies obtained in a larger basis set.¹⁰⁴ Extrapolation of the canonical CCSD energies obtained with AVTZ and AVQZ basis sets leads to a value that is close to the CCSD-F12b/AVQZ data, and the difference is smaller than 0.1 kJ/mol, as shown in Table II. When the CCSD-F12b values are also extrapolated, the difference to extrapolated CCSD decreases to 0.06 kJ/mol. It is not clear which of the two numbers is more precise without going to even larger basis sets. The MP2 data, which we obtained also with the AV5Z basis set, do not help to identify which of the extrapolated values is closer to the CBS. In fact, the change between AVTZ \rightarrow AVQZ and AVQZ \rightarrow AV5Z extrapolated values is similar for MP2 and MP2-F12 so that neither of them can be considered more precise than the other. Keeping these uncertainties in mind, we use the CCSD-F12b/AVTZ \rightarrow AVQZ extrapolated values as the reference.

In the Δ MP2 approach, the two-body CCSD energy obtained with a basis set X , E_X^{CCSD} , is corrected with the basis set incompleteness error of MP2, $E_{X^{\text{CBS}}}^{\text{MP2}} - E_X^{\text{MP2}}$, estimated for the same basis set. Interestingly, for the ethylene dimers, the Δ MP2 scheme is less accurate even when the largest basis sets are used, as shown in Table II. This is caused by the different convergence rates of the MP2 and CCSD energies. Performing the Δ correction with the F12 methods leads to more consistent results. However, this is more likely due to the fact that the energy differences between different basis sets are smaller when the F12 corrections are used.

The last component of the CCSD(T) energy is the triples (T) contribution. Note that for triples, we use the scaling procedure proposed by Knizia et al. to reduce its basis-set size dependence.¹⁰¹ There is only a small basis-set dependence of the (T) energy both for proximate and distant dimers of ethylene. The (T) contribution of the proximate dimers obtained with the AVTZ and AVQZ basis sets differs only by around 0.01 kJ/mol (Table II). The distant dimers contribute by less than 0.1 kJ/mol to the two-body (T) energy and even the small AVDZ basis set is sufficient to obtain the contribution with an error of less than 0.01 kJ/mol, as shown in Table III.

We observe similar basis-set convergence trends also for the other systems. The largest uncertainty comes from the evaluation of the two-body CCSD energy of proximate dimers. Using AVTZ \rightarrow AVQZ extrapolation for canonical CCSD and CCSD-F12b leads to values that are within 0.05 kJ/mol of each other. The uncertainty due to HF is almost an order of magnitude smaller, and the change in the two-body HF+CABS energies is below 0.01 kJ/mol upon going from the AVTZ to AVQZ basis set. All the energy components have a negligible basis-set dependence for the distant dimers.

While analyzing the data of acetylene dimers, we noted that the F12 and CABS corrections introduce numerical errors into the two-body energies when the AVQZ basis set is used. The magnitude of the errors is in the order of few tenths of kJ/mol for F12 and one or two orders less for CABS. Specifically, for the distant dimers of cubic acetylene, we find a two-body CCSD-F12b/AVTZ energy of -0.40 kJ/mol and the same value for CCSD in either the AVTZ or AVQZ basis set. However, the CCSD-F12b/AVQZ contribution is -0.54 kJ/mol. Similar issues were observed before, and they likely stem from finite precision errors and the need to sum contributions of a large number of fragments.^{3,7} The numerical errors in the AVQZ basis are marginal for ethane and ethylene.

Our final reference two-body energies are obtained with the setup that follows. For the proximate dimers, we use CABS-corrected HF in the AVQZ basis set together with AVTZ \rightarrow AVQZ-extrapolated CCSD-F12b and scaled (T) contribution obtained with the AVQZ basis set. For distant dimers, we take the values obtained with the AVTZ basis set, without extrapolation, but with the use of the CABS and F12b corrections. The magnitude of the CABS and F12b corrections is, however, small in the AVTZ basis set, close to 0.001 of kJ/mol for CABS and below 0.005 kJ/mol for F12b. The final two-body energies obtained with the aforementioned setup for the different CCSD(T) energy components are given in Table IV.

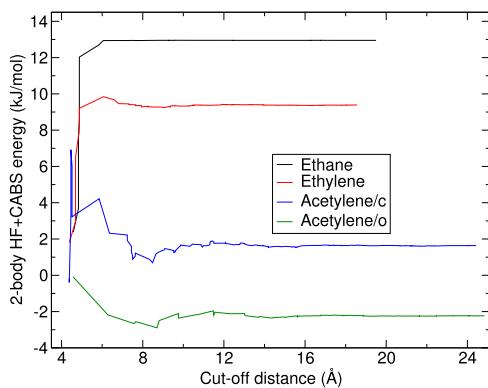
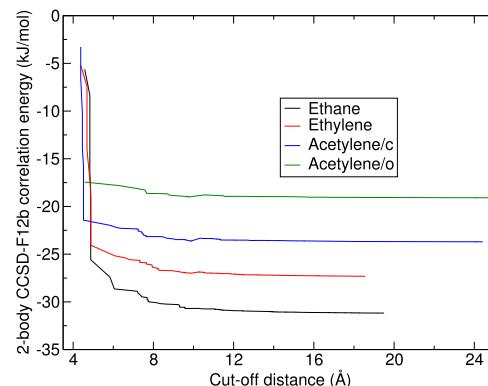
TABLE IV. Contributions of proximate and distant dimers to two-body HF, CCSD, and (T) energies. Data in kJ/mol.

System	Proximate			Distant		
	HF	CCSD	(T)	HF	CCSD	(T)
Ethane	12.95	-30.69	-5.55	0.00	-0.51	-0.10
Ethylene	9.32	-26.99	-5.31	0.06	-0.47	-0.10
Acetylene/c	1.66	-23.64	-5.16	-0.03	-0.40	-0.09
Acetylene/o	-2.36	-19.01	-4.19	0.13	-0.35	-0.08

We now turn to the convergence of the two-body energies with the cutoff distance. The convergence of the HF energies is shown in Fig. 2. The convergence is very fast for ethane, while we observe an oscillatory convergence of the energy for the two forms of acetylene. The oscillatory behavior is caused by the different electrostatic moments of the molecules, especially the quadrupole moment: ethane has a zero moment, while for acetylene, it is around 4 a.u.^{105,106} Similar behavior of the cutoff dependence of the two-body energies was observed for other systems.⁷

The contributions to the two-body HF energies of ethane are dominated by Pauli repulsion, which is short-ranged (decaying exponentially) and repulsive. Pauli repulsion dominates initially for ethylene as well, but attractive electrostatic interactions start to dominate above 6.5 Å and they somewhat reduce the repulsive terms by ≈ 1 kJ/mol. For acetylene, the repulsive and attractive interactions almost cancel each other so that the total two-body HF energy is close to zero, as shown in also Table IV.

The two-body CCSD and (T) correlation energies show the same convergence trend for all the systems, as shown in Fig. 3, Fig. S1, and Table IV. The contributions of molecules at a small cutoff distance dominate and there are minimal or no oscillations for larger cutoffs. This is expected as the correlation interactions decay proportionally to $-r^{-6}$ with the intermolecular distance r . However, the contributions of the distant dimers to the two-body CCSD energies are around -0.4 kJ/mol and, thus, cannot be neglected. Due to the $-r^{-6}$ decay, the convergence of the two-body energy with the cutoff

**FIG. 2.** Cutoff distance convergence of the two-body HF+CABS energy obtained with the AVQZ basis set.**FIG. 3.** Cutoff distance convergence of the two-body CCSD-F12b correlation energy obtained with the AVQZ basis set.

distance r_{cut} is proportional to $-r_{\text{cut}}^{-3}$. This can be used to extrapolate the two-body energy to the infinite cutoff. The extrapolation is not much sensitive to the choice of the fitting interval. For example, extrapolating the energies using data between 10 and 15 Å or between 12 and 18 Å leads to a difference of around 0.03 kJ/mol for the CCSD-F12b/AVTZ energy of ethylene. The extrapolated two-body CCSD correlation energies differ from those obtained for a finite cutoff by 0.08 kJ/mol for ethane and ethylene and around 0.03 kJ/mol for the two forms of acetylene. Therefore, they would be still relevant when energies converged with the cutoff distance were sought.

The (T) contribution for distant dimers is close to -0.1 kJ/mol for all the systems, which is also not negligible. In fact, the (T) terms are around 1/5 to 1/6 of the two-body CCSD correlation energy and this ratio is similar to that obtained for the proximate dimers. Therefore, the relative importance of the (T) contribution does not decay with distance. The possible reason for this is that the (T) terms change the response properties of the molecules. This can be thought of as a change of the effective C_6 coefficients.

Overall, when the two-body HF and CCSD(T) correlation energies are added together, they are rather similar for all the systems, with values between -23.49 kJ/mol for ethylene and -27.66 kJ/mol for acetylene/c, as shown in Table V. One can see that the binding at the HF level increases when going from ethane to acetylene/o, as shown in Table VI, while it decreases for CCSD and (T). The similar binding energies for the different systems are, thus, a result of a compensation between the mean-field and correlation contributions.

TABLE V. The two-, three-, and four-body CCSD(T) binding energies and their sum for all the considered systems. The energies contain the CABS and F12b corrections and are in kJ/mol.

Systems	Two-body	Three-body	Four-body	Total
Ethane	-23.89	1.45	-0.10	-22.54
Ethylene	-23.49	1.56	-0.11	-22.04
Acetylene/c	-27.66	0.97	-0.10	-26.77
Acetylene/o	-25.86	1.15	-0.03	-24.74

TABLE VI. The two-, three-, and four-body contributions to the coupled cluster interaction energies for all the considered systems; data in kJ/mol.

Systems	HF+CABS			CCSD-F12b			(T)		
	Two-body	Three-body	Four-body	Two-body	Three-body	Four-body	Two-body	Three-body	Four-body
Ethane	12.95	-0.62	0.03	-31.20	1.87	-0.12	-5.64	0.19	-0.01
Ethylene	9.38	-0.38	0.02	-27.46	1.74	-0.12	-5.41	0.20	-0.01
Acetylene/c	1.64	-0.90	0.03	-24.03	1.64	-0.11	-5.25	0.23	-0.02
Acetylene/o	-2.24	-0.15	0.05	-19.35	1.13	-0.07	-4.27	0.18	-0.01

2. Three-body terms

We now turn to the three-body terms for which we first analyze the basis set errors taking, again, ethylene as a representative case. The total three-body contributions obtained for ethylene using a cutoff distance of 26.3 Å and different methods and basis sets are given in Table VII and we discuss the main findings as follows.

The three-body HF+CABS contribution shows very little dependence on the basis-set size, changing by 0.004 kJ/mol between the AVDZ and AVQZ basis sets (Table VII). This is consistent with previous results obtained for other systems.^{23,69} Interestingly, the CABS corrections are essentially negligible. Due to the fast convergence with the basis-set size, we use the AVTZ basis set to evaluate the three-body HF energies for the other systems as well. To obtain the reference three-body HF energy, we include the CABS corrections.

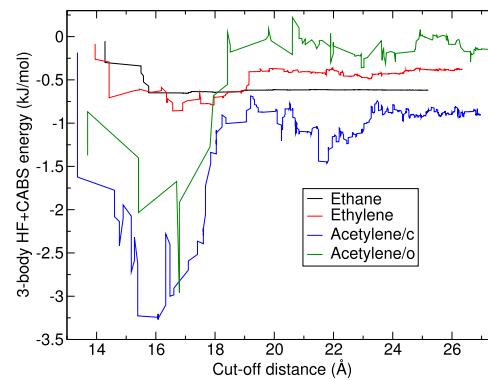
The non-additive correlation energies converge also faster with the basis-set size than the two-body terms. For example, the three-body MP2 energies change only by ~0.1 kJ/mol when going from the AVDZ to the AVQZ basis set. Extrapolating the AVTZ and AVQZ values for MP2 leads to a value that differs by only ~0.01 kJ/mol from the AVQZ data and, thus, extrapolation is hardly necessary. The three-body MP2-F12 energy shows even a smaller dependence on the basis-set size, and the data could be considered converged already in the AVDZ basis set. Note that, however, this does not necessarily guarantee that the convergence of the CCSD or (T) energies would be also fast as MP2 is missing the three-body correlation. In any case, the CCSD and CCSD-F12b energies converge also quickly with the basis-set size, although with a different rate from their MP2 equivalents. Unexpectedly, the F12b-corrected CCSD energy depends more strongly on the basis-set size than the canonical CCSD variant. The changes between AVDZ and AVTZ are, however, only

some hundredths of kJ/mol. The (T) terms, which we evaluate without scaling,^{23,101} change by even a smaller amount, around 0.004 kJ/mol. We, therefore, use the AVTZ basis set to evaluate the three-body CCSD and (T) energies for all the systems and we include the F12b corrections.

We now discuss the convergence of the three-body energies with the cutoff distance for all the systems. The data are shown in Fig. 4 for HF+CABS, Fig. 5 for CCSD-F12b, and in Fig. S2 for the (T) contribution. We note that even though the values are not completely converged with the cutoff, the set contains compact trimers formed by molecules in contact and trimers with molecules separated by almost 10 Å. It, therefore, contains sufficient data for assessing the accuracy of other methods.

The convergence of the three-body HF energies clearly depends on the magnitude of the electrostatic moments, as was the case for the two-body energy. The convergence is very fast for ethane, with terms above $r_{\text{cut}} > 16$ Å contributing by less than 0.05 kJ/mol to the final value of -0.62 kJ/mol. In contrast, one can see a slow convergence for both forms of acetylene. There are large negative and positive terms for distances below 20 Å that lead to changes of several kJ/mol. Beyond that distance, the three-body energies are converged to within a few tenths of kJ/mol for all the systems including acetylene. Interestingly, despite the very different convergence with the cutoff, the final three-body HF energy is between 0 and -1 kJ/mol for all the systems.

The three-body CCSD correlation energies of all the systems are dominated by contributions of trimers with distances below ~20 Å,

**FIG. 4.** Cutoff distance convergence of the three-body HF+CABS/AVTZ energies for all the considered systems.**TABLE VII.** Basis-set convergence of the three-body term in the MP2 and CCSD(T) calculations for ethylene; data in kJ/mol.

Methods	AVDZ	AVTZ	AVQZ	AVTZ/AVQZ
HF	-0.375	-0.378	-0.379	...
HF+CABS	-0.375	-0.378	-0.379	...
MP2	0.708	0.791	0.807	0.819
MP2-F12	0.815	0.828	0.815	...
CCSD	1.729	1.742
CCSD-F12b	1.707	1.738
(T) _{unscaled}	0.193	0.197

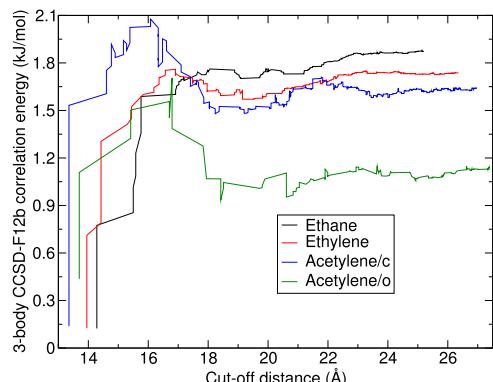


FIG. 5. Cutoff distance convergence of the three-body CCSD-F12b/AVTZ correlation energies for all the considered systems.

and the values then change by only ~ 0.2 kJ/mol between 20 Å and the cutoff used. The convergence is again affected by the magnitude of the electrostatic moments: it is almost monotonic for ethane, while there are significant positive and negative contributions for acetylene. The final three-body CCSD energies (within the cutoffs) are between 1 and 2 kJ/mol. One can also note that their magnitude decreases when going from ethane to acetylene/o; thus, the ordering is the same as for the two-body energies (Table VI). The three-body (T) energies are rather small, around 0.2 kJ/mol for all the systems, and show a similar convergence to the three-body CCSD energies (Fig. S2).

Overall, we find that the total three-body contributions are repulsive for all the systems, with values close to 1 kJ/mol, as shown in Table V. As with the two-body terms, the similar final values are due to a partial cancellation between the HF and correlation contributions.

3. Four-body terms

We now turn to the four-body terms starting with their basis-set convergence. As for the two- and three-body energies, we assessed the convergence in more detail for ethylene. Previous works have shown that the basis-set convergence of the four-body terms is fast.^{23,69} Indeed, we observe that for ethylene, the HF and MP2 values change by at most a few thousandths of kJ/mol when going from the AVDZ to the AVTZ basis set (Table VIII). We were only able to perform the CCSD(T) calculations in the AVDZ basis set. Based on the MP2 data and the convergence of CCSD(T) for the three-body energy, we expect that the basis-set error for AVDZ is negligible, below 0.01 kJ/mol. This is also supported by the small effect of the F12b corrections. We use the HF+CABS/AVTZ and CCSD(T)-F12b/AVDZ values as the reference data.

We compare the total four-body CCSD(T) energies for all the systems in Table V and their components in Table VI. Clearly, all the components of the four-body CCSD(T) energy have, within the cutoffs used, a small magnitude. The HF and correlation contributions have opposite signs and partially cancel each other for all the systems. The final four-body CCSD(T) energy is close to 0.1 kJ/mol for all the systems except for acetylene/o where it is -0.03 kJ/mol.

The convergence of the four-body energies with distance is shown in Figs. 6 and 7 for HF+CABS and CCSD-F12b, respectively.

TABLE VIII. Basis-set convergence of the four-body term in the MP2 and CCSD(T) calculations for ethylene; data in kJ/mol.

Methods	AVDZ	AVTZ
HF	0.020	0.021
HF+CABS	0.022	0.022
MP2	-0.063	-0.068
MP2-F12	-0.069	-0.070
CCSD	-0.128	...
CCSD-F12b	-0.122	...
(T) _{unscaled}	-0.010	...

As with the two- and three-body energies, the distance convergence of the four-body HF+CABS energy varies depending on the electrostatic moments of the molecule: for ethane, the values stay between 0 and 0.05 kJ/mol, while for acetylene/c the energy first reaches a value of almost 0.5 kJ/mol at a cutoff of around 30 Å, and then it returns to zero. In contrast, the four-body correlation energies show a similar convergence pattern for all the systems.

Note that the four-body energies are not completely converged with the cutoff distance. We have tested that extending the cutoff by around 2 Å changes the total four-body CCSD(T) energies by -0.03 to -0.08 kJ/mol, but we do not include these data due to possible numerical issues. We note that regardless of the cutoff convergence issues, the CCSD(T) results form a valid reference dataset for assessing the performance of approximate methods.

4. Summary

The n -body CCSD(T) contributions and their sum are given in Table V. We note that due to the use of finite cutoffs and basis sets, the n -body energies are not completely converged. We estimate that the deviations from the converged values are only few tenths of kJ/mol. A part of the difference comes from the basis-set incompleteness error. This is most likely significant only for the two-body energies where the difference between extrapolated CCSD and CCSD-F12b was around 0.05 kJ/mol, and the basis-set errors are negligible (close to 0.01 kJ/mol) for the non-additive terms.

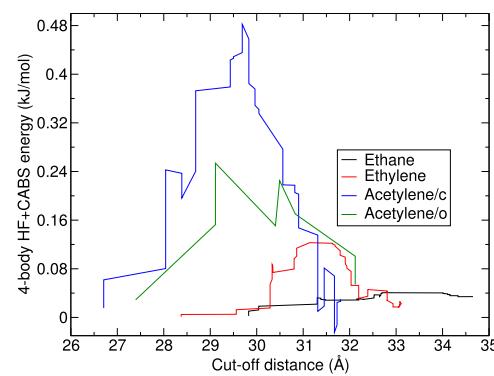


FIG. 6. Cutoff distance convergence of the four-body HF+CABS/AVTZ energy for all the considered systems; note the small scale on the y axis.

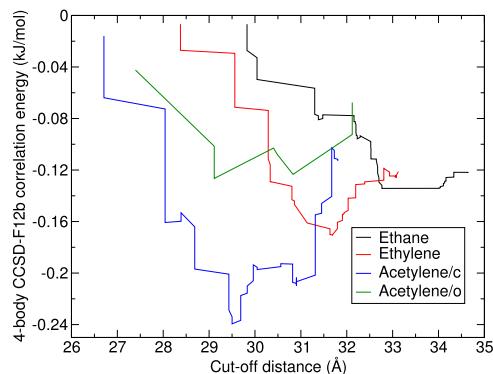


FIG. 7. Cutoff distance convergence of the four-body CCSD-F12b/AVDZ correlation energy for all the considered systems. Note that the y axis scale is approximately one-half compared to the HF case in Fig. 6.

The error due to the finite distance cutoff can be almost avoided by extrapolation for the two-body energies. This leads to an uncertainty of the two-body terms well below 0.1 kJ/mol. The convergence with the cutoff distance is more problematic for the three- and four-body energies, especially for systems with strong electrostatic interactions. However, their convergence with the distance cutoff suggests that they are converged to some tenths of kJ/mol.

B. Accuracy of RPA and MP2

We now use the CCSD(T) energies as a reference to examine the accuracy of the RPA and MP2 methods for predicting the total n -body energies and the contributions of the individual fragments. We reiterate that our aim here is to understand the good accuracy of RPA with singles corrections observed for molecular solids within periodic boundary conditions,²² gain more insight into the difference between PBE and SCAN input states for non-additive energies,⁶⁴ and test the suitability of RPA for the subtractive embedding scheme. Before that, we briefly comment on the basis-set convergence of RPA and MP2 and the numerical setup used to perform the calculations.

The basis-set convergence of MP2 is similar to that of CCSD(T), as discussed previously. We, therefore, use an identical setup to that used for CCSD(T) for the two- and three-body terms. For the four-body MP2 terms, we use the AVTZ basis set instead of the AVDZ that we used for CCSD(T). However, the results obtained with AVDZ and AVTZ basis sets differ only marginally (by around 0.001 kJ/mol for ethylene), as shown in Table VIII.

The RPA calculations used the same setup for all the n -body terms. The EXX and RSE energies were obtained with the AVQZ basis set. All the RPA correlation energies were evaluated by AVTZ → AVQZ extrapolation. The extrapolation is also performed for the three- and four-body energies as they show a stronger dependence on the basis-set size than the HF-based methods, as discussed below and shown in Tables S5–S10.

1. Two-body terms

We first discuss the mean-field (or single-determinant) contributions, i.e., HF, EXX, and RSE. The data are shown in Tables IX and X for the proximate and distant dimers, respectively. One can

TABLE IX. The two-body mean-field contributions of proximate dimers (with an intermolecular distance below 10 Å); data in kJ/mol.

Systems	PBE			SCAN	
	HF	EXX	EXX+RSE	EXX	EXX+RSE
Ethane	12.95	19.45	14.99	18.02	14.89
Ethylene	9.32	14.17	10.95	12.01	10.54
Acetylene/c	1.66	6.60	3.49	3.84	2.89
Acetylene/o	-2.36	1.47	-0.90	-0.82	-1.40

see that EXX gives more repulsive two-body energies than HF for both PBE and SCAN input states. This is consistent with previous observations.^{21,24} The repulsion is, however, much smaller for EXX based on SCAN, which agrees with previous calculations for molecular clusters.^{23,64} When the RSE corrections are added to EXX, the difference to HF is reduced to ~1–2 kJ/mol for both inputs with the SCAN-based values still closer to HF than the PBE-based data.

The mean-field contributions of the distant dimers are small, i.e., below 0.2 kJ/mol (Table X). Interestingly, there is a close agreement between the HF values and EXX(SCAN) values for all the systems, and the RSE corrections for EXX(SCAN) are negligible. The EXX(PBE) values differ from EXX(SCAN) and HF for both forms of acetylene but the differences are reduced upon the addition of the RSE correction.

We now add the two-body correlation energies to the mean-field data for the proximate dimers to compare the methods. Note that this is necessary due to the different two-body mean-field energies for the post-HF methods and RPA for the proximate dimers. For MP2 we find the expected behavior: the difference to CCSD(T) increases significantly when going from the aliphatic ethane to the molecules with delocalized π -electron systems (Table XI).^{27,107} The errors are ~1 kJ/mol for ethane and as large as 5 kJ/mol for acetylene/c.

The RPA correlation energies based on the PBE and SCAN states differ by around 1 kJ/mol from each other with RPA(PBE) giving a stronger binding (Table S4). The differences in the correlation energies then partly cancel the differences in the mean-field EXX energies (Table IX) so that EXX+RPA based on SCAN binds around 0.2–2 kJ/mol more strongly than EXX+RPA with the PBE input. The situation is reversed when the RSE corrections are added, as shown in Table XI. Overall, the total RPA two-body energies with the RSE corrections underestimate the CCSD(T) reference by around 1–3 kJ/mol, with PBE states giving smaller errors than the SCAN input.

TABLE X. The two-body mean-field contributions of dimers with an intermolecular distance larger than 10 Å (distant dimers); data in kJ/mol.

Systems	PBE			SCAN	
	HF	EXX	EXX+RSE	EXX	EXX+RSE
Ethane	0.00	0.00	0.00	0.00	0.00
Ethylene	0.06	0.06	0.06	0.06	0.06
Acetylene/c	-0.03	-0.07	-0.02	-0.03	-0.03
Acetylene/o	0.13	0.15	0.12	0.13	0.13

TABLE XI. The two-body total energy contributions of proximate dimers in kJ/mol.

Systems	CCSD	CCSD(T)	MP2	RPA(PBE)	RPA(SCAN)	RPA(PBE)+RSE(PBE)	RPA(SCAN)+RSE(SCAN)
Ethane	-17.74	-23.29	-22.19	-17.58	-17.78	-22.04	-20.91
Ethylene	-17.67	-22.99	-25.05	-17.48	-18.44	-20.70	-19.91
Acetylene/c	-21.98	-27.13	-31.89	-21.03	-23.01	-24.14	-23.96
Acetylene/o	-21.37	-25.56	-29.27	-20.54	-22.11	-22.91	-22.69

For distant dimers, the EXX+RSE are almost identical regardless of the input states and agree with HF. The differences in binding are then entirely due to the correlation part. For all systems considered, using SCAN states as the input for RPA leads to a close agreement with CCSD for the long-range interactions, and the differences are within 0.01 kJ/mol, as shown in **Table XII**. The differences are somewhat larger for RPA(PBE). The close agreement of CCSD and RPA can be also seen in **Fig. 8(a)** for ethane and **Fig. 8(b)** for acetylene/c, which shows the cutoff dependence of the binding energy. Note that the graph shows the sum of contributions of dimers with distance larger than the cutoff on the x axis. As expected, the two-body MP2 correlation energies of distant dimers do not show consistent behavior. MP2 is very close to CCSD(T) for ethane [**Fig. 8(a)**], but the binding is overestimated for acetylene/c [**Fig. 8(b)**].

The total two-body energies of MP2 and the two RPA variants are compared to the CCSD(T) values for all the systems in **Fig. 9**. One

can see that there is little difference between RPA with RSE based on SCAN and PBE. Using SCAN states produces stronger binding from EXX and RSE but a weaker correlation than PBE-based RPA. RPA and MP2 predict similar two-body energies for ethane. However, when going to ethylene and acetylene, the MP2 binding becomes too strong, while either of the RPA variants produces underestimated binding.

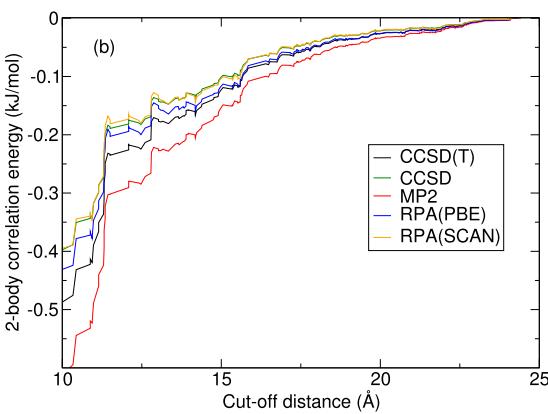
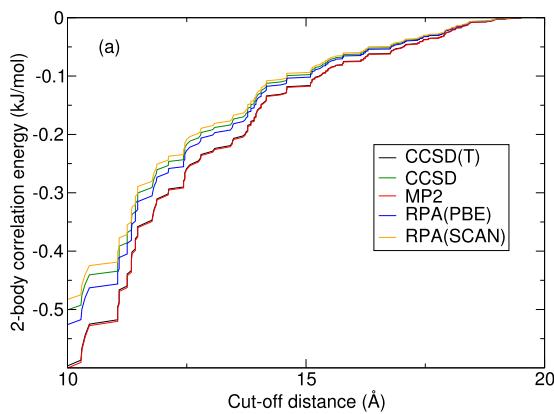
2. Three-body terms

We discuss first the mean-field energies, i.e., HF and EXX with and without RSE, as shown in **Table XIII**. The PBE- and SCAN-based EXX energies show a different behavior with respect to the HF values, the first gives a too strong binding, while the latter is more repulsive. As with the two-body terms, the EXX energies based on SCAN states are closer to the HF data than when PBE states are used. Consequently, the RSE corrections are also larger for PBE states than for the SCAN ones. Nevertheless, the three-body EXX+RSE energies based on PBE differ by around 0.8 kJ/mol from the HF values, and the difference is below 0.25 kJ/mol for EXX+RSE based on SCAN. This again hints at smaller many-body errors of SCAN.¹⁰⁸

We analyze the distance dependence of the PBE- and SCAN-based EXX and RSE components in **Fig. 10**. Clearly, all of the methods (i.e., including HF) tend to predict similar energies for larger distances but differ at small separations. The three-body HF energies at larger distances are especially well reproduced by EXX+RSE based on the SCAN states. Specifically, the difference between three-body EXX+RSE energy and three-body HF energy is converged to within

TABLE XII. The two-body correlation energy contributions for distant dimers in kJ/mol. F12 corrections were used for CCSD and MP2.

Systems	CCSD	CCSD(T)	MP2	RPA(PBE)	RPA(SCAN)
Ethane	-0.51	-0.61	-0.61	-0.54	-0.50
Ethylene	-0.47	-0.56	-0.62	-0.50	-0.46
Acetylene/c	-0.40	-0.49	-0.60	-0.46	-0.40
Acetylene/o	-0.35	-0.42	-0.55	-0.37	-0.36

**FIG. 8.** The two-body correlation energy calculated by different methods for (a) ethane and (b) acetylene/c. The energy, in kJ/mol, is a sum of two-body contributions of dimers with a distance larger than the cutoff given on the x axis and smaller than the largest cutoff used. The calculations used the AVTZ basis set.

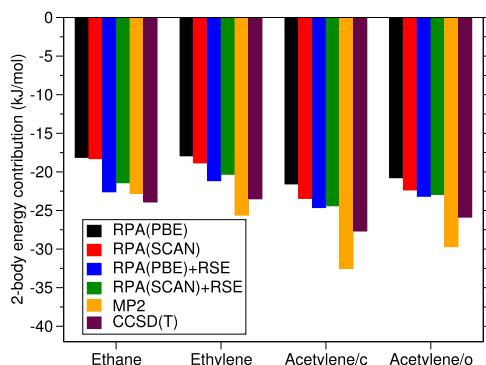


FIG. 9. The two-body contributions to the total RPA and MP2 binding energies compared to the CCSD(T) reference data.

0.01 kJ/mol at a cutoff distance of around 17 Å. When PBE states are used, the difference between EXX+RSE and HF converges more slowly, and the difference still changes by around 0.1 kJ/mol above 17 Å. We also observe the same trends for the other systems. Therefore, if one considers the three-body HF energies as a mean-field reference, then the three-body EXX and EXX+RSE energies based

on SCAN are superior to the PBE-based three-body energies. This is again consistent with the behavior observed for methane clathrate.²³

We now turn to the three-body correlation energies given in Table XIV, and the total three-body energies, shown in Fig. 11. The three-body MP2 energies are close to 50% of the CCSD or CCSD(T) values for all the systems, less than that for ethane and more than that for both forms of acetylene. This is likely due to the missing three-body correlations in MP2.

The three-body mean-field (EXX and EXX+RSE) energies gave stronger and weaker binding, respectively, than the three-body HF energies for all the systems. Therefore, if the total RPA

TABLE XIV. The three-body correlation energies in kJ/mol. The MP2 and CCSD energies include F12 corrections and are in the AVTZ basis set, and the (T) contribution is "unscaled" and also in the AVTZ basis set. The RPA values were obtained by AVTZ → AVQZ extrapolation.

Systems	MP2	CCSD	(T)	RPA(PBE)	RPA(SCAN)
Ethane	0.68	1.87	0.19	1.89	0.93
Ethylene	0.83	1.74	0.20	1.07	0.47
Acetylene/c	0.98	1.64	0.23	0.46	0.43
Acetylene/o	0.63	1.13	0.18	0.23	0.14

TABLE XIII. The three-body mean-field energies in kJ/mol. HF values do not include CABS corrections and were obtained with the AVTZ basis set, while for EXX and RSE, the values are based on the PBE and SCAN states and obtained in the AVQZ basis set.

Systems	HF	PBE-based			SCAN-based		
		EXX	RSE	EXX+RSE	EXX	RSE	EXX+RSE
Ethane	-0.62	-2.95	1.46	-1.49	-0.10	-0.36	-0.46
Ethylene	-0.38	-2.16	1.02	-1.14	0.46	-0.60	-0.14
Acetylene/c	-0.91	-2.48	0.80	-1.67	-0.15	-0.67	-0.82
Acetylene/o	-0.16	-1.36	0.60	-0.76	0.54	-0.58	-0.04

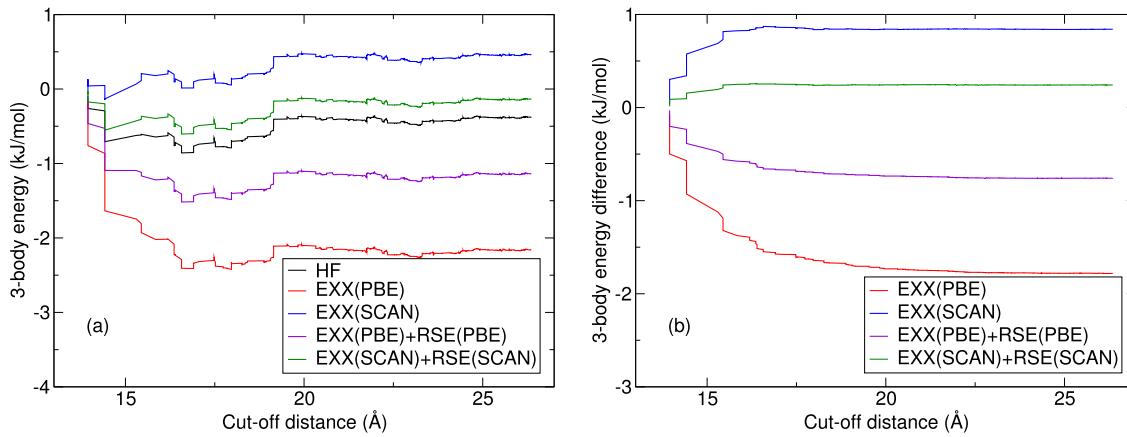


FIG. 10. (a) The three-body energies of ethylene obtained by HF and EXX without and with RSE corrections using both PBE and SCAN orbitals. (b) The differences of the EXX and EXX+RSE data with respect to HF.

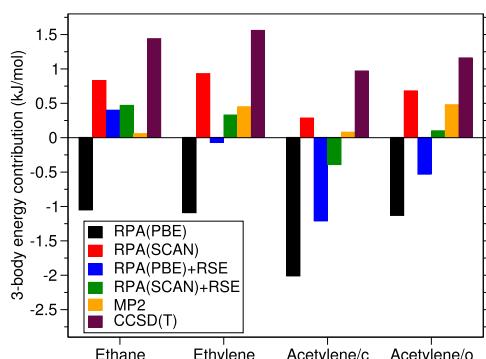


FIG. 11. The three-body contributions to the total RPA and MP2 binding energies compared to the CCSD(T) reference data.

energy was to recover the total CCSD or even the three-body CCSD(T) energy, the RPA(SCAN) correlation energies would need to be close to the CCSD correlation energies and the RPA(PBE) values even larger in magnitude (more repulsive). However, we observe neither. The RPA(PBE) correlation energy is only similar to the CCSD correlation for ethane. In all the other cases, both RPA(PBE) and RPA(SCAN) correlation energies are much smaller than the CCSD correlation energies. The total three-body energies are, therefore, underestimated for either of the RPA methods, as shown in Fig. 11.

We again plot the convergence of the three-body correlation energies with the cutoff distance to understand possible origins of the differences. The convergence shows similar trends for all the systems, and we, therefore, show only the convergence for ethylene in Fig. 12(a). For small distances, below ≈ 20 Å, the MP2 energies are very close to one-half of the CCSD energies. However, for larger cut-offs the MP2 data show much smaller variations than CCSD. This is most likely due to missing three-body correlations and the resulting error could likely be reduced by including a three-body correlation correction.^{9,109}

The convergence of the RPA correlation energy with the cut-off distance looks similar to that of CCSD(T) at first sight; however, there are two possible issues with the behavior of RPA. First, the three-body energies show a larger dependence on the basis-set size, and second, the contributions of the three-body fragments show different asymptotic behavior from the CCSD(T) data. The first issue is shown in Fig. 12(a) where one can see that the difference between the three-body energies obtained from AVTZ and CBS are close to 0.5 kJ/mol for both PBE- and SCAN-based RPA. We note that the difference to the CBS limit is still around 0.2 kJ/mol for the AVQZ basis set. This is likely a consequence of using input states based on DFT; we observe that PBE and SCAN also have a larger basis-set dependence than HF.

The second issue is shown in Fig. 12(b), which shows the differences between the distance-dependent three-body correlation energies obtained for the various methods and CCSD(T). Clearly, while the difference converges for CCSD and even MP2 within some tenths of kJ/mol above ~ 20 Å, the differences obtained for RPA show a much slower convergence. The slower convergence is not caused by some basis-set errors; the basis-set size is mostly relevant for trimers with small distances. We have also checked that the different convergence is not caused by numerical issues by comparing the values obtained by Molpro and the in-house code. Finally, the trend is also present when the total energies, and not only the correlation energies, are compared, so it is not a consequence of a different mean-field reference for RPA and CCSD(T).

Because of the potential use of RPA and MP2 in subtractive embedding schemes, it is important to understand the performance of those methods as a function of the separation of molecules in a cluster. To this end, we divided the ethylene trimers into four groups according to the number of contacts in the fragment. Two molecules are in contact when their intermolecular distance is below 6 Å in the case of ethylene. One can see in Table XV that as the number of contact decreases, the three-body contributions tend to decrease in magnitude. In the case of CCSD(T), the contributions are reduced by a factor of two when one contact is lost between the molecules. Interestingly, the (T) terms are important only for the compact trimers; their effect is minor already for the group

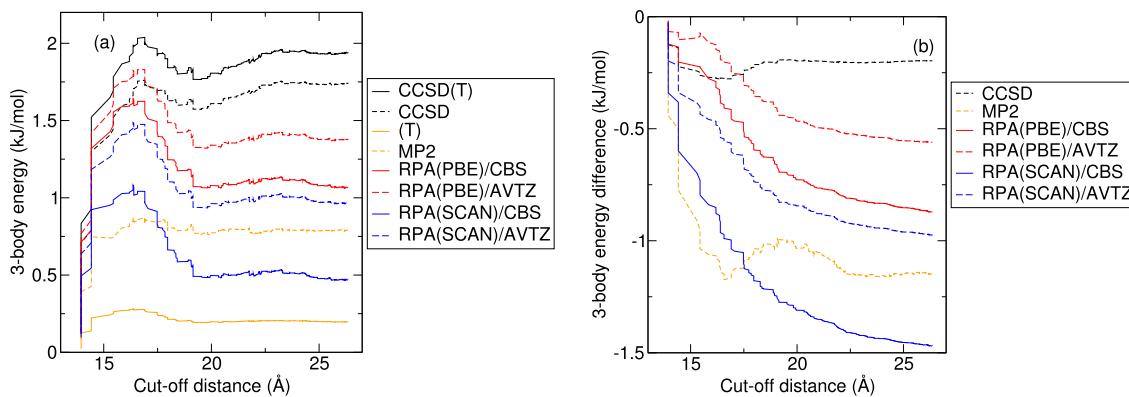


FIG. 12. (a) The distance cutoff convergence of the three-body correlation energies of ethylene obtained for different methods. Panel (b) shows the difference with respect to the CCSD(T) curve, i.e., $E(r_{\text{cut}}) - E^{\text{CCSD(T)}}(r_{\text{cut}})$.

TABLE XV. The three-body energies of ethylene obtained with CCSD(T), CCSD, MP2, HF, and two variants of RPA. The trimers are divided into groups according to the number of contacts between the molecules in the trimer. Data are in kJ/mol.

No. of contacts	CCSD(T)	CCSD	MP2	HF	RPA(PBE)	RPA(SCAN)
3	0.82	0.60	0.08	-0.71	0.23	0.37
2	0.42	0.46	0.27	0.26	-0.32	-0.10
1	0.22	0.22	0.07	0.03	-0.04	-0.01
0	0.09	0.09	0.04	0.04	0.06	0.07

TABLE XVI. The four-body correlation energies in kJ/mol. The AVTZ and AVDZ basis sets were used for the MP2 and CCSD(T) calculations, respectively. The RPA energies are extrapolated based on AVTZ and AVQZ data.

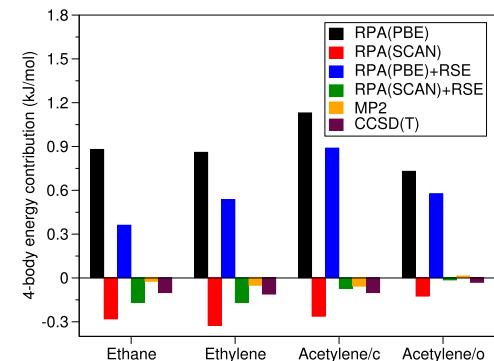
Systems	MP2	CCSD	(T)	RPA(PBE)	RPA(SCAN)
Ethane	-0.05	-0.12	-0.01	-0.01	-0.05
Ethylene	-0.07	-0.12	-0.01	0.24	-0.02
Acetylene/c	-0.09	-0.11	-0.02	0.54	0.00
Acetylene/o	-0.04	-0.07	-0.01	0.32	0.01

with two contacts. MP2 correlation energy is only around 50% of the CCSD(T) correlation energy for the group with three contacts and below 0.05 kJ/mol in the other groups. Consequently, it underestimates the three-body contributions for all the groups. Part of the error can be attributed to the missing three-body correlations in MP2.

The RPA energies show considerable differences from the reference for the groups with two and one contact, and the errors are small only for the group with zero contacts. Part of the error likely stems from the many-body errors of the DFT input states, e.g., the three-body energies for the group with two contacts are 2.33 and -1.22 kJ/mol for PBE and SCAN, respectively.²³ It is possible that these errors could be alleviated by going beyond our setup based on non-self-consistent RPA with (meta-)GGA input states, such as including exchange-correlation kernels, performing self-consistency, or utilizing HF input states. However, these approaches are currently either not available or more computationally demanding within periodic boundary conditions and, thus, less efficient for the subtractive embedding. In fact, the compact trimer groups with three and two contacts are finite and the erroneous RPA contributions can be replaced by CCSD(T) within the subtractive embedding approach.

3. Four-body terms

Finally, we compare the four-body MP2 and RPA energies with the CCSD(T) reference. The four-body CCSD(T) contributions obtained with our set of tetramers are almost negligible, close to -0.1 kJ/mol for all the systems (Table V). Considering first MP2, we find that the MP2 correlation energies are close to one-half of the CCSD(T) correlation for all the systems, as shown in Table XVI. Note that with our tetramer cutoffs, all the clusters can be still considered compact, without fully isolated molecules. Therefore, when a tetramer forms, the single-particle HF states change due to the overlap of the monomers. This changes the MP2 energy compared to the

**FIG. 13.** The four-body contributions to the total RPA and MP2 binding energies compared to the CCSD(T) reference data.

isolated monomers so that the four-body MP2 energy is non-zero despite there being no four-body correlation terms in MP2.

Moving to RPA, we see a stark difference between the PBE- and SCAN-based RPA energies, both in their mean-field and correlation components. The total four-body energies are small (below 0.2 kJ/mol) and close to the reference values for RPA based on SCAN, when RSE are included, as shown in Fig. 13. In contrast, RPA based on the PBE states gives too positive values with errors up to 1.0 kJ/mol for acetylene/c. These differences come from both the mean-field and correlated contributions. The four-body EXX(PBE) component is more repulsive than HF, while EXX(SCAN) is more attractive, as shown in Table XVII. The magnitude of the four-body contributions is reduced upon the addition of RSE. The four-body correlation energies are small in magnitude for RPA based on SCAN; in fact, they are even smaller than the MP2 values (Table XVI). In contrast, the RPA correlation energy based on PBE is close to zero only for ethane, while the values are few tenths of kJ/mol for the other systems.

TABLE XVII. The four-body mean-field energies in kJ/mol. HF values do not include CABS corrections and were obtained with the AVDZ basis set; for EXX and RSE, the values are based on the PBE and SCAN states and obtained in the AVQZ basis set.

Systems	HF	PBE-based			SCAN-based		
		EXX	RSE	EXX+RSE	EXX	RSE	EXX+RSE
Ethane	0.04	0.89	-0.52	0.37	-0.23	0.11	-0.12
Ethylene	0.02	0.62	-0.33	0.30	-0.31	0.16	-0.15
Acetylene/c	0.03	0.60	-0.24	0.35	-0.26	0.19	-0.07
Acetylene/o	0.04	0.41	-0.16	0.26	-0.13	0.11	-0.02

TABLE XVIII. The total MP2 and RPA energies in kJ/mol compared to the CCSD(T) reference.

Systems	CCSD(T)	PBE-based			SCAN-based	
		MP2	RPA	RPA+RSE	RPA	RPA+RSE
Ethane	-22.54	-22.75	-18.29	-21.82	-17.73	-21.11
Ethylene	-22.04	-25.21	-18.15	-20.67	-18.23	-20.14
Acetylene/c	-26.77	-32.51	-22.44	-24.94	-23.42	-24.85
Acetylene/o	-24.74	-29.22	-21.16	-23.12	-21.78	-22.83

4. Summary

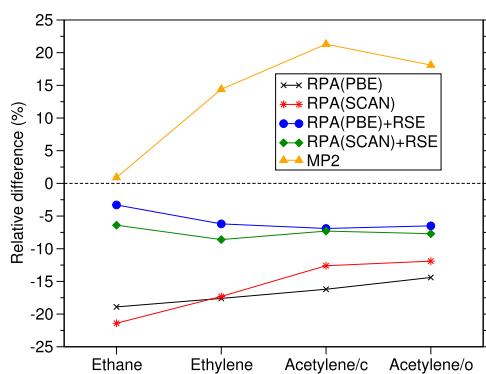
Finally, we summarize this section by comparing the total RPA and MP2 binding energies to the CCSD(T) reference. The binding energies are given in Table XVIII, and their relative deviations from the reference are shown in Fig. 14.

We start with MP2 for which we find that it predicts well the binding energy of ethane but overestimates the binding for the other systems by at least 14%. For ethane, the small error is a consequence of an error cancellation between a small error in two-body interactions (≈ 1 kJ/mol) and an error in the three-body interactions (≈ -1.4 kJ/mol), which is partly caused by the missing three-body correlations. For the other systems, the three-body errors are similar to those obtained for ethane, but the errors in the two-body

interactions are several kJ/mol. The negative error in both two- and three-body terms leads to a substantially overestimated total binding energy. The two-body error comes from an inaccurate description of systems with delocalized or π bonds within the second-order perturbation theory, and similar issues can be expected for related systems. The errors in four-body energies are below 0.1 kJ/mol for all the systems, and thus, they play only a minor role in the final deviation.

The binding energies obtained for SCAN- and PBE-based RPA are rather similar, while the relative deviations differ by only a few percent and exhibit essentially the same trends (Fig. 14). As expected, RPA without singles corrections gives binding energies that underestimate the reference data. The range of errors is $\sim -20\%$ to -15% for RPA(PBE), and larger, around -22% to -12% , for RPA(SCAN). For either method, the largest error occurs for ethane and the smallest for acetylene/o. When the singles are included, we find that RPA(PBE)+RSE performs somewhat better than RPA(SCAN)+RSE for all the considered systems. Specifically, the average difference to the reference data is 5.7% for the first method and 7.5% for the latter. These errors are consistent with those observed for RPA(PBE)+RSE for molecular solids bound dominantly by dispersion in Ref. 22.

While the total binding energies are similar for PBE- and SCAN-based RPA, they show significant differences in their many-body components, as discussed previously. For RPA based on PBE, the three-body error is negative and the four-body error is positive; therefore, there is a partial error cancellation between the three- and four-body errors. As the three-body errors are between -1 and -2 kJ/mol and the four-body errors around $1/2$ of that and positive, the sum of the three- and four-body errors is around -0.5 to -1 kJ/mol, i.e., too strong binding. The overestimated three- and four-body contributions then partly cancel the underestimated two-body terms leading to the observed underbinding.

**FIG. 14.** Relative difference of the binding energies with the RPA and MP2 methods with respect to the reference data.

The effect of error cancellation between n -body energies is smaller for the SCAN-based RPA. First of all, the mean-field and correlation contributions have opposite differences from the HF and CCSD(T) correlation for both the three- and four-body terms. Thus, these deviations partly cancel for SCAN-based RPA and do not add up as with RPA based on the PBE states. The total four-body energies are then close to the reference, while the positive three-body terms are again underestimated. As with the PBE-based RPA, this negative error in the three-body terms compensates part of the error of the two-body terms resulting in underestimated binding energies.

IV. CONCLUSIONS

In the present work, we obtained MBE contributions at the CCSD(T) level for four molecular solids and used them to assess the accuracy of MP2 and RPA. The CCSD(T) energies were obtained up to the fourth order of MBE with a finite distance cutoff, and we thoroughly tested their convergence with the basis-set size. In doing so, we identified strategies that can be used to save computational time without significantly sacrificing the precision of the results. First, a large basis set is required to obtain two-body energies at the CBS limit, but it is only necessary for dimers with a small intermolecular distance. The contributions of dimers with a large separation can be obtained using a small basis set such as AVDZ. Moreover, the correlation contributions of the distant dimers can be obtained by extrapolation of the correlation energy with the cutoff distance. The (T) terms need to be considered even for the distant contributions to the two-body energy as they affect the response properties. They can be evaluated using a small basis set (AVDZ) for the three-body contributions and can only be neglected in the four-body energy.

We have assessed the suitability of MP2 and RPA approaches for the subtractive embedding schemes for the computation of the total binding energy. The comparison against reference CCSD(T) data supports the following observations.

1. The performance of MP2 for distant two-body dimers is relatively good for aliphatic systems but deteriorates significantly for π -electron systems with errors.
2. The largest difference between MP2 and CCSD(T) occurs for the three-body interactions, around 1 kJ/mol. The cause of this error is the lack of three-body correlation and, thus, could be at least partially corrected using the three-body correlation Axilrod–Teller–Muto formula or similar.^{109–111}
3. MP2 recovers about 50% of the correlated contribution to the non-additive four-body energies. This is enough for accurate total binding energies as for all tested systems the magnitude of the four-body contribution is small (below 0.1 kJ/mol).
4. Compared to MP2, the performance of RPA does not deteriorate as strongly for π -electron systems.
5. RPA includes three-body correlation terms, but those contributions are of poor quality if there are close contacts in the trimer of molecules.
6. The MBE errors of RPA are clearly affected by the DFT states used to evaluate the RPA energy. While PBE-based RPA with RSE corrections leads to the smallest overall errors, it relies considerably on error cancellation between the different MBE

terms. Using SCAN instead of PBE increases the error of the binding energies by a few percent; however, the many-body errors are substantially reduced for RPA(SCAN). Therefore, the SCAN-based RPA is more suitable for the subtractive embedding strategy than using the PBE states.

7. On the technical side, the basis-set convergence of the three- and four-body energies at the RPA level is slower than for the HF-based methods, which increases the computational cost.

Overall, we conclude that dispersion-dominated systems remain a challenge for approximate electronic structure methods. The many-body-resolved binding energies allow obtaining much detailed information about the origin of errors for a given method compared to assessment based only on the total binding energies. Moreover, the data presented in this work are intended as a benchmark for the development of novel low-scaling approaches.

SUPPLEMENTARY MATERIAL

The [supplementary material](#) contains the structures used in this study, summary of convergence tests, and additional figures showing (T) energy convergence with distance cutoff.

ACKNOWLEDGMENTS

This work was supported by the European Research Council (ERC) under European Union's Horizon 2020 Research and Innovation Program (Grant Agreement No. 759721). We acknowledge the computational resources provided by the IT4Innovations National Supercomputing Center (Grant No. LM2015070), CESNET (Grant No. LM2015042), CERIT Scientific Cloud (Grant No. LM2015085), and e-Infrastruktura CZ (e-INFRA CZ Grant No. LM2018140) supported by the Czech Ministry of Education, Youth, and Sports. This research was supported, in part, by PLGrid infrastructure.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Khanh Ngoc Pham: Data curation (equal); Formal analysis (equal); Investigation (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal).
Marcin Modrzejewski: Investigation (equal); Resources (equal); Software (equal); Validation (equal); Writing – review & editing (equal).
Jiří Klimeš: Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Resources (equal); Supervision (equal); Visualization (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its [supplementary material](#) and openly available at https://github.com/klimes/Pham_MBE_CH, Ref. 90.

REFERENCES

- ¹S. L. Price, "Predicting crystal structures of organic compounds," *Chem. Soc. Rev.* **43**, 2098–2111 (2014).
- ²G. J. O. Beran, "Modeling polymorphic molecular crystals with electronic structure theory," *Chem. Rev.* **116**, 5567–5613 (2016).
- ³R. M. Richard, K. U. Lao, and J. M. Herbert, "Understanding the many-body expansion for large systems. I. Precision considerations," *J. Chem. Phys.* **141**, 014108 (2014).
- ⁴J. Yang, W. Hu, D. Usyat, D. Matthews, M. Schütz, and G. K.-L. Chan, "Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy," *Science* **345**, 640–643 (2014).
- ⁵J. Nyman and G. M. Day, "Static and lattice vibrational energy differences between polymorphs," *CrystEngComm* **17**, 5154–5165 (2015).
- ⁶C. Červinka and G. J. O. Beran, "Ab initio prediction of the polymorph phase diagram for crystalline methanol," *Chem. Sci.* **9**, 4622–4629 (2018).
- ⁷J. Hofierka and J. Klimeš, "Binding energies of molecular solids from fragment and periodic approaches," *Electron. Struct.* **3**, 034010 (2021).
- ⁸G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, "Towards an exact description of electronic wavefunctions in real solids," *Nature* **493**, 365–370 (2013).
- ⁹M. R. Kennedy, A. R. McDonald, A. E. DePrince, M. S. Marshall, R. Podeszwa, and C. D. Sherrill, "Communication: Resolving the three-body contribution to the lattice energy of crystalline benzene: Benchmark results from coupled-cluster theory," *J. Chem. Phys.* **140**, 121104 (2014).
- ¹⁰A. Zen, J. G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè, and A. Michaelides, "Fast and accurate quantum Monte Carlo for molecular crystals," *Proc. Natl. Acad. Sci. U. S. A.* **115**, 1724–1729 (2018).
- ¹¹K. Liao, X.-Z. Li, A. Alavi, and A. Grüneis, "A comparative study using state-of-the-art electronic structure theories on solid hydrogen phases under high pressures," *npj Comput. Mater.* **5**, 110 (2019).
- ¹²A. M. Reilly and A. Tkatchenko, "Understanding the role of vibrations, exact exchange, and many-body van der Waals interactions in the cohesive properties of molecular crystals," *J. Chem. Phys.* **139**, 024705 (2013).
- ¹³L. Goerigk and S. Grimme, "A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions," *Phys. Chem. Chem. Phys.* **13**, 6670 (2011).
- ¹⁴P. Jurečka, J. Šponer, J. Černý, and P. Hobza, "Benchmark database of accurate [MP2 and CCSD(T) complete basis set limit] interaction energies of small model complexes, DNA base pairs, and amino acid pairs," *Phys. Chem. Chem. Phys.* **8**, 1985–1993 (2006).
- ¹⁵A. Otero-de-la-Roza and E. R. Johnson, "A benchmark for non-covalent interactions in solids," *J. Chem. Phys.* **137**, 054103 (2012).
- ¹⁶L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, and S. Grimme, "A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions," *Phys. Chem. Chem. Phys.* **19**, 32184–32215 (2017).
- ¹⁷D. Lu, Y. Li, D. Rocca, and G. Galli, "Ab initio calculation of van der Waals bonded molecular crystals," *Phys. Rev. Lett.* **102**, 206411 (2009).
- ¹⁸M. Del Ben, J. Hütter, and J. VandeVondele, "Second-order Møller-Plesset perturbation theory in the condensed phase: An efficient and massively parallel Gaussian and plane waves approach," *J. Chem. Theory Comput.* **8**, 4177–4188 (2012).
- ¹⁹M. Del Ben, J. VandeVondele, and B. Slater, "Periodic MP2, RPA, and boundary condition assessment of hydrogen ordering in ice XV," *J. Phys. Chem. Lett.* **5**, 4122–4128 (2014).
- ²⁰M. Macher, J. Klimeš, C. Franchini, and G. Kresse, "The random phase approximation applied to ice," *J. Chem. Phys.* **140**, 084502 (2014).
- ²¹J. Klimeš, M. Kaltak, E. Maggio, and G. Kresse, "Singles correlation energy contributions in solids," *J. Chem. Phys.* **143**, 102816 (2015).
- ²²J. Klimeš, "Lattice energies of molecular solids from the random phase approximation with singles corrections," *J. Chem. Phys.* **145**, 094506 (2016).
- ²³M. Modrzejewski, S. Yourdkhani, S. Śmiga, and J. Klimeš, "Random-phase approximation in many-body noncovalent systems: Methane in a dodecahedral water cage," *J. Chem. Theory Comput.* **17**, 804–817 (2021).
- ²⁴X. Ren, A. Tkatchenko, P. Rinke, and M. Scheffler, "Beyond the random-phase approximation for the electron correlation energy: The importance of single excitations," *Phys. Rev. Lett.* **106**, 153003 (2011).
- ²⁵M. Del Ben, J. Hütter, and J. VandeVondele, "Electron correlation in the condensed phase from a resolution of identity approach based on the Gaussian and plane waves scheme," *J. Chem. Theory Comput.* **9**, 2654–2671 (2013).
- ²⁶M. Kaltak, J. Klimeš, and G. Kresse, "A cubic scaling algorithm for the random phase approximation: Defect calculations for large Si model structures," *Phys. Rev. B* **90**, 054115 (2014).
- ²⁷S. Grimme, "Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies," *J. Chem. Phys.* **118**, 9095–9102 (2003).
- ²⁸A. Grüneis, M. Marsman, and G. Kresse, "Second-order Møller-Plesset perturbation theory applied to extended systems. II. Structural and energetic properties," *J. Chem. Phys.* **133**, 074107 (2010).
- ²⁹M. Pitoňák and A. Heßelmann, "Accurate intermolecular interaction energies from a combination of MP2 and TDDFT response theory," *J. Chem. Theory Comput.* **6**, 168 (2010).
- ³⁰J. Řezáč, Y. Huang, P. Hobza, and G. J. O. Beran, "Benchmark calculations of three-body intermolecular interactions and the performance of low-cost electronic structure methods," *J. Chem. Theory Comput.* **11**, 3065–3079 (2015).
- ³¹J. Řezáč, C. Greenwell, and G. J. O. Beran, "Accurate noncovalent interactions via dispersion-corrected second-order Møller-Plesset perturbation theory," *J. Chem. Theory Comput.* **14**, 4711–4721 (2018).
- ³²J. Lee and M. Head-Gordon, "Regularized orbital-optimized second-order Møller-Plesset perturbation theory: A reliable fifth-order-scaling electron correlation model with orbital energy dependent regularizers," *J. Chem. Theory Comput.* **14**, 5203–5219 (2018).
- ³³T. Goldzak, X. Wang, H.-Z. Ye, and T. C. Berkelbach, "Accurate thermochemistry of covalent and ionic solids from spin-component-scaled MP2," *J. Chem. Phys.* **157**, 174112 (2022).
- ³⁴E. Keller, T. Tsatsoulis, K. Reuter, and J. T. Margraf, "Regularized second-order correlation methods for extended systems," *J. Chem. Phys.* **156**, 024106 (2022).
- ³⁵J. Harl and G. Kresse, "Accurate bulk properties from approximate many-body techniques," *Phys. Rev. Lett.* **103**, 056401 (2010).
- ³⁶L. Schimka, J. Harl, A. Stroppa, A. Grüneis, M. Marsman, F. Mittendorfer, and G. Kresse, "Accurate surface and adsorption energies from many-body perturbation theory," *Nat. Mater.* **9**, 741 (2010).
- ³⁷B. D. Nguyen, G. P. Chen, M. M. Agee, A. M. Burow, M. P. Tang, and F. Furche, "Divergence of many-body perturbation theory for noncovalent interactions of large molecules," *J. Chem. Theory Comput.* **16**, 2258–2273 (2020).
- ³⁸F. Furche, "Molecular tests of the random phase approximation to the exchange-correlation energy functional," *Phys. Rev. B* **64**, 195120 (2001).
- ³⁹J. Harl and G. Kresse, "Cohesive energy curves for noble gas solids calculated by adiabatic connection fluctuation-dissipation theory," *Phys. Rev. B* **77**, 045136 (2008).
- ⁴⁰H. Eshuis and F. Furche, "A parameter-free density functional that works for noncovalent interactions," *J. Phys. Chem. Lett.* **2**, 983–989 (2011).
- ⁴¹H. Eshuis and F. Furche, "Basis set convergence of molecular correlation energy differences within the random phase approximation," *J. Chem. Phys.* **136**, 084105 (2012).
- ⁴²J. Paier, X. Ren, P. Rinke, G. E. Scuseria, A. Grüneis, G. Kresse, and M. Scheffler, "Assessment of correlation energies based on the random-phase approximation," *New J. Phys.* **14**, 043002 (2012).
- ⁴³A. Grüneis, M. Marsman, J. Harl, L. Schimka, and G. Kresse, "Making the random phase approximation to electronic correlation accurate," *J. Chem. Phys.* **131**, 154115 (2009).
- ⁴⁴J. E. Bates and F. Furche, "Communication: Random phase approximation renormalized many-body perturbation theory," *J. Chem. Phys.* **139**, 171103 (2013).
- ⁴⁵T. Olsen and K. S. Thygesen, "Beyond the random phase approximation: Improved description of short-range correlation by a renormalized adiabatic local density approximation," *Phys. Rev. B* **88**, 115131 (2013).

- ⁴⁶T. Olsen and K. S. Thygesen, "Accurate ground-state energies of solids and molecules from time-dependent density-functional theory," *Phys. Rev. Lett.* **112**, 203001 (2014).
- ⁴⁷N. Colonna, M. Hellgren, and S. de Gironcoli, "Correlation energy within exact-exchange adiabatic connection fluctuation-dissipation theory: Systematic development and simple approximations," *Phys. Rev. B* **90**, 125150 (2014).
- ⁴⁸B. Mussard, D. Rocca, G. Jansen, and J. G. Ángyán, "Dielectric matrix formulation of correlation energies in the random phase approximation: Inclusion of exchange effects," *J. Chem. Theory Comput.* **12**, 2191–2202 (2016).
- ⁴⁹M. Hellgren, N. Colonna, and S. de Gironcoli, "Beyond the random phase approximation with a local exchange vertex," *Phys. Rev. B* **98**, 045117 (2018).
- ⁵⁰G. P. Chen, M. M. Agee, and F. Furche, "Performance and scope of perturbative corrections to random-phase approximation energies," *J. Chem. Theory Comput.* **14**, 5701–5714 (2018).
- ⁵¹F. Hummel, A. Grüneis, G. Kresse, and P. Ziesche, "Screened exchange corrections to the random phase approximation from many-body perturbation theory," *J. Chem. Theory Comput.* **15**, 3223–3236 (2019).
- ⁵²A. Görling, "Hierarchies of methods towards the exact Kohn-Sham correlation energy based on the adiabatic-connection fluctuation-dissipation theorem," *Phys. Rev. B* **99**, 235120 (2019).
- ⁵³P. Verma and R. J. Bartlett, "Increasing the applicability of density functional theory. II. Correlation potentials from the random phase approximation and beyond," *J. Chem. Phys.* **136**, 044105 (2012).
- ⁵⁴P. Bleiziffer, A. Heßelmann, and A. Görling, "Efficient self-consistent treatment of electron correlation within the random phase approximation," *J. Chem. Phys.* **139**, 084113 (2013).
- ⁵⁵N. L. Nguyen, N. Colonna, and S. de Gironcoli, "Ab initio self-consistent total-energy calculations within the EXX/RPA formalism," *Phys. Rev. B* **90**, 045138 (2014).
- ⁵⁶M. Hellgren, F. Caruso, D. R. Rohr, X. Ren, A. Rubio, M. Scheffler, and P. Rinke, "Static correlation and electron localization in molecular dimers from the self-consistent RPA and GW approximation," *Phys. Rev. B* **91**, 165110 (2015).
- ⁵⁷P. Bleiziffer, M. Krug, and A. Görling, "Self-consistent Kohn-Sham method based on the adiabatic-connection fluctuation-dissipation theorem and the exact-exchange kernel," *J. Chem. Phys.* **142**, 244108 (2015).
- ⁵⁸Y. Jin, D. Zhang, Z. Chen, N. Q. Su, and W. Yang, "Generalized optimized effective potential for orbital functionals and self-consistent calculation of random phase approximations," *J. Phys. Chem. Lett.* **8**, 4746–4751 (2017).
- ⁵⁹V. K. Voora, S. G. Balasubramani, and F. Furche, "Variational generalized Kohn-Sham approach combining the random-phase-approximation and Green's-function methods," *Phys. Rev. A* **99**, 012518 (2019).
- ⁶⁰A. Thierbach, D. Schmidtel, and A. Görling, "Robust and accurate hybrid random-phase-approximation methods," *J. Chem. Phys.* **151**, 144117 (2019).
- ⁶¹D. Graf and C. Ochsnerfeld, "A range-separated generalized Kohn-Sham method including a long-range nonlocal random phase approximation correlation potential," *J. Chem. Phys.* **153**, 244118 (2020).
- ⁶²S. Riemelmoser, M. Kaltak, and G. Kresse, "Optimized effective potentials from the random-phase approximation: Accuracy of the quasiparticle approximation," *J. Chem. Phys.* **154**, 154103 (2021).
- ⁶³J. M. Yu, B. D. Nguyen, J. Tsai, D. J. Hernandez, and F. Furche, "Selfconsistent random phase approximation methods," *J. Chem. Phys.* **155**, 040902 (2021).
- ⁶⁴M. Modrzejewski, S. Yourdkhani, and J. Klimeš, "Random phase approximation applied to many-body noncovalent systems," *J. Chem. Theory Comput.* **16**, 427–442 (2020).
- ⁶⁵J. G. Brandenburg, T. Maas, and S. Grimme, "Benchmarking DFT and semiempirical methods on structures and lattice energies for ten ice polymorphs," *J. Chem. Phys.* **142**, 124104 (2015).
- ⁶⁶H. Stoll, "On the correlation energy of graphite," *J. Chem. Phys.* **97**, 8449 (1992).
- ⁶⁷K. Doll, M. Dolg, P. Fulde, and H. Stoll, "Correlation effects in ionic crystals: The cohesive energy of MgO," *Phys. Rev. B* **52**, 4842–4848 (1995).
- ⁶⁸S. S. Xantheas, "Ab initio studies of cyclic water clusters (H_2O)_n, n = 1–6. II. Analysis of many-body interactions," *J. Chem. Phys.* **100**, 7523–7534 (1994).
- ⁶⁹U. Góra, R. Podeszwa, W. Cencek, and K. Szalewicz, "Interaction energies of large clusters from many-body expansion," *J. Chem. Phys.* **135**, 224102 (2011).
- ⁷⁰P. J. Bygrave, N. L. Allan, and F. R. Manby, "The embedded many-body expansion for energetics of molecular crystals," *J. Chem. Phys.* **137**, 164102 (2012).
- ⁷¹G. J. O. Beran and K. Nanda, "Predicting organic crystal lattice energies with chemical accuracy," *J. Phys. Chem. Lett.* **1**, 3480–3487 (2010).
- ⁷²S. Wen and G. J. O. Beran, "Accurate molecular crystal lattice energies from a fragment QM/MM approach with on-the-fly ab initio force-field parameterization," *J. Chem. Theory Comput.* **7**, 3733–3742 (2011).
- ⁷³C. Müller and D. Usvyat, "Incrementally corrected periodic local MP2 calculations: I. The cohesive energy of molecular crystals," *J. Chem. Theory Comput.* **9**, 5590–5598 (2013).
- ⁷⁴G. A. Dolgonos, O. A. Loboda, and A. D. Boese, "Development of embedded and performance of density functional methods for molecular crystals," *J. Phys. Chem. A* **122**, 708 (2018).
- ⁷⁵A. Hermann and P. Schwerdtfeger, "Ground-state properties of crystalline ice from periodic Hartree-Fock calculations and a coupled-cluster-based many-body decomposition of the correlation energy," *Phys. Rev. Lett.* **101**, 183005 (2008).
- ⁷⁶J. Košata, P. Merkl, P. Teeratchanan, and A. Hermann, "Stability of hydrogen hydrates from second-order Møller-Plesset perturbation theory," *J. Phys. Chem. Lett.* **9**, 5624–5629 (2018).
- ⁷⁷S. Wen, K. Nanda, Y. Huang, and G. J. O. Beran, "Practical quantum mechanics-based fragment methods for predicting molecular crystal properties," *Phys. Chem. Chem. Phys.* **14**, 7578–7590 (2012).
- ⁷⁸M. Marsman, A. Grüneis, J. Paier, and G. Kresse, "Second-order Møller-Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set," *J. Chem. Phys.* **130**, 184103 (2009).
- ⁷⁹T. Schäfer, B. Ramberger, and G. Kresse, "Quartic scaling MP2 for solids: A highly parallelized algorithm in the plane wave basis," *J. Chem. Phys.* **146**, 104101 (2017).
- ⁸⁰J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- ⁸¹J. Sun, A. Ruzsinszky, and J. P. Perdew, "Strongly constrained and appropriately normed semilocal density functional," *Phys. Rev. Lett.* **115**, 036402 (2015).
- ⁸²G. J. H. van Nes and A. Vos, "Single-crystal structures and electron density distributions of ethane, ethylene and acetylene. I. Single-crystal X-ray structure determinations of two modifications of ethane," *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **34**, 1947–1956 (1978).
- ⁸³G. J. H. van Nes and A. Vos, "Single-crystal structures and electron density distributions of ethane, ethylene and acetylene. III. Single-crystal X-ray structure determination of ethylene at 85 K," *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **35**, 2593–2601 (1979).
- ⁸⁴R. K. McMullan, Å. Kvick, and P. Popelier, "Structures of cubic and orthorhombic phases of acetylene by single-crystal neutron diffraction," *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **48**, 726–731 (1992).
- ⁸⁵C. R. Groom, I. J. Bruno, M. P. Lightfoot, and S. C. Ward, "The cambridge structural database," *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **72**, 171–179 (2016).
- ⁸⁶M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, "Van der Waals density functional for general geometries," *Phys. Rev. Lett.* **92**, 246401 (2004).
- ⁸⁷G. Román-Pérez and J. M. Soler, "Efficient implementation of a van der Waals density functional: Application to double-wall carbon nanotubes," *Phys. Rev. Lett.* **103**, 096102 (2009).
- ⁸⁸J. Klimeš, D. R. Bowler, and A. Michaelides, "Chemical accuracy for the van der Waals density functional," *J. Phys.: Condens. Matter* **22**, 022201 (2009).

- ⁸⁹J. Klimeš, D. R. Bowler, and A. Michaelides, "Van der Waals density functionals applied to solids," *Phys. Rev. B* **83**, 195131 (2011).
- ⁹⁰Outputs of calculations and processing script, https://github.com/klimes/Pham_MBE_CH.
- ⁹¹C. H. Borca, B. W. Bakr, L. A. Burns, and C. D. Sherrill, "CrystaLattE: Automated computation of lattice energies of organic crystals exploiting the many-body expansion to achieve dual-level parallelism," *J. Chem. Phys.* **151**, 144103 (2019).
- ⁹²H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, "Molpro: A general-purpose quantum chemistry program package," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 242–253 (2012).
- ⁹³Z.-h. Yang, H. Peng, J. Sun, and J. P. Perdew, "More realistic band gaps from meta-generalized gradient approximations: Only in a generalized Kohn-Sham scheme," *Phys. Rev. B* **93**, 205205 (2016).
- ⁹⁴A. P. Bartók and J. R. Yates, "Regularized scan functional," *J. Chem. Phys.* **150**, 161101 (2019).
- ⁹⁵J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, "Accurate and numerically efficient r²SCAN meta-generalized gradient approximation," *J. Phys. Chem. Lett.* **11**, 8208–8215 (2020).
- ⁹⁶R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," *J. Chem. Phys.* **96**, 6796–6806 (1992).
- ⁹⁷A. Halkier, W. Klopper, T. Helgaker, P. Jørgensen, and P. R. Taylor, "Basis set convergence of the interaction energy of hydrogen-bonded complexes," *J. Chem. Phys.* **111**, 9157–9167 (1999).
- ⁹⁸W. Kutzelnigg and W. Klopper, "Wave functions with terms linear in the inter-electronic coordinates to take care of the correlation cusp. I. General theory," *J. Chem. Phys.* **94**, 1985–2001 (1991).
- ⁹⁹H.-J. Werner, T. B. Adler, and F. R. Manby, "General orbital invariant MP2-F12 theory," *J. Chem. Phys.* **126**, 164102 (2007).
- ¹⁰⁰B. Brauer, M. K. Kesharwani, S. Kozuch, and J. M. L. Martin, "The S66x8 benchmark for noncovalent interactions revisited: Explicitly correlated ab initio methods and density functional theory," *Phys. Chem. Chem. Phys.* **18**, 20905–20925 (2016).
- ¹⁰¹G. Knizia, T. B. Adler, and H.-J. Werner, "Simplified CCSD(T)-F12 methods: Theory and benchmarks," *J. Chem. Phys.* **130**, 054104 (2009).
- ¹⁰²T. B. Adler, G. Knizia, and H.-J. Werner, "A simple and efficient CCSD(T)-F12 approximation," *J. Chem. Phys.* **127**, 221106 (2007).
- ¹⁰³J. Noga and J. Šimunek, "On the one-particle basis set relaxation in R12 based theories," *Chem. Phys.* **356**, 1–6 (2009).
- ¹⁰⁴P. Jurečka and P. Hobza, "On the convergence of the $(\Delta E^{\text{CCSD}(T)} - \Delta E^{\text{MP}2})$ term for complexes with multiple H-bonds," *Chem. Phys. Lett.* **365**, 89 (2002).
- ¹⁰⁵I. R. Dagg, A. Anderson, W. Smith, M. Missio, C. G. Joslin, and L. A. A. Read, "The quadrupole moment of acetylene from collision-induced absorption in a gaseous mixture with argon," *Can. J. Phys.* **66**, 453–459 (1988).
- ¹⁰⁶R. Lindh and B. Liu, "Accurate *ab initio* calculations of the quadrupole moment of acetylene. a combined study of basis set, correlation, and vibrational effects," *J. Chem. Phys.* **94**, 4356–4368 (1991).
- ¹⁰⁷P. Hobza, H. L. Selzle, and E. W. Schlag, "Potential energy surface for the benzene dimer. Results of *ab initio* CCSD(T) calculations show two nearly isoenergetic structures: T-Shaped and parallel-displaced," *J. Phys. Chem.* **100**, 18790–18794 (1996).
- ¹⁰⁸M. Hapka, Ł. Rajchel, M. Modrzewski, R. Schäffer, G. Chałasiński, and M. M. Szczęśniak, "The nature of three-body interactions in DFT: Exchange and polarization effects," *J. Chem. Phys.* **147**, 084106 (2017).
- ¹⁰⁹Y. Huang and G. J. O. Beran, "Reliable prediction of three-body intermolecular interactions using dispersion-corrected second-order Møller-Plesset perturbation theory," *J. Chem. Phys.* **143**, 044113 (2015).
- ¹¹⁰B. M. Axilrod and E. Teller, "Interaction of the van der Waals type between three atoms," *J. Chem. Phys.* **11**, 299 (1943).
- ¹¹¹Y. Muto, "Force between nonpolar molecules," *J. Phys. Math. Soc. Jpn.* **17**, 629 (1943).