

Blind test of density-functional-based methods on intermolecular interaction energies

DeCarlos E. Taylor

U.S. Army Research Laboratory, Aberdeen Proving Ground, MD 21005-5069

János G. Ángyán

*CNRS, CRM², UMR 7036, Vandœuvre-lès-Nancy, F-54506, France and
Université de Lorraine, CRM², UMR 7036, Vandœuvre-lès-Nancy, F-54506, France*

Giulia Galli, Cui Zhang, and Francis Gygi

Department of Chemistry, University of California Davis, Davis, CA 95616

Kimihiko Hirao

*RIKEN Advanced Institute for Computational Science,
Computational Chemistry Unit, Kobe, Hyogo 6500047, Japan*

O. Anatole von Lilienfeld

*General Chemistry (ALGC), Free University Brussels (VUB), Pleinlaan 2,
1050 Brussel, Belgium; Institute of Physical Chemistry and National Center
for Computational Design and Discovery of Novel Materials (MARVEL),
Department of Chemistry, University of Basel, 4056 Basel, Switzerland*

Rafal Podeszwa

Institute of Chemistry, University of Silesia, Szkolna 9, 40-006 Katowice, Poland

Ireneusz W. Bulik, Thomas M. Henderson, and Gustavo E. Scuseria

Department of Chemistry, Rice University, Houston, TX 77005

Julien Toulouse

*Laboratoire de Chimie Théorique, Sorbonne Universités,
Université Pierre et Marie Curie, CNRS, F-75005, Paris, France*

Roberto Peverati and Donald G. Truhlar

Department of Chemistry, University of Minnesota, Minneapolis, MN 55455

Krzysztof Szalewicz

Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716

In the past decade, a number of approaches have been developed to fix the failure of (semi)local density-functional theory (DFT) in describing intermolecular interactions. The performance of several such approaches with respect to highly accurate benchmarks are compared here on a set of separation-dependent interaction energies for ten dimers. Since the benchmarks were unknown before the DFT-based results were collected, this comparison constitutes a blind test of these methods.

I. INTRODUCTION

Computer modeling of materials has become a prominent means to reduce the cost associated with notional material synthesis, testing, and application. With many significant developments in electronic structure theory and computer science, as well as due to rapid increases in available computer power, material modeling has reached a level of maturity that yields reliable predictions and, in fact, the computed observables for a variety of properties often rival experiment in accuracy. There is a wide range of computational methods available. The most accurate methods use wave-functions (WF) and utilize many-body perturbation theory (MBPT) or coupled cluster (CC)

expansions. The price for this accuracy is steep scaling of the cost with system size which, even with access to powerful supercomputers and highly scalable software packages, renders these methods inapplicable to many problems of interest. Another group of computational approaches is based on density-functional theory (DFT). These approaches utilize only electron densities and occupied orbitals and therefore scale much better than WF-based methods. In fact, DFT is the only computationally viable option for first-principles predictions involving molecules beyond the reach of WF-based methods, as well as for condensed phase systems. DFT has been applied to virtually all molecular systems as well as to all types of condensed phases: from metals to molecular crystals. However, DFT has some reliability prob-

lems. Whereas it is in principle an exact theory [1], the exact DFT functional is unknown and one has to use one of the many approximate variants of DFT. Although such variants have been successfully applied to a variety of problems (many variants are specifically tailored for given types of problems), starting from mid 1980s cases were reported of even the best DFT methods of that time failing when applied to intermolecular interactions. These interactions, called also van der Waals (vdW) or noncovalent interactions, are at least one order of magnitude weaker than the chemical bonds and can be decomposed into contributions of different physical origins: electrostatic, induction, dispersion, and exchange. These components are defined in symmetry-adapted perturbation theory (SAPT) [2, 3]. While DFT was able to describe reasonably well interactions dominated by electrostatics and induction components, it was failing badly in calculations of the potential energy curves for pairs of rare-gas atoms, dominated by dispersion energies [4–6]. The dispersion energies originate from mutual correlations of electron motions between monomers. For a pair of rare-gas atoms at separations a couple times larger than the vdW minimum distance, the dispersion energy constitutes nearly 100% of the interaction energy, so it was clear that the failure is related to this component. One should note that while the dimers of rare-gas atoms are an extreme example of the importance of dispersion interactions, these interactions are significant for all other intermolecular complexes at all separations. As pointed out by Kristyan and Pulay [5] in 1994, one reason for the failure is the local (or semilocal) character of the majority of approximate DFT functionals. As a result, the “exchange-correlation holes” modeled by such functionals have a range on the order of about 1 Å, whereas dispersion interactions involve correlations of electron motion at distances of several Ångströms which is typical of separations between monomers. Another type of interactions where DFT was found to have problems are those with large charge-transfer effects in the induction energies [7].

The failures of DFT for intermolecular interactions have spawned intense research activity directed toward fixing this problem. The initial attempts were based on hopes that one of the variants of DFT will actually work for such cases. However, methods performing well on a few dimers were later found to perform poorly on others. Searches for the “best” DFT method for intermolecular interactions were later systematized to utilize sets of benchmarks including dozens of dimers representing various classes of intermolecular interactions. These tests have shown that such a DFT method does not exist within the set of approaches published up to that time. For example, Zhao and Truhlar [8] examined 45 different density functionals on a set of 28 dimers at their minimum configurations and found that the best performing functional had an average unsigned error of 0.46 kcal/mol which should be considered large since the interaction energies of the set range from -0.04 to -16.1 kcal/mol, with

the average of -3.8 kcal/mol. Investigations of this type have also shown that different functionals overbind, underbind, or do not bind at all, a given dimer. This fact indicates that the inability to recover dispersion energies is not the only problem of DFT. If this were the case, all methods should underbind, or not bind, since the dispersion energy is always negative (for all distances). Some authors pointed out to a lack of balance between kinetic and exchange interactions in Kohn-Sham DFT as a possible other problem [4, 9].

The proposed strategies for making DFT-based methods more predictive for intermolecular interactions generally fall into the following four categories: *(i)* Optimization of parameters in existing functionals on training sets including interaction energies; *(ii)* Addition of dispersion energies in the form of an asymptotic expansion to existing functionals, the so-called DFT+D approaches; *(iii)* Development of nonlocal correlation density functionals which should in principle be capable of describing dispersion interactions. Since the existing methods of this type either do not include couplings to the local parts of the functional, or the couplings result in practically negligible effects, methods of this group can be viewed as a type of DFT+D methods. *(iv)* Application of “post-DFT” approaches, i.e., approaches that include virtual orbitals.

Methods belonging to each of the four categories defined above were first formulated in the early 2000s and have been actively developed later on. We will refer to these methods as “intermolecular interactions cognizant” methods, whereas other DFT variants will be called standard methods.

With quite a number of intermolecular interactions cognizant DFT methods available, the question arises which of them is most suitable for application to problems in materials science. Obviously, the most relevant feature is *(a)* the best accuracy in all regions of the potential energy surface. The next factor to consider is *(b)* computational efficiency. Finally, a given method should also be *(c)* efficient when applied to condensed-phase systems using periodic boundary conditions, which gives preference to methods not employing the so-called “exact” exchange, i.e., not including the Hartree-Fock (HF) exchange operator computed with Kohn-Sham (KS) orbitals. Whereas there are a number of papers in the literature comparing the accuracy of various intermolecular interactions cognizant DFT methods, it is not easy for researchers not directly involved in such work to find out which methods are the best. Most of the published papers place emphasis on the methods developed by the authors of these papers and sometimes the test set of benchmarks is the same, or very similar, to the training set used to optimize parameters. The most extensive tests including a variety of methods and distance-dependent interaction energies have been published in Refs. 10–12.

The present paper is related to a workshop entitled “Dispersion Interactions and Density Functional Theory”

co-organized by the Army Research Laboratory (ARL) and the University of Delaware (UD). The workshop consisted of presentations (available online) [13] covering developments of DFT-based methods for reliable treatment of intermolecular interactions. A part of this workshop was a “blind” test of such methods. Each speaker received a set of dimer geometries and was asked to compute interaction energies applying DFT-based methods developed or used in their groups. Independently, the organizers of the workshop computed benchmark interaction energies which were not made known to the speakers prior to the workshop. At the workshop, the results were revealed and compared.

During the comparison of the data when working on the present manuscript, it became clear that the benchmarks cannot be accurate for large separation. The problem was first traced to a coding error in the coupled cluster triple excitation subroutines of the NWChem software package [14]. The error was fixed by the authors of NWChem and all the benchmark calculations were repeated. An analysis of the new results revealed another problem, resulting from loss of accuracy due to linear dependencies in the largest basis sets used by us. We found a way around this problem and repeated relevant calculations, as described in Sec. IV A. Both problems resulted in a significant delay of publishing our results.

As a benchmark set, we have chosen a set of dimers relevant to the field of energetic materials. This set includes 10 dimers which range in size from 6 to 32 atoms. For each dimer, we considered the equilibrium configuration, as well as several configurations with different separations between the centers of mass of the monomers (sampling the repulsive wall, minimum, and asymptotic regions), but with the same relative orientation as in the minimum configuration. In this way, we created a set of 80 benchmark interaction energies for the following dimers:

1. Water: $(\text{H}_2\text{O})_2$;
2. Ethanol: $(\text{C}_2\text{H}_5\text{OH})_2$;
3. Nitromethane: $(\text{H}_3\text{C}-\text{NO}_2)_2$;
4. Methylformate: $(\text{C}_2\text{H}_4\text{O}_2)_2$;
5. Benzene-methane: $\text{C}_6\text{H}_6-\text{CH}_4$;
6. Benzene-water: $\text{C}_6\text{H}_6-\text{H}_2\text{O}$;
7. Imidazole: $(\text{C}_3\text{H}_4\text{N}_2)_2$;
8. Nitrobenzene: $(\text{C}_6\text{H}_5\text{NO}_2)_2$;
9. 1,1-diamino-2,2-dinitroethylene (FOX-7): $(\text{C}_2\text{N}_4\text{O}_4\text{H}_4)_2$;
10. Ethylenedinitramine (EDNA): $(\text{C}_2\text{O}_4\text{N}_4\text{H}_6)_2$.

This benchmark set has several advantages compared to similar sets used in the literature. The latter sets tended to be dominated by dispersion-bonded and hydrogen-bonded dimers. Whereas our set does contain systems

belonging to these two categories, most of our dimers cannot be classified in this way. Furthermore, most of the dimers have not been used in previous sets, which allows us to test methods on systems not used in optimization of these methods. For each system, interaction energies were computed applying the CC method with single, double, and perturbative triple excitations, CCSD(T), and using complete basis set (CBS) extrapolations, following the approach of Ref. 15.

Although in principle the participants could have computed the benchmark interaction energies themselves, which would have violated the “blind” character of the test, in practice this was not possible except for the smallest dimers in our set. As described later, the calculations of the benchmarks required extraordinary amounts of computer resources and took several months to complete which was more than the time available to the participants.

We have included 12 different intermolecular interactions cognizant DFT methods, representing all the approaches (i)–(iv) described above. We believe this selection is broader than in any previous survey of this type. These methods are briefly described in Sec. II. Technical details of the DFT-based calculations such as the basis sets are specified in Sec. III. The methodology used to obtain the benchmarks is outlined in Sec. IV. The results and their graphical and statistical analysis is presented in Sec. V. In the Supplementary Material [16], we include Cartesian coordinates for all configurations as well as the benchmark and DFT interaction energies for all entries in the data set.

II. DFT METHODS COGNIZANT OF INTERMOLECULAR INTERACTIONS

To facilitate the comparisons of various methods, we briefly describe in this section the four types of intermolecular interactions cognizant DFT methods defined in the Introduction. A summary of the salient features of each method used in this work is presented in Table I, along with some details of the computational approach (basis set, use of counterpoise correction, etc.). In most cases the authors used their own codes to perform the calculations.

A. Reoptimized standard functionals

This category involves methods that utilize standard generalized gradient approximation (GGA) functionals, meta GGA functionals, and functionals including a fraction of the exact exchange. The GGA functionals employ the exchange-correlation terms depending only on the density and the magnitude of its gradient. Meta GGA functionals include factors dependent on the kinetic energy density τ , i.e., on the sum of the squared magnitudes of gradients of occupied KS orbitals. The

TABLE I. Summary of methods. See text for explanations of acronyms; The term “meta” indicates that the exchange-correlation functionals contain enhancement factors depending on the kinetic energy density.

Method	Correlation	Exchange	Exact exchange	Parameters	Optimized on E_{int}	Counterpoise	Basis set
<i>Reoptimized standard functionals</i>							
M06	meta LDA/B96/V5XC	meta LDA/PBE/V5XC	27%	36	yes	no	jun-cc-pVTZ
M11	meta LDA/PBE	RSH LDA, meta PBE/RPBE	42.8%	46	yes	no	jun-cc-pVTZ
DCACP	PBE	PBE	none	2/atom	yes	yes	plane waves 300 Ry cutoff
<i>DFT+D approaches^a</i>							
dIDF+D	meta modified PBE	meta LDA	61.4%	11 ^b	no ^c	yes	aug-cc-pVTZ
B3LYP-D3	B88	LYP	20%	2 ^d	yes	yes	aug-cc-pVXZ $X = Q, 5, 6^e$
LC- ω PBE-D3	RSH PBE	PBE	$\omega = 0.4$	2 ^d	yes	yes	aug-cc-pVQZ
LC-BOP12+LRD	RSH B88	OP	$\omega = 0.42^f$	2 ^g	yes	yes	aug-cc-pVTZ
LCgau-BOP+LRD	RSH B88	OP	$\omega = 0.42^h$	2 ⁱ	yes	yes	aug-cc-pVTZ
<i>Nonlocal functionals</i>							
vdW-DF2	PW86	P86	none	none	no	yes	plane waves 85 Ry cutoff
<i>Post-KS methods</i>							
SAPT(DFT)	PBE	PBE	25%	none	no	BSSE-free	aug-cc-pVTZ+mb
RSH+lrMP2	RSH PBE ^j	RSH PBE ^j	$\omega = 0.5$	none	no	yes	aug-cc-pVTZ
RSH+lrRPax-SO2	RSH PBE ^j	RSH PBE ^j	$\omega = 0.5$	none	no	yes	aug-cc-pVTZ ^k

^a For DFT+D approaches, the parameters in dispersion functions fitted to vdW constants or to dispersion energies are not included in the parameter count, but parameters fitted to total energies are included.

^b Number of parameters in dIDF functional.

^c The dIDF and D parts were separately optimized on the corresponding components of interaction energies.

^d Parameters in the switching function of the dispersion energy.

^e $X = 6$ for water dimer, $X = Q$ for FOX-7, nitrobenzene, and EDNA dimers, and $X = 5$ for the remaining dimers.

^f A parameter in the OP functional, denoted by $q_{\text{OP}}^{\alpha\beta}$, was reoptimized to the value of 2.46 to match this value of ω .

^g The LDR parameters were $\kappa = 0.216$ a.u., $R_0 = 4.760$ a.u., and $\lambda = 0.228$.

^h The Gaussian exponent in the attenuation function was $\alpha = 0.011$ and the linear coefficient $k = -18$. The parameter $q_{\text{OP}}^{\alpha\beta} = 2.37$.

ⁱ The LDR parameters were $\kappa = 0.248$ a.u., $R_0 = 4.690$ a.u., and $\lambda = 0.229$.

^j Short-range PBE exchange and correlation functionals from Ref. [17]. Note that this short-range PBE exchange functional is different from the one used in LC- ω PBE-D3.

^k For the nitrobenzene, FOX-7, and EDNA dimers, the results were obtained as $E_{\text{int}} = E_{\text{int}}^{\text{RSH+lrRPax-SO2}}(\text{aug-cc-pVDZ}) + [E_{\text{int}}^{\text{RSH+lrMP2}}(\text{aug-cc-pVTZ}) - E_{\text{int}}^{\text{RSH+lrMP2}}(\text{aug-cc-pVDZ})]$.

functionals incorporating a fraction of the exact exchange are called “hybrid” functionals. Most of the functionals of these types depend on a number of adjustable parameters and these parameters are optimized using a training data set containing some benchmark interaction energies. The first functional of this type was published by Xu and Goddard [18] in 2004 and a large number of such functionals appeared since then. We have included in our survey three such functionals, described in the subsequent subsections.

1. M06

The M06 functional [19], where M stands for Minnesota, belongs to the suite of methods developed by Truhlar and coworkers. It can be classified as a hybrid meta-GGA functional. The GGA component of the exchange part is taken from the Perdew-Burke-Ernzerhof (PBE) functional [20] and the factors containing kinetic energy densities are modelled in the spirit of the van Voorhis-Scuseria exchange-correlation functional (VSXC) [21]. The M06 exchange functional includes in addition to

the PBE term also a local-density approximation (LDA) term. The correlation functional is in a modified LDA form with enhancement factors following a functional introduced by Becke in 1996 (B96) [22] and the VSXC functional. M06 includes a modest amount (27%) of the exact exchange. The set of benchmark interaction energies from Ref. [8] was used in the optimization of M06.

2. M11

The M11 functional [23] adopts a range-separated hybrid (RSH) approach [24–26], i.e., the exact exchange is included to a varying degree depending on the inter-electron separation. This is achieved by partitioning the Coulomb operator into the long- and short-range components:

$$\frac{1}{r_{12}} = \frac{\text{erf}(\omega r_{12})}{r_{12}} + \frac{\text{erfc}(\omega r_{12})}{r_{12}}, \quad (1)$$

where erf is the error function and erfc = 1-erf. Since erf(r) is zero at small r and one at large r , when this function is used in the exchange integrals, it introduces

the exact exchange only at larger separations. The second term in Eq. (1) switches off the Coulomb interaction for large r and is used in properly modified DFT exchange functionals [25, 27–29]. The extent of the DFT versus the exact exchange is controlled by the parameter ω . The M11 functional attenuates the Coulomb interactions in the DFT part using the expression derived by Gill, Adamson, and Pople [27] and by Savin [25] within the LDA method. The attenuated LDA exchange energy density is multiplied by the sum of the PBE [20] and revised PBE (RPBE) [30] exchange factors, each of the two factors multiplied in turn by its own meta enhancement factor. The RPBE functional differs from the PBE functional by using a different form of the exchange enhancement factor. M11 includes in addition to the range-separated exact exchange also 42.8% of all-separation exact exchange. The correlation functional in M11 is a sum of the LDA term and of the gradient correction to this term taken from the PBE functional, both multiplied by their own meta enhancement factors. The set of benchmark interaction energies from Ref. [8] was used in the optimization of M11.

3. DCACP

In the dispersion-corrected atom-centered potentials (DCACP) approach [31, 32], the atomic pseudopotentials routinely used in DFT calculations with plane-wave bases are parametrized and the parameters (two per atom) are fitted to a training set of benchmark interaction energies. These pseudopotentials can be used with any DFT functional. In the present work, the DCACP approach was used with the PBE functional in a plane-wave basis with a kinetic energy cutoff of 300 Rydberg.

B. DFT+D approaches

Although, as discussed above, inability to recover dispersion energies is certainly not the only problem that DFT encounters in predicting intermolecular interaction energies, one can hope that the predictions will be improved if dispersion energies calculated using a non-DFT approach are added to DFT interaction energies for selected functionals that tend to underbind molecular complexes. This idea was quite an obvious one since a method consisting of adding dispersion energies to HF interaction energies (so-called HFD approach) was proposed by Scoles and coworkers [33, 34] already in 1970s. An extended HFD method proposed in Ref. 35 is competitive with the DFT+D approaches. Note that the addition of the dispersion energy to the HF interaction energy is rigorous, i.e., by the definitions of these quantities, it involves no double counting, which is not the case for DFT+D. The DFT+D method was applied for the first time by Gianturco *et al.* [36] in 1998 and by Wu *et al.* [37] in 2001 for some specific systems. A generally

applicable approach was formulated by Wu and Yang [38] in 2002. These authors selected a few functionals that tended to underbind intermolecular complexes and added to the dimer interaction energies given by these functionals dispersion energies in the form of a simple atom-atom function:

$$E_{\text{disp}} = - \sum_{a \in A, b \in B} f(r_{ab}) C_6^{ab} / r_{ab}^6, \quad (2)$$

where A and B refer to the two monomers, r_{ab} are interatomic distances, C_6^{ab} are constants, and $f(r_{ab})$ is a damping function. This form of dispersion energy was well known in work on vdW clusters and in biomolecular force fields. The interatomic coefficients C_6^{ab} were expressed in terms of atomic coefficients using simple combination rules. The latter coefficients were fitted to reproduce molecule-molecule C_6 coefficients known from experiments, resulting in a universal atom-atom dispersion function. Wu and Yang examined several damping functions, but have not optimized parameters of these functions on benchmark interaction energies. This critical ingredient of the method was introduced by Grimme [39] in 2004 who also included an overall scaling factor multiplying the dispersion energy. This approach led to a significant improvement of the results since the dispersion energy was de facto used not only to reproduce this component, but also to cancel various deficiencies of the DFT methods to which it was applied. On the other hand, the fact that such approaches rely on error cancellations is their weak point from a physics point of view. This problem was removed by the dispersionless density functional plus dispersion (dIDF+D) method proposed by Pernal *et al.* [40] in 2009. This method uses a density functional optimized to reproduce the interaction energies with the dispersion (and exchange-dispersion) energies subtracted. Then atom-atom functions representing the true dispersion energies at *all* intermolecular separations can be added to the dIDF interaction energies. In fact, one can add accurate dispersion and exchange-dispersion energies computed *ab initio* using SAPT [3]. One should note that DFT+D approaches provide improvements only for the energetics, whereas all the other properties of the complexes (for example dipole moments or polarizabilities) remain the same as given by the DFT methods used. In our study, we have included the following DFT+D methods: dIDF+D [40], B3LYP+D3 [41], LC- ω PBE+D3 [42], LC-BOP12+LDR [43, 44], and LCgau-BOP+LDR [44, 45].

1. dIDF+D

The dispersionless density functional (dIDF) [40] is a hybrid meta GGA functional with parameters optimized on intermolecular interaction energies with the dispersion and exchange-dispersion energies subtracted. The total interaction energies were computed for a set of dimers using the CCSD(T) method and the dispersion and

exchange-dispersion energies were computed using SAPT based on DFT description of monomers [SAPT(DFT)] [46–49] (see a description of the latter method below). In contrast to the functionals of Sec. II A, dIDF includes only the components of intermolecular interactions that can be described by (semi)local DFT methods. The dIDF functional form is analogous to that of the M05-2X functional [50], i.e., it uses the PBE exchange density multiplied by a meta enhancement factor for the exchange part and a modified meta-enhanced LDA form for the correlation part.

Due to the way dIDF was constructed, one can add to the dIDF interaction energies accurate dispersion and exchange-dispersion energies computed using SAPT(DFT), and this approach leads to accurate total interaction energies for all intermonomer separations [40]. However, since this post-KS approach is more time consuming than DFT calculations, one can alternatively add an atom-atom function fitted to the SAPT(DFT) dispersion and exchange-dispersion energies: such functions have been developed in Refs. 35 and 40. The functions from Ref. 35 used in the present test have the form:

$$E_{\text{disp}} = - \sum_{a \in A, b \in B} \sum_{n=6,8} f_n(\beta_{ab} r_{ab}) \frac{C_n^{ab}}{r_{ab}^n} \quad (3)$$

where the Tang-Toennies [51] damping functions $f_n(r)$ are defined as

$$f_n(r) = 1 - \exp(-r) \sum_{k=0}^n \frac{r^k}{k!}, \quad (4)$$

and the interatomic constants are expressed in terms of atomic ones using the following combination rules: $\beta_{ab} = \sqrt{\beta_a \beta_b}$ and $C_n^{ab} = \sqrt{C_n^a C_n^b}$. The damping parameters are the same for $n = 6$ and 8 . All the parameters in this expression were fitted only to the sum of the dispersion and exchange-dispersion energies, so that these functions are DFT-independent and can be added to any dispersion-free interaction energy. Notice that the dIDF+D method, in contrast to most of the other approaches considered here, involves no fitting to the total interaction energies. The concept of dispersionless interaction energies was recently developed in alternative ways by Rajchel *et al.* [52] and Austin *et al.* [53].

2. B3LYP-D3 and LC- ω PBE-D3

The B3LYP-D3 [41] and LC- ω PBE-D3 [42] approaches use the standard DFT functionals for the DFT part and add the D3 atom-atom dispersion function from Ref. 41 with different dispersion switching parameters. The hybrid GGA B3LYP functional [54, 55] consists of the B88 Becke’s exchange functional [56] of 1988 and of the Lee, Yang, and Parr (LYP) [57] correlation functional. B3LYP contains three parameters (indicated by the numeral 3 in the acronym) fitted to a set of atomization

energies, ionization potentials, proton affinities, and total atomic energies.

The LC- ω PBE functional [58], where LC stands for “long-range corrected”, is an RSH version of the PBE functional. Its exchange-correlation functional is a sum of the PBE correlation functional, of the short-range [attenuated according to Eq. (1)] PBE exchange functional, and of the long-range exact exchange. The former exchange component used the expression developed by Heyd *et al.* [29, 59] based on the exchange hole model. Note that despite the use of this expression, LC- ω PBE is a functional which is completely different from the functionals of Ref. [29, 59] which use a certain amount of exact exchange only at short range and therefore have a wrong asymptotic behavior of the exchange potential. LC- ω PBE is actually similar in spirit to the long-range corrected functionals of Hirao and coworkers [28]. These relations explain the origin of the LC in the LC- ω PBE acronym, which may seem redundant since the use of range separation alone, indicated by the parameter ω , assures long-range correctness for exchange functionals of this type. Note that whereas RSH functionals lead to correct asymptotics of exchange-correlation potentials, the more important electron densities produced by such functionals still have wrong asymptotics with the standard values of ω [60]. LC- ω PBE was used here with the recommended value of $\omega = 0.4$ bohr⁻¹ optimized for a broad set of thermochemical data.

The D3 dispersion energy [41] is of similar character but different in several details from that defined by Eqs. (3) and (4). First, the following quantity is computed for the dimer AB and the monomers A and B:

$$E_X^{(2)} = - \sum_{a < b} \sum_{n=6,8} s_n f_n(r_{ab}) \frac{C_n^{ab}}{r_{ab}^n} \quad (5)$$

(D3 can also include a third-order term which has not been used here) where the switching function is of the form

$$f_n(r_{ab}) = 1 / \left[1 + 6 \left(\frac{t_n r_{ab}^0}{r_{ab}} \right)^{\alpha_n} \right]. \quad (6)$$

The summation here is performed over all pairs of atoms in a system. The dispersion energy is then calculated using the supermolecular approach

$$E_{\text{disp}} = E_{AB}^{(2)} - E_A^{(2)} - E_B^{(2)}.$$

Note that all terms coming from intramonomer atom-atom pairs cancel in this expression, so that one could equivalently sum only over intermonomer pairs to obtain the dispersion energy directly. The scaling parameters s_n and t_n (the latter denoted as $s_{r,n}$ in the original paper) were chosen to minimize deviations of the predictions from benchmarks that include both intermolecular interaction energies and thermochemical data. The value of s_6 and t_8 are set to 1.0 for both B3LYP and LC- ω PBE functionals. Thus, only two scaling parameters,

s_8 and t_6 are optimized. The constants r_{ab}^0 are distances where atom-atom Heitler-London energies computed using Slater determinants built from KS orbitals reach a prescribed value. Thus, the product $t_n r_{ab}^0$ defines the value of atom-atom separation where the switching function changes rapidly from zero to one. We use the name “switching” rather than “damping” function for the f_n of Eq. (6) to emphasize that the main role of this functions is not to damp the asymptotic inverse power expansion of dispersion energy in order to account for the charge-overlap effects at finite intermonomer distances R , as done by the Tang-Toennies function, but rather to completely switch off the dispersion energies at some chosen distance (the dispersion energies damped by the functions of Eq. (4) remain finite even at $R = 0$). Note furthermore that whereas the function of Eq. (3) contains summation only over the pairs of atoms from different monomers, the function of Eq. (5) sums over all pairs of atoms in a dimer and in the corresponding monomers. Then the intermolecular dispersion energy is obtained by the supermolecular approach. The powers in Eq. (6) were chosen as $\alpha_6 = 14$ and $\alpha_8 = 16$.

The reference coefficients, $C_{6,\text{ref},mk}^{ab}$, for a pair of atoms $a = X$ and $b = Y$ in their hydrides $X_m H_n$ and $Y_k H_l$ are computed from coupled KS frequency-dependent polarizabilities using the Casimir-Polder formula

$$C_{6,\text{ref},mk}^{ab} = \frac{3}{\pi} \int_0^\infty d\omega \frac{1}{m} \left[\alpha^{X_m H_n}(i\omega) - \frac{n}{2} \alpha^{H_2}(i\omega) \right] \times \frac{1}{k} \left[\alpha^{Y_k H_l}(i\omega) - \frac{l}{2} \alpha^{H_2}(i\omega) \right]. \quad (7)$$

Different hydrides give different coefficients for a given pair of atoms which corresponds to atoms in different coordination states (e.g., corresponding to sp^3 , sp^2 , sp hybridizations). For atoms X and Y in actual molecules, their fractional coordination numbers are computed from their environments and then the final value of C_6^{ab} is obtained as a weighted sum of the reference values from Eq. (7).

The B3LYP interaction energies were obtained using the NWChem software package [14] and the D3 correction was evaluated with the DF-TD3 V2.1 program [41]. The LC- ω PBE-D3 calculations were performed with Gaussian09 codes [61]. In both cases only the two-body, second-order dispersion energy was used.

3. LC-BOP+LRD and LCgau-BOP+LRD

The LC-BOP method [28] is an RSH approach which employs the B88 exchange functional and the one-parameter progressive (OP) correlation functional [62]. One should emphasize that the DFT-based short-range exchange energy in LC-BOP has a very different form than that used in the LC- ω PBE functional discussed above. The recently modified [43] form of this functional, LC-BOP12 was utilized in the present work. The LCgau-BOP functional [45] differs from LC-BOP by the use a

of a more elaborate range partitioning function which in addition to the terms present in Eq. (1) includes two mutually cancelling Gaussian terms, one incorporated into the DFT part and the other one into the exact exchange part.

A “local-response dispersion” (LRD) function of Sato and Nakai [63] is added to the LC-BOP12/LCgau-BOP interaction energies. The form of this function is the same as given by Eq. (3), except that the $n = 10$ term is also included. However, in contrast to previously discussed approaches, the coefficients C_n^{ab} (van der Waals constants) are calculated on-the-fly from electron densities. The starting point is the generalized Casimir-Polder expression for the exact second-order dispersion energy [64]

$$E_{\text{disp}}^{(2)} = -\frac{1}{2\pi} \int_0^\infty \iiint \iiint \alpha_A(\mathbf{r}_1, \mathbf{r}'_1 | i\omega) \alpha_B(\mathbf{r}_2, \mathbf{r}'_2 | i\omega) \times \frac{d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 d^3 \mathbf{r}'_1 d^3 \mathbf{r}'_2}{|\mathbf{r}_1 - \mathbf{r}_2| |\mathbf{r}'_1 - \mathbf{r}'_2|} du \quad (8)$$

where $\alpha_X(\mathbf{r}, \mathbf{r}' | i\omega)$ is the frequency-dependent density-density response function (called also the frequency-dependent density susceptibility) of monomer X . Such functions can be computed very accurately using the time-dependent DFT (TD-DFT) method. Therefore, the use of TD-DFT response functions gives very accurate dispersion energies, as shown for the first time in Refs. 46 and 47.

The first key simplifying assumption of Sato and Nakai was the use of an approximate expression for the response function, which is a modified version of the expression proposed by Dobson and Dinte [65]:

$$\alpha_X(\mathbf{r}, \mathbf{r}' | i\omega) = \nabla_{\mathbf{r}} \cdot \nabla_{\mathbf{r}'} \left[\frac{\rho(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')}{\omega_0^2(\mathbf{r}) + u^2} \right] \quad (9)$$

where $\rho(\mathbf{r})$ is the density of system X . The quantity ω_0 is a parametrized function of the density and of the magnitude of its gradient

$$\omega_0(\mathbf{r}) = \frac{1}{3} k_F^2 [1 + \lambda s^2]^2, \quad (10)$$

where $k_F(\mathbf{r}) = (3\pi^2 \rho(\mathbf{r}))^{1/3}$ and $s(\mathbf{r}) = |\nabla \rho(\mathbf{r})| / (2k_F(\mathbf{r}) \rho(\mathbf{r}))$. The parameter λ was empirically optimized as described below.

The second simplifying assumption of Sato and Nakai was to expand the interelectronic distances in Eq. (8) in the multipole series, leading to the asymptotic expansion of the dispersion energy. The volume integrals give then frequency-dependent multipole polarizabilities. The subsequent integration over the frequency gives the van der Waals constants. In the simplest case, this expansion is a series in inverse powers of the distance between centers of mass of monomers, but one can also formulate the so-called distributed expansion where all the intermonomer atom-atom distances are used. To this end, one partitions the response function into atomic

domains and compute distributed polarizabilities. Sato and Nakai achieved such distribution using numerical integration grids centered around individual atoms. Recently it has been shown [66, 67] that one can formulate a faster-converging distributed expansion of dispersion energy which almost exactly reproduces the SAPT(DFT) dispersion energy in all regions where the overlap effects can be neglected.

The parameter λ in Eq. (10) has been originally optimized to reproduce the empirical C_6 constants for pairs of rare-gas atoms. It was recently reoptimized [44] for 1225 known empirical C_6 constants for atoms and small molecules.

To damp the asymptotic expansion at short intermonomer separations, Sato and Nakai used the following damping function:

$$f_n(r) = \exp \left[-\frac{n-4}{2} \left(\frac{r_0}{r} \right)^6 \right] \quad (11)$$

where r_0 is a constant specific for each pair of atoms and dependent on two parameters κ and R_0 :

$$r_0 = \kappa \left[\alpha_a^{1/3} + \alpha_b^{1/3} \right] + R_0 \quad (12)$$

where α_x is the average static dipole-dipole polarizability for atom x computed using the response function of Eq. (9). In the original work, these parameters were optimized to reproduce the empirical minima of dimers of rare-gas atoms. In the calculations presented here, the values recently reoptimized [44] on the so-called S66 set [68] of separation-dependent benchmark interaction energies of 22 dimers.

C. Nonlocal functionals

Our survey included one nonlocal functional, vdW-DF2 [69], which is a modified version of the earlier vdW-DF functional [70]. The vdW-DF2 functional adds to a standard DFT functional a nonlocal term dependent on the density and the magnitude of its gradient, aimed at reproducing the dispersion energy. Thus, vdW-DF2 is in this respect similar to the DFT+D approach. However, the nonlocal term includes explicit overlap effects, as in SAPT(DFT). One has to properly select the standard DFT functional used in the vdW-DF method and the choice was based on similarities of predictions to those given by the HF approach. The original choice in vdW-DF was the revPBE functional [71], a reparametrized version of PBE. In vdW-DF2, the PW86R functional was chosen. This functional is a refitted [72] version of the PW86 functional with the exchange part from Ref. 73 and the correlation part (P86) from Ref. 74. PW86 was chosen based on good agreement of its exchange-only version with HF interaction energies for several dimers. The refit concerned the parameters in the exchange enhancement factor and was done to make this factor have the

exact large- s and small- s behavior, which was not the case for PW86. The agreement of PW86 and PW86R exchange-only interaction energies with HF interaction energies is about the same.

The nonlocal term has been derived by making approximations to the response function in the spirit of Eq. (9). The resulting expression reads

$$E_c^{\text{nl}} = \iint d^3\mathbf{r} d^3\mathbf{r}' \rho(\mathbf{r})\rho(\mathbf{r}')\phi(\mathbf{r}, \mathbf{r}') \quad (13)$$

Here $\rho(\mathbf{r})$ is the density of the dimer, so that vdW-DF methods include also the nonlocal correlation energies within monomers. The kernel $\phi(\mathbf{r}, \mathbf{r}')$ is given by the expression

$$\phi(\mathbf{r}, \mathbf{r}') = \frac{2}{\pi^2} \int_0^\infty da \int_0^\infty db \frac{a^2 b^2 W(a, b) T(\nu(a), \nu(b), \nu'(a), \nu'(b))}{a^2 b^2 W(a, b) T(\nu(a), \nu(b), \nu'(a), \nu'(b))}. \quad (14)$$

where W and T are relatively simple functions of their arguments, given by Eqs. (15) and (16) in Ref. 70: W is a linear combination of products of sines and cosines of the variables a and b (so W does not depend on density), whereas T is a sum of terms containing inverse powers of the binomials of the arguments. The arguments of T are defined as

$$\nu(c) = \frac{c^2}{2 \left(1 - \exp \left[-\frac{4\pi c^2}{9q_0^2(\mathbf{r})|\mathbf{r}-\mathbf{r}'|^2} \right] \right)} \quad (15)$$

and ν' differs from ν by $q_0^2(\mathbf{r}) \Rightarrow q_0^2(\mathbf{r}')$. The quantity q_0 is given by

$$q_0(\mathbf{r}) = \frac{\epsilon_{\text{xc}}^0(\mathbf{r})}{\epsilon_{\text{x}}^{\text{LDA}}(\mathbf{r})} k_{\text{F}}(\mathbf{r}) \quad (16)$$

where $\epsilon_{\text{x}}^{\text{LDA}}(\mathbf{r}) = -3k_{\text{F}}/4\pi$ and

$$\epsilon_{\text{xc}}^0(\mathbf{r}) = \epsilon_{\text{c}}^{\text{LDA}}(\mathbf{r}) + \epsilon_{\text{x}}^{\text{LDA}}(\mathbf{r}) \left(1 - \frac{Z_{\text{ab}} s^2}{9} \right). \quad (17)$$

The LDA correlation energy density, $\epsilon_{\text{c}}^{\text{LDA}}(\mathbf{r})$, is an approximation to this quantity used in the associated standard DFT functional. The parameter Z_{ab} has been chosen based on theoretical constraints arguments and its value of -0.8491 used in vdW-DF was changed to -1.887 in vdW-DF2.

D. Post Kohn-Sham methods

All the methods discussed so far use only electron densities, their gradients, and occupied KS orbitals. These methods scale therefore as N^4 with system size, where system size can be defined as the number of electrons (the methods that do not use the exact exchange can easily be programmed to scale as N^3). One way to improve the results is to make use of virtual KS orbitals, a

step towards WF theory. This usually makes the scaling worse, but if it is only increased to N^5 , such an approach can still be applied to very large systems. The post-KS terms have to be properly added to the DFT interaction energy to avoid double counting.

1. SAPT(DFT)

In SAPT [3, 75], the interaction energy is computed using a perturbative approach. Starting from the wave functions of unperturbed monomers, the intermolecular interaction energy can be evaluated order-by-order as an expansion in powers of intermolecular interaction operator V collecting Coulomb interactions of all electrons and nuclei of monomer A with those of monomer B. The resulting wave functions have to be properly antisymmetrized which explains the phrase “symmetry-adapted” in the name of this method and produces the so-called exchange components of interaction energies such as the exchange-dispersion term mentioned earlier. The SAPT interaction energy is naturally and rigorously decomposed into electrostatic, exchange, induction, and dispersion contributions. Since the exact wave functions of monomers are unknown, the starting point in standard SAPT are actually products of HF determinants of monomers. These wave function are corrected by using MBPT/CC expansion in powers of the intramonomer correlation operators W_X , i.e., the so-called Møller-Plesset (MP) fluctuation potentials which are differences between the exact Hamiltonian of a monomer and the Fock Hamiltonian. The HF-based SAPT properly truncated at some power of W_A gives interaction energies of similar accuracy as those produced by the CCSD(T) method, but also scales in the same way (N^7) and therefore is similarly time-consuming. A much better scaling method can be obtained using a DFT description of the monomers. Since all components of interaction energies are computed as a post-KS process, there is by definition no double counting. An important ingredient of SAPT(DFT) is the use of asymptotically corrected exchange-correlation potentials [60] giving electron densities with correct long-range behavior. The dispersion energies are computed from Eq. (8) using TD-DFT response functions without any approximations. SAPT(DFT) [46–49, 76–78] utilizing density fitting techniques [46, 49, 79] scales as N^5 . If a nonhybrid DFT is used to describe the monomers and some small exchange terms are neglected, SAPT(DFT) scales in practice as N^4 , i.e., as a hybrid DFT, and can be applied to systems with hundreds of atoms [80] [one step, the integral transformation, scales formally as N^5 , but since this scaling is proportional to $o^2 v^2 N_{\text{aux}}$, where $o(v)$ is the number of occupied (virtual) orbitals and N_{aux} the number of auxiliary functions, the prefactor of this term is small].

SAPT(DFT) is usually applied at the second-order in V , but for systems with large polarization effects the po-

larization terms of higher order are approximated by the so-called $\delta_{\text{int}}^{\text{HF}}$ correction which is the difference between the supermolecular Hartree-Fock interaction energy and the sum of the first- and second-order SAPT(HF) terms that are accounted for by the HF description of the system. The criterion for including $\delta_{\text{int}}^{\text{HF}}$ was the ratio of the sum of the induction and exchange-induction energies to the total interaction energy for the distance where the interaction energy was the most negative. If this ratio was larger than 12.5%, $\delta_{\text{int}}^{\text{HF}}$ was included. By this criterion, it was included for all systems except those containing benzene and nitrobenzene. The threshold value was determined by analyzing systems studied by SAPT(DFT) in the past. It turned out in the post analysis that this criterion was not working well for the nitrobenzene dimer. Whereas for this system the large dipole moment of 2.1 a.u. does not result in a significant induction contribution near the minimum, the $\delta_{\text{int}}^{\text{HF}}$ term is very important for large intermolecular separations where it significantly improves the agreement with the benchmarks. However, in order not to violate the ‘blind’ character of this test, $\delta_{\text{int}}^{\text{HF}}$ was not used for the nitrobenzene dimer. In future applications of SAPT(DFT), the criterion for adding $\delta_{\text{int}}^{\text{HF}}$ should also take into account the dipole moments of the monomers.

The DFT calculations for monomers have been performed using the Tozer-Handy-Fermi-Amaldi asymptotic correction [81, 82]. The electrostatic energies were computed in quadruple precision arithmetics [83] and using JK-optimized auxiliary bases [84]. For all other terms, the correlation-optimized bases [85] were used.

2. RSH+lrMP2 and RSH+lrRPAx-SO2

As discussed earlier, the RSH methods mix DFT and exact (HF) exchange in a ratio dependent on the interelectronic distance, using the DFT exchange at short range (sr) and exact exchange at long range (lr). It is also possible to mix the correlation part in the same way. However, the long-range correlation part has to be obtained from some WF-type approach. The simplest such method is the second-order MBPT with the MP partition of the Hamiltonian, known under the name MP2. The MP2 expression in RSH methods is formally identical to that of the regular MBPT, except that it is computed with attenuated interelectronic interactions and with KS orbitals and orbital energies. A method of this type, dubbed RSH+lrMP2 was developed in Ref. [86]. In the first implementation, this method used the short-range LDA exchange-correlation functional of Ref [87]. In the present work, the short-range PBE exchange-correlation functional of Goll *et al.* [17] (which a modification of the functional of Ref. [88]) is used. Note that the short-range PBE exchange part of this functional is different from the one of Refs [29, 59] used in LC- ω PBE.

The next step beyond MP2 is the random-phase approximation (RPA) approach. RPA may be derived from

the CCD method by selecting from the CCD expression only terms including products of two-electron integrals that can be illustrated by the so-called ring diagrams (the simplest ring diagram represents MP2) [89]. The RPA amplitudes are obtained from equations of a similar type as used in calculations of response functions. There are several versions of RPA available. The method selected in our survey, RSH+lrRPax-SO2 [90], uses the Hartree kernel plus the Hartree-Fock exchange kernel in the equations for amplitudes. However, matrix elements multiplying these amplitudes are not antisymmetrized. This is variant 2 of a method introduced by Szabo and Ostlund [91], which is indicated by SO2 in the acronym. As for RSH+lrMP2, the calculations were done with the short-range PBE exchange-correlation functional of Goll *et al.* [17].

III. BASIS SETS

A. Choices of basis sets in DFT and post-KS calculations

Ideally, in order to have the most unambiguous comparisons, all calculations should be performed using CBS extrapolations from large basis sets. This approach, however, would be too time consuming for some methods. Furthermore, some methods have been optimized using specific basis sets and perform best if used with the same basis sets. Finally, DFT calculations are known to converge faster with respect to basis set size than post-KS ones. The organizers of the blind test suggested that the augmented triple-zeta basis set of Dunning and coworkers [92, 93], denoted as aug-cc-pVTZ, be used and most participants followed this suggestion. This basis set obviously could not be used in codes based on the plane-wave functions. The other basis sets used were mostly different members of the aug-cc-pVXZ family with $X = 2$ (marked by D), 4 (Q), and 5. The jun-cc-pVTZ basis set is the aug-cc-pVTZ basis with omitted diffuse f functions on all atoms and without any diffuse functions on hydrogens. We will discuss possible uncertainties resulting from the use of different basis sets in Sec. VD. SAPT(DFT) calculations used the so-called dimer-centered plus midbond basis set format [94].

B. Basis set superposition error

Interaction energies computed using the supermolecular approach, i.e., by subtracting the total monomer energies from the dimer’s total energy, suffer the so-called basis set superposition error (BSSE) resulting from the lowering of monomer energies in dimer calculations due to “borrowing” of the basis set from the interacting partner. One way to reduce this error is to compute such energies in the counterpoise (CP) corrected way [95], i.e., by performing calculations for monomers in the full dimer’s

TABLE II. Basis set convergence of calculations for benzene-methane at $R = 3.28, 3.8,$ and 8.8 \AA . Energies are in kcal/mol and aXZ denotes aug-cc-pVXZ+mb. All quantities in parentheses are obtained from extrapolation formulas as described in the text.

$R = 3.28$	aDZ	aTZ	aQZ	a5Z	CBS
$E_{\text{int}}^{\text{HF}}$	5.105	5.105	5.101	5.099	(5.099)
$\delta E_{\text{int}}^{\text{MP2}}$	-5.562	-5.761	-5.845	-5.882	(-5.921)
$\delta E_{\text{int}}^{\text{CCSD(T)}}$	0.725	0.775	(0.787) ^a	(0.791) ^a	(0.796) ^b
E_{int}^c	0.268	0.119	(0.043)	(0.009)	(-0.025)
$E_{\text{int}}^{\text{CBS } d}$					(-0.047)
$R = 3.8$					
$E_{\text{int}}^{\text{HF}}$	0.885	0.897	0.898	0.898	(0.897)
$\delta E_{\text{int}}^{\text{MP2}}$	-2.599	-2.663	-2.688	-2.699	(-2.711)
$\delta E_{\text{int}}^{\text{CCSD(T)}}$	0.310	0.337	(0.344) ^a	(0.346) ^a	(0.349) ^b
E_{int}^c	-1.403	-1.428	(-1.446)	(-1.455)	(-1.465)
$E_{\text{int}}^{\text{CBS } d}$					(-1.476)
$R = 8.8$					
$E_{\text{int}}^{\text{HF}}$	-0.0056	-0.0052	-0.0052	-0.0051	(-0.0051)
$\delta E_{\text{int}}^{\text{MP2}}$	-0.0145	-0.0149	-0.0150	-0.0149	(-0.0149)
$\delta E_{\text{int}}^{\text{CCSD(T)}}$	0.0026	0.0025	(0.0025) ^a	(0.0025) ^a	(0.0025) ^b
E_{int}^c	-0.0175	-0.0176	(-0.0177)	(-0.0176)	(-0.0176)
$E_{\text{int}}^{\text{CBS } d}$					(-0.0175)

^a Values estimated from formula (20) applied to $\delta E_{\text{int}}^{\text{CCSD(T)}}$ with the CBS energy and the constant B obtained from the (D,T) extrapolation.

^b (D,T) extrapolation analogous to formula (20).

^c Sum of values in preceding rows.

^d Equation (18).

basis set. The organizers of the blind test suggested that the CP correction will be used in all calculations, but some participants have chosen not to use it. Note that codes using plane-wave basis sets have to perform all calculations in the same basis set, so that the CP correction is an inherent feature of such methods. The SAPT approach is by definition free of BSSE since the interaction energy is computed directly, i.e., it does not involve any subtraction of dimer and monomer energies.

IV. BENCHMARK INTERACTION ENERGIES

A. Basis sets and extrapolations

The DFT-based methods described in Sec. II were assessed by comparisons to a set of benchmark interaction energies computed in the same way as in Ref. 15:

$$E_{\text{int}}^{\text{CBS}} = E_{\text{int}}^{\text{HF}}(\text{CBS}) + \delta E_{\text{int}}^{\text{MP2}}(\text{CBS}) + \delta E_{\text{int}}^{\text{CCSD(T)}}, \quad (18)$$

where $E_{\text{int}}^{\text{HF}}(\text{CBS})$ is the CBS-extrapolated HF interaction energy, $\delta E_{\text{int}}^{\text{MP2}}(\text{CBS})$ is the CBS-extrapolated dif-

ference between the MP2 and HF interaction energies, and $\delta E_{\text{int}}^{\text{CCSD(T)}}$ is the difference between CCSD(T) and MP2 interaction energies, computed in the same basis set, generally smaller than the bases used for the HF and MP2 extrapolations. All correlated calculations employed the frozen-core approximation. The HF and MP2 interaction energies were extrapolated to the CBS limit using two-point extrapolations, exponential in the HF case

$$E_{\text{int}}^{\text{HF}}(X) = E_{\text{int}}^{\text{HF}}(\text{CBS}) + Ae^{-\alpha X}, \quad (19)$$

with α fixed at 1.63 as recommended in Ref. 96, and inverse third power in the case of MP2

$$\delta E_{\text{int}}^{\text{MP2}}(X) = \delta E_{\text{int}}^{\text{MP2}}(\text{CBS}) + BX^{-3}, \quad (20)$$

where “(X)” denotes that a given energy was obtained using the basis aug-cc-pVXZ (plus midbond functions, see below). The CBS values were obtained from calculations in $X-1$ and X bases. For all systems, extrapolations were performed with $(X-1, X) = (4, 5)$ except for the water dimer, for which we used (5,6) and the EDNA dimer for which we used (3,4). The $\delta E_{\text{int}}^{\text{CCSD(T)}}$ contribution was obtained using the aug-cc-pVTZ basis for all systems except for the water dimer for which the aug-cc-pVQZ basis was used and for the EDNA dimer for which aug-cc-pVDZ was used. For each configuration, to improve basis set convergence, an additional set of $3s$ ($\alpha=0.9, 0.3, 0.1$), $3p$ ($\alpha=0.9, 0.3, 0.1$), $2d$ ($\alpha=0.6, 0.2$), and $2f$ ($\alpha=0.6, 0.2$) midbond (mb) functions was used to augment the principal basis, with the location of the midbond functions determined using the algorithm of Ref. 97. The advantages of using midbond functions in conjunction with CBS extrapolations were demonstrated in Ref. 98. The NWChem software package [14] was used to obtain all benchmark interaction energies except for some MP2 energies in large basis sets, see below.

For larger monomers, our largest basis sets become linearly dependent to the point that NWChem removes a number of linear combinations of atomic orbitals from the basis set. With the default threshold of 10^{-5} for the eigenvalues of the overlap matrix, it removes 23 (47) linear combinations from the benzene-water aug-cc-pVQZ+mb (aug-cc-pV5Z+mb) basis set at $R = 7$ Å. Since our calculations are performed in the same basis set for the dimer and both monomers, and this reduction of basis set size is identical in all three cases, the counterpoise procedure is still rigorously imposed. On the other hand, the CBS extrapolations may be affected. Worse, the interaction energies can still be inaccurate despite the reduction and we found this to be the case for NWChem with the threshold given above. As in the case of the triple excitation error in CCSD(T), we found this problem by comparisons with SAPT(DFT) results which are very reliable at large separations. We first tried to lower the threshold to 10^{-6} , but it resulted in even more erratic results and self-consistent field (SCF) iterations convergence problems in some cases (even with

the 10^{-5} threshold for the overlap matrix, the iterations would sometimes not converge if a tight SCF convergence threshold was used). We then repeated the calculations of MP2 energies using the GAMESS [99] and Gaussian [61] packages and found that the results are consistent and appear to be sufficiently accurate (probably to at least about 3 significant digits for the interaction energies, based on the level of consistency). Also the CBS extrapolations were consistent (although this could be due to the same number of linear combinations removed). Therefore, we repeated all the MP2 calculations using GAMESS.

The CCSD(T) calculations (NWChem) were run on a Cray XE6 (32 cores/node, 60 GB/node) at the U.S. Army Engineer Research and Development Center (ERDC) using 512 cores for all systems except the 3 largest dimers (FOX-7, EDNA, nitrobenzene) which used 4,096 cores. For the small/intermediate sized systems most calculations completed in 12-36 hours (wall clock) with the total time being a function of the size of the system, as well as the center of mass (COM) separation R , with configurations at large R completing faster due to the significantly reduced number of non-negligible intermonomer integrals remaining at large separation. The CCSD calculations for the large systems completed in 4-6 days (wall clock) on average and due to the 7-day maximum runtime queue limits at ERDC, the (T) energy was evaluated using NWChem’s restart option. The SCF convergence tolerance for all calculations was equal to 10^{-10} a.u. where the threshold corresponds to the magnitude of the norm of the orbital gradient in NWChem’s quadratically convergent SCF implementation. The CCSD convergence tolerance was equal to 10^{-8} hartree and by default, NWChem uses the same convergence tolerance for the CC energy and the root mean square (RMS) error of the amplitudes. For the three biggest dimers, this led to an unacceptably large number of CC iterations (> 45) thereby prolonging the total wall clock and CPU time of these already expensive calculations. However, it was observed that although the energy converged fairly quickly (15-20 iterations), the remaining iterations were spent reducing the amplitude error to an equivalent level with little to no change observed in the energy. Therefore, for the large systems, we modified NWChem so that it only used the energy convergence criterion, and not that of the amplitudes, to terminate the CCSD iterations. Although not used as a criterion for terminating the CC iterations, the root mean square error of the amplitudes was still monitored and for all cases the final RMS error was on the order of 10^{-6} or below.

The large basis MP2 calculations (using GAMESS) were run on a Cray XC40 (32 cores/node, 126 GB/node) at the U.S. Army Research Laboratory’s Defense Shared Resource Center using 2,048 cores for all dimers except EDNA which used 5,120 cores. The SCF energy was converged to 10^{-10} hartree and required 1-8 hours (wall clock) time per job, again depending on the system size and COM separation. For three of the required

energy components (MP2/aQZ energy of nitromethane monomer A at $R=4.131$ Å, MP2/aQZ energy of the methylformate dimer at $R=4.243$ Å, and MP2/a5Z energy of benzene-methane $R=8.80$ Å), the GAMESS SCF was not convergent and these energies were computed using Gaussian on the same Cray XC40 using 32 to 256 cores. In total, the computation of the benchmark energies required approximately 40 million CPU hours with most of that time resulting from the required recalculation of the benchmarks due to the issues described above.

B. Interaction Energy Curves

The equilibrium geometries of the dimers were taken from literature, when available, or optimized at the MP2/aug-cc-pVTZ level of theory. Complete geometry optimizations were performed, so that the two monomers have slightly different geometries in homogeneous dimers. The optimizations were not CP corrected. All the geometries and references to the appropriate literature sources are given in the Supplementary Material [16]. For each system, given the optimized configuration, radial potential energy curves were obtained by variation of the radial separation R between the COMs of the two monomers and keeping the mutual orientations and intramonomer coordinates unchanged. The interaction energies computed are the vertical ones, i.e., the geometry of each monomer in calculations for this monomer is exactly the same as in the dimer calculations. All interaction energies from the DFT-based methods were obtained for the same configurations as used to compute the benchmark energies, i.e., the geometries of the dimers were not optimized using these methods. The number of points per dimer ranged from six to sixteen (depending on the size of the molecule) with geometries chosen to sample the repulsive wall, potential well, and the asymptotic region. The entire reference data set contains a total of 80 points.

V. RESULTS

A. Benchmarks

All the results of calculations leading to our benchmarks are listed in the Supplementary Material [16]. These tables show that the CBS extrapolations only modestly influence our results: the median unsigned percentage error of the nonextrapolated CCSD(T) interaction energies computed in the applied basis sets (aug-cc-pVTZ+mb for all but two systems) with respect to the CBS interaction energies is 1.04%. The reason is our use of the bond functions which makes the results computed in the aug-cc-pVTZ+mb basis quite well converged. However, there are a number of percentage deviations of the CCSD(T)/aug-cc-pVTZ+mb results from the CBS ones, all occurring for the shortest distances, that are very large in magnitude, up to 357% (for the

benzene-methane dimer at $R = 3.28$ Å). Despite such a large discrepancy, we believe that the CBS results at these separations are still reliable, as can be seen from the following analysis. The large percentage errors are due a combination of the fact that $\delta E_{\text{int}}^{\text{CCSD(T)}}$ is quite large for these dimers and that the interaction energy at these distances, near the points where the potential curves cross zero, are small differences of large numbers. Table II illustrates these relations for the benzene-methane dimer. The 357% error quoted above is the for the difference between $E_{\text{int}}^{\text{CCSD(T)}}$ computed in the basis aug-cc-pVTZ+mb, equal to 0.119 kcal/mol, and $E_{\text{int}}^{\text{CBS}}$ equal to -0.047 kcal/mol. The MP2/CBS energy is still quite different: -0.822 kcal/mol. Although such dramatic differences may suggest a low reliability of our benchmarks, an analysis of Table II shows that this is not the case. Clearly, the observed behaviour results from the fact that $E_{\text{int}}^{\text{HF}}$ and $\delta E_{\text{int}}^{\text{MP2}}$, the former converged to better than 0.001 kcal/mol and the latter to about 0.04 kcal/mol, cancel to a large extent. Since $\delta E_{\text{int}}^{\text{CCSD(T)}}$ is close to the MP2 interaction energy but of opposite sign, adding it results in a near zero interaction energy, which obviously is then very sensitive in relative terms to basis set size and CBS extrapolations.

Additional support for the reliability of our benchmarks, even for those difficult cases, can be obtained from performing the (D,T) CBS extrapolation of $\delta E_{\text{int}}^{\text{CCSD(T)}}$. We do not use such extrapolations in our benchmarks since extrapolations from such low X are generally unreliable, but let us assume the (D,T) extrapolation works in this case. Table II shows that the extrapolation gives the value of 0.796 kcal/mol (in the CBS column) and increases $\delta E_{\text{int}}^{\text{CCSD(T)}}$ by 0.021 kcal/mol, so that the interaction energy including this extrapolation is -0.026 kcal/mol, 0.021 kcal/mol from our recommended value. Also the sequence of the interaction energies in the row E_{int} is seen to converge smoothly (since aug-cc-pVTZ+mb was the largest basis set that we could use at the CCSD(T) level, the $\delta E_{\text{int}}^{\text{CCSD(T)}}$ entries in columns aQZ and a5Z were obtained using the X^{-3} dependence with the constants calculated using the aug-cc-pVDZ+mb and aug-cc-pVTZ+mb bases). Whereas it is difficult to say whether the value of -0.025 kcal/mol that includes the (D,T) extrapolation of $\delta E_{\text{int}}^{\text{CCSD(T)}}$ is a better representation of the exact CCSD(T) value than $E_{\text{int}}^{\text{CBS}} = -0.047$ kcal/mol, it is clear from the convergence patterns that either of these results is more accurate than $E_{\text{int}}^{\text{CCSD(T)}}$ computed using aug-cc-pVTZ+mb and equal to 0.119 kcal/mol. Based on the analysis performed above, we can estimate the residual error of $E_{\text{int}}^{\text{CBS}}$ to be about 0.04 kcal/mol. Whereas it is a large error percentage wise, it is reasonably small in absolute terms.

The results for two other distances presented also in Table II show that in these cases there are no doubts about the adequate convergence of the interaction energies. At $R = 3.8$ Å, the residual error can be roughly estimated to amount to about 0.02 kcal/mol or about 1%

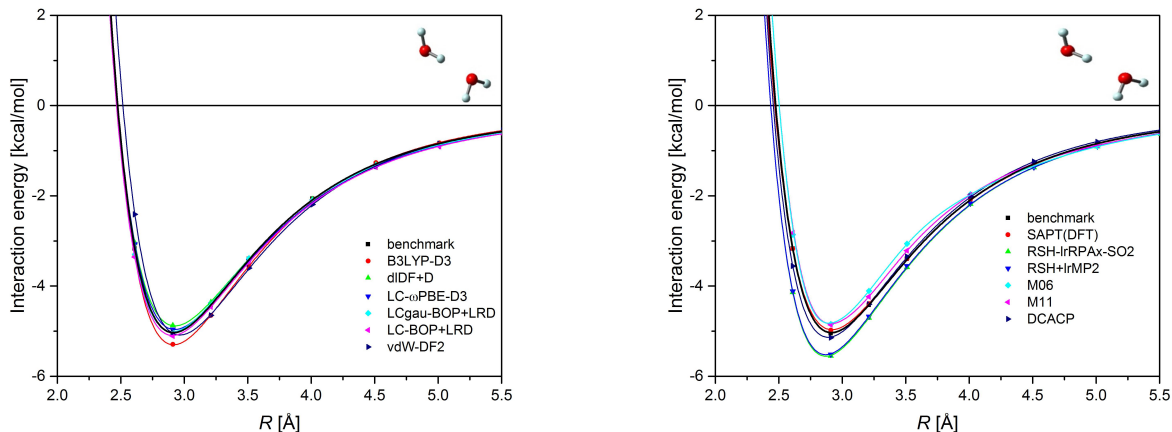


FIG. 1. Potential energy curves for the water dimer. The left panel compares the performance of DFT+D-type methods and of a nonlocal functional, whereas the right panel includes reoptimized standard functionals and methods using virtual orbital. Lines are drawn only to guide the eye.

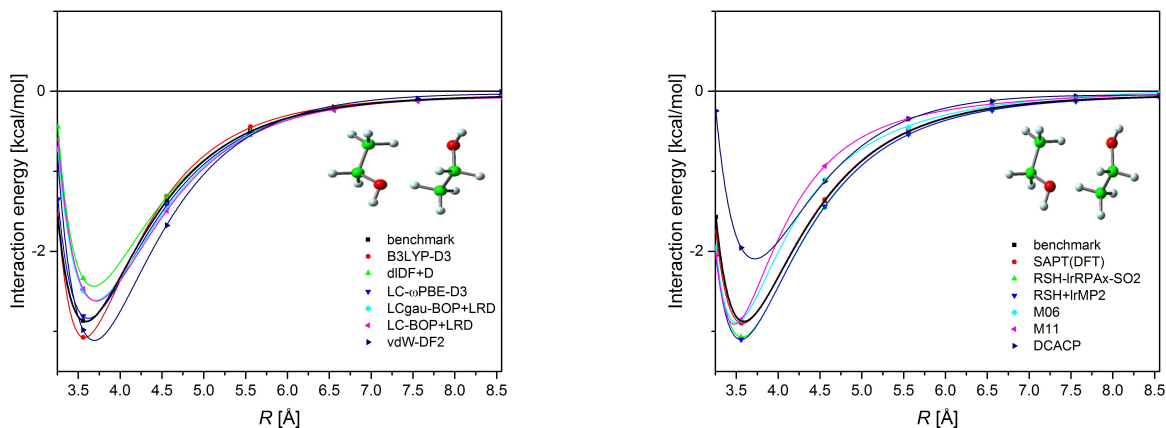


FIG. 2. Potential energy curves for the ethanol dimer. See Fig. 1 for explanations.

of interaction energy. This is the same relative accuracy as given in Ref. [15] for the set of benchmarks obtained there for systems involving only equilibrium dimer configurations. At $R = 8.8 \text{ \AA}$, the absolute error is smaller than 0.0001 kcal/mol and the relative one is smaller than 0.6% .

For our largest dimer, $(\text{EDNA})_2$ containing 32 atoms, we were able to use only the comparatively small aug-cc-pVDZ+mb basis in calculations of the $\delta E_{\text{int}}^{\text{CCSD(T)}}$ contribution. Table II shows that if this level of basis sets were applied to the benzene-methane dimer, the error relative to $E_{\text{int}}^{\text{CBS}}$ would be 0.15, 0.025, and 0.0001 kcal/mol for the three consecutive distances. The two latter errors are comparable to the estimated errors of $E_{\text{int}}^{\text{CBS}}$ for these distances, so the results will be less accurate by about a

factor of two. This decrease of accuracy makes a difference only for the best-performing DFT approaches. At $R = 3.28 \text{ \AA}$, the error increases about 4 times, so that at very small R the EDNA dimer benchmarks could be insufficiently accurate. However, the shortest R included for this system is well beyond the zero-crossing point, so that this issue does not become a problem.

B. Performance of DFT methods in different regions of potential energy curves

Plots of the potential energy surfaces for each complex are presented in Figs. 1 through 10 and a file listing all of the DFT and benchmark interaction energies is available

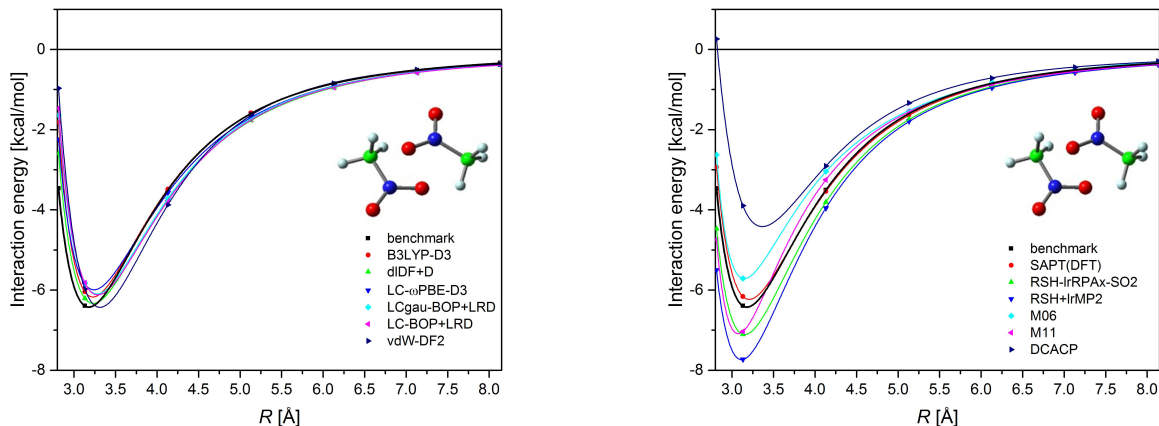


FIG. 3. Potential energy curves for the nitromethane dimer. See Fig. 1 for explanations.

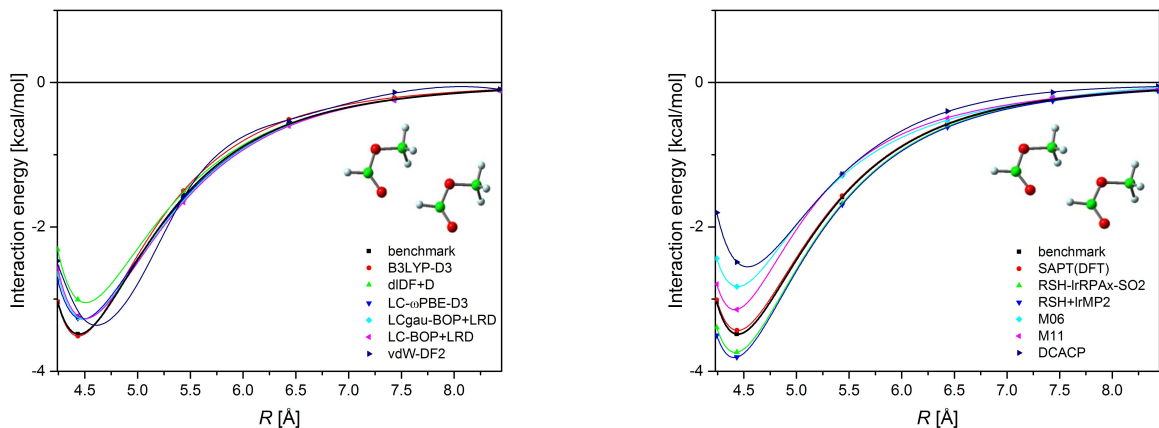


FIG. 4. Potential energy curves for the methylformate dimer. See Fig. 1 for explanations.

in the Supplementary Material [16]. In each figure, for clarity of presentation, the DFT methods were split into two groups with 6 methods in each group, each plotted in a separate panel. For a given system, both panels have the same scales on both axes, therefore curves residing in different panels can be compared directly.

Starting the comparisons from the asymptotic region, one can see on the figures that for most systems the majority of the methods seem perform very well. One reason is that all but two of our monomers have sizable dipole moments (in some cases very large: 3.6 a.u. for FOX-7 and 2.1 a.u. for nitrobenzene), so that the interaction energy decays like $1/R^3$. For benzene-water, the decay is still slow, as $1/R^4$, and only for benzene-methane the decay of the electrostatic energy is $1/R^6$, i.e., the same as that of the induction and dispersion energies. Since standard DFT methods give fairly accurate

multipole moments, one expects accurate reproduction of electrostatically dominated asymptotic decays. Another reason is just smallness of the interaction energy in this region: the typical percentage errors at the largest distances are actually not negligible and amount to a few percent for the systems with the $1/R^3$ decay. There are a couple of methods, DCACP, M06, M11, and vdW-DF2, which for several systems give still much larger errors, in the range 20%-230%. Since at these distances the electrostatic energy constitutes a major fractions of the interaction energy, this indicates that either the multipole moments of some monomer are not well reproduced by these approaches or there are some unphysical components appearing at these distances. For benzene-methane, the only dimer where electrostatics decays as fast as dispersion and the SAPT(DFT) energy decomposition shows that at large R the interaction energy consists of equal

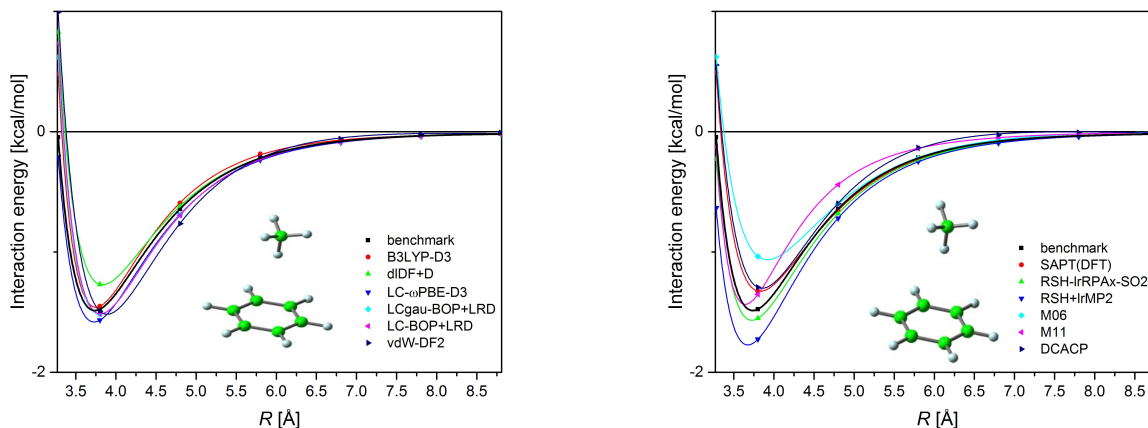


FIG. 5. Potential energy curves for the benzene-methane dimer. See Fig. 1 for explanations.

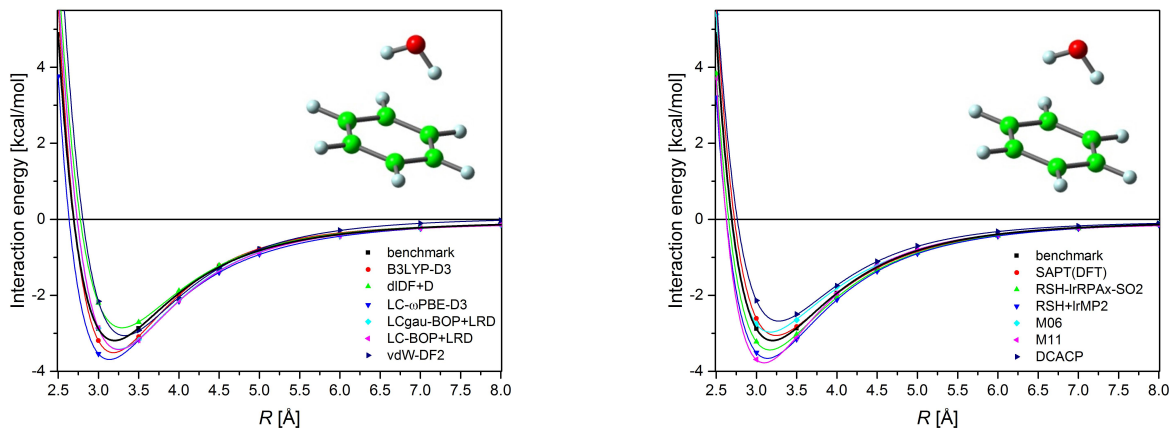


FIG. 6. Potential energy curves for the benzen-water dimer. See Fig. 1 for explanations.

fractions of these components, the four methods have errors between 52% and 77%, whereas the other methods perform overall similarly as for other systems. One should point out that while DCACP, M06, and M11 cannot produce the correct asymptotics of the dispersion energy, vdW-DF2 should be able to do so, so the large error of this method is unexpected.

Somewhat surprisingly, even for methods with potentially correct asymptotics, the relative unsigned errors often increase with R for large R . The observed behaviour may be due to unphysical effects in the DFT part resulting from the wrong decay of DFT densities. It is probably not due to insufficiently accurate C_6 coefficients, since for large enough R this error should be constant (but perhaps R is not large enough in our calculations).

Near the equilibria, the discrepancies visible in the figures are generally much larger than at large separations.

In particular, RSH+lrMP2 and RSH+lrRPax-SO2 systematically overbind, the former more strongly than the latter. In contrast, DCACP systematically underbinds. So does M06 and vdW-DF2, but to a much smaller extent. M11 is the only method which sometimes overbinds and sometimes underbinds. However, for each of the remaining methods there is a dimer or two where a given method clearly over or underbinds.

At the smallest R for each dimer, the deviations from the benchmarks are generally the largest. The main reason is that this point, for most systems, is near the region where the potential energy curves cross zero. Thus, the cancellations similar to those discussed in Sec. V A occur for most methods. The observed discrepancies are actually not that critical since these interaction energies are at the onset of the repulsive wall. As it is well seen on the example of the water dimer (Fig. 1), the walls produced

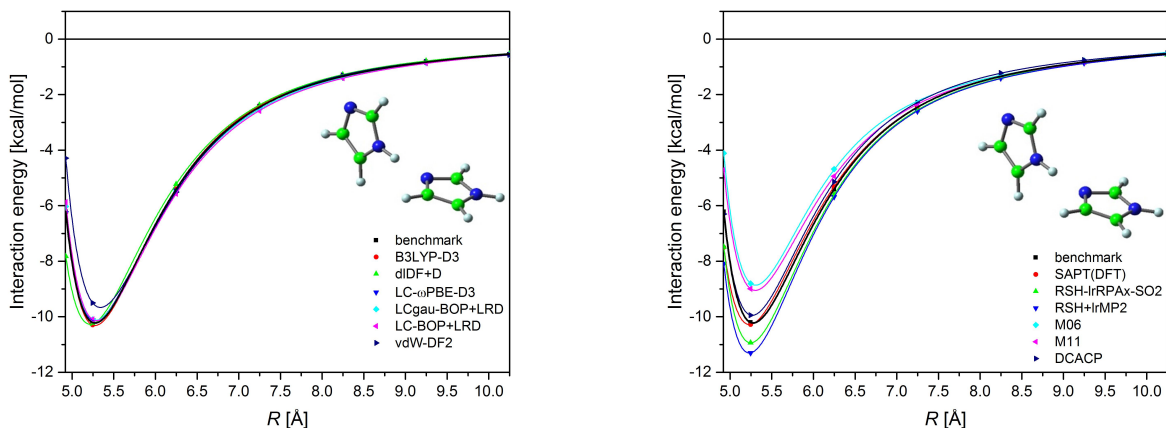


FIG. 7. Potential energy curves for the imidazole dimer. See Fig. 1 for explanations.

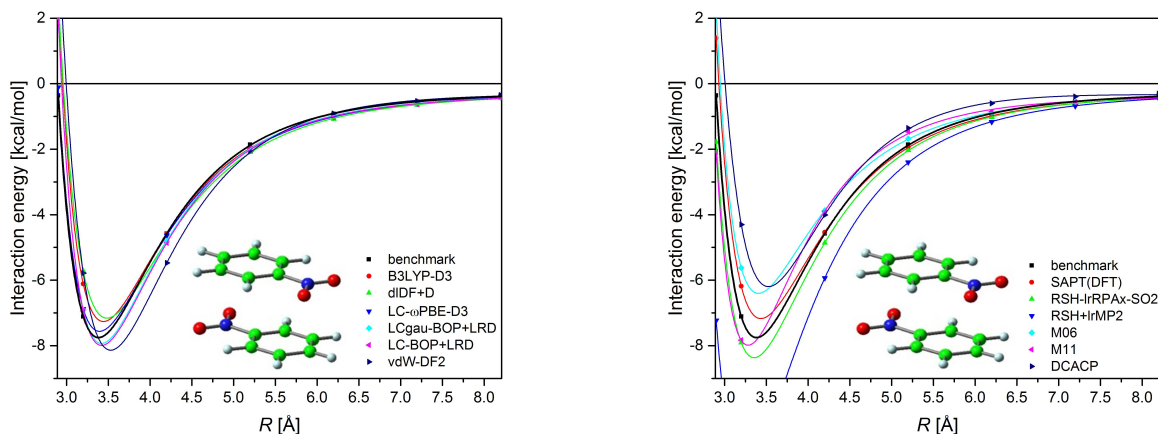


FIG. 8. Potential energy curves for the nitrobenzene dimer. See Fig. 1 for explanations.

TABLE III. Median unsigned percentage errors (MUPE) for individual dimers with respect to the benchmarks. The MP2 entry are the CBS extrapolated values obtained as described earlier.

	SAPT(DFT)	MP2	B3-D3 ^a	dIDF+D	LC-D3 ^b	gauBOP ^c	BOP ^d	vdw-DF2	M06	lrRPA ^e	M11	lrMP2 ^f	DCACP	Ave.
water dimer	1.38	1.04	1.61	1.39	4.44	3.28	6.23	5.56	4.48	7.03	5.32	7.63	7.06	4.34
ethanol dimer	1.94	1.59	3.96	2.31	4.49	6.75	10.10	16.68	22.65	8.22	29.64	9.18	39.12	12.05
nitromethane dimer	3.65	4.99	3.48	11.14	7.79	9.27	11.30	6.58	4.36	9.51	7.30	12.53	16.11	8.31
methylformate dimer	1.58	7.99	8.03	3.18	1.94	3.13	4.41	9.09	18.87	5.41	11.25	7.40	35.86	9.09
benzene-methane	3.47	13.93	7.84	6.16	9.11	7.73	8.23	28.02	29.77	8.87	40.98	16.40	67.74	19.10
benzene-water	4.38	3.31	4.90	5.56	11.09	6.95	8.08	25.03	7.54	7.65	5.17	9.96	18.26	9.07
imidazole dimer	0.70	4.38	1.20	4.11	2.66	3.04	5.38	2.17	6.85	4.23	4.01	5.70	6.79	3.94
nitrobenzene dimer	10.49	18.39	13.99	17.04	9.57	8.57	10.42	11.28	9.87	10.87	10.18	28.69	28.71	14.47
FOX-7 dimer	6.45	0.20	5.16	6.04	6.28	6.03	7.87	7.68	5.77	8.84	5.14	10.16	9.10	6.52
EDNA dimer	3.65	1.09	1.74	7.70	3.41	5.02	7.59	7.52	16.99	8.85	30.40	13.00	31.98	10.69
average of MUPEs	3.77	5.69	5.19	6.46	6.08	5.98	7.96	11.96	12.72	7.95	14.94	12.07	26.07	

^a B3LYP-D3; ^b LC- ω PBE-D3; ^c LCgau-BOP+LRD; ^d LC-BOP+LRD; ^e RSH-lrRPAx-SO2; ^f RSH-lrMP2.

by different methods are quite similar: the lateral shift of the wall between the extreme cases is less than 0.05 Å.

C. Performance of DFT methods on individual dimers

A summary of the performance of all methods on all dimers is presented in Table III. Papers comparing performance of various methods often use the average unsigned errors (AUE). To see that this approach does not provide sufficient information about the performance, one can recall the example given in the Introduction where the 0.46 kcal/mol AUE tells absolutely nothing about the performance of a method for the weakest bound system with the interaction energy of -0.04 kcal/mol. A much better measure, in the case when a benchmark set contains only energies at minima, is the average unsigned percentage error (AUPE), which works also very well in the asymptotic region. Unfortunately, AUPE cannot be used for the complete potential energy curves since the percentage errors in the region where a given curve crosses zero can be huge, as we have seen in Sec. V A. Thanthiriwatte *et al.* [11] proposed an elaborate procedure to evaluate such results. Another option is to use the median unsigned percentage error (MUPE) since a few very large errors at the edge of the range do not change the position of the median. We have used MUPE to compare the methods in Table III.

As Table III shows, the two “easiest” systems for all methods turned out to be the imidazole dimer and the water dimer, where the average MUPE was only 3.9% and 4.3% and the largest MUPE was 6.9% and 7.6%, respectively. The reason for good performance in the former case is that with the dipole moment of imidazole equal to 1.5 a.u., this interaction is strongly dominated by the electrostatic effects: for the largest R considered (10.25 Å), the dispersion energy is only 3.5% of electrostatic energy. The water dimer interaction is dominated by electrostatics as well, although to a lesser extent since the water dipole moment is only 0.78 a.u. However, one expects a good performance for water from methods with fitted parameters since the water dimer is included in most training data sets. Despite the low average MUPES, the differences in performance between different methods are significant. For example, for the water dimer, several curves [SAPT(DFT), LCgau-BOP+LRD, LC- ω PBE-D3] almost overlap the benchmark curve, while methods such as B3LYP-D3, RSH-lrMP2, RSH-lrRPAX-SO2, and vdW-DF2 are quite distinctly shifted from it.

The performance is only a bit worse (average MUPE 6.5% and maximum MUPE 10.2%) for the FOX-7 dimer, again related to the large FOX-7’s dipole moment of 3.6 a.u. As for the imidazole dimer, the dispersion energy is only about 3.5% of the electrostatic energy at the largest R (10.58 Å). One may ask why the importance of the dispersion energy is about the same for the two dimers despite the fact that imidazole has the dipole moment twice as large as FOX-7. The reason is that whereas electrostatic interactions coming from various regions of monomers are of both signs, and therefore tend to cancel, the dispersion contributions all add up. Therefore,

for large molecules, dispersion becomes relatively more important at the same level of polarity [100, 101]. Since FOX-7 is twice as large as imidazole, this partially explains the observed relations.

The next group of systems includes the nitromethane, benzene-water, methylformate, and EDNA dimers, with average MUPES in the range 8.3%–10.7% and maximum MUPES 16%–36%. The monomers have modest dipole moments per size of a given molecule, so the increase of difficulty compared to the group discussed above is expected. Exceptions are benzene-water with one dipole moment of zero and nitromethane dimer with the dipole moment of 1.6 a.u. The better than expected average performance on benzene-water can be due to the fact that this system is included in many training data sets. The nitromethane monomer has about the same size and dipole moment as the imidazole monomer, yet the performance of the DFT methods is about twice worse in the former case. This factor of two is quite consistently reflected by individual methods in both cases. The answer to this puzzle is provided by SAPT(DFT) components. Whereas at the largest separation, the ratio of the dispersion energy to the electrostatic energy (amounting to 9% for the nitromethane dimer) is larger than the 3.5% found in the case of the imidazole dimer, (but still very small), at the equilibria the ratios are 202% and 39%. Thus, overall the difficult to reproduce dispersion energy is much more important in the case of the nitromethane dimer. This relative behavior of the electrostatic and dispersion energies is related to the dimer geometry. The nitromethane dimer in a geometry with all atoms of a given monomer fairly close to all atoms of the interacting partner whereas the imidazole dimer has two rings roughly in the same plane, so only a few atoms from each monomer are in a close contact.

Somewhat surprisingly, for the ethanol dimer the average MUPE is 12.1% and the maximum MUPE is 39%, one of the largest. According to the tendencies discussed above, the ethanol dimer is expected to be a more difficult case than the water dimer since the ethanol dipole moment of 0.69 a.u. is slightly smaller than that of water and the number of atoms is twice as large. Also, in the vdW minimum region, the dispersion energy is almost two times larger in magnitude than the electrostatic energy in the former case, whereas in the latter case, the ratio is about one third, similar to the nitromethane–imidazole comparison. One may also notice that for the ethanol dimer about half of the methods perform very well with MUPES in the range 1.6% to 6.8%, whereas the remaining methods have errors from 8.2% to 39%. Thus, only a subset of methods makes the average MUPE so large.

The next to last most difficult system is the nitrobenzene dimer, with the average MUPE of 14.5% and the maximum MUPE of 29%. None of the methods performed well on this system, with all MUPES above 8.6%. This happens since despite the very large dipole moment of nitrobenzene, amounting to 2.1 a.u., the dispersion

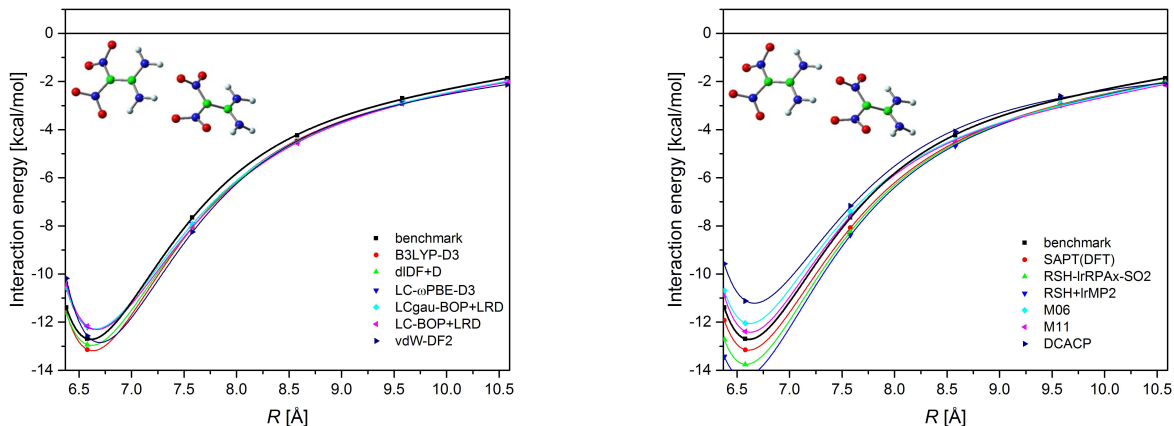


FIG. 9. Potential energy curves for the FOX-7 dimer. See Fig. 1 for explanations.

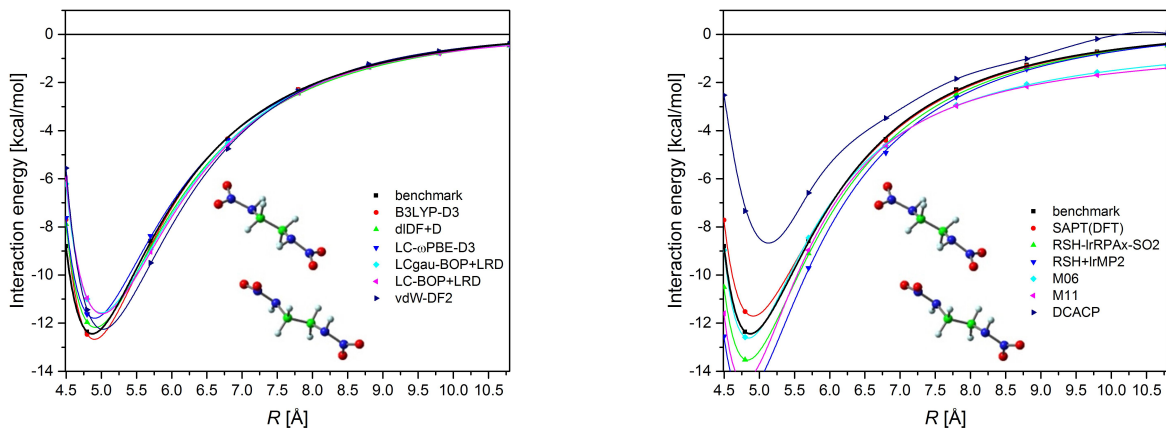


FIG. 10. Potential energy curves for the EDNA dimer. See Fig. 1 for explanations.

energy is about two times larger in magnitude than the electrostatic energy in the region of the vdW minimum. Even SAPT(DFT) gives a very large MUPE of 10.5% for this systems, however, as mentioned earlier, this is partly due to the fact that the $\delta_{\text{int}}^{\text{HF}}$ correction was mistakenly omitted. If it is added, SAPT(DFT)'s MUPE drops to 3.7%.

The largest average MUPE of 19.1% and the maximum MUPE of 68% was obtained for benzene-methane. This is expected as both monomers are nonpolar and the interaction is dispersion dominated: the ratio of the dispersion to the electrostatic energies is about 1 at the largest R , but it is 3.2 near the minimum. The performance of the methods falls into three groups. The best performing SAPT(DFT) method gives a MUPE of only 3.5%, the DFT+D methods and RSH-IrRPAx-SO2 give 6.2%–9.1%, MP2 and RSH-IrMP2 give 14% and 16%, re-

spectively, and the remaining errors are above 28%. The very large errors in the last group lead to the very large average MUPE.

In summary, the order of difficulty correlates reasonably well with the size of the dipole moments, i.e., the larger the moment the easier it is for most methods to obtain accurate predictions of interaction energies. However, for a given size of the dipole moment, the performance will be worse for larger systems since the dispersion energies increase with the number of atoms whereas electrostatic interactions from various regions of molecules tend to cancel. These simple rules explain most trends, some more puzzling cases were explained by analyzing the SAPT components.

D. Overall performance

TABLE IV. Errors of the investigated methods with respect to the benchmarks. AUPE: average unsigned percentage error; AUE: average unsigned error; MUPE: median unsigned percentage error computed for all dimers. The MP2 entry are the CBS extrapolated values obtained as described earlier.

Method	AUPE $R > 1.5 R_{\min}$	AUE (kcal/mol)	MUPE
SAPT(DFT)	3.3	0.17	2.6
MP2	4.3	0.30	3.0
B3LYP-D3	4.4	0.17	3.4
dIDF+D	5.5	0.26	4.2
LC- ω PBE-D3	6.3	0.15	5.3
LCgau-BOP+LRD	5.8	0.23	5.5
M06	20.6	0.31	7.3
vdW-DF2	14.5	0.42	7.5
LC-BOP12+LRD	8.1	0.27	7.7
RSH-lrRPAx-SO2	7.8	0.29	7.7
M11	22.7	0.34	8.5
RSH+lrMP2	11.3	0.57	10.2
DCACP	26.7	0.61	16.6

Table IV lists the overall errors of all computed interaction energies for each method. The methods are ordered using MUPE, but the AUE errors are also listed. The MUPEs are computed for the whole set of 80 configurations and therefore are different from the values in the last row of Table III which are averages of MUPEs for individual dimers. The MUPEs in Table IV range from 2.6% to 16.6%. Thus, the best methods perform really well. In fact, in routine calculations for these systems with the CCSD(T) method, where one would have to choose a much smaller basis set than used in calculations of our benchmarks, MUPE due to limited size of the basis set would certainly have been equal to a few percent, as the MUPE of CCSD(T) predictions in the largest basis sets used by us (without any CBS extrapolations) is 1.2%.

We have also included in Table IV the MP2 method at the basis set level used to calculate the benchmarks extrapolated to the CBS limit. The overall MUPE of MP2 is small, 3.0%, but the method gives very large errors for systems with π electrons, as seen in Table III. For example, the MUPE for the nitrobenzene dimer is as large as 18.4%.

For the purpose of systematizing the discussion, one can divide the methods into three groups using the MUPE ranges below 5%, 5% to 8%, and above 8%. The first group includes two methods that use virtual orbitals: SAPT(DFT) and MP2, and two DFT+D methods: B3LYP-D3 and dIDF+D. B3LYP-D3 calculations are fully converged in basis set size due to much larger basis set used than for other methods, see Table I, which

might have improved the relative performance of this method. Despite using virtual orbitals, for systems of the size included in the present comparison, a SAPT(DFT) calculation is about as time consuming as a DFT supermolecular calculation of the interaction energy [102]. Thus, for such systems, there is no cost difference between SAPT(DFT) and the two DFT+D approaches. The MP2 calculations presented in Table IV were, of course, orders of magnitude more expensive due to the use of very large basis sets. SAPT(DFT) also is overall most reliable since if the nitrobenzene dimer is excluded, the range of MUPE's is between 0.7% and 6.5%. The SAPT(DFT) MUPE of 2.6% is actually very close to the uncertainties resulting from basis set effects: the CBS benchmarks have about 1% uncertainties and a CBS extrapolation of SAPT(DFT) interaction energies would likely result in changes of the order of 1%. Thus, one may say that SAPT(DFT) gives results of CCSD(T) quality at DFT costs for dimers with up to a few dozens of atoms.

The second group includes three DFT+D methods, LC- ω PBE-D3, LCgau-BOP+LRD, and LC-BOP12+LRD, one reoptimized standard functional, M06, one nonlocal functional, vdW-DF2, and one method using virtual orbitals, RSH-lrRPAx-SO2. Thus, each type of approach has a representative in this group. One should point out that there are significant differences between these methods in terms of the largest MUPE (see Table III): the M06 and vdW-DF2 methods have the maximum MUPEs of 29.8% and 28.1%, respectively, whereas for the remaining approaches the largest MUPE is 11.3%.

The third group includes the M11, RSH+lrMP2, and DCACP methods, with maximum MUPEs (see Table III) amounting to 41%, 28%, and 68%, respectively.

The average unsigned errors (AUE), also shown in Table IV, allow for much less precise differentiation between various methods than MUPEs since the range is fairly narrow, from 0.15 to 0.61 kcal/mol. The main reason is that the asymptotic region has little effect on AUEs and they reflect mainly the minimum and the repulsive region results. For modelling of molecular crystals, the long-range behavior downplayed by the AUE criterion is critical since a major contribution to lattice energy comes from dimers with large intermonomer separations. Thus, AUE is not a good measure for applications in this field.

One more quantity listed in Table IV is the average unsigned percentage error for distances larger than 1.5 the minimum distance. For such separations, one can meaningfully use the simple average. These errors mainly show the quality of the dispersion function. It so happens that the four top methods in terms of MUPE, have also the lowest AUPE($R > 1.5 R_{\min}$). The AUPE of LC- ω PBE-D3 is 1.7% larger than that of B3LYP-D3, despite the fact both methods share asymptotically the same dispersion function. However, the short-range switching factors are different in the two methods and the dispersion interaction is not the only component influencing AUPE($R > 1.5 R_{\min}$), as discussed earlier.

One should also discuss here the influence of basis set size on the comparisons. The benchmarks and the MP2 results are at the CBS level. Since DFT is known to converge much faster in basis set size than the calculations using expansions in products of virtual orbitals, the standard DFT methods, the DFT methods with non-local functionals, and the DFT+D methods should all be fairly close to their CBS limits despite using a finite basis set and no extrapolations. In particular, the B3LYP-D3 calculations were performed in basis sets with the cardinal numbers up to 6, so these results are certainly fully converged in basis set size. In contrast, the SAPT(DFT), RSH-lrRPax-SO2, and RSH+lrMP2 predictions would probably improve somewhat if calculations were performed in larger basis sets and extrapolated to the CBS limit.

After having established the relative performance of the investigated approaches, one should consider the other properties of the methods discussed in the Introduction. The methods using virtual orbitals scale as N^5 for SAPT(DFT) and RSH-lrMP2, and as N^6 for RSH+lrRPax-SO2 in the present implementation. Despite such scaling, with proper programming SAPT(DFT) and MP2 are comparable in costs to hybrid DFT methods, scaling as N^4 , for systems of the size considered in the present test. However, for much larger system the better scaling of the latter method will result in a costs advantage. Only two DFT methods considered here, vdW-DF2 and DCACP, are non-hybrid and therefore scale as N^3 . Since the exact exchange is particularly costly in calculations with periodic boundary conditions, these two methods are most economical in such cases. This effectiveness has to be considered, however, in the context of accuracy of these methods.

VI. CONCLUSIONS

We have assessed the performance of a variety of DFT-based methodologies by comparisons to a set

of CCSD(T)/CBS interaction energies obtained for 10 dimers at varying intermonomer distances. This set avoids the typical bias of mixing similar amounts of hydrogen-bonded and purely dispersion-bonded systems and should be more representative for work in the broad field of molecular crystals. The benchmark data reported here should be useful in the development of new methodologies and both the benchmarks and the performance of DFT-based methods examined in this work should enable a convenient assessment of new approaches relative to currently existing techniques. Since our test was blind and used different monomers than considered in related literature, none of the surveyed methods was optimized on benchmarks similar to those included here (except for the water dimer).

ACKNOWLEDGMENTS

The ARL-UD workshop that led to the present paper was funded by ARO. We thank Drs. James Parker and Betsy Rice for their encouragement and constant support of this work. We than to Drs. Edoardo Apra and Karol Kowalski for fixing the errors in NWChem found by us. KS was partly supported by NSF grant Grant No. CHE-1152899. RP acknowledges support from the National Science Centre, Poland, grant No. 2015/17/B/ST4/03727 The work at Rice University was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Computational and Theoretical Chemistry Program under Award No.DE-FG02-09ER16053. GES is a Welch Foundation Chair (C-0036).

-
- [1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
 - [2] R. Eisenschitz and F. London, Z. Phys. **60**, 491 (1930).
 - [3] B. Jeziorski, R. Moszyński, and K. Szalewicz, Chem. Rev. **94**, 1887 (1994).
 - [4] J. Harris, Phys. Rev. B **31**, 1770 (1985).
 - [5] S. Kristyan and P. Pulay, Chem. Phys. Lett. **229**, 175 (1994).
 - [6] P. Hobza, J. Sponer, and T. Reschel, J. Comp. Chem. **16**, 1315 (1995).
 - [7] E. Ruiz, D. R. Salahub, and A. Vela, J. Phys. Chem. **100**, 12265 (1996).
 - [8] Y. Zhao and D. G. Truhlar, J. Chem. Theory Comput. **1**, 415 (2005).
 - [9] I. C. Gerber and J. G. Ángyán, J. Chem. Phys. **126**, 044103 (2007).
 - [10] L. A. Burns, A. Vázquez-Mayagoitia, B. G. Sumpter, and C. D. Sherrill, J. Chem. Phys. **134**, 084107 (2011).
 - [11] K. S. Thanthiriwatte, E. G. Hohenstein, L. A. Burns, and C. D. Sherrill, J. Chem. Theory Comput. **7**, 88 (2011).
 - [12] R. Podeszwa and K. Szalewicz, J. Chem. Phys. Comm. **136**, 161102 (2012).
 - [13] “Dispersion interactions and density functional theory,” <http://www.physics.udel.edu/~szalewic/AR0wks12>, Army Research Laboratory and University of Delaware Workshop, Newark, DE, August 1-3, 2012.
 - [14] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. A. deJong, Comp. Phys. Comm. **181**, 1477 (2010).

- [15] R. Podeszwa, K. Patkowski, and K. Szalewicz, *Phys. Chem. Chem. Phys.* **12**, 5974 (2010).
- [16] See supplemental material at XXX for Cartesian coordinates of all configurations (file Geometries.txt), the (T), CBS, and DFT interaction energies (file Energies.xlsx), and the Hartree-Fock, MP2 and (T) energies (file CBS-Extrap-Components.xlsx) used to obtain the CBS reference values.
- [17] E. Goll, H.-J. Werner, H. Stoll, T. Leininger, P. Gori-Giorgi, and A. Savin, *Chem. Phys.* **329**, 276 (2006).
- [18] X. Xu and W. A. Goddard, *Proc. Nat. Acad. Sci. USA* **101**, 2673 (2004).
- [19] Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.* **120**, 215 (2008).
- [20] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [21] T. van Voorhis and G. E. Scuseria, *J. Chem. Phys.* **109**, 400 (1998).
- [22] A. D. Becke, *J. Chem. Phys.* **104**, 1040 (1996).
- [23] R. Peverati and D. G. Truhlar, *J. Phys. Chem. Lett.* **2**, 2810 (2011).
- [24] H. Stoll and A. Savin, in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. D. Providencia (Plenum, New York, 1985) p. 177.
- [25] A. Savin, in *Recent Developments of Modern Density Functional Theory*, edited by J. M. Seminario (Elsevier, Amsterdam, 1996) pp. 327–357.
- [26] I. C. Gerber and J. G. Ángyán, *Chem. Phys. Lett.* **415**, 100 (2005).
- [27] P. M. W. Gill, R. D. Adamson, and J. A. Pople, *Mol. Phys.* **88**, 1005 (1996).
- [28] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, *J. Chem. Phys.* **115**, 3540 (2001).
- [29] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207 (2003).
- [30] B. Hammer, L. B. Hansen, and J. K. Nørskov, *Phys. Rev. B* **59**, 7413 (1999).
- [31] O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, *Phys. Rev. Lett.* **93**, 153004 (2004).
- [32] I.-C. Lin, M. D. Coutinho-Neto, C. Felsenheimer, O. A. von Lilienfeld, I. Tavernelli, and U. Rothlisberger, *Phys. Rev. B* **75**, 205131 (2007).
- [33] J. Hepburn, G. Scoles, and R. Penco, *Chem. Phys. Lett.* **36**, 451 (1975).
- [34] R. Ahlrichs, R. Penco, and G. Scoles, *Chem. Phys.* **19**, 119 (1977).
- [35] R. Podeszwa, K. Pernal, K. Patkowski, and K. Szalewicz, *J. Phys. Chem. Lett.* **1**, 550 (2010).
- [36] F. A. Gianturco, F. Paesani, M. F. Laranjeira, V. Vasilenko, M. A. Cunha, A. G. Shashkov, and A. F. Zolotoukhina, *Mol. Phys.* **94**, 605 (1998).
- [37] X. Wu, M. C. Vargas, S. Nayak, V. L. Lotrich, and G. Scoles, *J. Chem. Phys.* **115**, 8748 (2001).
- [38] Q. Wu and W. Yang, *J. Chem. Phys.* **116**, 515 (2002).
- [39] S. Grimme, *J. Comput. Chem.* **25**, 1463 (2004).
- [40] K. Pernal, R. Podeszwa, K. Patkowski, and K. Szalewicz, *Phys. Rev. Lett.* **103**, 263201 (2009).
- [41] S. Grimme, J. Antony, S. Elrich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- [42] L. Goerigk and S. Grimme, *Phys. Chem. Chem. Phys.* **13**, 6670 (2011).
- [43] J.-W. Song and K. Hirao, *Chem. Phys. Lett.* **563**, 15 (2013).
- [44] R. Kar, J.-W. Song, T. Sato, and K. Hirao, *J. Comp. Chem.* **34**, 2353 (2013).
- [45] J.-W. Song, S. Tokura, T. Sato, M. A. Watson, and K. Hirao, *J. Chem. Phys.* **127**, 154109 (2007), Errata: **131**, 059901 (2009).
- [46] A. J. Misquitta, B. Jeziorski, and K. Szalewicz, *Phys. Rev. Lett.* **91**, 033201 (2003).
- [47] A. Hesselmann and G. Jansen, *Chem. Phys. Lett.* **367**, 778 (2003).
- [48] A. J. Misquitta, R. Podeszwa, B. Jeziorski, and K. Szalewicz, *J. Chem. Phys.* **123**, 214103 (2005).
- [49] A. Hesselmann, G. Jansen, and M. Schütz, *J. Chem. Phys.* **122**, 014103 (2005).
- [50] Y. Zhao, N. E. Schultz, and D. G. Truhlar, *J. Chem. Theory Comput.* **2**, 364 (2006).
- [51] K. T. Tang and J. P. Toennies, *J. Chem. Phys.* **80**, 3726 (1984).
- [52] L. Rajchel, P. S. Zuchowski, M. M. Szczesniak, and G. Chalasiński, *Phys. Rev. Lett.* **104**, 163001 (2010).
- [53] A. Austin, G. A. Petersson, M. J. Frisch, F. J. Dobek, G. Scalmani, and K. Throssell, *J. Chem. Theory Comput.* **8**, 4989 (2012).
- [54] A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- [55] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Chem. Phys.* **98**, 11623 (1994).
- [56] A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- [57] C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- [58] O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.* **125**, 234109 (2006).
- [59] J. Heyd and G. E. Scuseria, *J. Chem. Phys.* **120**, 7274 (2004).
- [60] W. Cencek and K. Szalewicz, *J. Chem. Phys.* **139**, 024104 (2013), Erratum: **140**, 149902 (2014).
- [61] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, “Gaussian 09, Revision A.02,” (2009).
- [62] T. Tsuneda, T. Suzumura, and K. Hirao, *J. Chem. Phys.* **110**, 10664 (1999).
- [63] T. Sato and H. Nakai, *J. Chem. Phys.* **131**, 224104 (2009).
- [64] H. C. Longuet-Higgins, *Disc. Farad. Soc.* **40**, 7 (1965).
- [65] J. F. Dobson and B. P. Dinte, *Phys. Rev. Lett.* **76**, 1780 (1996).
- [66] F. Rob and K. Szalewicz, *Chem. Phys. Lett.* **572**, 146 (2013).
- [67] F. Rob and K. Szalewicz, *Mol. Phys.* **111**, 1430 (2013).

- [68] J. Řezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**, 2427 (2011).
- [69] K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, *Phys. Rev. B* **82**, 081101 (2010).
- [70] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, *Phys. Rev. Lett.* **92**, 246401 (2004).
- [71] Y. Zhang and W. Yang, *Phys. Rev. Lett.* **80**, 890 (1998).
- [72] E. D. Murray, K. Lee, and D. C. Langreth, *J. Chem. Theory Comput.* **5**, 2754 (2009).
- [73] J. P. Perdew and Y. Wang, *Phys. Rev. B* **33**, 8800 (1986).
- [74] J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
- [75] K. Szalewicz, *Wiley Interdisc. Rev.-Comp. Mol. Sci.* **2**, 254 (2012).
- [76] H. L. Williams and C. F. Chabalowski, *J. Phys. Chem. A* **105**, 646 (2001).
- [77] A. J. Misquitta and K. Szalewicz, *Chem. Phys. Lett.* **357**, 301 (2002).
- [78] A. Hesselmann and G. Jansen, *Chem. Phys. Lett.* **357**, 464 (2002).
- [79] R. Bukowski, R. Podeszwa, and K. Szalewicz, *Chem. Phys. Lett.* **414**, 111 (2005).
- [80] R. Podeszwa, W. Cencek, and K. Szalewicz, *J. Chem. Theory Comput.* **8**, 1963 (2012).
- [81] D. J. Tozer and N. C. Handy, *J. Chem. Phys.* **109**, 10180 (1998).
- [82] E. Fermi and G. Amaldi, *Mem. R. Accad. Italia* **6**, 119 (1934).
- [83] R. Podeszwa, R. Bukowski, and K. Szalewicz, *J. Chem. Theory Comput.* **2**, 400 (2006).
- [84] F. Weigend, *Phys. Chem. Chem. Phys.* **4**, 4285 (2002).
- [85] F. Weigend, A. Köhn, and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2002).
- [86] J. G. Ángyán, I. C. Gerber, J. Toulouse, and A. Savin, *Phys. Rev. A* **72**, 012510 (2005).
- [87] J. Toulouse, A. Savin, and H.-J. Flad, *Int. J. Quantum Chem.* **100**, 1047 (2004).
- [88] J. Toulouse, F. Colonna, and A. Savin, *J. Chem. Phys.* **122**, 014110 (2005).
- [89] R. Moszyński, B. Jeziorski, and K. Szalewicz, *Int. J. Quantum Chem.* **45**, 409 (1993).
- [90] J. Toulouse, W. Zhu, A. Savin, G. Jansen, and J. G. Ángyán, *J. Chem. Phys.* **135**, 084119 (2011).
- [91] A. Szabo and N. S. Ostlund, *J. Chem. Phys.* **67**, 4351 (1977).
- [92] T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- [93] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- [94] H. L. Williams, E. M. Mas, K. Szalewicz, and B. Jeziorski, *J. Chem. Phys.* **103**, 7374 (1995).
- [95] S. F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970).
- [96] A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, and J. Olsen, *Chem. Phys. Lett.* **302**, 437 (1999).
- [97] R. Podeszwa, R. Bukowski, and K. Szalewicz, *J. Phys. Chem. A* **110**, 10345 (2006).
- [98] M. Jeziorska, R. Bukowski, W. Cencek, M. Jaszuński, B. Jeziorski, and K. Szalewicz, *Coll. Czech. Chem. Commun.* **68**, 463 (2003).
- [99] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **14**, 1347 (1993).
- [100] R. Bukowski, K. Szalewicz, and C. Chabalowski, *J. Phys. Chem. A* **103**, 7322 (1999).
- [101] J. Hoja, A. F. Sax, and K. Szalewicz, *Chem. Europ. J.* **20**, 2292 (2014).
- [102] K. Szalewicz, *Acc. Chem. Res.* **47**, 3266 (2014).